SUPPLEMENTARY MATERIAL

Production of salidroside in metabolically engineered *Escherichia coli*

Yanfen Bai[#], Huiping Bi[#], Yibin Zhuang, Chang Liu, Tao Cai, Xiaonan Liu, Xueli Zhang, Tao Liu *, Yanhe Ma

Pathway and plasmids construction

All the primers used for gene amplification were listed as Table S2.

The pyruvate decarboxylase gene *ARO10*, was amplified from *Saccharomyces cerevisiae* S288 genome with primer pair aro-5FPNco and aro-3RPSac, the PCR fragment was digested with *Nco*I and *Sac*I , and ligated into pTrcHisB, generating the plasmid pTrc1. *73B6*^{syn} was amplified by PCR with primers UGT-5FPNco and UGT-3RPBam, and the resulting DNA fragment was cloned into vector pTrcHisB via restriction sites *Nco*I and *Bam*HI to give the plasmid pTrc-UGT. This plasmid was then used as template, and PCR was conducted with primers TrcUGT-5FPBgl/UGT-3RPBam to amplify the gene fragment *73B6*^{syn} under control of a trc promoter. Then this fragment was cloned into pTrc1 plasmid with *Bgl*II and *Bam*HI restriction sites to give the plasmid pTrc2.

The plasmid pBbA5c was modified in order to incorporate multiple cloning sites (MCS) designed to facilitate the addition of future genes. Briefly, the fragment containing *Bg*lII, *Pst*I, *Not*I, *Afl*II, *Pac*I and *Nhe*I restriction sites was amplified with the primers MCS-5FPBgl/MCS-3RPNhe using plasmid pACYCDuet-1 as the template. Then this fragment was digested by *Bg*lII and *Nhe*I, and ligated into the same restriction sites of plasmid pBbA5c to yield the plasmid pBb0, a pBbA5c derivative with novel MCS.

The *tyrA*^{*syn} was amplified with primer pair Ptet-tyrA-5FPPst/tyrA-3RPPst to generate a

DNA fragment with the upper end flanking with the constitutive P_{Ltet} promoter and RBS sequence. The fragment was then digested and ligated into pBb0 via the *Pst*I site to give the plasmid pBbA5c-tyrA^{*syn}. To facilitate cloning the following genes downstream of *tyrA*^{*syn}, the direction of gene *tyrA*^{*syn} on this plasmid was opposite to gene *lacI*. Subsequently, gene *aroG*^{*syn} was amplified with the primer pair aroG-5FPPst/aroG-3RPNot using the synthesized gene as the template, and the resulting gene fragment with individual RBS was then cloned into pBbA5c-tyrA^{*syn} via *Pst*I and *Not*I sites to yield the plasmid pBbA5c-tyrA^{*syn}-aroG^{*syn}. The gene *ppsA* was amplified using the primer pair ppsA-5FPNot/ppsA-3RPAfl, and cloned into the plasmid pBbA5c-tyrA^{*syn}-aroG^{*syn} via restriction sites *Not*I and *Afl*II, generating the plasmid pBb1. Thus the three genes *tyrA*^{*syn}, *aroG*^{*syn} and *ppsA* were driven by the P_{Ltet} promoter in an operon.

The gene *tktA* was amplified using primer pair tktA-5FPBgl and tktA-3RPPac, and the PCR product was digested wtih *Bgl*II and *Pac*I, and cloned into pBb0 under control of the P_{lacUV5} promoter. The P_{lacUV5}-tktA fragment was amplified from the above plasmid with primers Plac-tktA-5FPAfl/tktA-3RPPac by PCR, and the gene product was digested, and inserted into the plasmid pBb1 via *Afl*II and *Pac*I restriction sites. The resulting plasmid was designated as pBb1-tktA. The *aroE* gene was amplified from the genome of *E. coli* MG1655 by PCR with primers aroE-5FPPac and aroE-3RPNhe, and the resulting DNA fragment was subsequently cloned into pBb1-tktA via restriction sites *Pac*I and *Nhe*I yielding plasmid pBb2

The *aroD* gene was amplified from the genome of *E. coli* MG1655 by PCR with primers aroD-5FPBam and aroD-3RPPst, and the resulting DNA fragment was digested with *BamH*I and *Pst*I, and cloned into pBb0 downstream of the P_{lacUV5} promoter, generating pBbA5c-aroD. The codon-optimized variant of *aroB*^{op} gene fragment was generated by PCR with the primers

aroB-5FPPst/aroB-3RPNhe. The yielding DNA fragment containing the variant of *aroB*^{op} was digested with *Pst*I and *Nhe*I, and ligated into pBbA5c-aroD, generating plasmid pBbA5c-aroD-aroB^{op}. The P_{lacUV5} driven *aroD* and *aroB*^{op} fragment was amplified by PCR from above plasmid with primers Plac-5FPNhe/aroD-3RPNhe. The resulting DNA fragment was then digested with *Nhe*I, and inserted into plasmid pBb2, yielding pBb3.

^{1}H-NMR for salidroside

(600 MHz, DMSO-d_6): 9.15 (s, 1H); 7.02 (d, J = 10.0, H-C(10,14)); 6.64 (d, J = 10.0, H-C(11,13)); 4.95 (s, OH); 4.92 (s, OH); 4.88 (s, OH); 4.47 (t, J = 5.0, OH); 4.15 (d, J = 10.0, H-C(1)); 3.85 (dt, J = 5.0, 10.0, Ha-C(7)); 3.64 (dd, J = 5.0, 10.0, Ha-C(6)); 3.54 (dt, J = 5.0, 10.0, H_b-C(7)); 3.40 (m, H_b-C(6)); 3.09 (m, H-C(3)); 3.05 (m, H-C(5)); 3.02 (m, H-C(4)); 2.94 (m, H-C(2)); 2.71 (m, CH_2(8))

Fig. S1 ^{1}H spectrum of salidroside

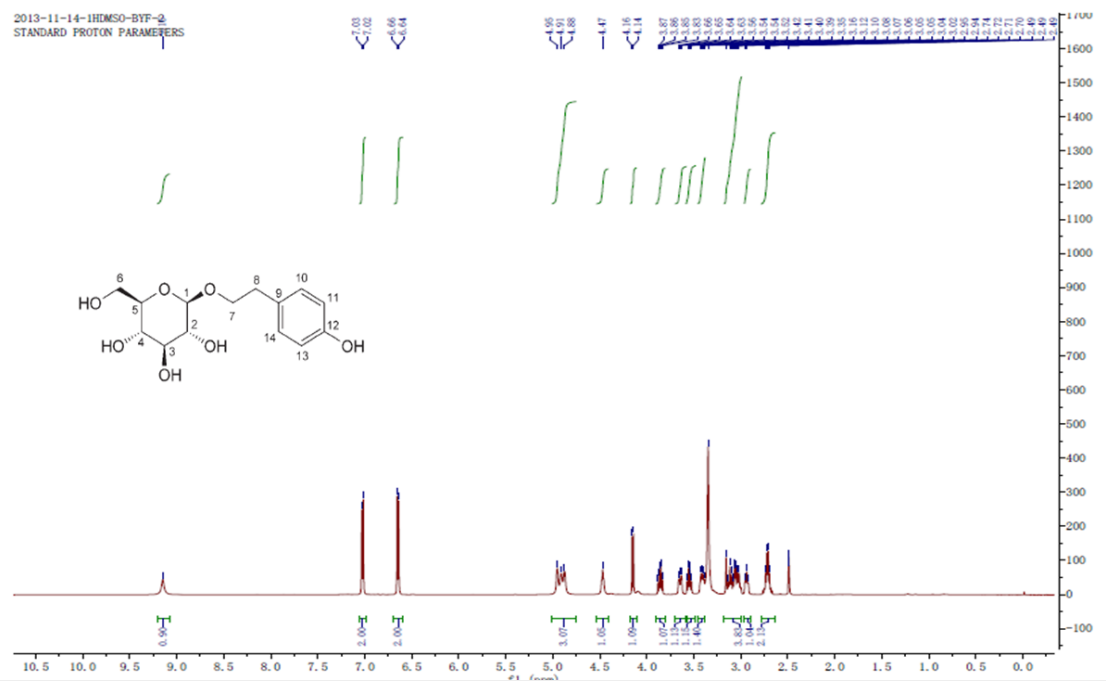

Table S1. Primers used for gene deletion in this study

Primers name	Sequence(5'→3')
feaB-5FPL	ATGACAGAGCCGCATGTAGCAGTATTAAGCCAGGTCC
	AACAGTTTCTCAGCGATTGTGTAGGCTGGAG
feaB-3RPL	TTAATACCGTACACACACCGACTTAGTTTCACACCAAC
	CGTCCAGCCATAACGGCTGACATGGGAATTAG
feaBGD-5FP	TACACCCGTGTAACGCTCAG
feaBGD-3RP	TTGACGAAACGCCAACGCTA
pykF-5FPL	AAGCAAGTTTCTCCCATCCTTCTCAACTTAAAGACTAA
	GACTGTCATGAGCGATTGTGTAGGCTGGAGCT
pykF-3RPL	TTACAGGACGTGAACAGATGCGGTGTTAGTAGTGCCG
	CTCGGTACCAGTAACGGCTGACATGGGAATTAG
pykFGD-5FP	CACGTTGGGCTGAGACACAAGC
pykFGD-3RP	CAGGATGCTTCCATCGGATTCATC
tyrR-5FPL	ATGCGTCTGGAAGTCTTTTGTGAAGACCGACTCGGTC
	TGACCCGCGAAAGCGATTGTGTAGGCTGGAG
tyrR-3RPL	TTACTCTTCGTTCTTCTTCTGACTCAGACCATATTCCCG
	CAACTTATTTAACGGCTGACATGGGAATTAG
tyrRGD-5FP	TTTTGCCGGAGCCAGCGAACTGGTGCG
tyrRGD-3RP	CTGCGGCGCAGAAGGTCTGAACAACG
pykA-5FPL	ATGTCCAGAAGGCTTCGCAGAACAAAAATCGTTACCA
	CGTTAGGCCCAAGCGATTGTGTAGGCTGGAG
pykA-3RPL	TTACTCTACCGTTAAAATACGCGTGGTATTAGTAGAAC
	CCACGGTACTTAACGGCTGACATGGGAATT
pykAGD-5FP	GCAATTACCCTCGACGTACC
pykAGD-3RP	ATGATGGCAAGACGCATCGTC
pheA-5FPL	AACACTATGACATCGGAACCATGGAACCCGTTACTGG
	CGCTGCGAGAGAGCGATTGTGTAGGCTGGAGCT
pheA-3RPL	TCAGGTTGGATCAACAGGCACCCATGGTACGTTCTCA
	CTTGGGTAACATAACGGCTGACATGGGAATTAG
pheAGD-5FP	TGAATGGGAGGCGTTTCGTCGT
pheAGD-3RP	GTTATTGCGTCAGGCGAATGAC

Table S2. Primers used for pathway construction and gene overexpression in this study

Primers name	Sequence(5’→3’)
MCS-5FPBgl	AGATCTAAGCTTATCGATCTGCAGGCGGCCGCCATA TGGGTACCCTTAAGCCCGGGGGCGAAACCCGACAG GACTATAAA
MCS-3RPNhe	GCTAGCGATATCTTAATTAAGTCGACCCCGGGGCGC GTAATCTCTTGCTCTGAAAA
Ptet-tyrA-5FPPst	AACTGCAGTCCCTATCAGTGATAGAGATTGACATCC CTATCAGTGATAGAGATACTGAGCACATCAGCAGG ACGCACTGACCGAATTCATTAAAGAGGAGAAAGGT ACCAAGGAGGAACAGACATGGTTGCTGAATTGACC GCATT
tyrA-3RPPst	AACTGCAGTTACTGGCGATTGTCATTCGCC
aroG-5FPPst	AACTGCAGAAGGAGGCCATCCATGAATTATCAGAA CGACGA
aroG-3RPNot	ATAAGAATGCGGCCGCTTACCCGCGACGCGCTTTTA CTG
ppsA-5FPNot	ATAAGAATGCGGCCGCAAGGAGGCCATCCATGTCC AACAATGGCTCGTCAC
ppsA-3RPAfl	ATAATCCTTAAGTTATTTCTTCAGTTCAGCCAG
tktA-5FPBgl	GAAGATCTAAGGAGATATACCATGTCCTCACGTAAA GAGCTTGCC
tktA-3RPPac	TCCTTAATTAATTACAGCAGTTCTTTTGCTTTCGC
Plac-tktA-5FPAfl	GCCTTAAGGTAGAGGATCGAGATCGTTTAGGC
aroE-5FPPac	TTAATTAAAAGGAGATATACCATGGAAACCTATGCT GTTTTTGGT
aroE-3RPNhe	CTAGCTAGCTCACGCGGACAATTCCTCCTGCAA
aroD-5FPBam	CGGGATCCAAGGAGATATACCATGAAAACCGTAACTGTA AAAGATC
aroD-3RPPst	TTCTGCAGTTATGCCTGGTGTAAAATAGTTAATACC
aroB-5FPPst	AACTGCAGAAGGAGATATACCATGGAGCGTATTGTC GTTACTCTGGGCGAACGTAGCTACCCAATTACCA
aroB-3RPNhe	CTAGCTAGCTTACGCTGATTGACAATCGGCAATG
Plac-5FPNhe	CTAGCTAGCGTAGAGGATCGAGATCGTTTAGGC
aroD-3RPNhe	CTAGCTAGCTTATGCCTGGTGTAAAATAGTTAATAC C
aro-5FPNco	ATACCATGGCACCTGTTACAATTGAAAAGTTCG
aro-3RPSac	CCGGAGCTCCTATTTTTTATTTCTTTTAAGTGCCG
UGT-5FPNco	ATACCATGGGCTCTGAAACTCGCCCGCTG
UGT-3RPBam	TCGGATCCTTAAACCTTCTTCAGCTTCAGTTC
TrcUGT-5FPBgl	GAAGATCTCCGACATCATAACGGTTCTG

Underlined parts indicated restriction sites.

Table S3 DNA and protein sequences of synthesized genes and proteins

Gene	DNA/Protein Sequence
TyrA$^{*\,syn}$ (DNA)	ATGGTTGCTGAATTGACCGCATTACGCGATCAAATTGATGAAGTCGATAAAGCGCT GCTGAATTTATTAGCGAAGCGTCTGGAACTGGTTGCTGAAGTGGGCGAGGTGAAA AGCCGCTTTGGACTGCCTATTTATGTTCCGGAGCGCGAGGCATCTATTTTGGCCTCG CGTCGTGCAGAGGCGGAAGCTCTGGGTGTACCGCCAGATCTGATTGAGGATGTTT TGCGTCGGGTGATGCGTGAATCTTACTCCAGTGAAAACGACAAAGGATTTAAAAC ACTTTGTCCGTCACTGCGTCCGGTGGTTATCGTCGGCGGTGGCGGTCAGATGGGA CGCCTGTTCGAGAAGATGCTGACCCTCTCGGGTTATCAGGTGCGGATTCTGGAGC AACATGACTGGGATCGAGCGGCTGATATTGTTGCCGATGCCGGAATGGTGATTGTT AGTGTGCCAATCCACGTTACTGAGCAAGTTATTGGCAAATTACCGCCTTTACCGAA AGATTGTATTCTGGTCGATCTGGCATCAGTGAAAAATGGGCCATTACAGGCCATGC TGGTGGCGCATGATGGTCCGGTGCTGGGGCTACACCCGATGTTCGGTCCGGACAG CGGTAGCCTGGCAAAGCAAGTTGTGGTCTGGTGTGATGGACGTAAACCGGAAGCA TACCAATGGTTTCTGGAGCAAATTCAGGTCTGGGGCGCTCGGCTGCATCGTATTAG CGCCGTCGAGCACGATCAGAATATGGCGTTTATTCAGGCACTGCGCCACTTTGCTA CTTTTGCTTACGGGCTGCACCTGGCAGAAGAAAATGTTCAGCTTGAGCAACTTCT GGCGCTCTCTTCGCCGATTTACCGCCTTGAGCTGGCGATGGTCGGGCGACTGTTTG CTCAGGATCCGCAGCTTTATGCCGACATCATTATGTCGTCAGAGCGTAATCTGGCGT TAATCAAACGTTACTATAAGCGTTTCGGCGAGGCGATTGAGTTGCTGGAGCAGGG CGATAAGCAGGCGTTTATTGACAGTTTCCGCAAGGTGGAGCACTGGTTCGGCGATT ACGTTCAGCGTTTTCAGAGTGAAAGCCGCGTGTTATTGCGTCAGGCGAATGACAA TCGCCAGTAA
TyrA$^{*\,syn}$ (Protein)	MVAELTALRDQIDEVDKALLNLLAKRLELVAEVGEVKSRFGLPIYVPEREASILASRR AEAEALGVPPDLIEDVLRRVMRESYSSENDKGFKTLCPSLRPVVIVGGGGQMGRLFE KMLTLSGYQVRILEQHDWDRAADIVADAGMVIVSVPIHVTEQVIGKLPPLPKDCILVD LASVKNGPLQAMLVAHDGPVLGLHPMFGPDSGSLAKQVVVWCDGRKPEAYQWFLE QIQVWGARLHRISAVEHDQNMAFIQALRHFATFAYGLHLAEENVQLEQLLALSSPIYR LELAMVGRLFAQDPQLYADIIMSSERNLALIKRYYKRFGEAIELLEQGDKQAFIDSFRK VEHWFGDYVQRFQSESRVLLRQANDNRQ
aroG^{*syn} (DNA)	ATGAATTATCAGAACGACGATTTACGCATCAAAGAAATCAAAGAGTTACTTCCTCC TGTCGCATTGCTGGAAAAATTCCCCGCTACTGAAAATGCCGCGAATACGGTTGCCC ATGCCCGAAAAGCGATCCATAAGATCCTGAAAGGTAATGATGATCGCCTGTTGGTT GTGATTGGCCCATGCTCAATTCATGATCCTGTCGCGGCAAAAGAGTATGCCACTCG CTTGCTGGCGCTGCGTGAAGAGCTGAAAGATGAGCTGGAAATCGTAATGCGCGTC TATTTTGAAAAGCCGCGTACCACGGTGGGCTGGAAAGGGCTGATTAACGATCCGC ATATGGATAATAGCTTCCAGATCAACGACGGTCTGCGTATAGCCCGTAAATTGCTGC TTGATATTAACGACAGCGGTCTGCCAGCGGCAGGTGAGTTTCTCAATATGATCACC CCACAATATCTCGCTGACCTGATGAGCTGGGGCGCAATTGGCGCACGTACCACCG AATCGCAGGTGCACCGCGAACTGGCATCAGGGCTTTCTTGTCCGGTCGGCTTCAA AAATGGCACCGACGGTACGATTAAAGTGGCTATCGATGCCATTAATGCCGCCGGTG CGCCGCACTGCTTCCTGTCCGTAACGAAATGGGGGCATTCGGCGATTGTGAATACC AGCGGTAACGGCGATTGCCATATCATTCTGCGCGGCGGTAAAGAGCCTAACTACAG CGCGAAGCACGTTGCTGAAGTGAAAGAAGGGCTGAACAAAGCAGGCCTGCCAGC

	ACAGGTGATGATCGATTTCAGCCATGCTAACTCGTCCAAACAATTCAAAAAGCAGA TGGATGTTTGTGCTGACGTTTGCCAGCAGATTGCCGGTGGCGAAAAGGCCATTATT GGCGTGATGGTGGAAAGCCATCTGGTGGAAGGCAATCAGAGCCTCGAGAGCGGG GAGCCGCTGGCCTACGGTAAGAGCATCACCGATGCCTGCATCGGCTGGGAAGATA CCGATGCTCTGTTACGTCAACTGGCGAATGCAGTAAAAGCGCGTCGCGGGTAA
aroG^{*syn} (protein)	MNYQNDDLRIKEIKELLPPVALLEKFPATENAANTVAHARKAIHKILKGNDDRLLVVI GPCSIHDPVAAKEYATRLLALREELKDELEIVMRVYFEKPRTTVGWKGLINDPHMDN SFQINDGLRIARKLLLDINDSGLPAAGEFLNMITPQYLADLMSWGAIGARTTESQVHR ELASGLSCPVGFKNGTDGTIKVAIDAINAAGAPHCFLSVTKWGHSAIVNTSGNGDCHI ILRGGKEPNYSAKHVAEVKEGLNKAGLPAQVMIDFSHANSSKQFKKQMDVCADVCQ QIAGGEKAIIGVMVESHLVEGNQSLESGEPLAYGKSITDACIGWEDTDALLRQLANAV KARRG
73B6 ^{syn} (DNA)	ATGGGCTCTGAAACTCGCCCGCTGAGCATCTTCTTTTTTCCGTTTATGGCGCATGGC CACATGATTCCGATGGTGGATATGGCACGTCTGTTTGCTTCTCAGGGTGTGCGTTG CACCATTGTTACCACTCCGGGTAACCAGCCGCTGATTGCTCGCTCTATCGGTAAGG TTCAGCTGCTGGGTTTTGAAATTGGTGTGACTACTATCCCGTTCCGCGGTACTGAG TTCGGCCTGCCGGATGGCTGTGAAAACCTGGATAGCGTGCCGAGCCCGCAGCATG TGTTTCATTTCTTTGAGGCAGCGGGTAGCCTGCGTGAGCCGTTTGAACAGCTGCTG GAAGAGCACAAACCGGACTGTGTTGTGGGCGATATGTTCTTTCCGTGGTCTACCG ACTCTGCGGCTAAATTCGGTATTCCGCGCCTGGTTTTCCACGGTACCTCCTACTTCG CGCTGTGCGCTGGCGAAGCAGTGCGTATTCATAAGCCGTACCTGTCTGTGTCTTCT GATGATGAACCGTTCGTTATTCCGGGCCTGCCGGACGAGATCAAACTGACCAAGT CCCAGCTGCCGATGCACCTGCTGGAGGGTAAGAAAGACTCTGTTCTGGCACAGCT GCTGGATGAGGTGAAAGAAACTGAGGTTTCCTCTTACGGTGTTATCGTTAACTCTA TCTACGAACTGGAACCGGCTTACGCAGATTACTTCCGTAACGTTCTGAAGCGCCGT GCGTGGGAGATCGGTCCGCTGTCTCTGTGTAACCGTGACGTTGAAGAGAAAGCGA TGCGTGGTATGCAGGCTGCTATCGATCAGCATGAATGCCTGAAATGGCTGGATTCC AAAGAACCGGATTCCGTTGTTTACGTTTGTTTTGGTAGCACTTGCAAATTCCCGGA TGATCAGCTGGCGGAAATCGCGTCTGGTCTGGAGGCAAGCGGCCAGCAGTTCATC TGGGTTATTCGCCGTATGTCTGACGACTCTAAGGAAGACTACCTGCCGAAAGGTTT CGAAGAGCGTGTTAAGGACCGTGCTCTGCTGATTCGCGGTTGGGCTCCGCAGGTT CTGATCCTGGACCATCAGTCCGTTGGCGGTTTTGTTTCTCACTGTGGTTGGAACTC TACCCTGGAAGGCATCAGCGCGGGTCTGCCGATGGTTACTTGGCCGGTGTTCGCT GAACAGTTCTACAACGAAAAACTGCTGACCGAGGTGCTGAAAATCGGTGTTGCA GTGGGTGCTCGTAAGTGGCGTCAGCTGGTGGGTGACTTCGTTCACAAAGACGCTA TTCAGCGTGCGGTGCGTGAAATTATGGAGGGCGAAGAGGCGGAGGAACGTCGTAT CATCGCGCGTCAGATGGGTAAAATGGCGAAACGCGCGGTGGAGAAGGACGGTAG CTCTTGGACCAACCTGAACAACCTGCTGCAGGAACTGAAGCTGAAGAAGGTTTA A
73B6 ^{syn} (protein)	MGSETRPLSIFFFPFMAHGHMIPMVDMARLFASQGVRCTIVTTPGNQPLIARSIGKVQ LLGFEIGVTTIPFRGTEFGLPDGCENLDSVPSPQHVFHFFEAAGSLREPFEQLLEEHKP DCVVGDMFFPWSTDSAAKFGIPRLVFHGTSYFALCAGEAVRIHKPYLSVSSDDEPFVI PGLPDEIKLTKSQLPMHLLEGKKDSVLAQLLDEVKETEVSSYGVIVNSIYELEPAYAD YFRNVLKRRAWEIGPLSLCNRDVEEKAMRGMQAAIDQHECLKWLDSKEPDSVVYV

	CFGSTCKFPDDQLAEIASGLEASGQQFIWVIRRMSDDSKEDYLPKGFEERVKDRALLI RGWAPQVLILDHQSVGGFVSHCGWNSTLEGISAGLPMVTWPVFAEQFYNEKLLTEV LKIGVAVGARKWRQLVGDFVHKDAIQRAVREIMEGEEAEERRIIARQMGKMAKRA VEKDGSSWTNLNNLLQELKLKKV