Development of novel riboswitches for synthetic biology in the green alga Chlamydomonas

Authors:
Payam Mehrshahi1, Ginnie Trinh D.T. Nguyen2, Aleix Gorchs Rovira5, Andrew Sayer1, Marcel Llavero Pasquina1, Michelle Huei Sin Lim4, Elliot J Medcalf1, Gonzalo I Mendoza Ochoa1, Mark A Scaife3, Alison G Smith1*

Current Address:

1Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK

2GTDTN: Glanbia Performance Nutrition Canada Inc., 3500 Lacey Road, Suite 1200, Downers Grove, IL 60515, USA
gnguyen@glanbia.com

3MAS: Mara Renewables Corporation (101A Research Drive, Dartmouth, Nova Scotia, B2Y 4T6) mscaife@maracorp.ca

4MHSL: John Swire & Sons (H.K.) Ltd. 33/F One Pacific Place, 88 Queensway, Hong Kong
michelle.lim205@gmail.com

5Coauthor deceased
AGR: This article is dedicated to our co-author and friend Aleix Gorchs Rovira who was instrumental to this story and was tragically killed in September 2019 before submission of this article.

*Corresponding author:
Prof Alison G Smith
Tel: +44-1223 333952
Fax: +44-1223 333953
Email: as25@cam.ac.uk
ORCID: 0000-0001-6511-5704
Abstract
Riboswitches are RNA regulatory elements that bind specific ligands to control gene expression. Because of their modular composition, where a ligand-sensing aptamer domain is combined with an expression platform, riboswitches offer unique tools for synthetic biology applications. Here we took a mutational approach to determine functionally important nucleotide residues in the thiamine pyrophosphate (TPP) riboswitch in the THI4 gene of the model alga Chlamydomonas reinhardtii, allowing us to carry out aptamer swap using THIC aptamers from Chlamydomonas and Arabidopsis thaliana. These chimeric riboswitches displayed a distinct specificity and dynamic range of responses to different ligands. Our studies demonstrate ease of assembly as 5’UTR DNA parts, predictability of output, and utility for controlled production of a high-value compound in Chlamydomonas. The simplicity of riboswitch incorporation in current design platforms will facilitate generation of genetic circuits to advance synthetic biology and metabolic engineering of microalgae.

Keywords
Riboswitch, Chlamydomonas, diterpenoid, synthetic biology, metabolic engineering, microalgae

Synthetic biology approaches are making significant advances in both the understanding of biological systems and the development of biological engineering. A key feature offered by synthetic biology is the ability to establish designer regulatory circuits for control of gene expression, which can generate predictable and reliable outputs in terms of protein or metabolite production. One way to do this is the use of synthetic promoters/transcription factors, for example integration of the binding site of designer transcription activator-like effectors (TALEs) that can be activated by custom-designed TALEs. Another more direct method, which does not require trans-acting factors, is RNA-based. Riboswitches are one such example; these are regions in the mRNA that modulate expression in response to a small molecule. Upon ligand binding to the aptamer region, the riboswitch undergoes a conformational change, which affects the second functional unit, the expression platform. In prokaryotes there are over 30 naturally-occurring classes of riboswitch, responding to a range of cofactors, amino acids, metal ions and other small metabolites. Many of these, together with computationally designed synthetic riboswitches, or molecularly evolved RNA components have been used for synthetic biology approaches in bacteria, including to allow selection of E. coli strains that overproduce metabolites, to serve as sensors of intracellular metabolites, and to uncover novel aspects of biosynthetic pathways.

In eukaryotes, all riboswitches identified so far respond to thiamine pyrophosphate (TPP). The TPP riboswitch aptamer is one of the best characterised, with several 3D structures known including those from E. coli and Arabidopsis thaliana. Two such riboswitches are found in introns of the THI4 and THIC genes of the green alga Chlamydomonas reinhardtii (hereafter Chlamydomonas). In cells grown without thiamine, the introns are spliced to produce mRNAs
encoding functional enzymes. On thiamine supplementation, cellular TPP levels rise, and binding of TPP to the riboswitches results in alternative splicing. For the THI4 mRNA, shown schematically in Figure 1a, retention of the first intron in the 5'UTR introduces an 81 nt upstream open reading frame (uORF; blue box), which interferes with translation of THI4. Alternative splicing of the THIC transcript results in retention of an intron with an in-frame stop codon. Thiamine (structure 1, Figure S1) is the product of a branched biosynthesis pathway, where condensation of a pyrimidine (structure 2, Figure S1) and a thiazole (structure 3, Figure S1) moiety makes thiamine monophosphate, which is ultimately phosphorylated to the active cofactor TPP. THI4 and THIC catalyse the first committed steps in the thiazole and pyrimidine branches respectively, so the riboswitches enable tight feedback inhibition of each branch of thiamine biosynthesis. Moreover, in addition to TPP, the Chlamydomonas THI4 riboswitch (CrTHI4_RS) responds to the thiazole precursor, 5-hydroxyethyl-4-methylthiazole (HET, structure 3 Figure S1), and CrTHIC_RS to the pyrimidine precursor 4-amino-5-hydroxymethyl-2-methylpyrimidine (HMP; structure 2 Figure S1)\textsuperscript{13}, providing a further level of refinement to the regulation of cellular thiamine homeostasis.

Chlamydomonas is increasingly being used as a chassis for metabolic engineering and for synthetic biology because of its ease of genetic manipulation\textsuperscript{14}, the availability of a comprehensive genome-wide mutant library\textsuperscript{15}, a standardised modular cloning (MoClo) kit\textsuperscript{16} and amenability to production of high value compounds\textsuperscript{17}. Here we dissect the different functional components of Chlamydomonas TPP riboswitches, including functionally important residues and ligand specificity, to characterise its modular parts \textit{in vivo}. Through rational application of this knowledge we generated a novel collection of riboswitches that respond to specific ligands in a dose-dependent manner, allowing transgene expression to be tuned. The possibility to control expression of metabolic enzymes to avoid toxicity or pathway bottlenecks thus provides the means for sophisticated metabolic engineering strategies.

\textbf{Results}

\textbf{Genetic characterisation of the CrTHI4 riboswitch}

To examine in detail the mechanism underlying \textit{CrTHI4}\textsubscript{RS}-mediated gene regulation, a reporter construct was generated in which the phleomycin resistance gene \textit{ble} fused to the green fluorescent protein (\textit{Ble-GFP}), which confers resistance to the antibiotic zeocin\textsuperscript{18}, was placed under the control of the \textit{CrTHI4} 5'UTR containing the riboswitch (Figure 1b, construct pPM1-044). Given that zeocin resistance is a direct reflection of the level of the Ble-GFP protein, the use of this resistance marker in this set up is ideal for testing of riboswitch functionality \textit{in vivo}. The construct was introduced into the Chlamydomonas strain UVM4\textsuperscript{19} and transformants were selected on zeocin. To test the riboswitch function, three independent transgenic lines were spotted on solid media supplemented with increasing concentrations of zeocin. In the absence of any supplementation (Figure 1c, left hand panel), the lines were resistant to 100 μg/ml zeocin (the upper limit tested). In contrast, inclusion in the medium of thiamine, which is phosphorylated to
TPP on uptake into the cell\textsuperscript{12}, resulted in zeocin-susceptibility of the \textit{CrTHI4\textsubscript{RS}}-\textit{Ble-GFP} lines, demonstrated by a reduction in growth, even at the lowest concentration of 10 μg/ml zeocin (Figure 1d). We screened 40 lines containing the reporter construct, and this response was seen in all lines (data not shown).

![Diagram](image1)

**Figure 1.** The Chlamydomonas \textit{THI4} riboswitch can regulate transgene expression in vivo

(a) Schematic of the \textit{CrTHI4} gene showing exon-intron structures as boxes and lines respectively. 5'UTR and 3'UTR regions are shown as black filled boxes, white boxes are coding regions and grey boxes show exons generated upon alternative splicing. When no ligand is present, the introns are completely spliced between the authentic splice sites (red dotted line), producing the short \textit{THI4} transcript, and translation is initiated from the black coloured start codon which results in a functional protein (black tear drop). If levels of TPP in cell rise the riboswitch aptamer within the first intron binds this and results in: alternative splicing between the cryptic splice sites (grey dotted lines), retention of the grey exon(s), and ultimately translation of the 81 nt upstream open reading frame (uORF, blue box), which is translated instead of the native enzyme-encoding \textit{CrTHI4} (its protein product is represented by blue tear drops). (b) The \textit{CrTHI4\textsubscript{RS}}-\textit{Ble-GFP} reporter cassette was made by making the following assembly from 5' to 3' HSP70-RBCS2 promoter\textsuperscript{20}, 22 nt RBCS2 5'UTR, \textit{CrTHI4} riboswitch, \textit{Ble-GFP} (green box) and the carbonic anhydrase 1 (CA1) terminator\textsuperscript{25}. Lines carrying this construct are hereafter referred to as \textit{CrTHI4\textsubscript{RS}}-\textit{Ble-GFP}. (c-f) Viability of three independent \textit{CrTHI4\textsubscript{RS}}-\textit{Ble-GFP} transgenic lines (numerated 1-3) was determined through spot tests on solid media containing 0 to 100 μg/ml of zeocin (concentration indicated at the bottom of panels). (c) In control plates with no supplements, the lines were resistant to all concentrations of zeocin. (d) Addition of 10 μM thiamine to the media resulted in the sensitivity of all lines to even the lowest concentration of zeocin (10 μg/ml). (e) In the presence of 10 μM HET, \textit{CrTHI4\textsubscript{RS}}-\textit{Ble-GFP} lines are sensitive to >10 μg/ml zeocin. (f) \textit{CrTHI4} riboswitch does not respond to HMP\textsuperscript{13}, thus in presence of 10 μM HMP \textit{CrTHI4\textsubscript{RS}}-\textit{Ble-GFP} lines showed resistance to all concentrations of zeocin. Images taken 7 days after plating. * = stop codon.

The endogenous \textit{THI4} gene, responsible for the biosynthesis of the thiazole precursor HET, is responsive to this intermediate, but not to HMP\textsuperscript{13}, so we asked whether the reporter construct \textit{CrTHI4\textsubscript{RS}}-\textit{Ble-GFP} showed the same responses. Inclusion in the media of 10 μM HET (which, like thiamine, is phosphorylated by the cell on uptake) revealed the sensitivity of the lines to zeocin above 30 μg/ml (Figure 1e). In contrast, addition of 10 μM HMP (again phosphorylated on uptake) had little or no effect (Figure 1f). The response of the native \textit{CrTHI4\textsubscript{RS}} to thiamine and HET was
assessed at the transcript level. Expression of the functional transcript, formed after splicing to remove the first intron (Figure 1a), was tested in one of the lines grown for 7 days in 10 µM of each of the compounds. As expected\textsuperscript{12,13}, significant repression of the CrTHI4 short transcript (CrTHI4s) was found in strains cultivated in media supplemented with thiamine and to a lesser extent with HET, but not with HMP (Figure S2a).

To investigate the \textit{CrTHI4}_{RS} aptamer further, a detailed mutagenesis study was conducted. One of the responsive \textit{CrTHI4}_{RS}-Ble-GFP reporter lines was treated with UV mutagenesis, and colonies resistant to 10 µg/ml zeocin were selected to discard lines with mutated and impaired promoter, \textit{Ble} reporter or 3′UTR. These mutagenized lines were then tested on thiamine and zeocin to identify mutants that had lost responsiveness to thiamine, and so were still zeocin resistant. Sequence analysis of the riboswitch in these lines found three with alterations (UV13, UV14 and UV15; Figure S3). These were all deletions of some or all of the aptamer and thus would result in disruption of metabolite-aptamer interaction. In particular, the effect of small 3 nt and 10 nt deletions in the UV13 and UV14 mutants respectively highlights the strict requirement to maintain nucleotides involved in ligand binding (denoted by triangles) and/or base-pairing\textsuperscript{10,11}.

More targeted mutations were introduced in the \textit{CrTHI4}_{RS} to alter specific nucleotides rather than deletions. Site-directed mutagenesis was used to remove the start codon in the uORF (Figure 2a, construct \textit{CrTHI4\_noATG}). After transformation into UVM4 using selection on zeocin, three independent transgenic lines were spot tested as before on media containing zeocin, with and without thiamine or HET (Figure 2b). None of the lines responded to either compound, demonstrating that the uORF was essential for the function of the riboswitch. A more random approach of error-prone PCR of the reporter construct resulted in several different variants, denoted EPa (Figure 2). Spot-tests showed that disruption of 1 out of 9 base pairs in the P1 stem (EPa-38) had little effect on the riboswitch response to thiamine or HET. This indicates the flexibility of the P1 stem with respect to riboswitch function, as was also demonstrated in \textit{Neurospora crassa NMT1} and \textit{A. thaliana THIC} riboswitches\textsuperscript{11,21}. In EPa-39, which has four mutations (A730U, G643A, C614U, U617C), base pairing in the P3 stem is affected and the sequence of the uORF is altered, although the length and reading frame are maintained. However, EPa-39 did not show impaired riboswitch function either. Thus, the sequence of the uORF is unlikely to be involved directly in the ligand-binding nor in the inhibition of expression of the downstream reporter gene. In contrast, an extra complementary base pair in the P5 stem (EPa-49) abolished completely the riboswitch response to either thiamine or HET. The extra base pair is likely to affect the interaction with the β-phosphate group of TPP, as shown in the \textit{E. coli thiM} riboswitch\textsuperscript{22}, as well as reduce the flexibility of the junction between the P4 and P5 stems.

Disruption to riboswitch function was also observed for EPa-50, which carries a single mutation in the conserved region CUGAGA (U606C). This mutant still responded to thiamine (Figure 2b). However, uniquely the mutated riboswitch was no longer responsive to HET. The CUGAGA region
is highly conserved in all TPP riboswitches\textsuperscript{23} and an integral part of the J2/3 bulge that has been shown to serve as the pyrimidine sensing region of the riboswitch\textsuperscript{11}. Interestingly, the U606C mutation in the \textit{EPa-50} line affects the stop codon of the uORF in the riboswitch yet retains responsiveness to thiamine. The most likely reason for this outcome is the utilisation of a downstream stop codon (TGA at -513; Figure S3a), located in frame with the uORF. This scenario would result in an uORF of 58 amino acids extending out of the aptamer, which again suggests that the length of the uORF is not critical, and that its location need not be restricted to the aptamer.

**Figure 2. Detailed molecular characterisation of the \textit{CrTHI4} riboswitch**

Mutagenesis of \textit{CrTHI4}_{rs}-\textit{Ble-GFP} reporter line allowed identification of aptamer residues responsible for the regulatory response of this riboswitch to its ligands. (a) Secondary structure of the \textit{CrTHI4} aptamer (predicted from the crystal structures of the TPP-binding aptamers of \textit{E. coli} and \textit{A. thaliana}\textsuperscript{10,11}). Coloured triangles indicate residues involved in: binding of the a and b phosphate group of TPP (light and dark red triangles), the pyrimidine group (purple triangles), and 3D folding of two sensor helices (blue triangles)\textsuperscript{10,11}. Black boxes show nucleotide substitutions generated by error-prone amplification (\textit{EPa}) site-directed mutagenesis of the \textit{CrTHI4}_{rs} mutation sites are indicated with \textit{EPa} identifiers. (b) Spot tests show the response of 3 independent lines with these variants grown on plates with 50 μg/ml zeocin only (No ligand) or supplemented with 10 μM Thi or HET. Lines tested carry the non-mutagenised \textit{CrTHI4}_{rs}-\textit{Ble-GFP} reporter construct, the \textit{EPa-riboswitch-Ble-GFP} variants (three independent transformants per construct) or the \textit{CrTHI4_noATG} construct that carries a mutated start codon (green arrow A-684U), which removes the native uORF. The images were taken 7 days post inoculation.

**Developing plug-and-play riboswitch devices**

The experiments above demonstrate that the \textit{CrTHI4}_{rs} uORF is essential for riboswitch regulation of the downstream gene, but its sequence composition and length are amenable to manipulation without impacting its responsiveness to thiamine. Moreover, given that the position of the uORF is not conserved among known TPP riboswitches that operate via this mechanism\textsuperscript{24}, we reasoned that the aptamer and uORF components of the riboswitch could be physically separated. This would be an important step in establishing a modular riboswitch design. We therefore introduced a new start codon at position -529 bp in the \textit{CrTHI4_noATG}_{rs}-\textit{Ble-GFP} construct, to move the uORF.
from the aptamer to the expression platform (Figure 3a). Transgenic lines containing this 
CrTHI4_4NRS-Ble-GFP construct were spot tested as before on media containing zeocin (Figure S4).
In the absence of any ligand the lines showed zeocin-resistance, whilst inclusion of thiamine or 
HET in the media resulted in reduced or no growth; HMP had no effect. This result confirmed 
recovery of the riboswitch response to ligands as a direct result of the re-introduction of the new 
uORF.

The modified CrTHI4_NRS riboswitch allowed us to test its flexibility further through aptamer 
swap experiments, replacing the CrTHI4RS aptamer with aptamers derived from other TPP 
riboswitches. Since splicing components are crucial for riboswitch function, we selected the well 
characterised aptamer domains of the Chlamydomonas CrTHICRS and A. thaliana THICRS, 
which contain similar alternative AG splice sites to that found in the CrTHI4RS aptamer. These 
were substituted for the aptamer in CrTHI4_4NRS (Figure 3b). The relationship to the splice sites was 
maintained, although it should be noted that the predicted 9 bp P1 stem of the CrTHI4_4NRS 
aptamer (Figure 2a) was reduced to 7 bp in the chimeric CrTHIC_4NRS and 3 bp in AtTHIC_4NRS. A 
进一步 modified riboswitch, CrEPa-50_4NRS, was generated in which the EPa-50 mutation was 
introduced into CrTHI4_4NRS.

The constructs with the riboswitch variants were transformed into Chlamydomonas, and 96 
independent zeocin-resistant colonies per construct were selected. These were inoculated into 
microtiter plates containing liquid TAP media with zeocin (10 µg/ml). This allowed a more 
quantitative measure of riboswitch activity in response to different ligands than the spot test 
method. Figure 3c shows the plates at 7 days post inoculation, and growth was quantified by 
measurement of optical density at 730 nm (Figure 3d). In the presence of thiamine and zeocin, 
growth of the majority of transformants for each construct was inhibited compared to zeocin 
alone, indicating that the modified riboswitches were functional. However, it is noteworthy that in 
zeocin-only media, the number of transformants carrying CrTHIC and AtTHIC aptamers that were 
resistant to the antibiotic was lower than lines carrying the CrTHI4_4N or the CrEPa-50_4N 
aptamers (Figure 3c), and the level of resistance was also less, with a median OD730 of ~0.75 for 
lines carrying CrTHIC and AtTHIC aptamers versus 0.9-1 for lines carrying CrTHI4_4N and CrEPa-
50_4N aptamers respectively (p-values ≤ 0.05) (Figure 3d). This might reflect inefficiency of splicing 
in the chimeric constructs. In testing the effect of the thiamine biosynthesis intermediates, 
CrTHIC_4NRS was responsive to HMP but not HET, whereas neither AtTHIC_4NRS nor CrEPa-
50_4NRS responded to the intermediates (Figure 3d; in pairwise analyses between zeocin-only and 
zeocin with HET or HMP, p-values ≥ 0.05). In A. thaliana, the response of the endogenous THIC 
gene to HET or HMP is unknown, but the behaviour of the chimeric CrTHIC_4NRS and AtTHIC_4NRS 
reflects the selectivity of each of these aptamers.

To capture the dynamics of the responses, a more detailed analysis over a 10-day time course 
was conducted with ~50 independent zeocin-resistant transformants per construct (Figure 3e). 
The consistency of the data demonstrates that, despite the fact that the transgenes are inserted
randomly in the Chlamydomonas nuclear genome, responses of the multiple transformants for each riboswitch variant are remarkably similar under each condition. In thiamine (green symbols), lines with aptamers from Chlamydomonas responded after 2 days, a lag that is likely due to the time to switch off the expression of the *Ble-GFP* gene and/or turnover of pre-existing Ble-GFP protein. The response of the *AtTHIC_4N<sub>RS</sub>* variant appears further delayed to day 3 and is less pronounced. The effect of HET (blue symbols) on *CrTHI4_4N<sub>RS</sub>* is clearly less than that of thiamine, and indeed is reminiscent of the thiamine/*AtTHIC_4N<sub>RS</sub>* behaviour. The slight effect of HET on *CrEPa-5O_4N<sub>RS</sub>* seen in Figure 3d, is more clearly shown in the time-course where there is a reduction in growth from day 2. HMP (red symbols) is just as effective as thiamine in switching off *CrTHIC_4N<sub>RS</sub>* but has no effect on any of the other aptamers.
Figure 3. Design and characterisation of plug and play riboswitches
Mutational analysis and development of a chimeric riboswitch expression platform (a) Schematic of the CrTHI4 5’UTR with highlighted mutations to test the impact of altered positioning of the uORF (blue box) within the CrTHI4 expression platform. Locations of uORF and splice sites are numbered according to the start codon of the Ble-GFP reporter gene at position 1. Riboswitch aptamer is shown as grey box and CrTHI4 5’UTR as black boxes. Mutations integral to the design of the chimeric riboswitches are shown above the boxes and include UUG (mutated native start codon A-684U), new uORF AUG (-532) and new uORF stop codon (-412). In the absence of thiamine, the splice site GU and AG are employed, while in abundance of thiamine, the gu/AG alternative splice sites are utilized. (b) Secondary RNA structures are those predicted for aptamers used in the different synthetic riboswitches, including the CrEpA-50, CrTHIC and AtTHIC aptamers. Black dots represent conserved nucleotides in all aptamers and the purple circle in the CrEpA-50 aptamer indicates the specific point mutation. Following transformation of UVM4 with the four different riboswitch-reporter variants, 96 zeocin resistant colonies were inoculated into four microtitre plates containing liquid media with zeocin (10 μg/ml). (c) After 3 rounds of weekly subculturing, response of transformants carrying the riboswitch chimeras was tested in liquid media supplemented with zeocin (10 μg/ml) and 10 μM of thiamine, HET or HMP. Microtitre plates were scanned 7 days post-inoculation. (d) To quantify response of the different lines to the different ligands, growth of cells was determined by measuring OD_{730} 7 days post inoculation. After filtering empty wells (that contained unstable transformants which died during the weekly subcultures), plate reading data was used to generate box and whisker plots that summarise the growth behaviour of the transformants in different conditions. CrTHI4_4Nrs, in which a new 41 amino acid uORF (blue box) was introduced outside of the aptamer and within the expression platform region resulted in responsive CrTHI4_4Nrs-Ble-GFP strains to both THI and HET. CrTHIC_4Nrs-Ble-GFP, AtTHIC_4Nrs-Ble-GFP and CrEpA-50_4Nrs-Ble-GFP lines only respond to thiamine. Letters above error bars indicate statistical groupings provided by Tukey’s test. (e) To determine response of transformants to ligands over time, growth of cells was also determined by measuring OD_{730} daily for 10 days post inoculation. Error bars indicate standard error of the mean for >50 biological replicates.

Recognition and responsiveness of native and synthetic riboswitch devices to natural ligand analogues
To allow more predictability for the behaviour of different ligands and riboswitch variants, we established a procedure to determine the ligand concentration required for 50% response (Effective Dose 50, ED50), and the maximum repression (MaxR) score, the reduction in growth of ligand-treated relative to control samples. Cells carrying each of the riboswitch variants were inoculated into microtitre plates containing growth media and different concentrations of the ligands at a known cell density (5 x 10^4 cells/ml), which was measured again after 5 days. By this approach, the ED50 of the native CrTHI4rs for thiamine was found to be 15.5 nM, and 22.7 nM for HET (Supplemental Figure S5; Table 1), whereas, as expected, no response was seen for HMP. MaxR for CrTHI4rs was calculated to be 0.75 for thiamine and 0.41 for HET. The ED50 and MaxR for the other riboswitch variants with the natural ligands confirmed the previous qualitative observations (Table 1). Of note is that CrTHI4_4Nrs showed the greatest MaxR of any riboswitch (0.98), and this was coupled with tighter binding for thiamine than the native CrTHI4rs. The ED50 for CrTHIC_4Nrs and AtTHIC_4Nrs was 20 times and 10 times higher than the CrTHI4_4Nrs aptamer, respectively.

Previous studies that probed TPP riboswitch-ligand interactions concluded that these riboswitches have evolved to bind different ligands with varying affinities and specificities. To determine the selectivity of the riboswitch variants, we tested their response to five chemical analogues shown in Figure S1: azide HMP (compound 5), 2-hydroxymethylpyrimidine (2HXMP, compound 6), triazole analogue (compound 4), 5-methylthiazole (compound 7) and 4,5-dimethylthiazole (compound 8). These analogues share certain chemical features with the natural ligands, but do not participate in the pathway. Prior to testing the riboswitch-reporter lines, we asked whether these analogues would induce TPP-like changes in the expression of the functional THI4s transcript (Figure S2) in untransformed Chlamydomonas cells. Supplementation of the
media with triazole analogue (compound 4) and azide-HMP (compound 5) resulted in the repression of CrTHI4s transcript levels, comparable to thiamine, whereas 2HXMP (compound 6), 5-methylthiazole (compound 7) and 4,5-dimethylthiazole (compound 8) had no effect (Figure S2). When lines containing any of the five riboswitch variants were screened for response to the analogues, none responded to these latter three compounds (data not shown). However, the triazole analogue (compound 4) repressed the CrTHI4_RS and CrTHI4_4N_RS, with an ED50 value of 5.5 µM and 2.2 µM, and MaxR scores of 1 and 0.8, respectively (Table 1). Although the ED50 values are ~400 times higher than for thiamine, the MaxR scores highlight the ability of this analogue to elicit a strong response by THI4_RS aptamers. ED50 values for the THIC_RS aptamers could not be assessed since the values were outside the 10 µM upper threshold set for this analysis. For azide-HMP (compound 5) a response was seen for all the riboswitches except CrTHIC_4N_RS (Table 1). ED50s ranged from 0.8-5 µM (Table 1), values that are 70-150 times the ED50 of these riboswitches with thiamine, respectively. As shown by the MaxR scores, the ability of azide-HMP to achieve complete response is similar to that of thiamine.

Table 1. Response of different riboswitch-Ble-GFP reporter variants to ligands
Maximum repression (MaxR) was determined for different riboswitch lines with different ligands. The OD730 of the compound treatments was normalised to the no-compound treatment for each line, compound and plate. These data were fitted to a 4-parameter logistic equation, and from that model the maximum repression was determined as 1 – lower asymptote of the model. This corresponds to the maximum predicted reduction in growth of the compound treated line relative to the no-compound treated line. Concentrations of ligands (nM) needed for half maximal growth repression (Effective Dose 50 or ED50) were calculated as described in the methods. Briefly, growth of the respective lines treated with different concentrations of ligands was measured by OD730 and normalised to a no-treatment control. For each compound the data were fitted to a four-parameter log-logistic model (and shown in Supplemental Figure S5) and presented with 95% confidence intervals (n=3). Riboswitch-treatment combinations that resulted in an ED50 greater than 6,000 nM were labelled as non-responsive (nr) as not enough data points were recorded above 10,000 nM to have confidence in this estimate; treatment combinations that were not examined for this table, but which have been tested and shown elsewhere (Figure 3) are labelled as not determined (nd).

| Lines               | Thiamine | HET       | HMP       | Triazole analogue | Azide-HMP  |
|---------------------|----------|-----------|-----------|-------------------|------------|
|                     | MaxR     | ED50      | MaxR      | ED50              | MaxR       | ED50      |
| CrTHI4_RS-Ble-GFP   | 0.75     | 15.52     | 0.41      | 22.7              | 1          | 5.5 x 10^3 |
|                    |          |           |           |                   | 0.86       | 1.1 x 10^3 |
| CrTHI4_4N_RS-Ble-GFP| 0.98     | 5.48      | 0.60      | 19.48             | 0.8        | 2.2 x 10^3 |
|                    |          |           |           |                   | 0.86       | 0.8 x 10^3 |
| CrTHIC_4N_RS-Ble-GFP| 0.40     | 106.96    | nd        | nd                | nd         | nd        |
|                    |          |           |           |                   | nd         | nd        |
| AtTHIC_4N_RS-Ble-GFP| 0.50     | 64.25     | nd        | nd                | nd         | nd        |
|                    |          |           |           |                   | nd         | nd        |
| EPa-50_4N_RS-Ble-GFP| 0.62     | 50.19     | nd        | nd                | nd         | nd        |
|                    |          |           |           |                   | nd         | nd        |

Exploiting the riboswitches to regulate transgenes
To begin to explore the versatility of the different riboswitches for synthetic biology purposes, we introduced two riboswitch-regulated transgenes into the same cell, firstly using fluorescent protein reporters. CrTHI4_4N_RS was used to regulate production of a nuclear localised Ble-GFP
fusion protein, and CrEPa-50_4N_RS was adopted to regulate Venus fluorescent protein targeted to the chloroplast (cpVenus). The cassette encoding both CrTHI4_4N_RS-Ble-GFP and CrEPa-50_4N_RS-cpVenus (Figure 4a) was transformed into Chlamydomonas. Figure 4b shows the result from one transgenic line expressing both reporters. In the absence of ligands, both GFP fluorescence in the nucleus and Venus fluorescence in the chloroplast were apparent (top row). The GFP and Venus signals were both repressed when the cells were grown with 10 µM thiamine (middle row), but only GFP was repressed by HET (bottom row).

![Diagram](image)

**Figure 4. Simultaneous regulation of differentially subcellular-targeted proteins by two different riboswitch variants**

(a) Schematic of construct containing a nucleus-targeted Ble-GFP fusion regulated by CrTHI4_4N_RS and a chloroplast-targeted Venus fluorescent protein (cpVenus) regulated by CrEPa-50_4N_RS used to transform Chlamydomonas. Both reporter genes contain the RBCS2 intron 1 (i1, grey box). Chloroplast targeting of cpVenus was achieved by addition of the PSAD chloroplast target peptide immediately after the start codon (shown as blue box).

(b) Confocal microscopy images of Chlamydomonas transformant encoding the above construct, grown in liquid TAP media only (Control) or TAP media supplemented with 10 µM thiamine or 10 µM HET. t = 7 days post inoculation. Scale bar represents 50 µm and the insets are 13.5 µm wide.
Reciprocal regulation was further tested by introducing a cassette encoding \( \text{CrTHI}_4 \text{N}_{RS} \) regulating expression of Venus, this time targeted to the cytosol, and Ble-GFP regulated by \( \text{CrTHIC}_4 \text{N}_{RS} \) (Figure S6a). A Chlamydomonas transformant carrying both reporter genes showed expected Venus fluorescence in the cytosol and GFP fluorescence in the nucleus (Figure S6b, top row, with no-ligand). The weak fluorescence signal for GFP, which we observed with multiple transformants, is likely the result of the reduced transgene expression with the \( \text{CrTHIC} \) aptamer, also seen in growth assays described earlier (Figure 3d). When grown in media containing thiamine, both Venus and GFP signals were repressed, an expected outcome that indicates complete response of both aptamers to this ligand (Figure S6b, second row, +Thi). Nuclear-localised GFP was detectable in the transformant cultured in media with no-ligand or with HET (Figure S6b, third row, +HET), showing that \( \text{CrTHIC}_4 \text{N}_{RS} \) did not respond to HET. In addition, HET had little or no effect on \( \text{CrTHI}_4 \text{N}_{RS} \)-regulated Venus. This result is consistent with our earlier observations that this riboswitch is less responsive to HET than thiamine (Figure 3c & d and Table 1). In contrast, when cells were grown in HMP, there was no detectable GFP regulated by the \( \text{CrTHIC}_4 \text{N} \) aptamer demonstrating that the specificity of the aptamer was retained (Figure S6b, bottom row, +HMP). The results in this section indicate that the combinations of the four riboswitch parts and different ligands can enable a gradient of gene regulation capability.

To demonstrate the utility of riboswitch parts as modular devices that can be incorporated in metabolic pathway engineering strategies, we used the \( \text{CrTHI}_4 \text{N}_{RS} \) to control the heterologous production of the bicyclic diterpene, casbene. This molecule is the first committed intermediate necessary for production of medicinal diterpenoids such as jolkinol-C and ingenol-3-angelate\(^{27,28}\). Casbene synthase (CBS) cyclises the prenylated intermediate geranylgeranyl pyrophosphate (GGPP) to produce casbene. A previous study showed that heterologous expression of \( \text{Ricinus communis} \) CBS in Chlamydomonas leads to production of casbene\(^{29}\), which is released from the cells and can be captured in a dodecane overlay. For our experiment, the \( \text{Jatropha curcas} \) CBS gene (\( \text{JcCBS} \)) was codon- and intron-optimised as previously described for enhanced expression in Chlamydomonas\(^{31}\). A construct was assembled in which the chloroplast targeted JcCBS protein was fused at C-terminal with GS linker peptide and a Venus fluorescent protein (Figure 5a). For riboswitch regulation, the \( \text{CrTHI}_4 \text{N}_{RS} \) was cloned between the promoter and chloroplast target peptide region of the cassette. This cassette was transformed into the Chlamydomonas nuclear genome followed by antibiotic selection, and transformants that expressed the Venus fluorescent protein were identified through fluorescence analysis in microtitre plates\(^{32}\). A representative Venus-expressing transformant was cultured in TAP media with a 10% n-dodecane overlay. Nine days post-inoculation, the overlay was mixed 1:1 with hexane and analysed using gas chromatography mass spectrometry (GC-MS). Casbene captured by the n-dodecane overlay was detected at the expected retention time, thereby confirming that the casbene synthase enzyme fused to Venus fluorescent protein was functional (Figure 5, black line). To determine whether the \( \text{CrTHI}_4 \text{N}_{RS} \) could regulate casbene production, the casbene producing transformant was cultivated in TAP media supplemented with 10 \( \mu \)M of thiamine. GC-MS analysis of the dodecane
overlay from 9-days old cultures showed no detectable casbene (Figure 5b, green line), thus confirming the capacity of the CrTHI4_4Nrs to fully suppress CBS gene expression when exogenous thiamine is added to the growth media. The tight regulation of a heterologously expressed biosynthesis enzyme by one of the riboswitch variants described in this work, serves to highlight the utility of riboswitches in design and implementation of a nuanced regulatory framework, which can be important for metabolic engineering.

*Figure 5. Riboswitch regulation of casbene production*

Casbene production was achieved by introducing the casbene synthase expression cassette into the nuclear genome of the Chlamydomonas UVM4 strain. (a) Schematic of the expression cassette shows the contributing parts assembled in the following order, HSP70/RBCS2 promoter, 22 nt RBCS2 5’UTR, CrTHI4-4Nrs, PSAD chloroplast target peptide (cTP shown as blue box), codon optimized casbene synthase (CBS) containing multiple copies of the Chlamydomonas RBCS2 intron 1 (i1) fused with a GS linker peptide (orange box) to Venus containing RBCS2 intron 2(i2), the 3’ UTR was that derived from CA1. (b) Casbene production in a Chlamydomonas transformant that expressed the casbene synthase transgene was assessed using Gas Chromatography Mass Spectrometry (GC-MS). A representative transformant was cultured in TAP media with a 10% n-dodecane overlay. Nine days post-inoculation, the overlay was analysed by GC-MS. Casbene captured by the n-dodecane overlay was detected at the expected retention time (black trace), thereby confirming that the casbene synthase enzyme fused to Venus fluorescent protein was functional. The GC-MS ion chromatogram (m/z 121) shows metabolites carrying a mass-to-charge ratio (m/z) of 121 +/- 0.5. Internal standard β-caryophyllene was detected at 12.32 min retention time (RT), while casbene was detected at 23.16 min RT. In agreement with previous reports, a selection of oxidised casbene molecules was also detectable between RT 25.6 and 26.7 RT.$^{30}$ The detection of casbene and its oxidised derivatives demonstrates capacity of this transformant to produce casbene. When this transformant was cultured in media containing 10 μM thiamine (green trace), casbene was not detected, a result that highlighted the utility of the riboswitch for regulation of transgene expression.
Discussion

Here we have dissected a complex eukaryotic genetic element that is the Chlamydomonas THI4 TPP riboswitch and demonstrated its versatility in regulation of transgenes. Our analyses captured attributes of the riboswitch that enable its modularity, namely that both the aptamer and the expression platform were able to tolerate specific modifications, which in turn allowed their initial overlap to be disentangled (Figure 2). This knowledge allowed the generation of a new riboswitch platform that is adaptable for aptamer-swap experiments (Figure 3a) and use of a high-throughput method for quantitative assessment of the riboswitch/ligand combinations revealed the dynamic response potential for the riboswitch variants to different ligands (Table 1). These data provide insight into molecular aspects of riboswitch regulation including base-pairings that impact ligand specificities. Of note is that this in vivo analysis approach has the advantage of moving beyond characterisation of riboswitches in isolation, instead analysing their performance within the complexity of the cellular metabolic milieu.

The modularity of riboswitches provides synthetic biology with a unique ‘plug-and-play’ quality at the level of secondary structures and functional units. Furthermore, the ratio of ligand to riboswitch can dictate the level of regulation, providing the potential for tuneable progression from ON to OFF. We verified the utility of these regulatory elements as part of multi-gene cassettes, whereby two different riboswitch-reporters were shown to be amenable to independent regulation within the same cell. Given the unique sensitivity of each riboswitch to different ligands, this experimental design enabled different chemical inputs (e.g. thiamine, HET or HMP) to achieve predictable output signals (e.g. differential regulation of the two fluorescent proteins Figure 4 and Figure S6). These results highlight the potential of the different riboswitches and how their ligand-response characteristics can be used as part of synthetic gene circuit designs. It is noteworthy that although the TPP riboswitches described here function as OFF switches, they have the potential to be incorporated into a double-switch series or repressor-of-repressor platform ultimately to achieve ON status.

In terms of the wider applications of TPP riboswitches, we propose their importance as tools for metabolic engineering and commercial applications. This is because thiamine, unlike many often-used chemical modifiers, is a naturally occurring vitamin that is non-toxic. The low concentration (nM to µM) of thiamine, HET and HMP required to regulate expression, combined with their relatively low-cost commercial availability and the ability to achieve near total repression (MaxR of CrTHI4_4N_RS for thiamine is 0.98, Table 1) make these compounds extremely attractive for large-scale low-cost continuous or semi-continuous commercial processes. As we look to develop microalgae further as industrial biotechnology platforms, tools such as the riboswitches described in this work take on a greater significance. This is because heterologously expressed proteins or compounds can cause toxicity and lead to undesirable allocation of metabolic substrates or reducing power when the cells are growing rapidly. In this study we showed the ease by which the
TPP riboswitch can be incorporated into the design and engineering of a Chlamydomonas strain that heterologously expresses the medically important diterpenoid, casbene (Figure 5). This is the first time that a riboswitch has been used to regulate the biosynthesis of any terpenoid in any species. This result highlights an important application for riboswitches in Chlamydomonas and potentially other eukaryotes. By adopting the different riboswitch-ligand combinations described here, metabolic engineers can set out to overcome key challenges that face reconstitution of complex terpenoid biosynthesis pathways in heterologous hosts, allowing optimisation of the expression of different pathway steps and balancing pull-push effects on the cell redox potential that ultimately determine terpenoid yields\textsuperscript{35,36}.

Insights from this study will allow the extension of the collection of RNA switches from a wide range of bacterial riboswitch aptamers or computational synthetic aptamers. Moreover, as we improve our understanding of different riboswitches in algae and higher plants (this study and\textsuperscript{37}), we will shed light on how riboswitches have evolved in eukaryotes and their important presence in the modern protein-based systems. Although our current understanding of RNA regulatory systems is limited to TPP riboswitches, yet to be identified RNA regulatory systems and RNA-ligand interactions will likely shape our view of the evolution of these ancient RNA regulators.
Methods

Unless otherwise specified all chemicals used in this study were sourced from Sigma-Aldrich.

Chlamydomonas strains and cultivation
Chlamydomonas strain UVM4\(^{19}\) was used for all experiments in this work. Unless otherwise noted, strains were cultivated in Tris acetate phosphate (TAP) medium\(^{38}\) in liquid culture or on agar plates (1.5% w/v), with 100–250 µmol photons m\(^{-2}\) s\(^{-1}\) at 25°C. All transformants were maintained on TAP plates with respective antibiotics (10 mg l\(^{-1}\) paromomycin, 20 mg l\(^{-1}\) hygromycin, and/or 5, 10 or 50 mg l\(^{-1}\) zeocin [Invivogen]), all liquid cultivations were performed in shake flasks or microtiter plates using TAP. For response assays thiamine (Melford Laboratories Ltd.), 4-methyl-5-(2-hydroxyethyl) thiazole (HET), 4-amino-5-hydroxymethyl-2-methylpyrimidine (HMP, Fluorochem UK) and their analogues (shown in Supplemental Table S1) were added to TAP media at the indicated concentrations.

Chlamydomonas transformation and screening by fluorescence plate reader
Transformation by electroporation was carried out as previously described\(^{16}\) with the following modifications. A TAP culture of 1–5 × 10\(^{6}\) cells/ml was concentrated 100 times in TAP containing 40 mM sucrose and 250 µl was incubated with 1 µg of linearized DNA for 10 min on ice in a 0.4 cm gapped cuvette (BioRad) prior to electroporation (BioRad Gene Pulser Xcell). The cells recovered in TAP containing 40 mM sucrose for 16 h in a shaking incubator (less than 10 µmol photon m\(^{-2}\) s\(^{-1}\) at 100 rpm) prior to plating on TAP-agar plates with the appropriate antibiotics. Screening for robust expression of fluorescent protein fusion was conducted using the ClarioStar fluorescence plate reader (BMG Labtech) as previously described\(^{29}\). Briefly, Venus fluorescence readings were obtained using the filters for excitation 515/10 nm and emission 550/10 nm. TAP medium was used as a blank. Fluorescence signals were normalized to chlorophyll fluorescence (excitation 440/9 nm, emission 680/20 nm).

Generation of constructs
All in silico sequence designs and analysis were performed with SnapGene\textsuperscript® software (GSL Biotech). All PCR reactions were performed using the Q5\textsuperscript® DNA polymerase (New England Biolabs [NEB]) following the manufacturer’s instructions for use with GC-rich DNA and using the GC enhancer solution. Primers were produced by Sigma-Aldrich, while synthesised parts were obtained from Integrated DNA Technologies, with the exception of codon-optimised CBS that was synthesised by GenScript Corporation (USA). \textit{noATG-CrTHI4RS, CrTHI4_4NRS, CrTHIC_4NRS and AtTHIC_4NRS} parts were amplified from inhouse plasmid templates generated and provided by the Smith laboratory (University of Cambridge) and assembled by Gibson assembly\(^{39}\) using the isothermal method, at 50 °C for 1 h, each part having been amplified with respective primers shown in Supplemental Table S5. Error prone mutagenesis constructs were generated as detailed under the heading ‘Error prone mutagenesis’ using MEGAWHOP cloning\(^{40}\). All other
constructs were assembled by GG cloning according to the MoClo system\textsuperscript{16,41,42}. PCR products were purified with the Monarch\textsuperscript{®} PCR & DNA Clean up kit (NEB) and cloned into corresponding vectors by GG-cloning using either BsaI (NEB) or BpiI (Thermo Fisher Scientific) depending on the level of the destination MoClo vectors, and T4 Ligase (Thermo Fisher Scientific). All DNA parts used in this study are described in Supplemental Tables S1 to S3. Level 0 GG-parts coding for pPM0-037 (CrTHI4\_4N\_RS), pPM0-038 (CrTHIC\_4N\_RS), pPM0-039 (AtTHIC\_4N\_RS) and pPM0-040 (Epa-50\_4N\_RS) were synthesised. Parts sourced from the Chlamydomonas MoClo Kit\textsuperscript{16} are indicated in Supplemental Table S1 to S3. Level 0 GG-parts, as well as Level 1 and Level 2 assembled plasmids, were transformed into E. coli (NEB 5-alpha competent cells) by heat shock, then transformants were plated on LB agar plates with selective additions as required, using the following concentrations: spectinomycin and ampicillin (50 μg/ml), carbenicillin (100 μg/ml), X-gal (40 μg/ml). Plasmids were isolated with the Monarch Plasmid\textsuperscript{®} preparation kit (NEB) and analysed by sequencing (generation of GG-modules, Gibson and MEGAWHOP-derived plasmids were confirmed by sequencing (Source BioSciences UK), while all other plasmids were verified by differential restriction.

The amino acid sequences of the Jatropha curcas casbene synthase (JcCBS, NCBI Reference Sequence: NP\_001292945.1) was codon optimized for expression from the Chlamydomonas nuclear genome and synthesised by GenScript Corporation (USA). To enhance transgene expression of large codon optimized constructs from the nuclear genome of Chlamydomonas, JcCBS was designed to contain multiple copies of the first intron of the Chlamydomonas ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) small subunit 2 (RBCS2i1, NCBI: X04472.1)\textsuperscript{31}. The PSAD chloroplast target peptide was cloned at the N-terminus of the synthesised JcCBS sequence, while a GS linker and Venus fluorescent protein were added to the C-terminus of the JcCBS, thereby allowing the fluorescence plate reader-based screening for protein abundance in transformants. The fully assembled Level 1 and Level2 constructs carrying the JcCBS transgene are listed in Supplemental Table S2 and S3 as pPM1-070 and pPM2-051, respectively.

**Error prone mutagenesis**

PCR-based error prone MnCl\textsubscript{2} mutagenesis was used to target the 171 bp aptamer of the CrTHI4 riboswitch as previously described\textsuperscript{43}. Three parts of the THI4 riboswitch were amplified separately with at least 25bp flanking ends that were homologous to the adjoining parts (Supplemental Figure S7) using the template plasmid promRBCS2::THI4 5\textsuperscript{-}UTR-Ble\textsuperscript{13}. Primers used are listed in Supplemental Table S5 and assembly conditions are as previously described\textsuperscript{13}. The EP1 and EP3 regions were amplified by normal PCR reactions whilst the PCR reactions of EP2 containing the aptamer region were supplemented with increasing concentrations of MnCl\textsubscript{2} so as to increase the possibility of generating mutations in the PCR products. As each plasmid would contain different mutations, 25 E. coli plasmids were extracted to identify those having mutations in the CrTHI4 riboswitch aptamer.
Random mutagenesis using UV light
Different concentrations of cells in exponential phase were plated on TAP agar plates supplemented with appropriate selection (metabolites and/or antibiotics). Petri dishes were exposed under the UV light in a flow hood for exactly 60 sec at 25°C at a distance of 38 cm from the UV lamp emitting radiation at 253.7 nm. Plates were wrapped in aluminium foil immediately to minimise light-induced cellular repair mechanisms and incubated at 25 °C for 2 weeks. The number of visible colonies was counted with the desired dose of less than 1% survival. Some 16 colonies were picked and re-streaked on 50 mg l⁻¹ zeocin and 10 μM thiamine supplemented media to determine the impact of UV-induced mutagenesis on responsiveness of the riboswitch.

Determination of the riboswitch response to different ligands
To quantify the response to different ligands of lines containing the Ble-GFP-reporter under control of riboswitch variants growth in zeocin was used as a proxy. Cells were inoculated in 96-well microtiter plates in a total volume of 200 µl and incubated under constant light (125 μmol photon m⁻² s⁻¹) at 25 °C. After 2 days the cultures were resuspended, and after a further 2 days the optical density at 730 nm was recorded using a FLUOstar OPTIMA plate reader (BMG Labtech). Plates were shaken for 6–10 s at 600 rpm before measurement. Blanks that were anomalous, for example if they contained algae resulting from pipetting inaccuracies, were excluded from subsequent analysis. Samples that were not treated with thiamine or related compounds were included on every plate and used as positive controls to ensure that the strain was growing as expected. If any no-compound samples for a given strain were growing in an incorrect manner, all data for the affected strain on the affected plate were excluded from the subsequent analysis. For each 96-well plate, the mean optical density of the blanks for that plate was subtracted from the optical density of the samples. For all compound treatments on a plate, the optical density was divided by the mean optical density of each no-compound treated strain, providing a value of the growth relative to no-treatment. This allowed for more accurate comparisons between plates that may have experienced slightly different environmental conditions.

Strain and treatment combinations were set up in technical triplicate. The mean of technical triplicates was calculated and used for modelling. Models of the data were made, fitting to a 4-parameter logistic equation previously described⁴⁴. Different metrics were derived from the models: maximum predicted repression (MaxR) and Effective Dose 50 (ED50). The maximum repression was defined as 1 - lower asymptote of logistic equation. This corresponds to the maximum predicted reduction in growth of the compound treated strain relative to the no-compound treated strain, expressed as a percentage. The ED50 was defined as the concentration of ligand that results in half maximal repression. Any strain-treatment combinations that resulted in an ED50 greater than 6,000 nM were labelled as NR (non-responsive) as not enough data points were recorded above 10,000 nM to have confidence in this estimate. In Table 1, CrTHI4ᵣₛ₅-Ble-GFP and CrTHI4ᵣ₄Nᵣₛ₅-Ble-GFP when treated with HMP were also recorded as NR due to observations outside this analysis. All analysis was performed in R.
**Statistical analyses**

Algal growth data were recorded at 730 nm over time. Statistical analysis was performed at t=7, first by performing an ANOVA to determine whether there were differences between strain and treatment combinations. As this was significant, Tukey’s test was performed to determine which combinations were different from one another. qPCR relative expression data were assessed for significant difference to the No Ligand treatment. ANOVA was first performed to determine if there were any significant treatment comparisons. If so, Dunnett’s test was performed using the multcomp package to identify treatments that differed significantly from the No Ligand treatment. A P-value of < 0.05 was considered statistically significant.

**Confocal laser scanning microscopy**

Chlamydomonas transformants carrying the pPM2-034 and pPM2-035 constructs (Supplemental Table S3) were grown for 7 days in TAP media without supplementation, or in the presence of 10 μM thiamine or 10 μM HET, before visualization in a confocal laser scanning microscope (TCS SP8, Leica Microsystems, Germany). The images were taken in the sequential mode enabled by the Leica LAS software. The first image was acquired with excitation from a white light source at 476 nm and emission detection between 485 and 518 nm for the GFP channel, and between 650 nm and 720 nm for the chlorophyll channel. The second image was acquired with excitation at 516 nm and emission detection between 526 and 592 nm for the Venus channel.

**Quantitative real-time PCR**

Total cellular RNA was extracted from 1.0 × 10^7 Chlamydomonas cells using the RNeasy plant mini kit (Qiagen) according to the manufacturer’s instructions. Contaminating genomic DNA was removed by treating RNA samples with TURBO DNase (Thermo Fisher Scientific) following the manufacturer’s instructions. RNA quality and concentration were measured using a Nanodrop (Thermo Fisher Scientific) spectrometer at 260 nm. Structural integrity of the RNA was checked with non-denaturing agarose gel and ethidium bromide staining. Reverse transcription was initiated with an oligo dT<sub>18</sub> primer with 1 μg total RNA as template and the reverse-transcriptase SuperScript III (Thermo Fisher Scientific), according to the manufacturer’s instructions. The resulting cDNAs were diluted 10-fold before use. Each reaction contained cDNAs corresponding to 100 ng total RNA, 0.25 μM each of forward and reverse primers, 2x SYBR<sup>®</sup> Green JumpStart<sup>™</sup> Taq ReadyMix<sup>™</sup> with Mg<sup>2+</sup> (Sigma-Aldrich), in a total volume of 10 μL. The following program was used: 95 °C for 5 min, followed by 40 cycles of 95 °C for 15 s, 55 °C for 60 s, in a Rotor-Gene Q real-time PCR machine (Qiagen) with fluorescence measurement after each 65 °C step; a melting curve analysis was performed between 65 and 95 °C to confirm specific amplification. Transcript abundances were calculated using the ΔΔCT method, in relation to the housekeeping genes RACK1. Primer sequences are listed in Supplemental Table S4.
Metabolite analysis using GC-MS
The capacity of Chlamydomonas transformants carrying the pPM2-051 expression cassette to produce casbene was assessed by cultivation of individual strains in 50 ml Nunc flasks (ThermoFisher Scientific) with a 10% dodecane overlay, with or without 10 μM thiamine. Nine days post inoculation, the dodecane overlay was separated from the culture, clarified by centrifugation and mixed 1:1 with hexane containing 10 μg/ml of the internal standard α-caryophyllene. One microlitre of this mixture was injected and analysed by GC-MS as previously described. Caryophyllene and casbene were determined with the mass range (m/z) of 120.5-121.5.

Supporting Information:

- Supplementary Table S1-S3 construct design and details of assembled parts (XLSX)
- Supplementary Figures S1-S7 and Supplementary Tables 4 and 5 (PDF)
- Annotated sequences of plasmids (ZIP)
Contributions

PM conceived part of the project, designed experiments, obtained and analysed data and wrote the manuscript
GNDTN conceived part of the project, designed experiments, obtained and analysed data, and commented on the manuscript
AGR obtained and analysed data, and contributed to the manuscript
AS obtained and analysed data, and contributed to the manuscript
MLP obtained and analysed data, and contributed to the manuscript
MLHS obtained and analysed data
EJM obtained and analysed data
GIMO commented on the manuscript
MAS obtained and analysed data
AGS conceived the project, obtained the funding, analysed the data and wrote the manuscript

Conflict of Interest

The authors declare no conflict of interests.

Acknowledgments

Financial support is acknowledged from the Biotechnology and Biological Sciences Research Council (BBSRC) of the UK for research grants (BB/I007660/1; MS & AGS), (BB/L014130/1; GIMO & AGS) and PhD studentships (AGR, AS, MLP), from the EU FP7 (Grant number 311956; MS & AGS), and from Fundacio Bancaria La Caixa, Spain (MLP). In addition, a part of this project has been funded by the Industrial Biotechnology Catalyst (Innovate UK, BBSRC, EPSRC) to support the translation, development and commercialisation of innovative Industrial Biotechnology processes (BB/M018210/01; PM, AGS). PM thanks HVCfP BBSRC NIBB for a training grant to analyse casbene from samples. We are grateful for laboratory support from Ms Sue Aspinall, Susan Stanley, Beata Sikora, Dr Fiona Taylor, Ms Catherine Sutherland and useful discussions with colleagues Dr Katrin Geisler, Dr Mathew Davey, Prof Saul Purton (UCL) and Dr Tomasz Czechowski and Prof Ian Graham (University of York). The coding sequence for the codon-optimised casbene synthase was synthesised by GenScript for GSK. C. reinhardtii strain UVM4 was obtained from Prof Dr Ralph Bock, MPI-MP, Golm, Germany.
References

(1) Saxena, P., Bojar, D., and Fussenegger, M. (2017) Design of synthetic promoters for gene circuits in mammalian cells, in Mammalian Synthetic Promoters. Methods in Molecular Biology, pp 263–273. Humana, New York, NY.
(2) Brückner, K., Schäfer, P., Weber, E., Grützner, R., Marillonnet, S., and Tissier, A. (2015) A library of synthetic transcription activator-like effector-activated promoters for coordinated orthogonal gene expression in plants. Plant J. 82, 707–716.
(3) Schreiber, T., and Tissier, A. (2018) Synthetic Transcription Activator-Like Effector-Activated Promoters for Coordinated Orthogonal Gene Expression in Plants, in Molecular Pharming, pp 25–42. John Wiley & Sons, Inc., Hoboken, NJ, USA.
(4) Etzel, M., and Mörl, M. (2017) Synthetic Riboswitches: From Plug and Pray toward Plug and Play. Biochemistry 56, 1181–1198.
(5) Win, M. N., and Smolke, C. D. (2007) A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. Proc. Natl. Acad. Sci. 104, 14283–14288.
(6) Chen, Z., Wilmanns, M., and Zeng, A. P. (2010) Structural synthetic biotechnology: From molecular structure to predictable design for industrial strain development. Trends Biotechnol. 28, 534–542.
(7) Yang, J., Seo, S. W., Jang, S., Shin, S.-I., Lim, C. H., Roh, T.-Y., and Jung, G. Y. (2013) Synthetic RNA devices to expedite the evolution of metabolite-producing microbes. Nat. Commun. 4, 1413.
(8) You, M., Litke, J. L., and Jaffrey, S. R. (2015) Imaging metabolite dynamics in living cells using a Spinach-based riboswitch. Proc. Natl. Acad. Sci. U. S. A. 112, E2756–E2765.
(9) Fowler, C. C., Brown, E. D., and Li, Y. (2010) Using a riboswitch sensor to examine coenzyme B12 metabolism and transport in E. coli. Chem. Biol. 17, 756–765.
(10) Serganov, A., Polonskaia, A., Phan, A. T., Breaker, R. R., and Patel, D. J. (2006) Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. Nature 441, 1167–1171.
(11) Thore, S., Leibundgut, M., and Ban, N. (2006) Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. Science 312, 1208–11.
(12) Croft, M. T., Moulin, M., Webb, M. E., and Smith, A. G. (2007) Thiamine biosynthesis in algae is regulated by riboswitches. Proc. Natl. Acad. Sci. U. S. A. 104, 20770–20775.
(13) Moulin, M., Nguyen, G. T. D. T., Scaife, M. A., Smith, A. G., and Fitzpatrick, T. B. (2013) Analysis of Chlamydomonas thiamin metabolism in vivo reveals riboswitch plasticity. Proc. Natl. Acad. Sci. U. S. A. 110, 14622–14627.
(14) Salomé, P. A., and Merchant, S. S. (2019) A series of fortunate events: Introducing chlamydomonas as a reference organism. Plant Cell 31, 1682–1707.
(15) Li, X., Patena, W., Fauser, F., Jinkerson, R. E., Saroussi, S., Meyer, M. T., Ivanova, N., Robertson, J. M., Yue, R., Zhang, R., Vilarrasa-Blasi, J., Wittkopp, T. M., Ramundo, S., Blum, S. R., Goh, A., Laudon, M., Srikumar, T., Lefebvre, P. A., Grossman, A. R., and Jonikas, M. C. (2019) A genome-wide algal mutant library and functional screen identifies genes required for eukaryotic photosynthesis. Nat. Genet. 51, 627–635.
(16) Crozet, P., Navarro, F. J., Willmund, F., Mehrshahi, P., Bakowski, K., Lauersen, K. J., Pérez-
Pérez, M. E., Auroy, P., Gorchs Rovira, A., Sauret-Gueto, S., Niemeyer, J., Spaniol, B., Theis, J., Trösch, R., Westrich, L. D., Vavitsas, K., Baier, T., Hübner, W., De Carpentier, F., Cassarini, M., Danon, A., Henri, J., Marchand, C. H., De Mia, M., Sarkissian, K., Baulcombe, D. C., Peltier, G., Crespo, J. L., Kruse, O., Jensen, P. E., Schröda, M., Smith, A. G., and Lemaire, S. D. (2018) Birth of a Photosynthetic Chassis: A MoClo Toolkit Enabling Synthetic Biology in the Microalga Chlamydomonas reinhardtii. *ACS Synth. Biol.* 7, 2074–2086.

(17) Lauersen, K. J. (2019) Eukaryotic microalgae as hosts for light-driven heterologous isoprenoid production. *Planta* 249, 155–180.

(18) Fuhrmann, M., Oertel, W., and Hegemann, P. (1999) A synthetic gene coding for the green fluorescent protein (GFP) is a versatile reporter in Chlamydomonas reinhardtii. *Plant J.* 19, 353–361.

(19) Neupert, J., Karcher, D., and Bock, R. (2009) Generation of Chlamydomonas strains that efficiently express nuclear transgenes. *Plant J.* 57, 1140–1150.

(20) Schröda, M., Blöcker, D., and Beck, C. F. (2000) The HSP70A promoter as a tool for the improved expression of transgenes in Chlamydomonas. *Plant J.* 21, 121–131.

(21) Cheah, M. T., Wachter, A., Sudarsan, N., and Breaker, R. R. (2007) Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature* 447, 497–500.

(22) Ontiveros-Palacios, N., Smith, A. M., Grundy, F. J., Soberon, M., Henkin, T. M., and Miranda-Ríos, J. (2008) Molecular basis of gene regulation by the THI-box riboswitch. *Mol. Microbiol.* 67, 793–803.

(23) Wachter, A., Tunc-Ozdemir, M., Grove, B. C., Green, P. J., Shintani, D. K., and Breaker, R. R. (2007) Riboswitch control of gene expression in plants by splicing and alternative 3′ end processing of mRNAs. *Plant Cell* 19, 3437–3450.

(24) Nguyen, G. T. D. T., Scaife, M. A., Helliwell, K. E., and Smith, A. G. (2016) Role of riboswitches in gene regulation and their potential for algal biotechnology. *J. Phycol.* 52, 320–328.

(25) Bocobza, S., Adato, A., Mandel, T., Shapira, M., Nudler, E., and Aharoni, A. (2007) Riboswitch-dependent gene regulation and its evolution in the plant kingdom. *Genes Dev.* 21, 2874–2879.

(26) Chen, L., Cressina, E., Dixon, N., Erixon, K., Agyei-Owusu, K., Micklefield, J., Smith, A. G., Abell, C., and Leeper, F. J. (2012) Probing riboswitch-ligand interactions using thiamine pyrophosphate analogues. *Org. Biomol. Chem.* 10, 5924–5931.

(27) Vasas, A., and Hohmann, J. (2014) Euphorbia Diterpenes: Isolation, Structure, Biological Activity, and Synthesis (2008–2012). *Chem. Rev.* 114, 8579–8612.

(28) King, A. J., Brown, G. D., Gilday, A. D., Larson, T. R., and Graham, I. A. (2014) Production of bioactive diterpenoids in the Euphorbiaceae depends on evolutionarily conserved gene clusters. *Plant Cell* 26, 3286–3298.

(29) Lauersen, K. J., Wichmann, J., Baier, T., Kampranis, S. C., Pateraki, I., Möller, B. L., and Kruse, O. (2018) Phototrophic production of heterologous diterpenoids and a hydroxy-functionalized derivative from Chlamydomonas reinhardtii. *Metab. Eng.* 49, 116–127.

(30) King, A. J., Brown, G. D., Gilday, A. D., Forestier, E., Larson, T. R., and Graham, I. A. (2016) A Cytochrome P450-Mediated Intramolecular Carbon–Carbon Ring Closure in the Biosynthesis of Multidrug-Resistance-Reversing Lathyrane Diterpenoids. *ChemBioChem* 17, 1593–1597.

(31) Baier, T., Wichmann, J., Kruse, O., and Lauersen, K. J. (2018) Intron-containing algal
transgenes mediate efficient recombinant gene expression in the green microalga Chlamydomonas reinhardtii. *Nucleic Acids Res.* **46**, 6909–6919.

(32) Rasala, B. A., Barrera, D. J., Ng, J., Plucinak, T. M., Rosenberg, J. N., Weeks, D. P., Oyler, G. A., Peterson, T. C., Haerizadeh, F., and Mayfield, S. P. (2013) Expanding the spectral palette of fluorescent proteins for the green microalga Chlamydomonas reinhardtii. *Plant J.* **74**, 545–556.

(33) Liang, J. C., Bloom, R. J., and Smolke, C. D. (2011) Engineering Biological Systems with Synthetic RNA Molecules. *Mol. Cell** **74**, 545–556.

(34) Amendola, M., Giustacchini, A., Gentner, B., and Naldini, L. (2013) A double-switch vector system positively regulates transgene expression by endogenous microRNA expression (miR-ON vector). *Mol. Ther.* **21**, 934–946.

(35) Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M. D., Tai, A., Main, A., Eng, D., Polichuk, D. R., Teoh, K. H., Reed, D. W., Treynor, T., Lenihan, J., Jiang, H., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K. W., Fickes, S., Galazzo, J., Gaucher, S. P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L. F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Liedense, J., Covello, P. S., Keasling, J. D., Reiling, K. K., Renninger, N. S., and Newman, J. D. (2013) High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**, 528–532.

(36) Renault, H., Bassard, J. E., Hamberger, B., and Werck-Reichhart, D. (2014) Cytochrome P450-mediated metabolic engineering: Current progress and future challenges. *Curr. Opin. Plant Biol.* **19**, 27–34.

(37) Verhounig, A., Karcher, D., and Bock, R. (2010) Inducible gene expression from the plastid genome by a synthetic riboswitch. *Proc. Natl. Acad. Sci.* **107**, 6204–6209.

(38) Kropat, J., Hong-Hermesdorf, A., Casero, D., Ent, P., Castruita, M., Pellegrini, M., Merchant, S. S., and Malasarn, D. (2011) A revised mineral nutrient supplement increases biomass and growth rate in Chlamydomonas reinhardtii. *Plant J.* **66**, 770–780.

(39) Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345.

(40) Miyazaki, K. (2011) MEGAWHOP cloning: A method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. *Methods Enzymol.* 1st ed. Elsevier Inc.

(41) Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011) A Modular Cloning System for Standardized Assembly of Multigene Constructs. *PLoS One* (Peccoud, J., Ed.) **6**, e16765.

(42) Engler, C., Youles, M., Gruetzner, R., Ehnert, T. M., Werner, S., Jones, J. D. G., Patron, N. J., and Marillonnet, S. (2014) A Golden Gate modular cloning toolbox for plants. *ACS Synth. Biol.* **3**, 839–843.

(43) Miyazaki K, T. M. (2002) METHODS Short Technical Reports. *Biotechniques* **33**, 1036–1038.

(44) Ritz, C., Baty, F., Streibig, J. C., and Gerhard, D. (2015) Dose-response analysis using R. *PLoS One* **10**, 1–13.

(45) Hothorn, T., Bretz, F., and Westfall, P. (2008) Simultaneous inference in general parametric models. *Biometrical J.* **50**, 346–363.