

Computational analysis of the productivity-potential of CAM

Sanu Shameer¹, Kambiz Baghalian¹, C.Y. Maurice Cheung², R. George Ratcliffe¹, Lee J. Sweetlove¹

¹Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, OX1 3RB

²Yale-NUS College, 16 College Avenue West, Singapore, 138527, Singapore.

Abstract

There is considerable interest in transferring Crassulacean acid metabolism (CAM) to C_3 crops to improve their water use efficiency. However, because the CAM biochemical cycle is energetically costly, it is unclear what impact this would have on yield. Using diel flux balance analysis of the CAM and C_3 leaf metabolic networks we show that energy consumption is three-fold higher in CAM at night. However, this additional cost of CAM can be entirely offset by the carbon-concentrating effect of malate decarboxylation behind closed stomata during the day. Depending on the resultant rates of the carboxylase and oxygenase activities of rubisco, the productivity of the PEPCK-CAM subtype is 74-100% of the C_3 network. We conclude that CAM does not impose a significant productivity penalty and that engineering CAM into C_3 crops is likely to lead to a major increase in water-use efficiency without substantially affecting yield.

One of the greatest threats to the future productivity of many agricultural systems is increasing aridity¹. There is a pressing need to develop crop varieties that are able to maintain high productivity with less water. One way in which this might be achieved is to engineer Crassulacean acid metabolism (CAM) photosynthesis into conventional C₃ crops¹⁻³. CAM photosynthesis requires substantial reorganisation of central carbon metabolism to allow nocturnal CO₂ fixation by phosphoenolpyruvate carboxylase and diurnal decarboxylation of the resulting malate to supply CO₂ for photosynthesis⁴. The CAM cycle requires not only two CO₂ fixation steps but also the synthesis, storage, transport and degradation of a large amount of carbohydrate and organic acids (Fig. 1). As a result, CAM is thought to be metabolically more expensive than C₃ metabolism⁵ suggesting that transferring the CAM trait to C₃ plants would incur a penalty in terms of crop yield. Nevertheless, the fact that some CAM crops such as *Agave tequilana*, can achieve productivities comparable to C₃ and C₄ crops^{6,7} casts doubt on the idea that CAM is inherently less efficient.

Another aspect of CAM that is not fully understood, but which also relates to energetics, is the divergence from C₃ in the pathway used to mobilise transitory stores of starch at night⁴. CAM leaves use the phosphorolytic pathway as their main route of starch degradation in which the starch granule is degraded directly to glucose 1-phosphate (G1P) by a chloroplastic α -glucan phosphorylase. In contrast, C₃ leaves mainly use the hydrolytic pathway of starch breakdown in which the starch granule is degraded by the actions of β -amylase and the debranching enzyme isoamylase 3, producing mainly maltose which is exported to the cytosol⁸. It is clear that the hydrolytic pathway is more energy-expensive given its use of an ATP-dependent hexokinase (Fig. 1). This has led to the suggestion that the use of the phosphorolytic pathway in CAM is for energy-saving such that the leaf can meet the energy requirements of nocturnal accumulation of malate in the vacuole⁴.

Ultimately, to understand both of these issues, a detailed analysis of the energetics of CAM is required. Previous attempts to manually estimate CAM ATP-use stoichiometry were limited by the relatively small metabolic networks considered^{9,10}. Moreover, the energetics of a leaf depend on NAD(P)H as well as ATP¹¹. We have previously established an integrated diel (day-night) flux balance analysis modelling framework for leaf metabolism and this computational method allows for a full accounting of ATP and NAD(P)H production and expenditure across the entire metabolic network¹². Using this framework, an initial analysis of CAM energetics was made¹² but this neglected several important features of CAM, notably the proton-balancing / acidification of the vacuole at night¹³. Here we report the results of a more detailed analysis of the energetics and productivity of CAM.

Metabolic model construction and refinement

To generate a suitable metabolic model to compare the productivity / energetics of CAM and C₃ leaves, we refined our previously-established diel-flux balance analysis metabolic model¹². First, we reduced the previous Arabidopsis-based genome-scale stoichiometric model of metabolic reactions to a 'core' stoichiometric model of central plant metabolism¹⁴. This includes all the reactions and transport steps required for synthesis of major biomass components. The use of a core model is advantageous because it reduces the size of the model to a point where manual checking and curation is feasible. The core model established here consists of 641 reactions and 555 metabolites. Moreover, because central metabolism is highly conserved, the model can be generically applied for modelling any plant species. We

confirmed that the reactions of the core model are present in every species-specific network in the PlantCyc database¹⁵ which includes a range of C₃, C₄ and CAM species.

The second refinement to the model was to take into account the effect of organellar pH on the charged state of metabolites. CAM leaves acidify their vacuoles at night and this affects the charge state of the malate that accumulates and the number of protons required to balance that charge. To account for this, specific charged states of every metabolite were introduced into the model in stoichiometric proportions according to the pH of the subcellular compartment and the pK_a of the metabolite (see Materials and Methods for more details). All reactions were then checked to ensure that they were correctly charge- and proton-balanced. Subcellular compartmentation, intracellular transport steps, maintenance costs and the diel framework were all included in the model as before^{12,16}. The full model is provided as an SBML file (Supplementary Data 1).

To set up the model to represent mature leaves during a single diel cycle, a light input of 200 $\mu\text{mol m}^{-2} \text{s}^{-1}$ was specified for a 12 h light phase and the output of the model was specified as sucrose and a range of amino acids to be exported to the phloem (hereafter termed 'phloem output') based on the typical relative abundance of these compounds in phloem sap for a C₃ plant (tomato^{17,18}) and a CAM plant (*Opuntia ficus-indica*¹⁹); Supplementary Data 2. As before, the output was constrained to occur three times as fast during the day as at night²⁰. To predict fluxes, we used linear programming to compute the set of fluxes that maximised the phloem output as a primary objective while minimising the sum of fluxes as a secondary objective²¹. The secondary optimisation prevents futile cycles carrying unnecessarily high fluxes. To predict fluxes for C₃ leaves, CO₂ was allowed as an input during the day, whereas for CAM, CO₂ was only allowed as an input at night. As in the previous diel model¹² this approach generated appropriate fluxes through known pathways of both C₃ and CAM leaves (Supplementary Data 3). SBML files for the C₃ and CAM models, including all the constraints used here are provided as Supplementary Data 4 and 5.

The effect of nocturnal vacuolar acidification

The charge- and proton-balanced model allowed for the differences in pH between subcellular compartments, and for the acidification of the CAM leaf vacuole at night. The vacuolar pH was set to a typical value of 5.5²² in both models, except for the dark phase of the CAM model where the pH was set to 3.3²³. The pH values influence the population of the charged forms of ionisable metabolites according to their pK_a values. Taking malate as an example: this was represented in the model at pH 7.4 (cytosol) as 100% malate²⁻; at pH 5.5 as 70% malate²⁻, 30% Hmalate⁻ (net charge -1.7); and at pH 3.3 as 56% Hmalate⁻, 42% H₂malate (net charge of -0.56). Hence when malate²⁻ is moved from the cytosol into the acidic vacuole, the model must pump an appropriate quantity of protons into the vacuole using the tonoplast energy-dependent (PP_i or ATP) proton pumps to allow for the partial protonation of the malate.

The fluxes of the major reactions that consume or produce protons in the CAM model are shown in Fig. 2. It can be seen that the operation of the CAM cycle has major implications for proton balancing, particularly at night. In order to supply the PEP for nocturnal CO₂ fixation by PEP carboxylase, there is a high flux through the reactions of starch breakdown and glycolysis. This produces a substantial quantity of protons in the cytosol: 24.6 $\mu\text{mol H}^+ \text{m}^{-2} \text{s}^{-1}$. However, because of the high rate of malate accumulation in the vacuole at night, protons are pumped out of the cytosol into the vacuole at a rate of 16.4 $\mu\text{mol H}^+ \text{m}^{-2} \text{s}^{-1}$ (Fig. 2). Protons

are also consumed in the cytosol during the conversion of PEP to malate (at a rate of $4.3 \mu\text{mol H}^+ \text{m}^{-2} \text{s}^{-1}$). It can be seen that these reactions alone lead to a near-balance of cytosolic protons, with the remaining balance accounted for by the transport reactions responsible for moving respiratory substrates and cofactors between cytosol and mitochondria (Fig. 2). Note that a small amount of citrate accumulates in the vacuole at night in the CAM model which helps meet the proton balance during the day (protons are released when the carboxylic acid changes speciation state between vacuole and cytosol).

An interesting system-level prediction from the model is that the flux of the tonoplast ATPase in a CAM leaf at night is equal to the rate of malate accumulation in the vacuole. Since the tonoplast ATPase transfers 2 protons per ATP into the vacuole, the ratio of proton and malate influx into the vacuole is 2:1, which is exactly what is observed from measurements of malate content and titratable acidity of CAM plants^{24–26}. Note that this model prediction is solely the result of the net subcellular movement of carboxylic acids (both malate and citrate) and the effect of pH on the charged state of these metabolites and demonstrates that accounting for these charge balances and pH changes is important to generate a realistic model of CAM. If the acidification of the vacuole in the dark is ignored in the modelling framework, the CAM cycle leads to an excess of cytosolic protons at night. In this scenario, the model predicts that excess protons are pumped out of the cell into the apoplast leading to its acidification (Supplementary Fig. 1).

Further validation of the diel CAM model

The diel FBA approach used in this study has been shown to generate metabolic behaviour consistent with experimental data in C_3 leaves¹². The prediction of a 2:1 stoichiometry of vacuolar protons: malate suggests that the diel CAM model also generates realistic behaviour. To further validate the CAM model, it was constrained with the measured net CO_2 uptake rate²⁷ and the metabolite composition of the phloem sap¹⁹ for the CAM plant *Opuntia ficus-indica*. When these constraints were applied, the model predicted nocturnal malate accumulation at a rate of $10.37 \mu\text{mol m}^{-2} \text{s}^{-1}$ and daytime starch accumulation at a rate of $6.09 \mu\text{mol m}^{-2} \text{s}^{-1}$. Given a 12 h photoperiod, this corresponds to 448 mmol malate m^{-2} accumulated at night and 263 mmol starch (as glucose equivalents) m^{-2} during the day. Experimental measurements of dawn and dusk levels of malate and carbohydrate (starch + glucomannan) in chlorenchyma cells of this species¹⁹ allow the actual amount of nocturnal malate and daytime carbohydrate accumulated to be estimated as 422–629 mmol malate m^{-2} and 278–391 mmol carbohydrate (as glucose equivalents) m^{-2} . The model predictions for malate are hence within the range of experimental values and the starch prediction is within 5% of the lower end of the experimental range, demonstrating that the CAM model is capable of accurate and realistic predictions for CAM cycle fluxes.

Comparison of the energetic demands of CAM and C_3

The metabolic efficiency of CAM and C_3 leaves can be computed from the diel models by quantifying the total demand of the system for ATP / NAD(P)H and the net metabolic output of the system, which in this case is the total flux of sucrose and amino acids to the phloem. To allow comparison on a common basis, we fixed this output flux in both the C_3 and CAM models to a value $0.712 \mu\text{mol m}^{-2} \text{s}^{-1}$ which is the maximal output of the C_3 model when photon uptake was constrained to a value of $200 \mu\text{mol m}^{-2} \text{s}^{-1}$. Minimisation of sum of fluxes was then used as the optimisation objective for predicting fluxes. ATP consuming and producing processes were identified and quantitative comparisons were made (Fig. 3). ATP demand in the CAM

model was only slightly higher than the C₃ model during the day (1.2-fold) but a substantial 3-fold higher than C₃ at night. This was mainly due to high ATP consumption by the tonoplast ATP-dependent proton pump and the high flux through cytosolic phosphofructokinase (PFK), which accounted for 42% and 22% of the total ATP turnover, respectively (Fig. 3). In contrast to the C₃ network at night, the majority (50%) of the nocturnal ATP demand was met by substrate-level phosphorylation during glycolysis, with only 44% being synthesized by mitochondrial ATP synthase (Fig. 3). This is due to the fact that a high glycolytic flux is required to convert the hexose products of starch mobilisation into PEP. Nevertheless, because of the high ATP demand, the mitochondrial ATP synthase flux is still 1.6-fold higher than in the C₃ model. This suggests that attempts to engineer CAM into C₃ plants may need to consider mitochondrial respiratory capacity in addition to the reactions of the core CAM cycle if leaf productivity is to be maintained. In line with this, transcripts relating to mitochondrial biogenesis were among those that were observed to be strongly upregulated at night in an analysis of diel transcriptomic changes in the CAM plant *Agave americana*²⁸.

With respect to the demand for reducing power, a similar pattern emerges: the NAD(P)H demand of CAM during the day was similar to C₃ (1.1 times higher) but 3.2-fold greater than C₃ at night. (Supplementary Fig. 2). This additional reducing power is mainly required for synthesis of malate from oxaloacetate and is provided principally by the high flux through the cytosolic glyceraldehyde 3-phosphate dehydrogenase reaction. Interestingly, it has been observed that the NADPH: NADP⁺ ratio drops sharply towards the end of the night in *A. americana*²⁸. Hence, the high metabolic demand for nocturnal reducing power as predicted by our model, may leave the CAM plant less able to ameliorate oxidative stress at night and this ties in with a marked increase in ascorbate observed during the night in *Agave*²⁸. From these results, it is evident that CAM leaves require more energy to sustain their metabolic processes than C₃ leaves. Note that as observed previously¹² there are minor differences in overall energetic efficiency of CAM depending on the subtype (malic enzyme versus PEP carboxykinase) and choice of storage carbohydrate (starch versus sucrose) with the starch-storing, PEP carboxykinase subtype the most efficient (Supplementary Data 6)

Comparison of the productivity of CAM and C₃

The substantial additional energy cost of CAM opens the question of how the very high productivities achieved in some CAM crops can be achieved. One potential explanation is the CO₂ concentrating effect arising from high rates of malate decarboxylation behind closed stomata²⁹. If significant suppression of photorespiration occurs as a result then this would represent an energy saving. However, there is considerable uncertainty as to the degree of suppression of photorespiration in CAM plants because O₂ as well as CO₂ builds up behind closed stomata and a wide range of concentrations of the two gases have been reported³⁰. When these data were used to compute the ratio of the carboxylase / oxygenase activities of rubisco (v_c/v_o) using an enzyme kinetic model, it was concluded that photorespiration either occurs at the same rate as in C₃ plants or is suppressed by as much as 60%³¹. To test whether this would have a significant impact on the overall energetics of the CAM network, we used the model to predict the productivity of CAM (export of sucrose and amino acids to the phloem) with different v_c/v_o ratios and compared this to C₃. In the C₃ model, v_c/v_o was constrained to a value of 3 based on measured fluxes for *Arabidopsis*³². The results of the analysis are shown in Fig. 4. For fructan-storing CAM plants, there is some uncertainty as to whether phloem loading is symplastic or apoplastic^{19,33}, but this had little effect on the calculated productivity (Supplementary Fig. 4). The analysis shown in Fig. 4 demonstrates that reduced

photorespiration can be sufficient to offset the energetic cost of the CAM cycle. If there is no CO_2 -concentration ($v_d/v_o = 3$), then the model predicted that the productivity of CAM was 26-43% lower than C_3 (depending on the CAM subtype). However, as the CO_2 -concentrating effect was imposed (via an increase in the v_d/v_o) all CAM subtypes, except the fructan-storing malic-enzyme CAM subtype, the least productive of the subtypes, were able to exceed the productivity of C_3 . The most productive CAM subtype was the starch-storing PEP carboxykinase subtype which had an equivalent productivity to C_3 at a v_d/v_o of 7.2. To establish whether the estimated range of *in vivo* v_d/v_o ratios for CAM plants³¹ is sufficient to offset the cost of the CAM cycle, we superimposed this range on model-predicted curves (the green vertical bar in Fig. 4). From this, it can be seen that it is indeed possible that a CAM leaf (starch-storing, PEPCK subtype) could match the productivity of a C_3 leaf. More work is required to generate more reliable experimental estimates of v_d/v_o , for example by stationary ^{13}C metabolic flux analysis³², but we can conclude from this analysis that CAM can be highly productive, approaching the level of C_3 metabolism when compared under equivalent conditions.

To check that this conclusion is not sensitive to the precise parameters utilised in the models, we repeated the analysis under a range of different model input and output parameters: varying the photoperiod, the relative phloem output at day and night, maintenance cost and the metabolite composition of the phloem output (varying the sugar: amino acids ratio). The results of the analysis are shown in Supplementary Fig. 3 in which the productivity of the starch-storing PEPCK CAM subtype is computed under the various conditions and with varying v_d/v_o values. Without exception, the analysis supports the conclusion that the CAM network can be equally as productive as C_3 given a carbon-concentrating effect of CAM. In fact, the precise v_d/v_o ratio required for an equivalent productivity between CAM and C_3 was not altered under any of these conditions. Small changes in CAM productivity relative to C_3 were observed when carbon-concentration was absent in CAM ($v_d/v_o = 3$), but these were not more than a few percent. For example, increasing the photoperiod from 12 h to 18 h increased the diel output of both C_3 and CAM (as is to be expected given that the energy input to the model is greater), but without carbon-concentration, the longer day was slightly more beneficial for the CAM network. This can be explained due to the greater cost of the CAM network in the absence of carbon-concentration.

The model therefore leads to a robust conclusion with respect to the relative biochemical productivity of the C_3 and CAM metabolic networks. However, it is worth pointing out that other factors in addition to the biochemical efficiency of metabolism will constrain growth and productivity in C_3 and CAM plants. These include limitations in the availability of mineral nutrients, differences in the composition of the biomass of the plant tissues and genetic control of growth and development. Given the wide variety of plant growth forms and the very broad range of ecological niches occupied by plants with these two photosynthetic types, capturing the full spectrum of growth patterns and their relationship to underlying metabolism is outside the scope of the current study. What can reasonably be concluded from this study, is that if PEPCK-CAM is successfully engineered into a starch-storing C_3 leaf, then the productivity of the engineered leaf would be 74-100% of the original C_3 plant, under the same conditions. The effect of the type of diel carbon store on the productivity of CAM also allows us to infer which C_3 crops would be most productive when engineered for CAM: crops such as soybean and potato that utilise starch as their leaf carbon store are likely to suffer a smaller yield penalty

when running CAM than cereal crops such as wheat and barley that mainly utilise sugars for this purpose³⁴.

Starch breakdown pathway in CAM

The final aspect of CAM that we examined was the dominant use of the phosphorolytic pathway of starch breakdown in CAM leaves, in contrast to C_3 plants which mainly use a hydrolytic route. It has been suggested that this could be because of the lower ATP cost of the phosphorolytic route (due to avoiding the need to use ATP-dependent hexokinase to phosphorylate glucose for its entry into glycolysis). However, without a system-level energetic analysis, it is unclear whether this ATP saving is significant in the context of the overall energy budget of the leaf. To investigate this, we examined the effect of each route of starch breakdown in the C_3 and CAM diel models. For hydrolytic starch breakdown in CAM, we consider both export of glucose from the chloroplast and export of maltose. For the C_3 simulations a v_C/v_O of 3 was used as before; and for CAM the v_C/v_O ratio was set to 5.15 (the median value of the range of experimental estimates³¹). These simulations showed that the CAM system using the phosphorolytic route for starch breakdown exported 6-12% more sucrose and amino acids to the phloem than the CAM system using β -amylase. In contrast, the C_3 system using the phosphorolytic route showed only a negligible increase in phloem output (0.6-1.3%) over the hydrolytic route (Table 1). This can be explained by the respective effects on nocturnal ATP consumption (Table 1), with a 14-26% saving in ATP by use of the phosphorolytic route in CAM compared to only a 4-8% saving in C_3 . The difference in ATP saving between CAM and C_3 reflects the much higher rate of starch breakdown in CAM than C_3 , required to supply the PEP for nocturnal CO_2 fixation as well as supporting the carbon and energy requirements of the leaves at night. In our models, the rate of nocturnal starch breakdown in CAM was 8.7-fold higher than in C_3 (Supplementary data 3). Similar large increases in nocturnal carbohydrate store mobilisation have been measured experimentally: e.g. nocturnal starch mobilisation was 278 - 391 μmol glucose equivalents m^{-2} in *Opuntia ficus-indica*²⁷ compared to 19 ± 2 in *Arabidopsis*³⁵. Overall, this analysis demonstrates that the higher rate of starch breakdown in CAM makes the choice of starch degradation mechanism more energetically significant for nocturnal metabolism than in C_3 leaves and thus provides an explanation for the dominance of the phosphorolytic route in CAM.

Conclusion

The integrated diel flux balance analysis approach used in this study allowed us to make a system-level comparison of the energetics and productivity of the metabolic networks that operate in C_3 and CAM leaves, considering both day and night metabolism. From our analysis we can make three major conclusions: first, that a carbon-concentrating effect of daytime stomatal closure would be sufficient to offset the energetic cost of running the CAM cycle, implying that CAM can be expected to have comparable productivity to C_3 metabolism under the same conditions; secondly, that even without suppression of photorespiration due to carbon concentration, CAM productivity (per photon utilised) for the starch-storing PEPCK subtype is still 80% of that of C_3 metabolism; and thirdly, the use of the phosphorolytic route of starch breakdown in CAM leaves can be explained in terms of a 14-26% saving in nocturnal ATP consumption in CAM and a much smaller effect in C_3 leaves. This is due to the need for a much greater turnover of starch in CAM leaves. The analysis also highlights the greatly increased energy consumption of a CAM leaf at night compared to a C_3 leaf and this has implications for attempts to engineer CAM into C_3 plants. For example, it may be necessary to

increase the capacity for nocturnal respiration and mitochondrial ATP biosynthesis if the aim is to introduce CAM while maintaining a high productivity. There would also be a significant productivity benefit if the C₃ leaf was engineered to switch from hydrolytic to phosphorolytic starch breakdown.

Methods

Building a charge balanced metabolic model of primary metabolism in fully-developed plant leaves

A detailed description of the procedures used to build and analyse the model is provided in Supplementary Methods. Briefly, a core model of plant primary metabolism was built by reducing the previously published Arabidopsis genome-scale metabolic model^{12,16}. A draft of this core metabolic model was first generated by merging the reactions that were found to be active in parsimonious FBA simulations with glucose, sucrose or photons as the sole source of energy and maximization of biomass as the primary objective function. This draft was then manually curated with 54 metabolites added, removed or modified and 210 reactions added, removed or modified - see Supplementary Data 7 for additional details. All reactions in the core model were then duplicated and reactions representing the accumulation of starch, sucrose, nitrate, amino acids and organic acids were added to develop a diel leaf model as in our previous work¹².

Charge balancing the model required a knowledge of the pH values in the different subcellular compartments and the pK_a values of metabolites. The following representative pH values were used: cytosol, 7.4, peroxisome, 7.4, mitochondrial matrix, 8.0, chloroplast stroma, 8.0 and vacuole 5.5^{36–39} except in the CAM model in the dark phase where vacuolar pH was 3.3. For every metabolite in a compartment, the abundance of all charge states at the pH of the compartment was predicted using ChemAxon software. If the major charge state of a metabolite was present at less than 90% abundance, information on the fraction of each charged state was then used to update all reactions in the model and protons were added to the reactants or products of the reaction to balance the charges. See Supplementary Information 1 for a full description of the charge balancing of metabolic and transport reactions.

Pseudo-reactions representing oxidation of NADPH (cytosolic, mitochondrial and chloroplastic) and hydrolysis of ATP (cytosolic) were included in the model to represent non-growth associated maintenance (NGAM). These reactions were constrained to carry flux in a 1:3 ratio based on a previous estimate¹⁶. As in our previous leaf model¹², NGAM was assumed to be the same during day and night. Flux through NGAM pseudo reactions were estimated by gradually increasing them until night metabolism exhibited a carbon conversion efficiency of 50%.

Constraining flux ratios in flux balance analysis

Flux balance analysis was performed using the COBRApy package⁴⁰. While most constraints used in this study were straightforward, some of them required maintaining flux through multiple reactions in fixed ratios. The v_c/v_o constraint is an example of such a constraint. COBRApy does not offer a built-in function to set ratios of metabolic fluxes and so the steady state assumption of the metabolic system was exploited to implement these constraints. If P and Q are two non-zero sets of reactions whose sum of fluxes are to be constrained in the ratio m:n, then a pseudometabolite M is added to all reactions in P and Q such that the coefficient of M in all reactions in P and Q are n and -m respectively. Given the steady state assumption, this will result in an additional constraint to the solution space in the form of the equation below

$$n \times f_{p1} + n \times f_{p2} + \dots + n \times f_{pr} + (-m) \times f_{q1} + (-m) \times f_{q2} + \dots + (-m) \times f_{qs} = 0$$

where $\{p1, p2, \dots, pr\} = P$; $\{q1, q2, \dots, qs\} = Q$
and f_a is the flux through reaction a.

Rewriting the equation, it is evident that sum of fluxes of reactions in sets P and Q are now constrained to the ratio m:n.

$$\Rightarrow n(f_{p1} + f_{p2} + \dots + f_{pr}) = m(f_{q1} + f_{q2} + \dots + f_{qs})$$

$$\Rightarrow \frac{(f_{p1} + f_{p2} + \dots + f_{pr})}{(f_{q1} + f_{q2} + \dots + f_{qs})} = \frac{m}{n}$$

Database, Tools and packages used

PlantCyc and 22 plant BioCyc databases were downloaded from the Plant Metabolic Network (www.plantcyc.org). The Pathway Tools software suite⁴¹ with its API JavaCyc was used to extract data from the downloaded BioCyc databases. The Met4J package⁴² was used to add metadata to the core model and generate the diel leaf model. The parsimonious flux balance analysis (pFBA)²¹ function available in the COBRApy package⁴⁰ was used to perform linear optimization of the model using the GLPK solver.

Data Availability

The authors declare that all data and model files supporting the findings of this study are available within the paper and its supplementary information files. The code required to reproduce the results in this paper is provided at the Sweetlove laboratory GitHub repository (<https://github.com/ljs1002/Shameer-et-al-Productivity-potential-of-CAM>).

Correspondence and requests for materials should be addressed to lee.sweetlove@plants.ox.ac.uk; or george.ratcliffe@plants.ox.ac.uk

Acknowledgements

Funding from ERA-CAPS (Simultaneous manipulation of source and sink metabolism for improved crop yield; BO 1482/18-1 | FE 552/33-1 | RE 1351/2-1 | SW 122/2-1) is acknowledged. We would also like to thank Prof. Andrew Smith (Department of Plant Sciences, University of Oxford) for advice and discussions about CAM.

Author contributions

LJS, RGR and SS conceived the study and co-wrote the paper. SS, CYMC and KB constructed and curated the core model. SS did all subsequent computational analyses of the model and analysed the data generated.

References

1. Borland, A. M. et al. Engineering crassulacean acid metabolism to improve water-use efficiency. *Trends Plant Sci.* **19**, 327–38 (2014).
2. Borland, A. M., Griffiths, H., Hartwell, J. & Smith, J. A. C. Exploiting the potential of plants with crassulacean acid metabolism for bioenergy production on marginal lands. *J. Exp. Bot.* **60**, 2879–2896 (2009).
3. Yang, X. et al. A roadmap for research on crassulacean acid metabolism (CAM) to enhance sustainable food and bioenergy production in a hotter, drier world. *New Phytol.* **207**, 491–504 (2015).
4. Borland, A. M., Guo, H. B., Yang, X. & Cushman, J. C. Orchestration of carbohydrate processing for crassulacean acid metabolism. *Curr. Opin. Plant Biol.* **31**, 118–124 (2016).
5. Jordan, P. W., Nobel, P. S., Gazette, B. & Dec, N. Height distributions of two species of cacti in relation to rainfall, seedling establishment, and growth. *Bot. Gaz.* **143**, 511–517 (2007).
6. Yan, X. Y., Tan, D. K. Y., Inderwildi, O. R., Smith, J. A. C. & King, D. A. Life cycle energy and greenhouse gas analysis for agave-derived bioethanol. *Energy Environ. Sci.* **4**, 3110–3121 (2011).
7. Nobel, P. & Valenzuela, A. Environmental responses and productivity of the CAM plant, *Agave tequilana*. *Agric. For. Meteorol.* **39**, 319–334 (1987).
8. Smith, A. M. Starch in the Arabidopsis plant. *Starch* **64**, 421–434 (2012).
9. Nobel, P. S. Achievable productivities of certain CAM plants: basis for high values compared with C3 and C4 plants. *New Phytol.* **119**, 183–205 (1991).
10. Winter, K. & Smith, J. A. C. Crassulacean acid metabolism: current status and perspectives. in *Crassulacean acid metabolism. Biochemistry, ecophysiology and evolution* (eds. Winter, K. & Smith, J. A. C.) 389–426 (Springer-Verlag, 1996).
11. Cheung, C. Y. M., Ratcliffe, R. G. & Sweetlove, L. J. A method of accounting for enzyme costs in flux balance analysis reveals alternative pathways and metabolite stores in an illuminated Arabidopsis leaf. *Plant Physiol.* **169**, 1671–1682 (2015).
12. Cheung, C. Y. M., Poolman, M. G., Fell, D. A., Ratcliffe, R. G. & Sweetlove, L. J. A diel flux balance model captures interactions between light and dark metabolism during day-night cycles in C₃ and Crassulacean Acid Metabolism Leaves. *Plant Physiol.* **165**, 917–929 (2014).
13. Holtum, J. A. M., Smith, J. A. C. & Neuhaus, H. E. Intracellular transport and pathways of carbon flow in plants with crassulacean acid metabolism. *Funct. Plant Biol.* **32**, 429–449 (2005).
14. Arnold, A. & Nikoloski, Z. Bottom-up metabolic reconstruction of Arabidopsis and its application to determining the metabolic costs of enzyme production. *Plant Physiol.* **165**, 1380–1391 (2014).
15. Schläpfer, P. et al. Genome-wide prediction of metabolic enzymes, pathways, and gene clusters in plants. *Plant Physiol.* **173**, 2041–2059 (2017).
16. Cheung, C. Y. M. et al. A method for accounting for maintenance costs in flux balance analysis improves the prediction of plant cell metabolic phenotypes under stress conditions. *Plant J.* **75**, 1050–1061 (2013).
17. Valle, E. M., Boggio, S. B. & Heldt, H. W. Free amino acid composition of phloem sap and growing fruit of *Lycopersicon esculentum*. *Plant Cell Physiol.* **39**, 458–461 (1998).
18. Walker, A. J. & Ho, L. C. Carbon translocation in the tomato: effects of fruit temperature on carbon metabolism and the rate of translocation. *Ann. Bot.* **41**, 825–832 (1977).

19. Wang, N. & Nobel, P. S. Phloem transport of fructans in the Crassulacean Acid Metabolism species *Agave deserti*. *Plant Physiol.* **116**, 709–14 (1998).
20. Gibon, Y. et al. Adjustment of diurnal starch turnover to short days: depletion of sugar during the night leads to a temporary inhibition of carbohydrate utilization, accumulation of sugars and post-translational activation of ADP-glucose pyrophosphorylase in the followin. *Plant J.* **39**, 847–862 (2004).
21. Lewis, N. E., Nagarajan, H. & Palsson, B. O. Constraining the metabolic genotype-phenotype relationship using a phylogeny of *in silico* methods. *Nat. Rev. Microbiol.* **10**, 291–305 (2012).
22. Kurkdjian, A. & Guern, J. Intracellular pH: measurement and importance in cell activity. *Ann. Rev. Plant. Physiol. Plant. Mol. Biol.* **40**, 271–303 (1989).
23. Cheffings, C. M., Pantoja, O., Ashcroft, F. M. & Smith, J. A. C. Malate transport and vacuolar ion channels in CAM plants. *J. Exp. Bot.* **48**, 623–631 (1997).
24. Lüttge, U. & Ball, E. Electrochemical investigation of active malic acid transport at the tonoplast into the vacuoles of the CAM plant *Kalanchoë daigremontiana*. *J. Membr. Biol.* **47**, 401–422 (1979).
25. Luttge, U. & Ball, E. $2H^+ : 1\text{ malate}^{2-}$ stoichiometry during Crassulacean Acid Metabolism is unaffected by lipophilic cations. *Plant, Cell Environ.* **3**, 195–200 (1980).
26. Struve, I., Weber, A., Lüttge, U., Ball, E. & Smith, J. A. C. Increased vacuolar ATPase activity correlated with CAM induction in *Mesembryanthemum crystallinum* and *Kalanchoë blossfeldiana* cv. Tom Thumb. *J. Plant Physiol.* **117**, 451–468 (1985).
27. Nobel, P. S. & Hartsock, T. L. Relationships between photosynthetically active radiation, nocturnal acid accumulation, and CO_2 uptake for a Crassulacean Acid Metabolism plant, *Opuntia ficus-indica*. *Plant Physiol.* **71**, 71–75 (1983).
28. Abraham, P. E. et al. Transcript, protein and metabolite temporal dynamics in the CAM plant *Agave*. *Nat. Plants* **2**, 16178 (2016).
29. Luttge, U. CO_2 -concentrating: consequences in Crassulacean Acid Metabolism. *J. Exp. Bot.* **53**, 2131–2142 (2002).
30. Niewiadomska, E. & Borland, A. Crassulacean acid metabolism: a cause or consequence of oxidative stress in planta? *Prog. Bot.* **69**, 247–266 (2007).
31. Lüttge, U. Photorespiration in phase III of Crassulacean Acid Metabolism: evolutionary and ecophysiological implications. *Prog. Bot.* **72**, 371–384 (2010).
32. Ma, F., Jazmin, L. J., Young, J. D. & Allen, D. K. Isotopically nonstationary ^{13}C flux analysis of changes in *Arabidopsis thaliana* leaf metabolism due to high light acclimation. *Proc. Natl. Acad. Sci.* **111**, 16967–16972 (2014).
33. Pollock, C. J. & Cairns, A. J. Fructan metabolism in grasses and cereals. *Annu. Rev. Plant Physiol. Plant Mol. Biol* **42**, 77–1 (1991).
34. Sicher, R. C., Kremer, D. F. & Harris, W. G. Diurnal carbohydrate metabolism of barley primary leaves. *Plant Physiol.* **76**, 165–169 (1984).
35. Sun, J., Zhang, J., Larue, C. T. & Huber, S. C. Decrease in leaf sucrose synthesis leads to increased leaf starch turnover and decreased RuBP regeneration-limited photosynthesis but not Rubisco-limited photosynthesis in *Arabidopsis* null mutants of SPSA1. *Plant, Cell Environ.* **34**, 592–604 (2011).
36. Couldwell, D. L. et al. Response of cytoplasmic pH to anoxia in plant tissues with altered activities of fermentation enzymes: application of methyl phosphonate as an NMR pH probe. *Ann. Bot.* **103**, 249–258 (2009).
37. Mathieu, Y. et al. Regulation of vacuolar pH of plant cells: I. Isolation and properties of vacuoles suitable for ^{31}P NMR studies. *Plant Physiol.* **89**, 19–26 (1989).
38. Werdan, K., Heldt, H. W. & Milovancev, M. The role of pH in the regulation of carbon fixation in the chloroplast stroma. Studies on CO_2 fixation in the light and dark. *Biochim. Biophys. Acta* **396**, 276–92 (1975).
39. Casey, J. R., Grinstein, S. & Orlowski, J. Sensors and regulators of intracellular pH. *Nat. Rev. Mol. Cell Biol.* **11**, 50–61 (2010).
40. Ebrahim, A., Lerman, J. A., Palsson, B. O. & Hyduke, D. R. COBRApy: COstraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.* **7**, 74 (2013).

41. Karp, P. D., Paley, S. & Romero, P. The Pathway Tools software. *Bioinformatics* **18 Suppl 1**, S225-232 (2002).
42. Cottret, L. et al. MetExplore: a web server to link metabolomic experiments and genome-scale metabolic networks. *Nucleic Acids Res.* **38**, W132-137 (2010).

Tables and Figures

	C ₃			CAM		
	Hydrolytic starch degradation		Phosphorolytic starch degradation	Hydrolytic starch degradation		Phosphorolytic starch degradation
	maltose export	glucose export		maltose export	glucose export	
Photon input ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	200	200	200	200	200	200
Productivity (mmol metabolites exported to phloem $\text{m}^{-2} 24\text{h}^{-1}$)	44.4	44.1	44.7	38.4	35.8	40.9
Rate of nocturnal ATP usage ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	9.48	9.93	9.07	33.2	37.3	29.2

Table 1: The effect of alternative routes of starch breakdown on the productivity and energy use of CAM and C₃ leaves as predicted using the diel flux balance analysis models. V_c/V_o was constrained to 3.0 and 5.15 in C₃ and CAM models, respectively. Photon input was constrained to $200 \mu\text{mol m}^{-2} \text{s}^{-1}$ in both models and parsimonious optimisation was used to compute the flux distribution that maximised the amount of sucrose and amino acids exported to the phloem. To force the model to use a certain route of starch breakdown, the alternative route was constrained to zero flux. ATP usage was calculated by summing the fluxes of all reactions that consumed ATP.

Figure 1. Simplified schematic of the CAM cycle in starch-storing leaves. The two CAM subtypes utilize alternative malate decarboxylating pathways – the malic-enzyme (ME) pathway is shown in magenta and the phospho*eno*lpyruvate carboxykinase (PEPCK) pathway in green. Alternative routes of starch mobilization are shown in red (phosphorolytic route) and blue (hydrolytic route). Abbreviations: CA, carbonic anhydrase; DPE2, disproportionating enzyme 2; G1P, glucose 1-phosphate; G6P, glucose 6-phosphate; GLC, glucose; OAA, oxaloacetate; PEP, phospho*eno*lpyruvate; PEPC, phospho*eno*lpyruvate carboxylase; PEPCK, phospho*eno*lpyruvate carboxykinase; PGM, phosphoglucomutase; Pho1, plastidic α -glucan phosphorylase; PYR, pyruvate; G6PT, glucose 6-phosphate transporter; HK, hexokinase; MDH, malate dehydrogenase; ME, malic enzyme; Pho2, cytosolic α -glucan phosphorylase; PPK, pyruvate-phosphate dikinase.

Figure 2. Representation of major proton consuming and producing reactions in the CAM model. Photon uptake in CAM was constrained to $200 \mu\text{mol m}^{-2} \text{s}^{-1}$ and parsimonious optimisation was used to compute the flux distribution that maximised the amount of sucrose and amino acids exported to the phloem. All reactions producing and consuming protons with a flux greater than $0.5 \mu\text{mol H}^+ \text{m}^{-2} \text{s}^{-1}$ are shown. Red numbers indicate the number of protons consumed in that reaction (as $\mu\text{mol H}^+ \text{m}^{-2} \text{s}^{-1}$) and blue numbers indicate the amount of protons produced (as $\mu\text{mol H}^+ \text{m}^{-2} \text{s}^{-1}$). Blue boxes with dashed edges highlight reactions used to represent non-growth associated maintenance. Abbreviations: as for Fig. 1 with the following additional abbreviations: 2KG, 2-ketoglutarate; 3PG, 3-phosphoglycerate; A.A, amino acids; ACO, *cis*-aconitate; ADPG, ADP-glucose; CIT, citrate; DPG, 1,3-bisphosphoglycerate; F6P, fructose 6-phosphate; FBP, fructose 1,6-bisphosphate; GLCT, glycerate; GLN, glutamine; GLT, glutamate; GSA, glutamate semialdehyde; GT6P, 6-phosphogluconate; IsoCIT, isocitrate; MAL, malate; NAGP, N-acetylglutamyl-phosphate; NAGSA, N-acetyl-L-glutamate 5-semialdehyde; OHPYR, hydroxypyruvate; Pi, inorganic phosphate; PPi, inorganic pyrophosphate; RuBP, ribulose 1,5-bisphosphate; R5P, ribulose 5-phosphate; SUC, sucrose; SUC6P, sucrose 6-phosphate; TP, glyceraldehyde 3-phosphate.

Figure 3. ATP production and consumption in C₃ and CAM systems. C₃ and CAM leaf metabolic flux distributions were predicted by diel FBA and ATP-consuming and -producing fluxes during day and night were identified. The contribution of each flux as a percentage to total ATP consumed/produced is shown. The value for total ATP produced/consumed during day/night is shown above each bar (in units of $\text{mmol m}^{-2} 12\text{h}^{-1}$). Abbreviations: pETC, photosynthetic electron transport chain; mETC, mitochondrial electron transport chain; cPGK, cytosolic phosphoglycerate kinase; SucCoAS, succinyl-CoA synthase; cPK, cytosolic pyruvate kinase; pPGK, plastidic phosphoglycerate kinase; pPRK, plastidic phosphoribulokinase; NGAM, non-growth associated maintenance; cPEPCK, cytosolic phospho*eno*lpyruvate carboxykinase; pAGPase, plastidic ADP-glucose pyrophosphatase; pCPS, carbamoyl phosphate synthetase; pGK, plastidic glycerate kinase; mGS, mitochondrial glutamine synthase; NAGK, N-acetyl-L-glutamate kinase; pGS, plastidic glutamine synthetase; PM-ATPase, plasma membrane proton pump; V-ATPase, vacuolar proton pump; cPFK, cytosolic phosphofructokinase; cHK, cytosolic hexokinase; cNDPK, cytosolic nucleoside diphosphate kinase.

Figure. 4. Predicted productivity of CAM leaves with varying rubisco carboxylase / oxygenase (v_c/v_o) ratios. The range of v_c/v_o values estimated from published measurements of CO_2 and O_2 concentrations and a kinetic model of rubisco are highlighted in green. Productivity was calculated from the computed output of sucrose and amino acids exported to the phloem over a diel cycle, expressed relative to the equivalent value from the C_3 model. Note that a fixed v_c/v_o of 3 was used for C_3 but for ease of comparison a horizontal line is projected at the 100% value. CAM was simulated for the different subtypes (PEPCK, phosphoenolpyruvate carboxykinase; ME, malic enzyme) and storage carbohydrate as indicated.