Supplementary information

Biochemical Characterization of the Carotenoid 1,2-Hydratases (CrtC) from *Rubrivivax gelatinosus* and *Thiocapsa roseopersicina*

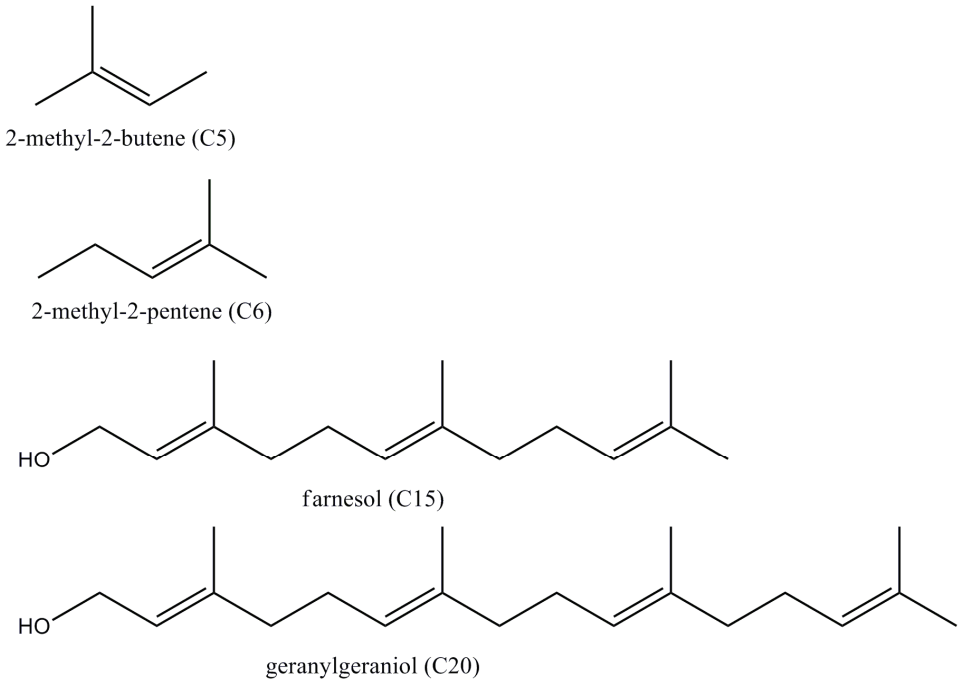
Aida Hiseni, Isabel W.C.E. Arends and Linda G. Otten

Biocatalysis and Organic Chemistry, Department of Biotechnology, Delft University of Technology, Julianalaan 136, Delft 2628 BL, The Netherlands. Tel: +31 15 2789969. Fax: +31 15 2781415. E-mail: L.G.Otten@tudelft.nl, URL: www.bt.tudelft.nl/boc

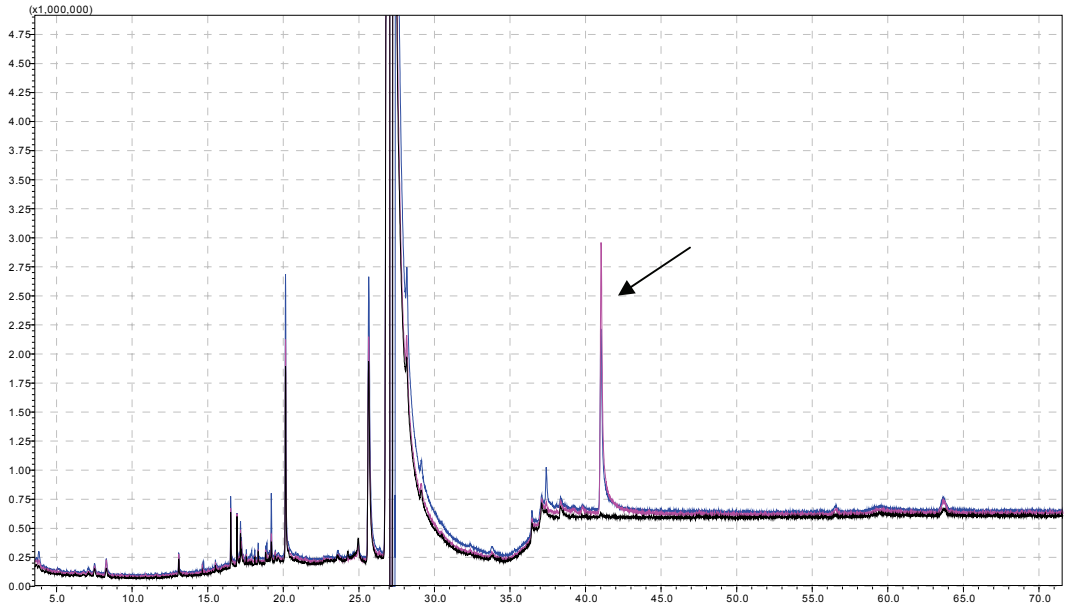
Supplementary Tab. 1 Bacterial strains and plasmids used in this study

Strain and plasmid	Relevant trait(s)	Source or reference
Strains		
E. coli BL21 (DE3)	F^- ompT gal dcm lon $hsdS_B(r_B^- m_B^-)$ λ(DE3)	Novagen
E. coli TOP10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ^-	Invitrogen
Plasmids		
pPQE30crtC$_{Rg}$	pPQ30; carries the *Bam*HI/*Kpn*I fragment with *crtC* from *R. gelatinosus*	(Steiger et al. 2003)
pTcrt3	pBluescript SK(+); carries the wild-type *Bam*HI-*Sac*I fragment of the *crtDC* operon of *T. roseopersicina*	(Kovacs et al. 2003)
pET15-b	*E. coli* general expression vector with N-terminal His tag; Amp^r	Novagen
pET15b_CrtC$_{Rg}$	pET15-b with 1252-bp *Nde*I/*Xho*I fragment from pPQE30crtC$_{Rg}$	this work
pET15b_CrtC$_{Tr}$	pET15-b with 1249-bp *Nde*I/*Xho*I fragment from pTcrt3	this work

Supplementary figures 1 – 5

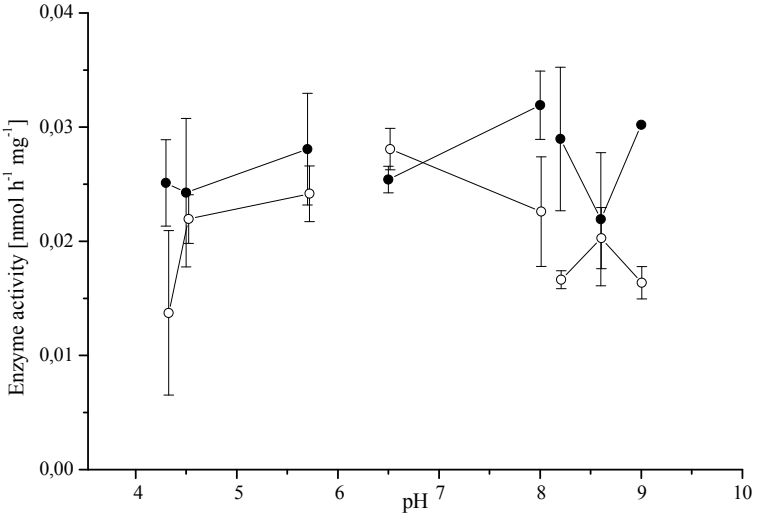


Supplementary Fig. 1 Structure of substrates used for substrate specificity studies of *R. gelatinosus* and *T. roseopersicina* carotenoid 1,2-hydratase

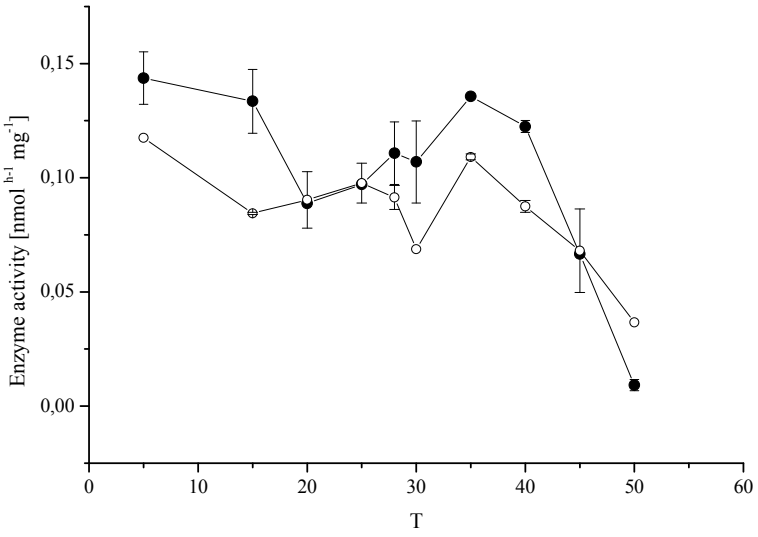


Supplementary Fig. 2 GC separation of products formed *in vitro* by *Escherichia coli* extract expressing *Rg*CrtC (*pink line*) and *Tr*CrtC (*blue line*) using the substrate geranylgeraniol (RT 27 min). Obtained product is indicated with arrow (RT 41 min). Extract with empty vector pET15-b served as negative control (*black line*)

a

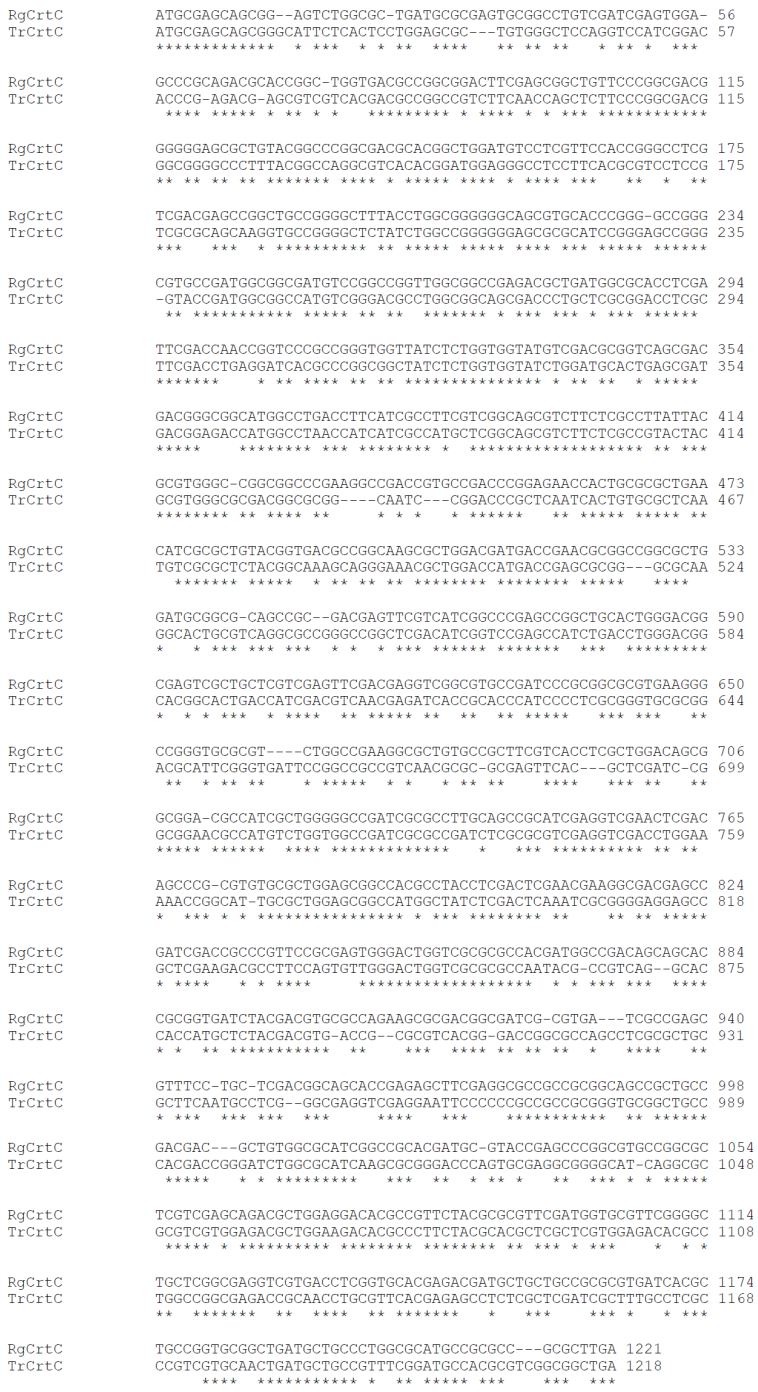


b



Supplementary Fig. 3 Stability of *Rg*CrtC (*filled circle*) and *Tr*CrtC (*empty circle*) at different pH (**a**) and temperature (**b**) values. The remaining activity was assayed under standard assay conditions after the cell-free extracts had been incubated in the corresponding buffers (pH 4.3 to pH 9.0) or at the indicated temperature (5–50ºC) in 50 mM Na_2HPO_4 sodium phosphate (pH 8.0) for 30 min

a

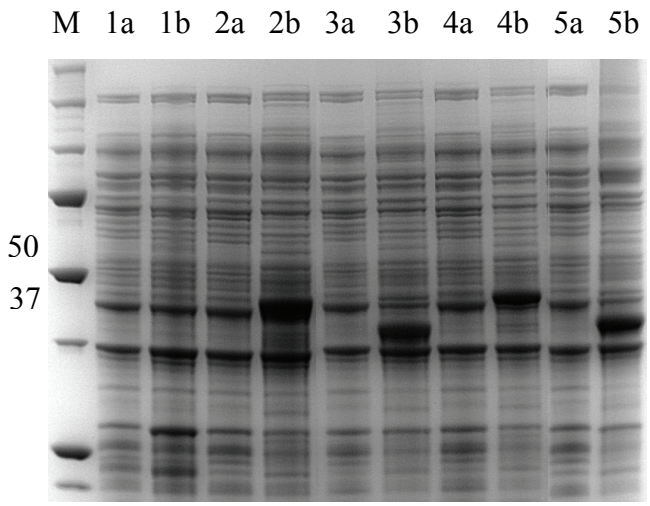

RgCrtC ATGCGAGCAGCGG--AGTCTGGCGC-TGATGCGCGAGTGCGGCCTGTCGATCGAGTGGA- 56
TrCrtC ATGCGAGCAGCGGGCATTCTCACTCCTGGAGCGC---TGTGGGCTCCAGGTCCATCGGAC 57
RgCrtC GCCCGCAGACGCACCGGC-TGGTGACGCCGGCGGACTTCGAGCGGCTGTTCCCGGCGACG 115
TrCrtC ACCCG-AGACG-AGCGTCGTCACGACGCCGGCCGTCTTCAACCAGCTCTTCCCGGCGACG 115
RgCrtC GGGGGAGCGCTGTACGGCCCGGCGACGCACGGCTGGATGTCCTCGTTCCACCGGGCCTCG 175
TrCrtC GGCGGGGCCCTTTACGGCCAGGCGTCACACGGATGGAGGGCCTCCTTCACGCGTCCTCCG 175
RgCrtC TCGACGAGCCGGCTGCCGGGGCTTTACCTGGCGGGGGGCAGCGTGCACCCGGG-GCCGGG 234
TrCrtC TCGCGCAGCAAGGTGCCGGGGCTCTATCTGGCCGGGGGGAGCGCGCATCCGGGAGCCGGG 235
RgCrtC CGTGCCGATGGCGGCGATGTCCGGCCGGTTGGCGGCCGAGACGCTGATGGCGCACCTCGA 294
TrCrtC -GTACCGATGGCGGCCATGTCGGGACGCCTGGCGGCAGCGACCCTGCTCGCGGACCTCGC 294
RgCrtC TTCGACCAACCGGTCCCGCCGGGTGGTTATCTCTGGTGGTATGTCGACGCGGTCAGCGAC 354
TrCrtC TTCGACCTGAGGATCACGCCCGGCGGCTATCTCTGGTGGTATCTGGATGCACTGAGCGAT 354
RgCrtC GACGGGCGGCATGGCCTGACCTTCATCGCCTTCGTCGGCAGCGTCTTCTCGCCTTATTAC 414
TrCrtC GACGGAGACCATGGCCTAACCATCATCGCCATGCTCGGCAGCGTCTTCTCGCCGTACTAC 414
RgCrtC GCGTGGGC-CGGCGGCCCGAAGGCCGACCGTGCCGACCCGGAGAACCACTGCGCGCTGAA 473
TrCrtC GCGTGGGCGCGACGGCGCGG----CAATC---CGGACCCGCTCAATCACTGTGCGCTCAA 467
RgCrtC CATCGCGCTGTACGGTGACGCCGGCAAGCGCTGGACGATGACCGAACGCGGCCGGCGCTG 533
TrCrtC TGTCGCGCTCTACGGCAAAGCAGGGAAACGCTGGACCATGACCGAGCGCGG---GCGCAA 524
RgCrtC GATGCGGCG-CAGCCGC--GACGAGTTCGTCATCGGCCCGAGCCGGCTGCACTGGGACGG 590
TrCrtC GGCACTGCGTCAGGCGCCGGGCCGGCTCGACATCGGTCCGAGCCATCTGACCTGGGACGG 584
RgCrtC CGAGTCGCTGCTCGTCGAGTTCGACGAGGTCGGCGTGCCGATCCCGCGGCGCGTGAAGGG 650
TrCrtC CACGGCACTGACCATCGACGTCAACGAGATCACCGCACCCATCCCCTCGCGGGTGCGCGG 644
RgCrtC CCGGGTGCGCGT----CTGGCCGAAGGCGCTGTGCCGCTTCGTCACCTCGCTGGACAGCG 706
TrCrtC ACGCATTCGGGTGATTCCGGCCGCCGTCAACGCGC-GCGAGTTCAC---GCTCGATC-CG 699
RgCrtC GCGGA-CGCCATCGCTGGGGGCCGATCGCGCCTTGCAGCCGCATCGAGGTCGAACTCGAC 765
TrCrtC GCGGAACGCCATGTCTGGTGGCCGATCGCGCCGATCTCGCGCGTCGAGGTCGACCTGGAA 759
RgCrtC AGCCCG-CGTGTGCGCTGGAGCGGCCACGCCTACCTCGACTCGAACGAAGGCGACGAGCC 824
TrCrtC AAACCGGCAT-TGCGCTGGAGCGGCCATGGCTATCTCGACTCAAATCGCGGGGAGGAGCC 818
RgCrtC GATCGACCGCCCGTTCCGCGAGTGGGACTGGTCGCGCGCCACGATGGCCGACAGCAGCAC 884
TrCrtC GCTCGAAGACGCCTTCCAGTGTTGGGACTGGTCGCGCGCCAATACG-CCGTCAG--GCAC 875
RgCrtC CGCGGTGATCTACGACGTGCGCCAGAAGCGCGACGGCGATCG-CGTGA---TCGCCGAGC 940
TrCrtC CACCATGCTCTACGACGTG-ACCG--CGCGTCACGG-GACCGGCGCCAGCCTCGCGCTGC 931
RgCrtC GTTTCC-TGC-TCGACGGCAGCACCGAGAGCTTCGAGGCGCCGCCGCGGCAGCCGCTGCC 998
TrCrtC GCTTCAATGCCTCG--GGCGAGGTCGAGGAATTCCCCCCGCCGCCGCGGGTGCGGCTGCC 989
RgCrtC GACGAC---GCTGTGGCGCATCGGCCGCACGATGC-GTACCGAGCCCGGCGTGCCGGCGC 1054
TrCrtC CACGACCGGGATCTGGCGCATCAAGCGCGGGACCCAGTGCGAGGCGGGGCAT-CAGGCGC 1048
RgCrtC TCGTCGAGCAGACGCTGGAGGACACGCCGTTCTACGCGCGTTCGATGGTGCGTTCGGGGC 1114
TrCrtC GCGTCGTGGAGACGCTGGAAGACACGCCCTTCTACGCACGCTCGCTCGTGGAGACACGCC 1108
RgCrtC TGCTCGGCGAGGTCGTGACCTCGGTGCACGAGACGATGCTGCTGCCGCGCGTGATCACGC 1174
TrCrtC TGGCCGGCGAGACCGCAACCTGCGTTCACGAGAGCCTCTCGCTCGATCGCTTTGCCTCGC 1168
RgCrtC TGCCGGTGCGGCTGATGCTGCCCTGGCGCATGCCGCGCC---GCGCTTGA 1221
TrCrtC CCGTCGTGCAACTGATGCTGCCGTTTCGGATGCCACGCGTCGGCGGCTGA 1218

b

```
RgCrtC          MRAAESGADARVRPVDRVEPADAPAGDAGGLRAAVPGDGGSAVRPGDARLDVLVPPGLVD 60
TrCrtC          MRAAGILTPGALWAPGPSDTRDERRHDAGRLQPALPGDGRGPLRPGVTRMEGLLHASSVA 60
                ****   : . : . .  :. *    *** *:.*:**** ..:*** :*:: *: .. *

RgCrtC          EPAAGALPGGGQRAPGAGRADGGDVRPVGGRDADGAPRFDQPVPPGGYLWWYVDAVSDDG 120
TrCrtC          QQGAGALSGRGERASGSRGTDGGHVGTPGGSDPARGPRFDLRITPGGYLWWYLDALSDDG 120
                : .****.* *:**.*:  :***.* . ** *.  .****  :.********:**:****

RgCrtC          RHGLTFIAFVGSVFSPYYAWAGGPKADRADPENHCALNIALYGDAGKRWTMTERGRRWMR 180
TrCrtC          DHGLTIIAMLGSVFSPYYAWAR--RRGNPDPLNHCALNVALYGKAGKRWTMTERGRKALR 178
                 ****:**::***********   : ...** ******:****.***********: :*

RgCrtC          RSRDEFVIGPSRLHWDGESLLVEFDEVGVPIPRRVKGRVRVWPKALCRFVTSLDSGGRHR 240
TrCrtC          QAPGRLDIGPSHLTWDGTALTIDVNEITAPIPSRVRGRIRVIPAAVNAREFTLDPAERHV 238
                :: ..: ****:* *** :* ::.:*: .*** **:**:** * *:     :**.. **

RgCrtC          WGPIAPCSRIEVELDSPRVRWSGHAYLDSNEGDEPIDRPFREWDWSRATMADSSTAVIYD 300
TrCrtC          WWPIAPISRVEVDLEKPALRWSGHGYLDSNRGEEPLEDAFQCWDWSRANTP-SGTTMLYD 297
                * **** **:**:*:.* :*****.*****.*:**:: .*: ******. . *.*:::**

RgCrtC          VRQKRDGDRVIAERFLLDGSTESFEAPPRQPLPTT-LWRIGRTMRTEPGVPALVEQTLED 359
TrCrtC          VTARHGTGASLALRFNASGEVEEFPPPPRVRLPTTGIWRIKRGTQCEAGHQARVVETLED 357
                *  ::. .  :* **  .*..*.* .***  **** :*** *  : *.*  * * :****

RgCrtC          TPFYARSMVRSGLLGEVVTSVHETMLLPRVITLPVRLMLPWRMPRRA- 406
TrCrtC          TPFYARSLVETRLAGETATCVHESLSLDRFASPVVQLMLPFRMPRVGG 405
                *******:*.: * **..*.***:: * *. :  *:****:**** .
```

Supplementary Fig. 4 DNA (**a**) and amino acid (**b**) sequence alignment of *Rg*CrtC and *Tr*CrtC. Identical bases/amino acids are highlighted with a star

a



b

```
RgCrtC              1 MRAAESGADARVRPVDRVEPADAPAGDAGGLRAAVPGDGGSAVRPGDARLDVLVPPGLVDEPAAGALPGGGQRAPGAGRA
Rg_trunc            1 ----------------------------------------------------------..........................

RgCrtC             81 DGGDVRPVGGRDADGAPRFDQPVPPGGYLWWYVDAVSDDGRHGLTFIAFVGSVFSPYYAWAGGPKADRADPENHCALNIA
Rg_trunc           25 ..........................................................................................

RgCrtC            161 LYGDAGKRWTMTERGRRWMRRSRDEFVIGPSRLHWDGESLLVEFDEVGVPIPRRVKGRVRVWPKALCRFVTSLDSGGRHR
Rg_trunc          105 ..........................................................................................

RgCrtC            241 WGPIAPCSRIEVELDSPRVRWSGHAYLDSNEGDEPIDRPFREWDWSRATMADSSTAVIYDVRQKRDGDRVIAERFLLDGS
Rg_trunc          185 ..........................................................................................

RgCrtC            321 TESFEAPPRQPLPTTLWRIGRTMRTEPGVPALVEQTLEDTPFYARSMVRSGLLGEVVTSVHETMLLPRVITLPVRLMLPW
Rg_trunc          265 ..........................................................................................

RgCrtC            401 RMPRRA
Rg_trunc          345 ......



TrCrtC              1 MRAAGILTPGALWAPGPSDTRDERRHDAGRLQPALPGDGRGPLRPGVTRMEGLLHASSVAQQGAGALSGRGERASGSRGT
Tr_trunc            1 ----------------------------------------------..............................

TrCrtC             81 DGGHVGTPGGSDPARGPRFDLRITPGGYLWWYLDALSDDGDHGLTIIAMLGSVFSPYYAWARRRGNPDPLNHCALNVALY
Tr_trunc           36 ..........................................................................................

TrCrtC            161 GKAGKRWTMTERGRKALRQAPGRLDIGPSHLTWDGTALTIDVNEITAPIPSRVRGRIRVIPAAVNAREFTLDPAERHVWW
Tr_trunc          116 ..........................................................................................

TrCrtC            241 PIAPISRVEVDLEKPALRWSGHGYLDSNRGEEPLEDAFQCWDWSRANTPSGTTMLYDVTARHGTGASLALRFNASGEVEE
Tr_trunc          196 ..........................................................................................

TrCrtC            321 FPPPPRVRLPTTGIWRIKRGTQCEAGHQARVVETLEDTPFYARSLVETRLAGETATCVHESLSLDRFASPVVQLMLPFRM
Tr_trunc          276 ..........................................................................................

TrCrtC            401 PRVGG
Tr_trunc          356 .....
```

Supplementary Fig. 5 SDS-PAGE (10%) analysis (**a**) and amino acid sequence alignment (**b**) of the wildtype and the truncated CrtC. *M* Precision plus protein standard; *a* whole cells before induction; *b* whole cells after induction with 0.1 mM IPTG and expression overnight at 25°C; *1* pET15-b; *2* *Tr*CrtC wildtype (38 and 44 kDa); *3* *Tr*CrtC truncated (39 kDa); *4* *Rg*CrtC wildtype (44 kDa); *5* *Rg*CrtC truncated (39 kDa)