Bottom-Up Metabolic Reconstruction of Arabidopsis and Its Application to Determining the Metabolic Costs of Enzyme Production

Anne Arnold and Zoran Nikoloski

Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam, Germany

Supplemental Text

List of Tables

S1 Model characteristics

The assembled model is a large-scale metabolic network of Arabidopsis following the bottom-up reconstruction process (Supplemental Data S2 – Arabidopsis core model). In contrast to genome-scale networks, the focus lies on the functionality and completeness of the metabolic objective, *e.g.*, amino acid synthesis, and the underlying pathways (Figure 2, Supplemental Data S3 – Reaction list). The included reactions are assigned according to AraCyc 11.5 (Mueller *et al.*, 2003). Exceptions are the light-dependent reactions, the importer and exporter, and the internal transport reactions. Moreover, the consideration of different environmental conditions and their effects on the cellular scenario are a main aspect of this study. The latter issue is modeled by different biomass compositions, one for each cellular scenario.

S1.1 Light-dependent reactions

Metabolic databases, such as AraCyc (Mueller *et al.*, 2003) and KEGG (Kanehisa and Goto, 2000), do not incorporate the proton pump into the thylakoid lumen in the context of the electron transport chain as well as the one of the ATP synthase back into the stroma. Four electrons enter the noncyclic electron transport chain per dissociated water molecule, and are utilized to form two NADPH molecules. Moreover, 12 protons in total are pumped into the lumen (Allen, 2002). On the other hand, the corresponding ATP synthase needs 14 protons for a full rotation forming three molecules of ATP (Seelert *et al.*, 2000). In contrast, the mitochondrial ATP synthase has a 12-fold symmetry of proton-powered turbine forming turbine.

S1.2 Import and export reactions

The corresponding reactions are initially assigned as reversible and are then restricted to achieve a minimal in- and efflux of the system. For the first time, the uptake of the required ions nitrate, sulfate and phosphate is coupled with energy demand. It is known that the uptake of these ions is a proton coupled co-transport. To maintain the intercellular pH value those protons have to be exported, again. Assuming that the pH value is only regulated by an ATP-driven proton pump, the the costs of exporting each proton can be approximated by one ATP. The stoichiometry of the co-transport differs for the different ions: 2 H^+/NO_3^- (Pate and Layzell, 1990; Meharg and Blatt, 1995), 3 H^+/SO_4^{2-} (Buchner *et al.*, 2004), and 2–4 $H^+/H_2PO_4^-$ (Raghothama, 1999). In the case of Pi, the determined physiological charge formula is HPO_4^{2-} which entails an adaption of the co-transport stoichiometry. Moreover, it is experimentally confirmed that the lower the phosphate concentration, the greater the ratio of protons and phosphate ions (Sakano, 1990). For this purpose, we assigned to the import of phosphate a cost of three ATP.

S1.3 Internal transporter

For the incorporation of transport reactions the aim is twofold: include a minimal set of reactions, and avoid biologically unverified transporters. Therefore, we started with confirmed transport reactions, extended the list as far as required, and inspected it subsequently to remove redundant transporter if possible. The respective references are given in the reaction list (Supplemental Data S3 – Reaction list).

S1.4 Biomass reactions

To reinforce the biochemical reliability of a metabolic model, the careful assembly of the metabolic function to examine is of great importance. Therefore, we assembled three biomass compositions simulating growing cells exposed under carbon-limiting, nitrogen-limiting and optimal growth conditions. We would like to highlight again that, in contrast to the biomass compositions utilized in the other Arabidopsis reconstructions, the here presented compositions are experimentally obtained predominantly from plant fresh weight and, subsequently, converted into dry weight (DW) in a condition-specific manner (Table S1). We accounted for all major cell components,

Table S1: Condition-specific dry weight content (DWC; $\left[\frac{\text{g DW}}{\text{g FW}}\right]$) of Arabidopsis.

Condition	DWC	Reference
Optimal growth	0.088	(Tschoep *et al.*, 2009, high N)
Carbon limitation	0.088	(Pyl *et al.*, 2012, data set S2: 24°C/16°C)
Nitrogen limitation	0.083	(Tschoep *et al.*, 2009, low N)

namely cell wall, proteins, lipids, soluble metabolites, starch, DNA and RNA. All of these components are incorporated via representative metabolites or precursors (Tables S2–S5). For instance, cell wall and starch are represented by cellulose and amylose, respectively, while a Glc dimer precursor of each is incorporated. Similarly, lipids, proteins, DNA and RNA are included. Regarding the lipids, we considered palmitic acid as a fatty acid representative and incorporated malonyl-acyl carrier protein (ACP) as the corresponding precursor, whereas for cell protein, and DNA and RNA only components, namely amino acids and nucleotides, respectively, are considered as representatives.

The experimental basis is the condition-specific metabolomics data set of Sulpice *et al.* (2013) providing data for total protein, soluble metabolites and starch (Tables S2–S4). The soluble metabolite data are measured by GC-MS which, consequently, first have to be converted into absolute levels by means of calibration curves for each metabolite (Table S2, Supplemental Data S4 – GC-MS data). For the protein composition, we used a proteomics data set of Mooney *et al.* (2006). To this end, we determined the amino acid composition of the identified proteins in Arabidopsis leaves and divided the total amount of proteins into amino acid fractions (see Supplemental Data S5 – 2D-Gel data). Combined with the total protein data of Sulpice *et al.* (2013), we assigned the allocation to the amino acids (Table S3). As there is no condition-specific proteomics data set available, the resulting partition into amino acids was used for all conditions. The ultimate amino acid fraction in the biomass composition is the combination of free and protein-bound amino acids (Tables S2 and S3, Supplemental Data S3 – Reaction list). In a similar condition-unspecific manner also the remaining components were incorporated, namely cell wall, lipids, and DNA and RNA. The cell wall fraction was approximated by 118 µg cellulose mg^{-1} DW (DeBolt *et al.*, 2009, Fig. 3A) which was converted into 363.88 µmol Glc units of cellulose g^{-1} DW by considering a molar mass of anhydroglucose, 162.14 g mol^{-1} (Smith and Zeeman, 2006) (Table S4). For the lipid content, the amount of fatty acids was used amounting to 3.3 mg g^{-1} FW in leaves of wild type Arabidopsis (Doermann *et al.*, 1995). Based on the specific dry weight content (Table S1), the molar mass of palmitic acid as the major fatty acid, 256.42 g mol^{-1}, and its number of carbon atoms, 16, we calculated the corresponding amount of the three-carbon fatty acid precursor malonyl-ACP, 146.24 and 155.05 µmol g^{-1} DW, respectively, which ultimatively serves as the lipid representative (Table S4). Finally, 0.092 mg DNA g^{-1} FW is measured in Arabidopsis leaves (Sharrock and Clack, 2002) which we also assumed to be an appropriate level for RNA. Based on the DNA and cDNA sequence of Arabidopsis, we determined the nucleotide fraction which was converted into the respective nucleotide levels by means of the molar mass of each nucleotide (Table S5).

Overall, our findings indicate that the three biomass compositions differ in their coverage in a range of 495.08 to 695.78 µg mg^{-1} DW (Table S7). This coverage specifies the fraction of measured biomass components to the reference fresh weight and results from converting the concentrations of all constituting metabolites into mass fraction and, subsequently, summing them up.

S1.5 Extended GPR associations

Beyond the common features of a metabolic reconstruction, the Arabidopsis core model provides the stoichiometric subunit compositions of the included enzymes. In the existing genome-scale models, the GPR associations comprise the different genes assigned for each reaction and their linkage, whether they encode alternative or necessary proteins. Accordingly, the relation between isoforms are denoted by logical OR-operators and subunit

Table S2: Soluble metabolite representatives measured in FW and converted into DW for model integration. The complete data set is provided as Supplemental Data S4 – GC-MS data.

	Metabolite	Carbon			Nitrogen			Optimal		
		$\left[\frac{\mu g}{g\ DW}\right]$	$\left[\frac{nmol}{g\ DW}\right]$	$\left[\frac{nmol}{g\ FW}\right]$	$\left[\frac{\mu g}{g\ DW}\right]$	$\left[\frac{nmol}{g\ DW}\right]$	$\left[\frac{nmol}{g\ FW}\right]$	$\left[\frac{\mu g}{g\ DW}\right]$	$\left[\frac{nmol}{g\ DW}\right]$	$\left[\frac{nmol}{g\ FW}\right]$
free Amino acids	*Ala*	59.20	664.43	58.47	10.26	115.16	9.56	11.18	125.47	11.04
	Arg	606.75	3483.04	306.51	6882.91	39511.56	3279.46	1218.57	6995.22	615.58
	Asn	196.53	1487.49	130.90	1119.68	8474.71	703.40	433.54	3281.42	288.76
	Asp	95.62	718.39	63.22	243.81	1831.76	152.04	98.56	740.46	65.16
	Cys									
	Gln	154.58	1057.76	93.08	1216.50	8324.18	690.91	214.16	1465.46	128.96
	Glu	182.32	1239.17	109.05	527.34	3584.19	297.49	371.84	2527.29	222.40
	Gly	88.36	1177.02	103.58	9.19	122.35	10.16	60.08	800.32	70.43
	His									
	Ile	15.28	116.52	10.25	42.24	322.04	26.73	17.31	131.97	11.61
	Leu	12.19	92.97	8.18	30.70	234.06	19.43	12.73	97.06	8.54
	Lys	35.02	239.58	21.08	275.01	1881.16	156.14	43.04	294.44	25.91
	Met	69.14	463.36	40.78	157.36	1054.64	87.54	39.21	262.82	23.13
	Phe	89.17	539.83	47.50	223.73	1354.40	112.41	125.92	762.28	67.08
	Pro	51.64	448.53	39.47	42.46	368.78	30.61	60.72	527.37	46.41
	Ser	128.72	1224.90	107.79	83.13	791.07	65.66	87.52	832.79	73.29
	Thr	1035.03	8688.94	764.63	777.43	6526.40	541.69	778.63	6536.54	575.22
	Trp	52.10	255.11	22.45	302.80	1482.62	123.06	68.84	337.09	29.66
	Tyr	76.86	424.19	37.33	112.54	621.13	51.55	55.39	305.70	26.90
	Val	20.32	173.43	15.26	31.03	264.85	21.98	21.18	180.79	15.91
Sugars	Fru	2661.71	14774.15	1300.13	1922.14	10669.09	885.53	2168.21	12034.90	1059.07
	Glc	6098.31	33849.42	2978.75	4240.82	23539.17	1953.75	3329.12	18478.68	1626.12
	Mas	779.56	2277.42	200.41	1841.45	5379.65	446.51	893.26	2609.58	229.64
	Suc	15539.82	45398.23	3995.04	29785.27	87015.11	7222.25	10497.78	30668.37	2698.82
TCA	Fum	9005.16	77583.91	6827.38	1344.95	11587.40	961.75	7800.75	67207.27	5914.24
	Mal	5110.44	38111.99	3353.86	1461.39	10898.56	904.58	6838.11	50996.41	4487.68
	SCA	153.75	1301.99	114.57	192.48	1629.92	135.28	116.56	987.09	86.86
Others	GABA	15.78	153.05	13.47	60.14	583.23	48.41	9.69	94.01	8.27
	Orn	1930.50	14607.27	1285.44	1274.92	9646.82	800.69	209.68	1586.54	139.62
	SA	54.32	311.89	27.45	33.19	190.59	15.82	66.05	379.24	33.37
	Tre	438.18	1158.18	101.92	1068.54	2824.36	234.42	218.43	577.36	50.81
	urea	332.23	5531.64	486.78	622.54	10365.35	860.32	170.60	2840.43	249.96

connections by logical AND-operators.

For the calculation of enzyme costs, the exact complex structure of an enzyme is required so that we additionally provide the stoichiometry of the subunit assembly. This is easily realized by an additional factor for each multimer, *e.g.*, the large and small subunit term of RuBisCO is complemented by factor eight (Equation S1).

$$8 * (\texttt{ATCG00490})\ \texttt{AND}\ 8 * (\texttt{AT1G67090 OR AT5G38430 OR AT5G38420 OR AT5G38410}) \tag{S1}$$

The complex structure information is only rarely available and, moreover, often contradicting across the different databases, AraCyc (Mueller *et al.*, 2003), UniProt (The UniProt Consortium, 2012) and BRENDA (Schomburg *et al.*, 2013). As far as possible, we tried to confirm the utilized information by an additional reference (Supplemental Data S3 – Reaction list).

S2 Protein assembly costs

The costs for the enzyme synthesis can be decomposed into two parts, (1) the amino acid biosynthesis and (2) the protein assembly. Obviously, the costs for the biosynthesis of each amino acid differ with the underlying

Table S3: Fractional amount of protein-bound amino acids of Arabidopsis leaf based on measured total protein (Sulpice *et al.*, 2013, last row).

Amino acid	Fraction	Carbon			Nitrogen			Optimal		
	[%]	$\left[\frac{\text{mg}}{\text{g DW}}\right]$	$\left[\frac{\mu\text{mol}}{\text{g DW}}\right]$	$\left[\frac{\mu\text{mol}}{\text{g FW}}\right]$	$\left[\frac{\text{mg}}{\text{g DW}}\right]$	$\left[\frac{\mu\text{mol}}{\text{g DW}}\right]$	$\left[\frac{\mu\text{mol}}{\text{g FW}}\right]$	$\left[\frac{\text{mg}}{\text{g DW}}\right]$	$\left[\frac{\mu\text{mol}}{\text{g DW}}\right]$	$\left[\frac{\mu\text{mol}}{\text{g FW}}\right]$
Ala	8.4253	18.42	206.80	18.20	31.26	350.90	29.12	20.54	230.55	20.29
Arg	4.8366	10.58	60.71	5.34	17.95	103.02	8.55	11.79	67.69	5.96
Asn	3.6452	7.97	60.33	5.31	13.53	102.38	8.50	8.89	67.26	5.92
Asp	5.476	11.97	89.97	7.92	20.32	152.66	12.67	13.35	100.30	8.83
Cys	1.5465	3.38	27.91	2.46	5.74	47.36	3.93	3.77	31.12	2.74
Gln	2.9852	6.53	44.67	3.93	11.08	75.80	6.29	7.28	49.80	4.38
Glu	6.872	15.03	102.14	8.99	25.50	173.31	14.38	16.75	113.87	10.02
Gly	8.3967	18.36	244.59	21.52	31.16	415.04	34.45	20.47	272.68	24.00
His	2.1157	4.63	29.82	2.62	7.85	50.60	4.20	5.16	33.24	2.93
Ile	4.826	10.55	80.46	7.08	17.91	136.52	11.33	11.77	89.69	7.89
Leu	8.5374	18.67	142.33	12.52	31.68	241.51	20.05	20.81	158.67	13.96
Lys	6.1944	13.55	92.66	8.15	22.98	157.23	13.05	15.10	103.30	9.09
Met	2.0413	4.46	29.92	2.63	7.57	50.76	4.21	4.98	33.35	2.93
Phe	4.4653	9.76	59.11	5.20	16.57	100.30	8.33	10.89	65.90	5.80
Pro	5.0772	11.10	96.43	8.49	18.84	163.63	13.58	12.38	107.51	9.46
Ser	6.3223	13.83	131.56	11.58	23.46	223.23	18.53	15.41	146.67	12.91
Thr	6.3592	13.91	116.74	10.27	23.60	198.09	16.44	15.50	130.15	11.45
Trp	1.3851	3.03	14.83	1.31	5.14	25.17	2.09	3.38	16.53	1.45
Tyr	3.4726	7.59	41.91	3.69	12.89	71.12	5.90	8.47	46.72	4.11
Val	7.0199	15.35	131.04	11.53	26.05	222.35	18.46	17.11	146.09	12.86
Total	100	218.68	1803.92	158.75	371.06	3060.97	254.06	243.79	2011.09	176.98

Table S4: Starch, cell wall and lipid precursors.

Biomass precursor	Carbon			Nitrogen			Optimal		
	$\left[\frac{\mu\text{g}}{\text{g DW}}\right]$	$\left[\frac{\mu\text{mol}}{\text{g DW}}\right]$	$\left[\frac{\mu\text{mol}}{\text{g FW}}\right]$	$\left[\frac{\mu\text{g}}{\text{g DW}}\right]$	$\left[\frac{\mu\text{mol}}{\text{g DW}}\right]$	$\left[\frac{\mu\text{mol}}{\text{g FW}}\right]$	$\left[\frac{\mu\text{g}}{\text{g DW}}\right]$	$\left[\frac{\mu\text{mol}}{\text{g DW}}\right]$	$\left[\frac{\mu\text{mol}}{\text{g FW}}\right]$
Starch (Glc dimer)	73.72	227.35	20.01	108.81	335.54	27.85	95.62	294.86	25.95
Cellulose (Glc dimer)	118.00	363.88	32.02	118.00	363.88	30.20	118.00	363.88	32.02
Fatty acid (Malonyl-ACP)	37.50	146.24	12.87	39.76	155.05	12.87	37.50	146.24	12.87

Table S5: Fractional amount of nucleotides of DNA and RNA.

Nucleotide	Total	w/o N[b]	Carbon			Nitrogen			Optimal		
	[%]	[%]	$\left[\frac{\text{mg}}{\text{g DW}}\right]$	$\left[\frac{\text{nmol}}{\text{g DW}}\right]$	$\left[\frac{\text{nmol}}{\text{g FW}}\right]$	$\left[\frac{\text{mg}}{\text{g DW}}\right]$	$\left[\frac{\text{nmol}}{\text{g DW}}\right]$	$\left[\frac{\text{nmol}}{\text{g FW}}\right]$	$\left[\frac{\text{mg}}{\text{g DW}}\right]$	$\left[\frac{\text{nmol}}{\text{g DW}}\right]$	$\left[\frac{\text{nmol}}{\text{g FW}}\right]$
DNA	99.84	100	1.05			1.11			1.05		
dATP	31.94	31.99	0.33	680.91	59.92	0.35	721.93	59.92	0.33	680.91	59.92
dCTP	18.01	18.04	0.19	403.66	35.52	0.20	427.98	35.52	0.19	403.66	35.52
dGTP	17.99	18.02	0.19	371.41	32.68	0.20	393.78	32.68	0.19	371.41	32.68
dTTP	31.90	31.95	0.33	692.81	60.97	0.35	734.55	60.97	0.33	692.81	60.97
RNA[c]	99.99	100	1.05			1.11			1.05		
ATP	28.41	28.41	0.30	585.59	51.53	0.31	620.86	51.53	0.30	585.59	51.53
CTP	22.94	22.94	0.24	496.38	43.68	0.25	526.28	43.68	0.24	496.38	43.68
GTP	20.26	20.26	0.21	404.78	35.62	0.22	429.16	35.62	0.21	404.78	35.62
UTP	28.39	28.39	0.30	613.07	53.95	0.31	650.00	53.95	0.30	613.07	53.95

[b] – IUPAC nucleotide code: any base

[c] – based on cDNA

biochemical pathways and, consequently, this former cost component varies for each protein which shows difference in the amino acid sequence. Therefore, we estimated the costs based on the provided metabolic network. In contrast, the costs for the protein can be approximated per amino acid or polypeptide as either the

Table S6: Elemental composition of the three biomass functions.

Cellular scenario	Elemental composition $\left[\frac{mmol}{g\ DW}\right]$					
	C	H	N	O	P	S
Optimal growth	20.621	35.505	2.595	15.697	0.013	0.065
Carbon limitation	19.192	32.767	2.350	14.721	0.013	0.058
Nitrogen limitation	26.737	48.002	4.104	19.261	0.014	0.099

Table S7: Biomass coverage and free nitrate content for all three conditions.

	Carbon limiting		Nitrogen limiting		Optimal growth	
	$\left[\frac{\text{mg}}{\text{g DW}}\right]$	$\left[\frac{\text{mmol}}{\text{g DW}}\right]$	$\left[\frac{\text{mg}}{\text{g DW}}\right]$	$\left[\frac{\text{mmol}}{\text{g DW}}\right]$	$\left[\frac{\text{mg}}{\text{g DW}}\right]$	$\left[\frac{\text{mmol}}{\text{g DW}}\right]$
Total biomass	495.08	3.44	695.78	4.84	533.03	3.67
Free NO_3^-	38.05	0.61	4.45	0.07	70.30	1.13

process is the same for each amino acid or proceeded only once per polypeptide. Besides the actual protein biosynthesis, we took into account costs arising from protein maturation.

As mentioned above, the biosynthesis comprises four energy demanding steps: (1) the amino acid activation, (2) the formation of the initiation complex, (3) the amino-acid-sequence elongation, and (4) the finalizing of the polypeptide within the termination phase. Before an amino acid is capable of being included into protein biosynthesis, it has to be activated by forming the corresponding aminoacyl-tRNA. This activation process requires one molecule of ATP and releases one molecule of AMP and pyrophosphate, respectively. The recycling of AMP into ADP requires a second molecule of ATP, so that for convenience the cost can be estimated as two molecules ATP (Formulas S2–S4, Table S8):

$$\text{aa} + \text{ATP} + \text{tRNA} \rightleftharpoons \text{aa-tRNA} + \text{AMP} + \text{PPi}, \tag{S2}$$

$$\text{AMP} + \text{ATP} \rightleftharpoons 2\,\text{ADP}, \tag{S3}$$

$$\text{PPi} + \text{H}_2\text{O} \rightarrow 2\,\text{Pi}. \tag{S4}$$

The actual biosynthesis starts with the formation of the initiation complex. Thereby, first the small ribosome subunit, the translation initiation factors and the ternary complex including one molecule GTP form the 43S preinitiation complex (Jackson *et al.*, 2010). By means of further translation initiation factors the mRNA is ATP-dependent activated and, subsequently, attached to build the 48S initiation complex. Finally, the large ribosome subunit and a further initiation factor comprising a further GTP are added to form the 80S initiation complex whereby the associated initiation factors are displaced (Jackson *et al.*, 2010). In the end, both GTP are hydrolyzed and released and the nucleoside-diphosphate kinase (EC 2.7.4.6) catalyzes the recycling of GTP via ATP such that the costs for the initiation can be approximated by three molecules ATP per polypeptide (Formulas S5 and S6, Table S8):

$$40\text{S} + 60\text{S} + \text{mRNA} + \text{eIFs} + 2\,\text{GTP} + \text{ATP} \rightarrow 80\text{S complex} + \text{eIFs} + 2\,\text{GDP} + \text{ADP} + 3\,\text{Pi}, \tag{S5}$$

$$2\,\text{GDP} + 2\,\text{ATP} \rightleftharpoons 2\,\text{GTP} + 2\,\text{ADP}. \tag{S6}$$

During an elongation cycle, first an elongation factor binds the aminoacyl-tRNA forming the ternary complex and, subsequently, releases the aminoacyl-tRNA into the so-called aminoacyl (A) site of the ribosome and catalyzes thereby the peptide bond formation. Secondly, another elongation factor catalyzes the translocation of the peptidyl-tRNA into the peptidyl (P) site while the deacylated tRNA is pushed to the exit (E) site. The application of both elongation factors requires one molecule of GTP released as GDP, respectively. By means of the nucleoside-diphosphate kinase the cost of the elongation cycle can be estimated with two molecules ATP per

peptide bound (Formulas S7 and S8, Table S8):

$$80\text{S complex} + \text{aa-tRNA} + \text{eEFs} + 2\,\text{GTP} \rightarrow \text{ribosome} + \text{eEFs} + 2\,\text{GDP} + 2\,\text{Pi}, \tag{S7}$$

$$2\,\text{ATP} + 2\,\text{GDP} \rightleftharpoons 2\,\text{ADP} + 2\,\text{GTP}. \tag{S8}$$

The finalizing release of the polypeptide sequence in the termination phase requires the binding of a termination factor. Again, this factor is associated with a molecule of GTP which is hydrolyzed before the factor is released. Recycling the GDP to GTP, the costs for the termination phase can be determined as one molecule ATP per polypeptide (Formulas S9 and S10, Table S8):

$$\text{ribosome} + \text{eRF} + \text{GTP} \rightarrow \text{post-termination complex} + \text{peptide} + \text{eRF} + \text{GDP} + \text{Pi}, \tag{S9}$$

$$\text{ATP} + \text{GDP} \rightleftharpoons \text{ADP} + \text{GTP}. \tag{S10}$$

Furthermore, we accounted for the costs of protein maturation including the energy demand for maintenance processes, the secretion of signaling sequences and post-translational modifications. Altogether, those processes consume approximately one molecule ATP per amino acid as specified in Table S8. Consequently, the overall costs for the protein assembly can be estimated as five molecules of ATP hydrolyzed per amino acid of synthesized protein.

Table S8: Itemization of the protein assembly costs comprising the protein biosynthesis and the maturation.

Process	Costs [ATP]		References
Protein biosynthesis			
Amino acid activation	2	per amino acid	(Zerihun *et al.*, 1998)
Initiation	3	per polypeptide	(Jackson *et al.*, 2010)
Elongation – peptide bond formation and translocation	2	per peptide bond	(Zerihun *et al.*, 1998)
Termination	1	per polypeptide	(Zerihun *et al.*, 1998)
Tool maintenance	0.16	per amino acid	(Zerihun *et al.*, 1998)
Signal sequences – $0.7 \cdot 0.18 \cdot$ (costs from biosynthesis and biodegradation of polypeptide)	0.64	per amino acid	(Zerihun *et al.*, 1998)
Amino acid turnover	0	per peptide bond	(de Visser *et al.*, 1992)
Post-translational processing			
Methylation, acetylation, etc.	0.1	per peptide bond	(de Visser *et al.*, 1992)
Phosphorylation, *e.g.* enzyme (de)activation	0.1	per peptide bond	(de Visser *et al.*, 1992)
Protein biodegradation	1	per peptide bond	(Zerihun *et al.*, 1998)
Total	≈ 5	per amino acid	(Noguchi *et al.*, 2001)

S3 Analyses and supplementary findings

The analyses within this study are based on constraint-based modeling techniques, in particular flux balance analysis (FBA; Orth *et al.*, 2010). By means of FBA, the optimal synthesis and/or consumption rate of specific metabolites or combinations thereof can be determined. The the optimization of a flux, v with respect to an objective, c, only requires: (1) the stoichiometry of the biochemical reactions, summarized in matrix S, (2) the reaction directionality and reversibility, denoted by lower and upper bounds, v_{min} and v_{max} respectively, and (3) the steady-state assumption meaning that there is no accumulation or depletion of any metabolites, $S \cdot v = 0$. The mathematical formulation of the underlying linear program reads as follows:

$$\begin{array}{ll} \max\,/\,\min & c^T v \\ \text{s.t.} & S \cdot v = 0 \\ & v_{min} \leq v \leq v_{max} \end{array} \quad . \tag{S11}$$

The result of the computation is one set of reaction fluxes which obey these constraints and produce the maximal or minimal objective value, one flux distribution. Importantly, the solution may not be unique, *i.e.*, the optimal value might arise for different flux distributions.

S3.1 Constraints and boundaries of the flux balance analyses

All optimizations are based on the photoautotrophic scenario, namely only inorganic carbon, nitrogen, sulfur and phosphorus sources are available, and light is the only energy source. The exact import and export model boundaries are listed in Table S9.

Table S9: Model boundary constraints for photoautotrophic scenario. Columns marked with *lb* and *ub* represent the default lower and upper boundary, respectively.

	Reaction	*lb*	*ub*
98	`G6PDH_h`	0	0
136	`PPIF6PK_c`	0	0
446	`Im_hnu`	0	Inf
447	`Im_CO2`	-Inf	Inf
448	`Im_H2O`	-Inf	Inf
449	`Im_Pi`	0	Inf
450	`Im_NO3`	0	Inf
451	`Im_NH4`	0	Inf
452	`Im_SO4`	0	Inf
453	`Im_H2S`	0	Inf
454	`Ex_O2`	-Inf	Inf
455–534	`Ex_Ala_c` – `Ex_Val_p`	0	0
535–542	`Ex_starch` – `Ex_Tre`	0	0
546–549	`Bio_AA` – `Bio_opt`	0	0

To compare the compartmentalized Arabidopsis models, we determined the maximum biomass production based on 1000 µmol CO_2 (Tables 1 and S10, no. 1). Beforehand, we deactivated the unneeded import and export reactions to restrict the model boundaries to reactions listed in table S9. Unfortunately, a comparison based on restricted photon import is not possible for the model of Mintz-Oron *et al.* (2012) since this restriction has no effect on the production of biomass in this model. Surprisingly, even without any photon import (and no other energy source) this model is able to produce biomass.

The determination of the minimum ATP consumption for cell performance, amino acid and RuBisCO costs was realized by a three-step optimization (Algorithm S1). First, the minimum amount of required energy to produce the metabolite(s) of interest, in terms of photons, was computed (Algorithm S1 LP1; Table S10, no. 7, 10, 18, 21) and utilized as additional constraint in the subsequent optimization steps. Thereby, the amount of required light energy was translated into consumed metabolic energy equivalents. To this end, we maximized the sum of the nongrowth-associated maintenance functions (NGAM) representing the unspecified ATP consuming reactions in the system. On the one hand, by disabling the metabolite's production of interest we determined the maximum conversion rate of photons into ATP (Algorithm S1 LP2; Table S10, no. 8, 11, 19, 22), and, on the other hand, by enabling the metabolite's production we were able to calculate the ATP surplus which can be utilized for other processes (Algorithm S1 LP3; Table S10, no. 9, 12, 20, 23). Accordingly, the number of consumed ATP results from the maximum amount of ATP that could be synthesized by the provided photons from which the ATP excess by producing the metabolite(s) of interest is subtracted. We note that such multiple-step optimizations may lead to solutions for the second optimization step which are suboptimal in comparison to the case when the solution of the second step would be obtained independently of the first.

Algorithm S1: Three-step optimization of minimum ATP consumption for synthesizing the metabolite(s) of interest.

WLOG:
Import and export reactions are assigned as left-to-right operating such that the transport has positive flux
Input:
Metabolic network of the Arabidopsis core model, including
stoichiometric matrix, $\mathbf{S}$,
vector of lower and upper flux boundaries, $\mathbf{v}^{min}$ and $\mathbf{v}^{max}$
Initial constraints, *cond* (Table S9)
Output: minimum ATP consumption for synthesizing the metabolite(s) of interest, $A_{\min}$

Begin

(De)Activate reactions according to *cond*,
Assign the flux of the export reaction of the respective metabolite(s) of interest, v_s, to one,
Minimize the flux through the photon import reaction, $v_e = \{\texttt{'Im\_hnu'}\}$:

$$z_1 = \min \mathbf{c}^T\mathbf{v}\ , \qquad c_i = \begin{cases} 1 & \text{if } i \in e \\ 0 & \text{otherwise.} \end{cases}$$
$$s.t. \quad \mathbf{S}\cdot\mathbf{v} = 0 \qquad \text{(LP1)}$$
$$v_s^{min} = v_s^{max} = 1$$
$$\mathbf{v}^{min} \leq \mathbf{v} \leq \mathbf{v}^{max}$$

Define *minimum energy requirement* as $E_{\min} = z_1$

Assign the flux of the photon import reaction, v_e, to $E_{\min}$,
Maximize the flux through the maintenance reactions, v_m:

$$z_2 = \max \mathbf{c}^T\mathbf{v}\ , \qquad c_i = \begin{cases} 1 & \text{if } i = m \\ 0 & \text{otherwise.} \end{cases}$$
$$s.t. \quad \mathbf{S}\cdot\mathbf{v} = 0 \qquad \text{(LP2)}$$
$$v_e^{min} = v_e^{max} = E_{\min}$$
$$\mathbf{v}^{min} \leq \mathbf{v} \leq \mathbf{v}^{max}$$

Define *maximum conversion rate of photons int ATP* as $A_{\text{tot}} = z_2$,

Assign the flux of the export reaction of the respective metabolite(s) of interest, v_s, to one,
Assign the flux of the photon import reaction, v_e, to $E_{\min}$,
Maximize the flux through the maintenance reactions, v_m:

$$z_3 = \max \mathbf{c}^T\mathbf{v}\ , \qquad c_i = \begin{cases} 1 & \text{if } i \in m \\ 0 & \text{otherwise.} \end{cases}$$
$$s.t. \quad \mathbf{S}\cdot\mathbf{v} = 0 \qquad \text{(LP3)}$$
$$v_s^{min} = v_s^{max} = 1$$
$$v_e^{min} = v_e^{max} = E_{\min}$$
$$\mathbf{v}^{min} \leq \mathbf{v} \leq \mathbf{v}^{max}$$

Define *ATP surplus* as $A_{\text{sur}} = z_3$,
Set $A_{\min} = A_{\text{tot}} - A_{\text{sur}}$

End

S3.2 Flux variability analysis

As a result of the bottom-up design, the model relies only on annotations from Arabidopsis and does not comprise any dead-ends and blocked reactions. On the other hand, it covers only a limited number of alternative pathways which means that the predicted flux distributions are not as flexible as genome-scale models (de Oliveira Dal'Molin

Table S10: Compendium of performed flux balance analyses. Columns marked with c, lb and ub represent the objective coefficients, the lower and upper boundary for a minimization with the glpk solver, respectively. The initial model boundaries are given in Table S9.

Description		Objective		Constraints			Comment	No.
		rxn	c	rxn	lb	ub		
Compare maximal biomass production		AraBio[a]	-1	`Im_CO2`	1000	1000		1
Cell performance	Minimize photon consumption	`Im_hnu`	1	Bio[b]	1	1		2
	Minimize CO_2 consumption	`Im_CO2`	1	Bio[b]	1	1		3
	Minimize NO_3^- consumption	`Im_NO3`	1	Bio[b]	1	1		4
				`Im_NH4`	0	0	only NO_3^-	
	Minimize PO_4^{3-} consumption	`Im_PO4`	1	Bio[b]	1	1		5
	Minimize SO_4^{2-} consumption	`Im_SO4`	1	Bio[b]	1	1		6
				`Im_H2S`	0	0	only SO_4^{2-}	
	Minimize ATP consumption	`Im_hnu`	1	Bio[b]	1	1	optimal value x	7
		NGAM[c]	-1	`Im_hnu`	x	x		8
				Bio[b]	0	0		
		NGAM[c]	-1	`Im_hnu`	x	x		9
				Bio[b]	1	1		
AA costs	Minimize ATP consumption	`Im_hnu`	1	AA[d]	1	1	optimal value x	10
		NGAM[c]	-1	`Im_hnu`	x	x		11
				AA[d]	0	0		
		NGAM[c]	-1	`Im_hnu`	x	x		12
				AA[d]	1	1		
RuBisCO costs	Minimize photon consumption	`Im_hnu`	1	`Bio_AA`[e]	1	1		13
	Minimize CO_2 consumption	`Im_CO2`	1	`Bio_AA`[e]	1	1		14
	Minimize NO_3^- consumption	`Im_NO3`	1	`Bio_AA`[e]	1	1		15
				`Im_NH4`	0	0	only NO_3^-	
	Minimize PO_4^{3-} consumption	`Im_PO4`	1	`Bio_AA`[e]	1	1		16
	Minimize SO_4^{2-} consumption	`Im_SO4`	1	`Bio_AA`[e]	1	1		17
				`Im_H2S`	0	0	only SO_4^{2-}	
	Minimize ATP consumption	`Im_hnu`	1	`Bio_AA`[e]	1	1	optimal value x	18
		NGAM[c]	-1	`Im_hnu`	x	x		19
				`Bio_AA`[e]	0	0		
		NGAM[c]	-1	`Im_hnu`	x	x		20
				`Bio_AA`[e]	1	1		
	Minimize ATP consumption	`Im_hnu`	1	`Ex_Suc`	1	1	optimal value x	21
		NGAM[c]	-1	`Im_hnu`	x	x		22
				`Ex_Suc`	0	0		
		NGAM[c]	-1	`Im_hnu`	x	x		23
				`Ex_Suc`	1	1		

[a] AraBio: `Bio_opt` with stoichiometric biomass coefficient of the model of de Oliveira Dal'Molin *et al.*
[b] Bio: an optimization for each biomass function (`Bio_opt`, `Bio_CLim`, `Bio_NLim`)
[c] NGAM: for each maintenance function (`NGAM_c`, `NGAM_h`, `NGAM_m`) objective coefficient -1 is assigned
[d] AA: an optimization for each of the 20 amino acids
[e] `Bio_AA`: an optimization for each of the 3003 amino acid compositions as stoichiometric coefficients

et al., 2010; Mintz-Oron *et al.*, 2012). This is confirmed by a flux variability analysis (FVA; Mahadevan and Schilling, 2003) of the photoautotrophic setting (Table S9) maximizing the biomass function of the model of de Oliveira Dal'Molin *et al.* (2010) (de Oliveira Dal'Molin *et al.*, 2010). By means of FBA, the actual flux range of each reaction is determined and is used for predicting the flux variability of the reactions. A reaction is denoted as variable if the actual flux range is more than 1 % of the total flux range meaning maximal minus minimal reaction boundary (Equation S12):

$$\text{variability frequency} = \frac{\text{Actual flux range}}{\text{Total flux range}} = \frac{\text{FVA}_{\max} - \text{FVA}_{\min}}{v_{\max} - v_{\min}} > 1\%. \tag{S12}$$

Importantly, the maximal and minimal reaction boundary are arbitrarily set to 1000 and -1000, respectively. The variability threshold was chosen to be 1 % as the frequency of variable reactions for all three models can be meaningfully compared within this range (Figure S1).

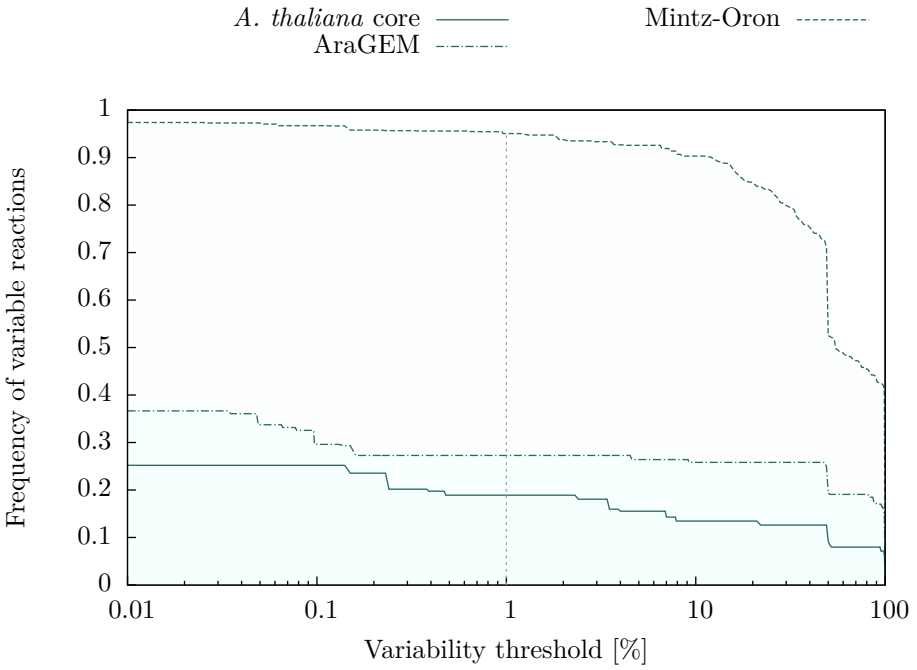

Figure S1: Fraction of variable reactions depending on the variability threshold. The gray line denotes the chosen threshold.

S3.3 Flux coupling analysis

Another approach to make statements regarding the flexibility of metabolic networks is flux coupling analysis (FCA; Marashi and Bockmayr, 2011). It describes the functional relation between two reactions, more precisely of their fluxes. By means of FBA, the effects of changing the flux of one coupling partner on the second reaction can be investigated (Larhlimi *et al.*, 2012). Thereby, one distinguishes three types of couplings, namely fully-, partially- and directionally-coupled reactions. Two reactions are fully-coupled, c_{ful}, if changing the reaction flux of one coupling partner results in a fixed proportional change of the flux of the other reaction. On the other hand, for partially-coupled reactions, c_{par}, such a change only causes an arbitrarily change of the corresponding coupling partner. Importantly, those effects are mutual meaning the change of each of both reactions induce the respective change of the second reaction flux. On the contrary, directionally-coupled reactions, c_{dir}, comprise reactions whose changes only affect a second reaction flux but may not be affected *vice versa*. All reactions which do not fall into these categories are named uncoupled.

As the computation of flux couplings only relies on the stoichiometry of the metabolic network, constraints such as flux boundaries are ignored. The compared models dramatically differ in the number of import and export reactions which highly influence the coupling of reactions. Consequently, we removed all importers and exporters except for those listed in Table S9 to establish a comparable framework. Moreover, we removed all blocked reactions, including also those arising from the removal the importers and exporters (in accordance to the algorithm's procedure; Larhlimi *et al.*, 2012). Accordingly, we arrive at functional networks for which we identify the number of couplings (Table 1). To determine the frequency of coupled reactions, we divided the sum of all coupled pairs of reactions, by the total number of reaction pairs (Equation S13)

$$\text{coupling frequency} = \frac{\text{number of coupled pairs}}{\text{total number of pairs}} = \frac{\#c_{\text{ful}} + \#c_{\text{par}} + \#c_{\text{dir}}}{\frac{n(n-1)}{2}}, \tag{S13}$$

where n denotes the number of reactions of the functional network.

S3.4 Conversion of RuBisCO costs

For the cost estimation of RuBisCO all possible amino acid sequences were considered. As the eight small subunits are encoded by four genes with overall seven splice variants the total number of combinations is $3003 = \left(\!\binom{7}{8}\!\right) = \binom{7+8-1}{8} = \binom{7+8-1}{7-1} = \frac{(7+8-1)!}{8!(7-1)!}$.

To get a better feel for the magnitude of RuBisCO costs, we provided several alternative cost measures. By means of the standard Gibbs free energy of ATP synthesis (Turina *et al.*, 2003) we converted the amount of ATP into kilojoule (kJ; Equations S14 and S15)

$$x\,\text{nmol ATP} \cdot 36\,\tfrac{\text{kJ}}{\text{mol ATP}}, \tag{S14}$$

$$x\,\text{molecules of ATP} \cdot 36\,\tfrac{\text{kJ}}{\text{mol ATP}} \div 6.02214129e23\,\tfrac{1}{\text{mol}}. \tag{S15}$$

The costs in terms of sucrose were determined via FBA by computing the ATP costs for producing one molecule of sucrose (Table S10, no. 21–23). As the optimization is a linear problem, the number of molecules sucrose can be achieved easily by cross multiplication with the 97 molecules of ATP per molecule sucrose.

Similarly, we determined the amount of biomass which can be produced for the energy 1 nmol *de novo* synthesized RuBisCO requires. Within the scope of cell performance, we also determined the amount of ATP required to synthesize one unit of biomass. Here, we relied on the biomass composition for optimal growth conditions. Together with the fact that one unit biomass represent 1 g FW, the amount of alternatively synthesized biomass was calculated by cross multiplication with the RuBisCO costs in terms of ATP.

Finally, we estimated the total number of reactions of a complex RuBisCO and, accordingly, the ATP costs per reaction. Therefore, we calculate the turnover number for RuBisCO by considering that each fourth reaction proceeds with O_2 instead of CO_2 (Peterhansel and Maurino, 2011). Thereby we obtain a turnover number, $k_{\text{cat}}^{\text{RBC}}$, of 2.66 to 2.88 s^{-1} (Equation S17). Assuming a half life of RuBisCO of about seven days (Piques *et al.*, 2009), we estimated a catalysis capability of 1,605,744 to 1,741,824 reactions for one complex RuBisCO (Equation S19). Eventually, by dividing the ATP costs per RuBisCO by the total number of reactions one complex of RuBisCO is able to catalyze, we achieved the estimated costs of one RuBisCO reaction (Equation S21)

$$k\text{cat}^{\text{RBC}} = \tfrac{3}{4} \cdot k\text{cat}^{\text{RBC - CO}_2} + \tfrac{1}{4} \cdot k\text{cat}^{\text{RBC - O}_2}, \tag{S16}$$

$$= \tfrac{3}{4} \cdot 3.4 \text{ to } 3.7 + \tfrac{1}{4} \cdot 0.42 = 2.655 \text{ to } 2.88\ \left[\tfrac{1}{s}\right], \tag{S17}$$

$$r_{\text{tot}} = t_{1/2} \cdot k\text{cat}^{\text{RBC}}, \tag{S18}$$

$$= 7\ [d] \cdot 2.655 \text{ to } 2.88\ \left[\tfrac{1}{s}\right] = 1,605,744 \text{ to } 1,741,824, \tag{S19}$$

$$C_{\text{RBC}}^{\text{rxn}} = \frac{C_{\text{RBC}}}{r_{\text{tot}}}, \tag{S20}$$

$$= \frac{243,287.9 \text{ to } 269,133.9}{1,605,744 \text{ to } 1,741,824} = 0.140 \text{ to } 0.168. \tag{S21}$$

S3.5 Comparison of amino acid cost measures

The complete comparison of the amino acid cost measures are provided in Table S11.

S3.6 Extended RuBisCO costs

The minimal and maximal costs and precursor requirements for all 3003 possible amino acid compositions of RuBisCO are provided in Table S12. Interestingly, except for SO_4^{2-} the amino acid sequence denoting the minimal and maximal consumption of precursors and ATP is the same, respectively. The minimal costs correspond to the combination $8 * (\texttt{ATCG00490.1})$ `AND` $8 * (\texttt{AT1G67090.2})$, and the maximal costs to

Table S11: Comparison of amino acid cost measures and the corresponding rankings of the Arabidopsis core model (Our costs; for optimal growth conditions), Sajitz-Hermstein and Nikoloski (2010, SN), Craig and Weber (1998, CW), Akashi and Gojobori (2002, AG), Wagner (2005, Wag), Seligmann (2003, Sel), and Barton *et al.* (2010, Bar). The 16th column denotes their averaged rank. The columns 17 to 25 capture the amino acid composition with respect to the size of the carbon-skeleton, and the number of incorporated amine groups and sulfur as well as the essential pathways for the amino acid precursors. The columns labeled with CBC, Glyc, PRPP, Pyr, Shik and TCA represent the pathways Calvin-Benson cycle, glycolysis, phosphoribosyl pyrophosphate synthesis, pyruvate decarboxylation, Shikimate and TCA cycle, respectively.

Amino	Our costs		SN		CW		AG		Wag		Sel		Bar		Average	Carbon	Amine	Sulfur	Involved biochemical pathways					
acid	[ATP]	rank	[ATP]	rank	[ATP]	rank	[ATP]	rank	[ATP]	rank	$\left[\frac{g}{mol}\right]$	rank	$\boxed{\frac{mmol}{g\ DW}}$	rank	rank	skeleton	groups	presents	CBC	Glyc	PRPP	Pyr	Shik	TCA
Gly	25.16	2	13.33	1	14.5	8	11.7	1	14.5	3	75.07	1	0.31	1	1	2	1	0	✓	✓				✓
Ala	22.48	1	19.33	2	12.5	6	11.7	1	14.5	3	89.09	2	0.5	3	2	3	1	0		✓				
Ser	26.83	5	19.33	2	15	9	11.7	1	14.5	3	105.09	3	0.49	2	3	3	1	0	✓	✓				
Asp	25.16	2	25.33	5	1	1	12.7	4	15.5	7	133.10	11	0.61	4	4	4	1	0						✓
Asn	26.34	4	26.66	7	4	2	14.7	5	18.5	8	132.12	10	0.79	7	5	4	2	0						✓
Thr	31.93	7	25.33	5	6	3	18.7	8	21.5	10	119.12	6	0.69	5	6	4	1	0						✓
Glu	46.01	8	31.33	8	8.5	4	15.3	6	9.5	1	147.13	14	0.86	8	7	5	1	0						✓
Pro	47.90	11	31.33	8	12.5	6	20.3	9	14.5	3	115.13	4	0.99	11	8	5	1	0				✓		✓
Gln	46.32	9	32.66	12	9.5	5	16.3	7	10.5	2	146.15	12	0.92	9	9	5	2	0						✓
Cys	30.13	6	19.33	2	24.5	14	24.7	11	26.5	11	121.16	7	0.75	6	10	3	1	1	✓	✓		✓		
Val	46.34	10	31.33	8	25	15	23.3	10	29	12	117.15	5	0.96	10	11	5	1	0		✓				
Met	55.75	13	31.33	8	18.5	10	34.3	16	36.5	15	149.21	15	1.25	14	12	5	1	1	✓	✓		✓		✓
Ile	56.50	14	37.33	13	20	13	32.3	15	38	17	131.17	8	1.21	12	13	6	1	0						✓
Leu	65.51	15	37.33	13	33	16	27.3	12	37	16	131.17	8	1.21	12	13	6	1	0		✓		✓		
Lys	53.00	12	38.66	15	18.5	10	30.3	14	36	14	146.19	13	1.31	15	15	6	2	0		✓				✓
Arg	65.67	16	41.32	17	18.5	10	27.3	12	20.5	9	174.20	18	1.39	16	16	6	4	0				✓		✓
His	69.46	17	39.99	16	33	16	38.3	17	29	12	155.16	16	1.46	17	17	6	3	0			✓			
Tyr	86.74	18	55.33	18	56.5	18	50	18	59	18	181.19	19	1.77	18	18	9	1	0					✓	
Phe	87.57	19	55.33	18	63	19	52	19	61	19	165.19	17	1.84	19	19	9	1	0					✓	
Trp	117.36	20	68.66	20	78.5	20	74.3	20	75.5	20	204.23	20	2.39	20	20	11	2	0	✓	✓	✓		✓	

$8 * (\texttt{ATCG00490.1})$ AND $8 * (\texttt{AT5G38410.3})$.

Table S12: Minimal and maximal requirements of precursors and energy to synthesize one complex RuBisCO. The number of required Pi are the same for all 3003 possible combinations to assemble RuBisCO.

Encoding genes		Costs [ATP]		Minimal precursor consumption											
				Photon		CO_2		H_2O		PO_4^{3-}		NO_3^-		SO_4^{2-}	
		min	max	min	max	min	max	min	max	min	max	min	max	min	max
RbcL	ATCG00490.1	8	8	8	8	8	8	8	8	8	8	8	8	8	8
RbcS1A	AT1G67090.1	–	–	–	–	–	–	–	–			–	–	–	–
	AT1G67090.2	8	–	8	–	8	–	8	–			8	–	–	8
RbcS1B	AT5G38430.1	–	–	–	–	–	–	–	–			–	–		–
RbcS2B	AT5G38420.1	–	–	–	–	–	–	–	–	8	8	–	–		–
RbcS3B	AT5G38410.1	–	–	–	–	–	–	–	–			–	–	8	–
	AT5G38410.2	–	–	–	–	–	–	–	–			–	–		–
	AT5G38410.3	–	8	–	8	–	8	–	8			–	8		–
Minimal costs		243,287.9		231,066.4		24,024		13,536		0		6,520		216	
Maximal costs		269,133.9		256,262.3		26,520		29,128		0		7,112		240	

References

Akashi, H, Gojobori, T (2002) Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. Proc Natl Acad Sci U S A **99**: 3695–3700

Allen, J (2002) Photosynthesis of ATP-electrons, proton pumps, rotors, and poise. Cell **110**: 273–276

Barton, MD, Delneri, D, Oliver, SG, Rattray, M, Bergman, CM (2010) Evolutionary systems biology of amino acid biosynthetic cost in yeast. PLoS One **5**: e11935

Buchner, P, Takahashi, H, Hawkesford, MJ (2004) Plant sulphate transporters: co-ordination of uptake, intracellular and long-distance transport. J Exp Bot **55**: 1765–1773

Craig, CL, Weber, RS (1998) Selection costs of amino acid substitutions in ColE1 and ColIa gene clusters harbored by *Escherichia coli*. Mol Biol Evol **15**: 774–776

DeBolt, S, Scheible, WR, Schrick, K, Auer, M, Beisson, F, Bischoff, V, Bouvier-Navé, P, Carroll, A, Hematy, K, Li, Y, Milne, J, Nair, M, Schaller, H, Zemla, M, Somerville, C (2009) Mutations in UDP-Glucose:sterol glucosyltransferase in Arabidopsis cause transparent testa phenotype and suberization defect in seeds. Plant Physiol **151**: 78–87

Doermann, P, Hoffmann-Benning, S, Balbo, I, Benning, C (1995) Isolation and characterization of an Arabidopsis mutant deficient in the thylakoid lipid digalactosyl diacylglycerol. Plant Cell **7**: 1801–1810

Jackson, RJ, Hellen, CUT, Pestova, TV (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol **11**: 113–127

Kanehisa, M, Goto, S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res **28**: 27–30

Larhlimi, A, David, L, Selbig, J, Bockmayr, A (2012) F2C2: a fast tool for the computation of flux coupling in genome-scale metabolic networks. BMC Bioinformatics **13**: 57

Mahadevan, R, Schilling, CH (2003) The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab Eng **5**: 264–276

Marashi, SA, Bockmayr, A (2011) Flux coupling analysis of metabolic networks is sensitive to missing reactions. BioSystems **103**: 57–66

Meharg, AA, Blatt, MR (1995) NO3- transport across the plasma membrane of *Arabidopsis thaliana* root hairs: kinetic control by pH and membrane voltage. J Membr Biol **145**: 49–66

Mintz-Oron, S, Meir, S, Malitsky, S, Ruppin, E, Aharoni, A, Shlomi, T (2012) Reconstruction of Arabidopsis metabolic network models accounting for subcellular compartmentalization and tissue-specificity. Proc Natl Acad Sci U S A **109**: 339–344

Mooney, BP, Miernyk, JA, Michael Greenlief, C, Thelen, JJ (2006) Using quantitative proteomics of Arabidopsis roots and leaves to predict metabolic activity. Physiologia Plantarum **128**: 237–250

Mueller, LA, Zhang, P, Rhee, SY (2003) AraCyc: a biochemical pathway database for Arabidopsis. Plant Physiol **132**: 453–460

Noguchi, K, Go, CS, Miyazawa, SI, Terashima, I, Ueda, S, Yoshinari, T (2001) Costs of protein turnover and carbohydrate export in leaves of sun and shade species. Functional Plant Biol. **28**: 37–47

de Oliveira Dal'Molin, CG, Quek, LE, Palfreyman, RW, Brumbley, SM, Nielsen, LK (2010) AraGEM, a genome-scale reconstruction of the primary metabolic network in Arabidopsis. Plant Physiol **152**: 579–589

Orth, JD, Thiele, I, Palsson, BO (2010) What is flux balance analysis? Nat Biotechnol **28**: 245–248

Pate, JS, Layzell, DB (1990) Energetics and biological costs of nitrogen assimilation. In Conn, PM, Miflin, BJ, Stumpf, W, Lea, PJ , eds, The Biochemistry of Plants: Intermediary Nitrogen Metabolism, 1 Ed Vol 16. Academic Press, San Diego, pp 1–42

Peterhansel, C, Maurino, VG (2011) Photorespiration redesigned. Plant Physiol **155**: 49–55

Piques, M, Schulze, WX, Hoehne, M, Usadel, B, Gibon, Y, Rohwer, J, Stitt, M (2009) Ribosome and transcript copy numbers, polysome occupancy and enzyme dynamics in Arabidopsis. Molecular Systems Biology **5**: 314

Pyl, ET, Piques, M, Ivakov, A, Schulze, W, Ishihara, H, Stitt, M, Sulpice, R (2012) Metabolism and growth in Arabidopsis depend on the daytime temperature but are temperature-compensated against cool nights. Plant Cell **24**: 2443–2469

Raghothama, KG (1999) Phosphate acquisition. Annu Rev Plant Physiol Plant Mol Biol **50**: 665–693

Sajitz-Hermstein, M, Nikoloski, Z (2010) A novel approach for determining environment-specific protein costs: the case of *Arabidopsis thaliana*. Bioinformatics **26**: i582–i588

Sakano, K (1990) Proton/Phosphate Stoichiometry in Uptake of Inorganic Phosphate by Cultured Cells of *Catharanthus roseus* (L.) G. Don. Plant Physiol **93**: 479–483

Schomburg, I, Chang, A, Placzek, S, Söhngen, C, Rother, M, Lang, M, Munaretto, C, Ulas, S, Stelzer, M, Grote, A, Scheer, M, Schomburg, D (2013) BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. Nucleic Acids Res **41**: D764–D772

Seelert, H, Poetsch, A, Dencher, NA, Engel, A, Stahlberg, H, Müller, DJ (2000) Structural biology. Proton-powered turbine of a plant motor. Nature **405**: 418–419

Seligmann, H (2003) Cost-minimization of amino acid usage. J Mol Evol **56**: 151–161

Sharrock, RA, Clack, T (2002) Patterns of expression and normalized levels of the five Arabidopsis phytochromes. Plant Physiol **130**: 442–456

Smith, AM, Zeeman, SC (2006) Quantification of starch in plant tissues. Nat Protoc **1**: 1342–1345

Sulpice, R, Nikoloski, Z, Tschoep, H, Antonio, C, Kleessen, S, Larhlimi, A, Selbig, J, Ishihara, H, Gibon, Y, Fernie, AR, Stitt, M (2013) Impact of the carbon and nitrogen supply on relationships and connectivity between metabolism and biomass in a broad panel of Arabidopsis accessions. Plant Physiol **162**: 347–363

The UniProt Consortium (2012) Update on activities at the Universal Protein Resource (UniProt) in 2013. Nucleic Acids Research **41**: D43–D47

Tschoep, H, Gibon, Y, Carillo, P, Armengaud, P, Szecowka, M, Nunes-Nesi, A, Fernie, AR, Koehl, K, Stitt, M (2009) Adjustment of growth and central metabolism to a mild but sustained nitrogen-limitation in Arabidopsis. Plant Cell Environ **32**: 300–318

Turina, P, Samoray, D, Graeber, P (2003) H+/ATP ratio of proton transport-coupled ATP synthesis and hydrolysis catalysed by CF0F1-liposomes. EMBO J **22**: 418–426

de Visser, R, Spitters, C, Bouma, T (1992) Energy costs of protein turnover: Theoretical calculation and experimental estimation from regression of respiration on protein concentration of fullgrown leaves. In Lambers H, VdPL , ed, Molecular, Biochemical & Physiological Aspects of Plant Respiration. SPB Academic Publishing, Den Haag, pp 493–508

Wagner, A (2005) Energy constraints on the evolution of gene expression. Mol Biol Evol **22**: 1365–1374

Zerihun, A, McKenzie, BA, Morton, JD (1998) Photosynthate costs associated with the utilization of different nitrogen-forms: influence on the carbon balance of plants and shoot-root biomass partitioning. New Phytologist **138**: 1–11