

This doc file includes:

Supplementary Table S1 to Table S4

Supplementary Figure S1 to Figure S8

Supplementary materials and methods

Supplementary Tables

Table S1. Strains and plasmids used in this study.

Plasmids or strains	Relevant characteristics
Plasmids	
pET-28a	pBR322 ori, Kan ^R , P _{T7} , 6*His, T _{T7}
Strains	
<i>E.coli</i> JM109	For plasmid clone
<i>E. coli</i> BL21(DE3)	For enzyme expression and purification

Table S2. The primers for gene amplification.

Name	Primer (5' – 3')
<i>Bm</i> -F	ATGGGTCGCGGATCCGAATTCATGAGTAACAAGTATAGAGTTAGAAAAACGT
<i>Bm</i> -R	GCAAGCTTGTCGACGGAGCTCTTATGAGGAACGTTTGTATTTCTTAA
<i>Ec1</i> -F	AGCAAATGGGTCGCGGATCCATGAGTAAGCTGATTGTCCC
<i>Ec1</i> -R	AGCAAATGGGTCGCGGATCCATGAGTAAGCTGATTGTCCC
<i>Ec2</i> -F	CAGCAAATGGGTCGCGGATCCATGAGTAAGCTGATTGTCCCGC
<i>Ec2</i> -R	TTGTCGACGGAGCTCGAATTCCTTATATCAGCCAAATGAACTTGCC
<i>Bs1</i> -F	GTGGACAGCAAATGGGTCGCGGATCCATGAATACATATCACCCATT
<i>Bs1</i> -R	CAAGCTTGTCGACGGAGCTCGAATTCCTTATGCCTCCTTCATTCCGA
<i>Bs2</i> -F	CTCGAGTGCGGCCGCAAGCTTTTATTTATGGGGATCAGTTATATCCATC
<i>Bs2</i> -R	CAGCAAATGGGTCGCGGATCCATGACACTTGAAAAATTTGTGGATG
<i>An1</i> -F	ATGGGTCGCGGATCCGAATTCATGCACTTGCATACTATCCT
<i>An1</i> -R	AGTGCGGCCGCAAGCTTTTAGATACCAGAATCATCCTCCTCAA
<i>An2</i> -F	AGCAAATGGGTCGCGGATCCATGACAATCTTTTGTACTCCTT
<i>An2</i> -R	TCGAGTGCGGCCGCAAGCTTTTATATACCCGACTCGTAT

Table S3. The primers for the saturated mutation.

Name	Primer (5' – 3')	Name	Primer (5' – 3')
E158D-F	ACCAAAGATGCGCCGACCCTG	V217D-F	GGTCAGATGGGCGATGTGCCG
E158D-R	CGGCGCATCTTTGGTCGCGCC	V217D-R	CGCGGTCGGCACATCGCCCAT
E157Q-F	ACCAAACAGGCGCCGACCCTG	V217L-F	GGTCAGATGGGCCTGGTGCCG
E157Q-R	CGGCGCCTGTTTGGTCGCGCC	V217L-R	CGCGGTCGGCACAGGCCCAT
F197W-F	CAGCTGGAAGGCTGGATTAAC	V217K-F	GGTCAGATGGGCAAAGTGCCG
F197W-R	CGGGCCGTTAATCCAGCCTTC	V217K-R	CGCGGTCGGCACTTTGCCCAT
P201A-F	TTTATTAACGGCGCGCAGCTG	V217A-F	GGTCAGATGGGCGCGGTGCCG
P201A-R	GTTATGCAGCTGCGCGCCGTT	V217A-R	CGCGGTCGGCACCGCGCCCAT
P201S-F	TTTATTAACGGCAGCCAGCTG	V217R-F	GGTCAGATGGGCCGCGTGCCG
P201S-R	GTTATGCAGCTGGCTGCCGTT	V217R-R	CGCGGTCGGCACGCGGCCCAT
R209A-F	AACCGCGTGCATGCGTGGGTG	V217E-F	GGTCAGATGGGCGAAGTGCCG
R209A-R	ACCGCCACCCACGCATGCAC	V217E-R	CGCGGTCGGCACTTCGCCCAT
R209N-F	AACCGCGTGCATAACTGGGTG	V218A-F	CAGATGGGCGTGCGCCGACC
R209N-R	ACCGCCACCCAGTTATGCAC	V218A-R	CGGCGCGGTCGGCGCCACGCC
R209S-F	AACCGCGTGCATAGCTGGGTG	V218L-F	CAGATGGGCGTGCTGCCGACC
R209S-R	ACCGCCACCCAGCTATGCAC	V218L-R	CGGCGCGGTCGGCAGCACGCC

Table S4. Data collection and refinement statistics.

Parameter	<i>Apo-BmTyr</i>
Unit cell dimensions	
Space group	P4321
a, b ,c (Å)	111.51, 111.51, 139.66
α , β , γ (Å)	90.00, 90.00, 90.00
Resolution (Å)	46.96-2.30
R _{merge}	0.167
R _{pim}	0.112
<i>I</i> / σ <i>I</i>	4.08
Completeness (%)	99.4
No. reflections	39392
R _{work} /R _{free}	0.190/0.237
No. atoms	4998
B factors (Å ²)	87.14
bond lengths (Å)	0.0375
bond angles (°)	1.4670

Supplementary Figures

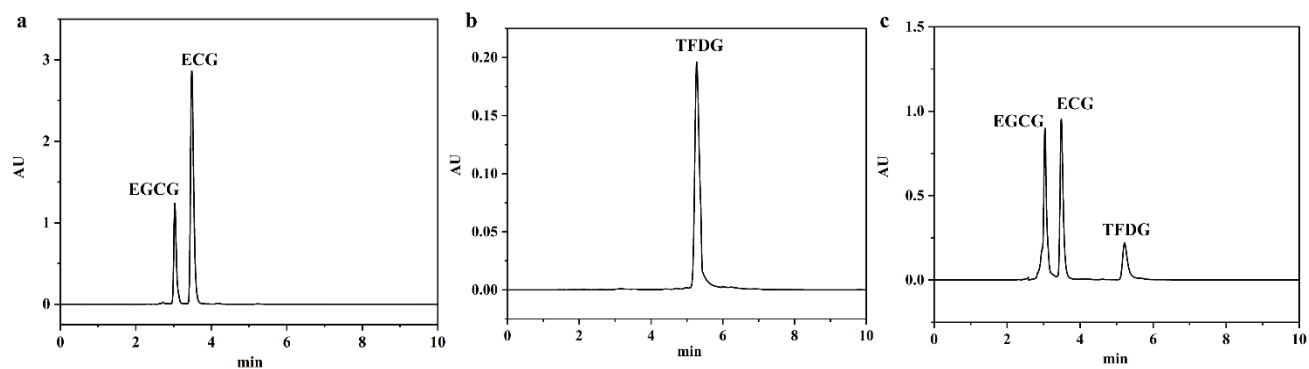


Figure S1. The HPLC chromatograms of TFDG reaction products. (A) The chromatographic peak of ECG, EGCG standard solution; (B) The chromatographic peak of TFDG standard solution; (C) Synthesis of TFDG from ECG and EGCG by *BmTyr*.



Figure S2. The whole structure of *BmTyr* (PDB ID:8HPI).

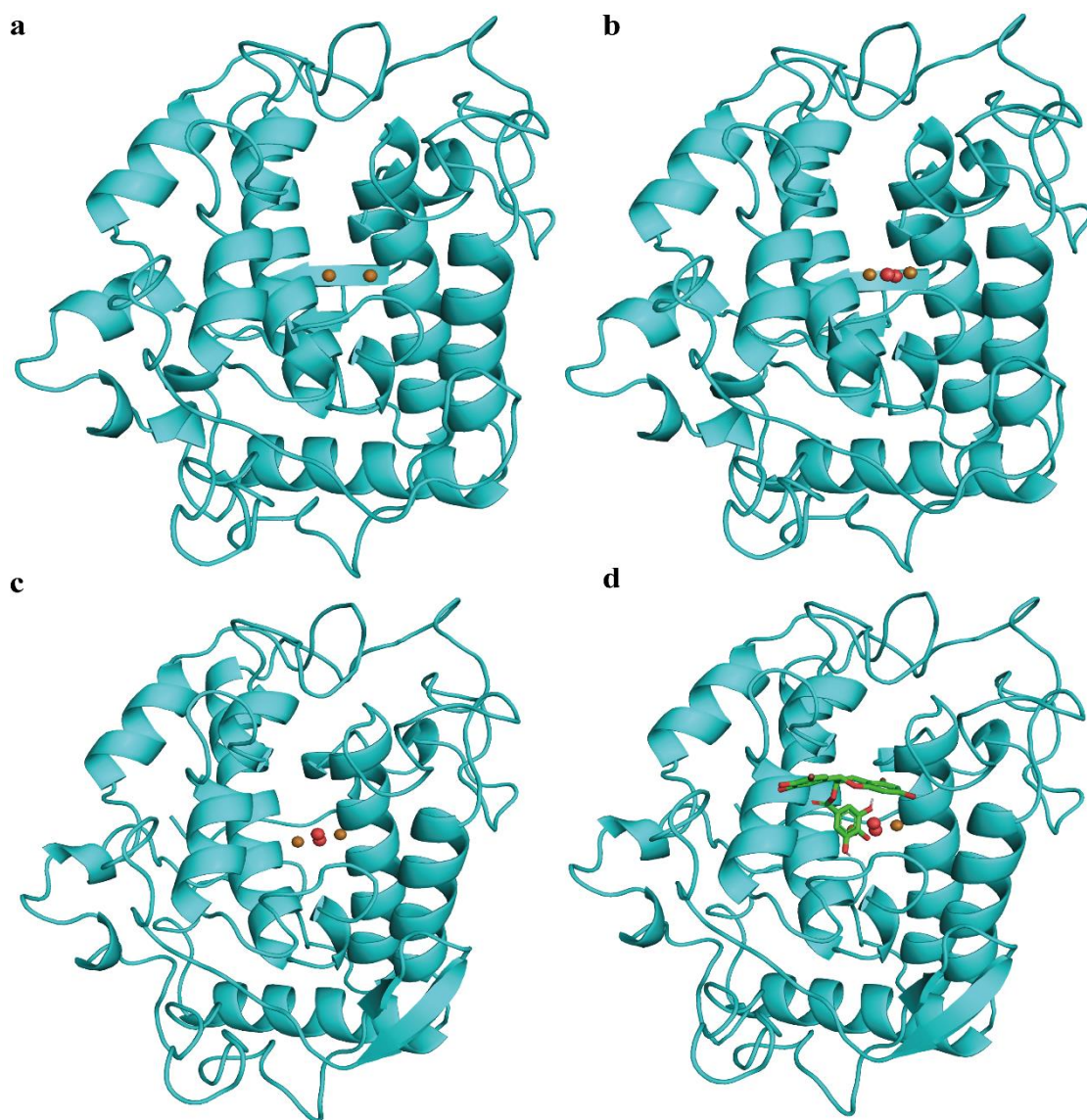


Figure S3. The construction process of the whole theoretical complex structure. (A) The whole structure of 3nm8; (B) The reactive binding mode M1 based on 3nm8; (C) The structure cluster analysis (M2); (D) The theoretical complex structure of BmTyr/Cu²⁺-EGCG (M3).

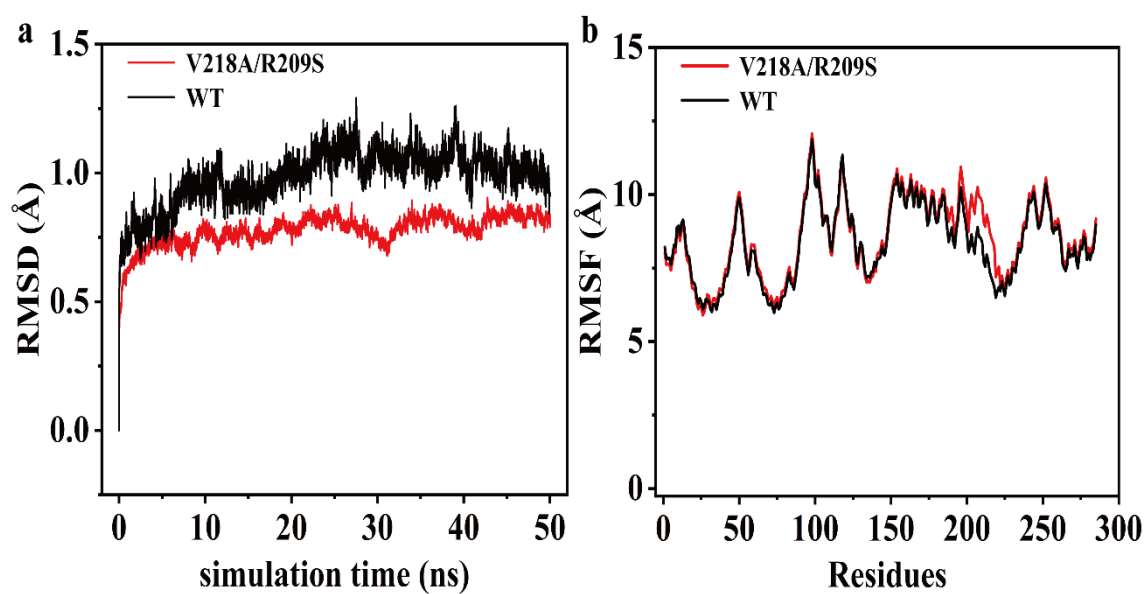


Figure S4. RMSD and RMSF of *BmTyr*^{WT} and *BmTyr*^{V218A/R209S}. (A) RMSD values calculated from MD simulations of *BmTyr*^{WT} and *BmTyr*^{V218A/R209S}; (B) RMSF values calculated from MD simulations of *BmTyr*^{WT} and *BmTyr*^{V218A/R209S}.

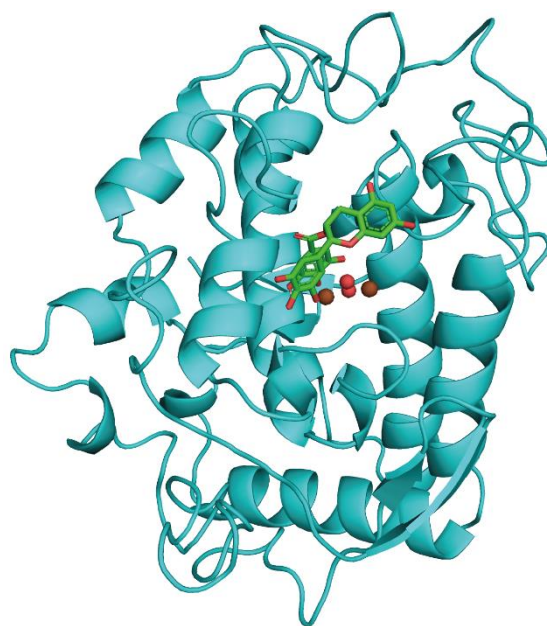


Figure S5. The whole TS2 state model.

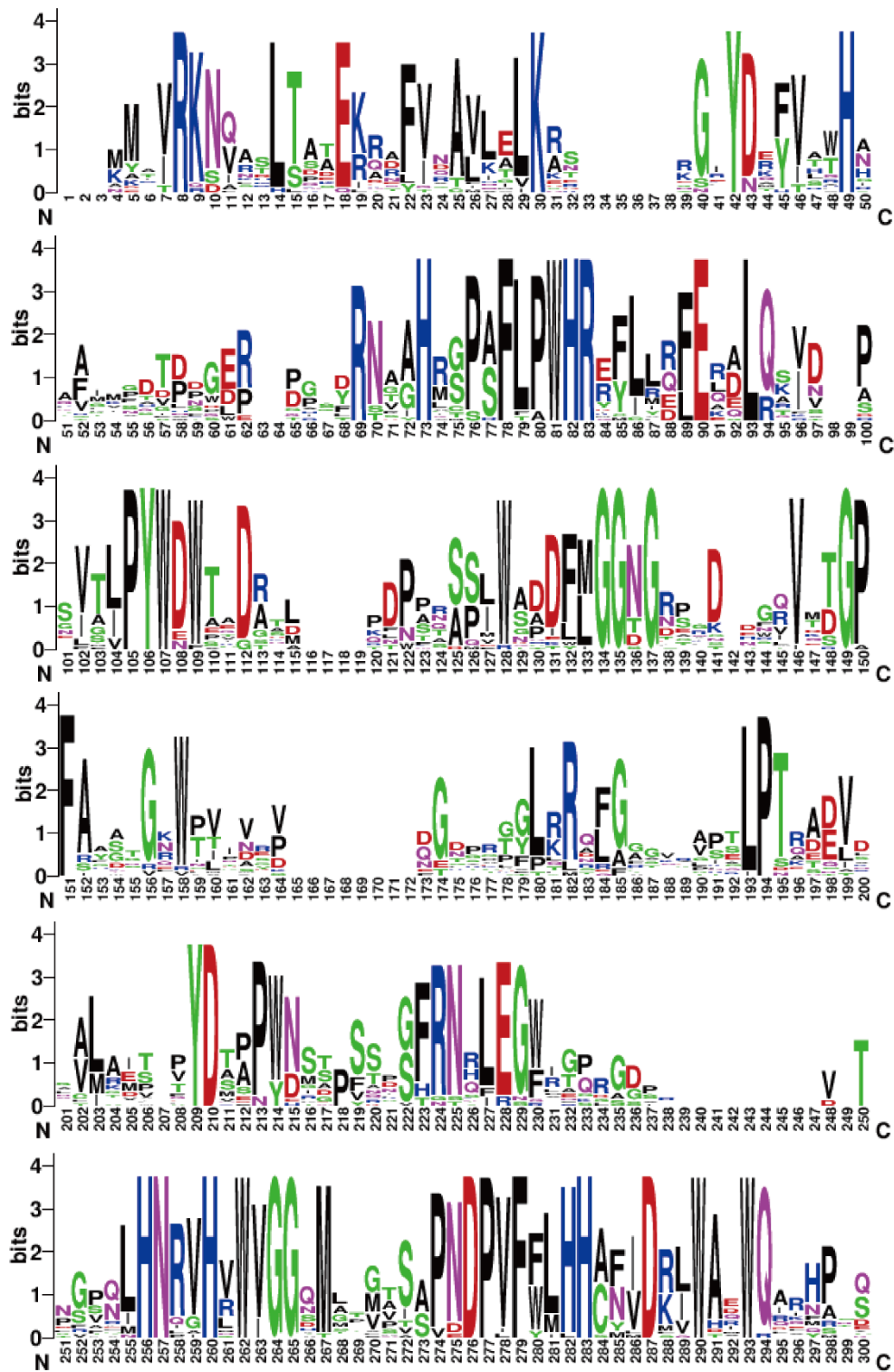


Figure S6. Multiple sequence alignment of *BmTyr* with homologous tyrosinase sequences from superfamily.

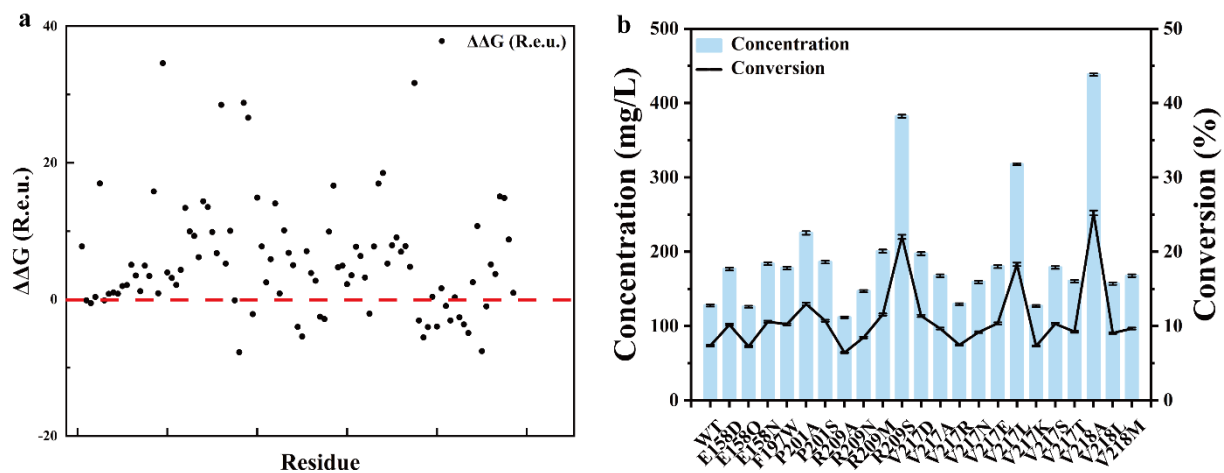


Figure S7. Redesign of *BmTyr*. (A) The energy difference between the wild type and the single point mutant was calculated ($\Delta\Delta G$). (B) The activity of the designed enzymes on the TFDG production (5 mM substrates in Mcilvaine Buffer (0.2 M, pH 5.0), 30 °C, 0.1 mM Cu²⁺). The data represent mean \pm s.d., as determined from three independent experiments.

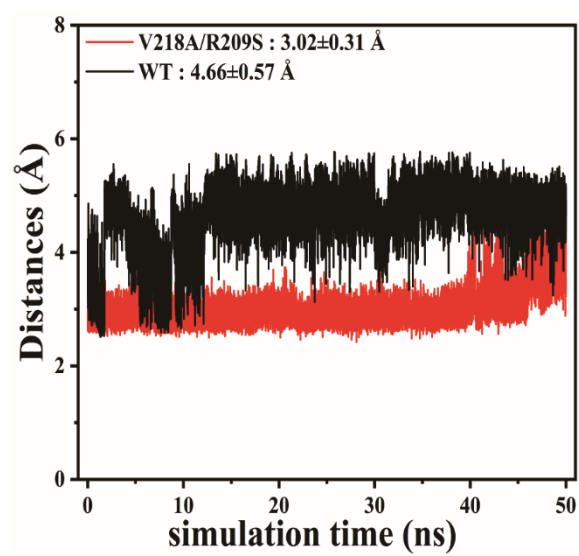


Figure S8. The distance between the -O11H of the hydroxy oxygen atom of EGCG and the O1 of the bridging oxygen atom of active centre.

Supplementary Material and Methods

Genetic constructions

The expression plasmid pET28a were purchased from Novagen (Madison, WI, U.S.A.). The host strain *E. coli* BL21(DE3) was purchased from Invitrogen (Carlsbad, CA, U.S.A.). All genetic constructions were carried out by using standard molecular biology techniques. All chemicals and reagents were pure and obtained commercially. Molecular biology reagents such as the restriction enzymes, PrimeSTAR, DNA polymerase and polymerase chain reactions (PCR) reagents were supplied by TaKaRa (Dalian, China). The cloning inserts were created via PCR of ORFs of interest from their respective genomic or synthesis by GENEWIZ (Suzhou, China). DNA and protein sequence.

The *Ec1* was synthesized and codon optimized for *E. coli*. The optimized DNA sequence is as follow:

```
ATGAGTAAGCTGATTGTCCCGCAGTGGCCGCAGCCAAAAGGTGTTGCGGCCTGTAGCTC
CACTCGTATCGGCGGCGTGAGCTTGCCCCGTATGACTCACTCAACCTCGGTGCCCATTG
TGCGGATAACCCGGATCACGTTGAGGAGAATCGCAAGCGACTTTTTGCTGCGGGCAATTT
GCCTTCTAAACCGGTCTGGCTTGAGCAGGTACACGGCAAAGATGTGCTTAAGCTCACTGG
CGAACCTTATGCCTCAAAACGGGCGGATGCCTCTTATAGCAATACGCCCGGCACGGTTTG
CGCAGTGATGACTGCCGACTGCCTCCCTGTGCTGTTTTGCAATCGAGCGGGAACGGAAG
TCGCCGCCGCTCATGCTGGCTGGCGTGGACTGTGCGCAGGCGTGCTGGAAGAGACGGTT
TCCTGTTTTGCTGATAATCCGGAAAATATTCTCGCCTGGTTAGGGCCGGCAATTGGTCCAC
GCGCGTTCGAAGTGGGGGGGGAGGTTTCGCGAGGCGTTTATGGCAGTAGACGCTAAAGCA
AGTGCAGCTTTCATTCAGCATGGTGATAAGTATCTGGCGGATATTTATCAGCTTGCCCGGC
AGCGTCTGGCGAACGTGGGTGTTGAGCAAATTTTCGGCGGCGACCGTTGTACATATACGG
AAAATGAGACTTTCTTCTCTTATCGTCGCGACAAGACCACCGGTCGTATGGCAAGTTTCA
```

TTTGGCTGATATAA

Protein sequence:

MSKLIVPQWPQPKGVAACSSSTRIGGVSLPPYDSLNLGAHCGDNPDHVEENRKRLFAAGNLPS
KPVWLEQVHGKDVLKLTGEPYASKRADASYSNTPGTVCAVMTADCLPVLFCNRAGTEVAA
AHAGWRGLCAGVLEETVSCFADNPENILAWLGPAIGPRAFEVGGGEVREAFMAVDAKASAAF
IQHGDKYLADIYQLARQRLANVGVEQIFGGDRCTYTENETFFSYRRDKTTGRMASFIWLI

The *Ec2* was synthesized and codon optimized for *E. coli*. The optimized DNA sequence is as follow:

ATGAGTAAGCTGATTGTCCCGCAGTGGCCGCTGCCAAAGGGTGTTGCGGCTTG TAGCTCT
ACTCGTATCGGCGGCGTGAGCTTGCCCCGTATGACTCGCTCAACCTCGGTGCCCATTTG
GGCGATAACCCGGATCACGTTGAGGAGAATCGCAAGCGACTTTTTGCTGCGGGCAATTTG
CCTTCTAAACCGGTCTGGCTTGAGCAGGTACACGGCAAAGATGTGCTTAAGCTTACTGGC
GAACCTTATGCCTCAAAACGAGCAGATGCCTCTTATAGCAATACGCCTGGCACGGTTTGC
GCAGTGATGACTGCCGACTGCCTCCCTGTGCTGTTTTGCAATCGAGCGGGAACGGAAGT
CGCCGCCGCTCATGCTGGCTGGCGTGGACTGTGCGCAGGCGTGCTGGAAGAGACAGTTT
CCTGTTTTGCTGATAATCCGGA AAAATATTCTCGCCTGGTTAGGGCCGGCAATTGGTCCACG
CGCGTTCGAAGTGGGGGCGGAGGTTCGCGAGGCGTTTATGGCAGCAGACGCCAAAGCG
AGTACAGCTTTCATTCAGCATGGTGATAAGTATCTGGCGGATATTTATCTGCTTGCCCGGC
AGCGTCTGGCGAGCGTTGGTGTTGAACAAATTTTCGGCGGCGACCGTTGTACATATACGG
AAAATGAGACTTTCTTCTCTTATCGTCGCGACAAGACCACGGGCGTGATGGCAAGTTTCA
TTTGGCTGATATAA

Protein sequence:

MSKLIVPQWPLPKGVAACSSSTRIGGVSLPPYDSLNLGAHCGDNPDHVEENRKRLFAAGNLPS

KPVWLEQVHGKDVCLKLTGEPYASKRADASYSNTPGTVCAVMTADCLPVLFCNRAGTEVAA
AHAGWRGLCAGVLEETVSCFADNPENILAWLGPAIGPRAFEVGAEVREAFMAADAKASTAF
IQHGDKYLADIYLLARQLASVGVEQIFGGDRCTYTENETFFSYRRDKTTGRMASFIWLI

The *Bm* was synthesized and codon optimized for *E. coli*. The optimized DNA sequence is as follow:

ATGAGCAACAAATATCGCGTGCGCAAAAACGTGCTGCATCTGACCGATACCGAAAAACG
CGATTTTGTGCGCACCGTGCTGATTCTGAAAGAAAAAGGCATTTATGATCGCTATATTGCG
TGGCATGGCGCGGGCGGGCAAATTCATACCCCGCCGGGCAGCGATCGCAACGCGGGCGCA
TATGAGCAGCGCGTTTCTGCCGTGGCATCGCGAATATCTGCTGCGCTTTGAACGCGATCTG
CAGAGCATTAAACCCGGAAGTGACCCTGCCGTATTGGGAATGGGAAACCGATGCGCAGAT
GCAAGATCCGAGTCAGAGTCAGATTTGGAGCGCGGATTTTATGGGCGGCAACGGCAACC
CGATTAAAGATTTTATTGTGGATACCGGCCCCGTTTGCGGCGGGCCGCTGGACCACCATTG
ATGAACAAGGCAACCCGAGCGGCGGCCTGAAACGCAACTTTGGCGCGACCAAAGAAGC
GCCGACCCTGCCGACCCGCGATGATGTGCTGAACGCGCTGAAAATTACGCAGTATGATAC
CCCGCCGTGGGATATGACGAGTCAGAACAGCTTTCGCAATCAGCTGGAAGGCTTTATTAA
CGGCCCCGCAGCTGCATAACCGCGTGCATCGCTGGGTGGGCGGTCAGATGGGCGTGGTGC
CGACCGCGCCGAACGATCCGGTGTTTTTCTGCATCATGCGAACGTGGATCGCATTGTTGGG
CGGTGTGGCAGATTATTCATCGCAATCAGAACTATCAGCCGATGAAAAACGGCCCCGTTTG
GTCAGAACTTTCGCGATCCGATGTATCCGTGGAACACCACCCCGGAAGATGTGATGAACC
ATCGCAAACCTGGGCTATGTGTATGATATTGAACTGCGCAAAAGCAAACGCAGCAGCTAA

Protein sequence:

MSNKYKVRKNVLHLTDTEKRDFIRAVLILKEKGIYDRYIAWHGAAGKFHTPPGSDRNAAHM
SSAFLPWHREYLLRFERDLQSINPEVTLPYWETDAQMQDPSQSQIWSADFMGGNGNPK

KDFIVDTGPFAAGRWTIDEQGNPSGGLKRNFGATKEAPTLPTRDDVLNALKITQYDTPPWD
MTSQNSFRNQLEGFINGPQLHNRVHRWVGGQMGVVPTAPNDPVFFLHHANVDRIWAVWQI
VHRNQNYQPMKNGPFGQNFRDSMYPWDTPKDVMDHRKLGYYVDIELRKSKRSS

The *An1* was synthesized and codon optimized for *E. coli*. The optimized DNA sequence is as follow:

ATGACAATCTTTTTGCTACTCCTTGGGCTTCTAGCCCCACTGGTCCTCTGCAGTCCAATCA
GACACGTTACTCCGCATCTTTCCCACTCTACCCCAATACCGAGAGATATGCAAGGCAATTC
TTCGCAGTCACCCAATACCCCATGGCAGGGATACGATATAAACACGAAC TACTATGAAAC
CATCCACAGACCAATGTTGTCCGCGAATACTGGTTTGATATCGTCAACACGACCGCTGC
TCTTGATGGAGTTGAGAGACCAGTGCTTCTGGTCAACGGCCAGTTTCCAGGACCTACTAT
TGAGGCAAAC TGGGGTGACACCGTTAAAGTTCATGTTACGAATCGCATGGAAAACAATG
GCACGGCCATCCACTTCCACGGAATTCGACAGCTCTACAACAACCAAATGGATGGCGTAG
CAGCGTTGACTCAGTGCCCTGTTCTCCGAATTCCAGTTATACCTATGTCTGGCGAGCGGA
AGAATATGGGTCGAGTTGGTATCACTCACACTTCTCTCTCCAGGCATGGGAAGGGGTTTT
CGGAGGCATTTTGATTCACGGCCCTTCAACAGCGGAATATGACCACGACTTGGGAATGGT
CTTTCTCAATGACTGGTCGCATCAAACAGTCGACGAGATGTATCAATCTGTACTCGAGAG
TCAA AATCCGCCGCATTTTCAGACAGGCCTTATCAACGGAAGCAATATCTGGGTCACTGC
AGATAACCAAACCGTTGGTCGCAGATTCCAAACAGAGTTTGTACCAGGCCAGAGATACC
GGCTTCGGTTGGTCAATGCCGCCATGCATACCCATTTTAGGTTCTCAATCGATAATCATGA
CCTTACCGTTATTGCCAGTGATTTTGTTCTATTGTACCATTACGACGAACAACGTTCCC
ATCGGCATGGGACAACGGTATGATATTATCGTGACAGCAAACCAAGCTCCGGATAACTATT
GGATTCGTGCAATCCCACAGTCATTCTGCAGTGACAACGCGAATTCAGACAATATTAAAG
GAGTCCTCCATTACGAAGGTGCTGCCGATAACAGCGATCCCACGAGCACAAAGTGGGAT

TATGGTGACGACATTCAGTGCCTGGATTTCTCATTAGATGAGCTTGTCCCCTGGCTTGCCC
TTGATGCAGATATTGGCGGTGCCCAGATGGCCGAGTCAGACGTGGACTTCACTCCGTTCG
GGGATGTTCTTTGTATTTGTGGACAATGGGAGGCAATGCCCTTAACATCTCGTGGAAG
ATCCACGTTACAGCAAACGTTCTGAAGATCCCGACAAAATGGACTGGAAAGCAAGCCAG
GGCGTGATAGAGGCTGCGATCCCGAATAAATGGACCGTGTTGGTTGTCCAGACAGATCTT
CCTGTTCTCACCTATTCATTTGCATGGGCATGATTTCTATCTCCTTGCTCAGGGATTCTGG
CCAATTCAATCCACAGAACGTCACGCTCAAAACGCATAACCCTCCCCGCCGCGACACTGC
GCTCATGACCGCTGCCACTCCTGAAAATGGTGGCGGGCGGCTATATGGTTATTGGCTTCCCC
GCCGACAACCCGGGTGTATGGCTGATCCACTGCCACATTGGATTCCATGCAACAGAAGGA
TTCGCGCAGCAGATTGTCGAGCGACAGAGTGAGTTCAACACGTTCTTCAGCGAAGACCT
GCTGGAGAACACTTGTGATGCGTGGGATGAGTATGCCAAGGTGAATCCCTACGGACATCA
GTACCGGGCTCTCGCTGGGCCATACGAGTCGGGTATATAA

Protein sequence:

MTIFLLLGLLAPLVLCSPIRHVTPHLSHSTPIPRDMQGNSSQSPNTPWQGYDINTNYETIPQ
TNVVREYWFDIVNTTAALDGVERPVLLVNGQFPGPTIEANWGDTVKVHVITNRMENNGTAI
HFHGIRQLYNNQMDGVAALTQCPVPPNSSYTYVWRAEEYGSSWYHSHFSLQAWEGVFGGIL
IHGPSTAEYDHDLMVFLNDWSHQTVDEMYQSVLESQNPPHFQTGLINGSNIWVTADNQTV
GRRFQTEFVPGQRYRLRLVNAAMHTHFRFSIDNHDLTVIASDFVPIVPFTTNNVPIGMGQRYD
IIVTANQAPDNYWIRAIPQSFCSDNANSJNIKGVLYHYEGAADNSDPTSTKWDYGDDIQCLDF
SLDELVPWLALDADIGGAQMAESDVDFTPFGDVPLYLWTMGGNALNISWKDPTLQQTFEDP
DKMDWKASQGVIEAAIPNKWTVLVVQTDLPVPHPIHLHGHDYLLAQGFQGNPQNVTLK
THNPPRRDTALMTAATPENGGGGYMVGFPADNPGVWLIHCHIGFHATEGFAQQIVERQSEF

NTFFSEDLLENTCDAWDEYAKVNPYGHQYRALAGPYESGI

The *An2* was synthesized and codon optimized for *E. coli*. The optimized DNA sequence is as follow:

ATGCACTTGCATACTATCCTGGCTGCGGGCTGGCTCGCAGCCAGCGCCTATGCCGGGGTT
TCCTCTCGTTCCACGCAAGCATGCGCAGGAAACACACCGCACACCCGTGGCCAGTGGTG
TGATTTCGACATTCACACCGACTACTATGACGAAGCTCCGGACACTGGGGTAGTTCGCGA
ATATTGGTTGGAAGTAGGAAACACCACCGCCGCTCCCGACGGCGTCCCTCGTAATGTCCT
AACCGTCAACGGGACGACTCCAGGGCCTACTCTATACGCCGATTGGGGTGACTGGGTCC
GAGTTCATGTATACAATGGTCTTGAAAACAATGGAACCAGCATCCATTGGCACGGTGTCC
GCCAGAACCACACCAACCCTCAAGACGGCACAACTCCATTACCCAATGTCCCACGCCC
CCCGGAGGATCCGTCACGTACGAATGGCGCGCCACGCAGTACGGCACAACTACTACCA
CTCCCACCTTTGCTCTGCAAGCATGGGAGGGCGTCTACGGTGGTATTGTGATCAACGGACC
TGCCAGCTCGAATTACGACGTCGACGCTGGCGTGCTTTATCTCATGGACTGGTCGCACCG
CACAGCCGACCAGATGTATTCCACCGCTGAAACGGACGGTGAACCAGAGATGACCACGG
GTTTGATTAATGGAACGAACATCTGGGTTAAGGAGAACAACGAAACGGTCGGTTCGCGC
TTCCTTATGAATGTGACTGCTGGCCAGTCTTACCGTCTTCGACTTATCAATGGTGGAATGA
GCAACTTGTTCCGATTCATGATCGATGATCATACTTGACGGTTATTGCTGCGGACTTTGT
CCCCATCGAGCCTTACAATACGACTTCGCTTAATATTGGTGTGGACAACGGTACGATGTG
ATCGTCACGGCAAACCAAGCCCAGGCTGGGTCTGACTTCTGGATGCGTGCTATCCCCCAG
GAGTCCTGTGCCGAGATTCAAAACACCGACATCAAGGGAATTTCCATTATGAGCAGCCT
GACACTGCTGTAGCGACGCCCACAACAACAGCATACAACCTATAACCGACAGCTGCTACGA
CGAGCCATTGAGTAGCTTGGTGCCGGTAGTTGCCATCGACGCGTCTGGCGTCGGCTACAA
AAGTCATAACGAGGCGAACGTCATCATGAACAGTGCAGGCCTGTACAAGTGGTACATGG

GCCCGACCACGTTCAAGGCAGAATGGGACAACCCGACTTTGCTCCAGATCGCTAACGGC
AACACTTCATGGACTACACATCGCACGTTGTTGAAGTTGAAGGAAACGGCAAGTGGGTG
ATTGTTGACATTGAAATGGAAATCAATGTTCTCATCCTATTCATCTCCACGGCCACGACT
TCTTCATCCTCGCCCAAGGCAAGGGAAGCTACACGTCGAACACGTCTCTGAACACTTAC
AACCACCCCCGTCGGGACACGGCCATGCTCCCTGCCCTCGGATATCTGGTCATCGCATTTA
AGACCGACAACCCTGGCGCGTGGCTTCTGCATTGCCATATTGGATGGCACCAATCGGAGG
GCTTCGCCATGCAGTTCGTGGAGAGCATCTCTCAGCTAGAGCCTATGATTGATGCTTCCGC
ACTCGAAAAGAACTGTGCTGCCTGGGATAAATATGCTGCTGCTAACCGCATTGAGGAGGA
TGATTCTGGTATCTAA

Protein sequence:

MHLHTILAAGWLAASAYAGVSSRSTQACAGNTPHTRGQWCDFDIHTDYYDEAPDTGVVRE
YWLELGNTTAAPDGVPRNVLTVNGTTPGPTLYADWGDWVRVHVYNGLENNGTSHWHGV
RQNHTNPQDGTNSITQCPTPPGGSVTYEWRAATQYGTITYYHSHFALQAWEGVYGGIVINGPA
SSNYDVDAGVLYLMDWSHRTADQMYSTAETDGEPEMTTGLINGTNIWVKENNETVGSRFL
MNVTAGQSYRLRLINGGMSNLFRFMIDDHTLTVIAADFVPIEPYNTTSLNIGVGQRYDVIVTA
NQAQAGSDFWMRAIPQESCAEIQNTDIKGIFHYEQPDTAVATPTTTAYNYTDSCYDEPLSSLV
PVVAIDASGVGYKSHNEANVIMNSAGLYKWYMGPTTFKAEDNPTLLQIANGNTSWTNTS
HVVEVEGNGKWVIVDIEMEINVPHPIHLHGHDFFILAQGKGSYTSNTSLNTYNPPRRDTAML
PALGYLVIAFKTDNPGAWLLHCHIGWHQSEGFAMQFVESISQLEPMIDASALEKNCAAWDK
YAAANRIEEDDSGI

The *Bs1* was synthesized and codon optimized for *E. coli*. The optimized DNA sequence is as follow:

ATGACACTTGAAAAATTTGTGGATGCTCTCCCAATCCCAGATACACTAAAGCCAGTACAG

CAATCAAAAGAAAAACATACTACGAAGTCACCATGGAGGAATGCACTCATCAGCTCCAT
CGCGATCTCCCTCCAACCCGCCTGTGGGGCTACAACGGCTTATTTCCGGGACTGACCATT
GAGGTAAAAGAAATGAAAACGTATATGTAAAATGGATGAATAACCTTCCTTCCACGCATT
TCCTTCCGATTGATCACACCATTTCATCACAGTGACAGCCAGCATGAAGAGCCCGAGGTAA
AGACTGTTGTTCAATTACACGGCGGGCTCACGCCAGATGATAGTGACGGGTATCCGGAGG
CTTGGTTTTCCAAAGACTTTGAACAAACAGGACCTTATTTCAAAGAGAGGTTTTTCATT
ATCCAAACCAGCAGCGCGGGGCTATATTGTGGTATCACGATCGCGCCATGGCGCTCACCA
GGCTAAATGTCTATGCCGGACTTGTGCGGTGCATATATCATTCATGACCCAAAGGAAAAACG
CTTAAACTGCCTTCAGACGAATACGATGTGCCGCTTCTTATCACAGACCGCACGACCAA
TGAGGATGGTTCTTTGTTTTATCCGAGCGCACCGGAAAACCCTTCTCCGTCCTGCCTAAT
CCTTCAATCGTTCCGGCTTTTTGCGGAGAAACCATACTCGTCAACGGGAAGGTATGGCCA
TACTTGGAAGTCGAGCCAAGGAAATACCGATTCCGTGTCATCAACGCCTCCAATACAAGA
ACCTATAACCTGTCACTCGATAATGGCGGAGATTTTATTCAGATTGGTTCAGATGGAGGGC
TCCTGCCGCGATCTGTAAACTGAATTCTTTCAGCCTTGCGCCTGCTGAACGTTACGATAT
CATCATTGACTTCACAGCATATGAAGGAGAATCGATCATTTTGGCAAACAGCGCGGGCTG
CGGCGGTGACGTCAATCCTGAAACAGATGCGAATATCATGCAATTCAGAGTCACAAAACC
ATTGGCACAAAAGACGAAAGCAGAAAGCCGAAGTACCTCGCCTCATACCCTTCGGTAC
AGCATGAAAGAATACAAAACATCAGAACGTTAAAACTGGCAGGCACCCAGGACGAATAC
GGCAGACCCGTCCTTCTGCTTAATAACAAACGCTGGCACGATCCCGTCACAGAAACACC
AAAAGTCGGCACAACTGAAATATGGTCCATTATCAACCCGACACGCGGAACACATCCGAT
CCACCTGCATCTAGTCTCCTTCCGTGTATTAGACCGGCGGCCGTTTGATATCGCCCGTTATC
AAGAAAGCGGGGAATTGTCCTATACCGGTCCGGCCGTCCCGCCGCCGCAAGTGAAAAG

GGCTGGAAAGACACCATTC AAGCGCATGCAGGTGAAGTCCTGAGAATCGCGGCGACATT
CGGTCCGTACAGCGGACGATACGTATGGCATTGCCATATTCTAGAGCATGAAGACTATGAC
ATGATGAGACCGATGGATATAACTGATCCCCATAAATAA

Protein sequence:

MTLEKFVDALPIPD TLKPVQQSKEKTY YEVTMEECTHQLHRDLPPTRLWGYNGLFPGLTIEV
KRNENVYVKWMNNLPSTHFLPIDHTIHHSDSQHEEPEVKTVVHLHGGVTPDDSDGYPEAWF
SKDFEQTGPFYFKREVFHYPNQQRGAILWYHDRAMALTRLNVYAGLVGAYIIHDPKEKRLKL
PSDEYDVPLLITDR TTNEDGSLFYPSAPENPSPSLPNPSIVPAFCGETILVNGKVWPYLEVEPR
KYRFRVINASNTRTYNLSLDNNGGDFIQIGSDGGLPRSVKLNSFSLAPAERYDIIIDFTAYEGES
IILANSAGCGGDVNPETDANIMQFRVTKPLAQKDESRKPKYLASYPSVQHERIQNIRTLKLAG
TQDEYGRPVL LLNNKRWHDPVTETPKVGTTEIWSIINPTRGTHPIHLHLVSFRVLDRRPF DIA
RYQESGELSYTGPAVPPPPSEKGWKDTIQAHAGEVLRIAATFGPYSGRYVWHCHILEHEDYD
MMRPM DITDPHK

The *Bs2* was synthesized and codon optimized for *E. coli*. The optimized DNA sequence is as follow:

ATGAATACATATCACC CATTCAGTCTTACCACACCCTCGACACTCATGATACAAGACTGGG
CTCAAACGAATCAAAACAACAAAGAGGTCATTGCCGGATTTACGACAAAAAACGGCGGT
GTCAGCCAAAAGCCTTTTGAATCGTTAAATACAGGATTGCACGTT CATGACAAAGATGCA
GATGTAGTTAAAAATCGTGAATATATTGCCGATATGTTTAATACTGATTTGCAGTCTTGGGT
ATTCGCTGATCAGACACATGATAATCGCGTTCAGAAAGTGACGCAGAGGGATAGGGGAA
AAGGCGCCCGTGAGTATCACACGGCTCTAAAAGCAACGGACGGGCTCTATACAAATGAA
AAAAATGTATTTTTTAGCATTATGCTTTGCTGATTGTGTGCCTCTTTTCTTTTATGATCCGGTT
AAGTCGCTTGTCGGAGTCGCCCATGCCGGGTGGAAAGGCACCGTCAAACAGATTGGCAG

AGAAATGGTGAAGCAATGGACTGAGAAGGAAGGTTCAAATCTCTCAGATATTTACGCTGT
TATTGGCCCGTCTATCAGCGGAGCATGCTATACGGTAGACGACCGCGTCATGGATGCTGTT
CGTGCATTGCCGGTTTCAGCAGACCTTGCCGCCAATCAGACGGCAAAGGCACAATATCAG
CTTGATCTGAAAGAGCTGAACCGTCTTATACTGATGGACAGCGGTTTGGCAAGTGAACAA
ATTTCTGTCAGCGGTTTATGCACGGAAAGCGAGCCGTCTCTTTTCTATTCTCATCGCCGCG
ATCAGGGGAAAACCTGGACGGATGATGTCCTTTATCGGAATGAAGGAGGCATAA

Protein sequence:

MNTYHPFSLTTPSTLMIQDWAQTNQNNKEVIAGFTTKNGGVSQKPFESLNTGLHVHDKDAD
VVKNREYIADMFNLDLQSWVFADQTHDNRVQKVTQRDRGKGAREYHTALKATDGLYTNE
KNVFLALCFADCVPPLFFYDPVKSLVGVAHAGWKGTVKQIGREMVKQWTEKEGSNLSDIYAV
IGPSISGACYTVDDRVM DAVRALPVSADLAANQTAKAQYQLDLKELNRLILMDSGLASEQIS
VSGLCTESEPSLFYSHRRDQGKTGRMMSFIGMKEA