Enzyme-fusion strategies for redirecting and improving carotenoid synthesis in S. cerevisiae

Hery Rabeharindranto, Sara Castaño-Cerezo, Thomas Lautier, Luis Garcia-Alles, Christian Treitz, Andreas Tholey, Gilles Truan

Enzyme-fusion strategies for redirecting and improving carotenoid synthesis in *S. cerevisiae*

Hery Rabeharindranto[a$], Sara Castaño-Cerezo[a$], Thomas Lautier[a], Luis F. Garcia-Alles[a], Christian Treitz[b], Andreas Tholey[b], Gilles Truan[a*]

[a] LISBP, Université de Toulouse, CNRS, INRA, INSA, Toulouse, France.

[b] Systematic Proteome Research and Bioanalytics, Institute for Experimental Medicine, Christian-Albrechts-Universität zu Kiel, Kiel, Germany.

[$] These authors contributed equally to this work

*** Corresponding author:** gilles.truan@insa-toulouse.fr; LISBP, 135 Avenue de Rangueil, 31077 Toulouse Cedex 04, FRANCE.

1 Abstract

Spatial clustering of enzymes has proven an elegant approach to optimize metabolite transfer between enzymes in synthetic metabolic pathways. Among the multiple methods used to promote colocalisation, enzyme fusion is probably the simplest. Inspired by natural systems, we have explored the metabolic consequences of spatial reorganizations of the catalytic domains of *Xanthophyllomyces dendrorhous* carotenoid enzymes produced in *Saccharomyces cerevisiae*. Synthetic genes encoding bidomain enzymes composed of CrtI and CrtB domains from the natural CrtYB fusion were connected in the two possible orientations, using natural and synthetic linkers. A tridomain enzyme (CrtB, CrtI, CrtY) harboring the full β-carotene producing pathway was also constructed. Our results demonstrate that domain order and linker properties considerably impact both the expression and/or stability of the constructed proteins and the functionality of the catalytic domains, all concurring to either diminish or boost specific enzymatic steps of the metabolic pathway. Remarkably, the yield of β-carotene production doubled with the tridomain fusion while precursor accumulation decreased, leading to an improvement of the pathway efficiency, when compared to the natural system. Our data strengthen the idea that fusion of enzymatic domains is an appropriate technique not only to achieve spatial confinement and enhance the metabolic flux but also to produce molecules not easily attainable with natural enzymatic configurations, even with membrane bound enzymes.

Keywords: Metabolic engineering, synthetic biology, metabolic flux, enzyme spatial proximity, carotenoids, multidomain enzymes.

2 Introduction

Metabolic engineering aims to design and implement metabolic pathways to produce high value compounds or novel molecules that would otherwise be more expensive or difficult to attain following chemical organic synthesis (Chen and Nielsen, 2013). However, product yields in engineered microorganisms are often low. It is therefore important to devise strategies optimizing the efficiency of both individual enzymatic steps and the global synthetic pathway (Siddiqui et al., 2012). A major drawback originates from the accumulation of metabolic intermediates. This is particularly detrimental to the cell, not only because of the loss of metabolites and non-optimized consumption of resources but also because of the possible consumption of those intermediates by other competitive metabolic pathways, potentially leading to toxicity (Dahl et al., 2013).

Spatial proximity and appropriate disposition of sequential enzymatic domains inside the metabolic pathway are assumed to contribute critically in promoting rapid and efficient consumption of pathway intermediates (Srere, 1985). In natural enzymatic systems, spatial confinement is achieved using various strategies such as the creation of enzymes combining several covalently-linked catalytic domains, the recruitment of several enzymes on protein scaffolds or the colocalization of activities within lipid- or protein-based compartments (Pompon et al., 2017; Schmitt and An, 2017). When applied to metabolic engineering, several studies have demonstrated that the efficiency of a synthetic pathway can be improved via spatial confinement. For example in *E. coli*, a 77-fold improvement in product titer was achieved by promoting artificial interactions between the three first enzymes of the mevalonate pathway, using SH3, PDZ and GBD binding domains as scaffold and their respective ligands as recruiting domains (Dueber et al., 2009). Synthetic RNAs were also used as macromolecular scaffolds to recruit and promote spatial confinement of ferredoxin and hydrogenase enzymes, leading to a 24-fold increase in the hydrogen production compared to the strain expressing free diffusing enzymes (Delebecque et al., 2011). In *Saccharomyces cerevisiae*, a 38% improvement in 2,3-butanediol was achieved by scaffolding enzymes onto cohesin and dockerin domains (Kim et al., 2016)

and a three times increase in the production of geranylgeraniol was detected upon overexpression of a synthetic bidomain enzyme (Tokuhiro et al., 2009). Similarly, by connecting farnesyl pyrophosphate (FPP) synthase with amorphadiene and patchoulol synthases from plant origin, the biotransformation of the intermediate molecule FPP into the desired molecules of interest was improved (Albertsen et al., 2011; Baadhe et al., 2013). Other strategies promoted the use of natural compartment such as mitochondria to improve the production of branched alcohols in yeast (Avalos et al., 2013).

Carotenoids, a particular group of terpene-based molecules, have raised scientific attention due to their potential benefit for human health (Lu et al., 2014) and in several industrial applications (López et al., 2015). Biosynthesis of carotenoids begins with the production of phytoene from geranylgeranyl pyrophosphate (GGPP) by phytoene synthase (CrtB) (Figure 1). Phytoene desaturase (CrtI) then catalyzes four successive desaturation reactions to form lycopene. Finally, lycopene cyclase (CrtY) performs the cyclisation on both sides of lycopene to produce β-carotene. Eukaryotic CrtY enzymes are predicted to be transmembrane proteins. On the contrary, CrtB and possibly CrtI are supposedly cytosolic (Schaub et al., 2012). Importantly, in some fungi, the phytoene synthase and lycopene cyclase catalytic activities are regrouped within a multidomain protein named CrtYB in *Xanthophyllomyces dendrorhous* (Verdoes et al., 1999) and CarRA in *Phycomyces blakesleeanus* (Arrach et al., 2001).

The full β-carotene producing enzymatic system from *X. dendrorhous* has been expressed successfully in *S. cerevisiae* (Verwaal et al., 2007). Similarly, the natural tridomain fusion (CrtIBY) from *Schizotrium* sp. was expressed recently in *Yarrowia lipolytica* (Gao et al., 2017) but the β-carotene production yields were considerably lower than those obtained with the yeast-expressed *X. dendrorhous* configuration (Verwaal et al., 2007; Xie et al., 2015c). Although these heterologous systems, based on *X. dendrorhous enzymes,* are the most productive to date, accumulation of β-carotene precursors in the pathway, such as phytoene, was evidenced. Moreover, attempts to alleviate the presumed metabolic bottleneck by overexpression of CrtI was only partially efficient (Verwaal et al., 2007).

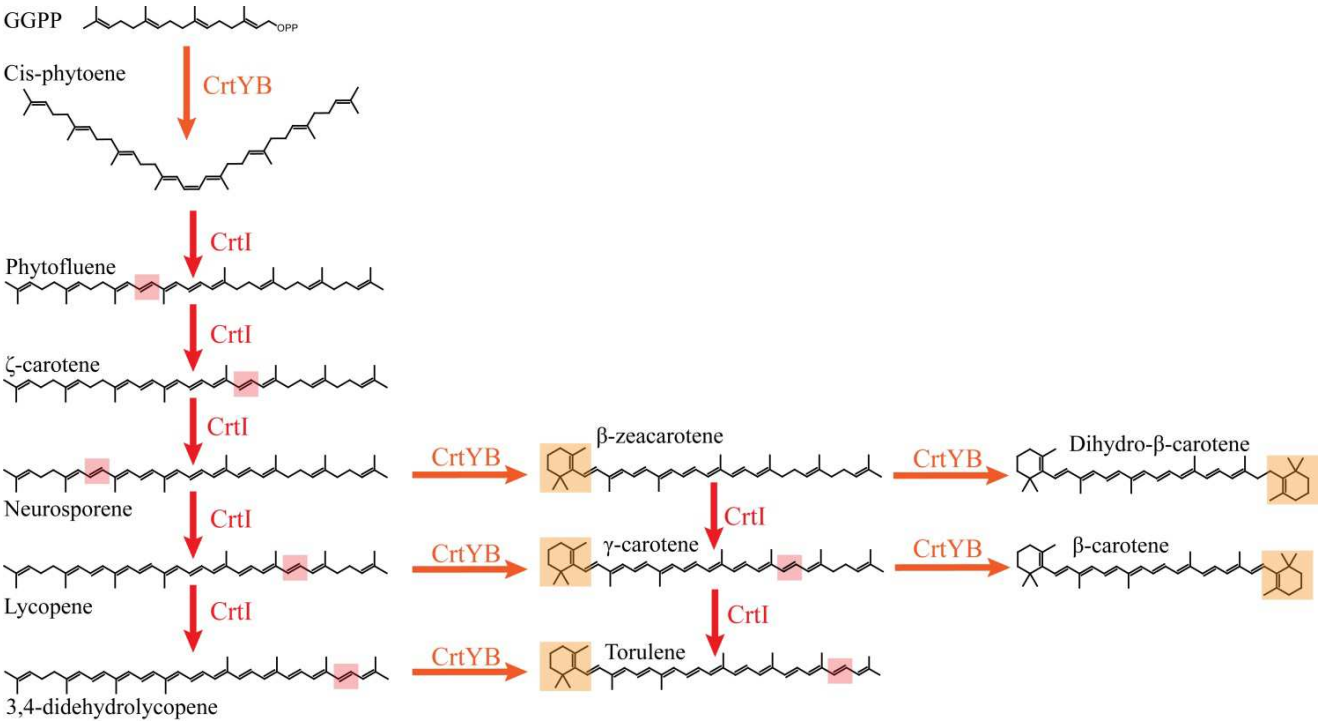

Figure 1. Presumed biosynthetic pathway for the synthesis of β-carotene in *X. dendrorhous* (Verdoes et al., 1999). The bifunctional lycopene cyclase/phytoene synthase (CrtYB) enzyme is depicted in orange, the phytoene desaturase enzyme (CrtI) is depicted in red. Orange or red boxes represent the localization of the modifications introduced by the CrtYB or CrtI enzymes respectively.

Several studies have demonstrated that the efficiency of colocalisation of metabolic reactions with membrane bound enzymes or transporters associated to cytosolic enzymes can be improved using synthetic protein scaffolds to recruit the various enzymes of the biosynthetic pathway (Lin et al., 2017; Somasundaram et al., 2017; Thomik et al., 2017). To test how spatial reorganization of the carotenoid enzymes could reduce the accumulation of intermediates and reorient the metabolic fluxes toward β-carotene production, we engineered a series of bifusion enzymes using the CrtB domain or the full CrtYB from *X. dendrorhous* connected to CrtI by means of various peptidic linkers providing variable spatial proximity and flexibility between domains. We then investigated how the relative orientation of interconnected domains and the properties of the linker (size, secondary structure, rigidity) influenced both the yield and pattern of product formation and, based on these results, we engineered a tridomain fusion that improved carotenoid production efficiency.

3 Materials and Methods

3.1 Plasmid construction

All the plasmids used in this study are presented in Supplementary Table S1. DNA Sequences for *CRTYB* and *CRTI* encoding genes were obtained by amplification of the open reading frames from the plasmid pYEplac195 *CRTYB/CRTI/CRTE* (Verwaal et al., 2007). For the construction of the different fusion proteins, the different linker sequences were obtained via commercial gene synthesis (GeneCust, Luxemburg). *CRTB* encoding gene was obtained via gene synthesis using codon optimization except for constructs using *CRTYB* where the natural sequence was used. All natural or recoded sequences used are presented in Supplementary Data S1, the linker sequences in Supplementary Data S2. The various reconstructed enzymes were expressed from the dual *GAL1-GAL10* promoters present in the pMRI31 and pMRI32 vectors (Xie et al., 2014), kindly provided by Prof. Hongwei Yu. Open reading frames encoding the various domains were cloned either at the N- or the C-terminus of one linker by classical restriction-ligation or isothermal assembly. All plasmid constructions and modified sequences were verified by sequencing. *E. coli* DH5α strain was used for all molecular biology protocols except for In-Fusion (Takara Bio USA, Mountain View, CA, USA) experiments for which Stellar competent cells were used according to the fabricant recommendations.

3.2 Strain construction

The strain BY4741 was used to construct the various strains bearing the fusion proteins. The first step was the disruption of the *S. cerevisiae GAL80* gene to alleviate the repression of glucose on the *GAL1-10* promoter obtaining strain yHR22. As carotenoid synthesis in *S. cerevisiae* is highly dependent on the level of GGPP (Misawa, 2011), a supplementary of the more efficient GGPP synthase (*CRTE*) of *X. dendrorhous* and the truncated form of *HMG1* were both expressed under the control of the p*GAL1-10*, obtaining the strain yHR21. All subsequent strains are derivatives of yHR21 and are presented in the Supplementary Table S2. The pMRI integrative plasmids, bearing the same dual promoter (*GAL1-GAL10*), were used for integration (Xie et al., 2014), avoiding as much as

possible transcriptional effects due to plasmid copy number variation or transcription interference (Carquet et al., 2015). The following strain naming was also adopted for clarity: each catalytic domain is represented by a capital letter (Y, B and I), each linker by a parenthesis defining the linker type (yb, cpr, ek and gs), the absence of linker by the absence of parenthesis and a / sign when the catalytic domain is expressed separately. As an example, the *S. cerevisiae* strain bearing the natural carotenoid system from *X. dendrorhous* is notated Y(yb)B/I.

The integration of the different constructions was performed using the high efficiency Gietz transformation method (Schiestl and Gietz, 1989) with some minor modifications: 1 µg of linearized DNA fragment digested by *Sf*iI (New England BioLabs, Ipswich, MA, USA) was added in a maximum volume of 10 µl. The mixture was incubated 30 min at 30°C after which cells were heat shocked 20 min at 42°C in a water bath. The cells were recovered during two hours in 1 mL of YPD at 30ºC and plated in YPD containing G418 at a concentration of 200 µg/ml.

Integrations were verified via colony PCR using two specific pairs of oligonucleotides. Each pair is composed of one oligonucleotide directed against the introduced gene or the Kan^R marker and the other one directed against the chromosomal region of insertion. In the case when the Kan^R resistance marker needed to be removed, correct colonies were then further transformed with the pSH47 vector to eliminate the Kan^R gene by the Cre-Lox P recombination system (Güldener et al., 1996). Transformed clones were grown overnight on 5 ml SD-URA Gal medium to maintain and express the Cre enzyme and plated on SG-URA plates (approximately 100 cells per plate) to allow for the loss of the *KanMX* gene. After three days growth, clones were individually tested for their Kan^R phenotype using both YPD and YPD containing 200 µg/ml of Geneticin (G418). Clones which are not growing on YPD G418 were grown overnight on YPD medium to allow for the loss of the pSH47 plasmid. Cells from the culture were plated onto YPD medium (approximately 100 cells per plate). After three days growth, clones were individually tested for their URA+ phenotype using both YPD and SG-URA medium.

3.3 Media composition

YPD or YPG media (20 g/L peptone, 10 g/L yeast extract, 20 g/L glucose or galactose respectively), YPD or YPG agar plates (YPD or YPG media + 20 g/L agar) were used to grow yeast. SD-URA or SG-URA media (1.7 g/L Yeast Nitrogen Base w/o amino acid, 5 g/L ammonium sulfate, 20 g/L glucose (SD) or galactose (SG), 0.77 g/L dropout mix containing all amino acids and nucleic acids without uracil (MP Biomedicals, Santa Ana, California, USA) and SD or SG -URA plates (SD or SG -URA medium with 17 g/L agar) were used to grow or select for URA^+ colonies. Geneticin (G418) was incorporated at 200 µg/ml in YPD media for the selection of Geneticin resistant colonies.

3.4 Carotenoid production in shake flasks

Cryoconserved strains were regrown on YPG plates. Typically, one isolated clone was used to inoculate a 5 ml preculture in YPG. After one night growth, the preculture was used to inoculate several 50 ml YPG flasks at 0.05 OD_{600nm} and further cultivated at 28°C, 140 rpm during 72 hours in an incubator INFORS HT with 50 mm shaking throw. Cells were then pelleted for 3 min at 4°C at 6000 g, washed in 2 ml of distilled water, and then transferred into a pre-weighted screw capped tube, and centrifuged again at 9300 g at 4°C for 3 min. Pelleted cells were stored at -80°C until further use.

3.5 Carotenoid extraction and analysis

The frozen cell pellets were lyophilized and resuspended in 500 µL of acetone. Apocarotenal (24 mg/L in hexane) was used as internal standard (100 µL per 0.025 g of dry cells). Extraction was performed as follows: 3 cycles of 20 s at 0.05 m/s with one minute of cooling on ice using a Fast Prep FP120 (Thermofisher, Waltham, MA, USA). The acetone phase was transferred to a new tube and the process repeated 3 times. Acetone extracts were pooled and dried under nitrogen flux and further suspended in 500 µL of hexane for HPLC analysis.

30 µl of carotenoids extract were injected on a Waters alliance HPLC system. Separation was done as follows: an isocratic gradient composed of isopropanol and acetonitrile/water (90:10, v:v) was applied for 49 min at a flow-rate of 0.3 ml/min into a 100 x 2.1 mm Brownlee RP18 column at 35°C. Absorbance spectra were recorded between 200 nm and 600 nm using a 996 Waters PDA detector. Peaks were characterized by their elution time and their spectra (Takaichi, 2000). Main

carotenoids peaks were manually integrated: at 282 nm (phytoene), at 475 nm (lycopene), at 454 nm (β-carotene), at 442 nm (neurosporene), 399 nm (ζ-carotene), 349 nm (phytofluene), 430 nm (dihydro-β-carotene), and 487nm (torulene). Peaks areas were used to calculate quantity based on a calibration curve performed with phytoene, lycopene and β-carotene standards solutions (Sigma-Aldrich, St. Louis, MO, USA) and neurosporene, zeta-carotene, phytofluene, β-zeacarotene and torulene (Carotenature, Münsingen, Switzerland) and the correction for the extraction yields were performed with the quantification of apocarotenal (Sigma-Aldrich, St. Louis, MO, USA). Attribution of dihydro-β-carotene peaks was done in reference to the elution order described in (Mizoguchi et al., 2015; Takaichi, 2000). As there is no available standard for dihydro-β-carotene, the quantification was performed with the β-zeacarotene standard curve as both products contain the same amount of conjugated double bonds. All other carotenoids were not quantified due to absence of commercial standard. Each extracted sample was injected and quantified twice (technical duplicate). A second set of batch cultures was also prepared and analyzed with the same protocol for duplication (biological duplicate). Error bars represent the standard deviation obtained from these different values.

3.6 Western blot analysis

Protein extracts were prepared following the protocol described by Zhang et al (Zhang et al., 2011). Briefly, pelleted cells from one 50 ml flask were resuspended in 1 ml PBS (Invitrogen, Carlsbad, CA, USA) by agitation. To normalize the quantity of cell before extraction, cell density was verified by OD_{600} measurements. 8 µL of the cell suspension were gently mixed on ice with 200 µl of a 2 M lithium acetate cold solution, and left to stand for 5 min, followed by 5 min centrifugation at 5000 g, 4°C. The supernatant was discarded and 200 µL of a 0.4 M solution of NaOH added. After gentle resuspension, and 5 min standing on ice, samples were centrifuged 5 min at 4°C. After discarding supernatants, pellets were vigorously vortexed with 100 µL of bromophenol Blue loading dye solution supplemented with 5% β-mercaptoethanol. After denaturation for 5 min at 90°C, 15 µL of each sample was deposited on a 10% SDS page gel. Semi-dry transfer were performed on PVDF membrane (Merck Millipore, Darmstadt, Germany) using a Trans-Blot® SD Cell BioRad apparatus (18V at 600 mA during 25 min). 5% bovine milk in TBS was used as blocking agent. Incubations with primary anti-Flag

(Thermofischer, Waltham, MA, USA), mouse anti-c-Myc (Invitrogen, Carlsbad, CA, USA), and secondary anti-mouse IgG coupled with alkaline phosphatase (Thermoscientific, Rockford, IL, USA), diluted following instructions from the provider, lasted one hour. Proteins were detected by the incubation of BCIP/NBT AP substrate buffer (Sigma-Aldrich, St. Louis, MO, USA).

3.7 Protein analysis by GeLC-MS

Crude cell lysates were diluted with reducing Laemmli buffer (Sambrook and Russell, 2006) to a concentration of 3.3 µg/µL and SDS-PAGE was performed on a 4-15 % precast gradient gel (Mini-PROTEAN TGX gel; Bio-rad, Munich, Germany). Gels were stained by Coomassie and protein bands were excised for subsequent in-gel digestion as indicated on the gel-scan (Supplementary Figure S1).

After reduction (20 mM dithiothreitol) and alkylation (55 mM chloroacetamide), proteins were digested with 0.1 µg of sequencing grade modified Trypsin (Promega, Mannheim, Germany) in 0.1 M triethylammonium bicarbonate buffer (TEAB) for 16h at 37 C. Peptides were extracted, the solvent was lyophilized to dryness and reconstituted in 20 µL of HPLC loading buffer (3% acetonitrile, 0.05% trifluoroacetic acid in water). Peptide extracts were analyzed by LC-MS on a Dionex U3000 uHPLC system coupled to a Q Exactive Plus mass spectrometer (Thermo Scientific, Dreieich, Germany). A volume of 4 µl of the samples were injected on a C18 PepMap 100 µ-precolumn (column dimensions: 300 µm i.d. x 5 mm) before separated on an Acclaim PepMap RSLC column (column dimension: 75 µm i.d. x 50 cm; Thermo Scientific). A binary gradient of eluent A (0.05% aqueous formic acid) and eluent B (80% acetonitrile, 0.04% formic acid) with a flow rate of 0.3 µL/min was used. Peptides were eluted over a gradient from 5 % eluent B to 50 % eluent B in 70 min, followed by an increased to 95 % eluent B in 6 min. After isocratic elution at 90 % eluent B for 10 min, the column was equilibrated for 10 min with 5 % eluent B. The LC-system was directly coupled to the Q Exactive Plus mass spectrometer. Full MS scans were acquired in positive ion mode with a resolution of 35,000, with a scan range from 240 m/z to 1800 m/z. Data dependent MS/MS spectra of the ten most intense precursor ions were acquired with a resolution of 17,500; Scan parameters were set to an isolation window of 1.6 m/z, a normalized collision energy of 27. Precursors with a charge states < 2 and > 6 were excluded and precursors were excluded form subsequent isolation for 10 s. MS raw files were searched against a

database including *Saccharomyces cerevisiae* (strain CEN.PK113-7D; 5,439 UniProt entries, 10.2018) in addition to the target proteins and common contaminants using the Sequest search algorithm and the Proteome discoverer software (Thermo Scientific). Searches with a full-tryptic and semi-tryptic protease specificity, allowing for variable oxidation of methionine and N-terminal acetylation, were combined. Peptide spectrum matches and protein identifications were performed with a false discovery rate below 1%. The complete list of identified proteins is provided as a Supplementary Excel file S1.

3.8 Sequence alignments

Alignments were performed with the Muscle software and with the sequence identification numbers boxed in parentheses from the Uniprot database. The names and Uniprot codes for the different aligned proteins are the following: C.t., *Chloracidobacterium thermophilum* (G2LGW3); M.sp, *Mycobacterium sp* (A0A1A0Q6A2); N.sp., *Nanosalina sp.* (G0QG73); C.f., *Cystobacter fuscus* (S9QG19); P.a., *Pantoea ananas* (P21687); A.n., *Aspergillus niger* (A2QM49); A.o., *Aspergillus oryzae* (Q2U4X9); A.g., *Arthroderma gypseum* (E4UPP6); *Neurospora crassa* (P37295); *Phaffia rhodozyma* (Q7Z859); *Podospora anserina* (B2ATB0); *Pyrenophora tritici-repentis* (B2WAQ3); *Phaeosphaeria nodorum* (Q0V6M5); *Pantoea ananas* (P21683); *Franconibacter pulveris* (C7C5F2); *Erwinia gerundensis* (A0A0U5L0T8); *Cronobacter turicensis* (C9Y0F0); *Enterobacter agglomerans* (Q47846).

4 Results and discussion

4.1 Design and construction of the various metabolic systems

4.1.1 Defining CrtB and CrtY domains

To define the limits of the CrtB and CrtY domains within CrtYB, we compared 57 bifunctional lycopene cyclase/phytoene synthase enzymes with 35 sequences of phytoene synthases and 55 sequences of single domain lycopene cyclases (Supplementary Figure S2). This alignment evidences a large gap between the end of the various single domains CrtY enzymes and the beginning of the various CrtB enzymes and is filled, in all bifunctional enzymes, with different linker sequences (Figure

2). As seen in Figure 2 (lower panel), the C-terminal segment of the CrtY domain of monofunctional lycopene cyclases and bifunctional CrtYB enzymes seems well conserved up to residue 254 (numbering refers to CrtYB from *X. dendrorhous*) where the homology starts decreasing between homologous enzymes. Furthermore, the alignment of CrtYB with monofunctional phytoene synthases confirmed that the residues of the linker have no equivalent in single domain CrtB or CrtY enzymes. The linker region seems to possess between 20 and 25 residues depending on the species considered, leading to a start of the CrtB domain around position 290. The end of the CrtY and the beginning of the CrtB domains were chosen with extra residues on both sides to avoid possible incorrect delimitation and were thus defined at the residue G260 (end of CrtY) and at the residue S286 (beginning of CrtB). The CrtY domain, when expressed as a single domain, contained the full linker sequence to avoid, as much as possible, any impact on activity.

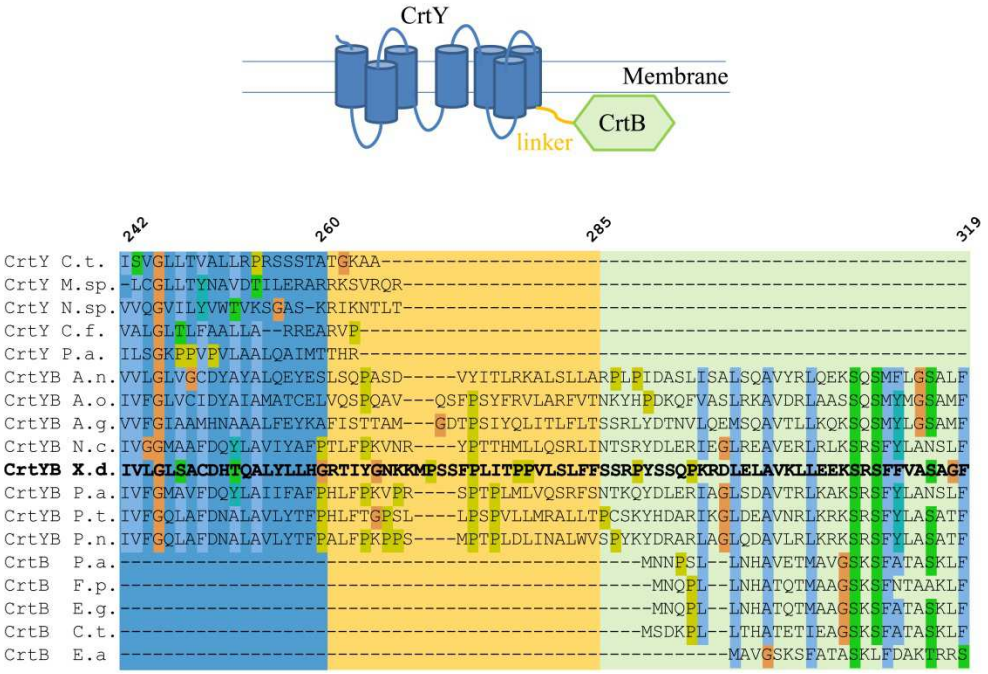

Figure 2. Localization and sequence of the linker of CrtYB from *X. dendrorhous*. Upper panel: graphical sketch of the putative structural organization of CrtYB in the membrane. The lycopene cyclase, phytoene synthase and linker domains are represented in blue, green and yellow, respectively. Eukaryotic lycopene cyclases enzymes are predicted to have seven transmembrane helices (http://www.uniprot.org/uniprot/Q7Z859) while phytoene synthases do not have any transmembrane segment (Schaub et al., 2012). Lower panel: partial alignment

between the CrtYB enzyme from *Xanthophyllomyces dendrorhous* (X.d.) (in bold) with other bifunctional phytoene synthases/lycopene synthases and monofunctional phytoene synthases or lycopene cyclases (See Materials and Methods).

4.1.2 Linkers properties and strains constructions

The combinatorial assembly of the domains of the 3 enzymes involved in the carotene synthesis was performed using either no linker or variable linkers. The choice of the linkers was dictated mostly to cover various physicochemical properties: size and dynamics. The longest linker used is the original linker from CrtYB, as defined previously (27 residues). Although there is no information on its dynamical properties, we expected this linker to be well tolerated when fused to the CrtB domain. We then selected a series of shorter linkers for which dynamical properties are well known. The second linker corresponds to the short segment (18 residues) connecting the FMN and FAD/NADPH domains of the human cytochrome P450 reductase enzyme (CPR) and for which high flexibility has been described by SAXS and NMR (Frances et al., 2015). The third and fourth linkers are short (20-residues long segments) and are completely synthetic sequences (Arai et al., 2004) that behave as flexible (four repetition of the GGGGS motif) or rigid α-helical connectors (four repetitions of the EAAAK motif). In the rest of the article, the linker in each construction can be inferred from the lower-case letters between the parenthesis: yb for the CrtYB linker, cpr for the connecting domain of CPR, ek and gs for the rigid or flexible small synthetic linkers corresponding to (EAAAK)4 or (GGGGS)4 respectively, the absence of linker by the absence of parenthesis. Furthermore, the name of the enzymatic domains was represented by a capital letter (Y for CrtY for example). A slash sign between catalytic domains indicated that domains on both sides of the sign were expressed as single domains. As an example, the *S. cerevisiae* strain bearing the natural carotenoid system from *X. dendrorhous* is notated Y(yb)B/I, corresponding to the production of CrtYB and CrtI.

	Strain name	Domain 1	Linker	Domain 2	Linker	Domain 3
Single domain	B	CrtB				
	B/I	CrtB		CrtI		

	B/I/Y	CrtB		CrtI		CrtY
Bidomain	BI	CrtB	-	CrtI		
	B(yb)I	CrtB	(yb)	CrtI		
	B(cpr)I	CrtB	(cpr)	CrtI		
	B(gs)I	CrtB	(gs)	CrtI		
	B(ek)I	CrtB	(ek)	CrtI		
	IB	CrtI	-	CrtB		
	I(yb)B	CrtI	(yb)	CrtB		
	I(cpr)B	CrtI	(cpr)	CrtB		
	I(gs)B	CrtI	(gs)	CrtB		
	I(ek)B	CrtI	(ek)	CrtB		
	Y(yb)B	CrtY	(yb)	CrtB		
	Y(yb)B/I	CrtY	(yb)	CrtB		CrtI
Tridomain	Y(yb)B(ek)I	CrtY	(yb)	CrtB	(ek)	CrtI

Table 1. Enzymatic systems used in this study. Columns after the strain names describe which domains (grey cells) and linkers (dark cells) are present in the strains. CrtY: lycopene cyclase; CrtB: phytoene synthase; CrtI: phytoene desaturase. No sign in the linker columns depicts expression of individual domains (white cell). The – sign represents a fusion without any linker present; (yb), (cpr), (gs) and (ek) represent the different linkers.

Sixteen different yeast strains containing combinations of the individual domains CrtB, CrtI and CrtY, bidomain or tridomain enzymes were constructed (Table 1). In the tridomain fusion, only the (ek) linker was tested and this choice was partly dictated by preliminary results obtained with the various bidomain enzymes (vide infra).

4.2 Enzyme detection by western blots in the various strains

Before analyzing the carotenoid content of the various strains, we first assessed the level of protein production of the different fusions and/or single domains. Figure 3 presents the western blots obtained for the full protein extracts of the strains in the same culture conditions used for carotenoids quantification, i.e. at 72 hours of culture. All single, double or triple domain constructs were tagged with the c-Myc peptide at the C-terminus, with the exception of the CrtY protein (when expressed alone) that carried a FLAG tag instead.

We first analyzed the production of single domain enzymes (Figure 3A). CrtB was invariably detected as intense bands, indicating a good expression in the strains B, B/I or Y/B/I. On the contrary, the CrtI protein was found at the expected size (67 kDa) only when coexpressed together with CrtY domain (strain B/I/Y). The presence of a faint band, with 50 kDa apparent molecular weight, in the B/I strain might indicate the presence of a truncated form of CrtI, reinforcing the idea of low production or stability of CrtI in the absence of CrtY. With regard to the CrtY single domain (33 kDa), its presence could not be revealed using our extractions conditions, suggesting an inadequate transfer of this integral membrane protein during blotting. Stability defects caused by the absence of the CrtB domain on the transmembrane portion are less likely, as suggested by the production of β-carotene by this construct (see below). Other constructs carrying the CrtY domain tagged with c-Myc peptide, and connected to a soluble domain (B), could be successfully detected (see below). Hence, independently of the precise reason why CrtY alone is not detected in our western blots, its expression can be inferred from the detection of the lycopene cyclase activity (vide infra) and also indirectly by to the abovementioned effect on CrtI expression.

Expression of the natural CrtYB was well detected (Figure 3B), although with a relatively lower size (65 kDa, as compared to 75 kDa theoretical value). This sort of discrepancies are commonly observed when considering the SDS-PAGE migration pattern of integral membrane proteins (Rath et al., 2009). This also led to the non-resolution of the two individual bands when CrtI (67 kDa) was coexpressed with CrtYB. Since CrtI was never well detected, the band at 65 kDa was cautiously assigned to CrtYB. Interestingly, the intensity of the 65 kDa band decreased when CrtI and CrtYB were coexpressed together, suggesting that CrtYB level is also reduced in Y(yb)B/I as compared to Y(yb)B. Again, a possible truncation product of CrtI was faintly detected in the Y(yb)B/I strain. As will be seen further, the expression of functional CrtI in this strain will be supported by its activity.

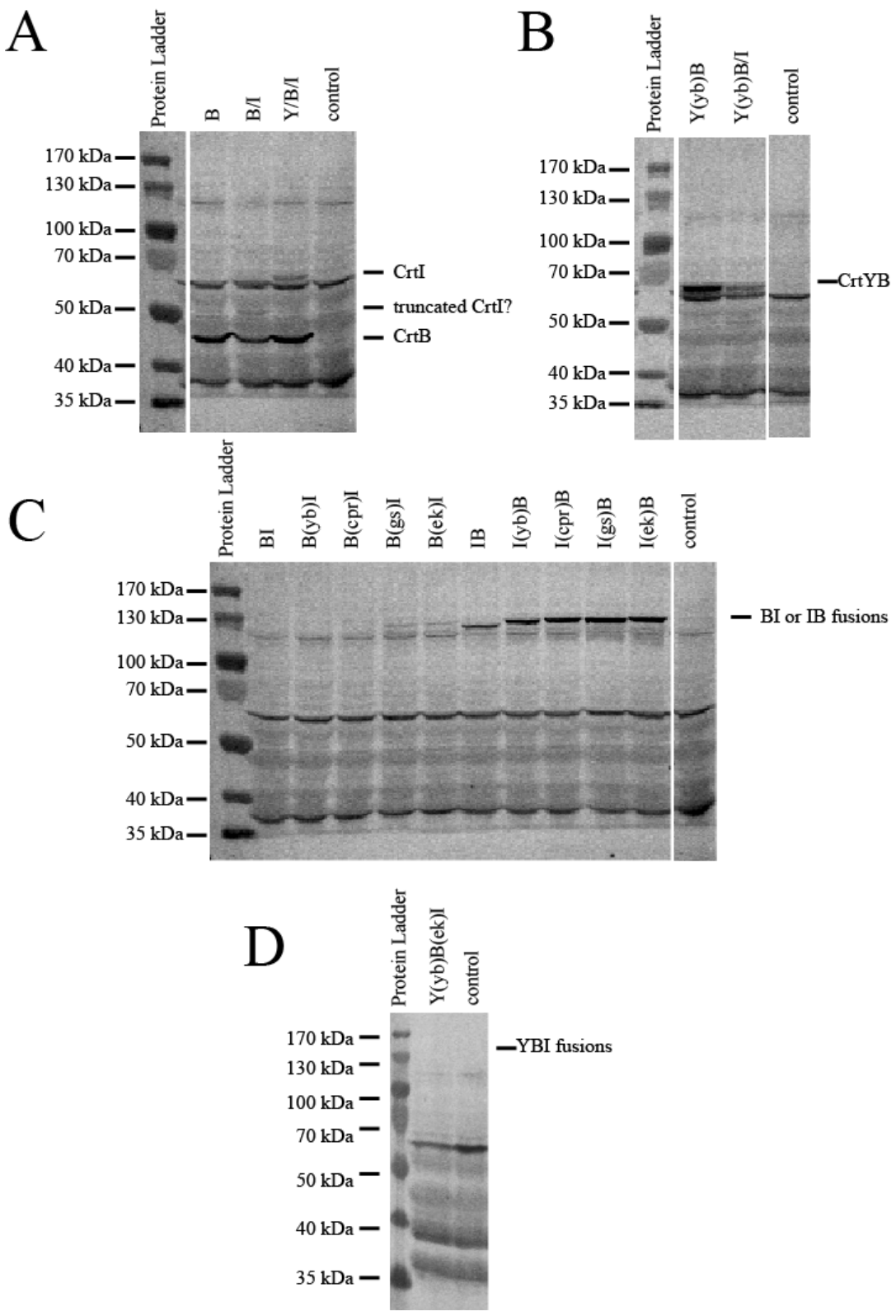

Figure 3. Expression of the different constructs in crude extracts detected in western blots. Loaded samples correspond to material withdrawn after 72 h of culture. The void space between lanes represents spatial arrangements of different lanes from the same blot. Control lanes are from extracts from the strain engineered

for GGPP overproduction but devoid of any carotenoid enzyme (yHR21). All components were blotted with antibody that detect specifically C-ter c-myc peptides, exception made of the single Y domain (third lane of panel A), which was detected by an antiFLAG antibody recognizing its C-ter FLAG peptide. **A.** Strains expressing monomeric enzymes. **B.** Strains expressing Y(yb)B based constructions. **C.** Strains with bifunctional synthetic enzymes. **D.** Strain expressing the trifunctional enzyme.

Bands for bidomain constructs with the CrtB domain connected at the N-ter of the CrtI domain (B to I orientation) were undistinguishable from those present in the control lane (compare lanes 2-6 with last one in Figure 3C). Specific protein bands only became visible for the fusions bearing the linkers (gs) and (ek) but still remained faint (lanes 5,6), suggesting that the B to I orientation is likely deleterious for expression. On the contrary, bidomain enzymes with the I to B orientation were perfectly detectable (lanes 7-11). As with the B to I constructs, the (gs) and (ek) linkers appeared most benefic for promoting higher amounts of fusions proteins (I(ek)B and I(gs)B. Remarkably, the different linkers induced apparent size shifts, in spite of their relatively small size (27 residues for the longest one). Overall, the results of the bidomain fusions expression indicate that the N- or C-terminal positioning of CrtI or CrtB catalytic domains strongly affects the protein production level. Expression of the tridomain fusion construct (strain Y(yb)B(ek)I) was also investigated without success. Different lytic and solubilization protocols were applied, but all efforts to establish its presence by western blot failed (Figure 3D).

The strong influence of domain orientation and linker composition on the protein expression level or the translocation through the periplasmic membrane of *E. coli* has already been described in the case of single chain antibodies composed of the fusion of light and heavy chains (Anand et al., 1991; Tsumoto et al., 1994). The oligomeric assembly and membrane association of the CrtI enzyme from *Pantoea ananatis* and *Oryza sativa* has been well studied, providing evidences that i), CrtI is associated to the membrane via peptidic regions that are not membrane spanning segments and, ii), that the N-terminus of CrtI likely plays a major role in directing the correct folding of the enzyme as well as its oligomerization (Gemmecker et al., 2015; Schaub et al., 2012). The fact that the I to B orientation is preferred may indicate that membrane interaction or oligomerization is indeed located

in the N-terminus of the CrtI domain and this N-terminal part, when placed between the CrtB and CrtI enzymes, might not properly promote its role in membrane interaction or oligomerization, reducing the amount of correctly folded, membrane localized CrtI and thus leading to its degradation within cells.

Overall, western blot analysis allowed us to define relative expression efficiencies for the various expression systems. Although semi-quantitative, such information is important to better understand some of the patterns of carotenoid molecules production.

4.3 Detection of the natural Y(yb)B/I and the Y(yb)B(ek)I trifusion by GeLC-MS

The unsuccessful detection of trifusion protein by western blot prompted us to perform shotgun proteomic analyses using gel electrophoresis followed by in-gel digestion and liquid chromatography coupled to mass spectrometry of the extracted peptides (GeLC-MS). After separating by SDS-PAGE the total protein extracts of yHR21, Y(yb)B/I and Y(yb)B(ek)I strains, gel regions around the MW of CrtYB(ek)I (145 kDa), CrtYB (76 kDa, but migrating as a 65 kDa protein (Figure 3B)) and CrtI (67 kDa) were treated in search for specific peptides of the Y, (yb), B, (ek) or I domains (Figure 4 and Supplementary Figure S1). A gel region around 40 kDa MW was also analyzed in both yHR21 and Y(yb)B(ek)I strains and peptides specific for CrtE were detected with high sequence coverage. A summary of all identified proteins including details for peptide evidence for the target proteins is presented in Supplementary Excel File S1.

A

CrtY(yb)B(ek)I

MTALAYYQIHLIYTLPILGLLGLLTSPILTKFDIYKISILVFIAFSATTPWDSWIIRNGAWTYP
SAESGQGVFGTFLDVPYEEYAFFVIQTVITGLVYVLATRHLLPSLALPKTR**SSALSLALK**ALIP
LPIIYLFTAHPSPSPDPLVTDHYFYMRALSLLITPPTMLLAALSGEYAFDWKSGRAKSTIAAIM
IPTVYLIWVDYVAVGQDSWSINDEKIVGWRLGGVLPIEEAMFFLLTNLMIVLGLSACDHTQALY
LLHGRTIYGNKKMPSSFPLITPPVLSLFFSSRPYSSQPK**RDLELAVK**LLEKKSR**SFFVASAGFP**
SEVRERLVGLYAFCRVTDDLIDSPEVSSNPHATIDMVSDFLTLLFGPPLHPSQPDK**ILSSPLLP**
PSHPSRPTGMYPLPPPPSLSPAELVQFLTER**VPVQYHFAFR**LLAK**LQGLIPRYPLDELLRGYTT**
DLIFPLSTEAVQARKTPIETTADLLDYGLCVAGSVAELLVYVSWASAPSQVPATIEER**EAVLVA**
SREMGTALQLVNIARDIKGDATEGR**FYLPLSFFGLR**DESK**LAIPTDWTEPRPQDFDKLLSLSPS**
STLPSSNASESFRFEWK**TYSLPLVAYAEDLAK**HSYK**GIDRLPTEVQAGMRAACASYLLIGR**EIK
VVWKGDVGERR**TVAGWR**RVRK**VLSVVMSGWEGQ**HMEAAAKEAAAKEAAAKEAAAKEAAAK**EDP**GKEQDQ
DKPTAIIVGCGIGGIATAARLAKEGFQVTVFEKNDYSGGRCSLIERDGYRFDQGPSLLLLPDLF
KQTFEDLGEKMEDWVDLIKCEPNYVCHFHDEETFTFSTDMALLKREVERFEGKDGFDRFLSFIQ
EAHRHYELAVVHVLQKNFPGFAAFLRLQFIGQILALHPFESIWTRVCRYFKTDRLRRVFSFAVM
YMGQSPYSAPGTYSLLQYTELTEGIWYPRGGFWQVPNTLLQIVKRNNPSAKFNFNAPVSQVLLS
PAKDRATGVRLESGEEHHADVVIVNADLVYASEHLIPDDARNKIGQLGEVKRSWWADLVGGKKL
KGSCSSLSFYWSMDRIVDGLGGHNIFLAEDFKGSFDTIFEELGLPADPSFYVNVPSRIDPSAAP
EGKDAIVILVPCGHIDASNPQDYNKLVARARKFVIQTLSAKLGLPDFEKMIVAEKVHDAPSWEK
EFNLKDGSILGLAHNFMQVLGFRPSTRHPKYDKLFFVGASTHPGTGVPIVLAGAKLTANQVLES
FDRSPAPDPNMSLSVPYGKPLKSNGTGIDSQVQLKFMDLERWVYLLVLLIGAVIARSVGVLAFV
DMEQKLISEEDLE

B

		yHR21			Y(yb)B/I			Y(yb)B(ek)I		
MW region (kDa)		120 to 150	70 to 80	40 to 50	120 to 150	70 to 80	60 to 70	120 to 150	70 to 80	40 to 50
# Peptides specific of	Y	0	0	0	0	0	0	1	0	0
	B	0	0	0	0	0	6	4	16	16
	(ek)	0	0	0	0	0	0	0	1	2
	I	0	0	0	0	0	14	2	11	17
	E	0	0	20	0	0	0	0	0	20

Figure 4. Detection of CrtY(yb)B(ek)I peptides by GeLC-MS analysis. **A.** Amino acid sequence of the CrtY(yb)B(ek)I construct. The different domains of the fusion protein are color coded: Y (red), (yb) (blue), B (green), (ek) (black), I (orange). Peptides that could be unequivocally identified are highlighted. Total sequence coverage by identified peptides for the fusion protein CrtY(yb)B(ek)I was 43%. **B.** Summary of GeLC-MS data. Numbers correspond to peptide hits identified for a given domain (horizontal lanes) within a given gel portion (MW region) which was selected as corresponding to the expected electrophoretic mobility for a given construct. The intensity of green coloring is related to the number of detected peptides.

Peptides identified for the trifusion protein covered about 43% of its sequence (Figure 4A), being highest for the stretch covering the B(ek)I portion. Thus, peptides from both the B and I domains could be detected in the *ca.* 145 kDa MW region. This, together with the successful identification of a peptide deriving from the Y domain (Figure 4A), strongly supported the occurrence

of the Y(yb)B(ek)I tridomain fusion. A higher coverage of the CrtYB(ek)I sequence was attained in measurements of MW gel regions from 40 to 80 kDa (Figure 4B). While most of the detected peptides correspond to either the B or I domains, three of them covered the (ek) linker, extending either towards the B or I domains (Figure 4A).

That a single peptide was detected for Y is likely explained by two parameters: i) the low ionizability of peptides generated from this hydrophobic domain, which consists of 7 membrane-spanning helices; ii) the relatively low abundance of trypsin-cleavage sites (13 potential sites over the 260 residues of the Y domain, as compared to 45 sites for the remaining 400 residues). Such difficulties were confirmed in measurements on the natural YB/I case, which failed to reveal any peptide from the Y domain in any gel region. Similarly, the analysis of the Y(yb)B/I strain shows that peptides covering both the B domain and CrtI were found in the region corresponding to the MW of CrtI but not in the region corresponding to the MW of CrtY(yb)B. This result is in agreement with the faster than expected electrophoretic migration of CrtY(yb)B observed in western blots (see above). Unfortunately, MS data did not permit to clarify whether this phenomena was due to proteolytic degradation of the natural fusion. However, semi-tryptic peptides (which would support the occurrence of non-tryptic proteolytic cleavage) were never be detected, thus indicating low occurrence of proteolysis of the Y(yb)B(ek)I protein. Overall, GeLC-MS investigations provided very good evidences of the presence of the full length Y(yb)B(ek)I protein.

4.4 Catalytic efficiency of the different natural or synthetic enzymatic systems

Quantification of all carotenoid molecules was performed using standards curves and authentic standards except for dihydro-β-carotene that does not exist as a standard but its relative elution order and spectrum are known (Mizoguchi et al., 2015; Takaichi, 2000) (Supplementary Figure S3). In the literature, there is no report on the absorption coefficient of dihydro-β-carotene. However the spectrum of dihydro-β-carotene and β-zeacarotene are relatively similar which signs relative homology between the double bonds resonances of the compounds. We then hypothesized that the two molar absorption coefficients were approximately equal and thus we could use β-zeacarotene calibration curves to quantify dihydro-β-carotene. The full quantification of carotenoids from all

strains was performed using three independent cultures and is presented in the Supplementary Table S3.

4.4.1 Functionality of single domain enzymes

The impact of splitting the CrtB domain from CrtYB on phytoene synthase activity was first evaluated (Figure 5). Phytoene production was severely decreased (40 times) in the B strain compared to the Y(yb)B strain, confirming the importance of the natural fusion between CrtY and CrtB for the phytoene synthase activity (Niklitschek et al., 2008; Verdoes et al., 1999; Xie et al., 2015a). A first hypothesis for the low activity of phytoene synthase when expressed alone could be associated with a faulty targeting of CrtB to the correct cellular environment when disconnected from CrtY, which is normally directed to the endoplasmic reticulum due to its transmembrane segments. As the low phytoene production with our CrtB construct is probably not due to defects in expression (Figure 3A), a second hypothesis is that splitting the two domains decreases the residence time of the phytoene synthase in the vicinity of the membrane, which may also hamper the attainment of the appropriate orientation required for efficient product release. Bearing in mind the difference in water solubility between the substrate (GGPP) and the product (phytoene) of CrtB, the overall effect would be a dramatic slowdown of the enzymatic process by product inhibition due to difficult water environment release.

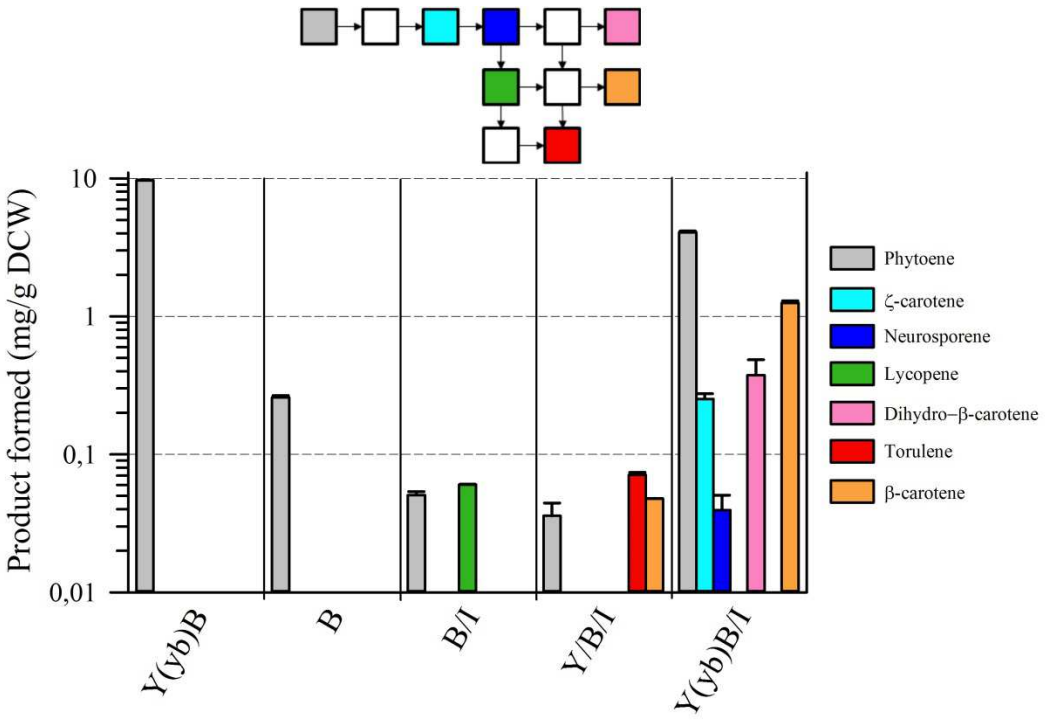

Figure 5. Quantification of carotenoids from the strains bearing single domains enzymes. Major carotenoids were quantified from strains expressing single domain enzymes (B, B/I, B/I/Y) and are compared to data obtained for strain Y(yb)B (left side) or Y(yb)B/I (right side). Metabolites are colored according to the legend in the right. The metabolic pathway from phytoene to carotene is depicted at the top with the same color code as previously, white squares correspond to intermediate metabolites that were not detected. Values represent the average ± S.D. of the quantification of each metabolite with three independent cultures.

The strain B/I produces a total amount of carotenoids around 100 µg/g DCW, roughly comparable to the one measured in the B strain. Therefore, the coexpression of CrtI with CrtB has almost no impact on the total flux of the carotenoid pathway which again confirms that the low activity of CrtB might be attributed to the slow release of phytoene.

We then reconstituted the full β-carotene producing pathway using the three separated domains (strain Y/B/I). The natural system (strain Y(yb)B/I) produces, in parallel to two major carotenoids (phytoene and β-carotene) several intermediates in detectable quantities such as ζ-carotene, neurosporene, dihydro-β-carotene but no lycopene, confirming data already reported by Verwaal and coll. (Verwaal et al., 2007). In the strain Y/B/I, the total production of carotenoids is reduced by a factor of 55 compared to the Y(yb)B/I strain corresponding roughly to the ratio between

phytoene production when comparing YB and B strains. Interestingly, we could evidence a small, yet visible, production of torulene (71 µg/g DCW) which, to our knowledge, has never been detected when expressing CrtYB and CrtI in *S. cerevisiae* although this molecule can be accumulated in *X. dendrorhous* (Verdoes et al., 2003). The relative levels of carotenoid production in the two metabolic systems (Y/B/I compared to Y(yb)B/I) strengthen the hypothesis of the phytoene synthase defect upon separation with the lycopene cyclase domain. As to why the Y/B/I accumulates small amounts of torulene which is absent when the metabolic flux is greater, the simplest explanation is that when low levels of γ-carotene are present, the competition between CrtI and CrtY might be in favor of the desaturation reaction, leading to an equal proportion of torulene and β-carotene. When the overall carotenoid flux is higher (higher amounts of neurosporene and lycopene due to the fusion of CrtY and CrtB), the competition between lycopene cyclase and phytoene desaturase might be in favor of CrtY, leading to the presence of dihydro-β-carotene rather than torulene.

4.4.1.1 Phytoene synthase/desaturase bidomain fusions

The major bottleneck in reconstituted β-carotene pathways is the reaction catalyzed by CrtI, leading to some accumulation of phytoene in all constructed strains. To study how the enzymatic reactions from phytoene to lycopene can be improved we constructed fusions between CrtB and CrtI domains in different orientations and using different linkers (Figure 6). From Figures 6A and 6B, there seems to be a clear distinction between the I to B and the B to I orientations: I to B fusions accumulate small quantities of intermediates up to neurosporene while the other orientation (with the exception of the BI fusion) seem to either accumulate ζ-carotene or the final product lycopene but not neurosporene. We can thus hypothesize that the geometrical constraints imposed by the order of domains as well as the linker properties influence not only the general activity of the bidomain fusions but also specific steps of the phytoene desaturase reactions.

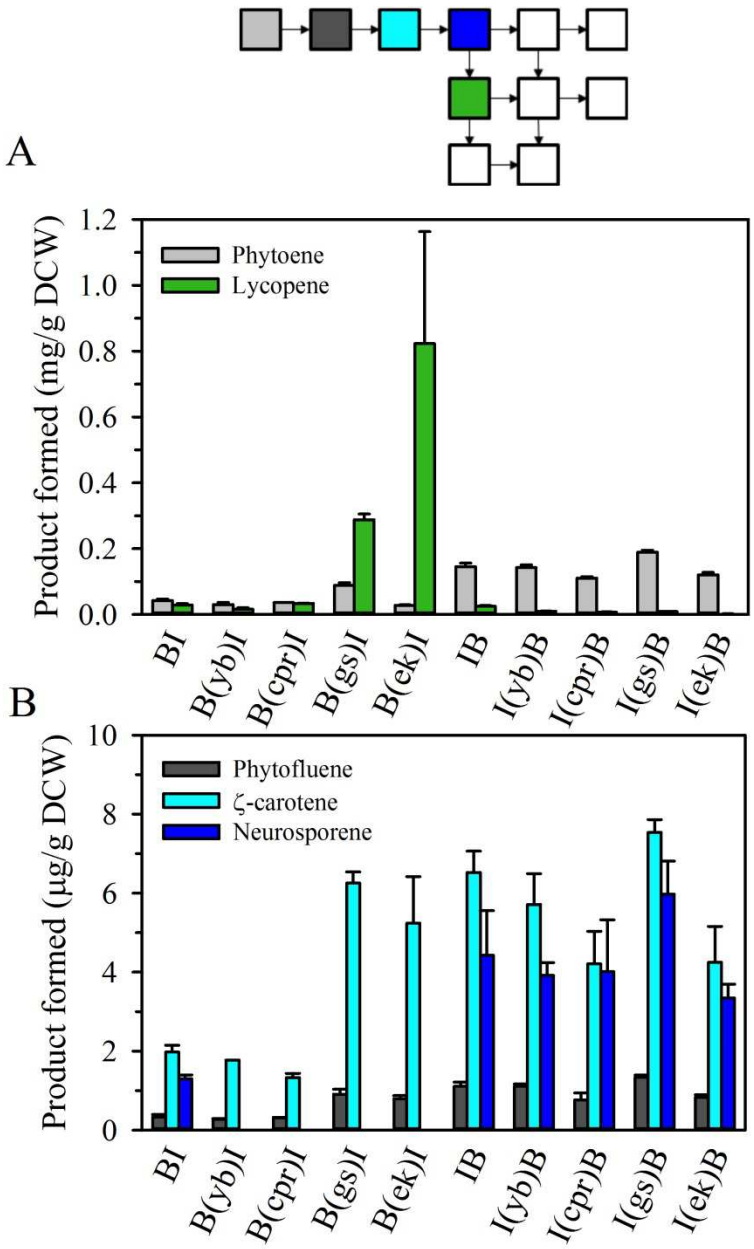

Figure 6. Quantification of carotenoids from the strains bearing bidomain fusions between phytoene synthase and phytoene desaturase enzymes. **A.** Quantification of major carotenes: phytoene (grey bars), lycopene (green bars). **B.** Quantification of minor carotenoid molecules: phytofluene (dark grey bars), ζ-carotene (cyan bars), neurosporene (blue bars). Metabolites are colored according to the legend inside the two panels. The metabolic pathway from phytoene to lycopene is depicted at the top with the same color code as previously, white squares correspond to intermediate metabolites that were not detected. Values represent the average ± S.D. of the quantification of each metabolite with three independent cultures.

Remarkably, the two most proficient bidomain fusions do not correlate with the most expressed enzymes (Figure 3C). Figure 6A evidences a pronounced distinction between B(ek)I and B(gs)I constructs from the rest of the bidomain fusions as both produce much higher amounts of total carotenoids and with a large preference for lycopene over any other intermediate. Furthermore, the stoichiometry between these two molecules is highly dependent on the linker type present between the two domains. Overall, in the B(gs)I and B(ek)I strains, the lycopene production attained 300 and 800 µg/g DCW levels respectively, being significantly higher than phytoene production. This result demonstrates that the phytoene desaturase activity in the B(gs)I and B(ek)I strains is high enough to convert, almost completely, phytoene into lycopene. Furthermore, the total carotenoid production of the B(ek)I strain (Figure 6) is approximately 3-4 times higher than the ones measured in the B or B/I strains (Figure 5), confirming the defect of phytoene synthase activity when expressed as a single domain. However, in the I to B fusions, the overall carotene production is lower than the one seen with the CrtYB natural enzyme which implies that the CrtI domain might not locate CrtB close enough to the membrane to efficiently release phytoene, but could potentially directly transfer the latter to phytoene desaturase which would then convert it into lycopene, reducing the potential amounts of intermediate carotenoids. This hypothesis is reinforced by the fact that the four linkers tested induce differential effects on the ratio between lycopene and phytoene, possibly evidencing that both proximity and geometry between phytoene synthase and phytoene desaturase domains are important. As to why a small, rigid linker is preferable over the others is still an open question.

4.4.2 Tridomain enzyme

The improvement of lycopene production attained with the B(ek)I fusion, when compared to the combination of individual domains (B/I), prompted us to explore the possibility of optimizing the metabolic efficiency towards β-carotene production with the corresponding tridomain fusion enzymes (Figure 7).

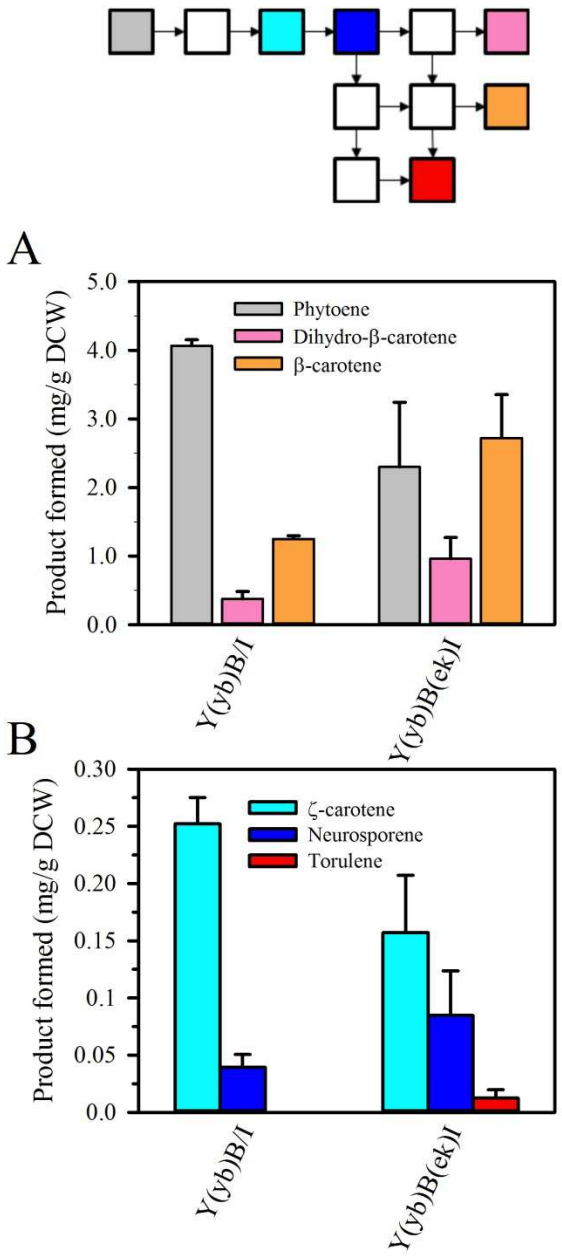

Figure 7. Quantification of carotenoids from the strain bearing tridomain fusion (Y(yb)B(ek)I) compared to the natural system (Y(yb)B/I). **A.** Quantification of major carotenoid molecules. **B.** Quantification of minor carotenoid molecules. Metabolites are colored according to the legend between the panels. The metabolic pathway from phytoene to β-carotene between the two panels with the same color code as previously, white squares correspond to intermediate metabolites that were not detected. Values represent the average ± S.D. of the quantification of each metabolite with three independent cultures.

The comparison between Y(yb)B/I and Y(yb)B(ek)I strains is striking. While the total carotenoid levels are the same in both strains (Supplementary Table S3) the relative amounts of intermediates decreased for carotenoids before ζ-carotene (phytoene and ζ-carotene) and increased for carotenoids after ζ-carotene (neurosporene, torulene, dihydro-β-carotene and β-carotene) (Figure 7A and 7B). As demonstrated above, the GeLC-MS analysis shows the presence of full-length Y(yb)B(ek)I and low evidence of specific proteolysis degradation that would lead to exact domain separation at the linker positions. Therefore, the change of quantity of late metabolites in the pathway suggests a better metabolic efficiency of the Y(yb)B(ek)I strain compared to Y(yb)B/I in pushing the overall flux toward β-carotene. Indeed, the Y(yb)B(ek)I strain was found to produce two times more β-carotene and dihydro-β-carotene than the strain carrying the natural configuration, while, on the contrary, phytoene and ζ-carotene levels dropped about two-fold. This shift in the ratio between the different carotenoid molecules, particularly those that are produced after ζ-carotene, indicates that metabolic bottlenecks have been alleviated. The absence of any accumulation of lycopene in Y(yb)B(ek)I, while the overall metabolic flux is enhanced, indicates that the last desaturation step is outcompeted by cyclisation with the CrtY enzyme. This competition is also evidenced by the accumulation of dihydro-β-carotene which should only originate through the formation of cyclisation of neurosporene followed by desaturation of β-zeacarotene. Interestingly, in the Y/B/I strain, in which the overall level of carotenoids is much smaller than in the natural system Y(yb)B/I or the trifusion Y(yb)B(ek)I, the metabolic flux is almost completely converted to torulene and β-carotene while no other intermediates could be seen. This might signify that the competition between the last desaturation step and cyclisation could depend on the amounts of these intermediates, favoring desaturation at low carotenoid flux and cyclisation when higher fluxes are present. This could also signify that the metabolic pathway to the production of β-carotene and torulene with the *X. dendrorhous* enzymatic system may favor the use of neurosporene over lycopene. Of course, as the trifusion protein should provide a exact 1:1 ratio between the cyclase (Y) and the desaturase (I) proteins, the observed “push” effect could be due to potential lower amounts of the I domain. This seems quite unlikely considering the western blot data and it is thus more probable that the assembly of the three enzymes promotes a faster cyclisation compared to the natural system.

5 Conclusion

This study shows not only the role of spatial proximity for the good performance of a branched enzymatic system but also demonstrates that such enhancements can be conducted in metabolic systems for which enzymes are present in two different subcellular compartments. The various productivities of the bidomain fusion in producing lycopene and the good efficiency of the tridomain fusion in producing β-carotene indicate that the geometry of the full assembly is a critical parameter for enzymatic activity and metabolites diffusion, both being the primary level of the efficiency of metabolic systems. As such, our system is still opened for further improvement, most notably at the characterization of the tridomain fusion but also in refining the linkers between domains, in improving individual enzymatic activities not to mention metabolic and process-based enhancements that have already shown to increase metabolic yields for carotenoids production (Xie et al., 2014, 2015b).

In summary, the system used for spatial enzymatic confinement shows two complementary advantages: it improves the pathway performance but also helps mapping branched metabolic pathways. In general, domain fusion is a fast method to test hypotheses on how spatial enzymatic proximity can improve metabolic pathway efficiency but it may be also complemented by other assembly methods in order to find the optimal combination of good expression level, high enzymatic activity and enzyme stoichiometry.

6 Conflict of Interest

The authors declare no conflict of interest.

7 Author Contributions

Experimental design: HR, TL, CT, AT, GT; experimental work: HR, SCC, TL, LFGA, CT, AT; data analysis and interpretation: all authors; writing, reviewing and editing of the manuscript: all authors; funding acquisition: GT.

8 Funding

This work was in part funded by the MICA and CEPIA department of INRA through the complete financial support of H.R. PhD salary and by the grant 11050480 “Accueil d’Equipes d’Excellence” from the Région Midi-Pyrénées.

9 Acknowledgments

We would like to thank Fayza Daboussi and Mathieu Fournié for critical reading of the manuscript and helpful discussions.

10 Abbreviations

DCW, Dry Cell Weight; HPLC, High Performance Liquid Chromatography; FPP, Farnesyl pyrophosphate; GGPP, geranyl geranyl pyrophosphate.

11 References

Albertsen, L., Chen, Y., Bach, L. S., Rattleff, S., Maury, J., Brix, S., et al. (2011). Diversion of Flux toward Sesquiterpene Production in Saccharomyces cerevisiae by Fusion of Host and Heterologous Enzymes. *Appl. Environ. Microbiol.* 77, 1033–1040. doi:10.1128/AEM.01361-10.

Anand, N. N., Mandal, S., MacKenzie, C. R., Sadowska, J., Sigurskjold, B., Young, N. M., et al. (1991). Bacterial expression and secretion of various single-chain Fv genes encoding proteins specific for a Salmonella serotype B O-antigen. *J. Biol. Chem.* 266, 21874–21879.

Arai, R., Wriggers, W., Nishikawa, Y., Nagamune, T., and Fujisawa, T. (2004). Conformations of variably linked chimeric proteins evaluated by synchrotron X-ray small-angle scattering. *Proteins Struct. Funct. Bioinforma.* 57, 829–838. doi:10.1002/prot.20244.

Arrach, N., Fernández-Martín, R., Cerdá-Olmedo, E., and Avalos, J. (2001). A single gene for lycopene cyclase, phytoene synthase, and regulation of carotene biosynthesis in Phycomyces. *Proc. Natl. Acad. Sci. U. S. A.* 98, 1687–1692. doi:10.1073/pnas.021555298.

Avalos, J. L., Fink, G. R., and Stephanopoulos, G. (2013). Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols. *Nat. Biotechnol.* 31, 335–341. doi:10.1038/nbt.2509.

Baadhe, R. R., Mekala, N. K., Parcha, S. R., and Prameela Devi, Y. (2013). Combination of *ERG9* Repression and Enzyme Fusion Technology for Improved Production of Amorphadiene in *Saccharomyces cerevisiae*. *J. Anal. Methods Chem.* 2013, 1–8. doi:10.1155/2013/140469.

Carquet, M., Pompon, D., and Truan, G. (2015). Transcription interference and ORF nature strongly affect promoter strength in a reconstituted metabolic pathway. *Synth. Biol.* 3, 21. doi:10.3389/fbioe.2015.00021.

Chen, Y., and Nielsen, J. (2013). Advances in metabolic pathway and strain engineering paving the way for sustainable production of chemical building blocks. *Curr. Opin. Biotechnol.* 24, 965–972. doi:10.1016/j.copbio.2013.03.008.

Dahl, R. H., Zhang, F., Alonso-Gutierrez, J., Baidoo, E., Batth, T. S., Redding-Johanson, A. M., et al. (2013). Engineering dynamic pathway regulation using stress-response promoters. *Nat. Biotechnol.* 31, 1039–1046. doi:10.1038/nbt.2689.

Delebecque, C. J., Lindner, A. B., Silver, P. A., and Aldaye, F. A. (2011). Organization of Intracellular Reactions with Rationally Designed RNA Assemblies. *Science*, 1206938. doi:10.1126/science.1206938.

Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V., et al. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* 27, 753–759. doi:10.1038/nbt.1557.

Frances, O., Fatemi, F., Pompon, D., Guittet, E., Sizun, C., Pérez, J., et al. (2015). A Well-Balanced Preexisting Equilibrium Governs Electron Flux Efficiency of a Multidomain Diflavin Reductase. *Biophys. J.* 108, 1527–1536. doi:10.1016/j.bpj.2015.01.032.

Gao, S., Tong, Y., Zhu, L., Ge, M., Jiang, Y., Chen, D., et al. (2017). Production of β-carotene by expressing a heterologous multifunctional carotene synthase in Yarrowia lipolytica. *Biotechnol. Lett.* 39, 921–927. doi:10.1007/s10529-017-2318-1.

Gemmecker, S., Schaub, P., Koschmieder, J., Brausemann, A., Drepper, F., Rodriguez-Franco, M., et al. (2015). Phytoene Desaturase from Oryza sativa: Oligomeric Assembly, Membrane Association and Preliminary 3D-Analysis. *PLOS ONE* 10, e0131717. doi:10.1371/journal.pone.0131717.

Güldener, U., Heck, S., Fielder, T., Beinhauer, J., and Hegemann, J. H. (1996). A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res.* 24, 2519–2524.

Kim, S., Bae, S.-J., and Hahn, J.-S. (2016). Redirection of pyruvate flux toward desired metabolic pathways through substrate channeling between pyruvate kinase and pyruvate-converting enzymes in Saccharomyces cerevisiae. *Sci. Rep.* 6. doi:10.1038/srep24145.

Lin, J.-L., Zhu, J., and Wheeldon, I. (2017). Synthetic Protein Scaffolds for Biosynthetic Pathway Colocalization on Lipid Droplet Membranes. *ACS Synth. Biol.* 6, 1534–1544. doi:10.1021/acssynbio.7b00041.

López, J., Essus, K., Kim, I., Pereira, R., Herzog, J., Siewers, V., et al. (2015). Production of β-ionone by combined expression of carotenogenic and plant CCD1 genes in Saccharomyces cerevisiae. *Microb. Cell Factories* 14, 1–13. doi:10.1186/s12934-015-0273-x.

Lu, M.-S., Fang, Y.-J., Chen, Y.-M., Luo, W.-P., Pan, Z.-Z., Zhong, X., et al. (2014). Higher intake of carotenoid is associated with a lower risk of colorectal cancer in Chinese adults: a case–control study. *Eur. J. Nutr.* 54, 619–628. doi:10.1007/s00394-014-0743-7.

Misawa, N. (2011). Pathway engineering for functional isoprenoids. *Curr. Opin. Biotechnol.* 22, 627–633. doi:10.1016/j.copbio.2011.01.002.

Mizoguchi, T., Harada, J., Yamamoto, K., and Tamiaki, H. (2015). Inactivation of bciD and bchU genes in the green sulfur bacterium Chlorobaculum limnaeum and alteration of photosynthetic pigments in the resultant mutants. *J. Photochem. Photobiol. Chem.* 313, 52–59. doi:10.1016/j.jphotochem.2015.06.012.

Niklitschek, M., Alcaíno, J., Barahona, S., Sepúlveda, D., Lozano, C., Carmona, M., et al. (2008). Genomic organization of the structural genes controlling the astaxanthin biosynthesis pathway

of Xanthophyllomyces dendrorhous. *Biol. Res.* 41, 93–108. doi:10.4067/S0716-97602008000100011.

Pompon, D., F. Garcia-Alles, L., and Truan, G. (2017). "Nanotechnology for Synthetic Biology: Crossroads Throughout Spatial Confinement," in *Nanotechnology in Agriculture and Food Science*, eds. M. A. Axelos and rcel H. V. de V. h.c (Wiley-VCH Verlag GmbH & Co. KGaA), 209–234. doi:10.1002/9783527697724.ch13.

Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G., and Deber, C. M. (2009). Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc. Natl. Acad. Sci.* 106, 1760–1765. doi:10.1073/pnas.0813167106.

Sambrook, J., and Russell, D. W. (2006). SDS-Polyacrylamide Gel Electrophoresis of Proteins. *Cold Spring Harb. Protoc.* 2006, pdb.prot4540. doi:10.1101/pdb.prot4540.

Schaub, P., Yu, Q., Gemmecker, S., Poussin-Courmontagne, P., Mailliot, J., McEwen, A. G., et al. (2012). On the Structure and Function of the Phytoene Desaturase CRTI from Pantoea ananatis, a Membrane-Peripheral and FAD-Dependent Oxidase/Isomerase. *PLoS ONE* 7, e39550. doi:10.1371/journal.pone.0039550.

Schiestl, R. H., and Gietz, R. D. (1989). High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr. Genet.* 16, 339–346.

Schmitt, D. L., and An, S. (2017). Spatial Organization of Metabolic Enzyme Complexes in Cells. *Biochemistry* 56, 3184–3196. doi:10.1021/acs.biochem.7b00249.

Siddiqui, M. S., Thodey, K., Trenchard, I., and Smolke, C. D. (2012). Advancing secondary metabolite biosynthesis in yeast with synthetic biology tools. *FEMS Yeast Res.* 12, 144–170. doi:10.1111/j.1567-1364.2011.00774.x.

Somasundaram, S., Tran, K.-N. T., Ravikumar, S., and Hong, S. H. (2017). Introduction of synthetic protein complex between Pyrococcus horikoshii glutamate decarboxylase and Escherichia coli GABA transporter for the improved production of GABA. *Biochem. Eng. J.* 120, 1–6. doi:10.1016/j.bej.2016.12.020.

Srere, P. A. (1985). The metabolon. *Trends Biochem. Sci.* 10, 109–110.

Takaichi, S. (2000). Characterization of carotenes in a combination of a C18 HPLC column with isocratic elution and absorption spectra with a photodiode-array detector. *Photosynth. Res.* 65, 93–99.

Thomik, T., Wittig, I., Choe, J., Boles, E., and Oreb, M. (2017). An artificial transport metabolon facilitates improved substrate utilization in yeast. *Nat. Chem. Biol.* 13, 1158–1163. doi:10.1038/nchembio.2457.

Tokuhiro, K., Muramatsu, M., Ohto, C., Kawaguchi, T., Obata, S., Muramoto, N., et al. (2009). Overproduction of Geranylgeraniol by Metabolically Engineered Saccharomyces cerevisiae. *Appl. Environ. Microbiol.* 75, 5536–5543. doi:10.1128/AEM.00277-09.

Tsumoto, K., Nakaoki, Y., Ueda, Y., Ogasahara, K., Yutani, K., Watanabe, K., et al. (1994). Effect of the Order of Antibody Variable Regions on the Expression of the Single-Chain HyHEL10 Fv Fragment in Escherichia coli and the Thermodynamic Analysis of Its Antigen-Binding Properties. *Biochem. Biophys. Res. Commun.* 201, 546–551. doi:10.1006/bbrc.1994.1736.

Verdoes, J. C., Krubasik, P., Sandmann, G., and Van Ooyen, A. J. J. (1999). Isolation and functional characterisation of a novel type of carotenoid biosynthetic gene from Xanthophyllomyces dendrorhous. *Mol. Gen. Genet. MGG* 262, 453–461.

Verdoes, J. C., Sandmann, G., Visser, H., Diaz, M., Mossel, M. van, and Ooyen, A. J. J. van (2003). Metabolic Engineering of the Carotenoid Biosynthetic Pathway in the Yeast Xanthophyllomyces dendrorhous (Phaffia rhodozyma). *Appl. Environ. Microbiol.* 69, 3728–3738. doi:10.1128/AEM.69.7.3728-3738.2003.

Verwaal, R., Wang, J., Meijnen, J.-P., Visser, H., Sandmann, G., Berg, J. A. van den, et al. (2007). High-Level Production of Beta-Carotene in Saccharomyces cerevisiae by Successive Transformation with Carotenogenic Genes from Xanthophyllomyces dendrorhous. *Appl. Environ. Microbiol.* 73, 4342–4350. doi:10.1128/AEM.02759-06.

Xie, W., Liu, M., Lv, X., Lu, W., Gu, J., and Yu, H. (2014). Construction of a controllable β-carotene biosynthetic pathway by decentralized assembly strategy in Saccharomyces cerevisiae. *Biotechnol. Bioeng.* 111, 125–133. doi:10.1002/bit.25002.

Xie, W., Lv, X., Ye, L., Zhou, P., and Yu, H. (2015a). Construction of lycopene-overproducing Saccharomyces cerevisiae by combining directed evolution and metabolic engineering. *Metab. Eng.* 30, 69–78. doi:10.1016/j.ymben.2015.04.009.

Xie, W., Lv, X., Ye, L., Zhou, P., and Yu, H. (2015b). Construction of lycopene-overproducing Saccharomyces cerevisiae by combining directed evolution and metabolic engineering. *Metab. Eng.* 30, 69–78. doi:10.1016/j.ymben.2015.04.009.

Xie, W., Ye, L., Lv, X., Xu, H., and Yu, H. (2015c). Sequential control of biosynthetic pathways for balanced utilization of metabolic intermediates in Saccharomyces cerevisiae. *Metab. Eng.* 28, 8–18. doi:10.1016/j.ymben.2014.11.007.

Zhang, T., Lei, J., Yang, H., Xu, K., Wang, R., and Zhang, Z. (2011). An improved method for whole protein extraction from yeast Saccharomyces cerevisiae. *Yeast* 28, 795–798. doi:10.1002/yea.1905.