



LUND UNIVERSITY

Chlorophyll biosynthesis in barley

Studies on the cyclase and chlorophyll synthase steps

Stuart, David

2024

Document Version:

Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):

Stuart, D. (2024). *Chlorophyll biosynthesis in barley: Studies on the cyclase and chlorophyll synthase steps*. [Doctoral Thesis (compilation), Department of Biology]. Lund University (Media-Tryck).

Total number of authors:

1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

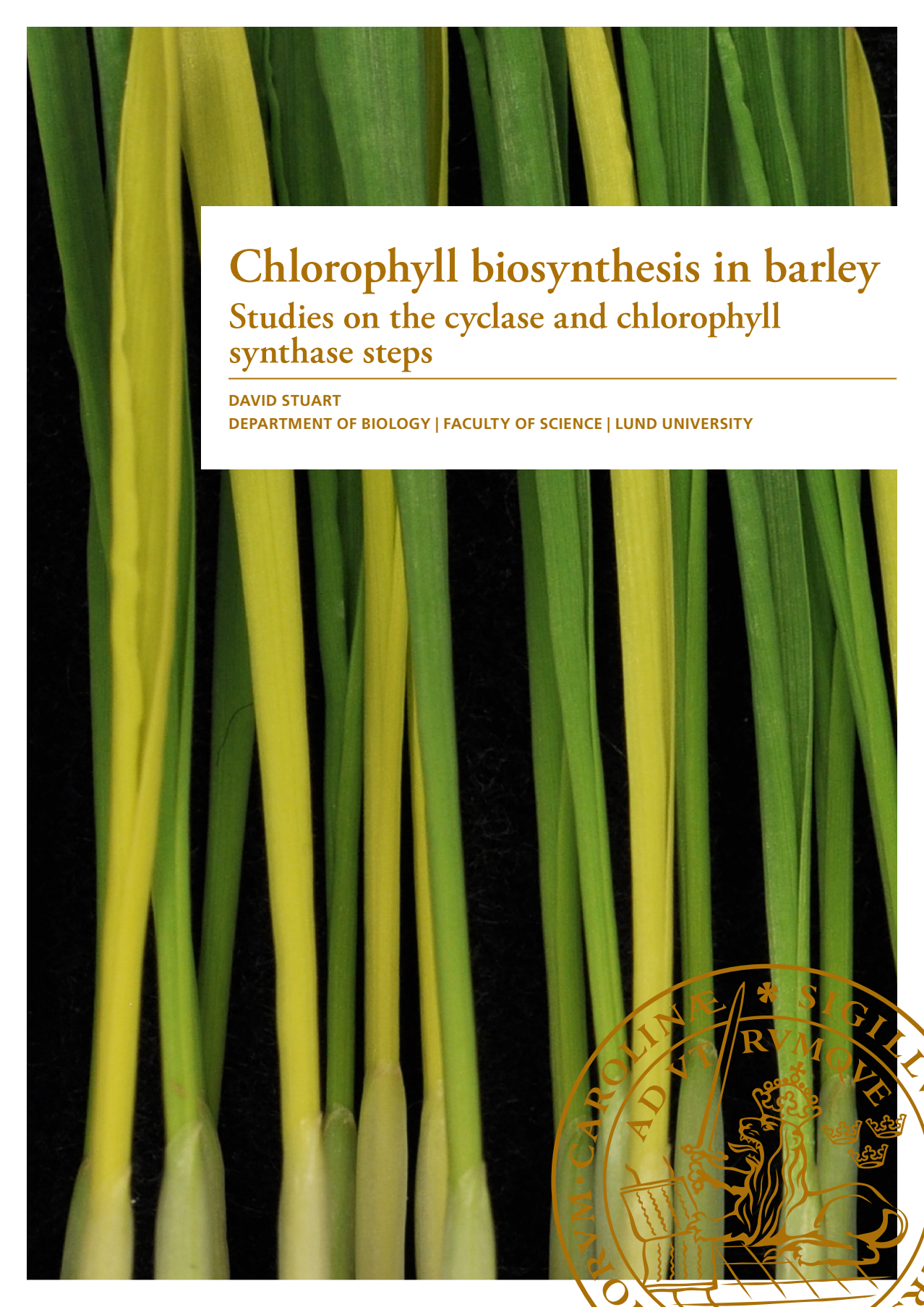
Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

A close-up photograph of several barley leaves against a black background. The leaves show a clear pattern of chlorophyll deficiency, with alternating green and yellowish-white stripes. The yellow stripes are most prominent in the lower half of the leaves, indicating a developmental or genetic defect in chlorophyll synthesis.

Chlorophyll biosynthesis in barley

Studies on the cyclase and chlorophyll synthase steps

DAVID STUART

DEPARTMENT OF BIOLOGY | FACULTY OF SCIENCE | LUND UNIVERSITY



Chlorophyll biosynthesis in barley
- studies on the cyclase and chlorophyll synthase steps

Chlorophyll biosynthesis in barley

Studies on the cyclase and chlorophyll synthase steps

David Stuart



LUND
UNIVERSITY

DOCTORAL DISSERTATION

Doctoral dissertation for the degree of Doctor of Philosophy (PhD) at the Faculty of Science at Lund University to be publicly defended on 13th of June 2024 at 09.00 in the Lecture Hall A, Department of Biology, Sölvegatan 35B, Lund.

Faculty opponent

Henrik Aronsson, University of Gothenburg

Organization: LUND UNIVERSITY

Document name: Doctoral dissertation

Date of issue 2024-06-13

Author(s): David Stuart

Sponsoring organization:

Title and subtitle: Chlorophyll biosynthesis in barley - studies on the cyclase and chlorophyll synthase steps

Abstract:

The well-known green color of plants is due to the chlorophyll molecules used for light harvesting during photosynthesis, which takes place in chloroplasts. Biosynthesis of chlorophyll *a* is performed by 15 enzymatic steps and chlorophyll *a* can be interconverted to chlorophyll *b* via the chlorophyll cycle. Barley is a valuable model organism for studying chloroplast development and chlorophyll biosynthesis thanks to the availability of many chlorophyll mutants with abnormal pigment accumulation. Many of these mutants are deficient in enzymes of chlorophyll biosynthesis. Using barley as a model system, I have performed studies of two enzymatic steps in the chlorophyll biosynthetic pathway; the aerobic magnesium-protoporphyrin IX monomethyl ester cyclase and the chlorophyll synthase. I have identified the cyclase as a ferredoxin dependent enzyme and established the first *in vitro* assay using recombinant cyclase enzyme. I have also utilized the barley *viridis-k* chlorophyll mutants, which have impaired chlorophyll synthesis due to deficiency at the cyclase step. This revealed that a particular ferredoxin isoform, FdC2, is likely the main *in vivo* electron donor to the cyclase. By developing a pipeline for identification of genes deficient in barley mutants and applying this on barley *xantha-j* mutants, I connected this locus to the chlorophyll synthase. An AlphaFold generated structural model of the chlorophyll synthase suggested the active site to be subdivided into a prenyl pyrophosphate tunnel, the catalytic cavity, and a tetrapyrrole-binding pocket. A deep multiple sequence alignment contributed to the understanding of amino-acid residues. I suggest that binding of the isoprenoid substrate is a prerequisite for stable maintenance of chlorophyll synthase in the plastid. I further suggest that chlorophyll synthase is a sensor for coordinating chlorophyll and isoprenoid biosynthesis.

Key words: barley, Chlorina, chlorophyll synthase, *Hordeum vulgare*, Mg-protoporphyrin IX monomethyl ester cyclase, mutant, Viridis, *vir-k*, Xantha, *xan-j*, *xan-l*, *ycf54*

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language

ISSN and key title:

ISBN: 978-91-8104-048-7 (print),

978-91-8104-049-4 (electronic)

Recipient's notes

Number of pages:44

Price

Security classification

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

David Stuart

Date 2024-04-24

Chlorophyll biosynthesis in barley

Studies on the cyclase and chlorophyll synthase steps

David Stuart



LUND
UNIVERSITY

Coverphoto by Mats Hansson

Copyright pp 1-44 David Stuart

Paper 1 © The authors (2020). Published by MDPI AG, Switzerland

Paper 2 © The authors (2021). Published by Oxford University Press on behalf of American Society of Plant Biologists

Paper 3 © The authors (2024). Published by Oxford University Press on behalf of American Society of Plant Biologists

Faculty of Sciences

Department of Biology

ISBN 978-91-8104-048-7 (print)

ISBN 978-91-8104-049-4 (electronic)

Printed in Sweden by Media-Tryck, Lund University

Lund 2024



Media-Tryck is a Nordic Swan Ecolabel certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

MADE IN SWEDEN 

Table of Contents

Abstract	8
Populärvetenskaplig sammanfattning	9
List of Papers.....	10
Author's contribution to the papers.....	11
Abbreviations	12
Introduction	13
Chlorophyll and photosynthesis	13
The microbial route to chlorophyll biosynthesis.....	15
Chlorophyll mutants in barley.....	16
Chlorophyll biosynthesis.....	17
Shared steps of chlorophyll and heme biosynthesis.....	17
Magnesium chelatase	19
Methyl transferase.....	22
Cyclase	23
Protochlorophyllide oxidoreductase.....	25
Divinyl reductase	26
Chlorophyll synthase and geranylgeranyl reductase.....	27
The chlorophyll cycle.....	28
Research results	31
Paper 1.....	31
Paper 2.....	32
Paper 3.....	33
Future perspectives	35
References	37

Abstract

The well-known green color of plants is due to the chlorophyll molecules used for light harvesting during photosynthesis, which takes place in chloroplasts. Biosynthesis of chlorophyll *a* is performed by 15 enzymatic steps and chlorophyll *a* can be interconverted to chlorophyll *b* via the chlorophyll cycle. Barley is a valuable model organism for studying chloroplast development and chlorophyll biosynthesis thanks to the availability of many chlorophyll mutants with abnormal pigment accumulation. Many of these mutants are deficient in enzymes of chlorophyll biosynthesis. Using barley as a model system, I have performed studies of two enzymatic steps in the chlorophyll biosynthetic pathway; the aerobic magnesium-protoporphyrin IX monomethyl ester cyclase and the chlorophyll synthase. I have identified the cyclase as a ferredoxin dependent enzyme and established the first *in vitro* assay using recombinant cyclase enzyme. I have also utilized the barley *viridis-k* chlorophyll mutants, which have impaired chlorophyll synthesis due to deficiency at the cyclase step. This revealed that a particular ferredoxin isoform, FdC2, is likely the main *in vivo* electron donor to the cyclase. By developing a pipeline for identification of genes deficient in barley mutants and applying this on barley *xantha-j* mutants, I connected this locus to the chlorophyll synthase. An AlphaFold generated structural model of the chlorophyll synthase suggested the active site to be subdivided into a prenyl pyrophosphate tunnel, the catalytic cavity, and a tetrapyrrole-binding pocket. A deep multiple sequence alignment contributed to the understanding of amino-acid residues. I suggest that binding of the isoprenoid substrate is a prerequisite for stable maintenance of chlorophyll synthase in the plastid. I further suggest that chlorophyll synthase is a sensor for coordinating chlorophyll and isoprenoid biosynthesis.

Populärvetenskaplig sammanfattning

Klorofyll är antagligen den mest synliga molekylerna på jorden – dess gröna färg syns överallt där det finns tillräckligt med vatten för att hålla växter vid liv. Det är på klorofyllmolekylen som alla större näringskedjor är baserade, då växterna är basproducenterna. Också jordbruket är helt beroende av klorofyll och därmed också vår egen existens. Allt detta har skött sig själv under många miljoner år utan att vi haft en aning om hur klorofyll tillverkas i växtcellerna. Men ska vi i framtiden kunna effektivisera våra grödor så är de kan utnyttja ännu mer av solens energi så måste vi först lära oss hur klorofyll tillverkas och hur detta hänger ihop med bildandet av växternas kloroplaster och själva fotosyntesen.

Jag har arbetat med en fantastisk samling av kornmutanter som är defekta i klorofyllsyntes och kloroplastens utveckling. Dessa togs fram genom strålnings- och kemikaliebehandling av svenska och danska växtförädlare och forskare mellan 1928-1983. Jag har utnyttjat mutanterna för att identifiera de gener som de är defekta i. Mycket av genens funktion avslöjas genom att jag lyckats knyta genen till syntes av klorofyll men jag har också gjort molekylära och biokemiska analyser av mutanterna och de inblandade proteinerna. På detta sätt har jag lyckats förstå genen och mutanten i ett större molekylärt sammanhang. Mer specifikt har jag identifierat den gen vars protein ger elektroner till cyclasenzymet i klorofyllbiosyntesen. Jag har också identifierat den gen som kodar för klorofyllsyntaset, vilket katalyserar det sista steget i syntesen av klorofyll.

Projektet är av grundvetenskaplig karaktär men kunskap om generna är en förutsättning för att veta vilka gener man ska använda sig av för att i framtiden modifiera de odlade växterna så att de kan öka sin fotosyntes. Att studien baseras på korn som är den mest odlade grödan i Sverige gör steget till tillämpningar kortare även om dessa antagligen ligger långt fram i tiden.

List of Papers

Paper I

Stuart, D., M. Sandström, Helmy M. Youssef, S. Zakhrabekova, P. E. Jensen, D. W. Bollivar and M. Hansson. 2020. Aerobic barley Mg-protoporphyrin IX monomethyl ester cyclase is powered by electrons from ferredoxin. *Plants* 9: 1157. <https://doi.org/10.3390/plants9091157>

Paper II

Stuart, D., M. Sandström, Helmy M. Youssef, S. Zakhrabekova, P. E. Jensen, D. Bollivar and M. Hansson. 2021. Barley *Viridis-k* links an evolutionary conserved C-type ferredoxin to chlorophyll biosynthesis. *Plant Cell* 33: 2834-2849. <https://doi.org/10.1093/plcell/koab150>

Paper III

Stuart, D., S. Zakhrabekova, M. Egevang Jørgensen, C. Dockter and M. Hansson. 2024. A pipeline for identification of causal mutations in barley identifies *Xantha-j* as the chlorophyll synthase gene. *Plant Physiol.* Published as an accepted manuscript. <https://doi.org/10.1093/plphys/kiae218>

Author's contribution to the papers

Paper I

Conceptualization, D.S., P.E.J., D.W.B. and M.H.; methodology, D.S., P.E.J., D.W.B. and M.H.; validation, D.S., P.E.J., D.W.B. and M.H.; formal analysis, D.S. and H.M.Y.; investigation, D.S., M.S., H.M.Y., S.Z., D.W.B. and M.H.; resources, M.H.; data curation, D.S. and M.H.; writing—original draft preparation, D.S., D.W.B. and M.H.; writing—review and editing, D.S., H.M.Y., S.Z., P.E.J., D.W.B. and M.H.; visualization, D.S., H.M.Y., S.Z. and M.H.; supervision, D.S., P.E.J., D.W.B. and M.H.; project administration, M.H.; funding acquisition, D.S., H.M.Y., S.Z., P.E.J., D.W.B. and M.H. All authors have read and agreed to the published version of the manuscript.

Paper II

D.S., D.B., M.H., and P.E.J. designed the experiments; D.S., H.M.Y., M.S., and S.Z. performed the experiments; D.S., D.B., and M.H. analyzed the data; D.S., D.B., M.H., and P.E.J. wrote the article. All authors read and approved the final article.

Paper III

DS and MH designed the research. DS, SZ, MEJ, CD and MH performed research. DS, SZ, MEJ, CD and MH analyzed data. DS and MH wrote the paper. All authors read and approved the final version of the manuscript.

Abbreviations

ALA	5-Aminolevulinic acid
ATP	Adenosine triphosphate
C5	Carbon 5
CAO	Chlorophyllide <i>a</i> oxygenase
CBR	Chlorophyll(ide) <i>b</i> reductase
CoA	Coenzyme A
CPOX	Coproporphyrinogen III oxidase
DPOR	Dark operative protochlorophyllide oxidoreductase
DVR	Divinyl reductase
FAD	Flavin adenine dinucleotide
Fd	Ferredoxin
FNR	Ferredoxin NADP ⁺ oxidoreductase
GluTR	Glutamyl-tRNA reductase
GSA-AT	Glutamate-1-semialdehyde aminotransferase
HCAR	7-Hydroxymethyl chlorophyll(ide) <i>a</i> reductase
kbp	Kilo base pairs
kDa	Kilo Daltons
LPOR	Light dependent protochlorophyllide oxidoreductase
Mbp	Mega base pairs
NADP	Nicotinamide-adenine-dinucleotide phosphate
NOL1	NYC1-like
NYC1	NON-YELLOW COLORING1
PBG	Porphobilinogen
PPOX	Protoporphyrinogen IX oxidase
PSI	Photosystem I
PSII	Photosystem II
SAH	S-adenosylhomocysteine
SAM	S-adenosylmethionine

Introduction

Chlorophyll and photosynthesis

Chlorophyll *a* and *b* are the light harvesting pigments used by plants for photosynthesis, a process that harvests and converts light energy into chemical energy. This process takes place in the endosymbiont derived organelle known as the chloroplast (Schwartz and Dayhoff, 1978, Sagan, 1967). These organelles have an envelope composed of an inner and outer membrane surrounding the stroma (Figure 1). Within the stroma is the thylakoid membrane system, which separates the stroma from the thylakoid lumen (Staehelein, 2003). The thylakoids are further organized as stacks known as grana and un-stacked stromal thylakoids. The chlorophyll-protein complexes responsible for the harvest and conversion of light energy are located in the thylakoid membranes (Figure 2).

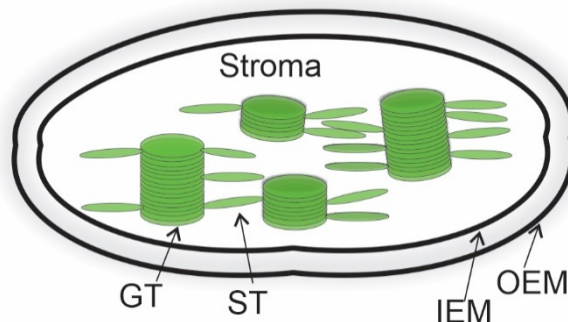


Figure 1 Schematic representation of a chloroplast. The organelle is enclosed by a double membrane consisting of the outer envelope membrane (OEM) and the inner envelope membrane (IEM) which surrounds the stroma. The thylakoid membrane system separates the stroma from the thylakoid lumen. Thylakoid membranes are organized into grana thylakoids (GT) which form stacks called grana and un-stacked stromal thylakoids (ST).

Chlorophylls are mainly found in the two photosystems, which perform charge separation (electron transfer) as well as in antenna complexes which help collect and funnel energy to the reaction centers in the photosystems (Caspary et al., 2020, Su et al., 2017). Electron transport begins at photosystem II where water is the electron donor at the lumen side of the membrane and the electrons are used to reduce plastoquinone to plastoquinol located within the thylakoid membrane (Hohmann-Marriott and Blankenship, 2011). Photosystem II contains a type II reaction center, a defining feature of which is using a quinone as the final electron acceptor (Hohmann-Marriott and Blankenship, 2011). The oxidation of water produces O_2 and H^+ at the lumen side and the protons required for the conversion of plastoquinone to plastoquinol are taken up from the stoma side. This results in a net decrease in pH within the lumen and this electrochemical gradient is used to drive ATP synthesis by the F_1F_0 ATP synthase also located in the thylakoid membranes (Hahn et al., 2018). Plastoquinol is oxidized back to plastoquinone by the cytochrome *b₆f* complex via the Q-cycle which releases protons into the lumen and takes up protons from the stroma, thus contributing to the electrochemical proton gradient across the thylakoid membrane (Stroebel et al., 2003). At the

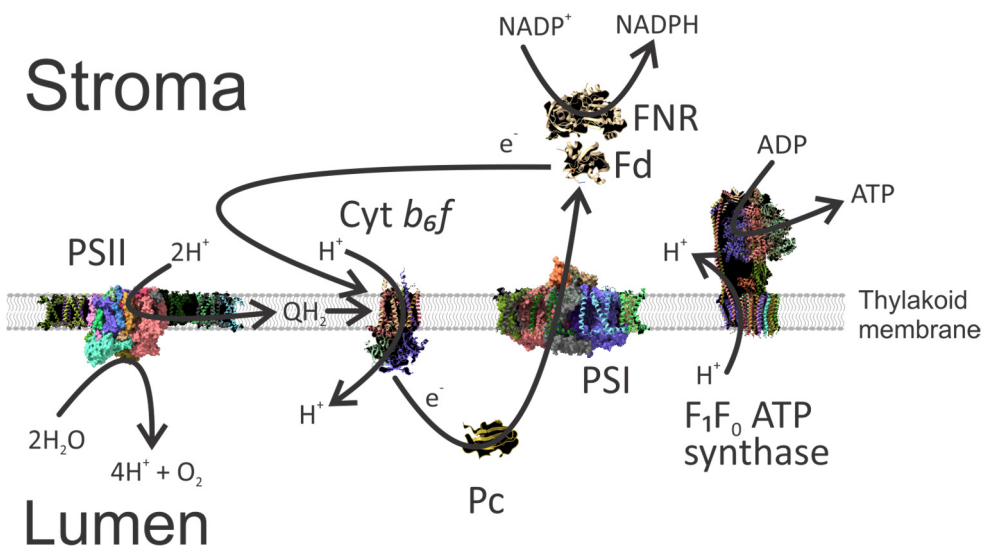


Figure 2 The main photosynthetic complexes located in the thylakoid membrane with the main routes of proton and electron transfer. PSII; photosystem II, image rendered from PDB 5XNL (Su et al., 2017). QH₂; plastoquinol. Cyt *b₆f*; cytochrome *b₆f* complex, image rendered from PDB 1Q90 (Stroebel et al., 2003). Pc; plastocyanin, image rendered from PDB 6YEZ (Caspary et al., 2020). PSI; photosystem I, image rendered from PDB 6YEZ. Fd; ferredoxin, image rendered from PDB 5H5J (Shinohara et al., 2017). FNR; Ferredoxin NADP⁺ oxidoreductase, image rendered from PDB 5H5J. F₁F₀ ATP synthase rendered from PDB 6FKF (Hahn et al., 2018).

cytochrome *b₆f* complex, electrons are passed to the soluble electron carrier plastocyanin which is located in the thylakoid lumen. Plastocyanin is the electron donor to photosystem I (PSI) which transfers the electrons to the soluble 2Fe-2S cluster containing ferredoxin protein. PSI has a type I reaction center, a defining feature of which is having an iron sulfur cluster as the terminal electron acceptor (Hohmann-Marriott and Blankenship, 2011). Ferredoxin can in turn donate electrons to ferredoxin NADP⁺ oxidoreductase (FNR) which reduces NADP⁺ to NADPH. Alternatively, ferredoxin can donate electrons directly to ferredoxin dependent enzymes. The generation of low potential electrons originating from water that are channeled into other metabolic processes is generally referred to as linear electron transport. However, electrons from ferredoxin can be channeled back into the plastoquinone pool by reducing plastoquinone in a process dependent on PGR5 and PGRL1 (Hertle et al., 2013). This is known as cyclic electron transport since the electrons cycle around the cytochrome *b₆f* complex and PSI. Since electron transfer through the cytochrome *b₆f* complex contributes to the proton gradient across the thylakoid membrane, this is a mechanism to balance the cellular need for ATP versus NADPH, as well as providing a mechanism to maintain the proper redox balance of the plastoquinone pool.

The microbial route to chlorophyll biosynthesis

While plants contain only chlorophyll *a* and *b*, there are a number of other chlorophylls and bacteriochlorophylls found in other photosynthetic organisms spread throughout the tree of life. Only cyanobacteria and eukaryotes with plastids perform oxygenic photosynthesis and contain both type I and type II reaction centers. Numerous bacterial groups such as *Acidobacteria*, *Chloroflexi*, *Chlorobi*, *Firmicutes*, *Gemmatinomadetes*, and *Proteobacteria* contain photosynthetic members that perform anoxygenic photosynthesis and contain only a single photosynthetic reaction center of either type I or type II (Hamilton, 2019). These bacteria generally produce bacteriochlorophylls but many of the biosynthetic steps use enzymes which are homologous to those found in cyanobacteria, algae and plants. Most of the genes involved in (bacterio)chlorophyll biosynthesis were thus first identified in bacterial species such as *Rhodobacter capsulatus* which contains all the genes required for photosynthesis in a 46 kbp cluster termed the photosynthesis gene cluster (Bollivar et al., 1994b). Many genes were thus named “*bch**” where the * is a letter designation for the gene. This naming convention was then transferred to chlorophyll synthesizing organisms as “*chl**”. Many genes, however, have alternative names that differ between organisms mainly for historical reasons since the genes have kept the names from the mutants, which were used for the discovery of the genes.

Chlorophyll mutants in barley

A rich variety of chlorophyll mutants are available in barley (*Hordeum vulgare* L.). The many mutants were induced by physical and chemical mutagenesis in breeding programs aiming for high-yielding barley cultivars. Mutants with low or no chlorophyll are not associated with high yield, but the chlorophyll mutants were easily observed in the M₂ generation already at the seedling stage and thus served as indicators for the success of the mutagenic treatment (Lundqvist, 1992). When the barley chlorophyll mutants were generated, they were classified based on their phenotype as *albina* (white), *xantha* (yellow), *viridis* (light green lethal), *chlorina* (light green viable), and *tigrina* (transverse striped) (Hansson et al., 2024). Mutants with the same phenotype were given an identifying number and then crossed to each other to identify allelic complementation groups which were designated by a letter. An example name would then be *xantha-l.35*. When it was first discovered in the field after mutagenesis with ethyleneimine in 1957 (Henningsen et al., 1993), it was given the name *xantha-35* because it was the 35th yellow mutant isolated. It was found that crosses to for example *xantha-81*, but not *xantha-27*, could generate yellow plants, which demonstrated that *xantha-35* and *xantha-81*, but not *xantha-27*, are deficient in the same gene. Their names were then extended to *xantha-l.35* and *xantha-l.81*, respectively, to indicate that they are allelic. For many years, very few molecular analyses could be performed with the barley mutants. The large 5.4 Mbp diploid genome was a major obstacle until a reference genome sequence was published (Mascher et al., 2017). The availability of a reference genome sequence has created a renaissance for the many available mutants since map-based cloning and marker analyses can now take a short-cut when transferring information between genetic and physical maps. Still, there is a challenge to study chlorophyll mutants since the recessive mutations, except for the light green *chlorina* mutants, are lethal in homozygous form. Therefore, the mutations must be kept in heterozygous stocks and each propagation has to be followed by segregation analyses in order to sort heterozygous green mutants from homozygous green wild-type plants. There are also advantages using barley as a model organism for chlorophyll biosynthesis. This is because the large barley seed can support also a lethal mutant with energy to form a reasonably large mutant seedling that can be used, not only for molecular studies, but also for biochemical analyses.

Chlorophyll biosynthesis

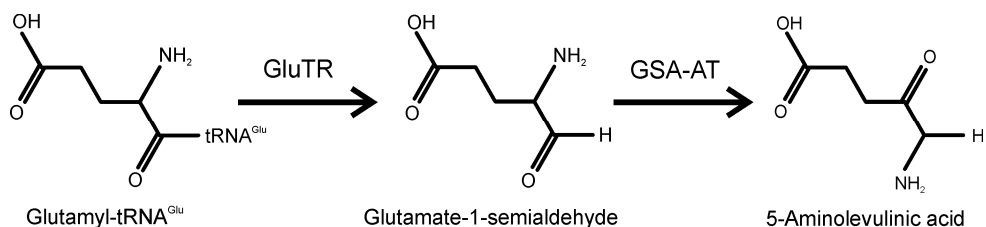


Figure 3 The C5 pathway for 5-aminolevulinic acid biosynthesis utilized by plants. Activated glutamate in the form of glutamyl-tRNA^{Glu} is first converted to glutamate-1-semialdehyde by glutamyl-tRNA reductase (GluTR) followed by conversion to 5-aminolevulinic acid by glutamate-1-semialdehyde aminotransferase (GSA-AT).

Chlorophylls are cyclic tetrapyrrole molecules and the biosynthetic pathway in plants involves numerous enzymatic steps. The first part of the pathway is shared with heme biosynthesis and the enzymatic steps are often called the “common” or “shared” steps since they are shared in common for both heme and chlorophyll biosynthesis until the red colored intermediate protoporphyrin IX. The first step specific for chlorophyll biosynthesis is insertion of a Mg^{2+} ion into protoporphyrin IX and the chlorophylls specific steps are therefore generally referred to as the “chlorophyll branch” or the “Mg branch” and ends with the synthesis of the characteristically green chlorophylls *a* and *b* in plants. Chlorophyll *a* and *b* can be interconverted via the chlorophyll cycle.

Shared steps of chlorophyll and heme biosynthesis

The common steps start with the synthesis of 5-aminolevulinic acid (ALA), the first dedicated intermediate of tetrapyrrole biosynthesis. There are two known pathways for ALA synthesis. One is the Shemin pathway that produces ALA by condensation of glycine and succinyl-CoA (Shemin et al., 1955). The second pathway is the carbon 5 (C5) pathway which produces ALA from glutamate (Beale and Castelfranco, 1973, Beale et al., 1975). Plants utilize the C5 pathway for ALA synthesis by the action of two enzymes (Figure 3), glutamyl-tRNA reductase

(GluTR) followed by glutamate-1-semialdehyde aminotransferase (GSA-AT). Activated glutamate in the form of glutamyl-tRNA^{Glu} (Schön et al., 1986, Huang et al., 1984) is first converted by GluTR into glutamate-1-semialdehyde by conversion of the carboxyl group to an aldehyde. GSA-AT next produces ALA by swapping the positions of the amino group and the carbonyl group (Kannangara and Gough, 1978, Grimm, 1990).

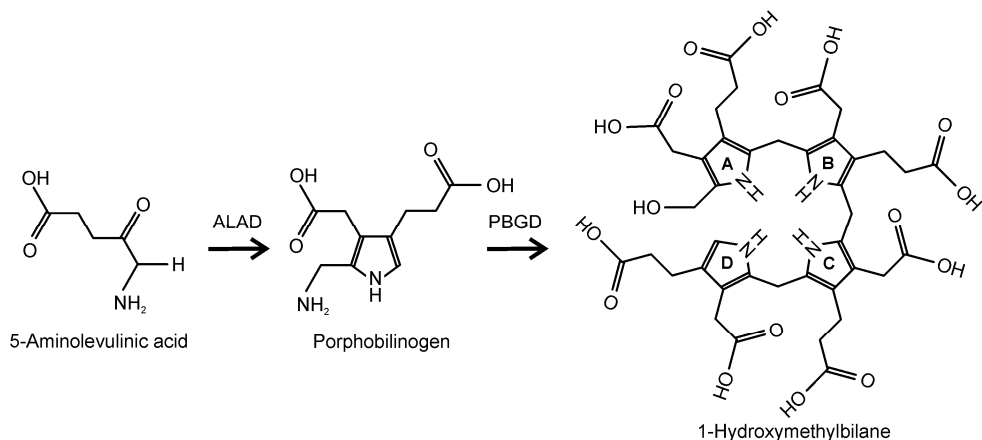


Figure 4 Biosynthesis of 1-hydroxymethylbilane (HMB) from 5-aminolevulinic acid (ALA). The enzyme ALA dehydratase (ALAD) forms the first pyrrole, porphobilinogen (PBG), by condensing two molecules of ALA. Four molecules of PBG are then combined by PBG deaminase (PBGD) to form the linear tetrapyrrole HMB. The four pyrrole rings are denoted as ring A, B, C, and D.

In the next step, two ALA molecules are condensed by ALA dehydratase to form porphobilinogen (PBG) which is a pyrrole (Boese et al., 1991, Schaumburg et al., 1992) (Figure 4). Further, PBG deaminase combines four PBG molecules to produce the linear tetrapyrrole 1-hydroxymethylbilane (Witty et al., 1993, Jones and Jordan, 1994). It thus takes a total of eight ALA molecules to produce one tetrapyrrole molecule. The next step is inversion of ring D and a ring closure catalyzed by uroporphyrinogen III synthase to form the first cyclic tetrapyrrole intermediate (Figure 5) (Tan et al., 2008). The enzyme uroporphyrinogen III decarboxylase next produces coproporphyrinogen III by decarboxylating the acetate groups on each of the four pyrrole rings (Mock et al., 1995). This is followed by conversion of the propionate side chains on rings A and B to vinyl groups by the oxygen-dependent coproporphyrinogen III oxidase (CPOX) to produce protoporphyrinogen IX (Madsen et al., 1993, Kruse et al., 1995). The last of the common steps is a six electron oxidation performed by protoporphyrinogen IX oxidase (PPOX) which converts the methylene bridges to methine bridges between the pyrrole rings to produce the first porphyrin of the pathway (Figure 6) (Jacobs

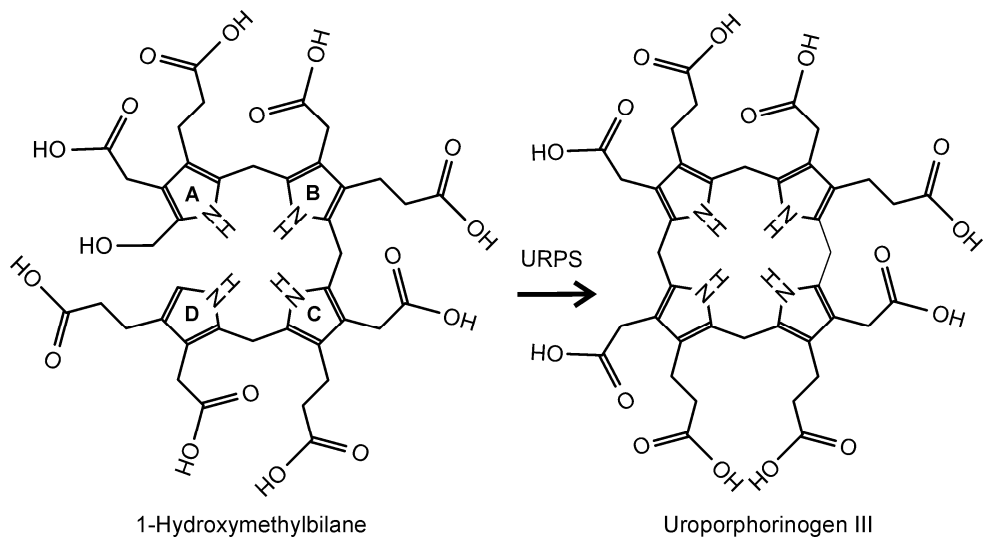


Figure 5 From a linear to a ring-closed tetrapyrrole. HMB is converted by uroporphyrinogen III synthase (URPS) to the first cyclic tetrapyrrole intermediate, uroporphyrinogen III. In the process, the D ring is turned leaving an isomer with two propionyl groups pointing down from rings C and D.

and Jacobs, 1987, Lermontova et al., 1997). At this point the pathway forks and metabolites are diverted to either heme biosynthesis after insertion of Fe^{2+} into the macrocycle by ferrochelatase or to chlorophyll biosynthesis after insertion of Mg^{2+} by the magnesium chelatase.

Magnesium chelatase

The insertion of Mg^{2+} into protoporphyrin IX (Figure 7) is performed by the magnesium chelatase which is a multi-subunit enzyme made up of three subunits. These are encoded by the *bchI/chlI*, *bchD/chlD*, and *bchH/chlH* genes (Bollivar et al., 1994b, Jensen et al., 1996a), which correspond to the *Xantha-h*, *Xantha-g*, and *Xantha-f* genes in barley (Jensen et al., 1996b) and have a mass of approximately 40 kDa, 70 kDa, and 140 kDa, respectively. The enzyme is located in the stroma but XanF can associate to thylakoid and envelope membranes depending on the buffer conditions, especially the Mg^{2+} concentration (Walker and Weinstein, 1995).

The reaction is ATP and Mg^{2+} dependent (Fuesler et al., 1981) and requires a Mg^{2+} concentration greater than that of ATP (Jensen et al., 1998). Early studies established that the reaction mechanism proceeds through a slow ATP dependent activation step followed by a faster, also ATP dependent metal chelation step (Walker and Weinstein, 1991). The activation step was shown to require the XanH/BchI and XanG/BchD subunits as well as ATP (Willows et al., 1996) and

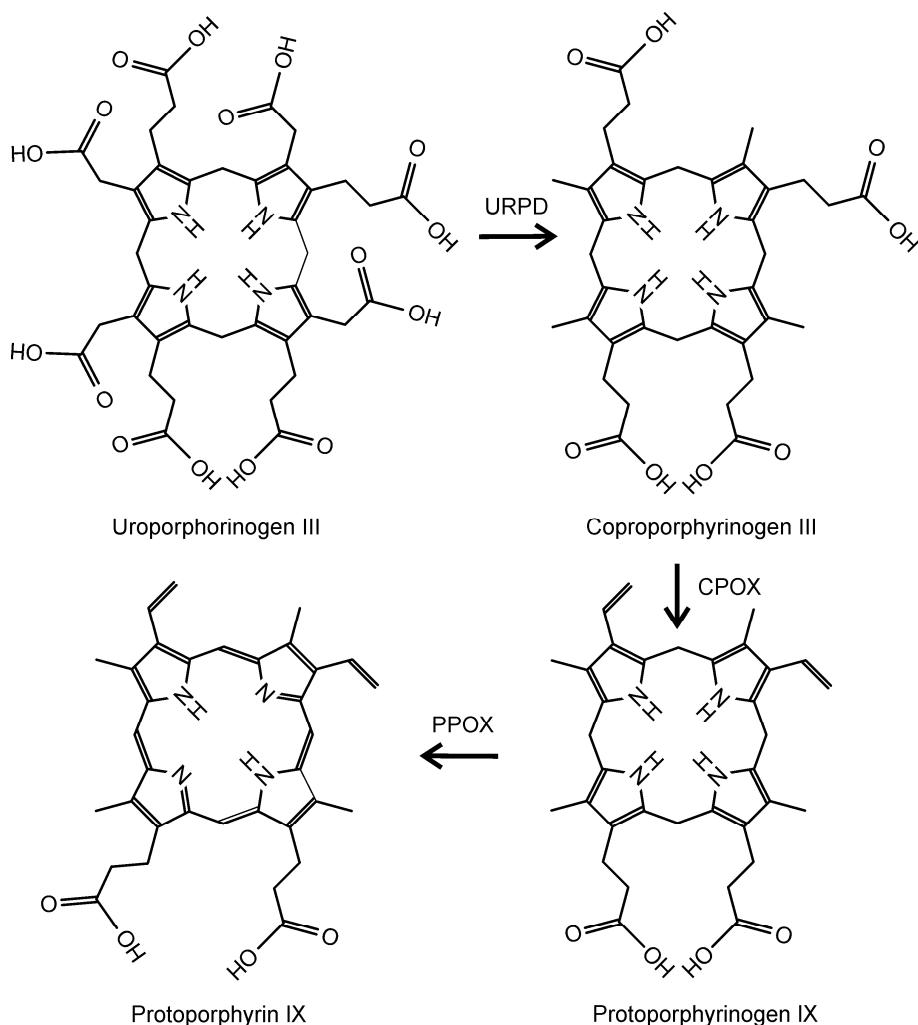


Figure 6 The three final enzymatic steps shared between chlorophyll and heme biosynthesis. Uroporphyrinogen III is converted to coproporphyrinogen III by uroporphyrinogen III decarboxylase (URPD). Next, protoporphyrinogen IX is produced by the action of the enzyme coproporphyrinogen III oxidase (CPOX). Finally, protoporphyrin IX is produced by the enzyme protoporphyrinogen IX oxidase (PPOX).

removal of Mg-protoporphyrin IX from the enzyme (Larkin et al., 2003). GUN4 also lowers the required Mg^{2+} concentration (Davison et al., 2005).

Methyl transferase

After magnesium chelation, the enzyme Mg-protoporphyrin IX O-methyl transferase, encoded by *bchM/chlM* (Bollivar et al., 1994a, Smith et al., 1996), forms a methyl ester on the propionate side chain of ring C (Figure 8) using S-adenosylmethionine (SAM) as the methyl donor (Tait and Gibson, 1961). No mutants in the methyl transferase gene of barley have been described. In *Arabidopsis* the enzyme has been localized to both the thylakoid and envelope membranes (Block et al., 2002).

The enzymatic reaction requires S-adenosylmethionine and is competitively inhibited by S-adenosylhomocysteine and S-adenosylmethionine (Shieh et al., 1978). It appears that activity of the methyl transferase is stimulated by the large subunit of the magnesium chelatase in both bacterial and eukaryotic systems (Hinchigeri et al., 1997, Alawady et al., 2005, Shepherd et al., 2005) which suggests a conserved substrate channeling mechanism. The enzyme has also been shown to be stimulated by addition of phospholipids to assay mixtures (Sawicki and Willows, 2007).

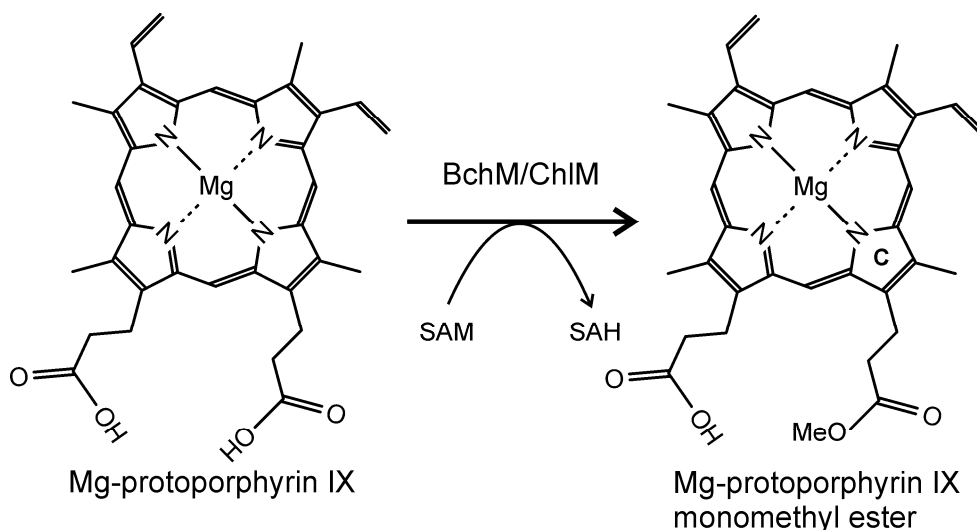


Figure 8 The methyl transferase reaction produces a methyl ester on propionate group on ring C. The reaction is performed by BchM/ChlM and uses S-adenosylmethionine (SAM) as the methyl donor. SAM is converted to S-adenosylhomocysteine (SAH).

Cyclase

After formation of the methyl ester the next enzyme produces the fifth ring E (Figure 9) which also results in the first green intermediate of the chlorophyll biosynthesis pathway (Bryant et al., 2020). This is a multi-step reaction that proceeds through addition of a carbonyl group to C13¹ followed by a carbon-carbon bond formation between C13² and the carbon bridging ring C and D which produces ring E. There are two unrelated enzymes performing this reaction. One enzyme is a hydratase encoded by *bchE* and is only found in photosynthetic bacteria. The other enzyme

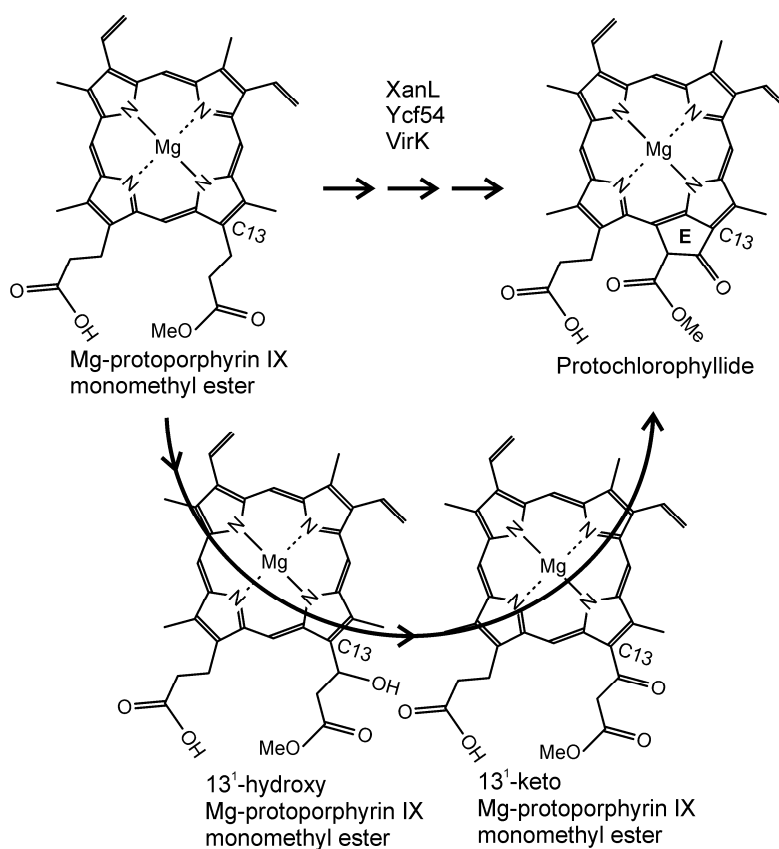


Figure 9 The cyclase reaction forms the fifth ring E of the chlorophyll molecule by converting Mg-protoporphyrin IX monomethyl ester to protochlorophyllide. The reaction proceeds through a C13¹ hydroxy followed by a C13¹ keto intermediate and all three transformations are catalyzed by XanL. An auxiliary protein, Ycf54, is important for the reaction *in vivo*. The enzyme requires electrons and these are provided by reduced ferredoxin. The VirK protein, the barley ortholog of Arabidopsis FdC2, is suggested to be the *in vivo* ferredoxin isoform responsible for donating electrons to XanL.

is a carboxylate bridged diiron monooxidase where the incorporated oxygen atom comes from molecular O₂. As this is the only enzyme found in plants all references to the cyclase from here on refer to the oxygenase.

In barley, the enzyme performing this reaction is encoded by the *Xantha-l* gene (Rzeznicka et al., 2005). In photosynthetic bacteria performing anoxygenic photosynthesis, the gene has been named *acsF* while in the cyanobacterium *Synechocystis*, the two homologs have been named *chlA_I* and *chlA_{II}* (Pinta et al., 2002, Minamizaki et al., 2008). In *Chlamydomonas reinhardtii* there are two homologs which have been named *CRD1* and *CTH1* while in *Arabidopsis* the gene is named *CHL27* (Tottey et al., 2003, Moseley et al., 2002, Moseley et al., 2000). In addition to the cyclase enzyme, plants contain an auxiliary protein, Ycf54, which is required for optimal function of the cyclase enzyme (Hollingshead et al., 2012). The role(s) of Ycf54 are poorly understood but it does not appear to be required for catalysis and there are cyclases in bacteria which are not dependent on Ycf54 (Chen et al., 2017). Studies of the cyclase have been hampered by the lack of catalytically active recombinant enzyme for *in vitro* activity assays – a hurdle which we recently overcame. When XanL is produced recombinantly in *Escherichia coli* it is not catalytically active unless co-expressed with Ycf54 (Stuart et al., 2020). Ycf54 did not co-purify with XanL, further supporting that it is not required for catalysis and suggesting a role in folding or maturation of XanL. The cyclase enzyme has been localized to both the inner envelope and thylakoid membranes in *Arabidopsis* (Tottey et al., 2003) and Ycf54 has been localized in soluble as well as membrane fractions from barley etioplasts (Bollivar et al., 2014). *In vitro* enzymatic assays have showed that the cyclase requires a reductant supplied as NADPH as well as both soluble and membrane fractions from etioplasts which suggested at least one more unidentified component required for cyclase activity (Rzeznicka et al., 2005). Other diiron monooxygenases generally do not utilize NAD(P)H directly so it was likely that the unidentified component is involved in electron transfer from NADPH to the cyclase. In developing a recombinant *in vitro* enzyme activity assay, we recently identified the soluble component as ferredoxin and that FNR is also required to reduce ferredoxin using NADPH (Stuart et al., 2020). This clearly demonstrated that the cyclase is a ferredoxin dependent enzyme. Plants contain a number of ferredoxin isoforms. Based on evidence from barley mutants at the *Viridis-k* and *Xantha-l* loci, the FdC2 ferredoxin isoform encoded by the *Viridis-k* gene in barley is the *in vivo* electron donor to the cyclase (Stuart et al., 2021). The FdC2 isoform is conserved in all plants and cyanobacteria and must thus have an indispensable function. The involvement of FdC2 in chlorophyll biosynthesis is the first clear role for the protein. Previous biochemical characterizations of FdC2 from rice and *Thermosynechococcus elongatus* indicate that FdC2 is most likely reduced by PSI and not via FNR *in vivo* (He et al., 2020, Schorsch et al., 2018). This would imply that most electrons for the cyclase step of chlorophyll biosynthesis are provided by PSI via FdC2 and would explain the cyclase deficient phenotype of the barley *vir-zb.63* mutant which lacks photosystem I activity (Steccanella et al., 2015).

Although a single enzyme is responsible for catalysis, the cyclase reaction proceeds through three steps, each of which is a two-electron oxidation. The enzyme requires a reductant in order to reduce the second oxygen atom at each step. The cyclase proceeds through two intermediates (Wong et al., 1985, Walker et al., 1988). First, the C13¹ carbon is hydroxylated. The hydroxyl group is then oxidized to a carboxy group. Finally, a carbon-carbon bond is formed to produce the E ring which completes the conversion of the red Mg-protoporphyrin IX monomethyl ester to the first green intermediate protochlorophyllide (Figure 9).

Protochlorophyllide oxidoreductase

The next step is the hydrogenation of the double bond between C17 and C18 in ring D to convert protochlorophyllide to divinyl chlorophyllide a (Figure 10). Two enzymes exist that perform this reaction. The first is the dark operative protochlorophyllide oxidoreductase (DPOR) which is related to nitrogenase and consists of three components BchL, BchN, and BchB (Burke et al., 1993). The second is the light dependent protochlorophyllide oxidoreductase (LPOR). Angiosperms only contain the LPOR enzyme and chlorophyll biosynthesis is thus halted at protochlorophyllide when plants are grown in the dark.

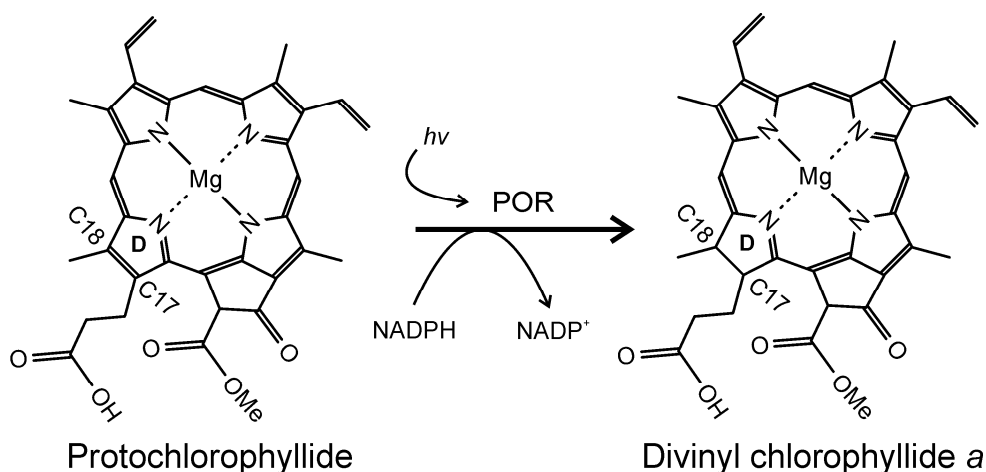


Figure 10 The conversion of protochlorophyllide to divinyl chlorophyllide a in angiosperms is performed by the light dependent protochlorophyllide oxidoreductase (LPOR). In addition to light ($h\nu$), the enzyme also requires NADPH to reduce the double bond between C18 and C17.

Dark grown angiosperms thus undergo a process known as etiolation where the chloroplasts halt development as etioplasts (Solymosi and Schoefs, 2010). The LPOR reaction can thus be considered as the first of the late steps of the Mg branch as flux through the pathway after LPOR only occurs in the light. Etioplasts accumulate a crystalline matrix known as a prolamellar body that contains protochlorophyllide:LPOR complexes. After illumination, the protochlorophyllide is converted to chlorophyllide and the prolamellar body starts to disperse and forms the thylakoid membranes as the plastids develop (Solymosi and Schoefs, 2010).

Divinyl reductase

Plants and most oxygenic phototrophs produce monovinyl chlorophylls where the C8 vinyl group has been reduced to an ethyl group (Figure 11). The enzyme performing this reaction is the divinyl reductase (DVR) and the plant enzyme was first identified by characterization of *Arabidopsis* mutants that accumulated divinyl chlorophylls (Nakanishi et al., 2005, Nagata et al., 2005) and later a rice mutant which also accumulated only divinyl chlorophylls (Wang et al., 2010). The enzyme has a putative C-terminal transmembrane domain (Nakanishi et al., 2005) and enzymatic activity has been isolated from chloroplast membrane fractions (Parham and Rebeiz, 1995).

Substrate specificity of DVR has been investigated using recombinant DVR from the monocots rice and maize, as well as dicots *Arabidopsis* and cucumber (Wang et

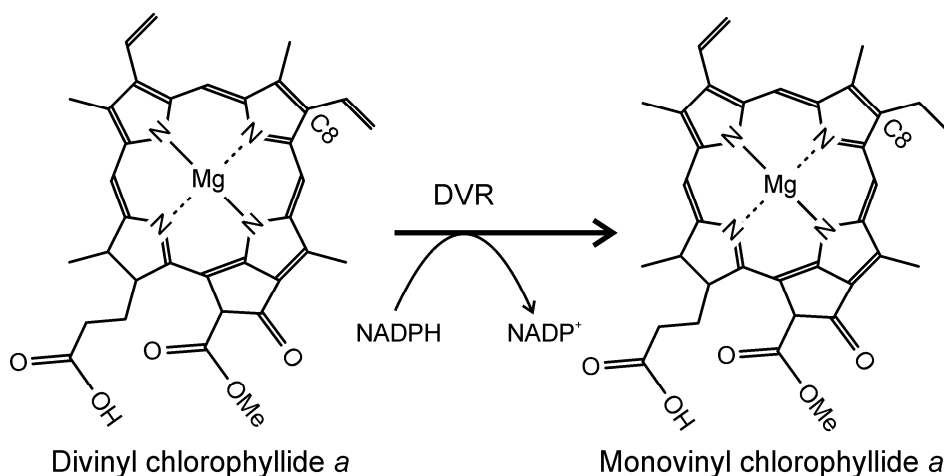


Figure 11 The reduction of the C8 vinyl group to an ethyl group is performed by the divinyl reductase (DVR) enzyme. The reaction requires NADPH.

al., 2013). Both monocots and dicots had highest activity with divinyl chlorophyllide *a* and activity decreased with earlier intermediates in the pathway but was detectable all the way back to divinyl Mg-protoporphyrin IX for monocot DVR and divinyl Mg-protoporphyrin IX monomethyl ester for dicot enzymes. The monocot enzymes had higher activity for a given substrate and also efficiently converted divinyl chlorophyll *a* (Wang et al., 2013) but not divinyl chlorophyll *b* or divinyl chlorophyllide *b* (Wang et al., 2010). In general, chlorophyllide is used synonymously with the monovinyl species if nothing is specified.

Chlorophyll synthase and geranylgeranyl reductase

The final conversion of chlorophyllide to chlorophyll is the formation of a phytol ester on the C17 propionate group (Figure 12). This is performed by the thylakoid bound enzyme chlorophyll synthase (ChlG) which can utilize either phytol pyrophosphate or geranylgeranyl pyrophosphate as a substrate (Oster et al., 1997). The geranylgeranyl moiety is reduced by geranylgeranyl reductase (ChlP) which can occur either before or after esterification to chlorophyllide (Bollivar et al., 1994c). Both oxygens of the ester originate from the C17 propionate group and the reaction is believed to occur by nucleophilic attack by the carboxylate anion on the

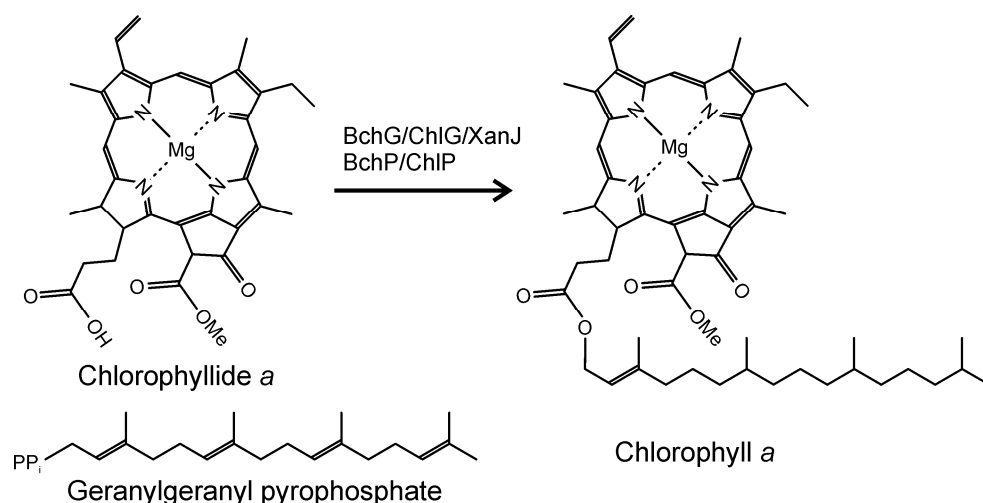


Figure 12 The conversion of chlorophyllide to chlorophyll involves esterification of a phytol “tail” on the C17 propionate group. Ester formation is performed by the chlorophyll synthase enzyme (BchG/ChlG/XanJ) which can use either geranylgeranyl pyrophosphate or phytol pyrophosphate. The enzyme geranylgeranyl reductase (BchP/ChlP) converts the geranylgeranyl chain to a phytol chain either before or after esterification.

pyrophosphoryl activated alcohol (Emery and Akhtar, 1987). Chlorophyll synthase has been shown to interact with the Alb3 membrane insertase and it is suggested that this is part of a mechanism to coordinate chlorophyll biosynthesis with the insertion of chlorophyll into chlorophyll binding proteins in the thylakoid membrane (Proctor et al., 2018, Chidgey et al., 2014).

Chlorophyll synthase from plants can utilize chlorophyllide *a* and chlorophyllide *b* (see chlorophyll cycle below) equally well (Benz and Rüdiger, 1981). Interestingly, ChlG is highly specific for chlorophyllide while BchG is highly specific for bacteriochlorophyllide, although the specificity can be altered with a single amino acid substitution (Kim et al., 2016). In barley, chlorophyll synthase is encoded by the *Xantha-j* gene (Stuart et al., 2024). As part of this thesis, we recently identified this gene using bulk segregant analysis by whole genome sequencing of an F₂-mapping population made by crossing the mutant *xan-j.59* with the cultivar Quench. The mutant *xan-j.64* has an amino-acid substitution that is likely to prohibit binding of the geranylgeranyl/phytyl pyrophosphate substrate and fails to accumulate the XanJ enzyme, suggesting that substrate binding stabilizes the enzyme *in vivo*. Treatment of plants with clomazone, which inhibits isoprenoid biosynthesis, resulted in up-regulation of chlorophyll synthase transcript levels but no accumulation of the protein. This further supports that chlorophyll synthase is degraded *in vivo* when the isoprenoid substrate is lacking. This may be a mechanism to balance the supply of isoprenoids between chlorophyll biosynthesis and carotenoid biosynthesis.

The chlorophyll cycle

The two chlorophylls present in plants, chlorophyll *a* and *b*, differ by having either a methyl group or a formyl group on C7 and can be interconverted via the chlorophyll cycle (Figure 13). The enzyme responsible for synthesis of chlorophyll *b* was first identified in *Chlamydomonas* and named chlorophyll *a* oxygenase (CAO) and contains a Rieske type iron sulfur cluster and a mononuclear iron cluster and is a monooxygenase enzyme (Tanaka et al., 1998). This is consistent with earlier isotope labeling experiments showing that the 7-formyl oxygen atom comes from molecular oxygen (Schneegurt and Beale, 1992, Porra et al., 1994). The enzyme was later renamed to chlorophyllide *a* oxygenase (still CAO) when it was shown that enzymatic activity of the recombinant enzyme was only obtained when chlorophyllide *a* was used as a substrate (Oster et al., 2000). This study also established that CAO is a ferredoxin dependent enzyme and identified 7-hydroxymethyl chlorophyllide *a* as an intermediate. In barley, 15 mutants in the

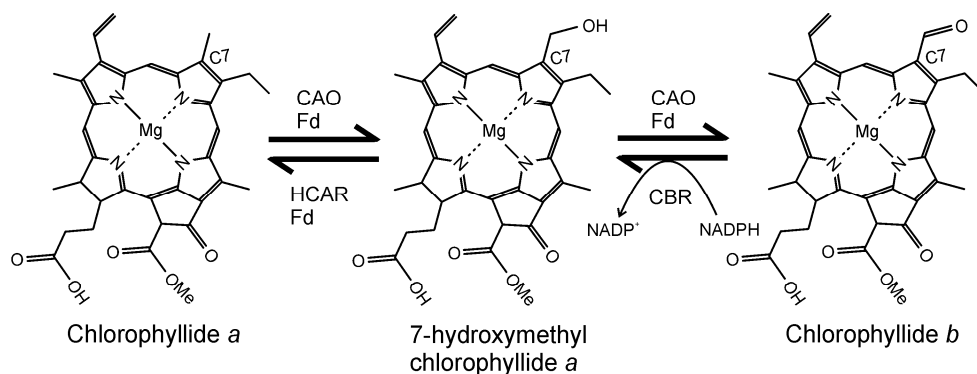


Figure 13 Chlorophyll *b* is synthesized from and converted back to chlorophyll *a* via the chlorophyll cycle. The enzyme chlorophyllide *a* oxygenase (CAO) converts chlorophyllide *a* to chlorophyllide *b* via the C7 hydroxymethyl intermediate and is not able to convert esterified pigments. Chlorophyll *b* is converted back to chlorophyll *a* by the sequential action of chlorophyll(ide) *b* reductase (CBR) followed by 7-hydroxymethyl chlorophyll(ide) *a* reductase HCAR. CBR uses NADPH and HCAR is ferredoxin dependent.

HvCAO gene have been characterized and the gene has been named *fch2* and *chlorina-f2* in different mutant collections (Mueller et al., 2012). Depending on the severity of the mutations the phenotype ranges from increased chlorophyll *a/b* ratios to accumulation of only chlorophyll *a*. Chlorophyll(ide) *b* can be converted back to Chlorophyll(ide) *a* by two sequential reactions catalyzed by chlorophyll(ide) *b* reductase (CBR) followed by 7-hydroxymethyl chlorophyll(ide) *a* reductase (HCAR) (Kusaba et al., 2007, Meguro et al., 2011).

Two genes implicated in the chlorophyll(ide) *b* reductase reaction were first identified in rice by analysis of stay-green mutants which are mutants that do not degrade chlorophyll during senescence. The first gene was named *NON-YELLOW COLORING1* (*NYC1*) and the second was identified by homology and named *NYC1-like* (*NOL*). Both encode short-chain dehydrogenase/reductase enzymes (Kusaba et al., 2007). The *NYC1* protein has three trans-membrane domains. *NOL* only contains one catalytic domain and enzymatic activity has only been obtained with recombinant *NOL* which requires NADPH as a reductant and is able to convert chlorophyll *b*, chlorophyllide *b*, pheophorbide *b*, pheophytin *b*, as well as light harvesting complex II bound chlorophyll *b* to the corresponding 7-hydroxymethyl species. *NOL* has been localized to the stromal surface of thylakoid membranes and while *NYC1* has not been localized it is presumably found in the thylakoid membrane based on the presence of the three transmembrane domains (Sato et al., 2009).

The final conversion of the 7-hydroxymethyl group to a methyl group is performed by HCAR which was first identified in the *Arabidopsis* genome based on its

homology to a *Synechocystis* DVR, which is unrelated to the DVR used by plants (Meguro et al., 2011). The enzyme is ferredoxin dependent and contains two 4Fe-4S clusters as well as binding FAD (Wang and Liu, 2016, Meguro et al., 2011). This inter-conversion between chlorophyll *a* and *b* is thought to be important for adapting to changing light conditions because chlorophyll *b* is only found in the light harvesting complexes, more of which are needed under low light conditions. In addition, the conversion of chlorophyll(ide) *b* to chlorophyll(ide) *a* is required for chlorophyll degradation which proceeds through pheophorbide *a* and is inhibited by pheophorbide *b* (Hörtensteiner et al., 1995). This also explains the stay-green phenotype of CBR and HCAR mutants.

Research results

My thesis work had two main tracks that were partly overlapping. The first track was the study of the cyclase step of chlorophyll biosynthesis. This step has been the least understood step of chlorophyll biosynthesis and all components required for catalysis were not known. Barley as a model organism for studying the cyclase has the advantage of a well-developed *in vitro* enzyme assay using fractionated barley plastids as well as chlorophyll mutants at two loci known to be involved in the reaction. The *Xantha-l* locus was known before the start of the project. *Xantha-l* encodes the main catalytic component of the cyclase. The *Ycf54* component was known but the function was not understood. The *Viridis-k* locus had not yet been identified but shown to be essential for the cyclase reaction. During my thesis work, I identified and analyzed the function of *Viridis-k* (Paper 2) and explored *Ycf54* to develop an *in vitro* cyclase enzyme assay based on recombinant proteins (Paper 1).

The second track of my thesis work was to develop a pipeline for the identification of the causal mutations behind the pigment deficient phenotypes in barley chlorophyll mutants. For many of these mutants little is known and once the causal mutation is identified the mutants can be used to study the biological role of the identified gene. The work on the identification of the *Viridis-k* gene (Paper 2) was important for the development of the pipeline, which was fine-tuned through the identification of the *Xantha-j* gene (Paper 3).

Paper 1

Paper 1 concerns the magnesium protoporphyrin IX monomethyl ester cyclase step of chlorophyll biosynthesis. We utilized *in vitro* enzymatic assays using fractionated barley etioplasts to establish that ferredoxin and FNR are likely involved in transferring electrons from NADPH to the cyclase enzyme XanL. This was first shown by using antibodies against ferredoxin and FNR as inhibitors of enzymatic activity. Using immunoblots, ferredoxin was localized to the soluble plastid fraction while FNR was found in both soluble and membrane fractions. Commercially available spinach ferredoxin was able to substitute for the soluble plastid fraction in enzymatic assays when combined with the membrane fraction. This showed that the cyclase is a ferredoxin dependent enzyme and that the previously unidentified soluble component is ferredoxin. Having identified an

electron donor for the cyclase, we sought to produce the catalytic component, XanL, recombinantly. When XanL was produced in *E. coli* it was not enzymatically active in assays. However, when it was co-expressed with Ycf54 the enzyme was active. Ycf54 did not co-purify with XanL suggesting a role other than being directly involved in catalysis. Although other studies also indicate that the cyclase reaction can occur without Ycf54 and that the auxiliary protein is likely to have regulatory roles (Chen et al., 2021b, Hollingshead et al., 2016), one cannot rule out that it can also have a direct although non-essential role in the reaction. Recombinant XanL and Ycf54 that were produced separately did not result in enzymatic activity when combined *in vitro*. This suggests that Ycf54 may have a role in folding or maturation of XanL *in vivo* although it may also be that the activation of XanL by Ycf54 is dependent on some factor not present in our enzymatic assays.

Paper 2

In paper 2 we identified the gene disrupted in barley *Viridis-k* mutants, which are deficient in the cyclase reaction. These mutants, although not completely blocked at the cyclase step, are able to produce only trace amounts of chlorophyll. The mutants were previously identified as deficient in the cyclase reaction because they accumulate substrate of the cyclase when fed with ALA. We identified the gene encoding the ortholog of the Arabidopsis FdC2 ferredoxin as a candidate gene and showed that the two allelic *vir-k.23* and *vir-k.170* mutants both contain mutations in the gene. That the correct gene had been identified was further confirmed by genetic complementation of the *vir-k.23* mutant by agroinfiltration, which restored chlorophyll synthesis around the infiltration site. Given that the *vir-k* mutants are deficient in the cyclase reaction, as well as the specific ortholog of FdC2, suggests that this may be the main *in vivo* electron donor. A pre-requisite for this is that FdC2/VirK must be able to transfer electrons to the cyclase. We therefore tested and confirmed that VirK can donate electrons to the cyclase in our recombinant *in vitro* assays. In addition, immunoblot analysis of the three available mutants of the cyclase enzyme, *xan-l.35*, *xan-l.81*, and *xan-l.82*, provided further evidence that VirK is associated with the cyclase enzyme in planta. The *xan-l.35* mutant accumulates less XanL protein and *xan-l.81* mutants accumulate more compared to the wild-type and both have missense mutations. The *xan-l.82* mutant has a nonsense mutation and thus accumulates no protein. In these mutants and the wild type, the protein levels of Ycf54 and VirK follow the levels of XanL which further supports that VirK is likely important for the cyclase since *in vivo* VirK levels depend on the cyclase levels. Furthermore, both *xan-l.81* and *xan-l.82* are completely blocked at the cyclase step so the changes in Ycf54 and VirK are likely to be an effect of altered cyclase enzyme levels rather than some general effect due to a block in chlorophyll biosynthesis at the cyclase step. No significant difference

in *Vir-k* mRNA levels were detected, further suggesting that the altered protein levels are not due to transcriptional changes but rather by post-transcriptional processes. VirK orthologs are conserved among oxygenic phototrophs which suggests that the protein has at least one function which is essential for oxygenic photosynthesis.

Paper 3

The use of genetic markers in plant breeding provides fast and efficient screening of breeding lines by genetic approaches compared to phenotypic methods. Preferably the genetic marker should be in the casual gene regulating a desired trait. In paper 3, we describe a relatively cheap and simple method for identification of mutated genes. We applied this on mutant *xan-j.59* but the method should be applicable for any locus where a mutant with a clear phenotype is available. The method is a variant of bulk-segregant analysis, which we apply on F₂-mapping populations. In short, a mapping population is produced and planted. In case of a recessive mutation, leaves of the homozygous mutant plants are collected and pooled. Genomic DNA is isolated from the pooled leaves and sent for genomic sequencing. SNPs linked to the mutation are searched for in the DNA sequence data, which reveals the chromosomal location of the mutation. The DNA sequence data is also likely to contain the DNA sequence information of the mutated gene. Therefore, the data is filtered for homozygous mutations in the identified linked region. In case of the *xan-j.59* sequence data, 4202 homozygous SNPs were found in the mapped region. Only two SNPs were located in high-confidence genes and caused changes of amino-acid residues in the corresponding proteins. One of the proteins was annotated as SKI family transcriptional corepressor 1 and the other as chlorophyll synthase. The later is a likely candidate gene of a yellow *xan-j.59* mutant deficient in chlorophyll biosynthesis since chlorophyll synthase catalyzes the last enzymatic step in the pathway of chlorophyll. The identification of *Xan-j* as the chlorophyll synthase gene was confirmed by sequencing of the other allelic mutants *xan-j.19* and *xan-j.64*. In *xan-j.19*, a one base pair deletion was identified in exon 13, which causes a frameshift in the gene and a truncated protein that contains 340 native amino-acid residues followed by 5 incorrect residues. Mutant *xan-j.59* has an early nonsense mutation that results in a peptide of 53 residues once the chloroplast transit peptide is removed. The missense mutation in *xan-j.64* causes an S212F substitution. An AlphaFold generated structure of the barley chlorophyll synthase showed the location of Ser-212 in relation to the binding site of the two substrates. Our model suggested that Ser-212 abolishes binding of the geranylgeranyl pyrophosphate/phytyl pyrophosphate substrate. By watering barley plants with clomazone, which is an inhibitor of chloroplastic isoprenoid biosynthesis, a white phenotype was obtained. By western blot analyses, no XanJ

protein could be detected in the white plants. We suggested that there might be a requirement of geranylgeranyl pyrophosphate/phytyl pyrophosphate binding to the active site of chlorophyll synthase to stabilize the enzyme. We further suggested that chlorophyll synthase might function as a sensor for coordinating chlorophyll and isoprenoid biosynthesis.

Future perspectives

Determining the mechanism behind how Ycf54 activates the cyclase enzyme is an important future direction of study that we made possible by the establishment of a recombinant production system. Although the system needs to be improved such that the components can be purified to homogeneity, it also opens up the possibility of further biochemical and structural studies. One interesting avenue to explore would be site-directed mutagenesis to produce variants of XanL that are unable to perform some of the reaction steps. Such mutants would be valuable for producing the reaction intermediates in usable quantities and studying the reaction mechanism of individual steps. The diiron cluster of XanL is coordinated by four glutamate residues and two histidine residues. A potential approach would be replacing the glutamates, individually and in all combinations, with aspartates in the hope that small disruptions in the diiron clusters localization may disrupt specific reaction steps. Chen et al. later published a recombinant assay system for the cyclase from the purple nonsulfur bacteria *Rubrivivax gelatinosus* (Chen et al., 2021a). Their system could also be used for studies of the cyclase reaction mechanism as it is likely to be conserved. However, as the *R. gelatinosus* cyclase is not dependent on Ycf54, their system is not suitable for studies on how Ycf54 regulates cyclase function in plants and cyanobacteria.

At first glance the results from paper 1 and paper 2 may seem contradictory since the first paper identifies canonical ferredoxin as the electron donor to the cyclase while paper 2 postulates that FdC2/VirK is the specific electron donor. The key difference is that in paper 1, ferredoxin was identified as the electron donor in *in vitro* assays with plastid extracts and also used in assays with the recombinant cyclase enzyme. These experiments are run under non-physiological conditions and tell us that regular ferredoxin can donate electrons to the cyclase enzyme. The assertion in paper 2 that VirK is likely the main electron donor comes from *in vivo* experiments utilizing barley mutants that are deficient in the cyclase enzyme or VirK. Mutations in either of these genes results in accumulation of cyclase substrate when plants are fed ALA (Steccanella et al., 2015). That combined with immunoblots showing that *in vivo* VirK accumulation is dependent on cyclase enzyme levels provides a strong indication that VirK is the main *in vivo* electron donor to the cyclase. It is likely that most, if not all, chloroplastic ferredoxins can donate electrons to the cyclase *in vitro*. We have tested and confirmed that this is possible with a number of different recombinant barley ferredoxins (Unpublished) and in assays with the recombinant *R. gelatinosus* cyclase reported by Chen et al., they used spinach or recombinant *Anabaena* ferredoxin. Future studies could thus use recombinant systems to test kinetics or binding energies for the different

chloroplastic ferredoxins which may provide some indication of the specificity of VirK as the donor to the cyclase enzyme. However, it is important to keep in mind that such experiments would not be representative of physiological conditions. *In planta*, there are likely to be many competing interaction partners so the efficiency with which a specific ferredoxin can donate electrons to the cyclase will be strongly influenced by the levels of competing electron acceptors. Therefore, such experiments would ideally be combined with *in vivo* experiments. For example, a systematic knockout of all ferredoxin genes could be done to confirm that the cyclase deficient phenotype is specific to FdC2/VirK mutants. Although beyond the scope of the present thesis work, such a collection of mutants would additionally be a valuable resource for future studies of specific roles of individual ferredoxins.

While VirK is likely the main *in vivo* electron donor to the cyclase enzyme it remains to be determined where those electrons originate from. One possibility is that the main electron donor to VirK *in vivo* is photosystem I. In the dark, chlorophyll biosynthesis halts, at least in angiosperms, at the step after the cyclase since POR is light dependent. Having a dedicated ferredoxin mainly reduced by PSI would thus decrease competition for reducing equivalents from NADPH in the dark when cyclase activity is not needed. Consistent with this idea, ALA feeding of the *vir-zb.63* mutant, that is deficient in PSI activity, indicates a deficiency at the cyclase reaction, particularly in the light (Steccanella et al., 2015). Further studies of this mutant could provide valuable insights into the involvement of photoreduced VirK in the cyclase reaction.

Compared to the cyclase enzyme, considerably more is known about the chlorophyll synthase. Several biochemical studies have been published in which an enzymatic assay was established based on recombinant enzymes (Oster et al. 1997). Substrate specificity (Oster et al. 1997), cofactor requirement and essential amino-acid residues were analyzed (Schmid et al. 2001). Also, the substrate binding order was determined (Schmid et al. 2002). It is striking that structural data of the chlorophyll synthase is not available, which is most likely because the protein is a membrane bound enzyme and probably challenging to crystalize. We used AlphaFold to predict a 3D model of the barley chlorophyll synthase, which can be regarded as a theoretical alternative to x-ray crystallography. In order to get more detailed information about the chlorophyll synthase and its reaction mechanism, future studies should focus on structural studies since that would reveal key information about substrate binding and important amino-acid residues in the active site and of structural importance.

References

- ALAWADY, A., RESKI, R., YARONSKAYA, E. & GRIMM, B. 2005. Cloning and expression of the tobacco *CHLM* sequence encoding Mg protoporphyrin IX methyltransferase and its interaction with Mg chelatase. *Plant Mol Biol*, 57, 679-691.
- AXELSSON, E., LUNDQVIST, J., SAWICKI, A., NILSSON, S., SCHRÖDER, I., AL-KARADAGHI, S., WILLOWS, R. D. & HANSSON, M. 2006. Recessiveness and dominance in barley mutants deficient in Mg-chelatase subunit D, an AAA protein involved in chlorophyll biosynthesis. *Plant Cell*, 18, 3606-3616.
- BEALE, S. I. & CASTELFRANCO, P. A. 1973. ¹⁴C incorporation from exogenous compounds into delta-aminolevulinic acid by greening cucumber cotyledons. *Biochem Biophys Res Commun*, 52, 143-149.
- BEALE, S. I., GOUGH, S. P. & GRANICK, S. 1975. Biosynthesis of delta-aminolevulinic acid from the intact carbon skeleton of glutamic acid in greening barley. *Proc Natl Acad Sci U S A*, 72, 2719-2723.
- BENZ, J. & RÜDIGER, W. 1981. Chlorophyll biosynthesis: various chlorophyllides as exogenous substrates for chlorophyll synthetase. *Z Naturforsch*, 36c, 51-57.
- BLOCK, M. A., TEWARI, A. K., ALBRIEUX, C., MARÉCHAL, E. & JOYARD, J. 2002. The plant S-adenosyl-L-methionine:Mg-protoporphyrin IX methyltransferase is located in both envelope and thylakoid chloroplast membranes. *Eur J Biochem*, 269, 240-248.
- BOESE, Q. F., SPANO, A. J., LI, J. M. & TIMKO, M. P. 1991. Aminolevulinic acid dehydratase in pea (*Pisum sativum* L.). Identification of an unusual metal-binding domain in the plant enzyme. *J Biol Chem*, 266, 17060-17066.
- BOLLIVAR, D., BRAUMANN, I., BERENDT, K., GOUGH, S. P. & HANSSON, M. 2014. The Ycf54 protein is part of the membrane component of Mg-protoporphyrin IX monomethyl ester cyclase from barley (*Hordeum vulgare* L.). *FEBS J*, 281, 2377-2386.
- BOLLIVAR, D. W., JIANG, Z. Y., BAUER, C. E. & BEALE, S. I. 1994a. Heterologous expression of the *bchM* gene product from *Rhodobacter capsulatus* and demonstration that it encodes S-adenosyl-L-methionine:Mg-protoporphyrin IX methyltransferase. *J Bacteriol*, 176, 5290-5296.
- BOLLIVAR, D. W., SUZUKI, J. Y., BEATTY, J. T., DOBROWOLSKI, J. M. & BAUER, C. E. 1994b. Directed mutational analysis of bacteriochlorophyll *a* biosynthesis in *Rhodobacter capsulatus*. *J Mol Biol*, 237, 622-640.
- BOLLIVAR, D. W., WANG, S., ALLEN, J. P. & BAUER, C. E. 1994c. Molecular genetic analysis of terminal steps in bacteriochlorophyll *a* biosynthesis: characterization of a *Rhodobacter capsulatus* strain that synthesizes geranylgeraniol-esterified bacteriochlorophyll *a*. *Biochemistry*, 33, 12763-12768.

- BRYANT, D. A., HUNTER, C. N. & WARREN, M. J. 2020. Biosynthesis of the modified tetrapyrroles-the pigments of life. *J Biol Chem*, 295, 6888-6925.
- BURKE, D. H., HEARST, J. E. & SIDOW, A. 1993. Early evolution of photosynthesis: clues from nitrogenase and chlorophyll iron proteins. *Proc Natl Acad Sci U S A*, 90, 7134-7138.
- CASPY, I., BOROVIKOVA-SHEINKER, A., KLAIMAN, D., SHKOLNISKY, Y. & NELSON, N. 2020. The structure of a triple complex of plant photosystem I with ferredoxin and plastocyanin. *Nat Plants*, 6, 1300-1305.
- CHEN, G. E., ADAMS, N. B. P., JACKSON, P. J., DICKMAN, M. J. & HUNTER, C. N. 2021a. How the O₂-dependent Mg-protoporphyrin monomethyl ester cyclase forms the fifth ring of chlorophylls. *Nat Plants*, 7, 365-375.
- CHEN, G. E., CANNIFFE, D. P. & HUNTER, C. N. 2017. Three classes of oxygen-dependent cyclase involved in chlorophyll and bacteriochlorophyll biosynthesis. *Proc Natl Acad Sci U S A*, 114, 6280-6285.
- CHEN, G. E., HITCHCOCK, A., MAREŠ, J., GONG, Y., TICHÝ, M., PILNÝ, J., KOVÁŘOVÁ, L., ZDVIHALOVÁ, B., XU, J., HUNTER, C. N. & SOBOTKA, R. 2021b. Evolution of Ycf54-independent chlorophyll biosynthesis in cyanobacteria. *Proc Natl Acad Sci U S A*, 118, e2024633118.
- CHIDGEY, J. W., LINHARTOVÁ, M., KOMENDA, J., JACKSON, P. J., DICKMAN, M. J., CANNIFFE, D. P., KONÍK, P., PILNÝ, J., HUNTER, C. N. & SOBOTKA, R. 2014. A cyanobacterial chlorophyll synthase-HliD complex associates with the Ycf39 protein and the YidC/Alb3 insertase. *Plant Cell*, 26, 1267-1279.
- DAVISON, P. A., SCHUBERT, H. L., REID, J. D., IORG, C. D., HEROUX, A., HILL, C. P. & HUNTER, C. N. 2005. Structural and biochemical characterization of Gun4 suggests a mechanism for its role in chlorophyll biosynthesis. *Biochemistry*, 44, 7603-7612.
- EMERY, V. C. & AKHTAR, M. 1987. Mechanistic studies on the phytylation and methylation steps in bacteriochlorophyll *a* biosynthesis: an application of the ¹⁸O-induced isotope effect in ¹³C NMR. *Biochemistry*, 26, 1200-1208.
- FUESLER, T. P., WRIGHT, L. A. & CASTELFRANCO, P. A. 1981. Properties of magnesium chelatase in greening etioplasts: Metal ion specificity and effect of substrate concentrations. *Plant Physiol*, 67, 246-249.
- GRIMM, B. 1990. Primary structure of a key enzyme in plant tetrapyrrole synthesis: glutamate 1-semialdehyde aminotransferase. *Proc Natl Acad Sci U S A*, 87, 4169-4173.
- HAHN, A., VONCK, J., MILLS, D. J., MEIER, T. & KÜHLBRANDT, W. 2018. Structure, mechanism, and regulation of the chloroplast ATP synthase. *Science*, 360, 620.
- HAMILTON, T. L. 2019. The trouble with oxygen: The ecophysiology of extant phototrophs and implications for the evolution of oxygenic photosynthesis. *Free Radic Biol Med*, 140, 233-249.
- HANSSON, A., KANNANGARA, C. G., VON WETTSTEIN, D. & HANSSON, M. 1999. Molecular basis for semidominance of missense mutations in the

- XANTHA-H (42-kDa) subunit of magnesium chelatase. *Proc Natl Acad Sci U S A*, 96, 1744-1749.
- HANSSON, M., YOUSSEF, H. M., ZAKHRABEKOVA, S., STUART, D., SVENSSON, J. T., DOCKTER, C., STEIN, N., WAUGH, R., LUNDQVIST, U. & FRANCKOWIAK, J. 2024. A guide to barley mutants. *Hereditas*, 161, 11.
- HE, L., LI, M., QIU, Z., CHEN, D., ZHANG, G., WANG, X., CHEN, G., HU, J., GAO, Z., DONG, G., REN, D., SHEN, L., ZHANG, Q., GUO, L., QIAN, Q., ZENG, D. & ZHU, L. 2020. Primary leaf-type ferredoxin 1 participates in photosynthetic electron transport and carbon assimilation in rice. *Plant J*, 104, 44-58.
- HENNINGSEN, K. W., BOYNTON, J. E. & WETTSTEIN, D. V. 1993. Mutants at *xantha* and *albina* loci in relation to chloroplast biogenesis in barley (*Hordeum vulgare* L.). *The Royal Danish Academy of Sciences and Letters*, 42, 1-349.
- HERTLE, A. P., BLUNDER, T., WUNDER, T., PESARESI, P., PRIBIL, M., ARMBRUSTER, U. & LEISTER, D. 2013. PGRL1 is the elusive ferredoxin-plastoquinone reductase in photosynthetic cyclic electron flow. *Mol Cell*, 49, 511-523.
- HINCHIGERI, S. B., HUNDLE, B. & RICHARDS, W. R. 1997. Demonstration that the BchH protein of *Rhodobacter capsulatus* activates S-adenosyl-L-methionine:magnesium protoporphyrin IX methyltransferase. *FEBS Lett*, 407, 337-342.
- HOHMANN-MARRIOTT, M. F. & BLANKENSHIP, R. E. 2011. Evolution of photosynthesis. *Annu Rev Plant Biol*, 62, 515-548.
- HOLLINGSHEAD, S., KOPECNA, J., ARMSTRONG, D. R., BUCINSKA, L., JACKSON, P. J., CHEN, G. E., DICKMAN, M. J., WILLIAMSON, M. P., SOBOTKA, R. & HUNTER, C. N. 2016. Synthesis of chlorophyll-binding proteins in a fully segregated delta-*yef54* strain of the cyanobacterium *Synechocystis* PCC 6803. *Front Plant Sci*, 7, 292.
- HOLLINGSHEAD, S., KOPECNA, J., JACKSON, P. J., CANNIFFE, D. P., DAVISON, P. A., DICKMAN, M. J., SOBOTKA, R. & HUNTER, C. N. 2012. Conserved chloroplast open-reading frame *yef54* is required for activity of the magnesium protoporphyrin monomethylester oxidative cyclase in *Synechocystis* PCC 6803. *J Biol Chem*, 287, 27823-27833.
- HUANG, D. D., WANG, W. Y., GOUGH, S. P. & KANNANGARA, C. G. 1984. Delta-aminolevulinic acid-synthesizing enzymes need an RNA moiety for activity. *Science*, 225, 1482-1484.
- HÖRTENSTEINER, S., VICENTINI, F. & MATILE, P. 1995. Chlorophyll breakdown in senescent cotyledons of rape, *Brassica napus* L.: Enzymatic cleavage of phaeophorbide *a in vitro*. *New Phytol*, 129, 237-246.
- JACOBS, J. M. & JACOBS, N. J. 1987. Oxidation of protoporphyrinogen to protoporphyrin, a step in chlorophyll and haem biosynthesis. Purification and partial characterization of the enzyme from barley organelles. *Biochem J*, 244, 219-224.
- JENSEN, P. E., GIBSON, L. C., HENNINGSEN, K. W. & HUNTER, C. N. 1996a. Expression of the *chlI*, *chlD*, and *chlH* genes from the cyanobacterium *Synechocystis* PCC6803 in *Escherichia coli* and demonstration that the three

- cognate proteins are required for magnesium-protoporphyrin chelatase activity. *J Biol Chem*, 271, 16662-16667.
- JENSEN, P. E., GIBSON, L. C. & HUNTER, C. N. 1998. Determinants of catalytic activity with the use of purified I, D and H subunits of the magnesium protoporphyrin IX chelatase from *Synechocystis* PCC6803. *Biochem J*, 334, 335-344.
- JENSEN, P. E., WILLOWS, R. D., PETERSEN, B. L., VOTHKNECHT, U. C., STUMMANN, B. M., KANNANGARA, C. G., VON WETTSTEIN, D. & HENNINGSEN, K. W. 1996b. Structural genes for Mg-chelatase subunits in barley: *Xantha-f*, *-g* and *-h*. *Mol Gen Genet*, 250, 383-394.
- JONES, R. M. & JORDAN, P. M. 1994. Purification and properties of porphobilinogen deaminase from *Arabidopsis thaliana*. *Biochem J*, 299, 895-902.
- KANNANGARA, C. G. & GOUGH, S. P. 1978. Biosynthesis of delta-aminolevulinate in greening barley leaves: Glutamate 1-semialdehyde aminotransferase. *Carlsberg Res Commun*, 43, 185-194.
- KIM, E. J., KIM, H. & LEE, J. K. 2016. The photoheterotrophic growth of bacteriochlorophyll synthase-deficient mutant of *Rhodobacter sphaeroides* is restored by I44F mutant chlorophyll synthase of *Synechocystis* sp. PCC 6803. *J Microbiol Biotechnol*, 26, 959-966.
- KRUSE, E., MOCK, H. P. & GRIMM, B. 1995. Coproporphyrinogen III oxidase from barley and tobacco - sequence analysis and initial expression studies. *Planta*, 196, 796-803.
- KUSABA, M., ITO, H., MORITA, R., IIDA, S., SATO, Y., FUJIMOTO, M., KAWASAKI, S., TANAKA, R., HIROCHIKA, H., NISHIMURA, M. & TANAKA, A. 2007. Rice NON-YELLOW COLORING1 is involved in light-harvesting complex II and grana degradation during leaf senescence. *Plant Cell*, 19, 1362-1375.
- LAKE, V., OLSSON, U., WILLOWS, R. D. & HANSSON, M. 2004. ATPase activity of magnesium chelatase subunit I is required to maintain subunit D *in vivo*. *Eur J Biochem*, 271, 2182-2188.
- LARKIN, R. M., ALONSO, J. M., ECKER, J. R. & CHORY, J. 2003. GUN4, a regulator of chlorophyll synthesis and intracellular signaling. *Science*, 299, 902-906.
- LERMONTOVA, I., KRUSE, E., MOCK, H. P. & GRIMM, B. 1997. Cloning and characterization of a plastidal and a mitochondrial isoform of tobacco protoporphyrinogen IX oxidase. *Proc Natl Acad Sci U S A*, 94, 8895-8900.
- LUNDQVIST, J., ELMLUND, H., PETERSON WULFF, R., BERGLUND, L., ELMLUND, D., EMANUELSSON, C., HEBERT, H., WILLOWS, R. D., HANSSON, M., LINDAHL, M. & AL-KARADAGHI, S. 2010. ATP-induced conformational dynamics in the AAA+ motor unit of magnesium chelatase. *Structure*, 18, 354-365.
- LUNDQVIST, U. 1992. *Mutation research in barley*. PhD, Swedish University of Agricultural Sciences.
- MADSEN, O., SANDAL, L., SANDAL, N. N. & MARCKER, K. A. 1993. A soybean coproporphyrinogen oxidase gene is highly expressed in root nodules. *Plant Mol Biol*, 23, 35-43.

- MASCHER, M., GUNDLACH, H., HIMMELBACH, A., BEIER, S., TWARDZIOK, S. O., WICKER, T., RADCHUK, V., DOCKTER, C., HEDLEY, P. E., RUSSELL, J., BAYER, M., RAMSAY, L., LIU, H., HABERER, G., ZHANG, X. Q., ZHANG, Q., BARRERO, R. A., LI, L., TAUDIEN, S., GROTH, M., FELDER, M., HASTIE, A., ŠIMKOVÁ, H., STAŇKOVÁ, H., VRÁNA, J., CHAN, S., MUÑOZ-AMATRIAIN, M., OUNIT, R., WANAMAKER, S., BOLSER, D., COLMSEE, C., SCHMUTZER, T., ALIYEVA-SCHNORR, L., GRASSO, S., TANSKANEN, J., CHAILYAN, A., SAMPATH, D., HEAVENS, D., CLISSOLD, L., CAO, S., CHAPMAN, B., DAI, F., HAN, Y., LI, H., LI, X., LIN, C., MCCOOKE, J. K., TAN, C., WANG, P., WANG, S., YIN, S., ZHOU, G., POLAND, J. A., BELLGARD, M. I., BORISJUK, L., HOUBEN, A., DOLEŽEL, J., AYLING, S., LONARDI, S., KERSEY, P., LANGRIDGE, P., MUEHLBAUER, G. J., CLARK, M. D., CACCAMO, M., SCHULMAN, A. H., MAYER, K. F. X., PLATZER, M., CLOSE, T. J., SCHOLZ, U., HANSSON, M., ZHANG, G., BRAUMANN, I., SPANNAGL, M., LI, C., WAUGH, R. & STEIN, N. 2017. A chromosome conformation capture ordered sequence of the barley genome. *Nature*, 544, 427-433.
- MEGURO, M., ITO, H., TAKABAYASHI, A., TANAKA, R. & TANAKA, A. 2011. Identification of the 7-hydroxymethyl chlorophyll *a* reductase of the chlorophyll cycle in *Arabidopsis*. *Plant Cell*, 23, 3442-3453.
- MINAMIZAKI, K., MIZOGUCHI, T., GOTO, T., TAMIAKI, H. & FUJITA, Y. 2008. Identification of two homologous genes, *chlAI* and *chlAII*, that are differentially involved in isocyclic ring formation of chlorophyll *a* in the cyanobacterium *Synechocystis* sp. PCC 6803. *J Biol Chem*, 283, 2684-2692.
- MOCK, H. P., TRAINOTTI, L., KRUSE, E. & GRIMM, B. 1995. Isolation, sequencing and expression of cDNA sequences encoding uroporphyrinogen decarboxylase from tobacco and barley. *Plant Mol Biol*, 28, 245-256.
- MOSELEY, J., QUINN, J., ERIKSSON, M. & MERCHANT, S. 2000. The *Crd1* gene encodes a putative di-iron enzyme required for photosystem I accumulation in copper deficiency and hypoxia in *Chlamydomonas reinhardtii*. *EMBO J*, 19, 2139-2151.
- MOSELEY, J. L., PAGE, M. D., ALDER, N. P., ERIKSSON, M., QUINN, J., SOTO, F., THEG, S. M., HIPPLER, M. & MERCHANT, S. 2002. Reciprocal expression of two candidate di-iron enzymes affecting photosystem I and light-harvesting complex accumulation. *Plant Cell*, 14, 673-688.
- MUELLER, A. H., DOCKTER, C., GOUGH, S. P., LUNDQVIST, U., VON WETTSTEIN, D. & HANSSON, M. 2012. Characterization of mutations in barley *fch2* encoding chlorophyllide *a* oxygenase. *Plant Cell Physiol*, 53, 1232-1246.
- NAGATA, N., TANAKA, R., SATOH, S. & TANAKA, A. 2005. Identification of a vinyl reductase gene for chlorophyll synthesis in *Arabidopsis thaliana* and implications for the evolution of Prochlorococcus species. *Plant Cell*, 17, 233-240.
- NAKANISHI, H., NOZUE, H., SUZUKI, K., KANEKO, Y., TAGUCHI, G. & HAYASHIDA, N. 2005. Characterization of the *Arabidopsis thaliana* mutant *pcb2* which accumulates divinyl chlorophylls. *Plant Cell Physiol*, 46, 467-473.

- OSTER, U., BAUER, C. E. & RÜDIGER, W. 1997. Characterization of chlorophyll *a* and bacteriochlorophyll *a* synthases by heterologous expression in *Escherichia coli*. *J Biol Chem*, 272, 9671-9676.
- OSTER, U., TANAKA, R., TANAKA, A. & RÜDIGER, W. 2000. Cloning and functional expression of the gene encoding the key enzyme for chlorophyll *b* biosynthesis (CAO) from *Arabidopsis thaliana*. *Plant J*, 21, 305-310.
- PARHAM, R. & REBEIZ, C. A. 1995. Chloroplast biogenesis 72: a [4-vinyl]chlorophyllide *a* reductase assay using divinyl chlorophyllide *a* as an exogenous substrate. *Anal Biochem*, 231, 164-169.
- PINTA, V., PICAUD, M., REISS-HUSSON, F. & ASTIER, C. 2002. *Rubrivivax gelatinosus* *acsF* (previously *orf358*) codes for a conserved, putative binuclear-iron-cluster-containing protein involved in aerobic oxidative cyclization of Mg-protoporphyrin IX monomethylester. *J Bacteriol*, 184, 746-753.
- PORRA, R. J., SCHÄFER, W., CMIEL, E., KATHEDER, I. & SCHEER, H. 1994. The derivation of the formyl-group oxygen of chlorophyll *b* in higher plants from molecular oxygen. Achievement of high enrichment of the 7-formyl-group oxygen from $^{18}\text{O}_2$ in greening maize leaves. *Eur J Biochem*, 219, 671-679.
- PROCTOR, M. S., CHIDGEY, J. W., SHUKLA, M. K., JACKSON, P. J., SOBOTKA, R., HUNTER, C. N. & HITCHCOCK, A. 2018. Plant and algal chlorophyll synthases function in *Synechocystis* and interact with the YidC/Alb3 membrane insertase. *FEBS Lett*, 592, 3062-3073.
- RZEZNICKA, K., WALKER, C. J., WESTERGREN, T., KANNANGARA, C. G., VON WETTSTEIN, D., MERCHANT, S., GOUGH, S. P. & HANSSON, M. 2005. *Xantha-1* encodes a membrane subunit of the aerobic Mg-protoporphyrin IX monomethyl ester cyclase involved in chlorophyll biosynthesis. *Proc Natl Acad Sci U S A*, 102, 5886-5891.
- SAGAN, L. 1967. On the origin of mitosing cells. *J Theor Biol*, 14, 255-274.
- SATO, Y., MORITA, R., KATSUMA, S., NISHIMURA, M., TANAKA, A. & KUSABA, M. 2009. Two short-chain dehydrogenase/reductases, NON-YELLOW COLORING 1 and NYC1-LIKE, are required for chlorophyll *b* and light-harvesting complex II degradation during senescence in rice. *Plant J*, 57, 120-131.
- SAWICKI, A. & WILLOWS, R. D. 2007. S-adenosyl-L-methionine:magnesium-protoporphyrin IX *O*-methyltransferase from *Rhodobacter capsulatus*: mechanistic insights and stimulation with phospholipids. *Biochem J*, 406, 469-478.
- SCHAUMBURG, A., SCHNEIDER-POETSCH, H. A. & ECKERSKORN, C. 1992. Characterization of plastid 5-aminolevulinic acid dehydratase (ALAD; EC 4.2.1.24) from spinach (*Spinacia oleracea* L.) by sequencing and comparison with non-plant ALAD enzymes. *Z Naturforsch*, 47c, 77-84.
- SCHNEEGURT, M. A. & BEALE, S. I. 1992. Origin of the chlorophyll *b* formyl oxygen in *Chlorella vulgaris*. *Biochemistry*, 31, 11677-11683.
- SCHORSCH, M., KRAMER, M., GOSS, T., EISENHUT, M., ROBINSON, N., OSMAN, D., WILDE, A., SADAF, S., BRÜCKLER, H., WALDER, L., SCHEIBE, R., HASE, T. & HANKE, G. T. 2018. A unique ferredoxin acts as a player in the

- low-iron response of photosynthetic organisms. *Proc Natl Acad Sci U S A*, 115, E12111-E12120.
- SCHWARTZ, R. M. & DAYHOFF, M. O. 1978. Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. *Science*, 199, 395-403.
- SCHÖN, A., KRUPP, G., GOUGH, S., BERRY-LOWE, S., KANNANGARA, C. G. & SÖLL, D. 1986. The RNA required in the first step of chlorophyll biosynthesis is a chloroplast glutamate tRNA. *Nature*, 322, 281-284.
- SHEMIN, D., RUSSELL, C. S. & ABRAMSKY, T. 1955. The succinate-glycine cycle. I. The mechanism of pyrrole synthesis. *J Biol Chem*, 215, 613-626.
- SHEPHERD, M., MCLEAN, S. & HUNTER, C. N. 2005. Kinetic basis for linking the first two enzymes of chlorophyll biosynthesis. *FEBS J*, 272, 4532-4539.
- SHIEH, J., MILLER, G. W. & PSENAK, M. 1978. Properties of S-adenosyl-L-methionine-magnesium-protoporphyrin IX methyltransferase from barley. *Plant Cell Physiol*, 19, 1051-1059.
- SHINOHARA, F., KURISU, G., HANKE, G., BOWSHER, C., HASE, T. & KIMATA-ARIGA, Y. 2017. Structural basis for the isotype-specific interactions of ferredoxin and ferredoxin: NADP⁺ oxidoreductase: an evolutionary switch between photosynthetic and heterotrophic assimilation. *Photosynth Res*, 134, 281-289.
- SMITH, C. A., SUZUKI, J. Y. & BAUER, C. E. 1996. Cloning and characterization of the chlorophyll biosynthesis gene *chlM* from *Synechocystis* PCC 6803 by complementation of a bacteriochlorophyll biosynthesis mutant of *Rhodobacter capsulatus*. *Plant Mol Biol*, 30, 1307-1314.
- SOLYMOSSI, K. & SCHOEFS, B. 2010. Etioplast and etio-chloroplast formation under natural conditions: the dark side of chlorophyll biosynthesis in angiosperms. *Photosynth Res*, 105, 143-166.
- STAEHELIN, L. A. 2003. Chloroplast structure: from chlorophyll granules to supra-molecular architecture of thylakoid membranes. *Photosynth Res*, 76, 185-196.
- STECCANELLA, V., HANSSON, M. & JENSEN, P. E. 2015. Linking chlorophyll biosynthesis to a dynamic plastoquinone pool. *Plant Physiol Biochem*, 97, 207-216.
- STROEBEL, D., CHOQUET, Y., POPOT, J. L. & PICOT, D. 2003. An atypical haem in the cytochrome *b₆f* complex. *Nature*, 426, 413-418.
- STUART, D., SANDSTRÖM, M., YOUSSEF, H. M., ZAKHRABEKOVA, S., JENSEN, P. E., BOLLIVAR, D. & HANSSON, M. 2021. Barley *Viridis-k* links an evolutionarily conserved C-type ferredoxin to chlorophyll biosynthesis. *Plant Cell*, 33, 2834-2849.
- STUART, D., SANDSTRÖM, M., YOUSSEF, H. M., ZAKHRABEKOVA, S., JENSEN, P. E., BOLLIVAR, D. W. & HANSSON, M. 2020. Aerobic barley Mg-protoporphyrin IX monomethyl ester cyclase is powered by electrons from ferredoxin. *Plants*, 9, 1157.
- STUART, D., ZAKHRABEKOVA, S., EGEVANG JØRGENSEN, M., DOCKTER, C. & HANSSON, M. 2024. A pipeline for identification of causal mutations in barley identifies Xantha-j as the chlorophyll synthase gene. *Plant Physiol*, In press.

- SU, X., MA, J., WEI, X., CAO, P., ZHU, D., CHANG, W., LIU, Z., ZHANG, X. & LI, M. 2017. Structure and assembly mechanism of plant C₂S₂M₂-type PSII-LHCII supercomplex. *Science*, 357, 815-820.
- TAIT, G. H. & GIBSON, K. D. 1961. The enzymic formation of magnesium protoporphyrin monomethyl ester. *Biochim Biophys Acta*, 52, 614-616.
- TAN, F. C., CHENG, Q., SAHA, K., HEINEMANN, I. U., JAHN, M., JAHN, D. & SMITH, A. G. 2008. Identification and characterization of the *Arabidopsis* gene encoding the tetrapyrrole biosynthesis enzyme uroporphyrinogen III synthase. *Biochem J*, 410, 291-299.
- TANAKA, A., ITO, H., TANAKA, R., TANAKA, N. K., YOSHIDA, K. & OKADA, K. 1998. Chlorophyll *a* oxygenase (CAO) is involved in chlorophyll *b* formation from chlorophyll *a*. *Proc Natl Acad Sci U S A*, 95, 12719-12723.
- TOTTEY, S., BLOCK, M. A., ALLEN, M., WESTERGRENN, T., ALBRIEUX, C., SCHELLER, H. V., MERCHANT, S. & JENSEN, P. E. 2003. *Arabidopsis* CHL27, located in both envelope and thylakoid membranes, is required for the synthesis of protochlorophyllide. *Proc Natl Acad Sci U S A*, 100, 16119-16124.
- WALKER, C. J., MANSFIELD, K. E., REZZANO, I. N., HANAMOTO, C. M., SMITH, K. M. & CASTELFRANCO, P. A. 1988. The magnesium-protoporphyrin IX (oxidative) cyclase system. Studies on the mechanism and specificity of the reaction sequence. *Biochem J*, 255, 685-692.
- WALKER, C. J. & WEINSTEIN, J. D. 1991. In vitro assay of the chlorophyll biosynthetic enzyme Mg-chelatase: resolution of the activity into soluble and membrane-bound fractions. *Proc Natl Acad Sci U S A*, 88, 5789-5793.
- WALKER, C. J. & WEINSTEIN, J. D. 1995. Re-examination of the localization of Mg-chelatase within the chloroplast. *Physiol Plant*, 94, 419-424.
- WANG, P., GAO, J., WAN, C., ZHANG, F., XU, Z., HUANG, X., SUN, X. & DENG, X. 2010. Divinyl chlorophyll(ide) *a* can be converted to monovinyl chlorophyll(ide) *a* by a divinyl reductase in rice. *Plant Physiol*, 153, 994-1003.
- WANG, P., WAN, C., XU, Z., WANG, P., WANG, W., SUN, C., MA, X., XIAO, Y., ZHU, J., GAO, X. & DENG, X. 2013. One divinyl reductase reduces the 8-vinyl groups in various intermediates of chlorophyll biosynthesis in a given higher plant species, but the isozyme differs between species. *Plant Physiol*, 161, 521-534.
- WANG, X. & LIU, L. 2016. Crystal structure and catalytic mechanism of 7-hydroxymethyl chlorophyll *a* reductase. *J Biol Chem*, 291, 13349-13359.
- WILLOWS, R. D., GIBSON, L. C., KANANGARA, C. G., HUNTER, C. N. & VON WETTSTEIN, D. 1996. Three separate proteins constitute the magnesium chelatase of *Rhodobacter sphaeroides*. *Eur J Biochem*, 235, 438-443.
- WITTY, M., WALLACE-COOK, A. D., ALBRECHT, H., SPANO, A. J., MICHEL, H., SHABANOWITZ, J., HUNT, D. F., TIMKO, M. P. & SMITH, A. G. 1993. Structure and expression of chloroplast-localized porphobilinogen deaminase from pea (*Pisum sativum* L.) isolated by redundant polymerase chain reaction. *Plant Physiol*, 103, 139-147.
- WONG, Y. S., CASTELFRANCO, P. A., GOFF, D. A. & SMITH, K. M. 1985. Intermediates in the formation of the chlorophyll isocyclic ring. *Plant Physiol*, 79, 725-729.

Paper I



Article

Aerobic Barley Mg-protoporphyrin IX Monomethyl Ester Cyclase is Powered by Electrons from Ferredoxin

David Stuart ¹, Malin Sandström ¹, Helmy M. Youssef ^{1,2}, Shakhira Zakhrebekova ¹, Poul Erik Jensen ³, David W. Bollivar ⁴ and Mats Hansson ^{1,*}

¹ Department of Biology, Lund University, Sölvegatan 35B, 22362 Lund, Sweden;

david.stuart@biol.lu.se (D.S.); malinsandstrom8@hotmail.com (M.S.); helmy.youssef@biol.lu.se (H.M.Y.); Shakhira.zakhrebekova@biol.lu.se (S.Z.)

² Faculty of Agriculture, Cairo University, Giza 12613, Egypt

³ Department of Food Science, University of Copenhagen, Rolighedsvej 26, DK-1958 Frederiksberg, Denmark; peje@food.ku.dk

⁴ Department of Biology, Illinois Wesleyan University, Bloomington, IL P.O. Box 2900, USA; dbolliva@iwu.edu

* Correspondence: mats.hansson@biol.lu.se; Tel.: +46-46-2224980

Received: 11 August 2020; Accepted: 4 September 2020; Published: 8 September 2020



Abstract: Chlorophyll is the light-harvesting molecule central to the process of photosynthesis. Chlorophyll is synthesized through 15 enzymatic steps. Most of the reactions have been characterized using recombinant proteins. One exception is the formation of the isocyclic E-ring characteristic of chlorophylls. This reaction is catalyzed by the Mg-protoporphyrin IX monomethyl ester cyclase encoded by *Xantha-1* in barley (*Hordeum vulgare* L.). The *Xantha-1* gene product (XanL) is a membrane-bound diiron monooxygenase, which requires additional soluble and membrane-bound components for its activity. XanL has so far been impossible to produce as an active recombinant protein for in vitro assays, which is required for deeper biochemical and structural analyses. In the present work, we performed cyclase assays with soluble and membrane-bound fractions of barley etioplasts. Addition of antibodies raised against ferredoxin or ferredoxin-NADPH oxidoreductase (FNR) inhibited assays, strongly suggesting that reducing electrons for the cyclase reaction involves ferredoxin and FNR. We further developed a completely recombinant cyclase assay. Expression of active XanL required co-expression with an additional protein, Ycf54. In vitro cyclase activity was obtained with recombinant XanL in combination with ferredoxin and FNR. Our experiment demonstrates that the cyclase is a ferredoxin-dependent enzyme. Ferredoxin is part of the photosynthetic electron-transport chain, which suggests that the cyclase reaction might be connected to photosynthesis under light conditions.

Keywords: *acsF*; *bchE*; *CHL27*; chlorophyll biosynthesis; *CRD1*; FNR; *Hordeum vulgare*; XanL; *Xantha-1*

1. Introduction

Chlorophylls are the most abundant light-harvesting pigments on Earth and essential for the process of photosynthesis. They belong to a large family of tetrapyrrole molecules also including hemes, bilins, and corrins. Chlorophylls are distinguished from the other tetrapyrroles by a centrally chelated magnesium ion and a fifth isocyclic ring, named the E ring. Chlorophyll biosynthesis is a major anabolic pathway divided into 15 enzymatic steps, each catalyzed by a unique enzyme [1]. The initial nine biosynthetic reactions are common to both chlorophyll and heme. In the first unique step of the chlorophyll biosynthetic pathway, Mg²⁺ is inserted into protoporphyrin IX. Subsequently, a methyl group is transferred to the carboxyl group of the propionate on the C ring of Mg-protoporphyrin IX, generating Mg-protoporphyrin IX monomethyl ester (MPE), which is the substrate of the MPE cyclase

in focus of the present study. The cyclase catalyzes the formation of the isocyclic E ring by insertion of oxygen and attaching the methylated propionate to the methene bridge between pyrrole rings C and D, forming protochlorophyllide. Chlorophyll is obtained after additional reactions involving a light-dependent oxidation of protochlorophyllide to chlorophyllide, reduction of the vinyl group on the B ring, and, finally, addition of a polyisoprene tail [1].

Most steps for chlorophyll biosynthesis have been studied using recombinant proteins in defined reactions, but the MPE cyclase is an exception. Instead, present knowledge concerning the cyclase is based on studies using fractionated cell extracts and genetic analyses of bacteria, algae, and plants [2–11]. Two distinct enzymes have been identified that catalyze the cyclase reaction, originally distinguished by the source of the incorporated oxygen. The enzyme that catalyzes the cyclase reaction in the absence of molecular oxygen (anaerobic) derives the oxygen from water and is encoded by *bchE* in facultative photosynthetic bacteria like *Rhodobacter sphaeroides* [12]. A reliable in vitro assay for this enzyme has not been developed yet. In the oxygen-requiring (aerobic) reaction, one of the oxygen atoms from molecular oxygen is incorporated into the substrate and the other is reduced to water [10,12]. Thus, the anaerobic enzyme functions as a hydratase, whereas the aerobic cyclase is an oxygenase. The aerobic cyclase was first identified in the purple nonsulfur photosynthetic bacterium *Rubrivivax gelatinosus* and named AcsF [6]. *R. gelatinosus* also has the *bchE* gene to enable photosynthesis under various oxygen conditions [13]. The discovery of *acsF* opened the possibility to identify the orthologous gene in other organisms, e.g., *Hordeum vulgare* and *Cth1* in *Chlamydomonas reinhardtii* [14,15]; *chlA1*, *chlA2*, *cycl*, and *cyclII* in *Synechocystis* sp. PCC 6803 [16,17]; *CHL27* in *Arabidopsis thaliana* [9]; and *Xantha-I* in barley (*Hordeum vulgare* L.) [7,18].

The aerobic cyclase belongs to the family of diiron carboxylate-bridged proteins characterized by the iron-binding motif E-X_n-E-X-X-H-X_n-E-X_n-E-X-X-H [19]. Detailed studies on the enzyme have been impaired by the absence of a recombinant expression system. Fractionation of cell extracts of cucumber (*Cucumis sativus* L.) [20,21], *C. reinhardtii*, *Synechocystis* [22], and barley [7,23] revealed that cyclase activity requires both additional soluble and membrane-bound fractions, but all involved components have not been identified. The discovery of Ycf54 and LCAA in *Synechocystis* [24] and tobacco *Nicotiana tabacum* L. [2], respectively, indicated that this protein affected the cyclase enzyme function in organisms that perform oxygenic photosynthesis, although the protein is not required for cyclase activity by the *acsF*-encoded aerobic cyclase found in organisms that perform anoxygenic photosynthesis [3]. Despite the lack of an active recombinant enzyme in vitro, a reaction mechanism of the aerobic cyclase has been proposed. These models suggested transformation via a β -oxidation of the propionate side-chain, similar to the β -oxidation of fatty acids, via β -hydroxy and β -keto intermediates [11,25,26]. Thus, it can be speculated that the additional soluble and membrane-associated components are involved in the redox chemistry of the reaction.

In the present work, we report a completely recombinant assay based on the barley *Xantha-I* gene product encoding the aerobic cyclase XanL. Production of active XanL strictly requires co-expression of Ycf54. We further demonstrate that assays performed with barley extract are inhibited by antibodies raised against Fd and ferredoxin-NADPH oxidoreductase (FNR). Finally, cyclase activity was obtained by combining the recombinantly expressed XanL with Fd and FNR. Our work establishes all components necessary for the functional recombinant cyclase and serves as a platform for further detailed studies on the aerobic enzyme, where Fd is the direct electron donor.

2. Results

2.1. Inhibition of Cyclase Activity with Antibodies

Fd and FNR are plastid-localized redox components, and we sought to test their role in the cyclase reaction. Antibodies are known to inhibit enzymatic reactions by steric interference [27], and antibodies against FNR have previously been shown to be effective FNR inhibitors [28]. Therefore, we used antibodies against either FNR (α -FNR) or Fd (α -Fd) to test for inhibition of cyclase activity in assays

using well-established barley plastid fractions prepared according to Bollivar et al. [23]. Addition of α -FNR or α -Fd resulted in a severe reduction in enzymatic activity (Figure 1). To ensure that there was no general inhibitory effect of antibodies, a brassinosteroid receptor (BRI1) antibody directed toward the barley brassinosteroid receptor [29] was included as a control but had no effect on cyclase activity (Figure 1).

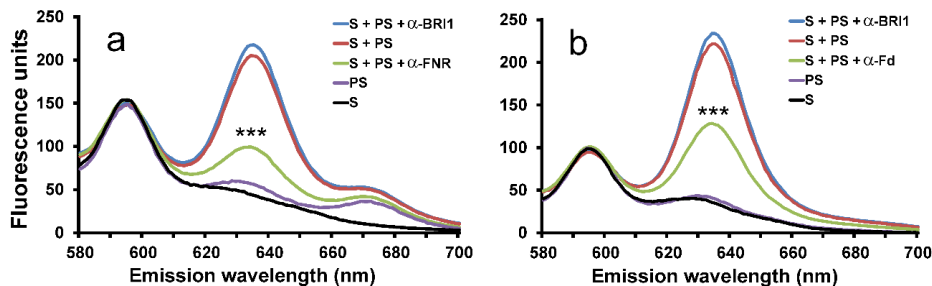


Figure 1. Effect of antibodies on enzymatic cyclase activity. Soluble (S) and solubilized membrane (PS) fractions of barley plastids were combined to obtain cyclase activity (red). Addition of antibodies raised against ferredoxin-NADPH oxidoreductase (FNR) (a, green) or Fd (b, green) significantly inhibited activity. Assays with addition of antibodies raised against barley brassinosteroid receptor BRI1 were used as controls to ensure that there was no general inhibitory effect of antibodies. Three parallel assays were performed, and the displayed curves are averages of the three replicates. ***, $p < 0.001$.

As the cyclase enzyme system from barley plastids was inhibited by the FNR and Fd antibodies, thus suggesting a role for these proteins in the cyclase reaction directly, immunoblot analysis using the same antibodies was performed on the soluble and membrane fractions used for the cyclase assays. FNR was primarily found in the membrane fraction though a small amount was also detected in the soluble fraction (Figure 2a). Fd, on the other hand, was only detected in the soluble fraction and was, therefore, a likely candidate for the previously unidentified soluble component (Figure 2b). Therefore, commercially available spinach (*Spinacia oleracea*) Fd was tested in cyclase assays to see if it could replace the soluble barley plastid fraction. When spinach Fd was combined with the membrane fraction of barley plastids, activity was indeed observed. The observed cyclase activity increased with Fd concentration (Figure 3). The results clearly demonstrate that Fd is the necessary soluble component of the cyclase enzyme system.

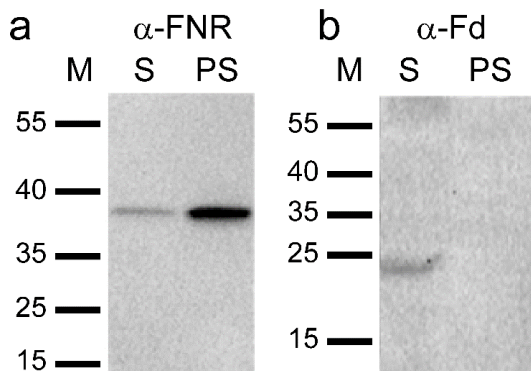


Figure 2. Immunoblot analysis of the soluble (S) and solubilized membrane (PS) fractions used in the assays. (a) FNR is predominantly detected in PS. (b) Fd is only detected in S.

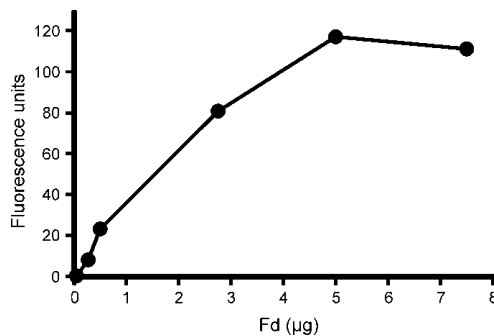


Figure 3. The effect of spinach ferredoxin (Fd) on cyclase activity. Assays were performed with an increasing amount of Fd added and a constant amount of solubilized membrane (PS) fractions of barley plastids. Fd could replace the soluble plastid fraction (S). Product formation increased with the amount of added Fd until saturation at 5 µg added Fd ($y = 23.3x + 6.35$, $R^2 = 0.980$, $p = 0.0012$). 1 µg Fd corresponds to 97 pmole.

2.2. Development of a Recombinant Cyclase Assay

In a recent study, heterologous expression of *R. gelatinosus* *acsF* in *Escherichia coli* demonstrated that cyclase activity could occur *in vivo* in this non-photosynthetic organism [3]. While this study suggested that any remaining parts of the enzyme are commonly present in *E. coli* cells, it did not fully define the essential components necessary for a completely recombinant system where all required partners are present. Based on the evidence presented above, it seemed that all that was missing for a reconstituted enzyme assay with defined components was an electron transfer system composed of Fd and FNR. However, attempts to simply produce the recombinant XanL protein and recombinant Ycf54 protein, separately, and then combine them with Fd and FNR in an *in vitro* assay did not result in any activity (Figure 4c). Given that Ycf54 did not show similarity to any previously described enzymes, it seemed plausible that the role of this protein was not catalytic but rather structural. Therefore, expression constructs using a vector that allows for co-expression of two genes were created. In the first plasmid construct, *Xantha-1* was placed in cloning site one (producing a His-tagged protein, XanL[coYcf54]) and Ycf54 in cloning site two (without a tag). As controls, a second plasmid construct was created that had only *Xantha-1* present in cloning site one (producing His-tagged XanL). A construct for expression of Ycf54 alone was already available [23].

Enzymatic assays combining spinach Fd and FNR with XanL[coYcf54] showed high activity (Figure 4a), whereas no enzymatic activity was detected in assays with recombinant XanL that was produced alone without Ycf54 (Figure 4b). Immunoblot analyses were performed and showed a similar amount of XanL produced from both expression constructs (Figure 5). As it was possible that Ycf54 was co-purifying with XanL, immunoblot analyses to detect Ycf54 levels were performed. Only trace amounts, if any, of Ycf54 were co-purified along with XanL when they were co-expressed (Figure 5a). Assays were also performed with recombinant XanL and Ycf54 that had been produced separately and then combined in the assay; however, no enzymatic activity could be detected (Figure 4c). Overall, this suggests that Ycf54 is not required for catalysis but fulfills a role during production of XanL within the *E. coli* cell. This conclusion is also supported by a very recent study of *Synechocystis*, *Chlamydomonas*, and *Arabidopsis* cyclases showing that the *in vivo* cyclase activity in *E. coli* was dependent on the co-expression of Ycf54 [30]. In our control experiments, it was further shown that activity was only seen when XanL[coYcf54] was combined with both Fd and FNR, and no activity was detected when either Fd or FNR was omitted from the reaction, nor assays with Fd and FNR but without XanL[coYcf54] reinforcing the direct role of both Fd and FNR in the catalytic cycle.

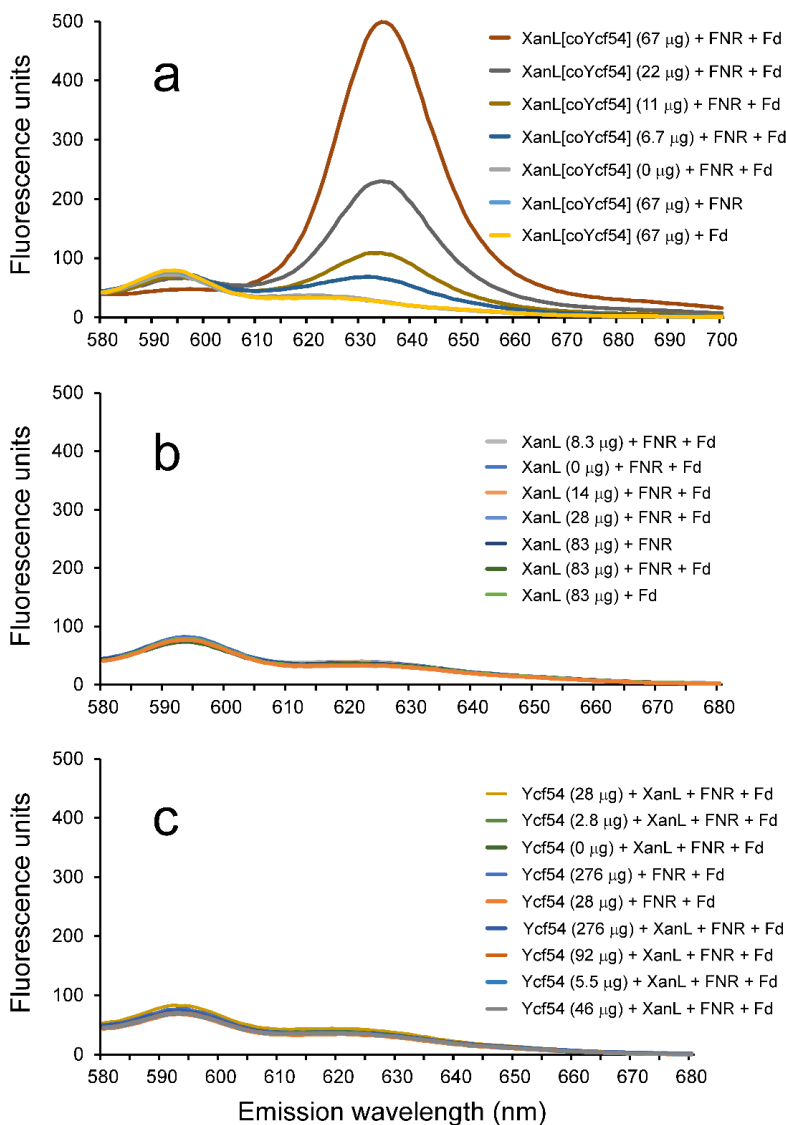


Figure 4. Enzymatic activity of recombinant XanL in combination with spinach Fd (ferredoxin) and spinach FNR (ferredoxin-NADPH oxidoreductase). Used concentrations of Fd and FNR were 0.5 µg/µL and 0.75 milli-units/µL, respectively. (a) Cyclase activity assays with recombinant XanL co-expressed with Ycf54 (XanL[coYcf54]). Product formation increased linearly with the amount of added XanL[coYcf54] ($y = 6.99x + 12.7$, $R^2 = 0.983$, $p = 0.0077$). No activity was obtained when Fd or FNR was omitted from the assay. (b) Assays performed with recombinant XanL expressed without Ycf54. No activity was detected. (c) Assays performed with recombinant XanL (83 µg) and Ycf54 expressed separately. No activity was detected.

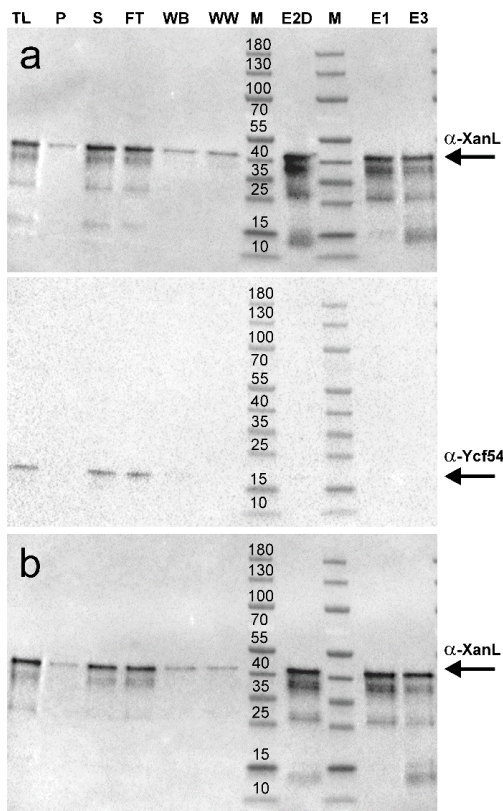


Figure 5. Immunoblots of recombinant barley XanL and Ycf54 produced in *E. coli* ArticExpress (DE3). (a) His-tagged XanL and non-His-tagged Ycf54 were co-produced from plasmid pETDuet-1. (b) Only His-tagged XanL was produced from pETDuet-1. TL, total lysate; P, pellet; S, supernatant; FT, flow through; WB, wash with binding buffer; WW, wash with wash buffer; M, marker; E2D, elution fraction 2 (desalted); M, marker; E1, elution fraction 1; E3, elution fraction 3. The molecular weight of the marker proteins is indicated (kDa). To ensure comparable loading between experiments, samples S, E2D, E1, and E3 were standardized to a protein concentration of 1 mg/mL prior to loading. Samples TL, P, and FT were loaded on an equal volume basis as the corresponding S.

3. Discussion

In the mid-1980s, Wong et al. were the first to report that both soluble and membrane-associated chloroplast components are required for the cyclase enzyme [11]. The cyclase is a carboxylate-bridged diiron monooxygenase [31] and, as such, it requires molecular oxygen as well as a reductant to supply electrons [10,31]. However, due to previous difficulties in producing active recombinant cyclase, the enzyme is poorly understood and its electron donor has so far not been identified. In oxygenic photosynthesis, Fd transfers electrons to FNR for production of NADPH. The reverse reaction can also occur [32]. Recently, it was proposed that FNR could act as a direct electron donor to the cyclase or create an NADPH-rich environment at the cyclase complex [33]. Our *in vitro* assay requires Fd, FNR, and NADPH for activity, which suggests that Fd is the true donor of electrons to XanL as cyclase activity was not obtained with only NADPH added or with NADPH in combination with FNR. By identifying Fd as the unknown soluble component and showing that the cyclase is an Fd-dependent enzyme, we have answered one of the key questions about basic plastid physiology that has been an enigma for

almost 40 years, and we can expand the current model of the aerobic cyclase reaction with the electron donor components (Figure 6).

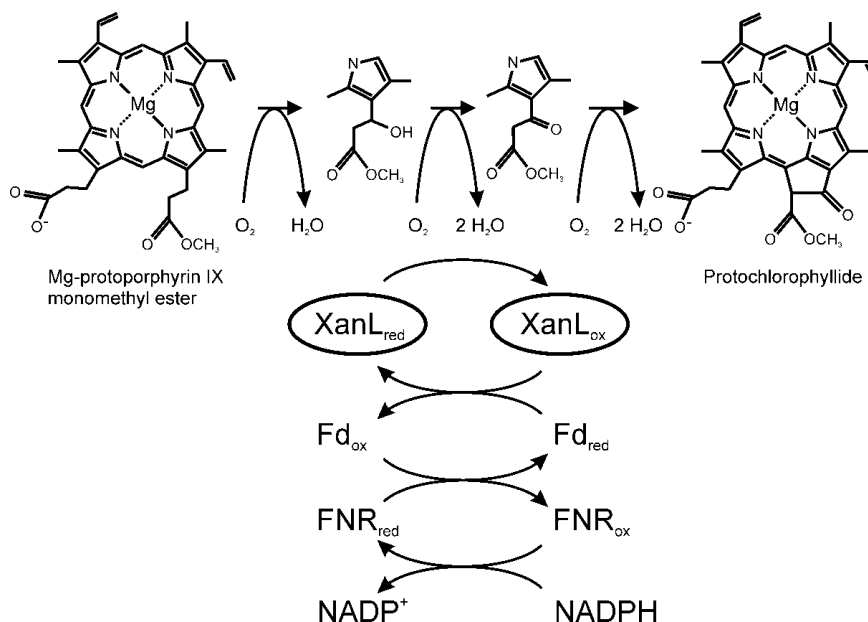


Figure 6. Reaction mechanism of the aerobic cyclase reaction. The cyclase reaction is a six-electron redox reaction suggested to proceed via β -hydroxy and β -keto intermediates. We showed that the electrons are provided by Fd. In the present study, we used NADPH as the source of electrons transferred to Fd via FNR. In green tissue, Fd can also be reduced by photosystem I.

The early biochemical experiments suggested a reaction mechanism via a β -oxidation of the propionate side-chain, similar to the β -oxidation of fatty acids [11,25,26]. It was shown that both the β -hydroxy and β -keto intermediates could function as substrates for the aerobic cyclase [11,34]. While the multi-step nature of the reaction makes investigating individual steps challenging, previous studies have shown that at least the first and last steps require both molecular oxygen and a reductant [10,11,21,34]. The second step, oxidation of the hydroxyl group to a keto group, is likely to have the same requirements. One possible scenario for this middle reaction is that the hydroxylation reaction is simply performed again to produce a geminal diol intermediate that spontaneously dehydrates to form the β -keto intermediate in a fashion similar to that which has been proposed for the chlorophyllide *a* oxygenase reaction [35], although this mechanism has yet to be verified for either enzyme. What has been shown, however, is that only one of the stereoisomers of the β -hydroxy intermediate is an active substrate for the cyclase, which indicates that the first hydroxylation must be stereospecific [34]. Which of the two isomers is the true intermediate has yet to be determined.

The evidence from this study and those supported by other recent studies [3,4,30,36] indicate that the accessory protein Ycf54 is unlikely to be directly involved in the catalytic reaction but likely facilitates proper maturation of the cyclase. One possibility is that Ycf54 facilitates proper folding and maturation of the XanL protein. This hypothesis is in agreement with a recent study showing that a *ycf54* knockout in the cyanobacterium *Synechocystis* PCC 6803 accumulates a low level of chlorophyll and has reduced levels of the XanL orthologue [36]. It was further shown that the formation of protochlorophyllide *in vivo* in *E. coli* cells expressing the genes of magnesium chelatase, Mg-protoporphyrin IX methyl transferase, and the XanL orthologues of *Synechocystis*, *Chlamydomonas*, and Arabidopsis was strictly

dependent on the co-expression of Ycf54 [30]. Presently, only limited amounts of active XanL can be produced facilitated by the co-expression of Ycf54, as shown in our present study and also reported by Chen and Hunter [30]. Thus, further refinement of the recombinant expression systems is required, which is likely to also deepen our understanding of the role and mechanism of Ycf54 in the process of XanL maturation. Accessory proteins connected to various binuclear iron monooxygenases have been reported. For example, a propane monooxygenase encoded by *mimABCD* in *Mycobacterium goodii* required co-expression of *mimG* to be obtained in active form [37]. The *mimG* gene product is similar to the chaperonin GroEL. Another gene, *phk*, encoding an accessory protein of a phenol hydroxylase of *Pseudomonas* sp. OX1, did not appear to be required for the production of active recombinant phenol hydroxylase but, rather, involved in increasing the affinity of the hydroxylase for iron [38]. It is obvious that accessory proteins are common in binuclear iron monooxygenases, and they might shed light on the function of Ycf54 in the process of XanL maturation, although they show no sequence similarity.

While the reductant in our assay system is NADPH, the source of reducing electrons may well be different in planta and even at different developmental stages. In green tissue, Fd is mainly reduced by photosystem I as a result of linear electron transport. Thus, under light conditions, Fd might connect photosynthesis with chlorophyll biosynthesis at the cyclase reaction step. A barley mutant, *viridis-zb.63*, which has only 2% photosystem I activity compared to wild type, behaves like a cyclase mutant, i.e., accumulates the cyclase substrate MPE when fed the chlorophyll biosynthetic precursor 5-aminolevulinic acid [8,39]. In this context, greater amounts of MPE are accumulated by plants incubated in the light as opposed to darkness [8]. This result can be explained if the cyclase receives electrons from photo-reduced Fd in the light, whereas in darkness, Fd would be reduced by FNR. Another stage where Fd is likely to be reduced by FNR is during de-etiolation of seedlings as there are not yet any active photosynthetic complexes that could photoreduce Fd. In both cases, Fd is reduced by FNR with electrons from NADPH generated in the pentose phosphate pathway. Identification of the required components and establishment of an in vitro assay using recombinant XanL have set the stage for future studies to decipher the structure and biochemistry of this enigmatic enzyme, as well as the source of electrons during different developmental stages and light conditions.

4. Materials and Methods

4.1. Barley Plastid Preparation

Barley (*Hordeum vulgare* L.) etioplasts were prepared and fractionated into soluble and solubilized membrane fractions from cultivar Bonus as described previously [23].

4.2. Cyclase Assays

The magnesium protoporphyrin IX monomethyl ester (MPE) substrate for cyclase assays was prepared by the method of [40]. Enzymatic activity assays were performed as described in Bollivar et al. [23] with minor alterations. Unless otherwise specified, the reaction conditions for assays containing barley plastid fractions were 20 mM Tricine and 10 mM HEPES pH 8.1, 1 mM EDTA, 25 mM MgCl₂, 1 mM DTT, 10 mM glucose-6-phosphate, 0.03 units/μL glucose-6-phosphate dehydrogenase, 0.5 mM NADPH, 90 μg/μL catalase, 0.026% Triton X-100, and 10 μM MPE. Reactions with recombinant XanL were the same except that the EDTA was omitted and contained 0.75 milli-units/μL *Spinacia oleracea* FNR (Sigma-Aldrich, Stockholm, Sweden) and 0.5 μg/μL *S. oleracea* ferredoxin (Sigma-Aldrich, Stockholm, Sweden). Assays were incubated in the dark at 30 °C and 750 rpm for one hour in a Thermomixer comfort (Eppendorf Nordic A/S, Hørsholm, Denmark). Assays were stopped by addition of 80% acetone with 0.32% NH₃ and centrifuged for 5 min at 29,000× g to pellet-precipitated proteins. Formation of protochlorophyllide was measured using an RF-5301 PC spectrofluorophotometer (Shimadzu, Kista, Sweden) with an excitation wavelength of 440 nm and an emission spectrum between 570 and 700 nm with slit widths of 10 nm for both excitation and emission. Product formation was estimated as the relative fluorescence emission at 634 nm for samples minus the emission for the negative control

sample. The linear model, $lm()$, function in R version 3.6.1 was used for statistical comparisons between categorical treatments, as well as for fitting regression lines to numerical treatments.

4.3. Antibodies Used

Polyclonal rabbit antibodies were raised against barley proteins XanL [23], Fd [41], FNR [42], and BRI1 [29]. The polyclonal rabbit antibodies against Ycf54 were produced by immunization with the synthetic peptide CAETVEEALASNPAL linked to KLH (Keyhole limpet hemocyanin) as a carrier protein (Agrisera, Vännäs, Sweden).

4.4. Plasmid Constructs

The plasmids pETDuet-1HvXanL and pETDuet-1HvXanLYcf54 containing *Xantha-l* and *Ycf54* that were codon-optimized for expression in *E. coli* were from GenScript Biotech Corporation (Leiden, The Netherlands). The *Xantha-l* gene was placed in multiple cloning site one between the *HindIII* and *NotI* sites, while the *Ycf54* gene was placed between the *NdeI* and *BglII* sites. In order to improve purification, the initial constructs were modified by PCR-based mutagenesis [43], resulting in a His tag with 10 residues instead of 6, as well as insertion of a TEV site between the tag and the *Xantha-l* gene. Constructs were verified by Sanger sequencing after mutagenesis. The primers used for mutagenesis can be found in Table 1. For expression and purification of Ycf54, the plasmid pET15bYcf54 was used [23].

Table 1. List of oligonucleotides used.

Name	Sequence	Description
Duet_XanL_TEV_Forward	5'-TGAGAATCTTTATTTTCAGGGCAA GCCGGGTAGCCCGAAGAAACGTGGCA	Forward primer to add TEV site to pETDuet-1HvXanL constructs
Duet_XanL_TEV_Reverse	5'-CCCTGAAAATAAAGATTCTCAAG CTTGTCGACCTGCAGGCGCGCCGAGCT	Reverse primer to add TEV site to pETDuet-1HvXanL constructs
DuetTEV2x_His_Forward	5'-CACCAACATCATAGCCAGGATCC GAATTCGAGCTCGGCGCCCTGCAGGT	Forward primer to add 2 extra Histidine residues to make a 8x His tag
DuetTEV2x_His_Reverse	5'-GGATCCTGGCTATGATGGTGGTG ATGATGGTGATGGCTGCTGCCCATGGT	Reverse primer to add 2 extra Histidine residues to make a 8x His tag
DuetTEV4x_His_Reverse	5'-GGATCCTGGCTATGATGGTGGTGATG ATGGTGATGATGATGGCTGCTGCCCATGGT	Reverse primer to add 2 extra Histidine residues to make a 10x His Tag. This primer was run with DuetTEV2xHis_Forward

4.5. Protein Production

Plasmids were transformed into *E. coli* ArcticExpress (DE3). Overnight cultures were grown at 37 °C at 200 rpm in LB media supplemented with 100 µg/mL ampicillin, 20 µg/mL gentamycin, and 1% glucose. The next day, 6 × 1 L LB in 5 L baffled flasks supplemented with antibiotics were inoculated to an OD₆₀₀ of 0.1. The flasks were placed in a Multitron Standard incubator (Infors HT, Farsta, Sweden) at 37 °C at 200 rpm for 20 min after which the temperature was changed to 15 °C and the cultures were allowed to grow as the temperature gradually decreased. After approximately 3 h when the cultures had an OD₆₀₀ between 0.5 and 0.6, the cultures were induced with IPTG to a final concentration of 1 mM. Cultures expressing *Xantha-l* were also supplemented with solid FeSO₄ to a final concentration of 1 mM at the time of induction. Cultures were then grown for an additional 68 h before the *E. coli* cells were harvested by centrifugation. Cell pellets were then frozen at −20 °C until use.

4.6. Protein Purification

For purification of XanL, the cell pellet from 2 L culture was re-suspended to 35 mL in binding buffer (20 mM imidazole, 500 mM NaCl, 20 mM Tris-HCl pH 8.0, and 15% (v/v) glycerol) with a few crystals of lysozyme and DNase I added. *E. coli* cells were lysed by passage through a French pressure cell three times at 12.4 MPa followed by centrifugation at 48,384× *g* for 30 min at 4 °C. The supernatant was then loaded on a 1 mL HisTrapFF crude column (GE Healthcare, Helsingborg, Sweden) equilibrated with binding buffer. The column was then washed with 25 mL binding buffer followed by 25 mL wash buffer (45 mM imidazole, 500 mM NaCl, 20 mM Tris-HCl pH 8, and 15% (v/v) glycerol). The proteins were eluted with elution buffer (250 mM imidazole, 500 mM NaCl, 20 mM Tris-HCl pH 8, and 15% (v/v) glycerol) and collected in 1 mL fractions. The elution fraction with the highest protein concentration was desalted over a NAP-10 column into 50 mM Tris-HCl pH 8.0 with 15% (v/v) glycerol. The protein was then stored at −80 °C until use. Protein concentration was estimated using Bradford reagent (BioRad, Solna, Sweden) using bovine serum albumin fraction V as a standard. Purification of Ycf54 using cell pellets from 1 L culture followed the same procedure except that glycerol was not present in the buffers during purification and the buffers were supplemented with 1 mM DTT. In addition, the column was only washed with 15 mL binding buffer and 10 mL wash buffer. The purified Ycf54 was desalted as above into 50 mM Tris-HCl pH 8.0, 15% (v/v) glycerol, and 1 mM DTT. Ycf54 was not stable when frozen and was used the same day it was purified. All protein purifications were performed at 4 °C.

4.7. SDS-PAGE and Immunoblotting

Proteins were separated on 4–15% Mini-Protean TGX precast gels (BioRad). Molecular weight markers (PageRuler™ Prestained Protein Ladder) were from Thermo Scientific. Gels were stained with Bio-Safe Coomassie G-250 (BioRad, Solna, Sweden) or, when used for immunoblots, transferred onto 2 μm PVDF membranes (BioRad). Membranes were blocked for 1 h in 5% non-fat powdered milk in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.4) followed by incubation for 1 h with primary antibody in 1% non-fat powdered milk in PBS. Membranes were then washed three times with PBST (PBS plus 0.1% Tween-20) followed by 1 h incubation with goat anti-rabbit horseradish peroxidase-conjugated secondary antibody (BioRad) in 3% non-fat powdered milk in PBS. The membrane was then washed twice with PBST followed by one wash in PBS before being developed using a Pierce ECL immunoblotting substrate (Thermo Scientific, Stockholm, Sweden) and imaged using a ChemiDoc MP system (BioRad).

Author Contributions: Conceptualization, D.S., P.E.J., D.W.B. and M.H.; methodology, D.S., P.E.J., D.W.B. and M.H.; validation, D.S., P.E.J., D.W.B. and M.H.; formal analysis, D.S. and H.M.Y.; investigation, D.S., M.S., H.M.Y., S.Z., D.W.B. and M.H.; resources, M.H.; data curation, D.S. and M.H.; writing—original draft preparation, D.S., D.W.B. and M.H.; writing—review and editing, D.S., H.M.Y., S.Z., P.E.J., D.W.B. and M.H.; visualization, D.S., H.M.Y., S.Z. and M.H.; supervision, D.S., P.E.J., D.W.B. and M.H.; project administration, M.H.; funding acquisition, D.S., H.M.Y., S.Z., P.E.J., D.W.B. and M.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the VILLUM Foundation (13363 to P.E.J.), the German Research Foundation (DFG YO 304/1-1 to H.M.Y.), the Swedish Research Council (VR 2018-05117 to M.H.), the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS 2018-01026 to M.H.), the Erik Philip-Sörensen Foundation (M.H.), the Royal Physiographic Society in Lund (D.S., H.M.Y., M.H., S.Z.), and an Artistic and Scholarly Development grant from Illinois Wesleyan University (D.W.B.).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tanaka, R.; Tanaka, A. Tetrapyrrole biosynthesis in higher plants. *Annu. Rev. Plant Biol.* **2007**, *58*, 321–346. [[CrossRef](#)] [[PubMed](#)]

2. Albus, C.A.; Salinas, A.; Czarnecki, O.; Kahlau, S.; Rothbart, M.; Thiele, W.; Lein, W.; Bock, R.; Grimm, B.; Schottler, M.A. LCAA, a novel factor required for magnesium protoporphyrin monomethylester cyclase accumulation and feedback control of aminolevulinic acid biosynthesis in tobacco. *Plant Physiol.* **2012**, *160*, 1923–1939. [[CrossRef](#)] [[PubMed](#)]
3. Chen, G.E.; Canniffe, D.P.; Barnett, S.F.H.; Hollingshead, S.; Brindley, A.A.; Vasilev, C.; Bryant, D.A.; Hunter, C.N. Complete enzyme set for chlorophyll biosynthesis in *Escherichia coli*. *Sci. Adv.* **2018**, *4*, eaaq1407. [[CrossRef](#)] [[PubMed](#)]
4. Chen, G.E.; Canniffe, D.P.; Hunter, C.N. Three classes of oxygen-dependent cyclase involved in chlorophyll and bacteriochlorophyll biosynthesis. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 6280–6285. [[CrossRef](#)] [[PubMed](#)]
5. Chereskin, B.M.; Castelfranco, P.A. Effects of iron and oxygen on chlorophyll biosynthesis. II. Observations on the biosynthetic pathway in isolated etiochloroplasts. *Plant Physiol.* **1982**, *69*, 112–116. [[CrossRef](#)]
6. Pinta, V.; Picaud, M.; Reiss-Husson, F.; Astier, C. Rubrivivax gelatinosus *acsF* (previously orf358) codes for a conserved, putative binuclear-iron-cluster-containing protein involved in aerobic oxidative cyclization of Mg-protoporphyrin IX monomethylester. *J. Bacteriol.* **2002**, *184*, 746–753. [[CrossRef](#)]
7. Rzeznicka, K.; Walker, C.J.; Westergren, T.; Kannangara, C.G.; von Wettstein, D.; Merchant, S.; Gough, S.P.; Hansson, M. Xantha-I encodes a membrane subunit of the aerobic Mg-protoporphyrin IX monomethyl ester cyclase involved in chlorophyll biosynthesis. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5886–5891. [[CrossRef](#)]
8. Steccanella, V.; Hansson, M.; Jensen, P.E. Linking chlorophyll biosynthesis to a dynamic plastoquinone pool. *Plant Physiol. Biochem.* **2015**, *97*, 207–216. [[CrossRef](#)]
9. Tottey, S.; Block, M.A.; Allen, M.; Westergren, T.; Albrieux, C.; Scheller, H.V.; Merchant, S.; Jensen, P.E. Arabidopsis *CHL27*, located in both envelope and thylakoid membranes, is required for the synthesis of protochlorophyllide. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 16119–16124. [[CrossRef](#)]
10. Walker, C.J.; Mansfield, K.E.; Smith, K.M.; Castelfranco, P.A. Incorporation of atmospheric oxygen into the carbonyl functionality of the protochlorophyllide isocyclic ring. *Biochem. J.* **1989**, *257*, 599–602. [[CrossRef](#)]
11. Wong, Y.S.; Castelfranco, P.A.; Goff, D.A.; Smith, K.M. Intermediates in the formation of the chlorophyll isocyclic ring. *Plant Physiol.* **1985**, *79*, 725–729. [[CrossRef](#)] [[PubMed](#)]
12. Porra, R.J.; Schafer, W.; Gadón, N.; Katheder, I.; Drews, G.; Scheer, H. Origin of the two carbonyl oxygens of bacteriochlorophyll a. Demonstration of two different pathways for the formation of ring E in *Rhodobacter sphaeroides* and *Roseobacter denitrificans*, and a common hydratase mechanism for 3-acetyl group formation. *Eur. J. Biochem.* **1996**, *239*, 85–92. [[CrossRef](#)]
13. Ouchane, S.; Steunou, A.S.; Picaud, M.; Astier, C. Aerobic and anaerobic Mg-protoporphyrin monomethyl ester cyclases in purple bacteria: A strategy adopted to bypass the repressive oxygen control system. *J. Biol. Chem.* **2004**, *279*, 6385–6394. [[CrossRef](#)] [[PubMed](#)]
14. Eriksson, M.; Moseley, J.L.; Tottey, S.; Del Campo, J.A.; Quinn, J.; Kim, Y.; Merchant, S. Genetic dissection of nutritional copper signaling in *Chlamydomonas* distinguishes regulatory and target genes. *Genetics* **2004**, *168*, 795–807. [[CrossRef](#)] [[PubMed](#)]
15. Moseley, J.; Quinn, J.; Eriksson, M.; Merchant, S. The *Crd1* gene encodes a putative di-iron enzyme required for photosystem I accumulation in copper deficiency and hypoxia in *Chlamydomonas reinhardtii*. *EMBO J.* **2000**, *19*, 2139–2151. [[CrossRef](#)]
16. Minamizaki, K.; Mizoguchi, T.; Goto, T.; Tamiaki, H.; Fujita, Y. Identification of two homologous genes, chlAI and chlAII, that are differentially involved in isocyclic ring formation of chlorophyll a in the cyanobacterium *Synechocystis* sp. PCC 6803. *J. Biol. Chem.* **2008**, *283*, 2684–2692. [[CrossRef](#)]
17. Peter, E.; Salinas, A.; Wallner, T.; Jeske, D.; Dienst, D.; Wilde, A.; Grimm, B. Differential requirement of two homologous proteins encoded by *sll1214* and *sll1874* for the reaction of Mg protoporphyrin monomethylester oxidative cyclase under aerobic and micro-oxic growth conditions. *Biochim. Biophys. Acta* **2009**, *1787*, 1458–1467. [[CrossRef](#)]
18. Wettstein, D.V.; Kahn, A.; Nielsen, O.F.; Gough, S. Genetic regulation of chlorophyll synthesis analyzed with mutants in barley. *Science* **1974**, *184*, 800–802. [[CrossRef](#)]
19. Berthold, D.A.; Stenmark, P. Membrane-bound diiron carboxylate proteins. *Annu. Rev. Plant Biol.* **2003**, *54*, 497–517. [[CrossRef](#)]
20. Walker, C.J.; Castelfranco, P.A.; Whyte, B.J. Synthesis of divinyl protochlorophyllide. Enzymological properties of the Mg-protoporphyrin IX monomethyl ester oxidative cyclase system. *Biochem. J.* **1991**, *276*, 691–697. [[CrossRef](#)]

21. Wong, Y.S.; Castelfranco, P.A. Resolution and reconstitution of Mg-protoporphyrin IX monomethyl ester (oxidative) cyclase, the enzyme system responsible for the formation of the chlorophyll isocyclic ring. *Plant Physiol.* **1984**, *75*, 658–661. [[CrossRef](#)] [[PubMed](#)]
22. Bollivar, D.W.; Beale, S.I. The chlorophyll biosynthetic enzyme Mg-protoporphyrin IX monomethyl ester (oxidative) cyclase. Characterization and partial purification from *Chlamydomonas reinhardtii* and *Synechocystis* sp. PCC 6803. *Plant Physiol.* **1996**, *112*, 105–114. [[CrossRef](#)] [[PubMed](#)]
23. Bollivar, D.; Braumann, I.; Berendt, K.; Gough, S.P.; Hansson, M. The Ycf54 protein is part of the membrane component of Mg-protoporphyrin IX monomethyl ester cyclase from barley (*Hordeum vulgare* L.). *FEBS J.* **2014**, *281*, 2377–2386. [[CrossRef](#)] [[PubMed](#)]
24. Hollingshead, S.; Kopečna, J.; Jackson, P.J.; Canniffe, D.P.; Davison, P.A.; Dickman, M.J.; Sobotka, R.; Hunter, C.N. Conserved chloroplast open-reading frame ycf54 is required for activity of the magnesium protoporphyrin monomethylester oxidative cyclase in *Synechocystis* PCC 6803. *J. Biol. Chem.* **2012**, *287*, 27823–27833. [[CrossRef](#)]
25. Ellsworth, R.K.; Aronoff, S. Investigations of the biogenesis of chlorophyll a. IV. Isolation and partial characterization of some biosynthetic intermediates between Mg-protoporphine IX monomethyl ester and Mg-vinylpheoporphine a5, obtained from *Chlorella* mutants. *Arch. Biochem. Biophys.* **1969**, *130*, 374–383. [[CrossRef](#)]
26. Granick, S. The structural and functional relationships between heme and chlorophyll. *Harvey Lect.* **1948**, *44*, 220–245.
27. Warshawsky, I.; Hortin, G.L. Effect of substrate size on immunoinhibition of amylase activity. *J. Clin. Lab. Anal.* **2001**, *15*, 64–70. [[CrossRef](#)]
28. Schmidt, H.; Heinz, E. Involvement of ferredoxin in desaturation of lipid-bound oleate in chloroplasts. *Plant Physiol.* **1990**, *94*, 214–220. [[CrossRef](#)]
29. Dockter, C.; Gruszka, D.; Braumann, I.; Druka, A.; Druka, I.; Franckowiak, J.; Gough, S.P.; Janeczko, A.; Kurowska, M.; Lundqvist, J.; et al. Induced variations in brassinosteroid genes define barley height and sturdiness, and expand the green revolution genetic toolkit. *Plant Physiol.* **2014**, *166*, 1912–1927. [[CrossRef](#)]
30. Chen, G.E.; Hunter, C.N. Protochlorophyllide synthesis by recombinant cyclases from eukaryotic oxygenic phototrophs and the dependence on Ycf54. *Biochem. J.* **2020**, *477*, 2313–2325. [[CrossRef](#)]
31. Berthold, D.A.; Andersson, M.E.; Nordlund, P. New insight into the structure and function of the alternative oxidase. *Biochim. Acta* **2000**, *1460*, 241–254. [[CrossRef](#)]
32. Shinohara, F.; Kurisu, G.; Hanke, G.; Bowshe, C.; Hase, T.; Kimata-Arigo, Y. Structural basis for the isotype-specific interactions of ferredoxin and ferredoxin: NADP⁺ oxidoreductase: An evolutionary switch between photosynthetic and heterotrophic assimilation. *Photosynth. Res.* **2017**, *134*, 281–289. [[CrossRef](#)] [[PubMed](#)]
33. Herbst, J.; Girke, A.; Hajirezaei, M.R.; Hanke, G.; Grimm, B. Potential roles of YCF54 and ferredoxin-NADPH reductase for magnesium protoporphyrin monomethylester cyclase. *Plant J.* **2018**, *94*, 485–496. [[CrossRef](#)] [[PubMed](#)]
34. Walker, C.J.; Mansfield, K.E.; Rezzano, I.N.; Hanamoto, C.M.; Smith, K.M.; Castelfranco, P.A. The magnesium-protoporphyrin IX (oxidative) cyclase system. Studies on the mechanism and specificity of the reaction sequence. *Biochem. J.* **1988**, *255*, 685–692. [[PubMed](#)]
35. Oster, U.; Tanaka, R.; Tanaka, A.; Rudiger, W. Cloning and functional expression of the gene encoding the key enzyme for chlorophyll b biosynthesis (CAO) from *Arabidopsis thaliana*. *Plant J.* **2000**, *21*, 305–310. [[CrossRef](#)] [[PubMed](#)]
36. Hollingshead, S.; Kopečna, J.; Armstrong, D.R.; Bucinska, L.; Jackson, P.J.; Chen, G.E.; Dickman, M.J.; Williamson, M.P.; Sobotka, R.; Hunter, C.N. Synthesis of chlorophyll-binding proteins in a fully segregated Dycf54 strain of the cyanobacterium *Synechocystis* PCC 6803. *Front. Plant Sci.* **2016**, *7*, 292. [[CrossRef](#)]
37. Furuya, T.; Hayashi, M.; Semba, H.; Kino, K. The mycobacterial binuclear iron monooxygenases require a specific chaperonin-like protein for functional expression in a heterologous host. *FEBS J.* **2013**, *280*, 817–826. [[CrossRef](#)]
38. Izzo, V.; Leo, G.; Scognamiglio, R.; Troncone, L.; Biolo, L.; Di Donato, A. PHK from phenol hydroxylase of *Pseudomonas* sp. OX1. Insight into the role of an accessory protein in bacterial multicomponent monooxygenases. *Arch. Biochem. Biophys.* **2011**, *505*, 48–59. [[CrossRef](#)]

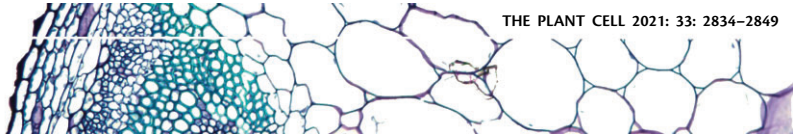
39. Nielsen, V.S.; Scheller, H.V.; Møller, B.L. The photosystem I mutant viridis-zb63 of barley (*Hordeum vulgare*) contains low amounts of active but unstable photosystem I. *Physiol. Plant.* **1996**, *98*, 637–644. [[CrossRef](#)]
40. Gough, S.P.; Rzeznicka, K.; Peterson Wulff, R.; de Cruz Francisco, J.; Hansson, A.; Jensen, P.E.; Hansson, M. A new method for isolating physiologically active Mg-protoporphyrin monomethyl ester, the substrate of the cyclase enzyme of the chlorophyll biosynthetic pathway. *Plant Physiol. Biochem.* **2007**, *45*, 932–936. [[CrossRef](#)]
41. Andersen, B.; Koch, B.; Scheller, H.V. Structural and functional analysis of the reducing side of photosystem I. *Physiol. Plant.* **1992**, *84*, 154–161. [[CrossRef](#)]
42. Andersen, B.; Scheller, H.V.; Møller, B.L. The PSI-E subunit of photosystem I binds ferredoxin:NADP+ oxidoreductase. *FEBS Lett.* **1992**, *311*, 169–173. [[CrossRef](#)]
43. Hansson, M.D.; Rzeznicka, K.; Rosenbäck, M.; Hansson, M.; Sirijovski, N. PCR-mediated deletion of plasmid DNA. *Anal. Biochem.* **2008**, *375*, 373–375. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Paper II





Barley *Viridis-k* links an evolutionarily conserved C-type ferredoxin to chlorophyll biosynthesis

David Stuart,¹ Malin Sandström,¹ Helmy M. Youssef,^{1,2,†} Shakhira Zakhrebekova,¹ Poul Erik Jensen,³ David Bollivar⁴ and Mats Hansson^{1,*}

¹ Department of Biology, Lund University, Lund 22362, Sweden

² Faculty of Agriculture, Cairo University, Giza 12613, Egypt

³ Department of Food Science, University of Copenhagen, Frederiksberg DK-1958, Denmark

⁴ Department of Biology, Illinois Wesleyan University, Bloomington, IL 61702-2900, USA

*Author for correspondence: mats.hansson@biol.lu.se

†Senior author.

†Present address: Faculty of Natural Sciences III, Institute of Agricultural and Nutritional Sciences, Martin Luther University Halle-Wittenberg, Halle 06120, Germany.

D.S., D.B., M.H., and P.E.J. designed the experiments; D.S., H.M.Y., M.S., and S.Z. performed the experiments; D.S., D.B., and M.H. analyzed the data; D.S., D.B., M.H., and P.E.J. wrote the article. All authors read and approved the final article.

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/plcell>) is: Mats Hansson (mats.hansson@biol.lu.se).

Abstract

Ferredoxins are single-electron carrier proteins involved in various cellular reactions. In chloroplasts, the most abundant ferredoxin accepts electrons from photosystem I and shuttles electrons via ferredoxin NADP⁺ oxidoreductase to generate NADPH or directly to ferredoxin dependent enzymes. In addition, plants contain other isoforms of ferredoxins. Two of these, named FdC1 and FdC2 in *Arabidopsis thaliana*, have C-terminal extensions and functions that are poorly understood. Here we identified disruption of the orthologous *FdC2* gene in barley (*Hordeum vulgare* L.) mutants at the *Viridis-k* locus; these mutants are deficient in the aerobic cyclase reaction of chlorophyll biosynthesis. The magnesium-protoporphyrin IX monomethyl ester cyclase is one of the least characterized enzymes of the chlorophyll biosynthetic pathway and its electron donor has long been sought. Agroinfiltrations showed that the *viridis-k* phenotype could be complemented in vivo by *Viridis-k* but not by canonical ferredoxin. VirK could drive the cyclase reaction in vitro and analysis of cyclase mutants showed that in vivo accumulation of VirK is dependent on cyclase enzyme levels. The chlorophyll deficient phenotype of *viridis-k* mutants suggests that VirK plays an essential role in chlorophyll biosynthesis that cannot be replaced by other ferredoxins, thus assigning a specific function to this isoform of C-type ferredoxins.

Introduction

The chlorophyll molecule is magnesium (Mg)-containing tetrapyrrole synthesized in several consecutive steps each catalyzed by a specific enzyme (Tanaka and Tanaka, 2007; Bryant et al., 2020). One of the least understood steps is the conversion of Mg-protoporphyrin IX monomethyl ester

(MPE) to protochlorophyllide, which is a six-electron oxidation. In this reaction, the fifth ring E is formed which is unique to chlorophylls. The reaction is catalyzed by an MPE cyclase, adding a carbonyl group to C13¹ as well as forming a carbon–carbon bond between C13² and the bridge carbon between rings C and D (Figure 1). There are two unrelated

Received May 25, 2021. Accepted May 28, 2021. Advance access publication May 29, 2021

© The Author(s) 2021. Published by Oxford University Press on behalf of American Society of Plant Biologists.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Open Access

IN A NUTSHELL

Background: Chlorophyll is the green color of plants. It is made in several consecutive steps each catalyzed by an enzyme. Mutants from photosynthetic organisms have been used over the years to elucidate the pathway of chlorophyll biosynthesis. In this way, we know that chlorophyll is made from eight molecules of the amino acid glutamate in 15 enzymatic steps. One of the least known enzymes in the pathway is the cyclase, which catalyzes the formation of a ring in the chlorophyll structure. In barley, the cyclase enzyme is named XanL and mutations in the cyclase gene, *Xantha-1*, causes a plant with yellow color because chlorophyll cannot be made. There are also two additional mutants in barley, *viridis-k.23* and *viridis-k.170*, which are blocked in the cyclase reaction.

Question: In the present study we aimed to identify the *Viridis-k* gene to reveal its function in the cyclase reaction.

Findings: The *viridis-k.23* and *viridis-k.170* mutants have a large chromosomal rearrangement and a point mutation, respectively, in a gene encoding a ferredoxin with a C-terminal extension. This type of ferredoxin is conserved in chlorophyll-producing plants, algae and bacteria, which suggests a common and important function of this protein. Barley has eight ferredoxins associated with the chloroplast. The yellow appearance of *viridis-k* mutants suggests that none of the other ferredoxins can replace the *VirK* ferredoxin and thereby complement a deficiency in the *Viridis-k* gene. Thus, our study revealed the long sought-after donor of electrons to the cyclase reaction.

Next steps: Ferredoxins transfer electrons in various cellular reactions. Their source of electrons is typically NADPH or photosystem I. Future experiments will focus on the source of electrons to *VirK* as well as the function of the C-terminal extension. This will be done by general biochemical characterizations and structural studies.

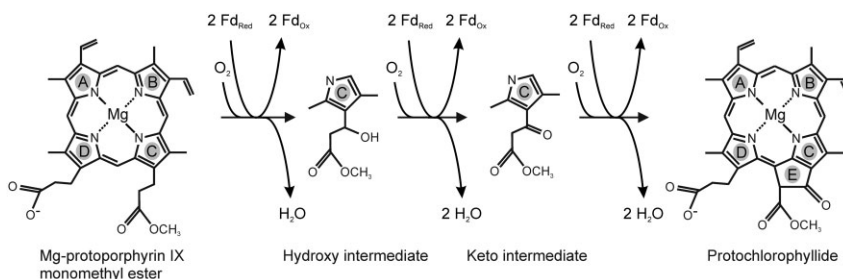


Figure 1 Proposed intermediates in the aerobic MPE cyclase reaction. The reaction has been proposed to proceed via β -hydroxy and β -keto intermediates. The first step is a hydroxylation reaction which bears some resemblance to the reaction catalyzed by soluble methane monooxygenase (Banerjee et al., 2019). Both the cyclase and the soluble methane monooxygenase use electrons from ferredoxin and their substrate (MPE and methane, respectively) to reduce molecular oxygen to water allowing for the oxidations leading to production of the hydroxy intermediate and methanol, respectively. The electrons in the further oxidation of the hydroxy intermediate and the keto intermediate to protochlorophyllide are also used together with additional electrons from ferredoxin to reduce oxygen.

enzymes that perform this reaction. One is an anaerobic enzyme where the C13¹ carbonyl group comes from water and the other is an oxygen-dependent enzyme where the carbonyl group instead comes from molecular oxygen (Walker et al., 1989; Porra et al., 1995; Wiessmann et al., 2020). It is the latter enzyme, from here on just the cyclase, which is found in plants (Totter et al., 2003; Chen et al., 2017) and is of special interest to this study. This enzyme is a carboxylate bridged diiron monooxygenase and contains a characteristic iron-binding motif EX_nEXXH_nEX_nEXXH where two iron atoms are bound by the glutamates and histidines (Berthold and Stenmark, 2003). This type of enzyme requires a reductant and the plant cyclase is reduced with electrons from ferredoxin (Stuart et al., 2020; Figure 1).

Plants and algae contain multiple chloroplast ferredoxins that are iron–sulfur proteins of the Fe₂S₂ type and catalyze single electron transfer reactions. The iron–sulfur cluster is coordinated by cysteines in a CX_nCX₂CX_nC motif. The pattern of chloroplast ferredoxin expression, in different parts and under different conditions, suggests that they play important functions in various metabolic pathways since they are central electron donors in chloroplast metabolism (Terauchi et al., 2009; Hanke and Mulo, 2013). The classical ferredoxin participates in photosynthesis and passes electrons to ferredoxin NADP⁺ oxidoreductase (FNR), which in turn reduces NADP⁺ to NADPH (Tagawa and Arnon, 1962). However, ferredoxins can directly donate electrons to ferredoxin dependent enzymes involved in processes such as

sulfur and nitrogen assimilation, tetrapyrrole metabolism, and fatty acid biosynthesis to name a few (Hanke and Mulo, 2013). Plants contain four major types of conserved ferredoxins (Hanke et al., 2004; Voss et al., 2011). The leaf-type is primarily expressed in photosynthetic tissues and accepts electrons from photosystem I as a part of photosynthetic electron transport while root-type is expressed mainly in nonphotosynthetic tissues and is tuned for accepting electrons from NADPH via FNR in order to support ferredoxin dependent enzymes (Shinohara et al., 2017). In addition to these canonical ferredoxins, plants contain two ferredoxins with an extended C-terminus and are thus referred to as FdC1 and FdC2 in *Arabidopsis thaliana* (Voss et al., 2011). Their functions are not well established. The FdC1 isoform has been shown to accept electrons from photosystem I or NADPH via FNR but does not seem to be able to reduce NADP⁺ via FNR (Voss et al., 2011; Guan et al., 2018). Similarly, FdC2 can also be reduced by photosystem I or NADPH via FNR and may be found in redox-dependent stromal ribonucleoprotein complexes (Kolton et al., 2011).

An *in vitro* cyclase assay utilizing recombinant proteins was recently developed (Stuart et al., 2020). The electrons for the reaction were provided by ferredoxin. In addition, it was found that the recombinant cyclase was only obtained in an active form when coexpressed with Ycf54 but Ycf54 itself did not seem to be required during catalysis (Stuart et al., 2020). Numerous *in vivo* studies have also demonstrated the importance of Ycf54 for the cyclase (Hollingshead et al., 2016; Chen et al., 2017, 2018; Chen and Hunter, 2020). In barley (*Hordeum vulgare* L.), the catalytic subunit of the cyclase is the XanL protein encoded by the *Xantha-1* gene. Three barley mutants, *xan-1.35*, *xan-1.81*, and *xan-1.82*, which are defective in the cyclase reaction due to mutations in the *Xantha-1* gene have been described previously (Rzeznicka et al., 2005). In addition, two mutants in the *Viridis-k* gene, *vir-k.23* and *vir-k.170* (Figure 2), have been isolated (Simpson and von Wettstein, 1980) and determined to be deficient in the cyclase reaction (Steccanella et al., 2015). Similar to the *xantha-1* mutations, the *viridis-k* mutations are lethal and have to be maintained in heterozygous stocks. Previous studies have established that the *Viridis-k* gene is clearly different from the *Xantha-1* gene and that it does not code for Ycf54 (Bollivar et al., 2014). We, therefore, set out to identify the *Viridis-k* gene to determine what role this additional component has in the cyclase reaction. As it turns out, the *Viridis-k* gene encodes the barley ortholog of *Arabidopsis* FdC2.

Results

Viridis-k is located on chromosome 4H

In order to map the location of the *Viridis-k* gene we constructed an F₂-mapping population by crossing the mutant *vir-k.23* to the barley cultivar Quench. As nothing was known previously about the location of the *Viridis-k* gene in the barley genome, we used a genotyping-by-sequencing approach on the F₂-mapping population to obtain single-

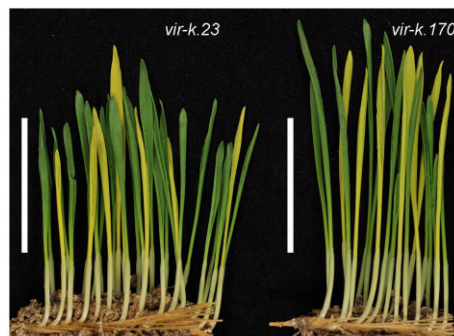


Figure 2 Segregation of *viridis-k* mutants. Ten-day-old seedlings germinated from a spike of heterozygous plants of *vir-k.23* and *vir-k.170*. Mutants homozygous for the recessive lethal *viridis-k* loss-of-function alleles are yellow due to deficiency in chlorophyll biosynthesis. The white bar is 5 cm.

nucleotide polymorphism (SNP) genotypes distributed throughout the whole genome. We sequenced 30 F₂ individuals (Supplemental Table S1) and obtained approximately 700 million paired-end reads in total. Raw reads were demultiplexed to assign reads to the correct individuals and to remove reads with errors in the barcode or the restriction enzyme recognition sequence. Overall, roughly 90% of reads were retained after demultiplexing. The analysis of SNPs showed a single region associated with the *viridis-k* phenotype on the short arm of chromosome 4H (Figure 3) which is expected from a trait that is controlled by a single locus. This region was approximately 9.6 Mbp long and located between base pair positions 6.5–16.1 Mbp on chromosome 4H. The region showed complete linkage to the *viridis-k* phenotype (Figure 2) and thus represented the interval where the *Viridis-k* gene is located.

Identification of a *Viridis-k* candidate gene

In order to narrow down the interval where the *Viridis-k* gene is located a larger F₂ population of roughly 300 individuals from the *vir-k.23* × Quench cross were genotyped with Cleaved Amplified Polymorphic Sequences (CAPSs) markers (Supplemental Table S2) designed for SNPs discovered during genotyping-by-sequencing. The recombination frequency between each marker and the *Viridis-k* locus was used to calculate the genetic distance measured in centi-Morgans (cM) between the marker and the *vir-k.23* mutation. This mapping was able to narrow down the interval to between 8.7–16.1 Mbp on chromosome 4H representing a region of ~7.4 Mbp. The two flanking markers located at 8.7 and 16.1 Mbp were calculated to be located at a genetic distance of 6.7 and 2.1 cM away from the *Viridis-k* gene, respectively. Surprisingly, a dramatic drop in recombination rate was observed that resulted in markers at 9.7, 10, and 13.7 Mbp showing a genetic distance of 0 cM to the *Viridis-k* gene (Supplemental Table S3). This region contained 149 high-

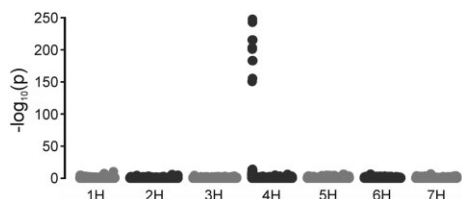


Figure 3 Manhattan plot of genome-wide association scan for chlorophyll deficiency in an F_2 -mapping population. The F_2 population originated from a cross between mutant *vir-k.23* and barley cultivar Quench. The significance level of SNP-marker trait association ($-\log_{10}$ of P -values) is shown as a function of the genomic location of the markers on the seven barley chromosomes (1H–7H). The Bonferroni corrected significance threshold is $-\log_{10}(P = 0.001/58903) > 7.8$. The analysis strongly suggests *Viridis-k* to be in the telomeric region of chromosome 4H.

confidence genes (Supplemental Data Set S1). The best candidate gene in the interval was the ferredoxin orthologous to Arabidopsis *FdC2*. As the cyclase is a ferredoxin-dependent enzyme, there is a direct functional link between *FdC2* and the biochemistry of the cyclase reaction (Stuart et al., 2020). We, therefore, investigated if the barley gene orthologous to *FdC2* contained any mutations in the two available *viridis-k* mutants.

Identification of genetic changes in *viridis-k* mutants

Sanger sequencing identified a single base pair substitution, a G–A transition, in the *vir-k.170* mutant. The transition changed a GCC codon to an ACC codon, which results in an alanine to threonine substitution (Ala-118-Thr; Figure 4). The G to A transition is consistent with the types of mutations that have previously been reported in barley after sodium azide mutagenesis (Olsen et al., 1993), which is the treatment used to generate *vir-k.170*. In the *vir-k.23* mutant the gene failed to be amplified by PCR using primers at the extreme ends of the gene but the 5′ and 3′ portions could be amplified separately (Figure 5). This suggested that the *vir-k.23* mutation is a larger chromosomal rearrangement that is consistent with the possible mutations induced by X-rays (Randolph, 1950; Hagberg and Tjio, 1951).

To further characterize the mutants, we performed a reverse-transcription PCR (RT-PCR) analysis of the *FdC2* ortholog in *vir-k.23* and *vir-k.170* as well as the control Bonus, which is the cultivar that the mutants were induced in. As the *vir-k.23* mutant appeared to have a chromosomal break in the middle of the gene, we designed two primer pairs, one pair on either side of the hypothesized break, to test for the presence of transcripts. This revealed that the 5′ portion of the gene was transcribed in both the mutants as well as the control. However, no transcript was detected covering the 3′ portion in the *vir-k.23* mutant even though this region could be amplified from genomic DNA and must therefore be present somewhere in the genome. The *vir-k.170* mutant and Bonus accumulated RNA

transcribed from the 3′ portion of the gene as expected (Figure 5).

The *viridis-k* phenotype can be complemented in vivo by *Viridis-k* but not by *Fd1*

To further confirm that the *viridis* phenotype was indeed due to disruption of the identified gene we set out to perform in vivo genetic complementation. Because of the lengthy process and genotype-specific requirements of performing stable transformations in barley (Harwood, 2014), we employed a transient expression approach based on agroinfiltration. This technique, has to the best of our knowledge, not previously been used to complement chlorophyll deficient mutants. To do this, we cloned full-length *Viridis-k* from cDNA obtained from the barley cultivar Bonus and inserted in a plant overexpression vector for *Agrobacterium tumefaciens* mediated transformation. As a control, the canonical photosynthetic ferredoxin *Fd1* was cloned into the same vector. The transient expression resulted in chlorophyll formation around the site of infiltration when *Viridis-k* was used but not *Fd1*. The successful complementation was best visualized with an imaging-PAM to reveal photosystem II chlorophyll fluorescence (Figure 6). This clearly showed presence of photosystem II bound chlorophyll, thus confirming that the suggested candidate gene, orthologous to Arabidopsis *FdC2*, is deficient in the *viridis-k* mutants and that *Viridis-k* is required for proper accumulation of chlorophyll in vivo. A search through the barley genome revealed eight plastid localized ferredoxins that are transcribed in the nucleus (Figure 7; Mayer et al., 2012; Colmsee et al., 2015; Mascher et al., 2017; Monat et al., 2019). The chlorophyll deficient phenotype of *viridis-k* mutants and the lack of complementation when leaf-type *Fd1* was used for agroinfiltration suggest that *Viridis-k* has a specific role that cannot be adequately performed by other ferredoxins in vivo.

Presence of cyclase components in known mutants

Immunoblot analysis was performed on total protein extracts from barley cyclase mutants, *xan-l.35*, *xan-l.81*, and *xan-l.82* as well as *vir-k.23* and *vir-k.170*. The *xan-l.35* mutant has a point mutation and a leaky phenotype producing small amounts of chlorophyll. The *xan-l.81* and *xan-l.82* mutants are both completely blocked at the cyclase step and have a single amino acid substitution and introduction of an early stop codon, respectively (Rzeznicka et al., 2005). Blots with antibodies against XanL showed a dramatic decrease of XanL in *xan-l.35*, while no XanL accumulated in *xan-l.82*. The *vir-k.23* and *vir-k.170* mutants accumulated XanL to similar levels as the control cultivar Bonus (Figure 8A). The same pattern was observed for immunoblots probed with antibodies raised against the cyclase-associated protein Ycf54, which shows that Ycf54 does not accumulate in the absence of XanL and that the levels of Ycf54 increase with increasing levels of XanL (Figure 8A). Immunoblots using an antibody that recognizes canonical ferredoxin show the presence in all lines. However, immunoblots using

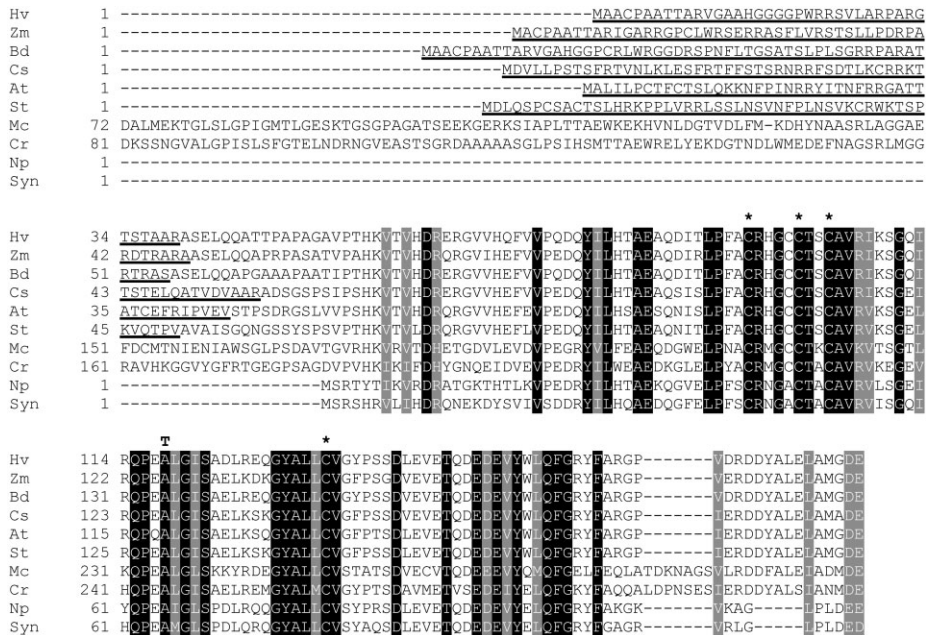


Figure 4 Amino-acid sequence alignment of VirK orthologs. Identical residues and residues with similar physical properties are boxed in black and gray, respectively. Cysteine residues participating in iron–sulfur cluster binding are indicated by asterisks. The Ala-118-Thr substitution caused by the *vir-k.170* mutation is indicated above the barley sequence. Sequences were aligned with Clustal Omega (Sievers et al., 2011). Chloroplast transit peptide sequences (underlined) were predicted by ChloroP (Emanuelsson et al., 1999). Hv, *H. vulgare* KAE8819615; Zm, *Z. mays* ACG28100; Bd, *Brachypodium distachyon* XP_003561243; Cs, *Cucumis sativus* XP_004148487; At, *A. thaliana* NP_0174533; St, *Solanum tuberosum* XP_006352847; Mc, *Micromonas commoda* XP_002501828 (the initial 71 residues have been omitted from the alignment); Cr, *C. reinhardtii* XP_001702961 (the initial 80 residues have been omitted); Np, *Nostoc punctiforme* WP_012407904; Syn, *Synechocystis* sp. PCC 6803 WP_010872870.

antibodies raised against VirK showed that no VirK protein accumulates in the *vir-k.23* mutant, whereas trace amounts were seen in the *vir-k.170* mutant (Figure 8A). Interestingly, VirK accumulation also increases with increasing XanL accumulation in the *xantha-1* mutants. That is, VirK decreased to hardly detectable levels in *xan-1.35* and failed to accumulate in the *xan-1.82* mutant that lacks XanL. The effect is apparently posttranslational since quantitative RT-PCR (RT-qPCR) analysis showed similar levels of *Viridis-k* mRNA in all mutants except *vir-k.23*, which has a large chromosomal rearrangement (Figure 8B). That VirK accumulation is apparently dependent on cellular accumulation of XanL indicates an important in vivo interaction.

VirK can function as an electron donor to the cyclase reaction

As the XanL is a ferredoxin-dependent enzyme and the VirK isoform is clearly associated with the cyclase reaction in vivo, it seems likely that VirK is the main electron donor to the enzyme. To test if VirK could provide electrons to drive the cyclase reaction we produced recombinant VirK and used it as the electron donor for cyclase enzyme activity

assays using recombinant XanL. The results clearly showed that VirK can drive the cyclase reaction with electrons provided by NADPH via FNR (Figure 9). Formation of the product, protochlorophyllide, increased with increasing VirK concentrations until saturation was reached, clearly showing that VirK can donate electrons to the cyclase enzyme. No product was formed if VirK was omitted from the reaction.

Discussion

Plants and other photosynthetic organisms have a rich variety of ferredoxins. A search of the barley nuclear genome revealed eight ferredoxins with a chloroplast transit peptide. A comparison of these to ferredoxins of maize (*Zea mays* L.) and Arabidopsis as well as the green alga *Chlamydomonas reinhardtii* and the cyanobacterium *Synechocystis* PCC6803 shows three classes of ferredoxins (Figure 7). Both of the plant ferredoxin isoforms with C-terminal extensions, FdC1 and FdC2, form well-supported subtrees. The FdC2 subtree contains VirK as well as orthologs in all species including *Synechocystis*. This is consistent with a highly conserved and ancient function which is supported by the essential nature of the gene in plants as well as cyanobacteria

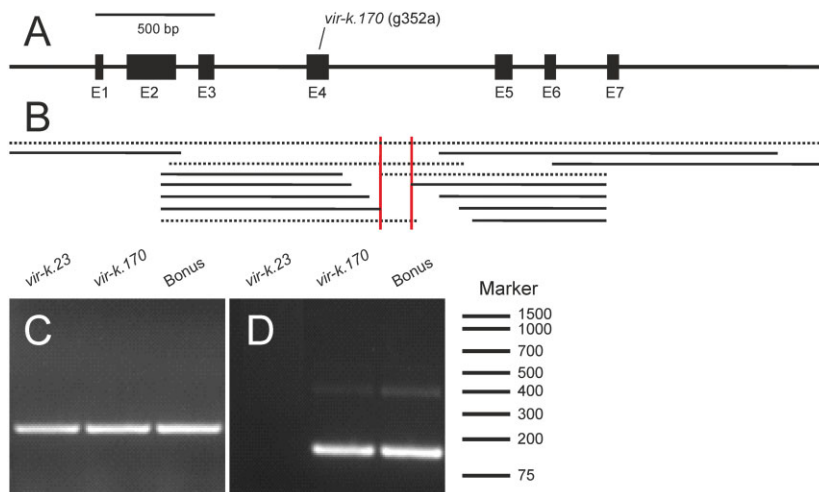


Figure 5 Gene structure of barley *Viridis-k*. A, The analyzed region of the *Viridis-k* gene consisting of seven exons; E1–E7. The *vir-k.170* point mutation is located in E4. B, The *vir-k.23* mutation was mapped by PCR using genomic DNA as template. Gene regions that could and could not be amplified are indicated by filled or dashed lines, respectively. The mapping and linkage analyses suggested a large chromosomal rearrangement interrupting the gene in the region between the two red vertical bars. C and D, RT-PCR amplicons separated by agarose gel electrophoresis and stained with Midori Green. First strand cDNA synthesis was performed with total RNA isolated from *vir-k.23*, *vir-k.170* and the cultivar Bonus, followed by end-point PCR. C, Amplification of a 237-bp cDNA fragment using a forward primer located in exon 1 and a reverse primer located in exon 2. D, Amplification of a 152-bp cDNA fragment using a forward primer located in exon 5 and a reverse primer located in exon 7. The 152-bp fragment could not be amplified from *vir-k.23*.

(Cassier-Chauvat and Chauvat, 2014; Schorsch et al., 2018). The FdC1 subtree does not contain a representative from *Synechocystis* but the Chlamydomonas FDX4 ortholog clusters into this group.

The canonical ferredoxin isoforms from plants were first characterized as “photosynthetic” and “nonphotosynthetic” but now are generally referred to as “leaf-type” and “root-type,” respectively. The separation into different categories was based on early studies which found that maize contains isoforms located primarily in green tissue (FdI and FdII) which are regulated by light and an isoform found primarily in roots (FdIII) where photosynthesis does not occur (Hase et al., 1991). Another apparent difference between these two categories is that the leaf-type ferredoxins generally have a more negative redox potential than the root type. Thus, the midpoint potential for NADPH is in-between that of leaf- and root-type ferredoxins. The midpoint potential of ferredoxin isoforms is likely an important biochemical property for predicting *in vivo* function since leaf-type ferredoxins are reduced by photosystem I in order to reduce NADP⁺ to NADPH via FNR, while root-type ferredoxins are reduced by NADPH via FNR to support ferredoxin dependent metabolism in the absence of photosynthesis (Shinohara et al., 2017). The leaf-type ferredoxins form a clear subtree and both Arabidopsis and barley have two apparent leaf-type ferredoxins. Maize has four leaf-type ferredoxins, which is likely due to a need for additional

specialization since maize has a C₄ metabolism. Consistent with this, FdI and FdII have been previously characterized as being specific for bundle sheath cells and mesophyll cells, respectively (Kimata and Hase, 1989; Matsumura et al., 1999). Arabidopsis contains, in addition to the two leaf-type ferredoxins, two more canonical ferredoxins which were classified as a root-type and an “other” type based on divergent sequence and a much higher redox potential than the root-type (Hanke et al., 2004). Both of the fall into the same subtree which also contains three maize ferredoxins including the root type. Barley also has three ferredoxins in this subtree. A third subtree separate from the leaf- and root-type ferredoxins contains only representatives from barley and maize, both of which are monocots. One of these, maize FdVI, has been shown to be induced in roots by nitrate (Matsumura et al., 1997). Ferredoxins in this subtree may thus have a monocot-specific role in nitrogen assimilation.

The Chlamydomonas and *Synechocystis* ferredoxins, except for SynFed3, are clearly most closely related to canonical plant ferredoxins although they are divergent enough to not cluster within the leaf and root types. In Chlamydomonas, the main photosynthetic ferredoxin is PETF/FDX1 that has a midpoint potential of −398 mV (Terauchi et al., 2009) which is similar to that of Arabidopsis and maize leaf-type ferredoxins which range from −433 to −406 mV (Hase et al., 1991; Hanke et al., 2004). The Chlamydomonas FDX2 is most closely related to PETF/FDX1 but has a midpoint

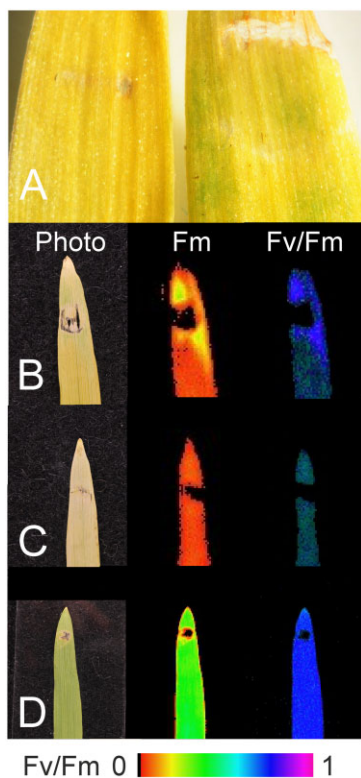


Figure 6 Complementation of the *vir-k.23* mutation by agroinfiltration. A. *tumefaciens* AGL0 harboring a plasmid containing either barley ferredoxin *Viridis-k* or ferredoxin *Fd1*, was infiltrated into seedling leaves of *vir-k.23*. Between 5 and 7 days postinfiltration, the leaves were inspected for presence of chlorophyll by visual inspection or by fluorescence using an imaging PAM. A. Green spots can be seen in a mutant seedling leaf injected with *Viridis-k* (right) but not with *Fd1* (left). B. Infiltration with *Viridis-k*. Chlorophyll was formed around the injection site. The maximum chlorophyll fluorescence (Fm) and the maximum quantum yield of photosystem II (Fv/Fm) are shown. A photo of the same leaf is shown to the left. C. Infiltration with *Fd1*. No chlorophyll could be detected. D. Photo and image of the mother cultivar Bonus infiltrated with *Viridis-k*.

potential of -321 mV (Terauchi et al., 2009) which is similar to that of Arabidopsis and maize root-type ferredoxins which have midpoint potentials of -337 and -321 mV, respectively (Hase et al., 1991; Hanke et al., 2004). Based on the midpoint potentials, FDX2 may have similar functions to plant root-type ferredoxins in supporting ferredoxin-dependent metabolism in the absence of photosynthesis.

Mutants are invaluable tools for revealing molecular processes in vivo. The yellow mutants of vascular plants are chlorophyll deficient and have been explored to learn about

genes and enzymes involved in chlorophyll biosynthesis. The biosynthetic step inhibited by a mutation is generally identified by accumulation of the substrate of the affected enzyme. However, a feedback mechanism prevents accumulation of chlorophyll biosynthetic intermediates unless the plants are fed with 5-aminolevulinic acid, which is the common precursor after the main regulatory step of the pathway. In this way, barley chlorophyll mutants were assigned to specific steps in the chlorophyll biosynthetic pathway (Gough, 1972). The *vir-k.23* and *vir-k.170* mutants were shown to be deficient in the cyclase reaction (Steccanella et al., 2015). In this study, we identified the *Viridis-k* locus as the gene orthologous to Arabidopsis *FdC2*, Chlamydomonas *FDX6*, and *Synechocystis* *SynFed2* encoding one of the two ferredoxin isoforms with C-terminal extension. Genetic mapping data identified a region containing 149 high-confidence genes including an *FdC2* ortholog, which appeared as the top candidate for several reasons. First, the known cyclase components are located in the chloroplast. Therefore, the gene product of the candidate gene should be localized to the chloroplast. The *FdC2* ortholog was one of 32 gene products in the mapping interval with a predicted chloroplast transit peptide. Second, the cyclase reaction requires reducing power and was recently demonstrated to be a ferredoxin-dependent enzyme (Stuart et al., 2020) making the *FdC2* ortholog the only chloroplast targeted candidate with an obvious functional association to the biochemistry of the cyclase enzyme. Previously, two mutants in the orthologous *FdC2* gene in rice had been isolated and were pale green due to decreased chlorophyll accumulation (Li et al., 2015; Zhao et al., 2015) but no further characterization or suggestions to their involvement in the cyclase reaction was reported. In addition, one would expect a gene required for chlorophyll biosynthesis to be highly conserved among oxygenic photosynthetic organisms. Recent studies on the evolution of photosystem I electron acceptors show that *FdC2* orthologs are conserved and found in all green plastids (green algae and land plants) as well as cyanobacteria from which chloroplasts derive (Karpowicz et al., 2011; Pierella Karlusich et al., 2015; Pierella Karlusich and Carrillo, 2017).

Sanger sequencing identified a single base-pair mutation resulting in an Ala-118-Thr mutation in *vir-k.170*. The change from alanine to threonine is a drastic change since alanine is a small nonpolar amino acid while threonine is bulkier and polar. In addition, Ala-118 is highly conserved across C-type ferredoxins from a range of species including land plants, green algae, and cyanobacteria (Figure 4) which indicates that substitution of this residue is not well tolerated.

The *vir-k.23* mutant has a lesion in the *Viridis-k* gene. The mutation is consistent with a large-scale genomic rearrangement such as an inversion, a large insertion or a translocation spanning several million base pairs, and that one of the break points is within the *Viridis-k* gene. The 5' and 3' regions of the gene could be amplified by PCR in the

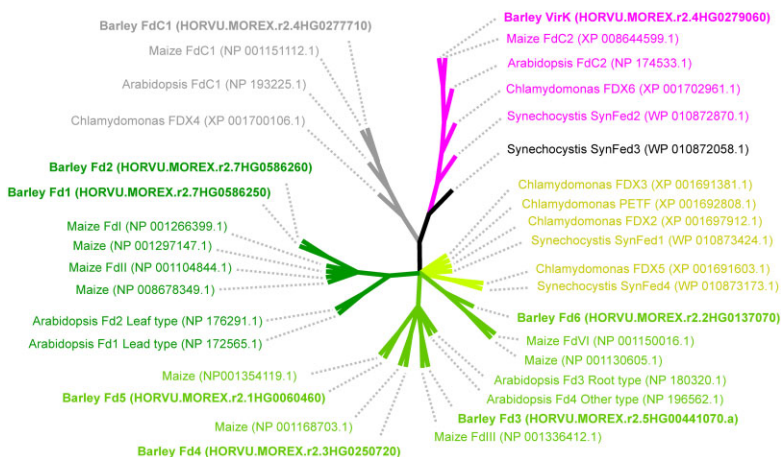


Figure 7 Maximum likelihood phylogenetic tree of plant-type ferredoxins. The eight barley ferredoxins are written in bold and used to identify ferredoxins by BLAST in maize, Arabidopsis, Chlamydomonas, and *Synechocystis* where ferredoxin isoforms have been previously characterized. The phylogeny is shown as an unrooted condensed bootstrap consensus tree generated in MEGA X and visualized using iTOL. Nodes with less than 50% bootstrap support were condensed. Branch lengths do not vary in bootstrap consensus trees generated in MEGA X and the tree should be interpreted as the maximum likelihood topology. Three main subtrees are apparent. These subtrees correspond to the canonical ferredoxins (dark green, light green, and lighter green) as well as the FdC1 (gray), and FdC2 (purple) types. The FdC2 subtree includes ferredoxins from all species while the FdC1 subtree does not contain a ferredoxin from *Synechocystis*. Canonical ferredoxins from plants have generally been classified as leaf-type and root-type. A leaf-type subtree of ferredoxins from plants is apparent (dark green) but the remaining canonical ferredoxins from plants (light green) do not form an obvious root-type subtree based on previously characterized ferredoxins. The main photosynthetic ferredoxins from Chlamydomonas and *Synechocystis* are PETF/FDX1 and SynFed1, respectively, located in the cluster indicated by lighter green.

vir-k.23 mutant from genomic DNA but amplification across the central portion was not possible in the mutant while it was possible in the control cultivar Bonus, as well as in *vir-k.170*. This indicates that both ends of the gene are present in the genome but that they are no longer positioned such that they are in the proper orientation or appropriate distance from each other to allow amplification by PCR or production of a full-length transcript. Combined with the CAPS mapping data of *viridis-k.23* which showed a drop in the recombination frequency to zero over about 4 Mbp, the most likely scenario is a large-scale genomic rearrangement which would be consistent with the mutagen used since ionizing radiation such as X-rays are known to be able to induce double-stranded DNA breaks (Randolph, 1950; Hagberg and Tjio, 1951). The presence of a lesion in the middle of the gene was also supported by RT-PCR which showed that the *vir-k.23* mutant accumulates transcripts from the 5' portion of the gene but not from the 3' portion. Immunoblots against VirK show that the protein does not accumulate in the *vir-k.23* mutant which is expected considering that the gene is disrupted by a large structural mutation. Although not completely absent, VirK accumulation was reduced in the *vir-k.170* mutant, which suggests that the Ala-118-Thr substitution affects the *in vivo* stability of the protein. The mutation could affect either the physical

stability of the protein and/or make it more prone to degradation by proteases. The level of *Viridis-k* mRNA in *vir-k.170* was not different from the mother cultivar Bonus (Figure 8).

Identification of multiple independent alleles is an efficient approach to validate a candidate gene. Over the years we have successfully used this approach where up to 215 alleles have been available; the *Eceriferum-cqu* locus involved in synthesis of epicuticular waxes as an example (Schneider et al., 2016). Unfortunately, only two *viridis-k* mutant alleles are available. To further support that we had identified the mutated gene in *viridis-k* mutants, we performed a transient expression of *Viridis-k* in seedling leaves of *vir-k.23* mutants. *A. tumefaciens* containing barley *Viridis-k* was infiltrated into the leaves. Typically, *Agrobacterium*-mediated infiltration is performed in dicot plants (Kapila et al., 1997; Wroblewski et al., 2005). After 5–7 days postinfiltration, chlorophyll formation could be seen around the site of infiltration conclusively demonstrating that the correct gene had been identified.

The *viridis-k* mutants are yellow, but trace amounts of chlorophyll can be detected (Steccanella et al., 2015). It is likely that this small amount of chlorophyll can be synthesized due to nonspecific action of other ferredoxins in the *vir-k.23* and *vir-k.170* mutants where VirK is absent or inactivated. Redundant cyclase activity supported by other ferredoxins might be expected since we recently showed that

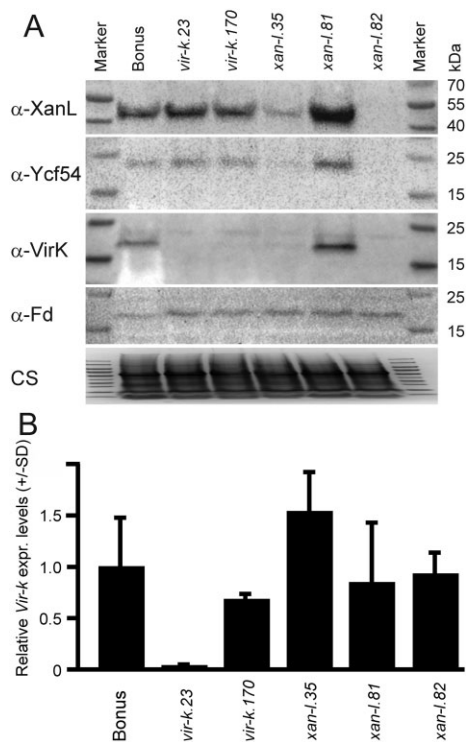


Figure 8 Immunoblots and RT-qPCR of barley mutants and the cultivar Bonus. A, Fifteen micrograms of total protein isolated from barley seedling leaves were separated by SDS-PAGE and transferred to PVDF membranes, which were incubated with antiserum raised against four different barley proteins as indicated in the figure. Coomassie-stained (CS) SDS-PAGE of the samples used in the immunoblot analyses showing that the same amount of protein has been loaded in each well. The CS image has been compressed. B, Total RNA was isolated and used for RT-qPCR analyses. The pair of primers used amplified the 5' region of the *Viridis-k* transcript (Supplemental Table S2). The relative expression of *Viridis-k* compared to Bonus is shown. Only the transcript level in *vir-k.23* is significantly different (Supplemental Table S6).

even leaf-type ferredoxin from spinach can drive the aerobic cyclase reaction in vitro using barley XanL (Stuart et al., 2020). Further, the aerobic cyclase, AcsF, of the purple bacterium *Rubrivivax gelatinosus* can be driven by ferredoxin from spinach and *Anabaena* (Chen et al., 2021). However, the *viridis-k* mutations are lethal and homozygous mutant seedlings die after approximately 2 weeks when the starch in the kernel is depleted. Thus, VirK plays a vital role in chlorophyll biosynthesis and its function cannot be meaningfully replaced by other ferredoxins in vivo. Additionally, *Viridis-k* is tightly connected to the cyclase enzyme system since we have shown that the steady state level of VirK is dependent

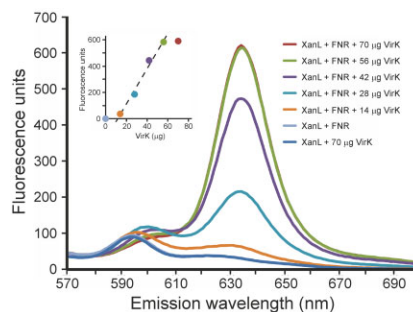


Figure 9 Enzymatic activity of recombinant VirK in combination with XanL and FNR. The XanL (40 μ g) used was coexpressed with Ycf54, and the FNR (13 μ g) was of the root type. Assay volumes were 30 μ L. Product formation increased linearly with the amount of added VirK in the range of 14–56 μ g (inset, $y = 13.6x - 163$, $R^2 = 0.986$, $P = 0.0071$). No activity was observed when VirK or FNR was omitted from the assay.

on the level of XanL. This suggests an intimate connection between XanL and VirK in vivo. The same phenomenon has been observed previously with other enzymes involved in chlorophyll biosynthesis. For example, the magnesium chelatase, which inserts Mg^{2+} into protoporphyrin IX consists of three subunits encoded in barley by *Xantha-h*, *Xantha-g*, and *Xantha-f* (von Wettstein et al., 1974; Jensen et al., 1996). The *xantha-h* mutants that do not produce a XanH protein also fail to accumulate the XanG protein which is known to interact directly with XanH (Hansson et al., 1999; Lake et al., 2004; Lundqvist et al., 2010). The decreased accumulation of VirK is unlikely to be an indirect effect due to the lack of chlorophyll biosynthesis since VirK accumulates in the *xan-l.81* mutant which is also completely blocked at the cyclase step but does accumulate XanL. In contrast, the *viridis-k* mutants accumulate canonical ferredoxins but still have a chlorophyll deficient phenotype, and the *vir-k.23* mutant could be complemented by introducing *Viridis-k* but not leaf-type *Fd1*. Taken together with the fact that XanL is a ferredoxin dependent enzyme and that VirK can donate electrons to XanL in vitro, it is likely that VirK is the electron donor to the cyclase reaction in vivo and that the other ferredoxin isoforms offer little redundancy when it comes to supplying electrons to the cyclase enzyme.

An alternative function of *Synechocystis* Fed2 has recently been suggested (Schorsch et al., 2018). In their study, SynFed2 was suggested to be involved in iron homeostasis. An explanation consistent with their data and the results presented in this paper is that the observed phenotypes are due to a deficiency in chlorophyll biosynthesis. The cyanobacterial study made use of C-terminal truncations since complete deletion of the gene was lethal. In the absence of a complete loss of function mutant in cyanobacteria it is

difficult to directly compare results, but the truncation mutants showed decreased growth and lower levels of chlorophyll accumulation.

As XanL is an iron-containing enzyme, one might speculate that the cyclase deficient phenotype in the *viridis-k* mutants is due to improper assembly of the diiron cluster of XanL. Improperly matured proteins tend to be degraded by the cell and this seems to be the case for improperly matured cyclase (Albus et al., 2012; Hollingshead et al., 2016; Chen and Hunter, 2020; Stuart et al., 2020). If the *viridis-k* mutations resulted in improper maturation of XanL we would expect to see decreased XanL levels in the *viridis-k* mutants that were not observed. In fact, the opposite was found—XanL is required for proper accumulation of VirK. Additionally, there was no apparent deficiency in the accumulation of canonical ferredoxin in the mutants which could also be expected if iron homeostasis was perturbed. The data are thus inconsistent with improperly matured XanL in the *viridis-k* mutants but are consistent with a function for VirK as the specific in vivo electron donor to the cyclase reaction.

Materials and methods

Plant material

The barley (*H. vulgare* L.) *vir-k.23* mutant was isolated in 1954 by X-ray mutagenesis of the cultivar Svalöv Weibull's Bonus (Simpson and von Wettstein, 1980). The *xan-L35* mutant was isolated in 1957 by ethyleneimine mutagenesis of Bonus while *xan-L81* and *-L82* mutants were isolated in 1975 by sodium azide mutagenesis of the *tigrina-d.12* mutant (Bonus genetic background) and screening for mutants that accumulated protochlorophyllide. The mutants were then backcrossed to Bonus in order to remove the *tigrina-d.12* mutation (Henningsen et al., 1993). Also in 1975, another mutant accumulating protochlorophyllide, originally named *xantha-83* was isolated which later turned out to be allelic to *vir-k.23* and thus renamed to *vir-k.170* (personal communications, Diter von Wettstein).

Mapping population

An F₂-mapping population was constructed by crossing the *vir-k.23* mutant to the cultivar Quench as the male parent. As the *vir-k.23* mutation is a recessive lethal mutation the female parent plants used were green and could be either homozygous wild-type or heterozygous for the *vir-k.23* mutation. Because of this, some F₁ plants were homozygous wild-type and their offspring could not be used for mapping. To eliminate these individuals, the seeds of F₁ plants were harvested on a per plant basis and one spike from each plant was grown in moist vermiculite for 9 days under a lab bench away from direct sunlight to check for segregation of the *vir-k.23* phenotype. Only seeds from F₁ plants whose offspring segregated were kept for the mapping population. The F₂-mapping population was grown under a lab bench away from direct sunlight. F₂-plant material was collected after 9 days.

Genomic DNA isolation

Genomic DNA for genotyping-by-sequencing library preparation or to be used as template for PCR was isolated by a modified CTAB protocol (Doyle, 1991). Between 100 and 200 mg leaf material was frozen under liquid nitrogen in screw cap tubes with two 4-mm glass beads and subsequently homogenized on a FastPrep 24 (MP Biomedicals, Santa Ana, CA) for 4 × 15 s at 4 m/s. Homogenized plant material was mixed with 1 mL 2 × CTAB buffer (2% [w/v] cetyltrimethylammonium bromide, 200 mM Tris-HCl pH 8, 20 mM EDTA, 1.4 M NaCl, 1% [w/v] polyvinylpyrrolidone [40 g/mol], 0.28 M β-mercaptoethanol) and incubated at 65°C for 30 min after which 800 μL chloroform:isoamylalcohol (24:1 [v/v]) was added and the samples were mixed on a rotary shaker for 15 min at room temperature. Samples were then centrifuged at 6,200g for 30 min at 4°C in order to separate the phases and 800 μL of the aqueous phase was transferred to a new tube and incubated at 37°C for 15 min with 50 μg RNase A. DNA was precipitated by addition of 560 μL isopropanol followed by centrifugation for 15 min at 15,700g and 4°C. The resulting pellet was then washed for 15 min in 76% (v/v) ethanol with 200 mM sodium acetate followed by a 5 min wash in 76% (v/v) ethanol with 10 mM ammonium acetate. After washing, the DNA pellet was allowed to dry at room temperature for 10 min prior to suspension in Low TE (10 mM Tris-HCl pH 8, 0.1 mM EDTA).

Genotyping-by-sequencing library preparation

Genotyping-by-sequencing library preparation for 30 F₂ seedlings was performed essentially as previously described (Poland et al., 2012) with minor modifications and using adaptor sequences as described in (Baird et al., 2008) with the modification that the P2 adaptors were modified to be compatible with *MspI* digested DNA ends. Briefly, 200 ng gDNA per individual was digested with 8 units *PstI*-HF (New England Biolabs) and 8 units *MspI* (New England Biolabs) in 1 × NEB buffer 4 (New England Biolabs, Ipswich, MA) by incubation for 2 h at 37°C. Sample specific inline barcoded P1 adaptors as well as a common P2 Y-adaptor that have overhangs complementary to *PstI* and *MspI*, respectively, were then ligated to the digested gDNA using 200 units T4 DNA ligase (New England Biolabs) in a reaction mixture containing all 200 ng restriction digested gDNA, 1 mM ATP, 2.5 nM P1 adaptor, and 375 nM P2 adaptor in 1 × NEB buffer 4. The ligation reaction was incubated for 2 h at 22°C. At this point samples were pooled and concentrated to 60 μL using a Qiagen Quiaquick PCR purification spin column (Qiagen, Hilden, Germany) following manufacturer instructions. Next, 20 μL was used as template for a 200 μL PCR, split into 8 × 25 μL reactions using NEB NEXT Q5 Ultra II master mix (New England Biolabs). PCR enriches for DNA fragments containing both P1 and P2 adaptors. Splitting of the PCR reaction mixture into multiple smaller reactions is done to minimize bias introduced by stochastic effects of PCR. After PCR the reactions were pooled and concentrated to 30 μL using a QiaQuick spin column (Qiagen) and run on

a 1.25% (w/v) agarose 1× TAE gel at 100 V for 1 h. The region containing molecules in the size range of 450–650 bp was cut out in order to size select the DNA fragments. DNA was purified from the excised gel using a JetSorb gel extraction kit (Genomed, Warszawa, Poland). The JetSorb gel extraction was performed according to the manufacturer's instructions with the modification that two extra washes with 80% (v/v) ethanol and 10 mM Tris–HCl pH 7.5 were added and the DNA was eluted in 60 µL Qiagen buffer EB. Sequencing was performed on an Illumina HiSeq 2500 with 125 cycles of paired-end sequencing with Illumina version 4 chemistry.

Genotyping-by-sequencing data analysis

Raw reads were demultiplexed using the `process_radtags` component of the STACKS (version 1.40) pipeline (Catchen et al., 2011). Phenotypes and index sequences for the F₂ individuals in the sequencing libraries can be found in Supplemental Table S1. `process_radtags` produced four files for each individual, two files with paired reads, and two files with unpaired read. The four files for each individual were combined into a single file for each individual. Reads were then aligned to the barley reference genome (Mascher et al., 2017) using BWA mem (version 0.7.13) to produce sam files (Li and Durbin, 2010). Reads aligned to the reference genome were then further processed through the stacks pipeline to output a list of SNPs in variant call format (.vcf).

The vcf output file was further analyzed using TASSEL (version 5.2.31) in order to locate SNPs that were associated with the *viridis* phenotype in our mapping population (Bradbury et al., 2007). The vcf file was loaded in TASSEL and sorted followed by filtering SNPs such that they needed to be present in 20 individuals and have a minimum allele frequency of 0.4 and a maximum allele frequency of 0.6 (an allele has an expected frequency of 0.5 in an F₂ population). Next, numerical values for phenotypes were assigned to the plants by setting wild-type to “0” and mutants to “100”. The General Linear Model (GLM) analysis was run with default settings except that “run permutations” was set to 10 and “Bi-allelic sites only” was selected since an SNP should have only two alleles in the mapping population. The output from the GLM analysis was then visualized to identify the genomic region linked to the mutant phenotype.

CAPS genotyping and genetic mapping

Additional F₂ individuals were genotyped using CAPS markers designed for SNPs discovered during genotyping-by-sequencing. For all markers, PCR was performed in 20 µL reactions using RedExtract-N-Amp (Sigma, St. Louis, MO, USA) following manufacturer's instructions. PCR product was then digested with restriction enzymes (New England Biolabs) specific for one allele of the SNP to be analyzed followed by separation on a 2% (w/v) agarose 1× TAE gel run for 22 min at 100 V. PCR primers used can be found in Supplemental Table S2. Five CAPS markers were tested located on chromosome 4H at 8.7 Mbp (4 units *Xba*I, Quench allele uncut), 9.7 Mbp (1 unit *Bts*CI, Quench allele cut),

10 Mbp (1 unit *Nco*I-HF, Quench allele cut), 13.7 Mbp (1 unit *Dde*I, Quench allele uncut), and 16.1 Mbp (1 unit *Nhe*I, Quench allele uncut). The recombination frequency between the SNPs and the *viridis-k* phenotype was calculated using a maximum likelihood approach (Fisher and Balmukand, 1928) and converted to genetic map units (cM) using the Kosambi mapping function (Kosambi, 1943).

PCR and sequencing of the *Viridis-k* gene

Sequencing of the *Viridis-k* gene was performed on overlapping PCR amplicons by Eurofins Genomics on an ABI 3730XL DNA Analyzer. PCR was run with either Phusion high fidelity DNA polymerase (New England Biolabs) or RedExtract-N-Amp Plant Tissue PCR Kit (Sigma) following manufacturer's guidelines. The PCR products were cleaned up using Illustra ExoProStar 1-Step (Cytiva, Marlborough, MA) following manufacturer's instructions except if unspecific bands were produced in which case the PCR products were run on a 2% (w/v) agarose 1× TAE gel for 30 min at 100 V. The correct bands were then cut out of the gel and cleaned up with NucleoSpin Gel and PCR clean-up columns (Macherey-Nagel, Bethlehem, PA) before sequencing.

RNA isolation, RT-PCR, and RT-qPCR

Plants were grown in vermiculite for 9 days under a lab bench away from direct sunlight and 300 mg material from a single leaf was harvested and immediately frozen in liquid nitrogen. Leaf material was homogenized under liquid nitrogen in a porcelain mortar. After homogenization total RNA was isolated using TRIzol reagent (Invitrogen Carlsbad, CA). Residual DNA was removed by treating total RNA preparations with DNase I (Thermo Scientific, Waltham, MA) after which cDNA was synthesized using 0.5 µg RNA per 20 µL reaction using RevertAid Reverse transcriptase (Thermo Scientific). The cDNA from *vir-k.23*, *vir-k.170* and Bonus were used for RT-PCR to detect *Viridis-k* transcripts and cDNA from Bonus was used for cloning of wild-type genes used in other experiments using Phusion high fidelity DNA polymerase (New England Biolabs).

For gene expression analysis using RT-qPCR, one leaf of three different plants was collected for RNA extraction per each biological replicate. Three technical replicates were performed on each biological sample. Total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions. RNA samples were treated by DNase on the columns using RNase-Free DNase Set (Qiagen). RNA integrity was checked on a 2% (w/v) agarose gel. The concentration and purity of the extracted RNA were measured spectrophotometrically. cDNA was synthesized from 1 µg of total RNA using the RevertAid First Strand cDNA synthesis kit (Thermo Scientific) following the manufacturer's instructions. Transcript levels were measured by using the SsoAdvanced Universal SYBR Green Supermix (Bio-Rad, Hercules, CA). Each reaction contained 20× diluted cDNA and primers at a final concentration of 0.25 µM. Samples were run on a CFX384TM Real-Time System (Bio-

Rad) using the following thermal cycling conditions: 95°C for 2 min, followed by 40 cycles of 95°C for 10 s and 60°C for 10 s. The used pair of primers amplified the 5' region of the *Viridis-k* transcript (Supplemental Table S2). For normalization, the housekeeping gene E3 ubiquitin-protein ligase *UPL6* gene (HORVU1Hr1G023480) was used. Results were expressed using the ΔC_t calculation method (Schmittgen and Livak, 2008).

In vivo genetic complementation by agroinfiltration

Full-length coding sequences of *Viridis-k* (MLOC_37911) and *Fd1* (MLOC_6135) were amplified from cDNA and cloned into the overexpression vector pK7WG2 (Karimi et al., 2002) by Gateway cloning (Invitrogen) to generate pK7WG2HvFdC2 and pK7WG2HvFd1. The pK7WG2 constructs were transformed into *A. tumefaciens* strain AGL0. Seeds of *vir-k.23* were planted in vermiculite, watered, and placed in a cold room (4°C), giving the seeds time to imbibe water and thereby synchronize germination. After 48 h the seeds were moved to a climate chamber (16/8 h of light/dark, 26.5°C, 35% relative humidity, and 8,900 Lux). Bacteria were inoculated into LB media (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl, pH 7) containing streptomycin (2 mg/mL). Cultures were incubated at 25°C with shaking for 48 h. New cultures were prepared by transferring bacteria from the two start cultures to two new cultures to OD₆₀₀ of 0.1 in LB media containing streptomycin (2 mg/mL) and acetosyringone (200 μM). Cultures were incubated overnight at 25°C with shaking. Bacteria were pelleted by centrifugation (7,400g for 40 min). Supernatants were discarded and pellets were resuspended in MM buffer (10 mM 2-(N-morpholino) ethanesulfonic acid pH 5.7 and 10 mM MgCl₂) containing 200 μM acetosyringone. The bacteria were washed by a second centrifugation (7,400g for 40 min). Supernatant was discarded and pellet resuspended to an OD₆₀₀ of 1.1 in MM buffer containing 300 μM acetosyringone. Cultures were then incubated for 1–2 h at room temperature. Ascorbic acid to 20 mM was added to cultures immediately prior to infiltration. Bacteria were delivered into the abaxial side of the leaf by first creating a small scratch with a needle and then pressure infiltrating bacteria with a needleless syringe. Infiltrated seedlings were placed in dark overnight, under a plastic hood to maintain humidity. The next day the infiltrated seedlings were moved to a climate chamber (same settings as before) for visual observations.

Chlorophyll fluorescence detection

Chlorophyll fluorescence was detected using an Imaging PAM M-series Maxi Version Chlorophyll Fluorometer (Heinz Walz, Effeltrich, Germany). The leaves were dark-adapted for a minimum of 10 min before analysis. The leaves were cut off directly prior to measurements, placed with adaxial side up under microscope slides to be held in position. Data were analyzed using ImagingWin (version2.47) software (Heinz Walz).

Immunoblots

Polyclonal rabbit antibodies against barley VirK were generated by Agrisera using recombinant VirK as the antigen. Other antibodies used were raised against barley proteins XanL (Bollivar et al., 2014), Ycf54 (Stuart et al., 2020), and purified ferredoxin (Andersen et al., 1992). Plants for total protein extraction were grown in vermiculite for 9 days under a lab bench away from direct sunlight and 200 mg tissue was snap frozen in liquid nitrogen in 2-mL screw cap vials with two 4 mm glass beads and 10 μL 2.5% (w/v) PMSF in isopropanol. Plant material was homogenized by grinding in a FastPrep 24 at 4 m/s for 10 s repeated 10 times with 4 min under liquid nitrogen between cycles to prevent material from thawing. Next, 500 μL protein extraction buffer (12 M urea, 2% [w/v] SDS, 100 mM DTT, 100 mM Tris-HCl pH 8, and 10 mM EDTA pH 8) was added, briefly heated to 60°C, and mixed by running four more cycles in the FastPrep 24. Samples were then incubated at 99°C and 1,400 rpm for 30 min in a Thermomixer comfort (Eppendorf, Hamburg, Germany) followed by centrifugation for 10 min at 16,000g. The supernatant was transferred to fresh tubes and centrifuged for another 30 min. The final supernatant was saved as the total protein extract. SDS-polyacrylamide gel (PAGE) and immunoblots were performed as described previously (Stuart et al., 2020).

Recombinant protein production and cyclase activity assay

Recombinant XanL was prepared as described previously (Stuart et al., 2020). The coding sequence of *Viridis-k* without the transit peptide as predicted by ChloroP (Emanuelsson et al., 1999) was codon optimized and cloned into pET15b by GenScript Biotech Corporation to generate pET15bHvFdC2. *Escherichia coli* BL21(DE3) transformed with pET15bHvFdC2 was inoculated to an OD₆₀₀ of 0.1 in 250 mL LB containing 100 μg/mL ampicillin and placed in an incubator at 20°C and 200 rpm. After 1 h the temperature was set down to 15°C. Once the culture reached an OD₆₀₀ of 0.3, IPTG was added to a final concentration of 1 mM and solid FeSO₄ was added to a final concentration of 2 mM. After 24 h another 2 mM equivalent of solid FeSO₄ was added and the culture was grown for an additional 36 h before the cells were harvested by centrifugation and stored at –80°C until use. For protein purification a cell pellet was resuspended to 50 mL in binding buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, 20 mM imidazole, and 5 mM DTT) supplemented with 3 M urea as well as a few grains of DNase I and lysozyme. The cell suspension was passed through a french press three times at 12.4 MPa after which the lysate was centrifuged 10 min at 48,384g. The supernatant was loaded on two 1 mL HisTrap FF crude (Cytiva) columns connected in series and washed four times with 15 mL binding buffer supplemented with decreasing concentrations of urea. The urea concentration was decreased from 3 M to no urea in 1 M increments. The column was then washed with 15 mL wash buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, 40 mM imidazole, and 5 mM DTT) followed by elution with 20 mM Tris–

HCl pH 8, 500 mM NaCl, 750 mM imidazole, and 5 mM DTT. The VirK protein was desalted over a NAP-10 (Cytiva) column into 50 mM Tris–HCl pH 8 with 1 mM DTT, after which 87% (v/v) glycerol was added to a final concentration of 15% (v/v). The VirK protein was aliquoted and stored at -80°C until use.

The coding sequence of barley root isoform of FNR (HvRFNR, MLOC_6838) was cloned from cDNA without the chloroplast transit peptide as predicted by ChloroP into the vector pDEST17 by Gateway cloning to produce pDEST17HvRFNR. *E. coli* Rosetta (DE3)pLys transformed with pDEST17HvRFNR was inoculated to an OD_{600} of 0.1 into 250 mL LB supplemented with 100 $\mu\text{g}/\text{mL}$ ampicillin and 25 $\mu\text{g}/\text{mL}$ chloramphenicol and grown at 25°C and 200 rpm until the culture reached an OD_{600} of 0.4, after which protein expression was induced by adding IPTG to a final concentration of 1 mM. Cells were harvested by centrifugation the next morning after 12–16 h of growth postinduction and cell pellets were frozen at -80°C until use. A cell pellet corresponding to 250 mL culture was resuspended to 20 mL in binding buffer supplemented with 100 $\mu\text{g}/\text{mL}$ lysozyme after which cells were disrupted by sonication on ice for a total of 4 min at 50% output. The lysate was centrifuged 8 min at 48,384g, after which the supernatant was loaded onto a 1 mL HisTrap FF crude column. The column was washed with 5 mL binding buffer followed by 10 mL wash buffer. The recombinant HvRFNR was eluted with 20 mM Tris–HCl pH 8, 500 mM NaCl, 250 mM imidazole, and 5 mM DTT and desalted over a NAP-10 column into buffer consisting of 1 mM DTT, 25 mM MgCl_2 , 1 mM EDTA, 20 mM Tricine, 10 mM HEPES, and adjusted to pH 8.1 with NaOH. After desalting, 87% (v/v) glycerol was added to a final concentration of 15% (v/v) and the protein was aliquoted and frozen at -80°C until use. Cyclase assays were run as described previously (Stuart et al., 2020) except that the above recombinant HvRFNR and VirK were used instead of spinach FNR and ferredoxin from Sigma. MPE was extracted from a *bchE* mutant of *Rhodobacter capsulatus* as previously described (Gough et al., 2007). A measurement series consisted of six enzymatic reactions with different amounts (0–70 μg) of added VirK but constant amounts of XanL (40 μg) and HvRFNR (13 μg) in total volumes of 30 μL containing 20 mM Tricine and 10 mM HEPES pH 8.1, 1 mM EDTA, 25 mM MgCl_2 , 1 mM DTT, 10 mM glucose-6-phosphate, 0.03 units/ μL glucose-6-phosphate dehydrogenase, 0.5 mM NADPH, 90 $\mu\text{g}/\mu\text{L}$ catalase, 0.026% (v/v) Triton X-100, and 10 μM MPE. Assays were incubated in the dark at 30°C and 750 rpm for 1 h in a Thermomixer comfort (Eppendorf). Assays were stopped by the addition of 80% (v/v) acetone with 0.32% (v/v) NH_3 and centrifuged for 5 min at $29,000 \times g$ to pellet precipitated proteins. Formation of protochlorophyllide was measured using an RF-5301 PC spectrofluorometer (Shimadzu, Kyoto, Japan) with an excitation wavelength of 440 nm and an emission spectrum between 570 and 700 nm with slit widths of 10 nm for both excitation and emission. Product formation

was estimated as the relative fluorescence emission at 634 nm for samples minus the emission for the negative control sample. Linear regression was performed with the $\text{lm}()$ function in R (version 3.6.1).

Phylogenetic analysis

Ferredoxin sequences for Arabidopsis, maize, Chlamydomonas, and *Synechocystis* were retrieved by BLAST against the NCBI Protein reference sequences database with an E value cutoff of 10^{-20} for a match using each of the barley ferredoxins as queries. In order to avoid bias from erroneous prediction of chloroplast transit peptides all sequences were N-terminally trimmed to the shortest *Synechocystis* ferredoxin after removal of the initiator methionine as these do not provide a phylogenetic signal. Sequences were aligned using ClustalW as implemented in MEGA-X (version 10.2.5). The alignment in FASTA format can be found in Supplemental File S1. MEGA-X was used to determine the best fitting maximum likelihood model after which a maximum likelihood tree was inferred using the LG +G model with 16 rate categories and gamma parameter equal to 1.6872. Node support was determined using 1,000 bootstrap replicates. Nodes with less than 50% support were condensed in the bootstrap consensus tree. The tree was visualized as an unrooted tree using iTOL (Letunic and Bork, 2021). The tree in Newick format can be found in Supplemental File S2.

Accession numbers

Sequence data from this article can be found in the NCBI Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/sra>) under accession no. PRJNA686392.

Supplemental data

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

Supplemental Table S1. Phenotype and index sequence for F_2 individuals included in the genotyping-by-sequencing libraries for mapping of the *Viridis-k* gene.

Supplemental Table S2. List of used DNA oligonucleotides.

Supplemental Table S3. Positions and genetic distances of markers to the *Viridis-k* gene.

Supplemental Table S4. Ct values from RT-qPCR of the 5' region of the *Viridis-k* gene.

Supplemental Table S5. Ct values from RT-qPCR of the housekeeping gene E3 ubiquitin-protein ligase *UPL6* (HORVU1Hr1G023480).

Supplemental Table S6. Calculations of relative expression of *Viridis-k* in the five mutants and their mother cultivar Bonus.

Supplemental Data Set S1. List of genes in the mapped interval for *Viridis-k*.

Supplemental File S1. Multiple sequence alignment of ferredoxin peptide sequences used for constructing phylogenetic relations in Figure 7.

Supplemental File S2. Phylogenetic tree (Figure 7) in Newick format.

Acknowledgments

We thankfully acknowledge the computer resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project SNIC 2019/3-378. We acknowledge SciLifeLab for next-generation sequencing.

Funding

This work was supported by the VILLUM Foundation (grant no. 13363 to P.E.J.), the German Research Foundation (grant no. DFG YO 304/1-1 to H.M.Y.), the Swedish Research Council (grant no. VR 2018-05117 to M.H.), the Swedish Research Council for Environment, Agricultural Sciences, and Spatial Planning (grant no. FORMAS 2018-01026 to M.H.), the Erik Philip-Sörensen Foundation (M.H.), the Royal Physiographic Society in Lund (D.S., H.M.Y., M.H., S.Z.), and an Artistic and Scholarly Development grant from Illinois Wesleyan University (D.B.).

Conflict of interest statement. None declared.

References

- Albus CA, Salinas A, Czarniecki O, Kahlau S, Rothbart M, Thiele W, Lein W, Bock R, Grimm B, Schottler MA (2012) LCAA, a novel factor required for magnesium protoporphyrin monomethyl-ester cyclase accumulation and feedback control of aminolevulinic acid biosynthesis in tobacco. *Plant Physiol* **160**: 1923–1939
- Andersen B, Koch B, Scheller HV (1992) Structural and functional analysis of the reducing side of photosystem I. *Physiol Plant* **84**: 154–161
- Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko W, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* **3**: e3376
- Banerjee R, Jones JC, Lipscomb JD (2019) Soluble methane monooxygenase. *Annu Rev Biochem* **88**: 409–431
- Berthold DA, Stenmark P (2003) Membrane-bound diiron carboxylate proteins. *Annu Rev Plant Biol* **54**: 497–517
- Bollivar D, Braumann I, Berendt K, Gough SP, Hansson M (2014) The Ycf54 protein is part of the membrane component of Mg-protoporphyrin IX monomethyl ester cyclase from barley (*Hordeum vulgare* L.). *FEBS J* **281**: 2377–2386
- Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**: 2633–2635
- Bryant DA, Hunter CN, Warren MJ (2020) Biosynthesis of the modified tetrapyrroles—the pigments of life. *J Biol Chem* **295**: 6888–6925
- Cassier-Chauvat C, Chauvat F (2014) Function and regulation of ferredoxins in the cyanobacterium, *synechocystis* PCC6803: recent advances. *Life (Basel)* **4**: 666–680
- Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci *de novo* from short-read sequences. *G3* **1**: 171–182
- Chen GE, Hunter CN (2020) Protochlorophyllide synthesis by recombinant cyclases from eukaryotic oxygenic phototrophs and the dependence on Ycf54. *Biochem J* **477**: 2313–2325
- Chen GE, Canniffe DP, Hunter CN (2017) Three classes of oxygen-dependent cyclase involved in chlorophyll and bacteriochlorophyll biosynthesis. *Proc Natl Acad Sci USA* **114**: 6280–6285
- Chen GE, Adams NBP, Jackson PJ, Dickman MJ, Hunter CN (2021) How the O₂-dependent Mg-protoporphyrin monomethyl ester cyclase forms the fifth ring of chlorophylls. *Nat Plants* **7**: 365–375
- Chen GE, Canniffe DP, Barnett SFH, Hollingshead S, Brindley AA, Vasilev C, Bryant DA, Hunter CN (2018) Complete enzyme set for chlorophyll biosynthesis in *Escherichia coli*. *Sci Adv* **4**: eaaq1407
- Colmsee C, Beier S, Himmelbach A, Schmutzer T, Stein N, Scholz U, Mascher M (2015) BARLEX - the barley draft genome explorer. *Mol Plant* **8**: 964–966
- Doyle J (1991) DNA protocols for plants taxonomy. In GM Hewitt, AWB Johnston, JPW Young eds, *Molecular Techniques*. Springer, Berlin, Heidelberg, Germany, pp 283–293
- Emanuelsson O, Nielsen H, von Heijne G (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci* **8**: 978–984
- Fisher RA, Balmukand B (1928) The estimation of linkage from the offspring of selfed heterozygotes. *J Genet* **20**: 79–92
- Gough S (1972) Defective synthesis of porphyrins in barley plastids caused by mutation in nuclear genes. *Biochim Biophys Acta* **286**: 36–54
- Gough SP, Rzeznicka K, Peterson Wulff R, da Cruz Francisco J, Hansson A, Jensen PE, Hansson M (2007) A new method for isolating physiologically active Mg-protoporphyrin monomethyl ester, the substrate of the cyclase enzyme of the chlorophyll biosynthetic pathway. *Plant Physiol Biochem* **45**: 932–936
- Guan X, Chen S, Voon CP, Wong KB, Tikkanen M, Lim BL (2018) FdC1 and leaf-type ferredoxins channel electrons from photosystem I to different downstream electron acceptors. *Front Plant Sci* **9**: 410.
- Hagberg A, Tjio JH (1951) Cytological studies on some X-ray mutants of barley. *An Estac Exp Aula Dei* **2**: 149–167
- Hanke G, Mulo P (2013) Plant type ferredoxins and ferredoxin-dependent metabolism. *Plant Cell Environ* **36**: 1071–1084
- Hanke GT, Kimata-Arigo Y, Taniguchi I, Hase T (2004) A post genomic characterization of Arabidopsis ferredoxins. *Plant Physiol* **134**: 255–264
- Hansson A, Kannangara CG, von Wettstein D, Hansson M (1999) Molecular basis for semidominance of missense mutations in the XANTHA-H (42-kDa) subunit of magnesium chelatase. *Proc Natl Acad Sci USA* **96**: 1744–1749
- Harwood WA (2014) A protocol for high-throughput *Agrobacterium*-mediated barley transformation. *Methods Mol Biol* **1099**: 251–260
- Hase T, Kimata Y, Yonekura K, Matsumura T, Sakakibara H (1991) Molecular cloning and differential expression of the maize ferredoxin gene family. *Plant Physiol* **96**: 77–83
- Henningsen KW, Boynton JE, Wettstein DV (1993) Mutants at *Xantha* and *Albina* Loci in Relation to Chloroplast Biogenesis in Barley (*Hordeum vulgare* L.). The Royal Danish Academy of Sciences and Letters, Copenhagen
- Hollingshead S, Kopečna J, Armstrong DR, Bucinska L, Jackson PJ, Chen GE, Dickman MJ, Williamson MP, Sobotka R, Hunter CN (2016) Synthesis of chlorophyll-binding proteins in a fully segregated *Dycf54* strain of the cyanobacterium *Synechocystis* PCC 6803. *Front Plant Sci* **7**: 292
- Jensen PE, Willows RD, Petersen BL, Vothknecht UC, Stummann BM, Kannangara CG, von Wettstein D, Henningsen KW (1996) Structural genes for Mg-chelatase subunits in barley: *Xantha-f*, *-g* and *-h*. *Mol Gen Genet* **250**: 383–394
- Kapila J, De Rycke R, Van Montagu M, Angenon G (1997) An *Agrobacterium*-mediated transient gene expression system for intact leaves. *Plant Sci* **122**: 101–108
- Karimi M, Inzé D, Depicker A (2002) GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci* **7**: 193–195

- Karpowicz SJ, Prochnik SE, Grossman AR, Merchant SS** (2011) The GreenCut2 resource, a phylogenomically derived inventory of proteins specific to the plant lineage. *J Biol Chem* **286**: 21427–21439
- Kimata Y, Hase T** (1989) Localization of ferredoxin isoproteins in mesophyll and bundle sheath cells in maize leaf. *Plant Physiol* **89**: 1193–1197
- Kolton M, Keren I, Shevtsov S, Shaya F, Peled-Zehavi H, Danon A, Ostersefer-Biran O** (2011) Plastidic redox switches: ferredoxins as novel RNA-binding proteins. *Endocyt Cell Res* **21**: 1–18
- Kosambi DD** (1993) The estimation of map distances from recombination values. *Ann Eugen* **12**: 172–175
- Lake V, Olsson U, Willows RD, Hansson M** (2004) ATPase activity of magnesium chelatase subunit I is required to maintain subunit D *in vivo*. *Eur J Biochem* **271**: 2182–2188
- Letunic I, Bork P** (2021) Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* **23**: 127–128
- Li C, Hu Y, Huang R, Ma X, Wang Y, Liao T, Zhong P, Xiao F, Sun C, Xu Z, et al.** (2015) Mutation of *FdC2* gene encoding a ferredoxin-like protein with C-terminal extension causes yellow-green leaf phenotype in rice. *Plant Sci* **238**: 127–134
- Li H, Durbin R** (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**: 589–595
- Lundqvist J, Elmlund H, Wulff RP, Berglund L, Elmlund D, Emanuelsson C, Hebert H, Willows RD, Hansson M, Lindahl M, et al.** (2010) ATP-induced conformational dynamics in the AAA+ motor unit of magnesium chelatase. *Structure* **18**: 354–365
- Mascher M, Gundlach H, Himmelbach A, Beier S, Twardziok SO, Wicker T, Radchuk V, Dockter C, Hedley PE, Russell J, et al.** (2017) A chromosome conformation capture ordered sequence of the barley genome. *Nature* **544**: 427–433
- Matsumura T, Sakakibara H, Nakano R, Kimata Y, Sugiyama T, Hase T** (1997) A nitrate-inducible ferredoxin in maize roots. Genomic organization and differential expression of two nonphotosynthetic ferredoxin isoproteins. *Plant Physiol* **114**: 653–660
- Matsumura T, Kimata-Ariga Y, Sakakibara H, Sugiyama T, Murata H, Takao T, Shimonishi Y, Hase T** (1999) Complementary DNA cloning and characterization of ferredoxin localized in bundle-sheath cells of maize leaves. *Plant Physiol* **119**: 481–488
- Mayer KF, Waugh R, Brown JW, Schulman A, Langridge P, Platzer M, Fincher GB, Muehlbauer GJ, Sato K, Close TJ, et al.** (2012) A physical, genetic and functional sequence assembly of the barley genome. *Nature* **491**: 711–716
- Monat C, Padmarasu S, Lux T, Wicker T, Gundlach H, Himmelbach A, Ens J, Li C, Muehlbauer GJ, Schulman AH, et al.** (2019) TRITEX: chromosome-scale sequence assembly of Triticeae genomes with open-source tools. *Genome Biol* **20**: 284
- Olsen O, Wang X, von Wettstein D** (1993) Sodium azide mutagenesis: preferential generation of A-T → G-C transitions in the barley *Ant18* gene. *Proc Natl Acad Sci USA* **90**: 8043–8047
- Pierella Karlusich JJ, Carrillo N** (2017) Evolution of the acceptor side of photosystem I: ferredoxin, flavodoxin, and ferredoxin-NADP⁺ oxidoreductase. *Photosynth Res* **134**: 235–250
- Pierella Karlusich JJ, Ceccoli RD, Graña M, Romero H, Carrillo N** (2015) Environmental selection pressures related to iron utilization are involved in the loss of the flavodoxin gene from the plant genome. *Genome Biol Evol* **7**: 750–767
- Poland JA, Brown PJ, Sorrells ME, Jannink JL** (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One* **7**: e32253
- Porra RJ, Schafer W, Katheder I, Scheer H** (1995) The derivation of the oxygen atoms of the 13¹-oxo and 3-acetyl groups of bacteriochlorophyll *a* from water in *Rhodospirillum rubrum* cells adapting from respiratory to photosynthetic conditions: evidence for an anaerobic pathway for the formation of isocyclic ring E. *FEBS Lett* **371**: 21–24
- Randolph LF** (1950) Cytological and phenotypical effects induced in maize by x-rays and the bikini test able atomic bomb. *J Cell Physiol Suppl* **35**: 103–117
- Rzeznicka K, Walker CJ, Westergren T, Kannangara CG, von Wettstein D, Merchant S, Gough SP, Hansson M** (2005) *Xantha-I* encodes a membrane subunit of the aerobic Mg-protoporphyrin IX monomethyl ester cyclase involved in chlorophyll biosynthesis. *Proc Natl Acad Sci USA* **102**: 5886–5891
- Schmittgen TD, Livak KJ** (2008) Analyzing real-time PCR data by the comparative C_T method. *Nat Protoc* **3**: 1101–1108
- Schneider LM, Adamski NM, Christensen CE, Stuart DB, Vautrin S, Hansson M, Uauy C, von Wettstein-Knowles P** (2016) The *Cer-cgu* gene cluster determines three key players in a β-diketone synthase polyketide pathway synthesizing aliphatics in epicuticular waxes. *J Exp Bot* **67**: 2715–2730
- Schorsch M, Kramer M, Goss T, Eisenhut M, Robinson N, Osman D, Wilde A, Sadaf S, Brückler H, Walder L, et al.** (2018) A unique ferredoxin acts as a player in the low-iron response of photosynthetic organisms. *Proc Natl Acad Sci USA* **115**: E12111–E12120
- Shinohara F, Kurisu G, Hanke G, Bowsher C, Hase T, Kimata-Ariga Y** (2017) Structural basis for the isotope-specific interactions of ferredoxin and ferredoxin: NADP⁺ oxidoreductase: an evolutionary switch between photosynthetic and heterotrophic assimilation. *Photosynth Res* **134**: 281–289
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, et al.** (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539
- Simpson DJ, von Wettstein D** (1980) Macromolecular physiology of plastids XIV. *Viridis* mutants in barley: genetic, fluoroscopic and ultrastructural characterisation. *Carlsberg Res Commun* **45**: 283–314
- Steccanella V, Hansson M, Jensen PE** (2015) Linking chlorophyll biosynthesis to a dynamic plastoquinone pool. *Plant Physiol Biochem* **97**: 207–216
- Stuart D, Sandström M, Youssef HM, Zakhrebekova S, Jensen PE, Bollivar DW, Hansson M** (2020) Aerobic barley Mg-protoporphyrin IX monomethyl ester cyclase is powered by electrons from ferredoxin. *Plants* **9**: 1157
- Tagawa K, Arnon DI** (1962) Ferredoxins as electron carriers in photosynthesis and in the biological production and consumption of hydrogen gas. *Nature* **195**: 537–543
- Tanaka R, Tanaka A** (2007) Tetrapyrrole biosynthesis in higher plants. *Annu Rev Plant Biol* **58**: 321–346
- Terauchi AM, Lu SF, Zabagnini M, Tappa S, Hirasawa M, Tripathy JN, Knaff DB, Farmer PJ, Lemaire SD, Hase T, et al.** (2009) Pattern of expression and substrate specificity of chloroplast ferredoxins from *Chlamydomonas reinhardtii*. *J Biol Chem* **284**: 25867–25878
- Totter S, Block MA, Allen M, Westergren T, Albrieux C, Scheller HV, Merchant S, Jensen PE** (2003) *Arabidopsis* CHL27, located in both envelope and thylakoid membranes, is required for the synthesis of protochlorophyllide. *Proc Natl Acad Sci USA* **100**: 16119–16124
- Walker CJ, Mansfield KE, Smith KM, Castelfranco PA** (1989) Incorporation of atmospheric oxygen into the carbonyl functionality of the protochlorophyllide isocyclic ring. *Biochem J* **257**: 599–602
- Wiesselmann M, Hebecker S, Borrero-de Acuña JM, Nitz M, Bollivar DW, Jänsch L, Moser J, Jahn D** (2020) Mg-protoporphyrin IX monomethyl ester cyclase from *Rhodospirillum rubrum*: Radical SAM-dependent synthesis of the isocyclic ring of bacteriochlorophylls. *Biochem J* **477**: 4635–4654
- von Wettstein D, Kahn A, Nielsen OF, Gough S** (1974) Genetic regulation of chlorophyll synthesis analyzed with mutants in barley. *Science* **184**: 800–802

- Voss I, Goss T, Murozuka E, Altmann B, McLean KJ, Rigby SE, Munro AW, Scheibe R, Hase T, Hanke GT** (2011) FdC1, a novel ferredoxin protein capable of alternative electron partitioning, increases in conditions of acceptor limitation at photosystem I. *J Biol Chem* **286**: 50–59
- Wroblewski T, Tomczak A, Michelmore R** (2005) Optimization of *Agrobacterium*-mediated transient assays of gene expression in lettuce, tomato and *Arabidopsis*. *Plant Biotechnol J* **3**: 259–273
- Zhao J, Qiu Z, Ruan B, Kang S, He L, Zhang S, Dong G, Hu J, Zeng D, Zhang G, et al.**(2015) Functional inactivation of putative photosynthetic electron acceptor ferredoxin C2 (FdC2) induces delayed heading date and decreased photosynthetic rate in rice. *PLoS One* **10**: e0143361

Supplemental Table S1. Phenotype and index sequence for F₂ individuals included in the genotyping-by-sequencing libraries for mapping of the *Viridis-k* gene.

F₂ Individual	Sequencing Library	Index sequence	Phenotype
MH84-134-1	1	ACATAGG	Wild Type
MH84-103-3	1	CGAACTG	Wild Type
MH84-106-3	1	GTACGTG	Wild Type
MH84-115-3	1	TCTCTCA	<i>vir-k.23</i>
MH84-119-3	1	ACGATTG	Wild Type
MH84-120-3	1	AGTCTGT	Wild Type
MH84-122-3	1	ATTACGG	Wild Type
MH84-126-3	1	CACTAGT	Wild Type
MH84-130-3	1	CATCGTA	<i>vir-k.23</i>
MH84-133-3	1	CGACAAT	Wild Type
MH84-134-3	1	GACCACA	Wild Type
MH84-103-5	1	GGTAACT	<i>vir-k.23</i>
MH84-106-5	1	GTCTTGG	Wild Type
MH84-119-5	1	TACGCTA	Wild Type
MH84-120-5	1	TCAGGAA	Wild Type
MH84-134-2	2	ACATAGG	Wild Type
MH84-103-4	2	CGAACTG	<i>vir-k.23</i>
MH84-106-4	2	GTACGTG	Wild Type
MH84-115-4	2	TCTCTCA	Wild Type
MH84-119-4	2	ACGATTG	Wild Type
MH84-120-4	2	AGTCTGT	<i>vir-k.23</i>
MH84-122-4	2	ATTACGG	Wild Type
MH84-126-4	2	CACTAGT	Wild Type
MH84-130-4	2	CATCGTA	<i>vir-k.23</i>
MH84-133-4	2	CGACAAT	<i>vir-k.23</i>
MH84-134-4	2	GACCACA	Wild Type
MH84-122-5	2	GGTAACT	<i>vir-k.23</i>
MH84-126-5	2	GTCTTGG	Wild Type
MH84-133-5	2	TACGCTA	Wild Type
MH84-134-5	2	TCAGGAA	<i>vir-k.23</i>

Supplemental Data. Stuart et al. (2021). Barley *Viridis-k* Links an Evolutionary Conserved C-Type Ferredoxin to Chlorophyll Biosynthesis. Plant Cell.

Supplemental Table S2. List of DNA oligonucleotides used in this study for cloning, polymerase chain reaction, and Sanger sequencing.

Name	Sequence	Description
HvFdC2_F1	GTGTACAAAGCCGTGAGCATTC	Forward primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_F2	TCGTCCACCAGTTCGTCGT	Forward primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_F3	GTAGGAAGAGAACGGATGTATGTG	Forward primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_F4	CACAAGGTGACCGTCCACGA	Forward primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_F5	CAAGCTTTGTGCGTCACAGT	Forward primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_F6	GCGAGGGCCTGTTGAAGTA	Forward primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_F7	TCATCCATTTTGCTTCGTATGTAAC	Forward primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_F8	AGGAAGAGAACGGATGTATGTGA	Forward primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_F9	TATCACGTGAATCGGGGTGC	Forward primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_F10	CCTACAGGCTTGCTTCGTGA	Forward primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_R1	CAATGCGAGAGCGGAACGTA	Reverse primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_R2	ATAGAGCACCCCGATTACCG	Reverse primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_R3	ACTTCTTCCATAAGCTTGGTC	Reverse primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_R4	GGGTTACATACGAAGCAAATGGAT	Reverse primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_R5	AATGAAAGACGGGCCGATT	Reverse primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_R6	GACGTGTAAAACGCAGACCG	Reverse primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_R7	GGTCCAACAGTCGAACCGAT	Reverse primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_R8	CTGAACGGGATCATCTGACAGT	Reverse primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_R9	ATTGCAGTACCGCACACAAAC	Reverse primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene
HvFdC2_R10	GGAAGGAGCATGAAGATCGGA	Reverse primer for PCR and Sanger sequencing of the <i>Viridis-k</i> gene

Supplemental Data. Stuart et al. (2021). Barley *Viridis-k* Links an Evolutionary Conserved C-Type Ferredoxin to Chlorophyll Biosynthesis. Plant Cell.

RT-HvFdC2_F1	AGTCCAAGGCCAGGATGG	Forward primer used for RT-PCR of <i>Viridis-k</i> 5' region
RT-HvFdC2_R1	CGACGAACTGGTGGACGA	Reverse primer used for RT-PCR of <i>Viridis-k</i> 5' region
RT-HvFdC2_F2	GCTATGCGTTGTTATGTGTTGG	Forward primer used for RT-PCR of <i>Viridis-k</i> 3' region
RT-HvFdC2_R2	CTCGTCTCCCATTGCGAG	Reverse primer used for RT-PCR of <i>Viridis-k</i> 3' region
HvSNP_8693162_F	GCACCTATTGTCAAAGACCA	Forward primer for CAPS SNP at 8.6 Mbp
HvSNP_8693162_R	CCGTCCTGTTCCATCGATCA	Reverse primer for CAPS SNP at 8.6 Mbp
HvSNP_9708039_F	TTCACCTTTGTCTACACCAC	Forward primer for CAPS SNP at 9.7 Mbp
HvSNP_9708039_R	AGGACGAAGACAAACTATGC	Reverse primer for CAPS SNP at 9.7 Mbp
HvSNP_10033984_F	CTCGTCTGAAACCTAATCC	Forward primer for CAPS SNP at 10 Mbp
HvSNP_10033984_R	GCAGGTGTAAGATAAGGCGAT	Reverse primer for CAPS SNP at 10 Mbp
HvSNP_13709787_F	GCGTCGGAACATTTGGTTGC	Forward primer for CAPS SNP at 13.7 Mbp
HvSNP_13709787_R	CTCATGGGGTGTGTTGAGTGC	Reverse primer for CAPS SNP at 13.7 Mbp
HvSNP_16087547_F	CAGTACATGCGAGCCGATTTG	Forward primer for CAPS SNP at 16.1 Mbp
HvSNP_16087547_R	CAACTGTTCTCTCGCAAATGG	Reverse primer for CAPS SNP at 16.1 Mbp
HvFdC2_GW_For_+cTP	GGGGACAAGTTTGTACAAAAAAGCAGG CTTCGAAGGAGATAGAACCATGGCGGC GTGCCCCGCCGCGAC	Forward primer for Gateway cloning of <i>Viridis-k</i> with cTP from cDNA
HvFdC2_GW_Rev	GGGGACCACTTTGTACAAGAAAGCTGG GTCTCCACCTCCGGATCACTCGTCTCCCA TTGCGAGCTCTAGC	Forward primer for Gateway cloning of <i>Viridis-k</i> with cTP from cDNA
HvLFd1_GW_For_+cTP	GGGGACAAGTTTGTACAAAAAAGCAGG CTTCGAAGGAGATAGAACCATGGCGGCC GCACTGAGC	Forward primer for Gateway cloning of <i>HvLFd1</i> with cTP from cDNA
HvLFd1_GW_Rev	GGGGACCACTTTGTACAAGAAAGCTGG GTCTTATGCGGGTGAAGTCTCTCTCTCT	Forward primer for Gateway cloning of <i>HvLFd1</i> with cTP from cDNA
HvRFNR_GW_For_no_cTP	GGGGACAAGTTTGTACAAAAAAGCAGG CTTCAGTCGAGCAAGAGCAAGGT	Forward primer for Gateway cloning of <i>HvRFNR</i> without cTP from cDNA
HvRFNR_GW_Rev	GGGGACCACTTTGTACAAGAAAGCTGG GTCCTAGTAAACCTCAACGTGCCATTG	Reverse primer for Gateway cloning of <i>HvRFNR</i> without cTP from cDNA
HvFdC2_qPCR_F1	CCCAGGACCAGTACATTCTG	Forward primer used for RT-qPCR of <i>Viridis-k</i> 5' region
HvFdC2_qPCR_R1	AGTCAGCCGATATCCAAGG	Reverse primer used for RT-qPCR of <i>Viridis-k</i> 5' region
UPL6_qPCR_F	AACGATCTGCCTTCTTGGA	Forward primer used for RT-qPCR of housekeeping gene
UPL6_qPCR_R	AAATCTGAGACGTCGCCCTT	Reverse primer used for RT-qPCR of housekeeping gene

Supplemental Data. Stuart et al. (2021). Barley *Viridis-k* Links an Evolutionary Conserved C-Type Ferredoxin to Chlorophyll Biosynthesis. Plant Cell.

Supplemental Table S3. Chromosome 4H position and genetic distance to the *Viridis-k* gene of cleaved amplified polymorphic sequence markers used for narrowing the mapping interval.

Position (Mbp)	Distance to <i>Viridis-k</i> (cM)	Number F₂ individuals genotyped
8.7	6.7	285
9.7	0	285
10	0	190
13.7	0	285
16.1	2.1	281

Supplemental Table S4. Ct values from RT-qPCR with forward primer HvFdC2_qPCR_F1 and reverse primer HvFdC2_pPCR_R1, which targets the 5' region of the *Viridis-k* gene. Three technical replicates on three biological replicates (three different plants) were performed for the five mutants and their mother cultivar Bonus. Ave, average; SD, standard deviation. This Supplemental Table supports Figure 8B.

	Technical replicates			Ave	SD
	1	2	3		
Bonus_1	21.76	21.45	21.67	21.63	0.1623
Bonus_2	21.37	21.47	21.41	21.42	0.0511
Bonus_3	24.22	22.23	23.94	23.46	1.0755
xan-l.35_1	22.15	21.75	22.79	22.23	0.5231
xan-l.35_2	21.83	22.00	21.72	21.85	0.1424
xan-l.35_3	21.43	21.57	21.76	21.59	0.1678
xan-l.81_1	23.06	21.25	26.50	23.60	2.6689
xan-l.81_2	22.54	21.46	21.20	21.73	0.7117
xan-l.81_3	23.16	23.40	23.09	23.22	0.1598
xan-l.82_1	21.61	21.09		21.35	0.3659
xan-l.82_2	24.00	22.53	21.58	22.70	1.2216
xan-l.82_3	22.68	21.52	21.12	21.78	0.8087
vir-k.23_1	28.01	29.41	27.67	28.36	0.9189
vir-k.23_2	29.38	30.64	27.97	29.33	1.3355
vir-k.23_3	28.21	28.57	27.67	28.15	0.4533
vir-k.170_1	24.18	22.44	21.77	22.80	1.2406
vir-k.170_2	24.12	21.65	21.47	22.42	1.4817
vir-k.170_3	22.41	22.41	22.57	22.46	0.0904

Supplemental Table S5. Ct values from RT-qPCR with forward primer UPL6_qPCR_F and reverse primer UPL6_qPCR_R, which targets the housekeeping gene E3 ubiquitin-protein ligase UPL6 (HORVU1Hr1G023480). Three technical replicates on three biological replicates (three different plants) were performed for the five mutants and their mother cultivar Bonus. Ave, average; SD, standard deviation. This Supplemental Table supports Figure 8B.

	Technical replicates			Ave	SD
	1	2	3		
Bonus_1	25.40	25.88	26.05	25.78	0.3386
Bonus_2	25.25	25.52	25.41	25.40	0.1346
Bonus_3	25.21	25.78	25.80	25.60	0.3367
xan-l.35_1	26.06	25.72	26.05	25.94	0.1902
xan-l.35_2	26.22	26.44	26.69	26.45	0.2358
xan-l.35_3	25.80	25.97	26.18	25.98	0.1888
xan-l.81_1	25.81	26.30	26.66	26.26	0.4259
xan-l.81_2	25.49	25.34	27.59	26.14	1.2612
xan-l.81_3	25.61	25.60	25.26	25.49	0.2011
xan-l.82_1	24.69	24.77	25.31	24.92	0.3382
xan-l.82_2	25.05	25.37	26.96	25.80	1.0224
xan-l.82_3	25.38	25.46	26.15	25.66	0.4218
vir-k.23_1	25.59	29.30	25.85	26.92	2.0736
vir-k.23_2	25.99	26.08	26.30	26.12	0.1596
vir-k.23_3	26.02	25.92	26.03	25.99	0.0616
vir-k.170_1	25.39	25.74	26.18	25.77	0.3975
vir-k.170_2		25.38	25.53	25.46	0.1008
vir-k.170_3	25.45	26.08	25.63	25.72	0.3259

Supplemental Table S6. Calculations of relative expression of *Viridis-k* in the five mutants and their mother cultivar Bonus using values from Supplemental Tables S4 and S5. Significant differences were tested by a two-sided t-test. SD, standard deviation. This Supplemental Table supports Figure 8B.

	F1xR1	House-keeping	Subtraction	2 ^{ΔΔCt}	1/x	Average	Relative Bonus	Relative Bonus	SD	p (t-test)
Bonus_1	21.63	25.78	-4.15	0.0563	17.78	12.64	1.00	1.4061	0.4665	
Bonus_2	21.42	25.40	-3.98	0.0634	15.77		1.247			
Bonus_3	23.46	25.60	-2.13	0.2282	4.38		0.3467			
xan-135_1	22.23	25.94	-3.71	0.0762	13.12	19.48	1.54	1.0379	0.3707	0.268
xan-135_2	21.85	26.45	-4.60	0.0412	24.27		1.9199			
xan-135_3	21.59	25.98	-4.40	0.0475	21.06		1.6659			
xan-181_1	23.60	26.26	-2.66	0.1586	6.31	10.80	0.85	0.4989	0.5871	0.797
xan-181_2	21.73	26.14	-4.41	0.0470	21.26		1.6816			
xan-181_3	23.22	25.49	-2.27	0.2072	4.83		0.3817			
xan-182_1	21.35	24.92	-3.57	0.0841	11.89	11.73	0.93	0.9406	0.2023	0.851
xan-182_2	22.70	25.80	-3.09	0.1173	8.52		0.6742			
xan-182_3	21.78	25.66	-3.89	0.0677	14.78		1.1692			
vir-k.23_1	28.36	26.92	1.45	2.7317	0.37	0.23	0.02	0.0290	0.0083	0.041
vir-k.23_2	29.33	26.12	3.20	9.2171	0.11		0.0086			
vir-k.23_3	28.15	25.99	2.16	4.4720	0.22		0.0177			
vir-k.170_1	22.80	25.77	-2.97	0.1274	7.85	8.54	0.68	0.6211	0.0575	0.384
vir-k.170_2	22.42	25.46	-3.04	0.1216	8.22		0.6503			
vir-k.170_3	22.46	25.72	-3.25	0.1048	9.55		0.7551			

Supplemental Data Set S1. Position of genes in the mapped interval for *Viridis-k* on chromosome 4H.

Genes with putative plastid localization based on presence of transit peptide sequences are marked in green.

The gene orthologous to Arabidopsis *FdC2* is marked in red.

Gene Name (2016 set)	Gene Start	Gene End	International Barley Genome Sequencing Consortium functional annotation
HORVU4HR1G003890.3	8719682	8723079	Ubiquitin-conjugating enzyme family protein
HORVU4HR1G003910.1	8746426	8749248	beta-xylosidase 2
HORVU4HR1G003930.1	8808440	8815115	Serine/threonine protein phosphatase 2A 57 kDa regulatory subunit B' alpha isoform
HORVU4HR1G003940.4	8815418	8820319	Transducin/WVD40 repeat-like superfamily protein
HORVU4HR1G003970.1	8851894	8852619	undescribed protein
HORVU4HR1G003990.3	8853810	8858789	RNA-binding protein 1
HORVU4HR1G004000.2	8861646	8863971	Major facilitator superfamily protein
HORVU4HR1G004010.8	8887725	8892842	Argonaute family protein
HORVU4HR1G004020.1	8960751	8961628	Ribonucleoside-diphosphate reductase
HORVU4HR1G004030.3	8972755	8980445	Argonaute family protein
HORVU4HR1G004040.2	9060996	9063638	Protein NRT1/ PTR FAMILY 8.3
HORVU4HR1G004050.1	9112207	9113541	HXXXD-type acyl-transferase family protein
HORVU4HR1G004070.1	9415930	9416931	RING/U-box superfamily protein
HORVU4HR1G004080.9	9417088	9421202	GATA transcription factor 18
HORVU4HR1G004090.6	9447869	9456317	ubiquitin-protein ligase 7
HORVU4HR1G004100.1	9456469	9460025	GDSL esterase/lipase
HORVU4HR1G004120.2	9468261	9474511	Magnesium transporter MRS2-B
HORVU4HR1G004130.3	9475096	9478254	RecQ-mediated genome instability protein 2
HORVU4HR1G004150.6	9479200	9486348	ARM repeat superfamily protein, putative isoform 5
HORVU4HR1G004160.6	9577553	9582803	Exostosin-2
HORVU4HR1G004170.2	9577605	9578487	Integral membrane HPP family protein
HORVU4HR1G004210.8	9657276	9660218	ferredoxin 3
HORVU4HR1G004230.14	9698628	9709684	Calcineurin-binding protein cabin-1
HORVU4HR1G004240.1	9727736	9731305	Cox19-like CHCH family protein
HORVU4HR1G004250.1	9727938	9728252	undescribed protein

HORVU4Hr1G004260.3	9733374	9737043	GMP synthase [glutamine-hydrolyzing]
HORVU4Hr1G004270.2	9758310	9760219	23S rRNA methyltransferase
HORVU4Hr1G004280.3	9773997	9779733	Tesmin/TSO1-like CXC domain-containing protein
HORVU4Hr1G004290.1	9796352	9797661	Pollen allergen Phl p 5a
HORVU4Hr1G004300.5	9801530	9801820	unknown function
HORVU4Hr1G004350.2	9823320	9823703	unknown function
HORVU4Hr1G004380.1	9832772	9834210	Pollen allergen Phl p 5a
HORVU4Hr1G004390.2	9848017	9848586	E3 ubiquitin-protein ligase
HORVU4Hr1G004410.4	9960255	9970108	glycine-rich protein
HORVU4Hr1G004440.7	9994775	9998711	Hexosyltransferase
HORVU4Hr1G004460.2	10014452	10016293	Protein NRT1/ PTR FAMILY 2.9
HORVU4Hr1G004470.10	10019781	10025221	5'-AMP-activated protein kinase beta-2 subunit protein
HORVU4Hr1G004520.1	10041369	10043986	unknown function
HORVU4Hr1G004540.4	10104425	10105938	calcium-dependent protein kinase 30
HORVU4Hr1G004550.1	10110833	10111517	glycine-rich protein
HORVU4Hr1G004590.4	10160470	10194672	Inositol hexakisphosphate and diphosphoinositol-pentakisphosphate kinase
HORVU4Hr1G004600.2	10165586	10166038	unknown function
HORVU4Hr1G004610.1	10166681	10166972	Myosin-J heavy chain
HORVU4Hr1G004690.1	10490037	10490687	LRR receptor-like serine/threonine-protein kinase EFR
HORVU4Hr1G004750.4	10523229	10527268	TBC1 domain family member 4
HORVU4Hr1G004780.2	10531204	10533758	Costars family protein ABRACL
HORVU4Hr1G004800.6	10547960	10551655	Tetratricopeptide repeat (TPR)-like superfamily protein
HORVU4Hr1G004810.2	10556111	10557749	Chalcone synthase 3
HORVU4Hr1G004820.8	10558662	10564781	H(+)-ATPase 11
HORVU4Hr1G004860.1	10594361	10595011	LRR receptor-like serine/threonine-protein kinase EFR
HORVU4Hr1G004880.6	10682463	10684236	RNA-binding region RNP-1
HORVU4Hr1G004890.1	10704210	10708712	Protein kinase superfamily protein
HORVU4Hr1G004900.1	10737907	10738421	NADH dehydrogenase subunit 5
HORVU4Hr1G004990.1	10788262	10789693	Thioesterase superfamily protein
HORVU4Hr1G005080.5	10900548	10905209	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
HORVU4Hr1G005090.2	11044852	11045925	Sec14p-like phosphatidylinositol transfer family protein

HORVU4HR1G005100.1	11218969	11302445	Tudor domain-containing protein 1
HORVU4HR1G005110.1	11328700	11334523	HIPL1 protein
HORVU4HR1G005140.2	11335301	11339009	Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase 7
HORVU4HR1G005150.1	11347318	11352922	unknown function
HORVU4HR1G005160.1	11445843	11446037	undescribed protein
HORVU4HR1G005180.2	11448617	11450175	undescribed protein
HORVU4HR1G005250.7	11539868	11544537	Domain of Uncharacterized protein function isoform 4
HORVU4HR1G005270.1	11571123	11573546	Beta-1,3-galactosyltransferase 15
HORVU4HR1G005290.1	11610715	11611822	Metallo-hydrolase/oxidoreductase superfamily protein
HORVU4HR1G005300.1	11626617	11631324	Galectin
HORVU4HR1G005310.1	11675072	11678750	Protein of unknown function, DUF538
HORVU4HR1G005320.1	11769492	11770713	Protein of unknown function, DUF538
HORVU4HR1G005350.1	11920685	11921261	undescribed protein
HORVU4HR1G005360.1	11932937	11937799	Protein of unknown function (DUF1336)
HORVU4HR1G005370.1	12238256	12239873	unknown function
HORVU4HR1G005380.1	12256558	12258042	germin-like protein 4
HORVU4HR1G005410.1	12260957	12261202	undescribed protein
HORVU4HR1G005440.1	12269769	12271098	germin-like protein 4
HORVU4HR1G005450.5	12280017	12284852	Protein of unknown function, DUF538
HORVU4HR1G005460.3	12284055	12286715	unknown function
HORVU4HR1G005480.1	12436135	12437672	unknown function
HORVU4HR1G005490.1	12442244	12444928	unknown function
HORVU4HR1G005500.49	12447997	12453224	50S ribosomal protein L18
HORVU4HR1G005510.2	12452719	12456622	Ribosomal protein S21 family protein
HORVU4HR1G005530.21	12461885	12465023	nucleobase-ascorbate transporter 7
HORVU4HR1G005540.1	12470902	12472023	germin-like protein 4
HORVU4HR1G005560.1	12527272	12528459	unknown function
HORVU4HR1G005580.1	12549380	12551474	Cytochrome P450 superfamily protein
HORVU4HR1G005590.3	12596860	12599303	chloride channel D
HORVU4HR1G005600.3	12602664	12605964	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
HORVU4HR1G005630.1	12659895	12660760	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein .

HORVU4HR1G005660.6	13249299	13251258	Cytokinin riboside 5'-monophosphate phosphoribohydrolase
HORVU4HR1G005670.3	13299705	13303184	Nuclear transcription factor Y subunit A-9
HORVU4HR1G005690.1	13535155	13536585	undescribed protein
HORVU4HR1G005710.39	13553274	13557838	ankyrin repeat family protein
HORVU4HR1G005730.2	13576936	13582500	Protein kinase superfamily protein
HORVU4HR1G005750.1	13704899	13705283	undescribed protein
HORVU4HR1G005770.4	13708524	13710921	Zinc finger CCCH domain-containing protein 24
HORVU4HR1G005780.5	13712643	13716016	Alkaline ceramidase 3
HORVU4HR1G005800.1	13738596	13739607	Oleosin Zn-II
HORVU4HR1G005810.7	13745769	13749602	Denticleless protein homolog
HORVU4HR1G005740.21	13804090	13808747	Polyketide cyclase/dehydrase and lipid transport superfamily protein .
HORVU4HR1G005870.3	13819158	13823485	Protein yippee-like
HORVU4HR1G005890.1	13862966	13865116	Adenylosuccinate synthetase
HORVU4HR1G005900.1	13912937	13918655	pescadillo-related
HORVU4HR1G005910.3	13919435	13921370	Adenylosuccinate synthetase
HORVU4HR1G005920.1	14084104	14090842	lipoxigenase 1
HORVU4HR1G005940.3	14172055	14174939	N-alpha-acetyltransferase 20
HORVU4HR1G005950.1	14285544	14295704	Importin-5
HORVU4HR1G005960.11	14394122	14396290	Shugoshin-1
HORVU4HR1G005970.2	14396370	14397964	Linoleate 9S-lipoxygenase 1
HORVU4HR1G005990.3	14415938	14417911	peptidoglycan-binding LysM domain-containing protein
HORVU4HR1G006020.1	14579545	14580983	Elongation of fatty acids protein 3-like
HORVU4HR1G006070.6	14590923	14591965	COX VIIa-like protein
HORVU4HR1G006080.3	14593761	14596492	Pentatricopeptide repeat-containing protein
HORVU4HR1G006090.1	14597509	14600465	unknown function
HORVU4HR1G006120.1	14631333	14632741	splicing factor Prp18 family protein
HORVU4HR1G006140.1	14636068	14636522	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
HORVU4HR1G006160.1	14656695	14657966	Tetrachloro-P-hydroquinone reductive dehalogenase
HORVU4HR1G006100.17	14696932	14701699	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
HORVU4HR1G006200.5	14733370	14739614	Phosphatidylinositol-4-phosphate 5-kinase family protein
HORVU4HR1G006210.1	14801169	14804331	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein .

HORVU4HR1G006250.1	14811417	14812144	NAD(P)H-quinone oxidoreductase subunit H, chloroplastic
HORVU4HR1G006260.5	14812375	14817154	Ethylene receptor
HORVU4HR1G006270.1	14869282	14870395	Cathepsin B-like cysteine proteinase 5
HORVU4HR1G006280.2	14950155	14951264	Cathepsin B-like cysteine proteinase 5
HORVU4HR1G006300.1	14973852	14978888	carotenoid cleavage dioxygenase 1
HORVU4HR1G006310.1	14988754	14989871	Cathepsin B-like cysteine proteinase 5
HORVU4HR1G006320.24	14991694	14994592	protein-protein interaction regulator family protein
HORVU4HR1G006330.1	15022368	15022943	Ppin1/SDK/mema/ protein conserved region containing protein
HORVU4HR1G006350.1	15054243	15056343	Glycosyltransferase
HORVU4HR1G006370.1	15100441	15101729	UDP-Glycosyltransferase superfamily protein
HORVU4HR1G006390.1	15102251	15102456	UDP-Glycosyltransferase superfamily protein
HORVU4HR1G006430.2	15186775	15187076	undescribed protein
HORVU4HR1G006450.1	15234706	15290488	Leucine-rich repeat receptor-like protein kinase family protein
HORVU4HR1G006460.5	15236333	15238170	UDP-Glycosyltransferase superfamily protein
HORVU4HR1G006480.5	15271854	15275739	beta glucosidase 43
HORVU4HR1G006490.1	15275939	15276619	undescribed protein
HORVU4HR1G006500.2	15277573	15279046	Eukaryotic peptide chain release factor subunit 1-3
HORVU4HR1G006510.1	15282085	15288070	Nucleotide-sugar transporter family protein
HORVU4HR1G006520.1	15293947	15295721	Zinc finger protein 593 homolog
HORVU4HR1G006530.2	15304130	15308120	beta glucosidase 43
HORVU4HR1G006540.1	15314005	15315030	Leucine-rich repeat receptor-like protein kinase family protein
HORVU4HR1G006620.1	15377728	15378636	Leucoanthocyanidin reductase
HORVU4HR1G006660.1	15419487	15425589	protein kinase family protein
HORVU4HR1G006670.2	15426863	15430534	unknown function
HORVU4HR1G006710.15	15505183	15511885	putative protein kinase 1
HORVU4HR1G006720.4	15512856	15520608	Histone-lysine N-methyltransferase ASHR1
HORVU4HR1G006730.7	15518371	15523943	Plastid-lipid associated protein PAP / fibrillin family protein
HORVU4HR1G006740.1	15524765	15527575	Transmembrane protein 41B
HORVU4HR1G006750.1	15544613	15553250	Oxysterol-binding protein-related protein 1C
HORVU4HR1G006760.1	15669209	15677845	Oxysterol-binding protein-related protein 1C
HORVU4HR1G006780.1	16018539	16023269	phosphatidylinositol-4-phosphate 5-kinase 1

Paper III



A pipeline for identification of causal mutations in barley identifies Xantha-j as the chlorophyll synthase gene

David Stuart¹, Shakhira Zakhrebekova¹, Morten Egevang Jørgensen², Christoph Dockter², Mats Hansson^{1*}

1 Department of Biology, Lund University, Sölvegatan 35B, 22362 Lund, Sweden

2 Carlsberg Research Laboratory, J. C. Jacobsens Gade 4, 1799 Copenhagen V, Denmark

Corresponding author:

Mats Hansson, Department of Biology, Lund University, Sölvegatan 35B, 22362 Lund, Sweden, +46 462224980, mats.hansson@biol.lu.se

Short title:

Barley chlorophyll synthase

Keywords: ChlG, Clomazone, Dimethazone, *Hordeum vulgare*, Xantha

Abstract:

Thousands of barley (*Hordeum vulgare* L.) mutants have been isolated over the last century, and many are stored in gene banks across various countries. In the present work, we developed a pipeline to efficiently identify causal mutations in barley. The pipeline is also efficient for mutations located in centromeric regions. Through bulked-segregant analyses using whole genome sequencing of pooled F₂ seedlings, we mapped two mutations and identified a limited number of candidate genes. We applied the pipeline on F₂-mapping populations made from *xan-j.59* (unknown mutation) and *xan-l.82* (previously known). The Xantha-j (*xan-j*) gene was identified as encoding chlorophyll synthase, which catalyzes the last step in the chlorophyll biosynthetic pathway: the addition of a phytol moiety to the propionate side chain of chlorophyllide. Key amino-acid residues in the active site, including the binding sites of the isoprenoid and chlorophyllide substrates, were analyzed in an AlphaFold2-generated structural model of the barley chlorophyll synthase. Three allelic mutants, *xan-j.19*, *xan-j.59*, and *xan-j.64* were characterized. While *xan-j.19* is a one-base pair deletion and *xan-j.59* is a nonsense mutation, *xan-j.64* causes an S212F substitution in chlorophyll synthase. Our analyses of *xan-j.64* and treatment of growing barley with clomazone, an inhibitor of chloroplastic isoprenoid biosynthesis, suggest that binding of the isoprenoid substrate is a prerequisite for the stable maintenance of chlorophyll synthase in the plastid. We further suggest that chlorophyll synthase is a sensor for coordinating chlorophyll and isoprenoid biosynthesis.

Barley (*Hordeum vulgare* L.) is one of the most economically important cereals, prompting intensive investigation into induced barley mutants since the 1920s (Lundqvist, 1992). Today, barley cultivars grown worldwide contain several mutations that were induced by physical (eg. x-ray) or chemical (eg. sodium azide) treatment. Thousands of mutants that were produced are today stored in various genebanks

(Hansson et al., 2024). These reservoirs of genetic diversity serve as a valuable asset for future plant breeding. The mutants are also of scientific interest since identifying defective genes associated with the observed mutant phenotypes provides a gateway to unraveling plant molecular processes. Traditionally, the identification of genes defective in barley mutants relied on findings from other organisms

(Hansson et al., 2018). This changed with the publication of the 5.3 Gbp barley reference genome (Mascher et al., 2017), coupled with advancements in genomic DNA sequencing techniques. There is now a great potential to harness barley mutants to identify and comprehend the causal genes behind observed traits. One of the more powerful gene identification techniques, which is often used for smaller genome species, is bulked-segregant analysis (BSA) (Michelmore et al., 1991) coupled with whole genome sequencing. Numerous specialized bioinformatics pipelines have been developed for such experiments (Schneeberger et al., 2009, Abe et al., 2012). The basic principle of this technique is to pool individuals from a mapping population based on presence or absence of the mutant phenotype. Whole genome re-sequencing is then performed on the “mutant” and “wild-type” pools. Genetic markers, usually SNPs, are then compared between the two pools. If, for example, the mutation is a single recessive allele, then markers completely linked to the mutant phenotype will be fixed for the allele from the mutant parent in the mutant pool. In the wild-type pool, the same markers will have an allele ratio of 1:2 for mutant parent to the wild-type parent alleles. On the other hand, markers that are completely unlinked will have a 1:1 allele ratio in both pools, while markers that show partial linkage will have ratios in-between. The power of this approach comes from the fact that one obtains both positional information from the allele frequency distributions to map the genomic region, as well as whole genome sequencing data that can be used to directly identify candidate mutations. As whole genome sequencing can still be prohibitively expensive for large genome species this technique has largely been limited to smaller genome species. In large genome species such as barley, various complexity reduction approaches have been applied to bulked-segregant analysis coupled with next generation sequencing. For example by performing exome sequencing (Mascher et al., 2014), RNA-seq (Liu et al., 2012), or restriction site associated DNA sequencing (RAD-seq) on the pooled samples (Xu et al., 2019). The complexity reduction techniques, although reducing the sequencing costs, have their own drawbacks. For example, exome capture may miss sequencing the causal

mutation if the gene is not in the exome capture panel. Furthermore, exome capture is not readily available for all species. RNA-seq may also miss the causal mutation if the gene is not expressed in the sampled tissue and the allele frequencies may be distorted by altered expression of genes between the mutant and wild-type individuals. RAD-seq is unlikely to provide direct sequence data of the causal mutation since only small fragments of the genome are sequenced. We applied the BSA-seq approach on two barley chlorophyll mutants, *xan-j.59* and *xan-l.82*. By not sequencing wild-type pools and instead using the other mutant pool as a control, cost was effectively decreased by 50% per sample. While the Xantha-1 (*xan-l*) gene has previously been identified as encoding the magnesium protoporphyrin IX monomethyl ester cyclase (Rzeznicka et al., 2005), the *xan-j* gene had not been identified. Three allelic mutants are available at the *xan-j* locus; *xan-j.19*, *xan-j.59* and *xan-j.64*. The mutants belong to the Xantha group of mutants, which are characterized by a yellow phenotype due to presence of carotenoids but absence of chlorophyll (Henningesen et al., 1993). We identified *xan-j* as chlorophyll synthase. This enzyme is responsible for attaching the phytol moiety (geranylgeranyl pyrophosphate (GGPP) or phytyl pyrophosphate (PhyPP)) to the carbon-17 propionate side chain of chlorophyllide, which is the last step of chlorophyll biosynthesis (Emery and Akhtar, 1987). We describe the three *xan-j* mutants and the effects of the mutations. We suggest that *in vivo* enzyme stability is dependent on binding of GGPP/PhyPP to the enzyme active site and propose that chlorophyll synthase functions as a sensor for coordinating chlorophyll and isoprenoid biosynthesis.

RESULTS

Identification of causal mutations

In order to map the chromosomal location of the mutations behind the yellow Xantha phenotype, the mutants *xan-j.59* and *xan-l.82* were crossed to the cultivar Quench. Since the mutations are

lethal in homozygous form, the plants used for crosses came from heterozygous stocks which segregated into green and yellow plants. While the yellow homozygous mutants soon died, the green plants, which are a mixture of heterozygous mutants and wild type, were used in the crosses. Because the green phenotypes of the

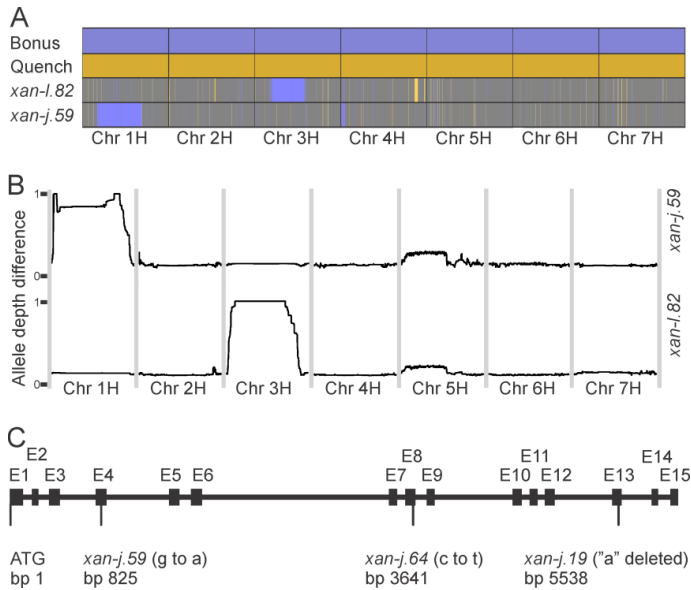


Fig 1. Gene mapping and gene structure. **A.** Mapping of *xan-j.59* and *xan-l.82* with phenotypic bulks from F₂-mapping populations using a 50k SNP chip. In the figure, the SNPs are placed in consecutive order on the respective chromosome (chr) and not according to their bp position in the physical map of the barley genome. The mutant genotypes were compared to those of Bonus and Quench. Bonus is the parental cultivar of the two mutants, which had been crossed to cultivar Quench to generate the mapping populations. Alleles from Bonus are colored blue and alleles from Quench are colored orange. Heterozygous SNPs are grey. The large blue regions on chromosome 3H and 1H indicate the relative position of the *xan-l.82* and *xan-j.59* mutation, respectively. **B.** Mapping of *xan-j.59* and *xan-l.82* using the allele depth difference for each bulk (the absolute value of the difference between read counts for two alleles divided by the sum of the read counts for the two alleles, giving values between zero and one). The *xan-j.59* mutation clearly maps to chromosome 1H with two small candidate regions with a median allele depth difference of 1. The *xan-l.82* mutation is known to be located in the centromere of chromosome 3H and this can be seen by the large centromeric region with an allele depth difference of 1 on chromosome 3H for the *xan-l.82* bulk. **C.** Gene structure of *xan-j*. The gene is 6052 bp and consists of 15 exons marked E1-E15. Base pairs are numbered from the first bp of the ATG start codon. The deduced polypeptide is 377 amino-acid residues including the transit peptide. The *xan-j.59* and *xan-j.64* mutations are single nucleotide substitutions causing an early truncation of the protein and substitution of Ser-212 to Phe, respectively. The *xan-j.19* mutation is a single base pair deletion adding a few non-native amino-acid residues followed by an early truncation.

Table 1. Filtration steps identifying candidate genes. Unique alleles for each phenotypic bulk were generated by subtracting all alleles shared between the two samples. These were further filtered to narrow down the mutation causing the respective Xantha phenotype.

Filtering level	<i>xan-j.59</i>	<i>xan-l.82</i>
Unique alleles on all chromosomes	1,426,075	795,865
Homozygous alleles on all chromosomes	12,283	18,472
Homozygous alleles on mapped chromosome	5,335	14,915
Homozygous alleles in mapped region	4,202	10,940
Homozygous alleles in mapped genes	33	36
Mapped genes with altered amino-acid sequence	13	21
Mapped high-confidence genes with altered amino-acid sequence	2	8

heterozygous mutants and the wild-type plants are indistinguishable, the progenies had to be sorted to get an F₂-mapping population based on heterozygous F₁ plants. To do this, the F₁ plants were grown to full maturity. One spike with F₂ seeds from each F₁ plant was planted. The appearance of yellow and green seedlings, germinating from a spike, demonstrated the heterozygous genotype of the F₁ plant. The remaining F₂ seeds of that plant were kept and pooled with F₂ seeds from other heterozygous F₁ plants to form the F₂-mapping population. For mapping, approximately 1,000 F₂ seeds were planted and grown at room temperature for one week. The plants were placed on the floor under the lab bench to avoid direct sunlight. For each mapping population 200-300 homozygous mutant F₂ seedlings were harvested. The leaf material was pooled, and DNA was extracted to create one mutant bulk sample for each mapping population. These bulks were first genotyped with the barley 50k SNP chip (Bayer et al., 2017) to ensure that the two causal mutations were on different chromosomes so that we could use the two bulks as control samples for each other when performing bulk segregant analysis by whole genome sequencing. This showed a linked region on chromosome 1H for the *xan-j.59* F₂ bulk and a linked region on chromosome 3H for the *xan-l.82* bulk (Fig. 1A). The *xan-l.82* mutation has been identified previously and is located on chromosome 3H (Rzeznicka et al., 2005).

For whole genome sequencing we aimed for 25-30x coverage using 150 bp paired end

sequencing. Sequencing yielded 444 and 633 million paired-end reads for the *xan-j.59* and *xan-l.82* bulks, respectively (Supplemental Table 1). To minimize noise due to incorrectly mapped reads, only reads with the maximum map quality were used for further analysis which corresponded to approximately half of the raw data (Supplemental Table 1). This resulted in a sequencing depth of 17x and 24x for the *xan-j.59* and the *xan-l.82* bulks, respectively (Supplemental Table 2). After variant calling, SNPs within a depth of approximately one standard deviation from the average were retained for further analysis (Supplemental Fig. 1). This was done because regions of low coverage will result in noisy allele frequency calculations and regions of high depth are likely to be artifacts due to mapping of similar repetitive sequences to the same position on the reference genome. To filter SNPs and map the location of the causal mutation, we used the allele depth difference calculated as the absolute value of the difference between read counts for two alleles divided by the sum of the read counts for the two alleles. This yields a value between zero and one. In each bulk a value of one indicates that a SNP is completely linked to the phenotype in the analyzed dataset. Since the two mutants, *xan-j.59* and *xan-l.82* were both induced in the cultivar Bonus, most SNPs should be due to genetic differences between Bonus and Quench which is what allows the two bulks to serve as control samples for each other. To map one of the bulks, the SNPs were first filtered by removing loci with an allele depth difference greater than 0.3 in the

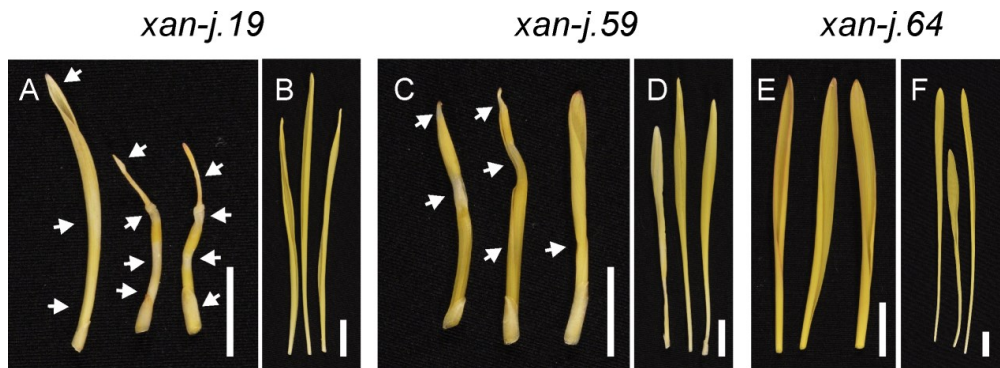


Fig 2. Seedling leaves of *xan-j* mutants grown in high light and low light. **A, C, E.** Plants grown in greenhouse (approximately 10,000 Lux). **B, D, F.** Plants grown under a lab bench (approximately 100 Lux). The seedling leaves of mutants *xan-j.19* and *xan-j.59* grown in high light show necrotic sections (indicated by arrows) corresponding to plant tissues developed during the night. No necrotic sections were observed if the plants were grown under low-light conditions. Mutant *xan-j.64* did not display any necrotic sections under high-light conditions. White bar, 1 cm.

control bulk as this was an indication that the SNP was not following Mendelian segregation and may thus be an artifact. Next, the allele depth differences were calculated for the bulk to be mapped. To denoise the data, a running median was calculated from the nearest 2,500 SNPs and plotted along each of the seven chromosomes (Fig. 1B). This showed two small regions on chromosome 1H that were completely linked to the *xan-j.59* phenotype and one large region consisting of the recombination coldspot around the centromere on chromosome 3H that was completely linked to the *xan-l.82* phenotype.

The next step was to identify candidate genes in the linked genomic regions. BCFtools was used for variant calling to identify SNPs and small indels. If a causal mutation is not identified by traditional base callers or is expected to be a large genomic structural rearrangement one could use specialized variant callers such as DELLY or BreakDancer to identify these even with short-read paired-end sequencing data (Chen et al., 2009; Rausch et al., 2012). Of the 1,426,075 alleles unique to the *xan-j.59* bulk, only 12,283 were homozygous (Table 1). The corresponding numbers for *xan-l.82* were 795,865 and 18,472. It was found that 4,202 and 10,940 of the homozygous alleles were in the mapped region of

xan-j.59 and *xan-l.82*, respectively (Table 1). Further analyses showed that only two and eight of these alleles were located in high-confidence genes and caused changes of amino-acid residues in the corresponding proteins (Table 1). The previously characterized magnesium protoporphyrin IX monomethyl ester cyclase gene was one of the eight candidate genes for *xan-j.59*. Of the two candidates for *xan-j.59*, one was annotated as a SKI family transcriptional corepressor 1 and the other as chlorophyll synthase (HORVU.MOREX.r2.1HG0049060, Supplemental Table 3). As chlorophyll synthase is required for biosynthesis of chlorophyll and the identified mutation resulted in an early stop codon this was a prime candidate.

Characterization of *xan-j.19*, *xan-j.59* and *xan-j.64*

To validate that the correct gene had been identified, the chlorophyll synthase gene was PCR amplified and Sanger sequenced from *xan-j.59* as well as the other two allelic mutants *xan-j.19* and *xan-j.64*. The mutation identified from whole genome sequencing of the *xan-j.59* bulk was verified and mutations were identified in *xan-j.19* and *xan-j.64*. Based on recommendations of the International Committee for Nomenclature and Symbolization of Barley

Genes, the chlorophyll synthase gene is therefore named Xantha-j with the symbol *xan-j* (Franckowiak and Lundqvist, 2019).

In *xan-j.19*, a one base pair deletion was identified in exon 13 and in *xan-j.64* a single base pair substitution was found in exon 8 (Fig. 1C). The early stop codon in *xan-j.59* results in a peptide of 53 residues once the chloroplast transit peptide is removed while the mutation in *xan-j.64* results in an S212F substitution (Supplemental Fig. 2). The deletion in *xan-j.19* results in a frameshift which causes the C-terminus of the protein to have 4 out of 5 incorrect residues after L340 followed by a premature stop codon (Supplemental Fig. 2).

Although all three mutants have defects in chlorophyll synthase, the *xan-j.19* and *xan-j.59* mutants showed light sensitivity while *xan-j.64* did not. When grown under low light (approximately 100 Lux), all three mutants were similar and produce yellow seedling leaves due to the lack of chlorophyll. However, when grown under high light (approximately 10,000 Lux), *xan-j.19* and *xan-j.59* tend to form bleached sections similar to Tigrina mutants, which show a striped phenotype due to accumulation of chlorophyll biosynthetic intermediates (Nielsen, 1974; Hansson et al., 1997). The *xan-j.64* mutant showed no signs of damage when grown under high light (Fig. 2). The difference in light sensitivity among the *xan-j* mutants suggests that regulation of the chlorophyll biosynthesis pathway is affected differently depending on the nature of the mutation in *xan-j*. Since the terminal step of a biosynthetic pathway often produces feedback signals to key regulatory steps in the pathway, we analyzed protein and mRNA levels for chlorophyll synthase as well as the large subunit of the magnesium chelatase encoded by Xantha-f (*xan-f*) in barley (Olsson et al., 2004). The magnesium chelatase reaction is the first unique step of the chlorophyll biosynthetic pathway and has been identified as the main regulator for flux of protoporphyrin IX towards chlorophyll biosynthesis (Tanaka and Tanaka, 2007). Interestingly, the amount of the large magnesium chelatase subunit was higher in all three *xan-j* mutants compared to the mother cultivars Bonus and Foma. The mRNA levels for

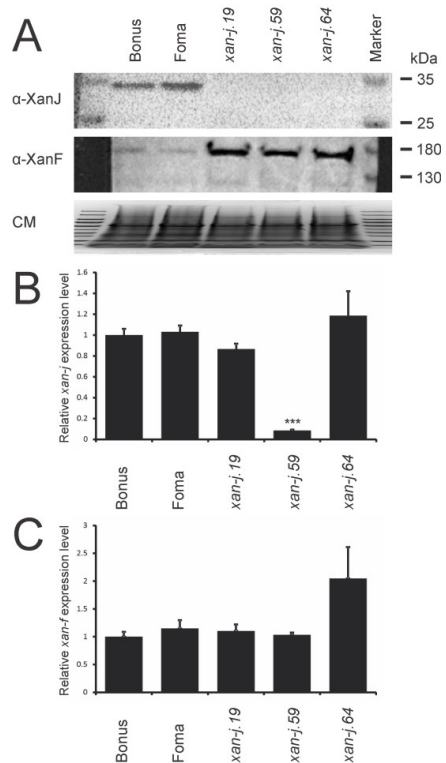


Fig 3. Presence of translation and transcription products in *xan-j* mutants. **A.** Immunoblot of *xan-j* mutants and the parental cultivars Foma and Bonus. All three mutants fail to accumulate XanJ. When probed with antibodies against the large subunit of the magnesium chelatase (XanF) the three *xan-j* mutants accumulate more XanF than the parental cultivars. As a loading control a Coomassie stained replica gel was run (CM). The Coomassie stained image has been compressed. **B.** Relative expression of *xan-j*. **C.** Relative expression of *xan-f*. **B.** and **C.** Expression levels are shown relative to Bonus and were determined by RT-qPCR from four biological replicates. Error bars show the standard error of the mean. T-test $P < 0.001$ is indicated by ***.

xan-f were not significantly different in any of the mutants (Fig. 3). The XanJ protein was not detected in any of the three *xan-j* mutants while the mRNA levels were only significantly decreased in *xan-j.59* (Fig. 3), which has a nonsense mutation in the beginning of the gene. The mRNA level of *xan-j* had a slight but statistically non-significant decrease in *xan-j.19*, which has a nonsense mutation near the end of the gene due to the frameshift caused by the one base pair deletion.

Natural variation in XanJ

Alleles affecting chlorophyll content may have an adaptive advantage in some natural or agricultural environments (Rotasperti et al., 2022). We therefore explored the natural variation in XanJ using an exome capture dataset encompassing 815 barley cultivars, landraces, and wild barley lines (Chen et al., 2022b). Within this dataset, we identified 241 variants within the chlorophyll synthase gene, of which eight cause alterations of amino-acid residues. All eight changes were observed solely in wild barley lines, with half affecting residues within the chloroplast transit peptide (Supplemental Table 4). The most common of the variants, not found in the chloroplast transit peptide, was the substitution of Q343 with lysine. This allele was found in 19 out of 266 wild barley accessions. The remaining three alleles were much less common. I86T and V273I substitutions were found in one wild barley accession each, whereas an L290F substitution was found in two accessions. The three rare alleles were all found in accessions from the area around the Sea of Galilee (Supplemental Fig. 3). These alleles likely represent natural variation that was not retained during domestication since they are exclusively found in wild but not domesticated barley lines. They may have been lost due to selection if they underperformed or more likely, they were simply not retained due to the bottleneck imposed by domestication. However, further analysis is required to determine whether these variations have a fine-tuning influence on chlorophyll content and photosynthetic efficiency.

Modeling of an AlphaFold2 generated structure of chlorophyll synthase

No structure has been solved for chlorophyll synthase from any organism. Since artificial intelligence based structural prediction is now able to generate high quality models *in silico*, we constructed AlphaFold2 (Jumper et al., 2021b; Jumper et al., 2021a) models for structural characterization of XanJ as well as variants affected by the mutations *xan-j.19* and *xan-j.64* with a truncated or modified (S212F) XanJ enzyme, respectively (Supplemental Fig. 4A). Most residues had a pLDDT (predicted local distance difference test) score above 90 (Supplemental Fig. 4B), which suggested that the AlphaFold2 generated models were of good quality. Residues with a pLDDT score above 90 generally have correct backbone placement and mostly correct side chain orientations (Tunyasuvunakool et al., 2021). Chlorophyll synthase belongs to the UbiA family of prenyl transferases (Cheng and Li, 2014). Two structures have previously been solved for related prenyl transferase enzymes; AfUbiA from *Archaeoglobus fulgidus* (Huang et al., 2014) and ApUbiA from *Aeropyrum pernix* (Cheng and Li, 2014). The structures were solved in their apo form as well as with bound geranyl-pyrophosphate (GPP), dimethylallyl-pyrophosphate (DMAPP) (Huang et al., 2014) and geranyl thiolopyrophosphate (Cheng and Li, 2014). Alignment of the XanJ model to the AfUbiA structure showed that the overall fold was conserved (Supplemental Fig. 5). The XanJ structure has nine transmembrane helices connected by longer loops on the predicted stromal side and by comparatively short loops on the lumen side of the membrane (Supplemental Fig. 6).

The active site of XanJ must accommodate both substrates; chlorophyllide and GGPP/PhyPP. To identify the active site, we started by aligning XanJ to the GPP bound AfUbiA structure. This placed GPP with the geranyl moiety in an elongated cavity with the pyrophosphate in a second more globular cavity. The globular cavity in turn connects to a distinctly flattened cavity near the pyrophosphate moiety (Supplemental Fig. 7). The elongated cavity of XanJ is likely to accommodate GGPP or

PhyPP, which are twice as long as GPP, and will be referred to as the prenyl pyrophosphate (PPP) tunnel. The globular cavity housing the pyrophosphate contains seven residues shown previously to be required for catalysis or binding of Mg^{2+} ions and prenyl substrates (Fig. 4) (Cheng and Li, 2014; Huang et al., 2014). Three residues are located in the end of three different transmembrane helices or in short breaks of the helices. The four other residues are found in loops on the stroma side of XanJ (Supplemental Table 5). The seven residues have been studied by site directed mutagenesis in various UbiA related enzymes. In all cases, the modifications had severe effects on the enzymes (Supplemental Table 5). This also included N145, D149, D153 and R160 (XanJ numbering), which were studied in chlorophyll synthase from oat (*Avena sativa* L.) (Schmid et al., 2002). Since the bonds between the pyrophosphate group and the prenyl moiety are broken followed by attachment of the prenyl group to chlorophyllide, the globular

region around the phosphate should be closest to where catalysis takes place. We therefore refer to this region of the active site as the catalytic cavity.

We next performed docking of chlorophyllide *a* to the XanJ structure using CB-Dock (Liu et al., 2020) to identify the tetrapyrrole-binding pocket. The CB-Dock engine first searches the enzyme surface for pockets and then uses AutoDock Vina to dock the substrate into the identified pockets (Trott and Olson, 2010). This docked chlorophyllide into the flattened pocket, which is adjacent to the catalytic cavity. Furthermore, the docked orientation positioned the carbon-17 propionate group of chlorophyllide *a* with the carboxyl moiety in close proximity to the reactive GPP carbon placed in the structure by alignment to AfUbiA (Fig. 5).

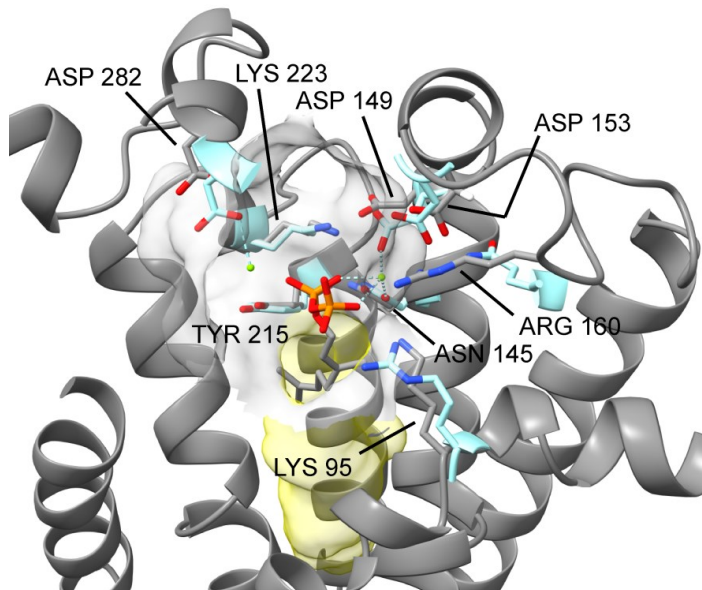


Fig. 4. AlphaFold 2 structural model of the globular cavity of XanJ. The pyrophosphate group of the prenyl substrate is located in the cavity (indicated in grey). The cavity contains several amino-acid residues of importance for catalysis, substrate binding or binding of a magnesium ion (green ball). The structural model has been aligned with the determined structure of AfUbiA containing a geranyl-pyrophosphate (GPP). The corresponding amino-acid residues of the AfUbiA structure are shown in turquoise. The prenyl pyrophosphate (PPP) tunnel is shown in yellow. Red, oxygen; orange, phosphorus; blue, nitrogen.

Previous work has shown that switching between isoleucine and phenylalanine in *R. sphaeroides* bacteriochlorophyll synthase (RsBchG) and *Synechocystis* chlorophyll synthase (SynChlG) at the position corresponding to P100 in XanJ alters the substrate specificity allowing RsBchG to utilize chlorophyllide *a* as a substrate and SynChlG to use bacteriochlorophyllide *a* (Kim et al., 2016). One would therefore expect P100 to be near a region of chlorophyllide *a* in which the structure differs from bacteriochlorophyllide *a*. Indeed, P100 is in close proximity to the carbon-8 ethyl group (Fig. 5). In chlorophyllide *a* there is a double bond between carbon-7 and carbon-8 while in bacteriochlorophyllide *a* this bond has

been reduced to a single bond. Therefore, the methylene bridge of the ethyl group is in the same plane as the macrocycle of chlorophyllide *a* but not in bacteriochlorophyllide *a*.

Support for correct substrate docking to XanJ can also be gained from heme O synthase, which is another prenyl transferase with a tetrapyrrole (heme) substrate. In heme O synthase, a histidine residue (H199) is proposed to be a ligand to Fe^{2+} of the heme substrate (Mogi, 2009; Rivett et al., 2021). The corresponding residue is I270 in XanJ. In our XanJ structural model, I270 forms part of the flattened pocket where chlorophyllide *a* was predicted to bind (Fig. 5). We designate the

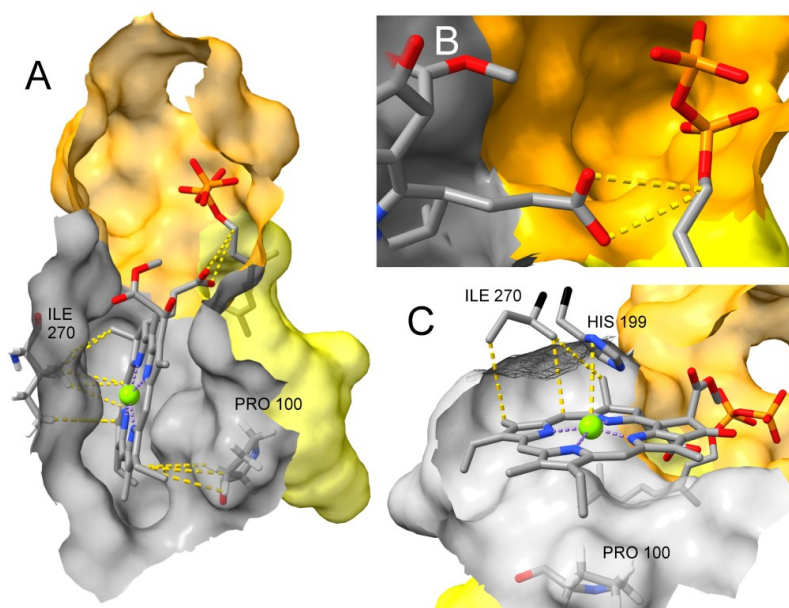


Fig. 5. Chlorophyllide *a* modelled into the active site. **A.** Chlorophyllide *a* docked into the flattened pocket adjacent to the catalytic cavity (orange) and the PPP tunnel (yellow). **B.** The carbon-17 propionate group of chlorophyllide is close to the reactive GGPP carbon placed in the structure by alignment to AfUbiA. **C.** I270 is part of the tetrapyrrole-binding pocket. I270 is predicted to be in close proximity to the chlorophyllide based on studies of heme O synthase. The shown histidine residue (H199) is from structural alignment with *Bacillus subtilis* heme O synthase (AlphaFold Protein Structure Database AF-O31652-F1-model_v4). Red, oxygen; orange, phosphorus; blue, nitrogen; green, magnesium.

flattened pocket as the tetrapyrrole binding pocket. In summary, the active site can be subdivided into the PPP tunnel, the catalytic cavity, and the tetrapyrrole-binding pocket (Fig. 6).

Evolutionary conserved amino-acid residues

To further characterize functionally important residues, we retrieved and aligned 3749 sequences of chlorophyll synthase (ChlG), bacteriochlorophyll synthase (BchG), and bacteriochlorophyll *c* synthase (BchK) from NCBI by blast. The alignment was trimmed to remove gaps relative to XanJ. We also trimmed residues before 41 since it has been shown that chlorophyll synthase from oat can be truncated in the N-terminus without loss of activity (Schmid et al., 2001). To cluster related sequences, a

maximum likelihood tree was generated. From this we designate eight groups as ChlG1, ChlG2, ChlG3, BchK1, BchK2, BchG1, BchG2, and BchG3 (Supplemental Fig. 8). Consensus sequences were generated using the most frequent amino-acid residue at each position for each of the eight groups. These were aligned and shaded based on conservation in the consensus sequence alignment (Fig. 7). For each position in this alignment, we also calculated three different conservation measurements based on the full alignment. The first of these was site-specific substitution rates based on the phylogenetic tree since evolutionary rates will be slower for structurally or functionally critical residues. The other two measures were sum of pairs using the blosum62 scoring matrix and entropy (Pei and Grishin, 2001; Pettersen et al., 2021). The

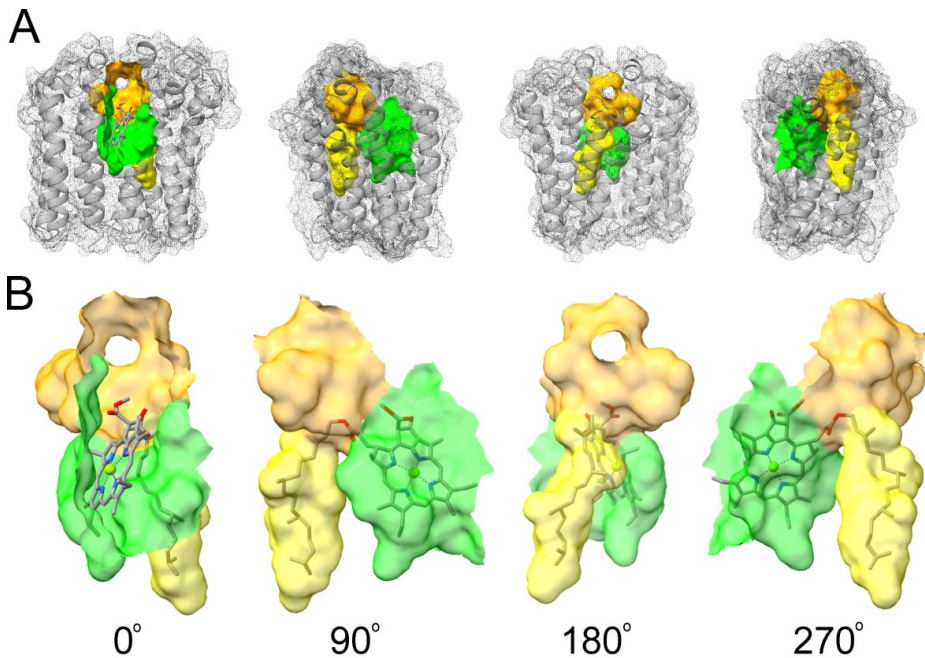


Fig. 6. Suggested active site of barley chlorophyll synthase. The active site can be subdivided in the catalytic cavity (orange), the PPP tunnel (yellow) and the tetrapyrrole-binding pocket (green). The model is shown in four different views. Red, oxygen; blue, nitrogen; green spheres, magnesium. **A.** Overall view of the active site in the enzyme. **B.** Close up of the active site with chlorophyll *a* docked to the model.

conservation values were grouped into 10 quantiles where zero represents the most conserved and 9 represents the least conserved residues. The conserved amino-acid residues were viewed in the light of the AlphaFold2-generated structural model of XanJ. As might be expected, conservation was highest at and around residues forming the active site of the enzyme with clearly higher conservation of residues of the catalytic cavity compared to the tetrapyrrole binding pocket and the PPP tunnel (indicated by orange, green and yellow, respectively in Fig. 7).

Two completely conserved residues, Q142 and Y237, and one semi-conserved residue, L233, are part of all three subcompartments (purple in Fig. 7). Inspection of the chlorophyllide *a* docked XanJ model suggests that Q142 and Y237 are hydrogen bond donors to the carbon-17 propionate oxygens which is the site of bond formation during catalysis (Supplemental Fig. 9). Of note, Y237 is located in a break of transmembrane helix 5 (TMH5) and the aromatic ring is placed in-between the tetrapyrrole-binding pocket and the PPP tunnel such that hydrophobic interactions between both substrates and Y237 are possible. The position of L233 is generally occupied by a non-polar residue and is likely to create a hydrophobic patch for interaction with ring D of chlorophyllide (Supplemental Fig. 9). The methyl group from a completely conserved residue, T98, forms part of the hydrophobic surface near the carbon-17 propionate carbon and is within contact distance of the side chains of Q142 and Y237 (Supplemental Fig. 9). Additionally, the backbone as well as the side chain hydroxyl group of T98, are H-bond donors to backbone carbonyl oxygen of K95 present in the catalytic cavity and which has been shown to be essential for prenyl-pyrophosphate binding in other UbiA family enzymes (Cheng and Li, 2014; Huang et al., 2014).

The residues which form the tetrapyrrole-binding pocket have predominantly hydrophobic side chains (green in Fig. 7). However, some of the enzyme clusters have positions with conservation of charged or polar residues, which is likely because some of the differences between the different tetrapyrrole substrates is the

presence or absence of oxygens on side groups. For example, position 238 has a completely conserved glutamate residue in BchG, which might bind the carbonyl oxygen present on carbon-2 of bacteriochlorophyllide *a* via a bridging water molecule. The residues N274 and K277 are completely conserved. K277 is located in loop 6-7 and is likely able to undergo relatively large positional changes during the catalytic cycle. Although speculative, one suggestion is that it may move far enough to act as an H-bond donor to the ring E carbonyl oxygen of the chlorophyllide, which would likely also seal off the active site during catalysis. N274 is located just at the junction of transmembrane helix 6 and loop 6-7. Slight motion could place the side-chain amide group in the catalytic cavity. The corresponding position in ApUbiA, Y178, has been implicated in decreased prenyl-pyrophosphate binding when substituted for alanine (Cheng and Li, 2014).

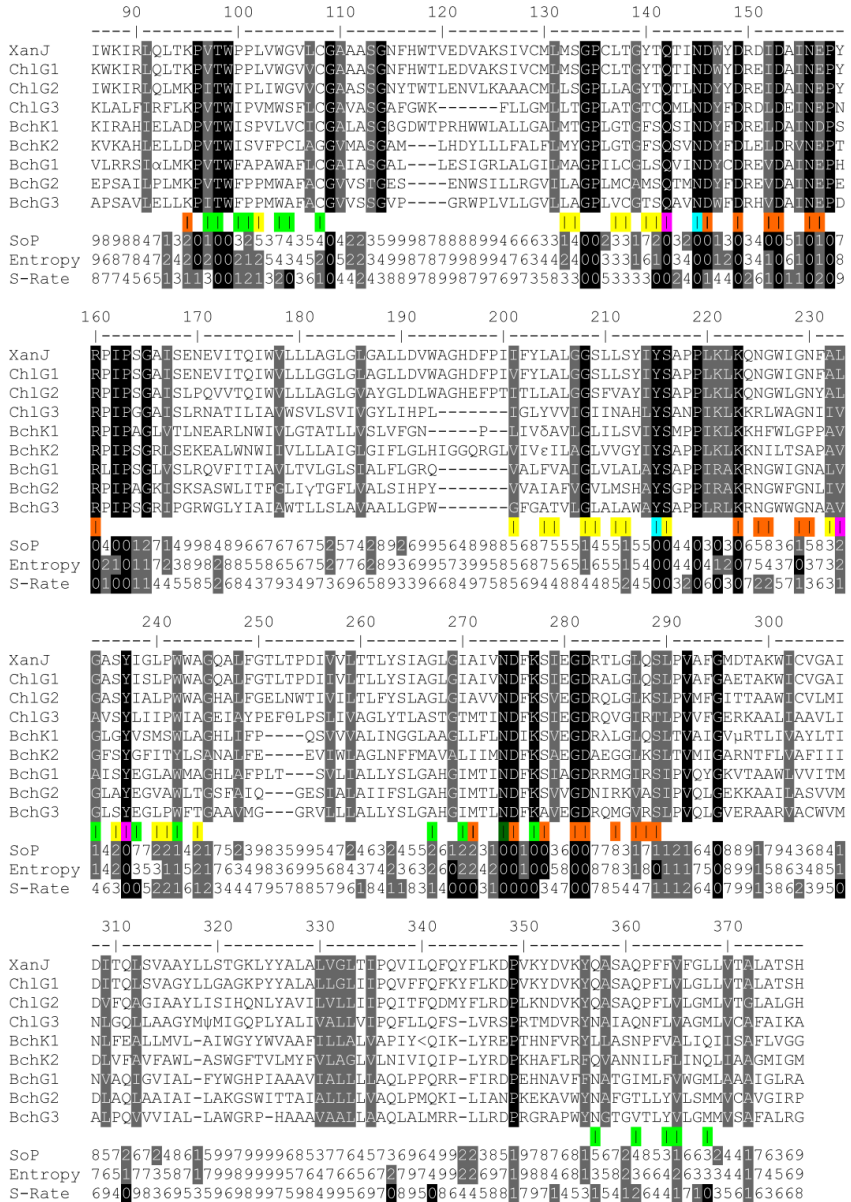
In general, the catalytic cavity has the most highly conserved residues. In addition to the residues shown in Fig. 5, positions corresponding to D146, N156, E157, D275, and S289 are highly conserved as charged or polar residues capable of hydrogen bonding. These residues are likely also essential for interactions with Mg²⁺ ions or the pyrophosphate group during the catalytic cycle.

Residues of the PPP tunnel tend to have hydrophobic side chains but otherwise most positions are fairly variable (yellow in Fig. 7). One position, S216, is completely conserved. In XanJ the hydroxyl group is likely able to act as an H-bond donor to the backbone oxygen of S212 and an H-bond acceptor from side chain of N145 (turquoise in Fig. 7). S212 is exchanged to phenylalanine by the mutation in *xan-j.64*, and N145 is a conserved residue which is part of the catalytic cavity as well as the PPP tunnel. S212 contributes to the surface of the PPP tunnel and is conserved as small amino-acid residues. The AlphaFold2 generated model of XanJ.64 clearly shows a steric clash between the bulky side chain of phenylalanine at position 212 and the isoprenoid substrate (Fig. 8).

Prenyl substrate is required for *in vivo* chlorophyll synthase accumulation

The effect of the *xan-j.64* mutation on the enzyme structure suggests that it may prevent GGPP/PhyPP binding to chlorophyll synthase,

and we could not detect the enzyme in the mutant (Fig. 3). To test if binding of the isoprenoid substrate may be required for stable maintenance of chlorophyll synthase *in vivo*, we induced GGPP/PhyPP deficiency by treating wild-type



plants with clomazone (2-[(2-chlorophenyl)methyl]-4,4-dimethyl-3-isoxazolidinone). Clomazone is a herbicide commonly used in agriculture. Clomazone has a dramatic effect on chloroplast pigmentation and results in white or pale green plants, depending on the amount applied (Duke and Kenyon, 1986; Duke and Paul, 1986). *In vivo*, clomazone is oxidized to 5-ketoclomazone (Ferhatoglu et al., 2005). 5-ketoclomazone inhibits deoxyxylulose-5-phosphate synthase, which is the first enzyme of the non-mevalonate isoprenoid synthesis pathway (Ferhatoglu and Barrett, 2006). This prevents the biosynthesis GGPP/PhyPP which in turn prevents synthesis of carotenoids and chlorophylls. A clear dose response was seen with treated seedlings being white, pale green, or green depending on clomazone concentration (Fig. 9). We noted that there were no obvious visual signs of light sensitivity in clomazone treated plants. We analyzed the treated plants for the presence of *xan-j* and *xan-f* gene products. In the white plants with the highest clomazone treatments, no chlorophyll synthase could be detected and there was an increased amount of the large magnesium chelatase subunit (Fig. 9). This is similar to the situation observed in the *xan-j* mutants. RT-qPCR analyses showed a statistically significant increase in the levels of *xan-j* mRNA in the white plants (Fig. 9).

DISCUSSION

The availability of barley genome sequence resources (Mascher et al., 2017; Monat et al., 2019) has provided a renaissance for the thousands of barley mutants available at various gene banks and seed stores around the world, since they provide excellent opportunities for functional studies of genes. By connecting a gene to a mutant phenotype, the function of the gene is revealed. In the present study, we used two chlorophyll deficient barley mutants and developed a pipeline for identifying causal mutations. We utilized only standard bioinformatic tools, which should be available on most computing clusters, and relatively simple calculations for the allele depth difference. The methodology should be easy for others to adopt and adapt.

After creating a mutant population, researchers or plant breeders are likely to have many mutants generated in the same genetic background that are of scientific or applied interest, which they would like to identify. The strategy of using two different populations of interest with the same genetic background as controls for each other effectively reduces costs by 50%, regardless of species or sequencing technology used. Further, our results can be used as a guideline for experiment design and as control data by other researchers for

Fig. 7. Protein sequence alignment of chlorophyll synthase (ChlG), bacteriochlorophyll synthase (BchG) and bacteriochlorophyll *c* synthase (BchK). The alignment is based on 3749 sequences which were clustered into eight groups (Supplemental Fig. 8) to generate a consensus sequence representing each group. Positions with alternative amino-acid residues have been indicated by Greek letters and symbols; α (A/T), β (R/Y), γ (I/L), δ (L/V), ϵ (G/S), θ (S/T), λ (K/Q), μ (K/R), ψ (F/L), ζ (N/S). Amino-acid residues forming the active site have been indicated by colors; orange (catalytic cavity), green (tetrapyrrole-binding pocket), yellow (GGPP tunnel), purple (all three subcompartments), turquoise (catalytic cavity and the GGPP tunnel), dark green (catalytic cavity and the tetrapyrrole-binding pocket). Conserved and almost conserved residues are marked in black and grey, respectively. Three different conservation values were calculated for each position; SoP (sum of pairs according the blosum62 scoring matrix); Entropy (based on amino-acid frequencies), S-Rate (site-specific substitution rate based on the phylogenetic tree). Zero represents the most conserved residues and 9 represents the least conserved residues.

identification of causal mutations in future studies where the genetic background of the mutant is the cultivar Bonus. This includes the approximately 3,000 classical mutants induced in Bonus that are curated at the Nordic Genetic Resource Center (www.nordgen.org) in Alnarp, Sweden.

Mutations located in recombination coldspot regions have traditionally been difficult to identify by map-based cloning, which relies on genetic recombination to narrow down the genomic interval until it is small enough to contain a low number of candidate genes. As a proof of principle to show that our methodology could efficiently be used for identifying mutations in recombination poor regions, the previously characterized *xan-l.82* mutant was included in this study which is known to be located in the centromeric region of chromosome 3H (Rzeznicka et al. 2005). The final list of candidates contained only eight genes, one of which was correct, i.e. the gene encoding magnesium protoporphyrin IX monomethyl ester cyclase with a TGG to TGA nonsense mutation (W291 to premature stop) (Rzeznicka et al. 2005). The chromosomal location of *xan-j* was previously unknown but found to be in the telomeric region of chromosome 1H. From analyses of the *xan-j.59* bulk, the candidate list was only two genes including the chlorophyll synthase gene. Thus, from 795,865 and 1,426,075 identified unique alleles in the *xan-l.82* and *xan-j.59* pools, respectively, it was possible to generate a very short list of candidate genes.

Chlorophyll synthase has been studied in multiple organisms such as *Synechocystis* (Oster et al., 1997), oat (Schmid et al., 2001; Schmid et al., 2002), rice (*Oryza sativa* L.) (Wu et al., 2007) and tobacco (*Nicotiana tabacum* L.) (Shalygo et al., 2009; Kim et al., 2023). Two of the studies also contained comparisons to the closely related bacteriochlorophyll synthase of *R. capsulatus* and *R. sphaeroides* (Oster et al., 1997; Kim et al., 2023). To further advance the knowledge about the enzymatic mechanism, structural information about the proteins would be most helpful. However, chlorophyll synthase is an intrinsic membrane protein, which has hampered such efforts. Recently, AlphaFold2 was published as a

machine learning approach for structure prediction with unprecedented accuracy (Jumper et al., 2021b). We used AlphaFold2 to develop a three-dimensional structure of the barley chlorophyll synthase. Like two related prenyl transferases, AfUbiA and ApUbiA, derived from two archaea, the chlorophyll synthase has nine transmembrane-spanning helices. The UbiA structures with bound prenyl pyrophosphate molecules helped us to identify the active site of the chlorophyll synthase, which consists of a tetrapyrrole-binding site and a PPP tunnel, which are connect to a catalytic cavity where the propionate group of carbon-17 of chlorophyllide is in close proximity to the reactive carbon of GGPP/PhyPP. A multisequence alignment involving 3749 sequences of chlorophyll and bacteriochlorophyll synthases, highlighted many conserved amino-acid residues. The function of most of the conserved residues could be understood from their location in the three-dimensional structure.

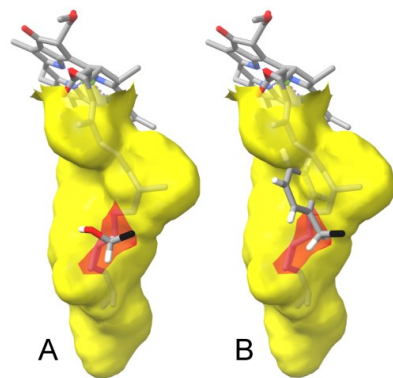


Fig. 8. The surface of the PPP tunnel (yellow) with a chlorophyll *a* molecule docked in the active site. Red, oxygen; blue, nitrogen; white, hydrogen; black, the carboxyl carbon of the amino-acid residue. **A.** S212 contributes to the surface of the PPP tunnel. **B.** In mutant *xan-j.64*, S212 has been changed to a phenylalanine. A steric clash occurs between the bulky side chain of phenylalanine and the phytol moiety.

The *xan-j.59* mutant has an early nonsense mutation that results in no detectable XanJ protein on western blots as well as decreased mRNA levels, likely due to nonsense-mediated decay (Raxwal and Riha, 2023). Similarly, the *xan-j.19* mutant also has a premature stop codon, but the stop is close to the end of the gene and the *xan-j* mRNA level was not significantly decreased. Probably the effect of nonsense-mediated decay is smaller with the mutation near the end since more of the mRNA would be protected by translating ribosomes. Since the mRNA level in *xan-j.19* is clearly close to that of wild-type, it is likely that a truncated protein is produced and then degraded since the enzyme was not detected on western blots. Although the mutation is near the end of the gene, it alters or removes 36 of the 37 terminal amino-acid residues. These changes affect two transmembrane helices, one of which is completely missing. It is likely that such a structural change prevents proper folding. Cellular machinery would likely detect this as a damaged protein and degrade it accordingly. The *xan-j.64* mutant also did not accumulate the XanJ protein while mRNA levels were unchanged. Thus, the amino-acid substitution is also likely to cause the enzyme to be degraded *in planta*.

The absence of XanJ protein according to western blot analyses in all three mutants makes it hard to understand why the *xan-j.19* and *xan-j.59* mutants are sensitive to high light intensity while *xan-j.64* is not. Possibly, a small amount of XanJ protein is still present in *xan-j.64*, although not detectable by western blot, which affects the regulation of the chlorophyll biosynthetic pathway. However, all three *xan-j* mutants showed an equally increased level of the large magnesium chelatase subunit XanF, which is at the main regulatory point for directing tetrapyrrole precursors toward chlorophyll biosynthesis. Thus, the magnesium chelatase levels alone are not likely to be the main regulatory variable explaining the difference in light sensitivity. However, the tetrapyrrole biosynthesis pathway is also regulated post-translationally (Wang et al., 2022) so the *in vivo* magnesium chelatase activity may be different between the light sensitive and insensitive *xan-j* mutants even if enzyme levels are similar. The

other main regulatory point affecting chlorophyll precursor levels is the formation of aminolevulinic acid (ALA), which is the first committed step for tetrapyrrole biosynthesis (Tanaka and Tanaka, 2007). A previous study on

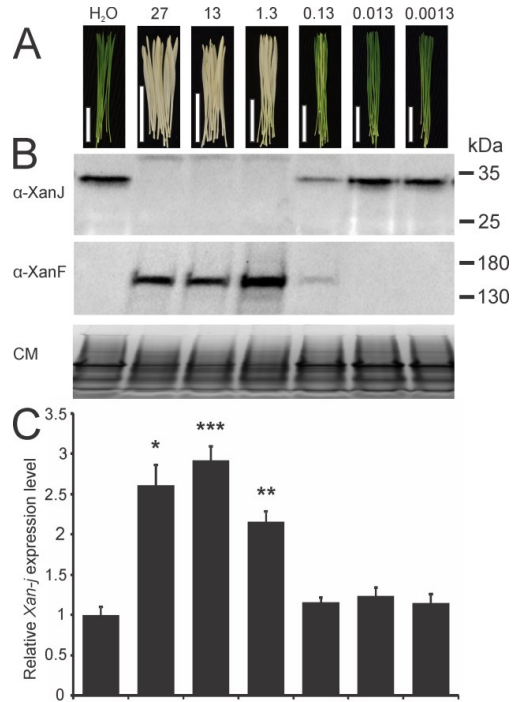


Fig. 9. Effect of clomazone treatment on growing barley seedlings. **A.** Visual effect of clomazone treatment on barley seedlings. The concentrations are in mg/l. White bar, 5 cm. **B.** Immunoblot analyses showed that XanJ was absent in seedlings where clomazone treatment had generated a white phenotype. In contrast, the large subunit of the magnesium chelatase (XanF) was more abundant in the same plants. As a loading control a Coomassie stained replica gel was run (CM). The Coomassie stained image has been compressed. **C.** RT-qPCR revealed an increased expression of the *xan-j* gene in clomazone affected seedlings. Expression levels are shown relative to water treated plants from four biological replicates. Error bars show the standard error of the mean. T-test $P < 0.001$ is indicated by ***, $P < 0.01$ **, $P < 0.1$ *.

a rice chlorophyll synthase mutant with decreased enzymatic activity showed increased levels of ALA, protoporphyrin IX, Mg-protoporphyrin IX, protochlorophyllide, as well as chlorophyllide, suggesting that regulation may be affecting the ALA synthesis step, although they noted decreased mRNA levels for glutamyl-tRNA reductase and the increased precursor levels might simply be due to the decreased conversion of chlorophyllide *a* to chlorophyll *a* (Wu et al., 2007). However, another study on tobacco chlorophyll synthase antisense lines revealed a tendency for decreased levels of intermediates as well as enzymatic activity of ALA formation and the early steps of the chlorophyll branch, suggesting that the pathway was down regulated by decreased chlorophyll synthase levels (Shalygo et al., 2009). The same study observed the opposite tendency in tobacco chlorophyll synthase overexpression lines. The inconsistencies between these two studies may be due to species differences or have to do with the mechanism behind how chlorophyll synthase was affected. In the rice study, the enzymatic activity of chlorophyll synthase was reduced *in vitro* due to the identified mutation while *in vivo* enzyme levels were not determined. In the tobacco study, chlorophyll synthase levels were reduced by RNAi but the function of the enzyme itself was not altered due to mutation. Neither of the two studies looked at a situation where there was essentially a complete block of chlorophyll synthase activity. The three barley *xan-j* mutants all result in a more severe block of the chlorophyll synthase step and do not accumulate chlorophyll synthase *in vivo*, yet they have differing light sensitivity phenotypes. We therefore looked at AlphaFold2 models of the *xan-j.64* mutant protein for clues as to how the S212F modification may affect chlorophyll synthase function. Although S212 is not strictly conserved, it is conserved as a small residue (serine, alanine or glycine), and is located in the PPP tunnel. We postulate that a bulky phenylalanine instead of S212 abolishes binding of GGPP/PhyPP to the active site of chlorophyll synthase. In order to analyze if the absence of XanJ in the *xan-j.64* mutant is connected to a requirement of GGPP/PhyPP binding, we exposed growing barley plants to clomazone, which is an inhibitor of chloroplastic isoprenoid biosynthesis

(Ferhatoglu and Barrett, 2006). This prevents synthesis of GGPP/PhyPP. We found that chlorophyll synthase was absent in clomazone treated plants. Thus, absence of GGPP/PhyPP binding in both the mutant *xan-j.64* and clomazone treatment of barley seedlings suggested that there might be a requirement of GGPP/PhyPP binding to the active site of chlorophyll synthase to stabilize the enzyme. Furthermore, *xan-j* mRNA levels were increased by clomazone treatment even though the protein was not detected. This suggests that either translation of the mRNA is repressed, or more likely that the chlorophyll synthase protein is rapidly degraded following translation. The levels of XanF were increased by clomazone treatment, which is similar to the increase in the *xan-j* mutants. Since chlorophyll synthase needs to coordinate chlorophyll synthesis with isoprenoid synthesis, the enzyme may act as a GGPP/PhyPP sensor by requiring the bound isoprenoid substrate for stability. In the AfUbiA structure it was noted that substrate binding resulted in an unstructured loop region becoming ordered (Huang et al., 2014). A similar observation was made with the ApUbiA structure, which resulted in a loop region becoming ordered and thereby protected from protease degradation upon binding the prenyl substrate (Cheng and Li, 2014). These loop regions correspond to regions of chlorophyll synthase that are likely to be exposed to the stroma and conformational changes could expose a degron to stromal proteases when chlorophyll synthase is not bound to GGPP/PhyPP. The lack of a light sensitive phenotype of *xan-j.64* further suggests that the mechanism of chlorophyll synthase degradation due to lack of the isoprenoid substrate might result in a signal to down-regulate flux through the chlorophyll synthesis pathway. This signal does not seem to be generated if chlorophyll synthase is degraded due to damage as in *xan-j.19* or completely missing as in *xan-j.59*. Thus, it is unlikely that this signal is simply the lack of chlorophyll synthase. The mammalian prenyl transferase UBIAD1, which utilizes GGPP for vitamin K synthesis, has been shown to act as a GGPP sensor to modulate flux through the cytosolic mevalonate pathway for isoprenoid biosynthesis (Nakagawa et al., 2010; Schumacher et al., 2015).

When sterols bind to 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGCR), which catalyzes the committed step of the mevalonate isoprenoid biosynthetic pathway, HMGCR is targeted for degradation by ubiquitination and binds UBIAD1. HMGCR is protected from degradation by binding UBIAD1. When UBIAD1 binds GGPP, it releases HMGCR which can then be degraded. UBIAD1 thus acts as a GGPP sensor to decrease flux through the mevalonate pathway when two different end products, sterols and GGPP are present. The interaction between HMGCR and UBIAD1 is mediated by interactions between transmembrane helices on both proteins and it seems that this interaction is only disrupted by supplying geranylgeraniol and not farnesol (Schumacher et al., 2015; Chen et al., 2022a). Induced regulatory conformational changes in membrane regions or surface exposed regions by isoprenoid substrate binding to UbiA family proteins thus seems to be a common feature (Nakagawa et al., 2010; Cheng and Li, 2014; Huang et al., 2014; Schumacher et al., 2015; Chen et al., 2022a). A GGPP/PhyPP sensing role for chlorophyll synthase is also consistent with the enzymatic mechanism of the chlorophyll synthase. The enzyme has previously been shown to follow a ping-pong mechanism with the isoprenoid substrate binding first and catalysis occurring upon chlorophyllide binding (Schmid et al., 2002). Since chlorophyll precursors are phototoxic, accumulation must be avoided. A mechanism where the chlorophyll biosynthetic pathway is downregulated when isoprenoid synthesis is lacking is thus necessary to avoid phototoxic effects. In the clomazone treated plants, chlorophyll synthase did not accumulate. However, there was an increased expression of *xan-j*. This suggests that the plants are poised to resume chlorophyll synthesis when GGPP/PhyPP becomes available. It is tempting to speculate that in this state there is a high turnover of chlorophyll synthase in a GGPP/PhyPP unbound confirmation with steady-state levels of chlorophyll synthase under the detection limit of our immunoblots. This population of chlorophyll synthase molecules may be the origin of the regulatory signal. Future studies will need to elucidate the molecular mechanism by which chlorophyll synthase acts as a GGPP/PhyPP sensor.

MATERIALS AND METHODS

Plant material

The barley (*Hordeum vulgare* L.) mutants *xan-j.19*, *xan-j.59*, *xan-j.64* and *xan-l.82* belong to the chlorophyll mutant collection previously kept at the Carlsberg Research Laboratory (Hansson, 2007). Mutants *xan-j.19*, *xan-j.59* and *xan-l.82* were derived from cultivar Bonus mutant populations in 1954, 1956 and 1975, respectively, whereas *xan-j.64* was induced in Foma in 1958. Mutant *xan-j.19* was induced by x-rays, *xan-j.59* by p-N-di(β -chloroethyl)phenylalanine, *xan-j.64* by diethyl sulfate, and *xan-l.82* by sodium azide (Henningsen et al., 1993). The mutants are kept as heterozygous stocks due to the lethal phenotype of the mutations. Mapping populations were constructed by crossing the *xan-j.59* and *xan-l.82* mutants to the cultivar Quench as the male parent.

Genotyping of F₂ mapping populations

Genomic DNA was isolated from leaves of pooled F₂ plants by a modified CTAB protocol as previously described (Stuart et al., 2021). Genomic DNA sequencing was performed by SciLifeLab (scilifelab.se) using an Illumina NovaSeq 6000 with 150 paired-end cycles. The library preparation was also performed by SciLifeLab using the TruSeq PCR-free DNA library preparation kit (Illumina Inc.). Sequencing reads were aligned to the barley MorexV2 reference genome using BWA mem (version 0.7.17) (Li and Durbin, 2010; Monat et al., 2019). Samtools fixmate (samtools version 1.10) was used to fill in mate coordinates. BCFtools (BCFtools version 1.17) was used for variant calling with flags -q 60 -Q 30 -D (Li, 2011). For mapping, only bi-allelic SNPs present in both bulks were used. Allele depth differences were calculated in UNIX using awk. The zoo package was used for calculating running ADD medians and plotted with ggplot in R version 4.3.1. To identify alleles observed in a mutant sample and not found in the control, the BCFtools contrast plugin was used. BCFtools was used to filter unique variants. Functional annotation of variants was performed with snpEff version 4.3t.

50k SNP genotyping was conducted at the James Hutton Institute, Dundee, UK, using the barley 50k iSelect SNP chip (Bayer et al., 2017).

Flapjack (Milne et al., 2010) was used for genotypic visualization of the 50k data.

General DNA techniques

Genomic DNA for PCR reactions was extracted from fresh leaves using the REDEExtract-N-Amp™ Plant PCR Kit (Sigma-Aldrich, St. Louis, MO, USA). PCR were performed by initial denaturation at 94°C/3 min, followed by 35 cycles of 94°C/30 s, 60°C/30 s and 72°C/60 s, with a final extension step of 72°C/5 min. Purification of PCR products was done by using Illustra ExoProStar 1-Step (Cytiva, Marlborough, MA, USA). Sanger sequencing was performed by Eurofins Genomics, Germany. RT-qPCR was performed from isolated total RNA as previously described (Stuart et al., 2021). Primers used in this study are listed in Supplemental Table 6. Expression levels were determined by RT-qPCR from four biological replicates as previously described (Stuart et al., 2021). Significant differences were calculated using a two sided Student's T-test.

Extraction of natural variation in XanJ

SNPs in the XanJ region (XanJ: HORVU1Hr1G059890 chr1H:435325093-435331638) were extracted from 815 exome sequenced cultivars, landraces and wild barleys (Chen et al., 2022) using BCFtools (Li, 2011). Google My Maps was used to visualize the origin of barley accession.

Modelling of a XanJ 3D structure

AlphaFold version 2.1.1 (Jumper et al., 2021b; Jumper et al., 2021a) was used to construct a 3D structural model of the *xan-j* gene product without the chloroplast transit peptide. The required databases (BFD, mgnify, pdb70, pdb_mmcif, small_bfd, uniclust30 and uniref90) for structural prediction were downloaded 2021-09-22. Structures were visualized with UCSF ChimeraX (Pettersen et al., 2021).

Chlorophyll synthase residue conservation

Sequences were aligned with Clustal Omega version 1.2.4 (Sievers and Higgins, 2018). IQ-Tree version 2.2.2.2 was used to determine the best fit model as LG+F+I+R10 which was used to infer the phylogenetic tree and site-specific substitution rates (Minh et al., 2020). Consensus

sequences of phylogenetic clusters were defined as the most frequent amino acid at each position and calculated with BioEdit version 7.2.5 and Excel (Hall, 1999). Conservation measures sum of pairs and entropy were calculated with AL2CO as implemented in ChimeraX (Pei and Grishin, 2001; Pettersen et al., 2021).

Western blot

Extraction of total proteins from barley leaves, separation of protein by SDS-polyacrylamide gel electrophoresis and immunoblots were performed as described previously (Stuart et al., 2020; Stuart et al., 2021). Polyclonal rabbit (*Oryctolagus cuniculus* L.) antibodies against barley XanJ were generated by Agrisera (Vännäs, Sweden) using synthetic peptides as antigen. The following synthetic peptides, derived from the XanJ amino-acid sequence, were used: CAKQEDNIWKIRLQL and YDRDIDAINEPYRPIPC.

Accession Numbers

Sequence data from this article can be found in the the NCBI Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/sra>) under accession number PRJNA1095161.

Supplementary Data

Supplementary Figure S1. The frequency distribution of sequencing depth for called SNPs

Supplementary Figure S2. Alignment of barley XanJ and the polypeptides resulting from the mutations in the three mutants *xan-j.19*, *xan-j.59* and *xan-j.64*

Supplementary Figure S3. Geographical distribution of barley lines with variation in the chlorophyll synthase gene *xan-j*

Supplementary Figure S4. Structural prediction of barley XanJ

Supplementary Figure S5. Alignment of the XanJ model to the UbiA structure of *Archaeoglobus fulgidus* determined by x-ray crystallography

Supplementary Figure S6. AlphaFold generated structural model of barley XanJ

Supplementary Figure S7. The GPP in the active site of XanJ after alignment of XanJ to the GPP bound AfUbiA structure

Supplementary Figure S8. Maximum likelihood phylogenetic tree

Supplementary Figure S9. Position of five key amino-acid residues in the active site of barley XanJ

Supplementary Table S1. Number of sequence reads

Supplementary Table S2. Sequence coverage and depth

Supplementary Table S3. Candidate genes containing homozygous mutations in the mapped chromosomal regions

Supplementary Table S4. Natural allelic variation in *xan-j* according to a panel of 815 barley cultivars, landraces and wild barley lines

Supplementary Table S5. Summary of amino-acid residues located in the active site of XanJ and/or previously studied in related enzymes

Supplementary Table S6. Primers used for PCR and RT-qPCR amplifications of *xan-j*

Funding information

This work was supported by the Swedish Research Council (grant no. VR 2018-05117 and VR 2022-03858 to M.H.), the Swedish Research Council for Environment, Agricultural Sciences, and Spatial Planning (grant no. FORMAS 2018-01026 to M.H.), the Erik Philip-Sörensen Foundation (M.H.), and the Royal Physiographic Society in Lund (D.S., S.Z.). Sequencing was performed by the SNP&SEQ Technology Platform in Uppsala. The facility is part of the National Genomics Infrastructure (NGI) Sweden and Science for Life Laboratory. The SNP&SEQ Platform is also supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation. The genomic data analyses were enabled by resources provided by the National Academic Infrastructure for

Supercomputing in Sweden (NAISS) and the Swedish National Infrastructure for Computing (SNIC) at Linköping University and Uppsala University, respectively, partially funded by the Swedish Research Council through grant agreements no. 2022-06725 and no. 2018-05973.

AUTHOR CONTRIBUTIONS

DS and MH designed the research. DS, SZ, MEJ, CD and MH performed research. DS, SZ, MEJ, CD and MH analyzed data. DS and MH wrote the paper. All authors read and approved the final version of the manuscript.

REFERENCES

- Bayer, M.M., Rapazote-Flores, P., Ganal, M., Hedley, P.E., Macaulay, M., Plieske, J., Ramsay, L., Russell, J., Shaw, P.D., Thomas, W., and Waugh, R.** (2017). Development and evaluation of a barley 50k iSelect SNP array. *Front Plant Sci* **8**: 1792.
- Chen, H., Qi, X., Faulkner, R.A., Schumacher, M.M., Donnelly, L.M., DeBose-Boyd, R.A., and Li, X.** (2022a). Regulated degradation of HMG CoA reductase requires conformational changes in sterol-sensing domain. *Nat Commun* **13**: 4273.
- Chen, K., Wallis, J.W., McLellan, M.D., Larson, D.E., Kalicki, J.M., Pohl, C.S., McGrath, S.D., Wendl, M.C., Zhang, Q., Locke, D.P., Shi, X., Fulton, R.S., Ley, T.J., Wilson, R.K., Ding, L., and Mardis, E.R.** (2009). BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods* **6**: 677-681.
- Chen, Y.Y., Schreiber, M., Bayer, M.M., Dawson, I.K., Hedley, P.E., Lei, L., Akhunova, A., Liu, C., Smith, K.P., Fay, J.C., Muehlbauer, G.J., Steffenson, B.J., Morrell, P.L., Waugh, R., and Russell, J.R.** (2022b). The evolutionary patterns of barley pericentromeric chromosome regions, as shaped by linkage disequilibrium and

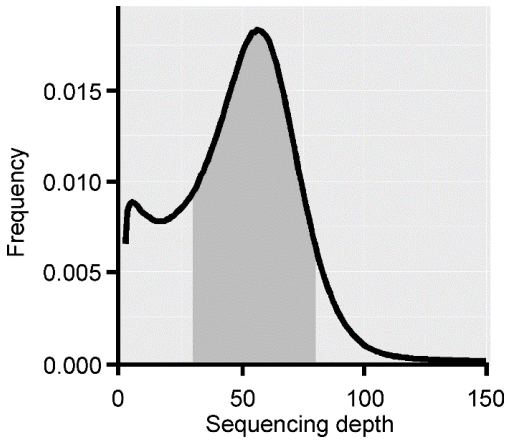
- domestication. *The Plant journal : for cell and molecular biology* **111**: 1580-1594.
- Cheng, W., and Li, W.** (2014). Structural insights into ubiquinone biosynthesis in membranes. *Science* **343**: 878-881.
- Duke, S.O., and Kenyon, W.H.** (1986). Effects of dimethazone (FMC 57020) on chloroplast development: II. Pigment synthesis and photosynthetic function in cowpea (*Vigna unguiculata* L.) primary leaves. *Pestic Biochem Physiol* **25**: 11-18.
- Duke, S.O., and Paul, R.N.** (1986). Effects of dimethazone (FMC 57020) on chloroplast development: 1. Ultrastructural effects in cowpea (*Vigna unguiculata* L.) primary leaves. *Pestic Biochem Physiol* **25**: 1-10.
- Ferhatoglu, Y., and Barrett, M.** (2006). Studies of clomazone mode of action. *Pestic Biochem Physiol* **85**: 7-14.
- Ferhatoglu, Y., Avdiushko, S., and Barrett, M.** (2005). The basis for the safening of clomazone by phorate insecticide in cotton and inhibitors of cytochrome P450s. *Pestic Biochem Physiol* **81**: 59-70.
- Franckowiak, J.D., and Lundqvist, U.** (2019). Rules for nomenclature and gene symbolization in barley. *Barley Genet Newslett* **49**: 1-5.
- Hall, T.A.** (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Nucleic acids symposium series* (Oxford), pp. 95-98.
- Hansson, M.** (2007). The Scandinavian barley chlorophyll mutation collection. *Barley Genet Newslett* **37**: 37-43.
- Hansson, M., Gough, S.P., Kannangara, C.G., and von Wettstein, D.** (1997). Analysis of RNA and enzymes of potential importance for regulation of 5-aminolevulinic acid synthesis in the protochlorophyllide accumulating barley mutant *tigrina-d¹²*. *Plant physiology and biochemistry : PPB* **35**: 827-836.
- Henningsen, K.W., Boynton, J.E., and Wettstein, D.v.** (1993). Mutants at *xantha* and *albina* loci in relation to chloroplast biogenesis in barley (*Hordeum vulgare* L.). *The Royal Danish Academy of Sciences and Letters* **42**: 1-349.
- Huang, H., Levin, E.J., Liu, S., Bai, Y., Lockless, S.W., and Zhou, M.** (2014). Structure of a membrane-embedded prenyltransferase homologous to UBIAD1. *PLoS biology* **12**: e1001911.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., and Hassabis, D.** (2021a). Applying and improving AlphaFold at CASP14. *Proteins* **89**: 1711-1721.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., and Hassabis, D.** (2021b). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**: 583-589.
- Kim, E.J., Kim, H., and Lee, J.K.** (2016). The photoheterotrophic growth of bacteriochlorophyll synthase-deficient mutant of *Rhodobacter sphaeroides* is restored by I44F mutant chlorophyll synthase of *Synechocystis* sp. PCC 6803. *J Microbiol Biotechnol* **26**: 959-966.
- Kim, J., Lee, J.K., and Kim, E.J.** (2023). Chlorophyll *a* synthesis in *Rhodobacter sphaeroides* by chlorophyll synthase of

- Nicotiana tabacum*. Biology (Basel) **12**: 573.
- Li, H.** (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**: 2987-2993.
- Li, H., and Durbin, R.** (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**: 589-595.
- Liu, Y., Grimm, M., Dai, W.T., Hou, M.C., Xiao, Z.X., and Cao, Y.** (2020). CB-Dock: a web server for cavity detection-guided protein-ligand blind docking. *Acta Pharmacol Sin* **41**: 138-144.
- Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S.O., Wicker, T., Radechuk, V., Dockter, C., Hedley, P.E., Russell, J., Bayer, M., Ramsay, L., Liu, H., Haberer, G., Zhang, X.Q., Zhang, Q., Barrero, R.A., Li, L., Taudien, S., Groth, M., Felder, M., Hastie, A., Šimková, H., Staňková, H., Vrána, J., Chan, S., Muñoz-Amatriaín, M., Ounit, R., Wanamaker, S., Bolser, D., Colmsee, C., Schmutzer, T., Aliyeva-Schnorr, L., Grasso, S., Tanskanen, J., Chailyan, A., Sampath, D., Heavens, D., Clissold, L., Cao, S., Chapman, B., Dai, F., Han, Y., Li, H., Li, X., Lin, C., McCooke, J.K., Tan, C., Wang, P., Wang, S., Yin, S., Zhou, G., Poland, J.A., Bellgard, M.I., Borisjuk, L., Houben, A., Doležel, J., Ayling, S., Lonardi, S., Kersey, P., Langridge, P., Muehlbauer, G.J., Clark, M.D., Caccamo, M., Schulman, A.H., Mayer, K.F.X., Platzer, M., Close, T.J., Scholz, U., Hansson, M., Zhang, G., Braumann, I., Spannagl, M., Li, C., Waugh, R., and Stein, N.** (2017). A chromosome conformation capture ordered sequence of the barley genome. *Nature* **544**: 427-433.
- Milne, I., Shaw, P., Stephen, G., Bayer, M., Cardle, L., Thomas, W.T., Flavell, A.J., and Marshall, D.** (2010). Flapjack-graphical genotype visualization. *Bioinformatics* **26**: 3133-3134.
- Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R.** (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* **37**: 1530-1534.
- Mogi, T.** (2009). Over-expression and characterization of *Bacillus subtilis* heme O synthase. *J Biochem* **145**: 669-675.
- Monat, C., Padmarasu, S., Lux, T., Wicker, T., Gundlach, H., Himmelbach, A., Ens, J., Li, C., Muehlbauer, G.J., Schulman, A.H., Waugh, R., Braumann, I., Pozniak, C., Scholz, U., Mayer, K.F.X., Spannagl, M., Stein, N., and Mascher, M.** (2019). TRITEX: chromosome-scale sequence assembly of Triticeae genomes with open-source tools. *Genome Biol* **20**: 284.
- Nakagawa, K., Hirota, Y., Sawada, N., Yuge, N., Watanabe, M., Uchino, Y., Okuda, N., Shimomura, Y., Suhara, Y., and Okano, T.** (2010). Identification of UBIAD1 as a novel human menaquinone-4 biosynthetic enzyme. *Nature* **468**: 117-121.
- Nielsen, O.F.** (1974). Macromolecular physiology of plastids. XII. *Tigrina* mutants in barley: genetic, spectroscopic and structural characterization. *Hereditas* **76**: 269-304.
- Olsson, U., Srijovski, N., and Hansson, M.** (2004). Characterization of eight barley *xantha-f* mutants deficient in magnesium chelatase. *Plant physiology and biochemistry* : PPB **42**: 557-564.
- Oster, U., Bauer, C.E., and Rüdiger, W.** (1997). Characterization of chlorophyll *a* and bacteriochlorophyll *a* synthases by heterologous expression in *Escherichia coli*. *J Biol Chem* **272**: 9671-9676.
- Pei, J., and Grishin, N.V.** (2001). AL2CO: calculation of positional conservation in a protein sequence alignment. *Bioinformatics* **17**: 700-712.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I.,**

- Morris, J.H., and Ferrin, T.E.** (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* **30**: 70-82.
- Rausch, T., Zichner, T., Schlattl, A., Stutz, A.M., Benes, V., and Korbel, J.O.** (2012). DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**: i333-i339.
- Raxwal, V.K., and Riha, K.** (2023). The biological functions of nonsense-mediated mRNA decay in plants: RNA quality control and beyond. *Biochem Soc Trans* **51**: 31-39.
- Rivett, E.D., Heo, L., Feig, M., and Hegg, E.L.** (2021). Biosynthesis and trafficking of heme *o* and heme *a*: new structural insights and their implications for reaction mechanisms and prenylated heme transfer. *Crit Rev Biochem Mol Biol* **56**: 640-668.
- Rotaspert, L., Tadini, L., Chiara, M., Crosatti, C., Guerra, D., Tagliani, A., Forlani, S., Ezquer, I., Horner, D.S., Jahns, P., Gajek, K., Garcia, A., Savin, R., Rossini, L., Tondelli, A., Janiak, A., and Pesaresi, P.** (2022). The barley mutant happy under the sun 1 (*hus1*): An additional contribution to pale green crops. *Environmental and Experimental Botany* **196**: 104795.
- Rzeznicka, K., Walker, C.J., Westergren, T., Kannangara, C.G., von Wettstein, D., Merchant, S., Gough, S.P., and Hansson, M.** (2005). *Xantha-1* encodes a membrane subunit of the aerobic Mg-protoporphyrin IX monomethyl ester cyclase involved in chlorophyll biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 5886-5891.
- Schmid, H.C., Oster, U., Kögel, J., Lenz, S., and Rüdiger, W.** (2001). Cloning and characterisation of chlorophyll synthase from *Avena sativa*. *Biol Chem* **382**: 903-911.
- Schmid, H.C., Rassadina, V., Oster, U., Schoch, S., and Rüdiger, W.** (2002). Pre-loading of chlorophyll synthase with tetraprenyl diphosphate is an obligatory step in chlorophyll biosynthesis. *Biol Chem* **383**: 1769-1778.
- Schumacher, M.M., Elsabrouty, R., Seemann, J., Jo, Y., and DeBose-Boyd, R.A.** (2015). The prenyltransferase UBIAD1 is the target of geranylgeraniol in degradation of HMG CoA reductase. *eLife* **4**: e05560.
- Shalygo, N., Czarnecki, O., Peter, E., and Grimm, B.** (2009). Expression of chlorophyll synthase is also involved in feedback-control of chlorophyll biosynthesis. *Plant Mol Biol* **71**: 425-436.
- Sievers, F., and Higgins, D.G.** (2018). Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci* **27**: 135-145.
- Stuart, D., Sandström, M., Youssef, H.M., Zakhrebakova, S., Jensen, P.E., Bollivar, D.W., and Hansson, M.** (2020). Aerobic barley Mg-protoporphyrin IX monomethyl ester cyclase is powered by electrons from ferredoxin. *Plants (Basel)* **9**: 1157.
- Stuart, D., Sandström, M., Youssef, H.M., Zakhrebakova, S., Jensen, P.E., Bollivar, D., and Hansson, M.** (2021). Barley *Viridis-k* links an evolutionarily conserved C-type ferredoxin to chlorophyll biosynthesis. *Plant Cell* **33**: 2834-2849.
- Tanaka, R., and Tanaka, A.** (2007). Tetrapyrrole biosynthesis in higher plants. *Annu Rev Plant Biol* **58**: 321-346.
- Trott, O., and Olson, A.J.** (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**: 455-461.
- Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Zidek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., Velankar, S., Kleywegt, G.J., Bateman, A., Evans, R., Pritzel, A., Figurnov, M., Ronneberger, O., Bates, R., Kohl, S.A.A., Potapenko, A., Ballard, A.J., Romera-Paredes, B.,**

- Nikolov, S., Jain, R., Clancy, E., Reiman, D., Petersen, S., Senior, A.W., Kavukcuoglu, K., Birney, E., Kohli, P., Jumper, J., and Hassabis, D.** (2021). Highly accurate protein structure prediction for the human proteome. *Nature* **596**: 590-596.
- Wang, P., Ji, S., and Grimm, B.** (2022). Post-translational regulation of metabolic checkpoints in plant tetrapyrrole biosynthesis. *J Exp Bot* **73**: 4624-4636.
- Wu, Z., Zhang, X., He, B., Diao, L., Sheng, S., Wang, J., Guo, X., Su, N., Wang, L., Jiang, L., Wang, C., Zhai, H., and Wan, J.** (2007). A chlorophyll-deficient rice mutant with impaired chlorophyllide esterification in chlorophyll biosynthesis. *Plant Physiol* **145**: 29-40.

SUPPLEMENTARY FIGURES



Supplementary Figure 1. The frequency distribution of sequencing depth (number of reads covering each SNP) for called SNPs. The shaded area corresponds to SNPs with a sequencing depth of 31x-80x for the two phenotypic bulks combined and represents 69% of the data or approximately one standard deviation from a normal distribution. These numbers divided by 2 were used for filtering the minimum and maximum sequencing depth SNPs for each phenotypic bulk.


```

xan-j      MATSHFLAAAAATSSSSTAFRPPLRFLSLPPPSLTLSRRRFPFPVCAADADAKETTKPKV 60
xan-j.19  MATSHFLAAAAATSSSSTAFRPPLRFLSLPPPSLTLSRRRFPFPVCAADADAKETTKPKV 60
xan-j.59  MATSHFLAAAAATSSSSTAFRPPLRFLSLPPPSLTLSRRRFPFPVCAADADAKETTKPKV 60
xan-j.64  MATSHFLAAAAATSSSSTAFRPPLRFLSLPPPSLTLSRRRFPFPVCAADADAKETTKPKV 60

xan-j      PEKAPAAGSSFNQLLGIKGAKQEDNIWKIRLQLTKPVTWPPLVWVLCGAAASGNFHWTV 120
xan-j.19  PEKAPAAGSSFNQLLGIKGAKQEDNIWKIRLQLTKPVTWPPLVWVLCGAAASGNFHWTV 120
xan-j.59  PEKAPAAGSSFNQLLGIKGAKQEDNIWKIRLQLTKPVT----- 98
xan-j.64  PEKAPAAGSSFNQLLGIKGAKQEDNIWKIRLQLTKPVTWPPLVWVLCGAAASGNFHWTV 120

xan-j      EDVAKSIVCMLMSGPCLTGYTQTINDWYDRDIDAINEPYRPIPSGAISENEVITQIWVLL 180
xan-j.19  EDVAKSIVCMLMSGPCLTGYTQTINDWYDRDIDAINEPYRPIPSGAISENEVITQIWVLL 180
xan-j.59  ----- 98
xan-j.64  EDVAKSIVCMLMSGPCLTGYTQTINDWYDRDIDAINEPYRPIPSGAISENEVITQIWVLL 180

xan-j      LAGLGLGALLDVWAGHDFPIIFYLALGGSLLSYIYSAPPLKQNGWIGNFALGASYIGL 240
xan-j.19  LAGLGLGALLDVWAGHDFPIIFYLALGGSLLSYIYSAPPLKQNGWIGNFALGASYIGL 240
xan-j.59  ----- 98
xan-j.64  LAGLGLGALLDVWAGHDFPIIFYLALGGSLLFYIYSAPPLKQNGWIGNFALGASYIGL 240

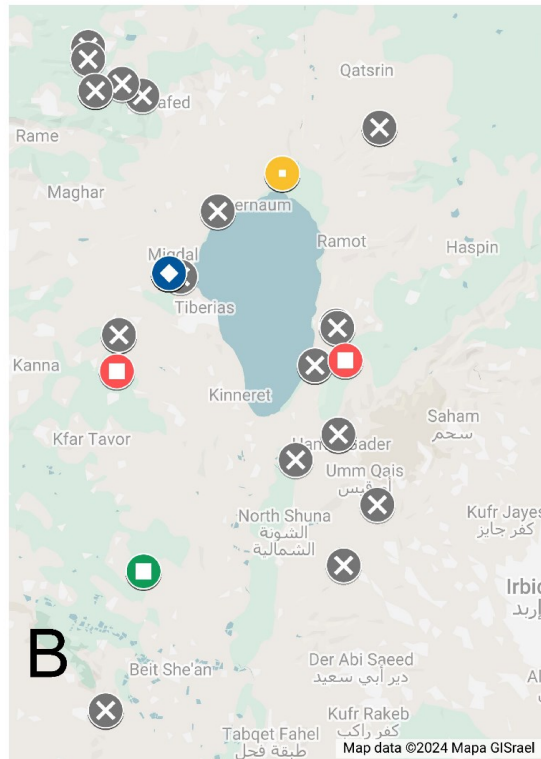
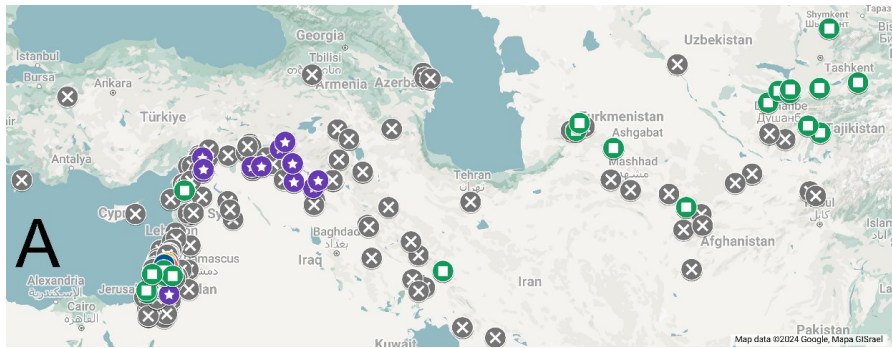
xan-j      PWWAGQALFGTLTPDIVVLTTLYSIAGLGIAIVNDFKSIEGDRTLGLQSLPVAFGMDTAK 300
xan-j.19  PWWAGQALFGTLTPDIVVLTTLYSIAGLGIAIVNDFKSIEGDRTLGLQSLPVAFGMDTAK 300
xan-j.59  ----- 98
xan-j.64  PWWAGQALFGTLTPDIVVLTTLYSIAGLGIAIVNDFKSIEGDRTLGLQSLPVAFGMDTAK 300

xan-j      WICVGAIDITQLSVAAYLLSTGKLYYALALVGLTIPQVILQFQYFLKDPVKYDVKYQASA 360
xan-j.19  WICVGAIDITQLSVAAYLLSTGKLYYALALVGLTIPQVILRFSTS----- 345
xan-j.59  ----- 98
xan-j.64  WICVGAIDITQLSVAAYLLSTGKLYYALALVGLTIPQVILQFQYFLKDPVKYDVKYQASA 360

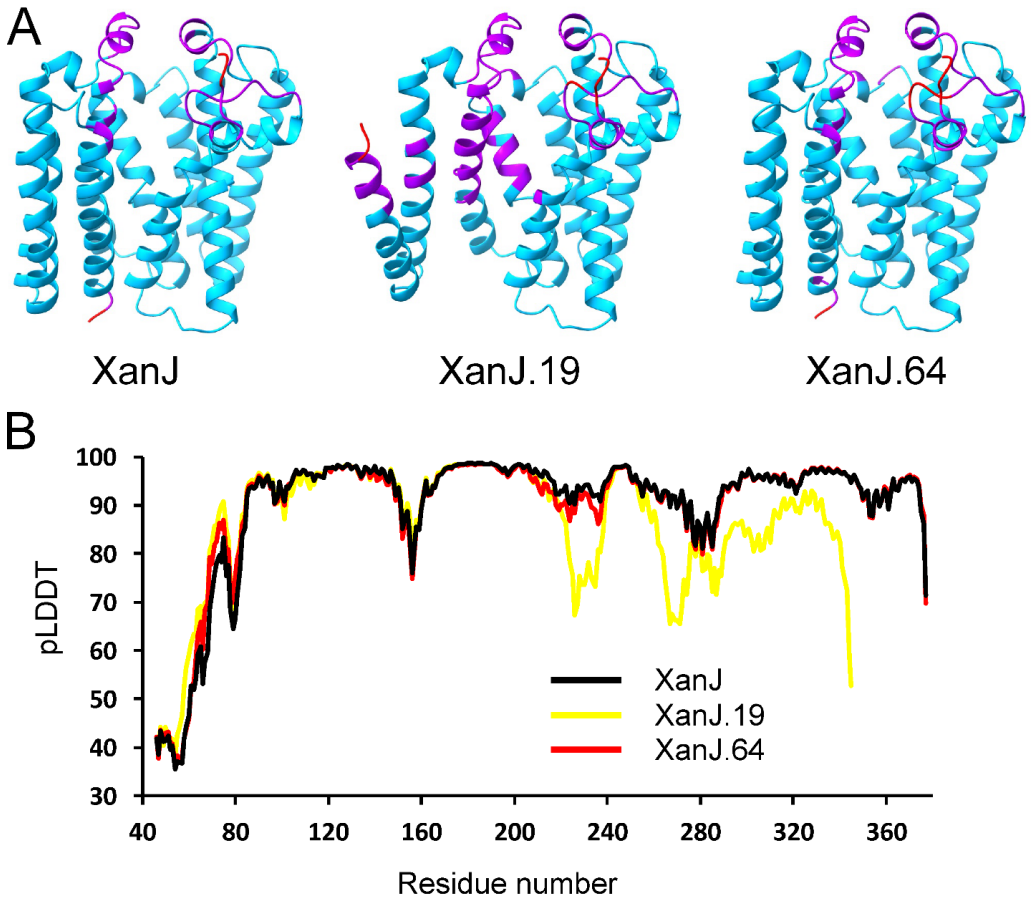
xan-j      QPFFVFGLLVTALATSH 377
xan-j.19  ----- 345
xan-j.59  ----- 98
xan-j.64  QPFFVFGLLVTALATSH 377

```

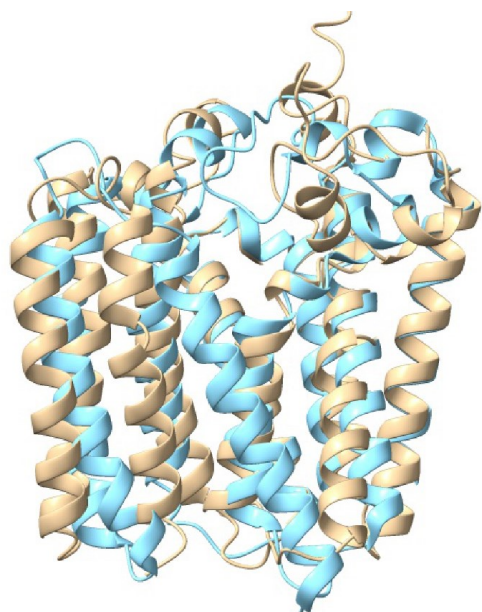
Supplementary Figure 2. Alignment of barley XanJ and the polypeptides resulting from the mutations in the three mutants *xan-j.19*, *xan-j.59* and *xan-j.64*. The predicted chloroplast transit peptide is underlined. Mutation *xan-j.59* results in an early stop codon. The *xan-j.64* mutation results in substitution of a serine with a phenylalanine. The *xan-j.19* mutation results in a frameshift. Non-native amino-acid residues are indicated by black boxes.



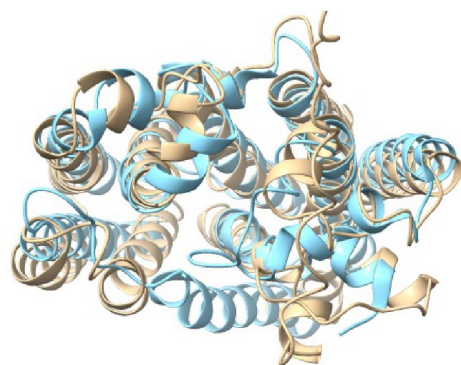
Supplementary Figure 3. Geographical distribution of barley lines with variation in the chlorophyll synthase gene *xan-j*. **A.** Lines with a Q343K substitution (purple with white star symbol) were only collected in regions on the west part of the map. The location of lines with no modifications of the polypeptide sequence or with modifications in the transit peptide have been indicated with black (with white cross) and green (with white square) dots, respectively. **B.** The wild barley lines with modifications I86T (blue with white square), V273I (yellow with white dot) and L290F (red with white square) were all collected around the Sea of Galilee.



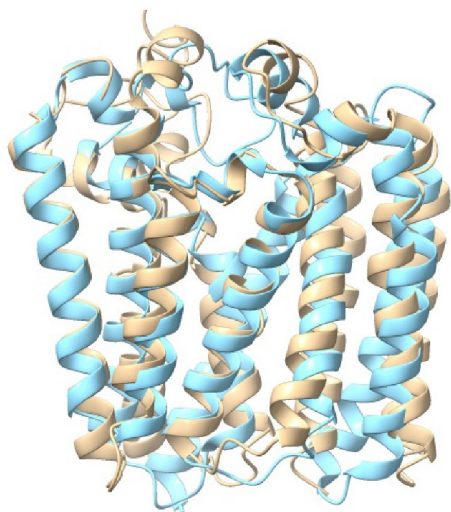
Supplementary Figure 4. Structural prediction of barley XanJ. **A.** AlphaFold generated structural models of XanJ and the proteins resulting from the mutations *xan-j.19* and *xan-j.64*. Blue parts of the structures have a pLDDT score above 90, purple between 70 and 90, red below 70. **B.** The pLDDT (predicted local distance difference test) score value for each amino-acid residues along the three polypeptides.



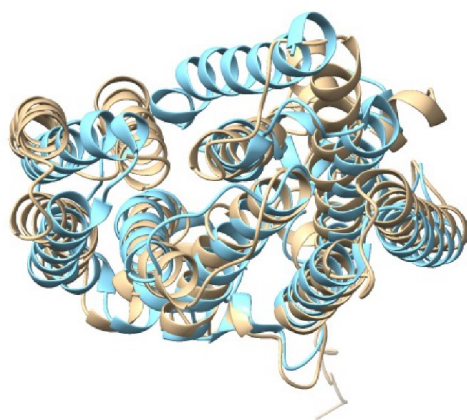
Front



Top

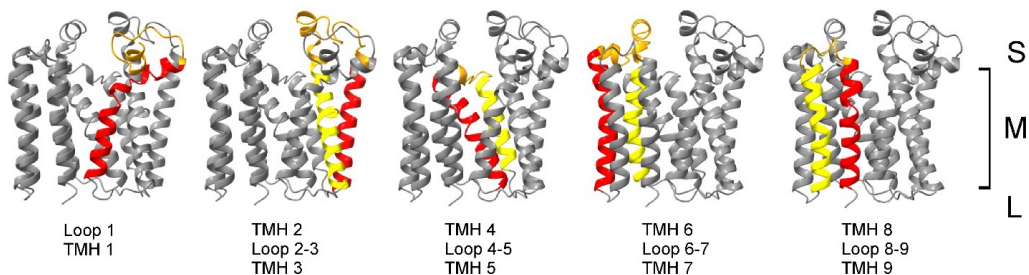


Back

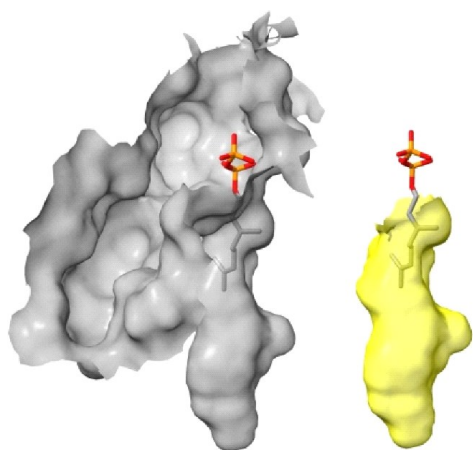


Bottom

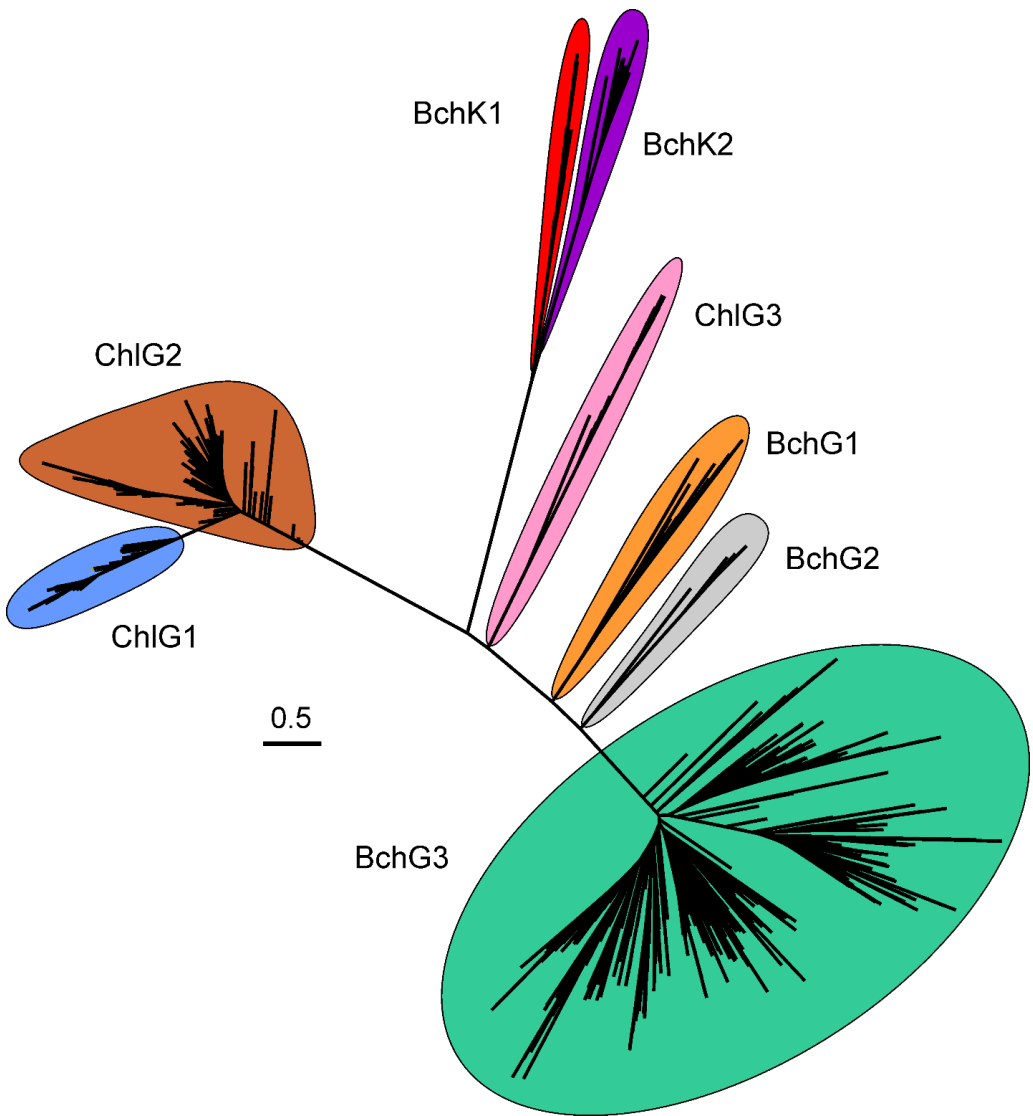
Supplementary Figure 5. Alignment of the XanJ model (brown) to the UbiA structure of *Archaeoglobus fulgidus* (blue) determined by x-ray crystallography.



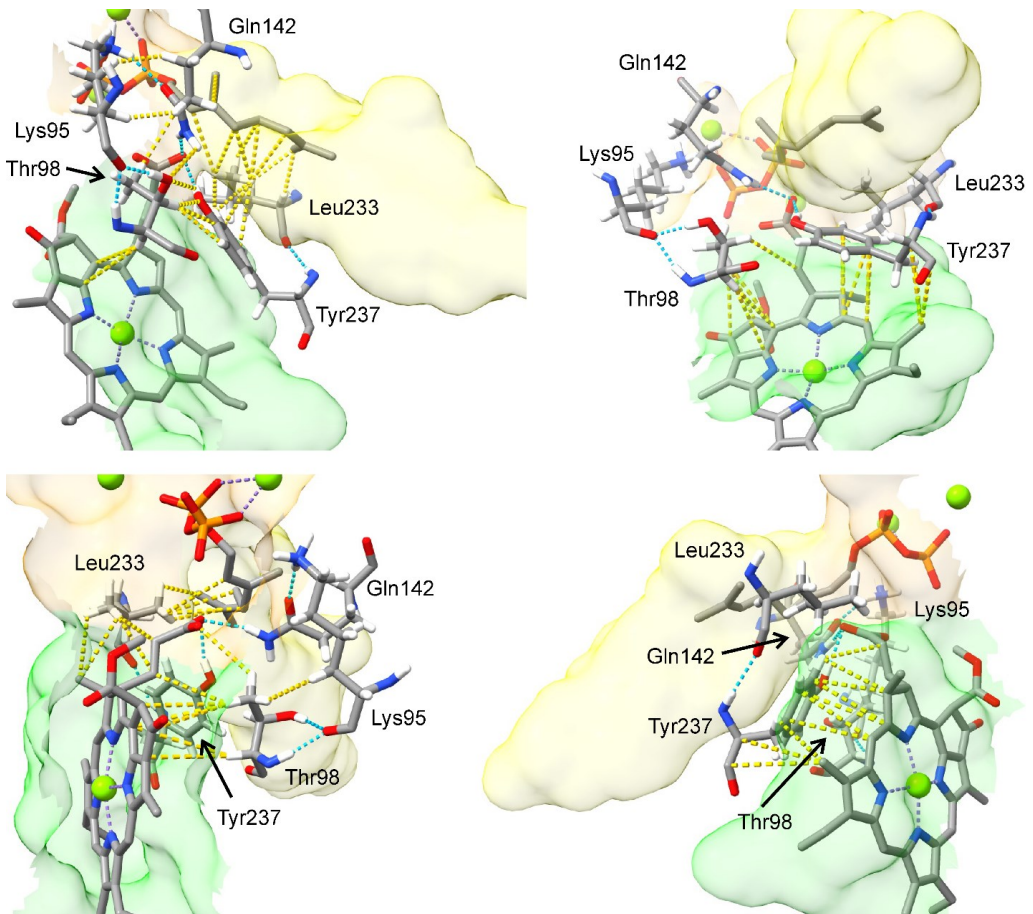
Supplementary Figure 6. AlphaFold generated structural model of barley XanJ. The five views show the consecutive order of the nine transmembrane helices (TMH). Odd and even numbered helices are colored red and yellow, respectively. The loops on the stroma (S) side of the membrane (M) are colored orange. L, lumen side.



Supplementary Figure 7. The GPP in the active site of XanJ after alignment of XanJ to the GPP bound AfUbiA structure. The location of the GPP molecule suggested the binding site of GGPP/PhyPP to an elongated cavity (yellow on the right) with the pyrophosphate in a second more globular cavity.



Supplementary Figure 8. Maximum likelihood phylogenetic tree based on 3749 sequences of chlorophyll synthase (ChIG), bacteriochlorophyll synthase (BchG) and bacteriochlorophyll c synthase (BchK). Eight groups were assigned; ChIG1, ChIG2, ChIG3, BchK1, BchK2, BchG1, BchG2, and BchG3. The scale shows Maximum likely distance.



Supplementary Figure 9. Position of five key amino-acid residues in the active site of barley XanJ. Y237, Q142 and L233 are part of all three subcompartments; the PPP tunnel (yellow), the catalytic cavity (orange) and the tetrapyrrole-binding pocket (green). Y237, Q142 and T98 are suggested to interact with the carbon-17 propionate oxygens of the tetrapyrrole substrate. Chlorophyllide α , GPP and magnesium ions (green spheres) are also displayed in the figure. Red, oxygen; orange, phosphorus; blue, nitrogen; white, hydrogen.

SUPPLEMENTARY TABLES

Supplementary Table 1. Total number of raw 150 bp paired end reads obtained for each phenotypic bulk and the percentage of reads retained after mapping to the barley genome as well as after only retaining reads with a map quality greater than 60 (MQ60).

	<i>xan-j.59</i>		<i>xan-l.82</i>	
	Total number	% of raw data	Total number	% of raw data
Raw paired end reads	443,954,686	100	632,738,242	100
Mapped reads	805,748,724	90.7	1,125,755,683	89.0
Mapped reads at MQ60	491,384,209	55.3	685,391,660	54.2

Supplementary Table 2. The percentage of the barley genome covered by sequencing reads (genome coverage) and how many reads cover each of those bases (sequencing depth). The numbers for the Raw paired end reads are theoretical since they assume complete coverage.

Mutant bulk	Reads used	Genome coverage (%)	Sequence depth
<i>xan-j.59</i>	Raw paired end reads	100	30.7
	Mapped reads	94.5	27.6
	Mapped reads at MQ60	78.8	17.0
<i>xan-l.82</i>	Raw paired end reads	100	43.7
	Mapped reads	94.8	38.6
	Mapped reads at MQ60	80.8	24.2

Supplementary Table 3. Candidate genes containing homozygous mutations in the mapped chromosomal region of mutants *xan-j.59* and *xan-l.82* that alter amino acid sequence of a protein.

	Gene name	Chromosome	Position (bp)	Confidence annotation	Functional annotation
<i>xan-j.59</i>	HORVU.MOR EX.r2.1HG0006030	chr1H	15,716,985	HC2	SKI family transcriptional corepressor 1
	HORVU.MOR EX.r2.1HG0049060	chr1H	411,433,841	HC1	Chlorophyll synthase
<i>xan-l.82</i>	HORVU.MOR EX.r2.3HG0198710	chr3H	66,274,901	HC2	Cytosolic Fe-S cluster assembly factor NBP35
	HORVU.MOR EX.r2.3HG0214090	chr3H	215,850,388	HC1	Magnesium-protoporphyrin IX monomethyl ester [oxidative] cyclase
	HORVU.MOR EX.r2.3HG0221700	chr3H	330,923,492	HC2	Disease resistance protein (TIR-NBS-LRR class) family
	HORVU.MOR EX.r2.3HG0224600	chr3H	362,737,081	HC1	Pentatricopeptide repeat-containing protein
	HORVU.MOR EX.r2.3HG0225130	chr3H	368,518,635	HC2	Serine/threonine protein phosphatase 7 long form isogeny
	HORVU.MOR EX.r2.3HG0227200	chr3H	390,534,682	HC2	D-3-phosphoglycerate dehydrogenase
	HORVU.MOR EX.r2.3HG0231670	chr3H	431,514,564	HC1	Rop guanine nucleotide exchange factor
	HORVU.MOR EX.r2.3HG0233580	chr3H	448,591,422	HC2	Protein phosphatase 2C

Supplementary Table 4. Natural allelic variation in *xan-j* according to a panel of 815 barley cultivars, landraces and wild barley lines (Chen et al., 2022). het – heterozygous allele.

Modification	Allele variant (ref/alt)	Bp position chr 1H	Lines
T3A	Acc/Gcc	435331513	B1K-32-02
A12AA	gcc/gccGCC	435331483	B1K-07-11 (het)
L24P	cTc/cCc	435331449	B1K-03-04, B1K-03-09 (het), FT31 (het), FT376, FT613, FT871, FT885, FT886 (het), WB-500, WBDC_007, WBDC_009, WBDC_207, WBDC_210, WBDC_212, WBDC_213, WBDC_216, WBDC_220, WBDC_225, WBDC_228, WBDC_233, WBDC_326, WBDC_336, WBDC004
P32H	cCt/cAt	435331425	B1K-29-13, WB-507
I86T	aTa/aCa	435330735	B1K-29-04
V273I	Gta/Ata	435326789	B1K-37-15
L290F	Ctt/Ttt	435326665	FT393, WB-511
Q343K	Cag/Aag	435325695	FT507, FT581, FT582 (het), FT584, FT595, FT604, FT632, FT660, FT741 (het), FT754, HID069, WB-489 (het), WB-490, WB-491, WB-497 (het), WB-498, WBDC_149, WBDC009, WBDC010

Supplemental Table 5. Summary of amino-acid residues located in the active site of XanJ and/or previously studied in related enzymes.

HvXanJ Residue	Topology	Secondary structure	Mutation	Enzyme	Functional information from Ref	Contribution to active site surface	Active site subcompartment	Ref
M1-C46	Removed				cTP	NA	NA	TargetP 2.0
A47-I86	Stroma	L 1	N-term truncation	AsChlG	46 % activity	NA	NA	Schmid et al., 2001
W87	Stroma / Membrane	TMH 1 - First residue	N-term truncation	AsChlG	Inactive	NA	NA	Schmid et al., 2001
R90	Stroma / Membrane	TMH 1	R91A	AsChlG	Inactive	NA	NA	Schmid et al., 2001
K95	Stroma / Membrane	TMH 1				Sidechain	Catalytic cavity	This study
			R22A	AfUbiA	No GPP binding			Huang et al., 2014
			K11A	EcCyoE	Inactive			Saiki et al., 1993
V97	Membrane	TMH 1				Backbone + Sidechain	TP pocket	This study
T98	Membrane	TMH 1				Backbone + Sidechain	TP pocket	This study
P100	Membrane	TMH 1				Backbone + Sidechain	TP pocket	This study

P101	Membrane	TMH 1				affinity for Behlride <i>a</i> by the mutation of F28I	Backbone + Sidechain	TP pocket	This study
L102	Membrane	TMH 1					Sidechain	GGPP tunnel	
W104	Membrane	TMH 1					Backbone + Sidechain (mostly)	TP pocket	This study
G105	Membrane	TMH 1					Backbone	TP pocket	
C108	Membrane	TMH 1					Sidechain	TP pocket	This study
					C109A	7% activity			Schmid et al., 2001
C129	Membrane	TMH 2			C130A	C130A 33% activity. NPM treatment 13% activity. C130A NPN 9% activity.			Schmid et al., 2001
M132	Membrane	TMH 2					Sidechain	GGPP tunnel	This study
S133	Membrane	TMH 2					Backbone + Sidechain	GGPP tunnel	This study
C136	Membrane	TMH 2			C137A	114% activity. Same inhibition by	NA	NA	Schmid et al., 2001

			D72A	AfUbiA	No GPP binding				Huang et al., 2014
			D61A	EcCyoE	Inactive				Saiki et al., 1993
R150	Stroma	L 2-3	R151A	AsChlG	24.14% activity	NA	NA	NA	Schmid et al., 2001
D151	Stroma	L 2-3	D63A	EcCyoE	Inactive	NA	NA	NA	Saiki et al., 1993
I152	Stroma	L 2-3				Sidechain		Catalytic cavity	This study
D153	Stroma	L 2-3				Sidechain		Catalytic cavity	This study
			D154A	AsChlG	Inactive				Schmid et al., 2002
			D75A	EcUbiA	Inactive				Cheng and Li, 2014
			D58A	ApUbiA	17% GSPP affinity				Cheng and Li, 2014
			D65A	EcCyoE	Inactive				Saiki et al., 1993
N156	Stroma	L 2-3				Sidechain (mostly)		Catalytic cavity	This study
			M68A	EcCyoE	Active				Saiki et al., 1993
E157	Stroma	L 2-3				Sidechain		Catalytic cavity	This study

				R80A	EcUbiA	7% activity			Cheng and Li, 2014
				R63A	ApUbiA	0% GSPP affinity			Cheng and Li, 2014
E157-Gap-P158	Stroma	L 2-3		R70A	EcCyoE	Inactive			Saiki et al., 1993
E157-Gap-P158	Stroma	L 2-3		T71A	EcCyoE	Active			Saiki et al., 1993
R160	Stroma	L 2-3					Sidechain	Catalytic cavity	This study
				R161A	AsChlG	No trace of activity			Schmid et al., 2001
				R161A	AsChlG	R161A 0% activity			Schmid et al., 2002
				R161K	AsChlG	34% activity			Schmid et al., 2002
				R161H	AsChlG	1% activity			Schmid et al., 2002
				R84A	EcUbiA	12% activity. partially complement <i>E. coli</i> growth.			Cheng and Li, 2014

			R67A	ApUbiA	2.41% GSPP affinity				Cheng and Li, 2014
			R74A	EcCyoE	Inactive				Saiki et al., 1993
G165	Stroma	L 2-3	G79A	EcCyoE	Active				Saiki et al., 1993
G187	Membrane	TMH 3	L101A	EcCyoE	Active				Saiki et al., 1993
P199	Membrane	HL3-4	P198S	OsChlG	22-35% activity				Wu et al., 2007
I201	Membrane	TMH 4					Backbone + Sidechain	GGPP tunnel	This study
F202	Membrane	TMH 4	C110A	EcCyoE	Active				Saiki et al., 1993
L204	Membrane	TMH 4					Backbone + Sidechain	GGPP tunnel	This study
A205	Membrane	TMH 4					Backbone + Sidechain	GGPP tunnel	This study
G207	Membrane	TMH 4	G217R	AtChlG	Temperature sensitivity				Lin et al., 2014
G208	Membrane	TMH 4					Backbone	GGPP tunnel	This study
S209	Membrane	TMH 4					Backbone + Sidechain	GGPP tunnel	This study
L211	Membrane	TMH 4					Backbone + Sidechain	GGPP tunnel	This study
S212	Membrane	TMH 4	S212F	<i>xan-j.64</i>	Xantha phenotype. Not light sensitive.		Backbone + Sidechain	GGPP tunnel	This study

Y215	Membrane	TMH 4	Y120A	EcCyoE	Inactive			Sidechain	GGPP tunnel and Catalytic cavity	Saiki et al., 1993 This study
			Y141L	EcMenA	Inactive					Huang et al., 2014
			Y132A	EcUbiA	Inactive					Cheng and Li, 2014
			Y115A	ApUbiA	5% GSPP affinity					Cheng and Li, 2014
			Y124A	EcCyoE	Inactive					Saiki et al., 1993
S216	Membrane	TMH 4 - Last residue						Sidechain	GGPP tunnel	This study
K223	Stroma / Membrane	L 4-5						Backbone + Sidechain	Catalytic cavity	This study
			K136A	EcUbiA	Inactive					Cheng and Li, 2014
			K119A	ApUbiA	Not purified					Cheng and Li, 2014
			K146A	AfUbiA	No GPP binding					Huang et al., 2014
			K129A	EcCyoE	Inactive					Saiki et al., 1993
Q224	Stroma / Membrane	L 4-5						Backbone + Sidechain	Catalytic cavity	This study

				R137A	EcUbiA	84% activity				Cheng and Li, 2014
				R130A	EcCyoE	Active				Saiki et al., 1993
Q224-Gap-N225	Stroma / Membrane	L 4-5	H131A	EcCyoE	Active					Saiki et al., 1993
N225	Stroma / Membrane	L 4-5 - Last Residue					Backbone	Catalytic cavity		This study
G226	Stroma / Membrane	TMH 5 - First Residue					Backbone	Catalytic cavity		This study
G229	Membrane	TMH 5					Backbone	Catalytic cavity		This study
N230	Membrane	TMH 5					Sidechain (mostly)	Catalytic cavity		This study
A232	Membrane	TMH 5					Backbone + Sidechain	GGPP tunnel		This study
L233	Membrane	TMH 5					Backbone + Sidechain	GGPP tunnel + TP pocket + Catalytic Cavity		This study
				G139A	EcCyoE	Active				Saiki et al., 1993
G234	Membrane	TMH 5					Backbone	TP pocket		This study
S236	Membrane	TMH 5					Backbone + Sidechain	GGPP tunnel		This study
Y237	Membrane	TMH 5					Sidechain (mostly)	Catalytic cavity + GGPP tunnel + TP pocket		This study
				G143A	EcCyoE	Inactive				Saiki et al., 1993

I238	Membrane	TMH 5 (in helix break)						Sidechain (mostly)	TP pocket	This study
L240	Membrane	TMH 5						Backbone + Sidechain	GGPP tunnel	This study
P241	Membrane	TMH 5						Backbone + Sidechain	GGPP tunnel	This study
			P146A	EcCyoE	Inactive					Saiki et al., 1993
W242	Membrane	TMH 5						Sidechain	TP pocket	This study
			P147A	EcCyoE	Active					Saiki et al., 1993
A242	Membrane	TMH 5						Sidechain	GGPP tunnel	This study
G245	Membrane	TMH 5	G150A	EcCyoE	Active					Saiki et al., 1993
Q246	Membrane	TMH 5	Y151A	EcCyoE	Inactive					Saiki et al., 1993
T261	Membrane	TMH 6	C262A	AsChlG	114% activity. Same inhibition by NPM as the wild-type.					Schmid et al., 2001
A266	Membrane	TMH 6	W172A	EcCyoE	Inactive					Saiki et al., 1993
G267	Membrane	TMH 6						Backbone	TP pocket	This study
G269	Membrane	TMH 6	P175A	EcCyoE	Inactive					Saiki et al., 1993
I270	Membrane	TMH 6						Sidechain (mostly)	TP pocket	This study

				H199A	BsCOXX1	Inactive. Proposed to be heme <i>b</i> binding residue.					Mogi 2009.
A271		Membrane	TMH 6	H176A	EcCyoE	Inactive		Backbone + Sidechain	TP pocket and Catalytic cavity		Saiki et al., 1993 This study
				D184A	EcUbiA	Inactive. Probably not functionally equivalent to A271.					Cheng and Li, 2014
				D175A	ApUbiA	5% GSPP affinity. Probably not functionally equivalent to A271.					Cheng and Li, 2014
N274		Membrane	TMH 6 - Last Residue	Y187A	EcUbiA	19% activity. Maybe not functionally		Backbone + Sidechain (mostly)	TP pocket + Catalytic cavity		This study Cheng and Li, 2014

						equivalent to N274.						
				Y178A	ApUbiA	20% GSPP affinity. Maybe not functionally equivalent to N274.						Cheng and Li, 2014
D275	Stroma / Membrane	L 6-7 - First Residue						Backbone + Sidechain	Catalytic cavity			This study
K277	Stroma	L 6-7						Backbone + Sidechain	TP pocket			This study
S278	Stroma	L 6-7						Backbone + Sidechain	Catalytic cavity			This study
				D204A	EcMenA	Inactive						Huang et al., 2014
				D191A	EcUbiA	Inactive						Cheng and Li, 2014
				D182A	ApUbiA	0% GSPP affinity						Cheng and Li, 2014
				D198A	AfUbiA	Decreased GPP binding						Huang et al., 2014
G281	Stroma	L 6-7						Backbone	Catalytic cavity			This study
				D187A	EcCyoE	Inactive						Saiki et al., 1993

C303	Membrane	TMH 7	C304A	AsChIG	Active				Schmid et al., 2001
Y325	Membrane	TMH 8	K3232A	EcCyoE	Active				Saiki et al., 1993
D348	Stroma	L 8-9	D256A	EcCyoE	Inactive				Saiki et al., 1993
P349	Stroma	L 8-9	D257A	EcCyoE	Inactive				Saiki et al., 1993
D353	Stroma	L 8-9 - Last residue					Backbone + Sidechain	Top ridge over TP pocket left side	This study
V354	Stroma / Membrane	TMH 9 - First Residue					Backbone + Sidechain	Top ridge over TP pocket left side	This study
Q357	Stroma / Membrane	TMH 9					Backbone + Sidechain	TP pocket	This study
			F265A	EcCyoE	Active				Saiki et al., 1993
A358	Membrane	TMH 9					Backbone + Sidechain	Top ridge over TP pocket left side	This study
Q361	Membrane	TMH 9					Sidechain (mostly)	TP pocket	This study
			S268A	EcCyoE	Inactive			TP pocket	Saiki et al., 1993
P362	Membrane	TMH 9					Backbone + Sidechain	Top ridge over TP pocket left side	This study
F364	Membrane	TMH 9					Sidechain	TP pocket	This study
V365	Membrane	TMH 9					Sidechain	TP pocket	This study

F366	Membrane	TMH 9				Sidechain	Top ridge over TP pocket left side	This study
L368	Membrane	TMH 9				Sidechain	TP pocket	This study
A374	Membrane	TMH 9	D282A	EcCyoE	Inactive			Saiki et al., 1993
S376	Membrane / Lumen	C term penultimate residue	S377 C-term truncation	AsChlG	No activity			Schmid et al., 2001
			S377A		63% activity			

References

- HUANG, H., LEVIN, E. J., LIU, S., BAI, Y., LOCKLESS, S. W. & ZHOU, M. 2014. Structure of a membrane-embedded prenyltransferase homologous to UBIAD1. *PLoS Biol.* 12, e1001911.
- KIM, E. J., KIM, H. & LEE, J. K. 2016. The photoheterotrophic growth of bacteriochlorophyll synthase-deficient mutant of *Rhodospirillum rubrum* is restored by 144F mutant chlorophyll synthase of *Synechocystis* sp. PCC 6803. *J Microbiol Biotechnol.* 26, 959-966.
- LIN, Y. P., LEE, T. Y., TANAKA, A. & CHARNG, Y. Y. 2014. Analysis of an Arabidopsis heat-sensitive mutant reveals that chlorophyll synthase is involved in reutilization of chlorophyllide during chlorophyll turnover. *Plant J.* 80, 14-26.
- MOGI, T. 2009. Over-expression and characterization of *Bacillus subtilis* heme O synthase. *J Biochem.* 145, 669-675.
- SAIKI, K., MOGI, T., HORI, H., TSUBAKI, M. & ANRAKU, Y. 1993. Identification of the functional domains in heme O synthase. Site-directed mutagenesis studies on the *cyoE* gene of the cytochrome *bo* operon in *Escherichia coli*. *J Biol Chem.* 268, 26927-26934.
- SCHMID, H. C., OSTER, U., KÖGEL, J., LENZ, S. & RÜDIGER, W. 2001. Cloning and characterisation of chlorophyll synthase from *Avena sativa*. *Biol Chem.* 382, 903-911.
- SCHMID, H. C., RASSADINA, V., OSTER, U., SCHOCH, S. & RÜDIGER, W. 2002. Pre-loading of chlorophyll synthase with tetraprenyl diphosphate is an obligatory step in chlorophyll biosynthesis. *Biol Chem.* 383, 1769-1778.
- WU, Z., ZHANG, X., HE, B., DIAO, L., SHENG, S., WANG, J., GUO, X., SU, N., WANG, L., JIANG, L., WANG, C., ZHAI, H. & WAN, J. 2007. A chlorophyll-deficient rice mutant with impaired chlorophyllide esterification in chlorophyll biosynthesis. *Plant Physiol.* 145, 29-40.

Supplementary Table 6. Primers used for PCR and RT-qPCR amplifications of *xan-j* (HORVU.MOREX.r2.1HG0049060) and RT-qPCR of *xan-f* (Olsson et al., 2004).

Primer name	Sequence (5' to 3')
ChlG F1	ATCCATTACACGCACGGACC
ChlG R1	CCCAATTCAACACATGCCCCG
ChlG F2	TTTGTTTTGGTGGGTTGGGC
ChlG R2	GAACGACCATTGCAAGCCAG
ChlG F3	TGTCTGGGGAGTGCTTTGTG
ChlG R3	ACATGGACCCGACATTAGCA
ChlG F4	ATCACGGCTTCTGTGCTCAT
ChlG R4	TATCCCACGCAACTAGCACC
ChlG F5	ACCGTCTTCGGGAGTGAGAT
ChlG R5	TGGGGGTAACAAAAGCCTCC
ChlG F6	AAACACAATGCACGCCTGAC
ChlG R6	AGACAAAGTCCCTGTGCAGC
ChlG F7	ACCGCATGCTCACAAGAAGT
ChlG R7	ACGATCGGAAGCTCGTCATT
ChlG F8	GCATAACTTGCCTGTCCCCA
ChlG R8	AGCCAATGTAACTCGCACCA
ChlG F9	GGGTGGGTCCTTGCTTTCTT
ChlG R9	AGTGGTGTATGCCTTGGAGC
ChlG F10	ACAACCCTAGCCCCGTGATTG
ChlG R10	GTCCCAGAGTTCTATCCCCCT
ChlG F11	TCTCTTTTGTGCAGGTGGGC
ChlG R11	ACTACGGCAAGCTTTGGGTG
ChlG F12	AGGAATGGGTTCGCACTTCAA
ChlG R12	CGTCGTACTTCACAGGGTCC
ChlG F13	GGCTTGTCAGACTGGCCTTT
ChlG R13	ATGCATGCAGCCGCTAATAC
Xan J QPCR-F2	CTTGTCTGGGGAGTGCTTT
Xan J QPCR-R2	GTAAGACATGGACCCGACAT
Xan F QPCR-F2	GGGTGTGTTTCAGAGACCTTT
Xan F QPCR-R2	TCATCTCGATTGGCTCTTCC



Department of Biology
Faculty of Science
Lund University
ISBN 978-91-8104-048-7

