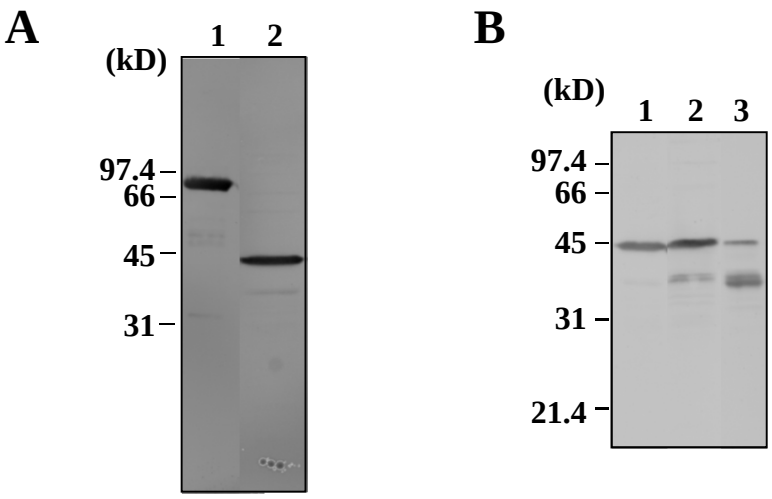
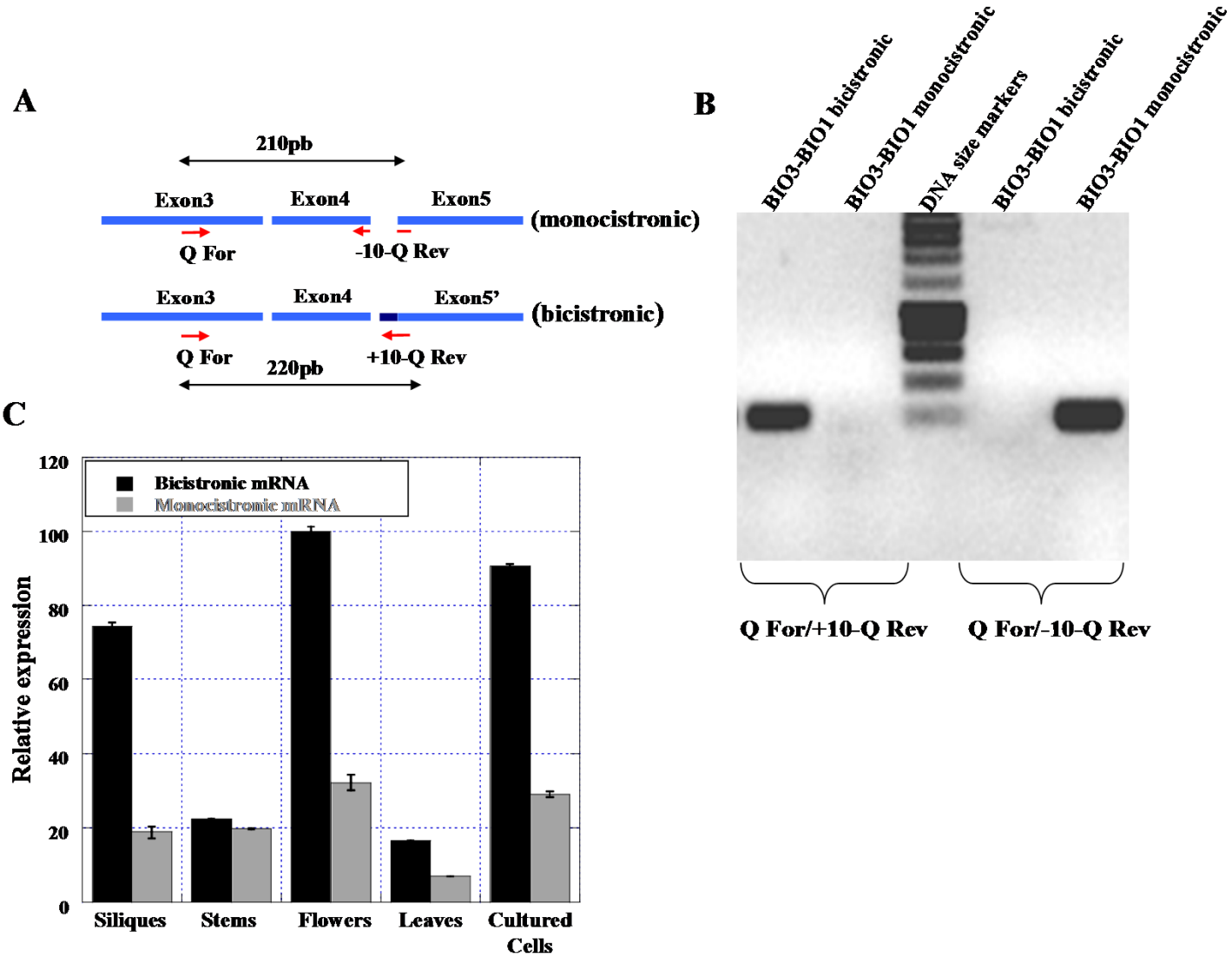


Supplemental Figure 1. Analysis of BIO3-BIO1 and BIO3 production in *E. coli* using polyclonal antibodies raised against recombinant *Arabidopsis* BIO3-BIO1 protein.

(A) Total proteins (10 µg per lane) from *E. coli* Rosetta 2 cells harbouring the pET-BIO3-BIO1 construct (Lane 1) or the pET-BIO3 (long) construct (Lane 2) grown in the presence of IPTG were analyzed by Immunoblot with affinity purified polyclonal antibodies raised against recombinant BIO3-BIO1. Position of molecular mass markers is given on the left. The pET-BIO3-BIO1 construct was obtained by subcloning the monocistronic *Arabidopsis BIO3-BIO1* cDNA in pET28b vector, as described in the Methods section. The pET-BIO3 (long) construct was obtained by PCR amplification of the BIO3 (long) ORF and cloning in pET28b vector, using the *Arabidopsis* bicistronic *BIO3-BIO1* cDNA as a template. Antibodies could not be tested against BIO1 (short) protein since attempts in overproduction of this protein failed. The pET-BIO1 (short) construct was obtained by amplification of the *BIO1* (short) ORF and cloning in pET28b vector, using the bicistronic *BIO3-BIO1* cDNA as a template.

(B) Documentation of *Arabidopsis* BIO3 (long) purification on nickel-nitrolotriacetic acid-agarose resin. Lane 1 and 2, total (2 µg) and soluble (5 µg) proteins, respectively, from *E. coli* cells producing BIO3 (long), and Lane 3, proteins eluted from the column (0.5 µg) were analysed by immunoblot with affinity-purified polyclonal antibodies raised against recombinant BIO3-BIO1. Position of molecular mass markers is given on the left.



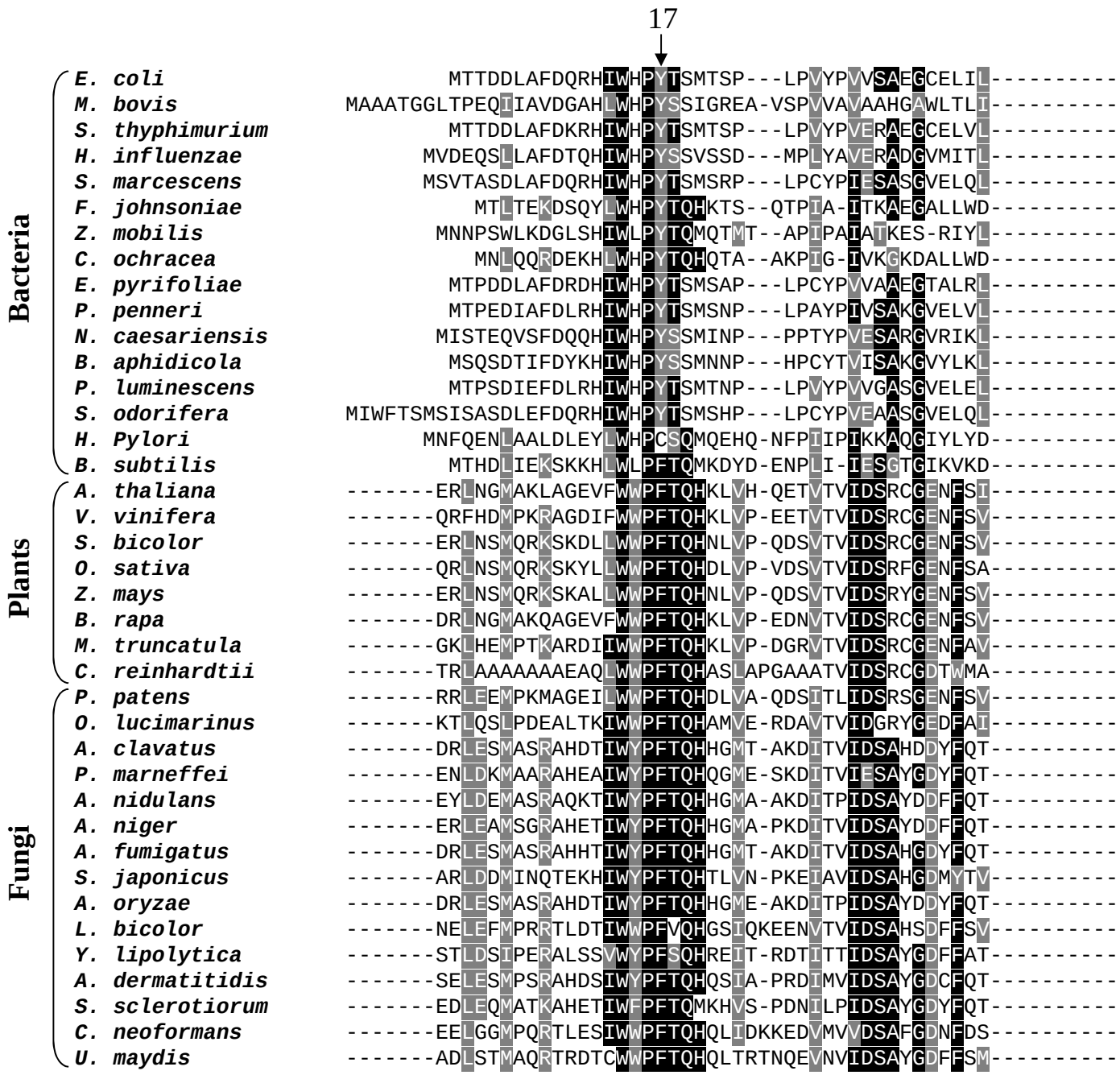
Supplemental Figure 2. (see legend on the next page)

Supplemental Figure 2. Expression analysis of *Arabidopsis* bicistronic and monocistronic *BIO3-BIO1* mRNAs by real-time PCR.

(A) Schematic representation of BIO3-BIO1-Q For, -10-Q Rev, and +10-Q Rev oligonucleotides positions on *BIO3-BIO1* cDNA variants. -10-Q Rev oligonucleotide enabled the quantification of monocistronic *BIO3-BIO1* mRNA. +10-Q Rev enabled the quantification of bicistronic *BIO3-BIO1* mRNA.

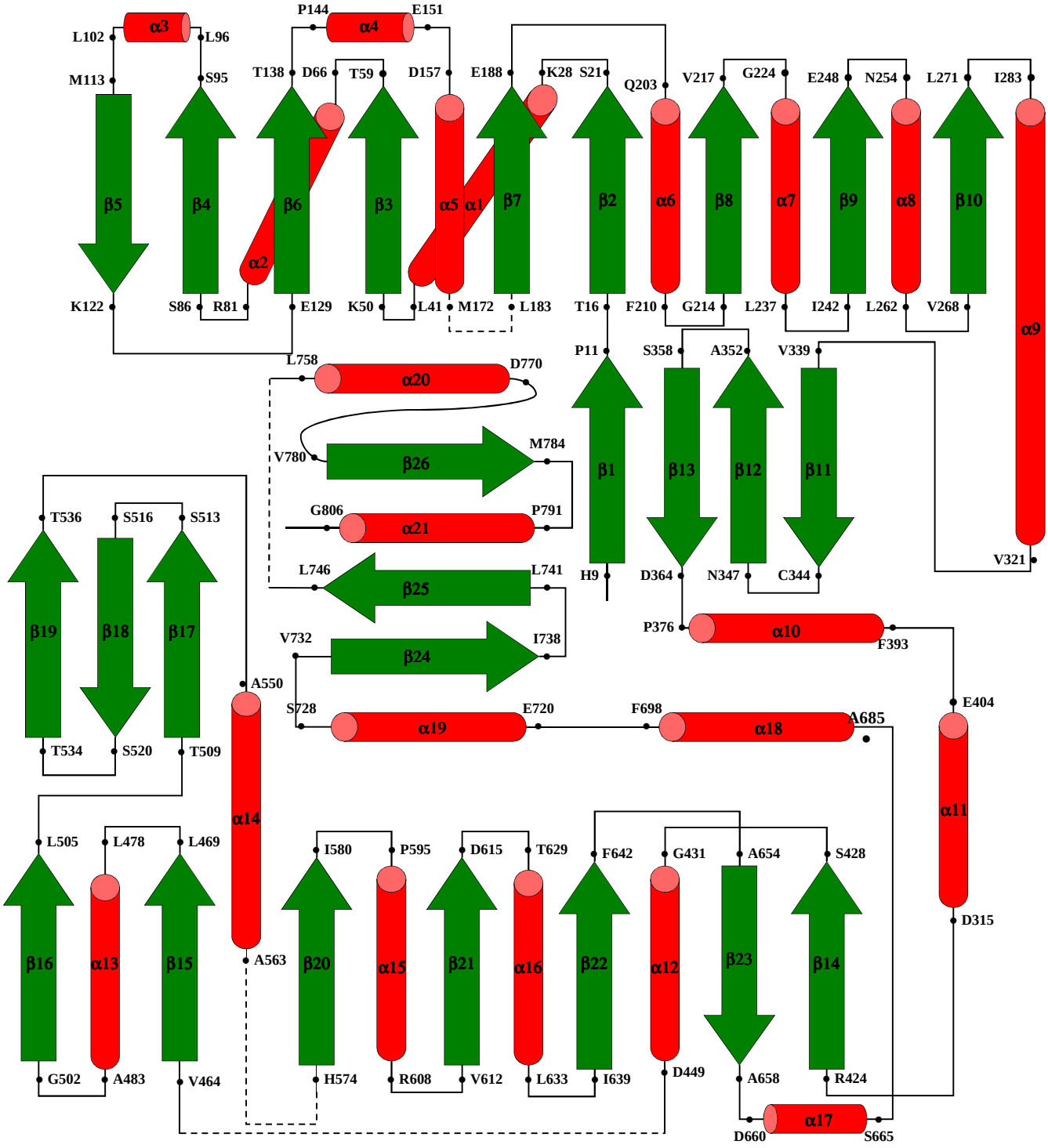
(B) Specificity of oligonucleotides used for quantification of *BIO3-BIO1* mRNA splice variants. PCR amplification of bicistronic and monocistronic cDNA fragments using Q-For/+10-Q Rev or Q-For/-10-Q Rev oligonucleotide pairs, and bicistronic or monocistronic *BIO3-BIO1* cDNAs as templates. Q-For and +10-Q Rev amplify a 220-bp product from the bicistronic cDNA but not from the monocistronic cDNA. Q-For and -10-Q Rev amplify a 210-pb product from the monocistronic cDNA but not from the bicistronic cDNA.

(C) Relative abundance of bicistronic and monocistronic *BIO3-BIO1* mRNA species in *Arabidopsis* organs. Relative quantification experiments were done by real-time RT-PCR using Rotor Gene System (Corbett research) and SYBR Green Jump Start Taq Readymix (Sigma-Aldrich). Experiments were performed on total RNA from various *Arabidopsis* organs (35-days old plants) and from *Arabidopsis* cultured cells, using splicing type-specific oligonucleotides. For each measurement 1µL of cDNA preparation was used as a template in 7.5µL Readymix with appropriate primers (used at a final concentration of 0.5 or 1µM). Amplification and detection were performed using the following profile: 95°C/2 min followed by 40 cycles of 95°C/15 s, 62°C/5 s, and 72°C/10 s. The specificity of the reaction was verified by melting curve analysis obtained by increasing temperature from 72°C to 95°C. Total RNA were prepared using RNeasy Mini kit from Qiagen followed by a treatment with RNase-free DNase I, quantified using a NanoDrop 2000 spectrophotometer (ThermoScientific), and controlled by gel electrophoresis. First strand cDNA was synthesized from 1µg of DNA-free RNA in a final volume of 20 µL using Oligo(dT)20 primers (Thermoscript RT-PCR System, Life Technologies) and used for real-time PCR analyses as described above. Control reactions omitting reverse transcriptase were run for all samples to ensure that genomic DNA contamination did not contribute to the amplified products. The efficiency of qPCR reactions (based on the slope of standard curves) ranged from 97% to 100%. Expression data were normalized to the constitutively expressed *ACTIN7* mRNA (At5g09810), which was used as internal standard of RNA integrity and cDNA preparation. Data are means of three biological replicates performed with four cDNA dilutions ±SD.

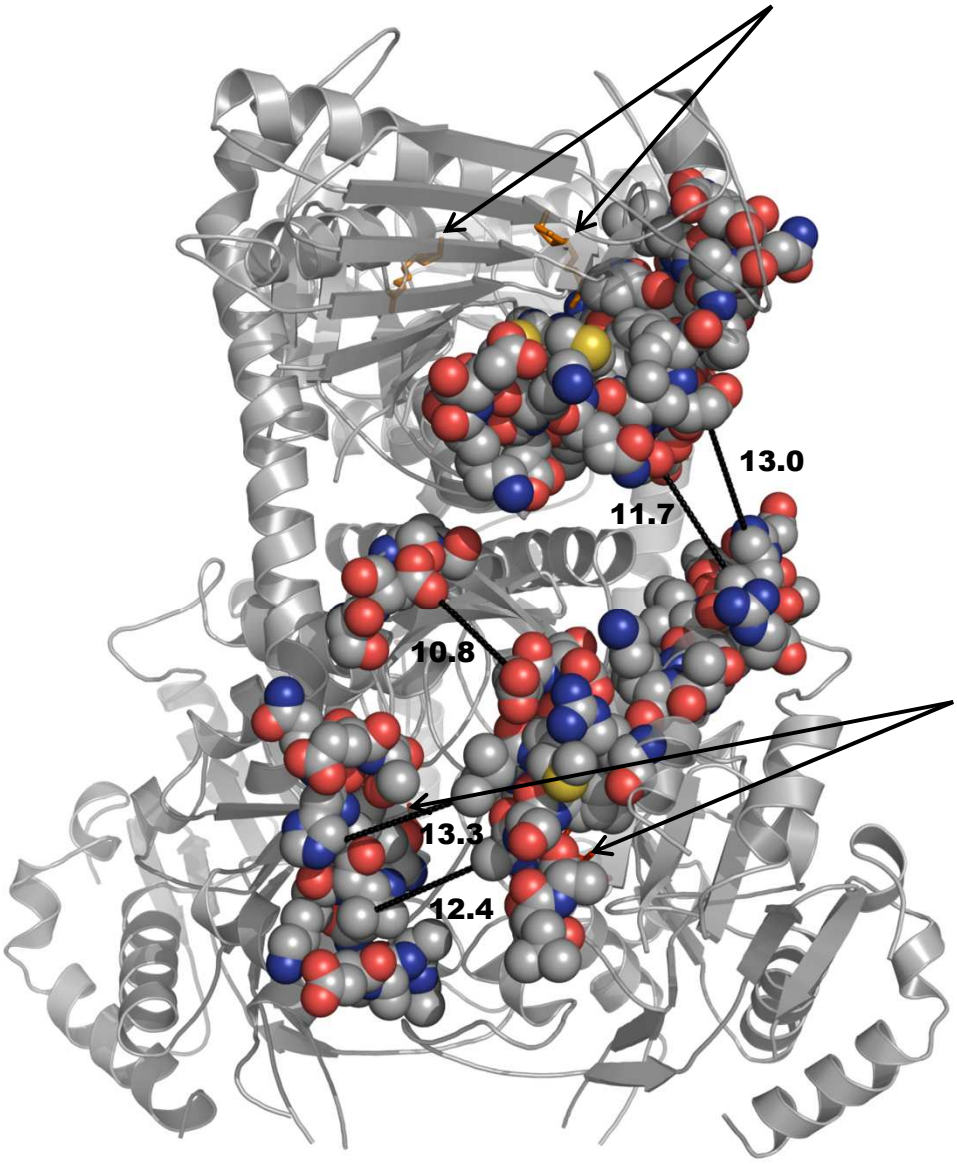


Supplemental Figure 3. (see legend on the next page)

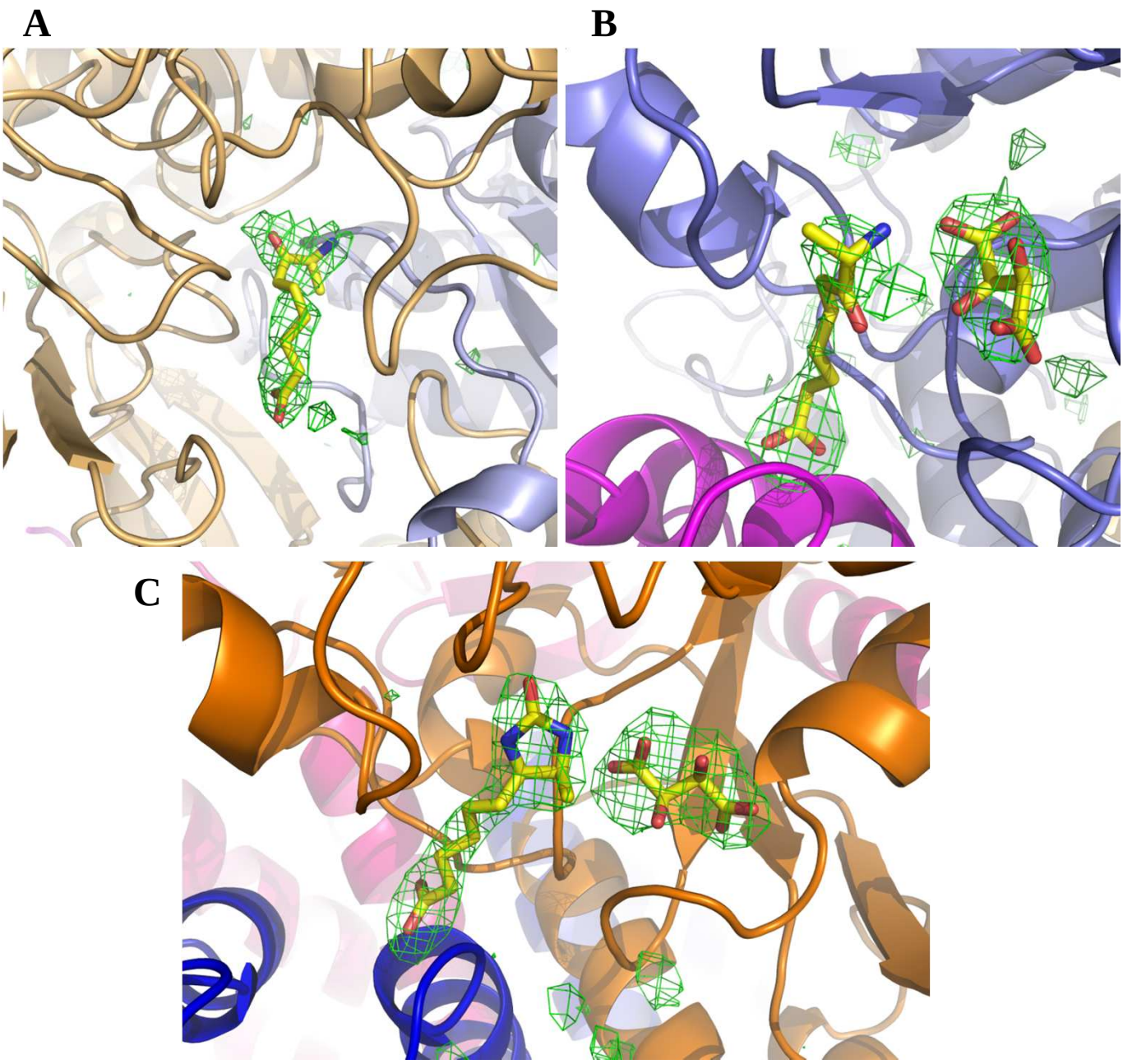
Supplemental Figure 3. Part of an amino acid alignment of DAPA aminotransferases from a selection of bacteria, plants and fungi, highlighting residu Tyr17 (*E. coli* numbering). Many bacterial sequences (monofunctional enzymes) have a tyrosine moiety at this position. All available plant and fungi sequences (bifunctional enzymes) contain a phenylalanine residue at this position (Phe326 for *Arabidopsis* mature protein). The following sequences from public databases were aligned with the program ClustalW (http://www.ch.embnet.org/software/ClustalW-XXL.html.): *Escherichia coli* (P12995.2), *Mycobacterium bovis* (P0A4X7.1), *Salmonella thyphimurium* (P12677.2), *Haemophilus influenzae* (P44426.1), *Serratia marcescens* (P36568.1), *Flavobacterium johnsoniae* (ABQ03844.1), *Zymomonas mobilis* (AAV90542.1) *Capnocytophaga ochracea* (ACU93703.1), *Erwinia pyrifoliae* (CAY74978.1), *Proteus penneri* (ZP_03806407.1), *Neptuniibacter caesariensis* (ZP_01167088.1), *Buchnera aphidicola* (*Acyrthosiphon pisum*) (P57379.1), *Photorhabdus luminescens* (CAE13777.1) *Serratia odorifera* (ZP_06189072.1), *Helicobacter pylori* (Q9ZKM5.1), *Bacillus subtilis* (P53555.1), *Arabidopsis thaliana* (EU089963.1), *Vitis vinifera* (XM_002270515.1), *Sorghum bicolor* (XM_002468273.1), *Oryza sativa* (NM_001067889.1), *Zea mays* (BT065649.1), *Brassica rapa* (AC189479.2), *Medicago truncatula* (AC174353.16), *Chlamydomonas reinhardtii* (XM_001690622.1), *Physcomitrella patens* (XM_001764409.1), *Ostreococcus lucimarinus* (XM_001422822.1), *Aspergillus clavatus* (XM_001270182.1), *Penicillium marneffei* (XM_002143123.1), *Aspergillus nidulans* (XM_659156.1), *Aspergillus niger* (XM_001396701.1), *Aspergillus fumigatus* (XM_742618.1), *Schizosaccharomyces japonicus* (XM_002171908.1), *Aspergillus oryzae* (XM_001816971.1), *Laccaria bicolor* (XM_001880692.1), *Yarrowia lipolytica* (XM_504233.2), *Ajellomyces dermatitidis* (XM_002627316.1), *Sclerotinia sclerotiorum* (XM_001590649.1), *Cryptococcus neoformans* (XM_569073.1), *Ustilago maydis* (XM_753969.1). Accession numbers are given in parentheses. White letters on black background designate conserved residues. White letters on grey backgrounds designate similar residues.



Supplemental Figure 4. Topology of mBIO3-BIO1.
The β-strands and the α-helices are represented by arrows and cylinders, respectively. The non-observed loops in electron density are drawn in dashed lines.



Supplemental Figure 5. View of the crevice at the surface of mBIO3-BIO1. The dimer is shown in grey cartoon and the flanking residues of the crevice are shown in spheres. Distances between flanking residues are indicated. The crevice width and length is around 12 Å and 95 Å, respectively. DTB observed in the DTBS catalytic site is drawn in orange stick. PLP in the DAPA-AT catalytic site is displayed in red stick.



Supplemental Figure 6. Simulated annealing omit maps for the mBIO3-BIO1 ligands.

(A) Fo-Fc omit map contoured at 3 sigma superimposed onto mBIO3-BIO1/KAPA calculated by omitting KAPA and tartrate in BIO1.

(B) Fo-Fc omit map contoured at 3 sigma superimposed onto mBIO3-BIO1/KAPA calculated by omitting KAPA and tartrate in BIO3.

(C) Fo-Fc omit map contoured at 3 sigma superimposed onto mBIO3-BIO1/DTB/tartrate calculated by omitting DTB and tartrate.

Supplemental Table 1. Synthetic oligonucleotides used in this study

Primer name	Primer sequence
BIO3-BIO1 cDNAs cloning	
BIO3-BIO1/*Nde*I	5’-CACATCCCACCTACCACATATGATACCCGTAACCGC-3’
mBIO3-BIO1/*Nde*I	5’-ACGCCACCGCATTCATATGAAATCCACCTCTGTTTCTC-3’
BIO1 short/*Nde*I	5’-GTTCTCTTAGCCAGCAATCATATGCTTGTGCAAGC-3’
BIO3-BIO1/*Sac*I	5’-AGCTGGAGAGAGAGCTCTGGGTTTCATGTTCTATTG-3’
BIO3 Long/*Sac*I	5’-GTGTAACCCATTTCTCTAGCGAGCTCAGCCTGAAA-3’
Mutagenesis	
F326Y For	5’-GTTTTCTGGTGGCCGTATACTCAGCATAAACTTGTGCATC-3’
F326Y Rev	5’-GATGCACAAGTTTATGCTGAGTATACGGCCACCAGAAAAC-3’
S360Y For	5'-TACAAGGCTTCCGATAACAGTTCTCTTTATCAGCAATTTGATGCTTGTGC-3'
S360Y Rev	5'-GCACAAGCATCAAATTGCTGATAAAGAGAACTGTTATCGGAAGCCTTGTA-3'
I793W For	5'-GGCCCTTGCACCTCGCCGGAATGGTGCCGCCGGTT-3'
I793W Rev	5'-AACCGGCGGCACCATTCCGGCGAGGTGCAAGGGCC-3'
BIO3-BIO1-GFP fusion cloning	
BIO3-BIO1/*Sal*I	5’-ATCCCACCTAGTCGACATGATACCCGTAAC-3’
BIO3-BIO1/*BspH*I	5’-GAGAGAGTTTTGTCATGAGTGTTCTATTGAATTCTCC-3’
Real time RT-PCR	
BIO3-BIO1-Q For	5’-GCATGGCTAAGCTAGCAGGA-3’
+10-Q Rev	5’-CAAGCTCAGCCTGAAACAGC-3’
-10-Q Rev	5’-CAAGCTCAGCCTGGAAAGTA-3’

Supplemental Table 2. Data Collection

	Native	Peak	Inflection	High energy remote	mBIO3-BIO1/KAPA	mBIO3-BIO1/DTB
Wavelength (Å)	0.979701	0.979760	0.979906	0.978123	0.979701	0.93340
Resolution (Å)	48.14-2.50 (2.55-2.50)	40.33-2.72 (2.78-2.71)	40.35-2.70 (2.77-2.70)	40.37-2.71 (2.78-2.71)	44.57-2.81 (2.88-2.81)	40.44-2.68 (2.75-2.68)
Space group	P1	C2			C2	C2
Cell parameters	a= 79.44 b= 80.07 c= 136.94 Å α= 99.958 β=107.125 γ= 97.25°	a= 233.67 b= 75.97 c= 88.63 Å β= 109.20°			a= 235.32 b=76.94 c=89.22 Å β=109.90°	a= 246.67 b= 76.63 c= 79.84 β= 108.02°
Total reflections	186134 (5705)	272398 (15082)	180558 (7567)	205945 (9284)	138512 (2500)	168286 (12504)
Unique reflections	95045 (2966)	76488 (5138)	75322 (4388)	75526 (4483)	36581 (4352)	40026 (2956)
Completeness (%)	88.0 (48.3)	98.2 (89.5)	94.9 (74.8)	95.7 (77.0)	99.0 (91.5)	99.7 (99.8)
Rsym (%)	5.0 (42.3)	7.9 (48.1)	7.5 (51.4)	8.9 (65.7)	10.9 (65.3)	15.1 (76.3)
I/sI	14.81 (1.91)	12.88 (2.11)	10.26 (1.28)	9.98 (1.21)	9.22 (1.61)	9.95 (1.83)

$R_{sym}=\sum\sum|I_i-I_m|/\sum\sum I_i$, where I_i is the intensity of the measured reflection and I_m is the mean intensity of this reflection. Values indicated in parentheses correspond to the statistics in the highest resolution shell.

Supplemental Table 3. Refinement Statistics

	[SeMet]-mBIO3-BIO1	mBIO3-BIO1	mBIO3-BIO1-KAPA	mBIO3-BIO1-DTB
Resolution (Å)	41.86-2.71 (2.78-2.71)	39.57-2.50 (2.53-2.50)	44.57-2.81 (2.88-2.81)	40.44-2.68 (2.75-2.68)
Number of reflections used for Rcrys calculation	37564 (2423)	90241 (1638)	34535 (2472)	38006 (2713)
Number of reflections used for Rfree calculation	1989 (133)	4750 (87)	1822 (134)	2006 (143)
Data cutoff F/σ_F	0.0	0.0	0.0	0.0
R (%)	19.96 (24.3)	17.72 (26.05)	19.34 (28.90)	18.44 (24.54)
Rfree (%)	26.10 (34.9)	23.91 (31.50)	26.19 (36.55)	25.90 (33.24)
Number of nonhydrogen protein atoms	11219	22809	11389	11524
Number of sulfate ions	4	10	6	4
Number of magnesium ions	0	2	0	0
Number of water molecules	80	429	0	121
Overall B factors ($Å^2$)	58.0	47.9	66.7	34.9
Wilson B ($Å^2$)	48.0	40.3	52.7	34.7
Mean B factor ligands ($Å^2$)	-	-	KAPA: 60.3	DTB: 32.4
Mean occupancy ligands			KAPA: 0.96	DTB: 1.0
Ramachandran plot				
Residues in most favored regions (%)	89.4	86.8	85.7	90.1
Residues in disallowed regions (%)	0.3	0.3	0.3	0.3
RMS differences from ideal geometry				
Bond length (Å)	0.009	0.011	0.011	0.009
Bond angle (°)	1.236	1.325	1.354	1.252

Values indicated in parentheses correspond to the statistics in the highest resolution shell.

$R_{cryst}=\sum||F_{obs}|-|F_{calc}||/\sum|F_{obs}|$. R_{free} (Brunger, 1992) is the same as R_{cryst} but calculated for n% data omitted from the refinement where n is 5 % for all the structure, approximately.

Reference

Brunger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature **355,** 472-475.