Synthetic metabolons for metabolic engineering

Chloe Singleton, Thomas P. Howard and Nicholas Smirnoff*

Biosciences, College of Environmental and Life Sciences, University of Exeter, Geoffrey Pope Building, Stocker Road, Exeter, EX4 4QD, Exeter, UK.

*To whom correspondence should be addressed. E-mail: N.Smirnoff@exeter.ac.uk; C.Singleton@exeter.ac.uk; T.P.Howard@exeter.ac.uk

Abstract

It has been proposed that enzymes can associate into complexes (metabolons) that increase the efficiency of metabolic pathways by channelling substrates between enzymes. Metabolons may increase flux by increasing local concentration of intermediates, decreasing the concentration of enzymes needed to maintain a given flux, directing the products of a pathway to a specific subcellular location or minimise the escape of reactive intermediates. Metabolons can be formed by relatively loose non-covalent protein-protein interaction, anchorage to membranes and (in bacteria) by encapsulation of enzymes in protein-coated microcompartments. Evidence that non-coated metabolons are effective at channelling substrates is scarce and difficult to obtain. In plants there is strong evidence that small proportions of glycolytic enzymes are associated with the outside of mitochondria and are effective in substrate channelling. More recently, synthetic metabolons, in which enzymes are scaffolded to synthetic proteins or nucleic acids, have been expressed in micro-organisms and these provide evidence that scaffolded enzymes are more effective than free enzymes for metabolic engineering. This provides experimental evidence that metabolons may have a general advantage and opens the way to improving the outcome of metabolic engineering in plants by including synthetic metabolons in the toolbox.

Key words: Bacterial microcompartments, cyanogenic glycosides, flavonoids, metabolic engineering, photosynthesis

Introduction

The concept of "synthetic biology" embraces the idea of redesigning organisms from the ground up. It is enabled by the standardisation of DNA parts and experimental procedures that allows automation, reliable output measurements and robust biological engineering; the thorough characterisation of DNA parts so that their behaviour in a given context is predictable; and the further modularisation of these parts into devices that perform well-defined functions. This latter concept allows further abstraction in which this information (e.g. DNA sequence, part or device data) is hidden thereby reducing the observable details allowing the researcher to focus on a few key concepts at a time (Freemont and Kitney, 2012). As an example, selectable marker cassettes may be considered devices - nowadays it is unusual for the researcher to look in detail at what resides within these markers, we know they confer resistance to a given antibiotic or herbicide and for most of the time that is all that is required. In applying this approach across the board the parts and devices can be used and re-used to provide predictable, highly controlled levels of gene expression to produce regulatory circuits or novel metabolic pathways with predictable outcomes. This approach often requires transformation with multiple genes and most of the examples are currently from readily transformed microorganisms. The effort is aided by improved gene cloning technology and the decreasing cost of synthesising genes *de novo*. In bacteria, the ability to produce novel organisms with streamlined genomes, in which the various patches and fixes installed during evolution are replaced or streamlined, has emerged (Gibson *et al.*, 2010) and will allow production of more efficient vehicles for production of useful end products. This review will not address the nuts and bolts of DNA manipulation and the problems of expressing multiple transgenes in plants at controlled levels in specific cell types but will rather focus on proteins

and enzymes and how these might be designed and manipulated more effectively for plant metabolic engineering. Metabolic engineering is well established in microorganisms and this is where synthetic biology approaches first emerged. There are metabolic engineering success stories in plants for example expression of carotene biosynthesis genes in rice grain to create "golden rice", multi-vitamin corn (Naqvi *et al.*, 2009) and altered seed lipids (DellaPenna, 2001; Mayer *et al.*, 2008; Napier and Graham, 2010; Ruiz-Lopez *et al.*, 2014; Ye *et al.*, 2000). Metabolic engineering could also be employed to improve crop yield by manipulating photosynthesis, starch synthesis and production of osmolytes and defensive compounds. However, the result of over-expression of a few judiciously chosen enzymes is often disappointing due to lack of understanding of metabolic regulation and possible toxicity of end products. Therefore use of approaches that go beyond enzyme over-expression should be considered. Improved photosynthesis, nutrient use and stress resistance would not only increase the efficiency of plant production but may also allow for the use of plants as chemical factories. Such synthetic biology approaches are emerging in microbial metabolic engineering and the purpose of this review is to consider one of these approaches- the creation of synthetic enzyme complexes (metabolons) and assess their potential application in plants.

Metabolons, microcompartments and metabolic channelling

The term "metabolon" was coined by Srere to describe complexes of enzymes that carry out sequential reactions (Srere, 1987, 2000). This definition is potentially wide and could include stable assemblies of enzymes that carry out complex or vectorial series of reactions such as ribosomes (protein synthesis), proteasomes (protein degradation), photosynthetic and respiratory electron transport complexes, ATP synthase, mammalian and fungal fatty acid

synthase, polyketide synthases, non-ribosomal peptide synthases and cellulosomes (cellulose degradation complexes in bacteria) (Hyeon *et al.*, 2013). These multi-protein complexes carry out series of reactions that require intermediates to be shepherded (channelled) precisely. The relatively complex series of reactions required for fatty acid synthesis are located in a dimeric multifunctional fatty acid synthase in mammals and fungi (Voet and Voet, 2004). Similarly, polyketide synthases and non-ribosomal peptide synthases consist of several proteins each with multiple active sites, allowing synthesis of complex molecules such as antibiotics (Khosla *et al.*, 1999). In the case of the pyruvate dehydrogenase complex, tryptophan synthase and carbamoyl phosphate synthetase, intermediates are effectively channelled between active sites (Voet and Voet, 2004). Apart from these examples of stable or close complexes, where intermediates are directly shuttled between active sites, the existence of functional metabolons based on loose or transitory enzyme-enzyme interactions is a matter of controversy. The proposed advantages of metabolons include increasing the concentration of intermediates at the active sites of sequential enzymes and minimizing escape of reactive or potentially toxic intermediates (Jørgensen *et al.*, 2005; Moller, 2010; Srere, 1987; Sweetlove and Fernie, 2013; Winkel, 2004). These possibilities are illustrated in the review by Chen and Silver (2012). The general argument against metabolons providing an advantage is that metabolites diffuse too quickly relative to reaction rates and cell size for there to be a significant concentration of intermediates in the vicinity of the complex. While demonstrating that complexes are likely to occur *in vitro* is relatively easy, assessing their functionality is much more difficult. To assess the significance of proposed complexes based on protein-protein interaction methods such as pull-downs (e.g. TAP-tagging), surface plasmon resonance and yeast2-hybrid assays, it is necessary to demonstrate that association occurs *in vivo*. This is probably best done with fluorescently-tagged proteins using techniques such as Förster

121 resonance energy transfer (FRET) and bimolecular fluorescence complementation (BIFC) (Ohad
122 *et al.*, 2007). However, methods with a higher throughput such as yeast 2-hybrid or pull-downs
123 followed by protein identification using mass spectrometry are ideal for initial screening. There
124 are numerous examples of proposed metabolons in animals and microorganisms (Chen and
125 Silver, 2012; Srere, 1987). Recent examples include glycolysis (Araiza-Olivera *et al.*, 2013;
126 Puchulu-Campanella *et al.*, 2013), the TCA cycle (Meyer *et al.*, 2011; Mitchell, 1996; Velot *et al.*,
127 1997), amino acid biosynthesis (de Cima *et al.*, 2012; Hutson *et al.*, 2011; Islam *et al.*, 2007),
128 acyl ester biosynthesis (Jiang and Napoli, 2013) and melanin biosynthesis (Sugumaran *et al.*,
129 2000). While most of the examples involve protein-protein interactions (with the possible
130 inclusion of a membrane-located anchoring proteins), other methods of complex formation
131 have been identified, for example association of glycolytic enzymes with F-actin (Araiza-Olivera
132 *et al.*, 2013) or targeting of enzymes to lipid droplets (Jiang and Napoli, 2013). In plants,
133 plastogobuli – lipid inclusions in chloroplasts- have enzymes of isoprenoid biosynthesis and
134 chlorophyll degradation associated with them (Lundquist *et al.*, 2012; Nacir and Bréhélin, 2013)
135 and so could organise enzymes that metabolise lipophilic substrates. Despite the relatively
136 abundant evidence for physical association of enzymes into metabolons, more definitive
137 evidence that pathway intermediates are effectively channelled (for example from isotope
138 dilution studies) is very rare (Graham *et al.*, 2007), leading to scepticism about their function.
139
140 Given the increasing number of metabolons proposed in other organisms, it is not surprising
141 that the evidence for association of enzymes into complexes in plants is accumulating. A
142 number of reviews cover this evidence (Jørgensen *et al.*, 2005; Sweetlove and Fernie, 2013;
143 Winkel, 2004). Pathways, or parts of pathways, in which enzymes are associated in complexes
144 include glycolysis and various biosynthesis pathways (polyamines, flavonoids,

phenylpropanoids, cyanogenic glycosides, sporopollenin, long chain alkanes, cyanogenic glycosides and indole acetic acid). The references and evidence for the existence of enzyme complexes and their functional significance are shown in Table 1 while other examples of possible secondary metabolite metabolons have been reviewed elsewhere (Jørgensen *et al.*, 2005). Mining the high throughput Arabidopsis yeast 2-hybrid data should indicate further pathways in which enzymes form complexes (Arabidopsis Interactome Mapping Consortium, 2011; Zhang *et al.*, 2010). It is possible that the search for metabolons in plants could also be guided by the conservation of protein-protein interactions across phylogenetic groups. Clearly, the demonstration of enzyme complexes does not prove that pathway flux is improved or that potentially toxic intermediates are confined. Evidence that this benefit occurs for some pathways was discussed in the previous section. In plants, this level of evidence is sparse. In the case of association of glycolytic enzymes with the outer mitochondrial membrane, isotope dilution experiments have shown that addition of unlabelled intermediate has little effect on the specific activity of products produced from a labelled precursor (Giege *et al.*, 2003; Graham *et al.*, 2007). This is key evidence that indicates not only are the enzymes spatially organised, but that there is also a degree of direct substrate channelling between them. In the case of glycolysis, it is important to note that only a small proportion of the enzymes locate to the mitochondria and therefore the effect is local.

The idea that enzymes of the Calvin-Benson cycle may also aggregate into metabolons has been around for some time (Anderson *et al.*, 2006) but their existence remains unproven. The Calvin-Benson cycle however does provide an alternative paradigm for regulation of central metabolism by protein complexes. A well-documented multi-enzyme complex exists between two, non-sequential Calvin-Benson cycle enzymes, phosphoribulokinase (PRK) and

glyceraldehyde 3-phosphate dehydrogenase (GAPDH) that is mediated by the small chloroplast protein CP12 (Wedel and Soll, 1998; Wedel *et al.*, 1997). When PRK and GAPDH are aggregated within this complex they are far less active and more susceptible to allosteric inhibition than their unbound forms. The PRK/GAPDH/CP12 complex, which is present in darkened leaves, is capable of almost instantaneous dissociation (activation) upon illumination, and similar, rapid association (deactivation) upon transfer to darkness (Howard *et al.*, 2008). This process, dependent upon the activity of the photosynthetic electron transport chain and mediated by thioredoxin (Howard *et al.*, 2008; Marri *et al.*, 2009) allows swift, dynamic activation/inactivation of enzyme activity in response to environmental changes. This mechanism provides an alternative function for synthetic metabolic protein complexes in line with the aims of making smarter metabolically engineered pathways that are responsive to cellular demands (Zhang *et al.*, 2012). Despite the lack of strong evidence for a more general Calvin-Benson cycle metabolons, recent detailed flux analysis using $^{13}CO_2$ labelling in *Arabidopsis thaliana* provides evidence for discrete pools of metabolites, a feature that could indicate some degree of metabolite channelling within the stroma (Szecowka *et al.*, 2013).

Bacterial microcompartments (metabolosomes/enterosomes) consist of enzymes that are encapsulated in self-assembling protein coats forming particles that are of the order of 100 nm in diameter. They are found in various bacteria and generally contain pathways that have reactive or toxic intermediates, for example in ethanol, ethanolamine and propanediol utilisation (Kerfeld *et al.*, 2010). Cyanobacterial carboxysomes, which are part of the carbon dioxide concentrating mechanism of these organisms, are a well-known example (Kerfeld *et al.*, 2010; Rae *et al.*, 2013). They contain ribulose bisphosphate carboxylase-oxygenase (RuBisCO) and carbonic anhydrase (CA). The coat proteins form suitably charged pores that allow ingress

of substrates (RuBP and bicarbonate). Within the carboxysome, CO_2 release from bicarbonate is catalysed by CA and then RuBisCO catalyses the carboxylation of RuBP to form phosphoglyceric acid, the first product of the Calvin-Benson cycle. The localised release of CO_2 in the carboxysome outcompetes the oxygenase reaction of RuBisCO which otherwise results in "wasteful" photorespiration (Bauwe *et al.*, 2012; Maurino and Peterhansel, 2010). Protein-coated microcompartments could therefore provide another route to controlling or improving metabolic flux and their introduction into C_3 chloroplasts, along with bicarbonate transporters has been suggested as an approach to decreasing photorespiration (Price *et al.*, 2013).

Synthetic metabolons

The previous discussion indicates that while a number of metabolons have been proposed, the evidence that they function in improving metabolic flux is scarce. Perhaps the most direct way to determine the utility of metabolons is to introduce them into cells and assess their effect on flux or product formation. The possibility of constructing synthetic enzymes complexes has been discussed in a number of recent reviews (Boyle and Silver, 2012; Lee *et al.*, 2012). While there are a number of ways that synthetic enzyme complexes could be assembled, the first approach to have been employed is to assemble enzymes onto a synthetic scaffold protein using protein interaction domains. This makes use of the reasonably high affinity and specific interactions between protein interaction modules and their peptide ligands derived from various scaffolding and signalling systems. A gene encoding a synthetic scaffold protein containing the protein interacting domains separated by spacer sequences in the required order and stoichiometry is synthesised. The enzymes are tagged with the corresponding peptide ligands. This approach has been used to scaffold three enzymes of the mevalonic acid biosynthesis pathway in *E. coli* (Dueber *et al.*, 2009). A synthetic scaffold protein was

constructed carrying three high affinity protein interaction domains: the Src homology domain 3 (SH3) from the mouse Crk adapter protein (Wu *et al.*, 1995) the PDZ domain from the α-syntrophin protein from mouse (Schultz *et al.*, 1998) and the GTPase-binding domain (GBD) from the neuronal Wiskott-Aldrich syndrome protein (N-WASP) (Kim *et al.*, 2000). The three enzymes were tagged with their cognate peptide ligands. When expressed together, the ligands reversibly bind to the scaffold protein and form a complex. Scaffolding increased product titre 77-fold compared to unscaffolded enzymes- and with lower overall protein expression (HMG-CoA intermediate is toxic) (Dueber *et al*). Similar approaches with a 3 step biosynthesis pathway for glucaric acid from glucose in *E. coli* yielded a 5 fold increase in yield of product (Moon *et al.*, 2010). Similarly, in yeast, scaffolding 4-coumarate:CoA ligase and stilbene synthase increased the yield of resveratrol by 5-fold compared to unscaffolded enzymes (Wang and Yu, 2012). For this two enzyme pathway, a fusion protein was also tested but was less effective than scaffolded enzymes perhaps because of folding problems (Zhang *et al.*, 2006).

Scaffolding has also been shown to work *in vitro*. In this case, triose phosphate isomerase, aldolase and fructose 1,6-bisphosphatase where scaffolded *in vitro* resulting in a 38 to 48-fold increase in reaction rate for scaffolds free in solution or bound to cellulose respectively (You and Percival Zhang, 2013). This example used a system based on the cellulosome complex from various cellulose degrading bacteria. The cellulosome consists of a scaffoldin protein containing dockerin domains to which proteins containing cohesin domains bind. In this example, the enzymes were tagged with three dockerin domains and the scaffold was constructed from the three respective cohesin domains with an N-terminal carbohydrate binding domain. The three dockerin/cohesin domains originated from three different cellulolytic microorganisms.

The factors underpinning increased product titre of these synthetic scaffold complexes have been investigated. Variation in linker length and protein positioning affect yield and support the conclusion that increased enzyme proximity is a key factor (Lee *et al.*, 2012). Also, it is evident that enzyme stoichiometry within the metabolons is important, revealing an issue for conventional metabolic engineering where enzyme over-expression levels are often not precisely controlled. Although the experiments with synthetic metabolons reveal an advantage, the explanation for this is not obvious. It has been argued that the product of an enzyme is likely to diffuse away very quickly along its concentration gradient because typical enzyme reaction rates are much slower than diffusion (Sweetlove and Fernie, 2013) so that proximity might not provide significantly increased substrate concentration. The actual diffusion rates in the cytoplasm are therefore important. Evidence from the mobility of fluorescently labelled glucose in *E. coli* cells suggests that it diffuses ~10 times more slowly *in vivo* (50 $\mu m^2 s^{-1}$) than in water (400 $\mu m^2 s^{-1}$), reflecting the intense macromolecular crowding of the cytoplasm (Mika *et al.*, 2010). Proteins, being larger, diffuse more slowly: 3-10 $\mu m^2 s^{-1}$ in the cytoplasm. Evidence for the existence of diffusion gradients would be difficult to produce if enzymes are evenly distributed in the cytoplasm but examples where a small molecule is produced in a localised manner suggest that concentration gradients can exist. Firstly, it is well-known that localised release of Ca^{2+} within cells leads to localised high concentration microdomains (Berridge, 2006). Secondly, hydrogen peroxide is produced locally in some cells by plasma membrane NADPH oxidase, an example being growing pollen tubes where a gradient of hydrogen peroxide forms. Visualised by a ROS sensitive fluorescent probe, hydrogen peroxide concentration is highest at the tip (Potocký *et al.*, 2007). Therefore, relative to the rate of production and breakdown, diffusion is apparently not fast enough to equilibrate

hydrogen peroxide concentration throughout the cell. It is also likely that the diffusion rate of small molecules will depend on their chemical properties (for example propensity to interact with proteins) as well as size. Interestingly, in a modification of the *in vitro* scaffolding of triose phosphate isomerase, aldolase and fructose 1,6-bisphosphatase, replacement of the original aldolase with another of lower specific activity removed the advantage of scaffolding (You and Zhang, 2013). This observation shows that the specifics of enzyme kinetics are critical. Modelling of the system is needed to clarify the important factors, particularly to determine if increased local substrate concentration is generally important. Because of relatively rapid diffusion, the size of the complex may also be critical: bigger complexes are predicted to capture and use more of the intermediates (Lee *et al.*, 2012). These authors suggested 100 nm would be ideal for the scaffolded mevalonate pathway. In this system, since some of the enzymes are multimeric, they could bind to multiple scaffolds producing a larger complex. Another potentially important effect could be control of pathway branch points and competition between different pathways for the same intermediates. In the latter case, a higher proportion of the product of a scaffolded enzyme might be passed on to its neighbour than to an unscaffolded competitor. Scaffolding also provides a potential means to control branch points by manipulating protein-protein interactions by signal molecules. As an example, one can envisage that binding sites for pathway products or precursors could be included in a synthetic scaffold such that binding causes conformational changes that alter the distance between the scaffolded enzymes or cause dissociation of the complex. This would be analogous to FRET-based glucose sensors (Jones *et al.*, 2013). There is already evidence that metabolic status can affect the conformation of the ER-bound flavonoid biosynthesis metabolons in plants, potentially directing intermediates down one branch of the pathway or another (Crosby *et al.*, 2011). It is very unlikely that any of the current examples have

289 introduced the more extreme form of substrate channelling shown by enzyme complexes such
290 as tryptophan synthetase where intermediates pass through tunnels lined with amino acid
291 residues of appropriate charge and never enter the external water. Introduction of direct
292 channelling would require more sophisticated protein engineering to create and orientate
293 channels. These examples provide evidence that synthetic scaffolds can improve flux and
294 product formation even though the mechanisms are not entirely clear and even counter-
295 intuitive given the fast diffusion rates of small molecules (Lee *et al.*, 2012; Sweetlove and
296 Fernie, 2013). Nevertheless, the results provide a rationale for the natural occurrence of
297 metabolons.
298
299 Since proteins can bind to nucleic acids with high affinity and in a nucleotide sequence-specific
300 manner (e.g. transcription factors), they could also act as scaffolds for enzymes. This has been
301 demonstrated in *E. coli* for introduced resveratrol, 1,2-propanediol and mevalonic acid
302 biosynthetic pathways (Conrado *et al.*, 2012). In this example, each enzyme was fused with a
303 DNA binding zinc finger domain. These were expressed in *E. coli* along with a plasmid
304 containing multiple copies of each Zn finger binding nucleotide sequence. Evidence that the
305 enzymes assembled on the plasmid was obtained and, in each case, product formation was
306 increased over a random scaffold control. Interestingly, the increase in mevalonic acid was
307 much less than the 77-fold reported for protein scaffolds (Dueber *et al.*, 2009). Problems with
308 plasmids will include sufficient copy number to scaffold available protein and supercoiling. This
309 method is currently not applicable to plants but assembly on RNA scaffolds is a possibility. RNA
310 molecules (aptamers) that fold into specific conformations and with sequences that can
311 specifically bind small molecules and proteins could be used as scaffolds (Delebecque *et al.*,
312 2011). Application of this approach to hydrogen production in *E. coli* using ferredoxin and

hydrogenase tagged with adapter proteins that bind to specific RNA aptamers resulted in stable scaffolded protein/RNA complexes and increased the rate of hydrogen production by up to 48 fold compared to unscaffolded. Clearly, in this case, electron transfer from reduced ferredoxin to hydrogenase requires very close proximity, but as with protein scaffolds, there is no reason to suppose this approach would not work for other pathways. A potential problem with expression of RNA in sufficient amounts is stability. However, advantages are the complex 3D structures that RNA can form and the possibility of designing specific protein binding aptamers. This could extend to the possibility of adding riboswitch-like control mechanisms in which small molecules also bind to the RNA scaffold and change its conformation. This could change the distance between scaffolded enzymes and modulate flux according to the concentration of an "allosteric" controlling molecule.

Bacterial microcompartments (BMCs) provide inspiration for another approach to creating synthetic metabolons in which enzymes are encapsulated in self assembling protein shells (Frank *et al.*, 2013; Retterer and Simpson, 2012). This provides a physical diffusion barrier to pathway intermediates thus improving local concentration of intermediates, but has the requirement that the shell proteins provide pores for substrate to enter and product to leave. Although the structure of some BMCs is complex, recent work shows that they can be assembled in host bacteria and protein cargo can be encapsulated. Thus, the shell proteins from the propanediol utilization BMC from *Citrobacter freundii* can be expressed in *E. coli* to form empty structures (Parsons *et al.*, 2010). Expression of a carboxysome operon from *Halothiobacillus neapolitanus,* including pore protein, coat protein, CA and Rubisco in *E. coli* produced functional carboxysomes (Bonacci *et al.*, 2012). Significantly, the *Salmonella enterica* ethanolamine utilization (*eut*) shell protein expressed in *E. coli* assembles to form a polyhedral

structure. These structures could be loaded with enhanced green fluorescent protein fused to an appropriate N-terminal signal sequence (Choudhary *et al.*, 2012). Therefore, in principle, it should be possible to encapsulate multiple enzymes in synthetic BMCs. Utility in metabolic engineering will depend not only on successful assembly and targeting of enzymes but also on understanding how to include entry and exit pores for substrate and product. Clearly, the way is open to attempt assembly of microcompartments in plant cells and the introduction of carboxysome-like compartments into chloroplasts to improve carboxylation efficiency is a potential target. Viral coat proteins offer another source of self assembling microcompartments. Empty cowpea mosaic virus particle (termed empty virus-like particles, eVLPs) can be assembled in plants by expressing the virus coat precursor protein and a protease that cleaves it into the final form (Saunders and Lomonossoff, 2013). Other approaches to assembling synthetic metabolons could include targeting enzymes to the cytoskeleton (Araiza-Olivera *et al.*, 2013) or using membrane anchored proteins as in the naturally-occurring flavonoid and glycolytic metabolons (Crosby *et al.*, 2011; Graham *et al.*, 2007) that tether the pathway enzymes. In this way metabolons could be targeted to specific membranes perhaps directing transport of products into specific subcellular compartments (e.g. the vacuole for storage of potentially toxic products). Positioning might also allow a pathway to efficiently access substrates emanating from an organelle.

Microcompartmentation and localisation of signalling complexes is achieved by anchoring complexes to lipid rafts. These are specialised membrane microdomains that are rich in cholesterol and sphingolipids into which GPI-anchored proteins insert thereby fostering close proximity of proteins (Simons and Ikonen, 1997). The NADPH oxidase (NOX) complex which includes NOX and various regulatory proteins that control the oxidative burst is an example from animals (Shao *et al.*, 2003) that may also operate in plants (Borner *et al.*, 2003).

Looking to the future: synthetic metabolons in plants.

The potential advantages of metabolons and microcompartments include increased flux (or less enzyme needed to maintain a given flux), containment of potentially toxic or biologically active intermediates, control over enzyme stoichiometry and the ability to control branchpoints and thereby direct metabolism. Based on the experience of the introduction of metabolons into microorganisms, it is clear that a similar strategy should be considered as part of the plant metabolic engineering toolbox. Since secondary metabolism provides clear examples of enzyme complexes (Jørgensen *et al.*, 2005) it is likely that metabolons will be of particular value in production of exotic compounds in plants. Pathways that have exotic or reactive intermediates may benefit from channelling. In the future, metabolons could incorporate switching mechanisms as proposed for flavonoid biosynthesis (Crosby *et al.*, 2011) in which metabolic status is sensed causing association or disassociation of the complexes (Michener *et al.*, 2012). It is interesting that, as well as associating into complexes, the enzymes of the flavonoid biosynthesis pathway are under tight transcriptional control suggesting that control of enzyme concentrations well as complex formation is important in this pathway. The appearance of an array of genetically-encoded fluorescent metabolite sensors also opens the way to monitoring metabolism in cells in "real time" and, in conjunction with flux analysis using labelled substrates, will allow improved understanding of metabolic control and bottlenecks (Berg *et al.*, 2009; Hung *et al.*, 2011; Jones *et al.*, 2013; Michener *et al.*, 2012; Tantama *et al.*, 2011).

The cyanobacterial carboxysome (and algal pyrenoids) are archetypal metabolons which bring about CO_2 concentration at the active site of RuBisCO by bringing carbonic anhydrase (CA) and RuBisCO into close proximity along with a supply of bicarbonate. Therefore, one possible way to improve photosynthesis would be to introduce mechanisms inspired by these systems into C_3 crop plants (Price *et al.*, 2013). This approach could include introduction of bicarbonate pumps into the chloroplast envelope and arrangement of RuBisCO and CA into scaffolded complexes or even into synthetic carboxysomes. The latter approach should be possible given the recent demonstrations that simple microcompartments can assemble in a host cell and that protein cargo can be included (Choudhary *et al.*, 2012; Lee *et al.*, 2012). The key trick is to ensure that the shell proteins contain suitable pores for transport of RuBP and bicarbonate into the microcompartment and 3-PGA out while trapping the released CO_2 effectively. An alternative strategy is to explore the utility of scaffolding RuBisCO and CA so that CO_2 is released in close proximity to RuBisCO, thus decreasing oxygenase activity (Singleton, Harmer, Porter and Smirnoff, unpublished results; http://magic.psrg.org.uk). This approach has theoretical problems because the inherently slow catalytic rate of RuBisCO prevents effective use of CO_2 before it diffuses away. Other synthetic biology approaches to improving photosynthesis such as expression of carboxysome-like structures in chloroplasts may also emerge in the near future.

Acknowledgements

Funding from the Biotechnology and Biological Research Council (BB/I020004/1 and BB/J004057/1) is gratefully acknowledged.

References

Achnine L, Blancaflor EB, Rasmussen S, Dixon RA. 2004. Colocalization of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis. *Plant Cell* **16**, 3098-3109.

Anderson LE, Chrostowski J, Carol AA. 2006. Enzyme co-localization with transketolase, xylulose-5-P 3-epimerase and phosphoriboisomerase in pea leaf chloroplasts. *Plant Science* **171**, 686-698.

Arabidopsis Interactome Mapping Consortium. 2011. Evidence for network evolution in an Arabidopsis interactome map. *Science* **333**, 601-607.

Araiza-Olivera D, Chiquete-Felix N, Rosas-Lemus M, Sampedro JG, Peña A, Mujica A, Uribe-Carvajal S. 2013. A glycolytic metabolon in *Saccharomyces cerevisiae* is stabilized by F-actin. *The FEBS Journal* **280**, 3887-3905.

Bassard J-E, Richert L, Geerinck J, Renault H, Duval F, Ullmann P, Schmitt M, Meyer E, Mutterer J, Boerjan W, De Jaeger G, Mely Y, Goossens A, Werck-Reichhart D. 2012. Protein-protein and protein-membrane associations in the lignin pathway. *Plant Cell* **24**, 4465-4482.

Bauwe H, Hagemann M, Kern R, Timm S. 2012. Photorespiration has a dual origin and manifold links to central metabolism. *Current Opinion in Plant Biology* **15**, 269-275.

Beeckmans S, Khan AS, Van Driessche E, Kanarek L. 1994. A specific association between the glyoxylic-acid-cycle enzymes isocitrate lyase and malate synthase. *European Journal of Biochemistry* **224**, 197-201.

Berg J, Hung YP, Yellen G. 2009. A genetically encoded fluorescent reporter of ATP:ADP ratio. *Nature Methods* **6**, 161-166.

Bernard A, Domergue F, Pascal S, Jetter R, Renne C, Faure J-D, Haslam RP, Napier JA, Lessire R, Joubès J. 2012. Reconstitution of plant alkane biosynthesis in yeast demonstrates that Arabidopsis ECERIFERUM1 and ECERIFERUM3 are core components of a very-long-chain alkane synthesis complex. *Plant Cell* **24**, 3106-3118.

Berridge MJ. 2006. Calcium microdomains: organization and function. *Cell Calcium* **40**, 405-412.

Bonacci W, Teng PK, Afonso B, Niederholtmeyer H, Grob P, Silver PA, Savage DF. 2012. Modularity of a carbon-fixing protein organelle. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 478-483.

Borner GHH, Lilley KS, Stevens TJ, Dupree P. 2003. Identification of glycosylphosphatidylinositol-anchored proteins in Arabidopsis. A proteomic and genomic analysis. *Plant Physiology* **132**, 568-577.

Boyle PM, Silver PA. 2012. Parts plus pipes: synthetic biology approaches to metabolic engineering. *Metabolic Engineering* **14**, 223-232.

Chen AH, Silver PA. 2012. Designing biological compartmentalization. *Trends in Cell Biology* **22**, 662-670.

Choudhary S, Quin MB, Sanders Ma, Johnson ET, Schmidt-Dannert C. 2012. Engineered protein nano-compartments for targeted enzyme localization. *PLoS ONE* **7**, e33342-e33342.

Conrado RJ, Wu GC, Boock JT, Xu H, Chen SY, Lebar T, Turnšek J, Tomšič N, Avbelj M, Gaber R, Koprivnjak T, Mori J, Glavnik V, Vovk I, Benčina M, Hodnik V, Anderluh G, Dueber JE, Jerala R, DeLisa MP. 2012. DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency. *Nucleic acids research* **40**, 1879-1889.

Crosby KC, Pietraszewska-Bogiel A, Gadella TWJ, Winkel BSJ. 2011. Förster resonance energy transfer demonstrates a flavonoid metabolon in living plant cells that displays competitive interactions between enzymes. *FEBS Letters* **585**, 2193-2198.

de Cima S, Gil-Ortiz F, Crabeel M, Fita I, Rubio V. 2012. Insight on an arginine synthesis metabolon from the tetrameric structure of yeast acetylglutamate kinase. *PLoS ONE* **7**, e34734-e34734.

Delebecque CJ, Lindner AB, Silver PA, Aldaye FA. 2011. Organization of intracellular reactions with rationally designed RNA assemblies. *Science (New York, N.Y.)* **333**, 470-474.

DellaPenna D. 2001. Plant metabolic engineering. *Plant Physiology* **125**, 160-163.

Dueber JE, Wu GC, Malmirchegini GR, Moon TS, Petzold CJ, Ullal AV, Prather KLJ, Keasling JD. 2009. Synthetic protein scaffolds provide modular control over metabolic flux. *Nature biotechnology* **27**, 753-759.

Frank S, Lawrence AD, Prentice MB, Warren MJ. 2013. Bacterial microcompartments moving into a synthetic biological world. *Journal of Biotechnology* **163**, 273-279.

Freemont RS, Kitney RI. 2012. *Synthetic Biology- A Primer*. London: Imperial College Press.

Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, Benders GA, Montague MG, Ma L, Moodie MM, Merryman C, Vashee S, Krishnakumar R, Assad-Garcia N, Andrews-Pfannkoch C, Denisova EA, Young L, Qi ZQ, Segall-Shapiro TH, Calvey CH, Parmar PP, Hutchison CA, Smith HO, Venter JC. 2010. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**, 52-56.

Giege P, Heazlewood JL, Roessner-Tunali U, Millar AH, Fernie AR, Leaver CJ, Sweetlove LJ. 2003. Enzymes of glycolysis are functionally associated with the mitochondrion in Arabidopsis cells. *Plant Cell* **15**, 2140-2151.

Graham JWA, Williams TCR, Morgan M, Fernie AR, Ratcliffe RG, Sweetlove LJ. 2007. Glycolytic enzymes associate dynamically with mitochondria in response to respiratory demand and support substrate channeling. *Plant Cell* **19**, 3723-3738.

Howard TP, Metodiev M, Lloyd JC, Raines CA. 2008. Thioredoxin-mediated reversible dissociation of a stromal multiprotein complex in response to changes in light availability. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 4056-4061.

Hung YP, Albeck JG, Tantama M, Yellen G. 2011. Imaging cytosolic NADH-NAD^+ redox state with a genetically encoded fluorescent biosensor. *Cell metabolism* **14**, 545-554.

Hutson SM, Islam MM, Zaganas I. 2011. Interaction between glutamate dehydrogenase (GDH) and L-leucine catabolic enzymes: intersecting metabolic pathways. *Neurochemistry international* **59**, 518-524.

Hyeon JE, Jeon SD, Han SO. 2013. Cellulosome-based, Clostridium-derived multi-functional enzyme complexes for advanced biotechnology tool development: Advances and applications. *Biotechnology Advances* **31**, 936-944.

Islam MM, Wallin R, Wynn RM, Conway M, Fujii H, Mobley JA, Chuang DT, Hutson SM. 2007. A novel branched-chain amino acid metabolon. Protein-protein interactions in a supramolecular complex. *Journal of Biological Chemistry* **282**, 11893-11903.

Jiang W, Napoli JL. 2013. The retinol dehydrogenase Rdh10 localizes to lipid droplets during acyl ester biosynthesis. *Journal of Biological Chemistry* **288**, 589-597.

Jones AM, Grossmann G, Danielson JAH, Sosso D, Chen LQ, Ho CH, Frommer WB. 2013. *In vivo* biochemistry: applications for small molecule biosensors in plant biology. *Current Opinion in Plant Biology* **16**, 389-395.

Jørgensen K, Rasmussen AV, Morant M, Nielsen AH, Bjarnholt N, Zagrobelny M, Bak S, Møller BL. 2005. Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Current Opinion in Plant Biology* **8**, 280-291.

Kerfeld CA, Heinhorst S, Cannon GC. 2010. Bacterial microcompartments. *Annual Review of Microbiology* **64**, 391-408.

Khosla C, Gokhale RS, Jacobsen JR, Cane DE. 1999. Tolerance and specificity of polyketide synthases. *Annual Review of Biochemistry* **68**, 219-253.
Kim AS, Kakalis LT, Abdul-Manan M, Liu GA, Rosen MK. 2000. Autoinhibition and activation mechanisms of the Wiskott-Aldrich syndrome protein. *Nature* **404**, 151-158.
Lallemand B, Erhardt M, Heitz T, Legrand M. 2013. Sporopollenin biosynthetic enzymes interact and constitute a metabolon localized to the endoplasmic reticulum of tapetum cells. *Plant Physiology* **162**, 616-625.
Lee H, DeLoache WC, Dueber JE. 2012. Spatial organization of enzymes for metabolic engineering. *Metabolic Engineering* **14**, 242-251.
Lundquist PK, Poliakov A, Bhuiyan NH, Zybailov B, Sun Q, van Wijk KJ. 2012. The functional network of the Arabidopsis plastoglobule proteome based on quantitative proteomics and genome-wide coexpression analysis. *Plant Physiology* **158**, 1172-1192.
Marri L, Zaffagnini M, Collin V, Issakidis-Bourguet E, Lemaire SD, Pupillo P, Sparla F, Miginiac-Maslow M, Trost P. 2009. Prompt and easy activation by specific thioredoxins of Calvin cycle enzymes of *Arabidopsis thaliana* associated in the GAPDH/CP12/PRK supramolecular complex. *Molecular Plant* **2**, 259-269.
Maurino VG, Peterhansel C. 2010. Photorespiration: current status and approaches for metabolic engineering. *Current Opinion in Plant Biology* **13**, 249-256.
Mayer JE, Pfeiffer WH, Beyer P. 2008. Biofortified crops to alleviate micronutrient malnutrition. *Current Opinion in Plant Biology* **11**, 166-170.
Meyer FM, Gerwig J, Hammer E, Herzberg C, Commichau FM, Völker U, Stülke J. 2011. Physical interactions between tricarboxylic acid cycle enzymes in *Bacillus subtilis*: evidence for a metabolon. *Metabolic Engineering* **13**, 18-27.
Michener JK, Thodey K, Liang JC, Smolke CD. 2012. Applications of genetically-encoded biosensors for the construction and control of biosynthetic pathways. *Metabolic Engineering* **14**, 212-222.
Mika JT, van den Bogaart G, Veenhoff L, Krasnikov V, Poolman B. 2010. Molecular sieving properties of the cytoplasm of Escherichia coli and consequences of osmotic stress. *Molecular Microbiology* **77**, 200-207.
Mitchell CG. 1996. Identification of a multienzyme complex of the tricarboxylic acid cycle enzymes containing citrate synthase isoenzymes from *Pseudomonas aeruginosa*. *Biochemical Journal* **313**, 769-774.
Moller BL. 2010. Dynamic Metabolons. *Science* **330**, 1328-1329.
Moon TS, Dueber JE, Shiue E, Prather KLJ. 2010. Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered *E. coli*. *Metabolic Engineering* **12**, 298-305.
Müller A, Weiler EW. 2000. IAA-synthase, an enzyme complex from Arabidopsis thaliana catalyzing the formation of indole-3-acetic acid from (S)-tryptophan. *Biological Chemistry* **381**, 679-686.
Nacir H, Bréhélin C. 2013. When proteomics reveals unsuspected roles: the plastoglobule example. *Frontiers in Plant Science* **4**, 114-114.
Napier JA, Graham IA. 2010. Tailoring plant lipid composition: designer oilseeds come of age. *Current Opinion in Plant Biology* **13**, 330-337.
Naqvi S, Zhu CF, Farre G, Ramessar K, Bassie L, Breitenbach J, Conesa DP, Ros G, Sandmann G, Capell T, Christou P. 2009. Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 7762-7767.
Nielsen KA, Tattersall DB, Jones PR, Møller BL. 2008. Metabolon formation in dhurrin biosynthesis. *Phytochemistry* **69**, 88-98.

Ohad N, Shichrur K, Yalovsky S. 2007. The analysis of protein-protein interactions in plants by bimolecular fluorescence complementation. *Plant Physiology* **145**, 1090-1099.

Panicot M, Minguet EG, Ferrando A, Alcazar R, Blazquez MA, Carbonell J, Altabella T, Koncz C, Tiburcio A. 2002. A polyamine metabolon involving aminopropyl transferase complexes in Arabidopsis. *Plant Cell* **14**, 2539-2551.

Parsons JB, Frank S, Bhella D, Liang M, Prentice MB, Mulvihill DP, Warren MJ. 2010. Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. *Molecular cell* **38**, 305-315.

Potocký M, Jones Ma, Bezvoda R, Smirnoff N, Zárský V. 2007. Reactive oxygen species produced by NADPH oxidase are involved in pollen tube growth. *The New Phytologist* **174**, 742-751.

Price GD, Pengelly JJL, Forster B, Du JH, Whitney SM, von Caemmerer S, Badger MR, Howitt SM, Evans JR. 2013. The cyanobacterial CCM as a source of genes for improving photosynthetic CO_2 fixation in crop species. *Journal of Experimental Botany* **64**, 753-768.

Puchulu-Campanella E, Chu H, Anstee DJ, Galan JA, Tao WA, Low PS. 2013. Identification of the components of a glycolytic enzyme metabolon on the human red blood cell membrane. *Journal of Biological Chemistry* **288**, 848-858.

Rae BD, Long BM, Whitehead LF, Forster B, Badger MR, Price GD. 2013. Cyanobacterial carboxysomes: microcompartments that facilitate CO_2 fixation. *Journal of Molecular Microbiology and Biotechnology* **23**, 300-307.

Rasmussen S, Dixon RA. 1999. Transgene-mediated and elicitor-induced perturbation of metabolic channeling at the entry point into the phenylpropanoid pathway. *Plant Cell* **11**, 1537-1551.

Retterer ST, Simpson ML. 2012. Microscale and nanoscale compartments for biotechnology. *Current Opinion in Biotechnology* **23**, 522-528.

Ruiz-Lopez N, Haslam RP, Napier JA, Sayanova O. 2014. Successful high-level accumulation of fish oil omega-3 long-chain polyunsaturated fatty acids in a transgenic oilseed crop. *Plant Journal* **10.1111/tpj.12378**.

Saunders K, Lomonossoff GP. 2013. Exploiting plant virus-derived components to achieve *in planta* expression and for templates for synthetic biology applications. *New Phytologist* **200**, 16-26.

Schultz J, Hoffmuller U, Krause G, Ashurst J, Macias MJ, Schmieder P, Schneider-Mergener J, Oschkinat H. 1998. Specific interactions between the syntrophin PDZ domain and voltage-gated sodium channels. *Nature Structural Biology* **5**, 19-24.

Shao DM, Segal AW, Dekker LV. 2003. Lipid rafts determine efficiency of NADPH oxidase activation in neutrophils. *FEBS Letters* **550**, 101-106.

Simons K, Ikonen E. 1997. Functional rafts in cell membranes. *Nature* **387**, 569-572.

Srere PA. 1987. Complexes of sequential metabolic enzymes. *Annual Review of Biochemistry* **56**, 89-124.

Srere PA. 2000. Macromolecular interactions: tracing the roots. *Trends in Biochemical Sciences* **25**, 150-153.

Sugumaran M, Nellaiappan K, Amaratunga C, Cardinale S, Scott T. 2000. Insect melanogenesis. III. Metabolon formation in the melanogenic pathway-regulation of phenoloxidase activityy by endogenous dopachrome isomerase (decarboxylating) from *Manduca sexta*. *Archives of Biochemistry and Biophysics* **378**, 393-403.

Suss KH, Arkona C, Manteuffel R, Adler K. 1993. Calvin cycle multienzyme complexes are bound to chloroplast thylakoid membranes of higher plants *in situ*. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 5514-5518.

Sweetlove LJ, Fernie AR. 2013. The spatial organization of metabolism within the plant cell. *Annual Review of Plant Biology* **64**, 723-746.
Szecowka M, Heise R, Tohge T, Nunes-Nesi A, Vosloh D, Huege J, Feil R, Lunn J, Nikoloski Z, Stitt M, Fernie AR, Arrivault S. 2013. Metabolic fluxes in an illuminated Arabidopsis rosette. *Plant Cell* **25**, 694-714.
Tantama M, Hung YP, Yellen G. 2011. Imaging intracellular pH in live cells with a genetically encoded red fluorescent protein sensor. *Journal of the American Chemical Society* **133**, 10034-10037.
Velot C, Mixon MB, Teige M, Srere PA. 1997. Model of a quinary structure between Krebs TCA cycle enzymes: a model for the metabolon. *Biochemistry* **36**, 14271-14276.
Voet D, Voet JD. 2004. *Biochemistry*: John Wiley and Sons.
Wang Y, Yu O. 2012. Synthetic scaffolds increased resveratrol biosynthesis in engineered yeast cells. *Journal of Biotechnology* **157**, 258-260.
Wedel N, Soll J. 1998. Evolutionary conserved light regulation of Calvin cycle activity by NADPH-mediated reversible phosphoribulokinase/CP12/glyceraldehyde-3-phosphate dehydrogenase complex dissociation. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 9699-9704.
Wedel N, Soll J, Paap BK. 1997. CP12 provides a new mode of light regulation of Calvin cycle activity in higher plants. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 10479-10484.
Winkel BSJ. 2004. Metabolic channeling in plants. *Annual Review of Plant Biology* **55**, 85-107.
Wu XD, Knudsen B, Feller SM, Zheng J, Sali A, Cowburn D, Hanafusa H, Kuriyan J. 1995. Structural basis for the specific interaction of lysine-containing proline-rich peptides with the N-terminal SH3 domain of C-CRK. *Structure* **3**, 215-226.
Ye XD, Al-Babili S, Kloti A, Zhang J, Lucca P, Beyer P, Potrykus I. 2000. Engineering the provitamin A (β-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* **287**, 303-305.
You C, Percival Zhang YH. 2013. Self-assembly of synthetic metabolons through synthetic protein scaffolds: one-step purification, co-immobilization, and substrate channeling. *ACS Synthetic Biology* **2**, 102-110.
You C, Zhang YHP. 2013. Annexation of a high-activity enzyme in a synthetic three-enzyme complex greatly decreases the degree of substrate channeling. *ACS Synthetic Biology*.
Zhang FZ, Carothers JM, Keasling JD. 2012. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nature biotechnology* **30**, 354-U166.
Zhang Y, Gao P, Yuan JS. 2010. Plant protein-protein interaction network and interactome. *Current genomics* **11**, 40-46.
Zhang Y, Li S-Z, Li J, Pan X, Cahoon RE, Jaworski JG, Wang X, Jez JM, Chen F, Yu O. 2006. Using unnatural protein fusions to engineer resveratrol biosynthesis in yeast and mammalian cells. *Journal of the American Chemical Society* **128**, 13030-13031.

Table 1. Evidence for metabolons and metabolic channeling in plants. The nature of the evidence is as follows: A, immunogold localisation; B, physical association (co-purification/pulldown/yeast 2-hybrid); C, co-localisation of fluorescently tagged proteins *in vivo* (e.g. FRET/BIFC); D, flux measurements.

Pathway	**Evidence**	**References**
Glycolysis	B, C, D	(Giege *et al.*, 2003; Graham *et al.*, 2007)
Calvin-Benson cycle	A, D	(Anderson *et al.*, 2006; Suss *et al.*, 1993; Szecowka *et al.*, 2013)
Phenylpropanoids	B, C, D	(Achnine *et al.*, 2004; Bassard *et al.*, 2012; Rasmussen and Dixon, 1999)
Flavonoids	B,C	(Crosby *et al.*, 2011)
Spermine/spermidine	B	(Panicot *et al.*, 2002)
Indole acetic acid	B	(Müller and Weiler, 2000)
Long chain alkanes	B, C	(Bernard *et al.*, 2012)
Sproropollenin	B, C	(Lallemand *et al.*, 2013)
Glyoxylate cycle	B	(Beeckmans *et al.*, 1994)
Dhurrin (cyanogenic glycoside)	B, C	(Nielsen *et al.*, 2008)