

Insulin-like growth factor 1 (IGF1) overexpression on a phycocyanin carrier protein in cyanobacteria

Bharat Kumar Majhi and Anastasios Melis*

Department of Plant and Microbial Biology, 111 Koshland Hall, MC-3102, University of California, Berkeley, CA 94720-3102, USA

SUPPLEMENTARY INFORMATION

*Correspondence: melis@berkeley.edu

Nucleotide and Protein Sequences

1. *cpcB*6xHis*S*tev*IGF1*

Sequence for 5'- homologous recombination

UPPER CASE

UPPER CASE BOLD

UPPER CASE UNDERLINE

LOWER CASE BOLD

UPPER CASE BOLD ITALICS

UPPER CASE UNDERLINE

LOWER CASE BOLD

UPPER CASE UNDERLINE ITALICS

- *cpcB* (without stop codon)

- His-tag

- Spacer (PAEKWAPGGS)

- *tev* cleavage sequence (ENLYFQG)

- codon-optimized IGF1 gene with stop codon

- intergenic sequence

- *CmR* (chloramphenicol selection cassette)

- Transcription terminator

Sequence for 3'- homologous recombination

LOWER CASE BOLD ITALICS

LOWER CASE UNDERLINE ITALICS

- *cpcB-cpcA* intergenic sequence

- *cpcA* (partial)

ATGTTGACGTATTCACTCGGGTTGTTCCCAAGCTGATGCTCGCGGCGAGTACCTCTCTGTTCTCAGTTAGATGCTTTGA
GCGCTACCGTTGCTGAAGGCAACAAACGGATTGATTCTGTTAACCGCATCACCGTAATGCTCCGCTATCGTTCCAAACG
CTGCTCGTGCTTTGTTGCGCCGAACAGCCCAATTAATCCAACCCGGTGAAACGCTACACCAGCCGTCGTATGGCTGCT
GTTTGCCTGACATGGAATCATCCTCCGCTATGTTACCTACGCAACCTTACCAGGCGACGCTCCGTTCTAGAAGATCGTT
GCTTGAACGGTCTCCGTGAAACCTACGTTGCCCTGGGTGTTCCCGGTGCTCCGTAGCTGCTGGCGTTCAAAAAATGAAA
GAAGCTGCCCTGGACATCGTTAACGATCCCAATGGCATCACCCGTGGTGATTGCAGTGCTATCGTTGCTGAAATCGCTGG
TACTTCGACCGCGCGCTGCTGCCGTAGCC**ACCACCACCACCACCACC**CCCGCGAAAAATGGGCCCGCGGCTCCga
aaacctgtattttcagggcACCCTCTGTGGGGCTGAGCTAGTGGATGCGCTGCAGTTTGTGGCGATCGCGGCTTCTACTTC
AACAAACCACTGGCTATGGCAGTTCTCCCGCGGGCGCCCAACGGGCATCGTGGAATGCTGCTTTCGCACT
GCGATCTGCGCGCTTGAAATGTA**CTGCGCTCCTCTCAAGCCAGCCAAATCCGCTTAA**TGATAACGGGAAGCCCTGGG
CCAACCTTTGGCGAAAAATGAGACGTTGATCGGCACGTAAGAGGTTCCAACCTTACCATAATGAAATAAGATCACTACC
GGGCGTATTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAAatggagaaaaaatcactggatataaccaccgtgat
atcccaatggcatgtaaagaacattttgaggcatttcagtcagttgctcaatgtacctataaccagaccgttcagctggatattacggccttttaagaccgt
aaagaaaaataagcacaagttttatccggcctttattcacattctgcccgcctgatgaatgctcatccggaattccgatggcaatgaagacggtgagctg
gtgatatgggatagttacccttgttacaccgttttccatgagcaaacgaaacgttttcatcgcctctggagtgaataccacgacgatttccggcagtttcta
cacatatattcgaagatgtggcgtgttacggtgaaaacctggcctatttccctaaaggtttattgagaatatttttctcagccaatccctgggtgagtt
tcaccagttttgatttaacgtggccaatattgacaacttcttccccgcttttaccatgggcaaatattatcgaaggcgacaaggtgctgatccgctg
gcgattcaggttcatcatgcccgtttgtgatgcttccatgctggcagaatgcttaataaataacagttactgcatgagtgaggcggggcgtaaTTTT
TTTAAGGCAGTTATTGGTGCCCTAAACGCCTGGTGCTACGCCTGAATAAGTGATAAtctggttattttaaaaaccaactttactcag
gttccatacccgagaaaaatccagcttaaaagctgacatatctaggaaaaattttcacatttcaacgggagataaccagaacaatqaaaaccctttaactg
aaqccgtttccaccqctgactctcaaggtcgttttctgaqcaqaccqaattqcaaatqctttcagtcgtctacgtcaagctaatqctggtttqcaagccg
taaaqctctqaccqacaatqccaqagcttqtaaatqgtqctqccaagcgtttataacaattcccctacaccacccaacccaagqcaacaacttt

gtcgcggaatcaacgggataaaqacaagtgtgcccgagacatcggtactactccgcacgttacctactgcttaqgtgtagtgataccggtccttga
tgaqtacttgatcgggtattgatgaaatcaaccgcacctttgacctctccccagctggtatgtt

Protein Sequence

UPPER CASE ITALICS - *cpcB* (without stop codon)
UPPER CASE BOLD - His-tag
UPPER CASE BOLD - Spacer (PAEKWAPGGS)
UPPER CASE UNDERLINE - tev cleavage sequence (ENLYFQG)
UPPER CASE BOLD ITALICS - codon-optimized IGF1 gene with stop codon

MFDVFTRVVSQADARGEYLSGSQLDALSATVAEGNKRIDSVNRITGNASAIVSNAARALFAEQPQLIQPGGNAYTSRRMAACL
RDMEILRYVYATFTGDASVLEDRLNGLRETYVALGVPASVAAGVQKMKEAALDIVNDPNGITRGDCSAIVAEIAGYFDRA
AAAVAHHHHHH**PAEKWAPGGS**ENLYFQGTLCGAE**LV**DA**LQVCGDRGFYFNKPTGYGSSRRAPQTGIVDECCFRSCD**
LRRLEMYCAPLKPAKSA*

2. *cpcB**6xHis*tev*IGF1

Sequence for 5'- homologous recombination

UPPER CASE - *cpcB* (without stop codon)
UPPER CASE BOLD - His-tag
LOWER CASE BOLD - tev cleavage sequence (ENLYFQG)
UPPER CASE BOLD ITALICS - codon-optimized IGF1 gene with stop codon
UPPER CASE UNDERLINE - intergenic sequence
LOWER CASE BOLD - *CmR* (chloramphenicol selection cassette)
UPPER CASE UNDERLINE ITALICS - Transcription terminator

Sequence for 3'- homologous recombination

LOWER CASE BOLD ITALICS - *cpcB-cpcA* intergenic sequence
LOWER CASE UNDERLINE ITALICS - *cpcA* (partial)

ATGTTTCGACGTATTCACTCGGGTTGTTTCCCAAGCTGATGCTCGCGGCGAGTACCTCTCTGGTTCTCAGTTAGATGCTTTGA
GCGCTACCGTTGCTGAAGGCAACAAACGGATTGATTCTGTTAACCGCATCACCGGTAATGCTCCGCTATCGTTTCCAAACG
CTGCTCGTGCTTTGTCGCCGAACAGCCCCAATTAATCCAACCCGGTGAAACGCCTACACCAGCCGTCGTATGGCTGCTT
GTTTGCCTGACATGGAATCATCCTCCGCTATGTTACCTACGCAACCTTACCGGCGACGCTCCGTTCTAGAAGATCGTT
GCTTGAACGGTCTCCGTGAAACCTACGTTGCCCTGGGTGTTCCCGGTGCTCCGTAGCTGCTGGCGTTCAAAAAATGAAA
GAAGCTGCCCTGGACATCGTTAACGATCCCAATGGCATCACCCGTGGTGATTGCAGTGCTATCGTTGCTGAAATCGCTGG
TACTTCGACCGCGCCGCTGCTGCCGTAGCC**CACCACCACCACCACC**gaaaacctgtat**tttcaggggACCCTCTGTGGGGCT**
GAGTAGTGGATGCGCTGCAGTTTGTGGCGATCGCGGCTTCTACTTCAACAAACCACTGGGCTATGGCAGTTCTTC
CCGGCGGGCGCCCAACGGGCATCGTGGATGAATGCTGCTTTCGAGCTGCGATCTGCGCCGCTTGAAATGTACTGC
GCTCCTCTAAGCCAGCCAAATCCGCTTAATGATACCGGGAAGCCCTGGGCCAACTTTTGGCGAAAATGAGACGTTGAT
CGGCACGTAAGAGGTTCCAACCTTACCATAATGAAATAAGATCACTACCGGGCGTATTTTTGAGTTATCGAGATTTTC

AGGAGCTAAGGAAGCTAAAatggagaaaaaatcactggatataccaccgttgatataccaatggcatgtaaagaacattttgaggcatttca
 gtcagttgctcaatgtacctataaccagaccgttcagctggatattacggccttttaagaccgtaaagaaaaataagcacaagttttatccggcctttattc
 acattcttccccctgatgaatgctcatccggaattccgatggcaatgaaagacggtagctgggatagttcaccctgttacaccgttttcc
 atgagcaaacgaaacgttttcacgctctggagtgaataccacgacgattccggcagtttctacacatatattcgcaagatgtggcgtgttacggtgaaaaac
 ctggcctatttccctaaagggttttatgagaatatttttctcctcagccaatccctgggtgagtttaccagttttgattaaacgtggccaatattgacaact
 tcttcgccccgttttaccatgggcaaatattatacgaaggcacaaggtgctgatccgctggcgattcaggttcatcatcgctttgtgatggcttccatg
 tcggcagaatgcttaataatgacaacagtagctgcatgagtgaggcagggcgggcgtaa **TTTTTTAAGGCAGTTATTGGTGCCTTAAACGC**
CTGGTGCTACGCCTGAATAAGTGATAAtctggttattttaaaccactttactcaggttccatacccagaaaaatccagcttaagctgac
 atatctaggaaaaattttcacattctaaccgggagataccagaaca atgaaaccctttaaactgaaqccgtttccaccqctgactctcaaggtcgttttctg
aqcaqaccqaattqcaaatgctttcggtcgtctacqcaagtaaatgctggttqcaaaccqctaaagctctqaccqacaatqcccaqagcttqgtaaa
tqgtqctqcccaagcgtttatacaaatccctacaccaccaaaccaagqcaacaactttgctqcgatcaacqaggtaaagacaagtqgcccq
ggacatcagctactacccqcatcgttacctactgctggtggtaccggtcctttgagatgagtagctgacccqgtattgagaaatcaaccgq
accttgacctctccccagctggtatgtt

Protein Sequence

<i>UPPER CASE ITALICS</i>	- <i>cpcB</i> (without stop codon)
UPPER CASE BOLD	- His-tag
<u>UPPER CASE UNDERLINE</u>	- tev cleavage sequence (ENLYFQG)
<i>UPPER CASE BOLD ITALICS</i>	- codon-optimized IGF1 gene with stop codon

MFDVFTRVVSQADARGEYLSGSQLDALSATVAEGNKRIDSVNRITGNASAIVSNAARALFAEQPQLIQPGGNAYTSRRMAACL
 RDMEILRYVYATFTGDASVLEDRCLNGLRETYVALGVPGASVAAGVQKMKEAALDIVNDPNGITRGDCSAIVAEIAGYFDRA
 AA~~AV~~AHHHHHHHENLYFQGTLCGAELVDALQVCGDRGFYFNKPTGYGSSRRAPQTGIVDECCFRSCDLRRLEMYCAPLKP
 AKSA*

3. *cpcBS*6xHis*hrv*IGF1**

Sequence for 5'- homologous recombination

<i>UPPER CASE</i>	- <i>cpcB</i> (without stop codon)
<u>UPPER CASE UNDERLINE</u>	- Spacer (PAEKWAPGGS)
UPPER CASE BOLD	- His-tag
LOWER CASE BOLD	- HRV3C cleavage sequence (LEVLFGQP)
<i>UPPER CASE BOLD ITALICS</i>	- codon-optimized reverse IGF1 gene with stop codon
<u>UPPER CASE UNDERLINE</u>	- intergenic sequence
LOWER CASE BOLD	- <i>CmR</i> (chloramphenicol selection cassette)
<u><i>UPPER CASE UNDERLINE ITALICS</i></u>	- Transcription terminator

Sequence for 3'- homologous recombination

<i>LOWER CASE BOLD ITALICS</i>	- <i>cpcB-cpcA</i> intergenic sequence
<i>LOWER CASE UNDERLINE ITALICS</i>	- <i>cpcA</i> (partial)

ATGTTGACGTATTCACTCGGGTTGTTCCCAAGCTGATGCTCGCGGCGAGTACCTCTCTGGTTCTCAGTTAGATGCTTTGA
 GCGCTACCGTTGCTGAAGGCAACAAACGGATTGATTCTGTTAACCGCATCACCGGTAATGCTTCCGCTATCGTTCCAACG
 CTGCTCGTGCTTTGTTGCGCCGAACAGCCCAATTAATCCAACCCGGTGAAACGCCTACACCAGCCGTCGTATGGCTGCTT
 GTTTGCCTGACATGGAAATCATCCTCCGCTATGTTACCTACGCAACCTTACCAGGCGACGCTTCCGTTCTAGAAGATCGTT
 GCTTGAACGGTCTCCGTGAAACCTACGTTGCCCTGGGTGTTCCCGGTGCTCCGCTAGCTGCTGGCGTTCAAAAAATGAAA
 GAAGCTGCCCTGGACATCGTTAACGATCCCAATGGCATCACCCGTGGTGATTGCAGTGCTATCGTTGCTGAAATCGCTGG
 TTACTTCGACCGCGCCGCTGCTGCCGTAGCCCCCGCCGAAAAATGGGCGCGGGCGGCTCC**CACCACCACCACCACC****Cct**
tgaggtactcttccagggtccg**GCCAGTAAGGCCCGAAATTACCAGCTTGTATATGGAGTTGCGTCTAGATTGTTCCCGT**
TTCTGTTGCAAGATGTTATCGGTACGCAACCAGCTCGCCGATCTTCTAGTGGTTATGGGACGCCAAAAATTTCTACTTC
GGACGCGATGGATGTGATTCCAACCTGGCTGATGTGCTCGAAGCGGGATGCTTGACCCGGCCAGGCTAATGATACCGG
GAAGCCCTGGGCCAACTTTTGGCGAAAAATGAGACGTTGATCGGCACGTAAGAGGTTCCAACCTTTACCATAATGAAAT
AAGATCACTACCGGGCGTATTTTTGAGTTATCGAGATTTTACAGGAGCTAAGGAAGCTAAA**atggagaaaaaatcactggat**
ataccaccgtgatataccaatggcatcgtaaagaacattttgaggcatttcagtcagttgctcaatgtacctataaccagaccgttcagctggatattacg
gccttttaagaccgtaaagaaaaataagcacaagttttaccgctttattcacattctgccgctgatgaatgctcatccggaattccgtatggcaatg
aaagacggtagctggtagatgggatagttcaccttggtaaccggtttccatgagcaactgaaacgtttcatcgctctggagtgaataccacgacga
ttccggcagtttctacacatatattcgcaagatgtggcgtgttacggtgaaaacctggcctatttcctaaagggtttattgagaatatgttttcgtctcagcc
aatccctgggtgagtttaccagtttgattaaacgtggccaatattggacaacttctcgccccgtttaccatgggcaaatattatacgaaggcga
ggtgctgatgccgtggcattcaggttcacatgccgtttgtagtgcttcacatgctggcagaatgcttaataacagtaactgctgatgagtgccagg
cgggggcgtaaTTTTTTAAGGCAGTTATTGGTGCCCTTAAACGCCTGGTGTACGCCTGAATAAGTGATAA**ctgggtatttttaa**
aaccaactttactcaggttccatacccgagaaaaatccagcttaaaagctgacatatctaggaaaaattttcacatttcaacgggagataccagaacaatg
aaaaccctttaactgaaqccgtttccaccgctgactctcaaggtcgctttctgagcaqaccqaattgcaaatgctttcggctcctacatcaaaqtaatg
ctggtttgcaagccgctaaagctctgaccgacaatgccagagcttgtaaatggtgctgccaagccgtttataacaaatcccctacaccacccaacc
caaggcaacaactttgctcggatcaacgggtaaaagacaagtatgcccggacatcggctactactccqcatcgttactactgcttagttgctggtg
gtaccgctccttggatgagtagctgacccggtattgatgaaatcaaccgacacctttgacctctcccccaqctggtatgtt

Protein Sequence

UPPER CASE ITALICS
UPPER CASE BOLD
UPPER CASE BOLD
UPPER CASE UNDERLINE
UPPER CASE BOLD ITALICS

- *cpcB* (without stop codon)
- **Spacer (PAEKWAPGGS)**
- **His-tag**
- **HRV3C cleavage sequence (LEVLFQGP)**
- **codon-optimized IGF1 gene with stop codon**

MFDVFTRVVSQADARGEYLSGSQLDALSATVAEGNKRIDSVNRITGNASAIVSNAARALFAEQPLIQPGGNAYTSRRMAA
CLRDMEILRYVYATFTGDASVLEDRCLNGLRETYVALGVPASVAAGVQKMKEAALDIVNDPNGITRGDCSAIVAEIAGYF
*DRAAAAVAPAEKWAPGGS***HHHHHH**LEVLFQGP**ASKAPKLPACYMELRRLCDSRFCCEDEVIGTQPARRSSSGYTPKNF**
YFGRDGCVFQLADVLEAGCLTRPG*

4. *cpcB*6xHis*hrv*IGF1*

Sequence for 5'- homologous recombination

UPPER CASE

UPPER CASE BOLD

LOWER CASE BOLD

UPPER CASE BOLD ITALICS

UPPER CASE UNDERLINE

LOWER CASE BOLD

UPPER CASE UNDERLINE ITALICS

- *cpcB* (without stop codon)

- His-tag

- HRV3C cleavage sequence (LEVLFQGP)

- codon-optimized reverse IGF1 gene with stop codon

- intergenic sequence

- *CmR* (chloramphenicol selection cassette)

- Transcription terminator

Sequence for 3'- homologous recombination

LOWER CASE BOLD ITALICS

LOWER CASE UNDERLINE ITALICS

- *cpcB-cpcA* intergenic sequence

- *cpcA* (partial)

ATGTTGACGTATTCACTCGGGTTGTTTCCCAAGCTGATGCTCGCGGCGAGTACCTCTCTGGTTCTCAGTTAGATGCTTTGA
GCGCTACCGTTGCTGAAGGCAACAAACGGATTGATTCTGTTAACCGCATCACCGGTAATGCTTCCGCTATCGTTTCCAACG
CTGCTCGTGCTTTGTTGCGCGAACAGCCCAATTAATCCAACCCGGTGAAACGCCTACACCAGCCGTCGTATGGCTGCT
GTTTGC GTGACATGGAATCATCCTCCGCTATGTTACCTACGCAACCTTACC GCGGACGCTTCCGTTCTAGAAGATCGTT
GCTTGAACGGTCTCCGTGAAACCTACGTTGCCCTGGGTGTTCCCGGTGCTTCCGTAGCTGCTGGCGTTCAAAAAATGAAA
GAAGCTGCCCTGGACATCGTTAACGATCCCAATGGCATCACCCGTGGTGATTGCAGTGCTATCGTTGCTGAAATCGCTGG
TACTTCGACCGCGCCGCTGCTGCCGTAGCC**CACCACCACCACCACC**cttgaggtactcttcagggtccg**GCCAGTAAGGCC**
CGAAATTACCAGCTTGTATATGGAGTTGCGTCGTCTAGATTGTTCCGTTTCTGTTGCGAAGATGTTATCGGTACGCAA
CCAGCTCGCGATCTTCTAGTGGTTATGGGACGCCAAAAATTTCTACTTCGGACGCGATGGATGTGTATTCCAAGTGGC
TGATGTGCTCGAAGCGGGATGCTTGACCGGCCAGGCTAATGATACCGGGAAGCCCTGGGCCAACTTTTGGCGAAAAT
GAGACGTTGATCGGCACGTAAGAGGTTCCAACCTTCCACATAATGAAATAAGATCACTACCGGGCGTATTTTTGAGTT
ATCGAGATTTTCAGGAGCTAAGGAAGCTAAAatggagaaaaaatcactggatataaccaccgttgatataatccaatggcatcgtaaagaac
atttgaggcatttcagtcagttgctcaatgtacctataaccagaccgttcagctggatattacggccttttaagaccgtaagaaaaataagcacaagtt
tatccggcctttattcacattcttgccgcctgatgaatgctcatccggaattccgatggcaatgaaagacggtagctggatggtgatgtttcacct
tgttacaccgttttccatgagcaaaactgaaacgttttcatcgctctggagtgaataaccagacgatttccggcagtttctacacatatattcgaagatgtggc
tgttacggtgaaaacctggcctatttccctaaagggttattgagaatatgttttctgctcagccaatccctgggtgagtttaccagttttgattaaacgtgg
ccaatatggacaacttcttcgccccgttttaccatgggcaaatattatacgcaaggcgacaaggctgctgatgccgtggcgattcaggttcatcatgccgtt
gtgatggcttccatgtcggcagaatgcttaatgaattacaacagtactgcatgagtgccagggcgggcgtaa **TTTTTTAAGGCAGTTATTGGTG**
CCCTAAACGCCTGGTGCTACGCTGAATAAGTGATAActggttatttttaaaaccaactttactcaggttccatacccgagaaaatccag
cttaagctgacatatctaggaatatttcaatttaacgggagataccagaaca atgaaaaaccctttaaactgaaqccgtttccaccqctgactctca
aggtcgtttctgagcaqaccqaattgcaaatgctttcgtctacgtcaagctaatgctggtttgcaagccqctaaagctgacccgacaatqccc
gagcttgtaaatggtgctqccaaqccgtttatacaaatccccacaccaccaaaaccaagqcaacaactttgctgagatcaacqgggtaaaq
caagtgtaccqggacatcgctactacctccgcatgtaacctactgcttagttgctggtggtaccggtcctttgagatgacttqatcqqgtattgt
gaaatcaaccgacctttqacctctccccagctggtatgtt

Protein Sequence

UPPER CASE ITALICS

UPPER CASE

UPPER CASE UNDERLINE

UPPER CASE BOLD ITALICS

- *cpcB* (without stop codon)

- **His-tag**

- **HRV3C cleavage sequence (LEVLFQGP)**

- **codon-optimized IGF1 gene with stop codon**

*MFDVFTRVVSQADARGEYLSGSQLDALSATVAEGNKRIDSVNRITGNASAIVSNAARALFAEQPQLIQGGNAYTSRRMAA
CLRDMEILRYVTYATFTGDASVLEDRCLNGLRETYVALGVPGASVAAGVQKMKEAALDIVNDPNGITRGDCSAIVAEIAGYF
DRAAAVAHHHHH**LEVLFQGP**ASKAPKLPACYMELRRLCDSRFCCEDVIGTQPARRSSSGYGTPKNFYFGRDGCVFQLAD
VLEAGCLTRPG**