PROFICS: A bacterial selection system for directed evolution of proteases

Christina Kröß1,2,§, Petra Engele1,2,§, Bernhard Sprenger1,2, Andreas Fischer1,3, Nico Lingg1,3, Magdalena Baier2, Christoph Öhlknecht1,2, Bettina Lier1,4, Chris Oostenbrink1,4, Monika Cserjan-Puschmann1,3, Gerald Striedner1,3, Alois Jungbauer1,3, Rainer Schneider1,2

List of material included

Additional experimental procedures for detection of protein expression in bacterial cells with Npro, cp caspase-2 and Mpro

- Figure S1 Expression control of cells containing cp ATCase and Npro fusion proteins
- Figure S2 Expression control of cells containing cp ATCase fusion variants and cp caspase-2
- Figure S3 Expression control of cells containing cp ATCase fusion variants and Mpro

Table S1 Selected cp caspase-2 mutants
Table S2 FRET assay results for cp caspase-2 S9 variant compared to cp caspase-2 (9) and in silico variant cp caspase-2 D323T H226A (25). Data provided in separate .xlsx format
Supporting materials and methods for P1’ ATCase selections

- Table S3 Activity of cp ATCase with different N-terminal amino acids
- Table S4 Sequences of used oligonucleotides for PCR
- Nucleotide and amino acid sequences of used constructs
- Table S5 Buffer and media compositions

Expression controls and cleavage of proteases in PROFICS selection system

Western blot (Npro-cp ATCase)

The proteins were separated by SDS-PAGE (NuPAGE 4-12 % Bis-Tris gel, Invitrogen, Thermo Fisher Scientific) and blotted onto a methanol activated PVDF membrane applying 30 V constant for 1 h. Blocking was executed in TBS-T (20 mM Tris Base, 150 mM NaCl, 0.05 % Tween, pH 7.6) with 5 % milk powder for one hour. Incubation with primary antibody (1:10,000 dilution of α-Npro antibody, produced in our laboratory, in blocking solution) was at 4 °C overnight. After washing with TBS-T, the membrane was incubated with secondary antibody (α-Goat IgG, Santa Cruz Biotechnology, Inc.; Dallas, TX, USA), diluted 1:4,000 in blocking solution, for 1 h. For detection SuperSignal® West Pico Luminol/Enhancer Solution and SuperSignal® West Pico Stable Peroxide Solution (Thermo Fisher Scientific) and the digital imaging system FUSION FX7 from Peqlab (VWR International; Radnor, PA, USA) were used.
**Figure S1:** SDS-PAGE and Western blot with antibody against N\(^{pro}\) of *E. coli* pyr- cells containing a fusion of active N\(^{pro}\) and cp ATCase (Npro-ATCase) or a fusion of inactive N\(^{pro}\) and cp ATCase (iNpro-ATCase). Samples were taken before and after induction of expression. Fusion proteins and the ATCase regulatory chain (R-chain) are expressed in both cell types. Only in the cells with active N\(^{pro}\) the fusion is cleaved and the ATCase catalytic chain (C-chain) and N\(^{pro}\) were detected separately.

**Figure S2:** SDS-PAGE of *E. coli* pyr- cells containing cp ATCase with different recognition sites (WEHD and VDVAD) in their fusion tags; cp ATCase without fusion and cp caspase-2 without cp ATCase were used as controls. Samples were taken before and after induction of expression. The regulatory chain is expressed in all ATCase containing cells. In addition, the cleaved cp c-chain is visible in the cells containing cp caspase-2 together with 6H-VDVAD-cp ATCase. The uncleaved (tagged) cp c-chain is visible in the cells without cp caspase-2 and in the cells with the WEHD recognition sequence.
Selected variants from caspase mutation

On selective plates colonies with varying sizes were obtained. Colonies larger in size, which were suspected to contain a caspase variant conferring an advantage over others because of the cells’ faster growth were analyzed in more detail.

Table S1: Mutations of variants selected from cp caspase-2 library. The library was transformed in E. coli pyr- cells containing tagged cp ATCase. Total volume of the transformation was 500 µl. Aliquots of 250 and 25 µl were pipetted into flasks with selective media (approach Fa and Fb), aliquots of 200 (Pa), and 20 µl (Pb) were plated on selective agar plates. 1 µl was plated on non-selective TY-agar and colonies sequenced to estimate the size of the library and the number of mutations.

<table>
<thead>
<tr>
<th>No.</th>
<th>Mutations</th>
<th>Silent mutations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colonies from non-selective conditions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>none</td>
<td>0</td>
</tr>
<tr>
<td>II</td>
<td>C44R T145A V238F</td>
<td>0</td>
</tr>
<tr>
<td>III</td>
<td>K89 F100S H116Q L194Q F217L</td>
<td>0</td>
</tr>
<tr>
<td>Agar plate assay Pa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>L45Q K136R</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>E105V</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>none</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>none</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>T126S</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>R35S Q144R</td>
<td>2</td>
</tr>
<tr>
<td>Agar plate assay Pb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>K26R</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>none</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>none</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>none</td>
<td>2</td>
</tr>
<tr>
<td>Flask assay Fa</td>
<td>7</td>
<td>E105V</td>
</tr>
<tr>
<td>---------------</td>
<td>----</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>E105V</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>E105V</td>
</tr>
<tr>
<td>Flask assay Fb</td>
<td>8</td>
<td>F147L</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>L149R V201A</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>none</td>
</tr>
</tbody>
</table>

FRET Results

Table S2: Michaelis Menten kinetics measured using FRET substrate VDVADXA at 1 µM enzyme concentration. (n.d. not determined), cp caspase-2 S9 variant compared to cp caspase-2 (data from (9)) and in silico variant cp caspase-2 D323T H226A (data from (25))

Data provided in separate .xlsx file

Additional experiments for P1’ selections

Construction of cp ATCase plasmids

For the cp MA-ATCase construct an alanine was inserted at the N-terminus of cp pyrB by site directed mutagenesis using primers MA-pyrB_forw and MA-pyrB_rev.

The original Thr\(^{228}\) was mutated to all 19 canonical amino acids with site directed mutagenesis using degenerate primers M-X-ATCase_forw and M-X-ATCase_rev, resulting in cp M-(T→X)-ATCase constructs.

For the deletion variant the Met\(^{227}\) in the cp pyrB was removed with site directed mutagenesis.

A gene library of N\(^{pro}\)-(M→X)-cp ATCase constructs was cloned with degenerate primers Npro-(M-X)-ATCase_forw and Npro-(M-X)-ATCase_rev.

Methionine deletion variants (6H-VDVAD-∆M-X-cp ATCase) were created with site directed mutagenesis.

Experiments to establish P1’ toolbox

For the P1’ toolbox experiments untagged cp ATCase constructs (cp MA-ATCase, cp M-(T→X)-ATCase), were transformed into *E. coli* pyr- cells without protease). Cell growth under selective conditions confirmed ATCase activity.

First an alanine was added to the N-terminus of cp ATCase (cp MA-ATCase) to see if the desired P1’ amino acid could simply be inserted upstream of the original terminus. The activity of this construct was tested with the agar plate assay, no colony formation was observed.

Next, we generated and selected a gene library in which we mutated the second residue Thr\(^{228}\) to all canonical amino acids (cp M-(T→X)-ATCase). Active cp ATCase variants with nine different amino acids in the mutated position were found (A, C, G, P, Q, R, S, T, V). Nearly all these residues are readily accepted by MAP (54).

To express cp ATCase variants without starting methionine we generated a gene library of N\(^{pro}\)-cp ATCase. The fusion enabled us to exchange the starting residue of the cp ATCase (N\(^{pro}\)-(M→X)-cp ATCase). This approach was unsuccessful, only variants starting with methionine survived the selection.

These experiments indicated that the activity of the cp ATCase depends on removal of the starting methionine rather than on the variation of the Thr\(^{228}\) residue (Table S3). Apparently, the spatial
restrictions of the structure and the charges introduced with the new termini do not tolerate much variation. We concluded that removal of the methionine could increase the flexibility of the protein chain. To test the activity of different P1' cp ATCase variants without their native N-terminal methionine the fusion protein with caspase cleavage tag was used, as it has another start codon upstream of the tag. The disadvantage is, that the process is not only influenced by proper cp ATCase folding but also by the P1’ tolerance of the caspase. The methionine deletion variants were generated with the ten P1’ residues not found in the cp M-(T→X)-ATCase selection. Co-expression with cp caspase-2 under selective conditions confirmed the activity of 5 VDVAD-ΔM-X-cp ATCase variants (X = D, E, I, L, and N). No cell growth was observed on selective plates when the P1’ was F, H, K, W, or Y after 48–72 hours of incubation.

To verify the results, a methionine deletion variant of Npro-cp ATCase as well as the VDVAD-ΔM-cp ATCase together with cp caspase-2 were tested in the shaking flask assay. Cell growth was drastically increased in comparison to cells expressing the respective proteins without the deletion (data not shown). This confirms that the initial methionine is not necessary for enzymatic function but possibly even detrimental and that the subsequent amino acid of cp ATCase (Thr228) can be mutated to nearly all canonical amino acids (Table S3), which allows the generation of a toolbox for selection of proteases with a specific P1’ tolerance.

Table S3: Activity of cp ATCase, X is any amino acid despite methionine.

<table>
<thead>
<tr>
<th>N-terminal amino acids of cp c-chain</th>
<th>Cell growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTRVQKERL…</td>
<td>Yes (original construct)</td>
</tr>
<tr>
<td>Npro-MTRVQKERL…</td>
<td>Yes</td>
</tr>
<tr>
<td>Npro-XTRVQKERL…</td>
<td>No, except when X=M</td>
</tr>
<tr>
<td>Npro-TRVQKERL…</td>
<td>Yes, improved compared to original construct</td>
</tr>
<tr>
<td>MA-MTRVQKERL…</td>
<td>No</td>
</tr>
<tr>
<td>MXRVQKERL…</td>
<td>Yes, when X accepted by MAP</td>
</tr>
<tr>
<td>MHGSGVDVAD-MTRVQKERL…</td>
<td>Yes, when co-expressed with cp caspase-2</td>
</tr>
<tr>
<td>MHGSGVDVAD-TRVQKERL…</td>
<td>Yes, when co-expressed with cp caspase-2 improved compared to MHGSGVDVAD-MTRVQKERL…</td>
</tr>
<tr>
<td>MHGSGVDVAD-XRVQKERL…</td>
<td>Yes, when co-expressed with cp caspase-2 except X is F, H, K, W, Y</td>
</tr>
</tbody>
</table>

Sequences

Primer Sequences

Table S4: Oligonucleotides used for cloning experiments

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence 5’→3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>pyrI_genome_forw</td>
<td>TATACCATGGGCACACACGATAATAAATTGCAG</td>
</tr>
<tr>
<td>pyrI_genome_rev</td>
<td>TATAGCGGCCGCTTAATTGGCCAGCACCAC</td>
</tr>
<tr>
<td>MA-pyrB_forw</td>
<td>atggccATGACCCCGCTGGAACAAAGGA</td>
</tr>
<tr>
<td>MA-pyrB_rev</td>
<td>ATGTATATCTCTCTCTATATCTAATATATAC</td>
</tr>
<tr>
<td>M-X-ATCase_forw</td>
<td>TATACATATGnnnCGCGTGCAAAAAAGAG</td>
</tr>
<tr>
<td>M-X-ATCase_rev</td>
<td>TCTCCTTCTTATATCTAATATATAC</td>
</tr>
<tr>
<td>Npro-(M-X)-ATCase_forw</td>
<td>CACTAGTTGCn nhnACC CGTGC</td>
</tr>
<tr>
<td>Npro-(M-X)-ATCase_rev</td>
<td>ACCCACAGTGACAGTTAG</td>
</tr>
<tr>
<td>AVLQS_forw</td>
<td>gcagageCGCGTGCAAAAAAGAGCGT</td>
</tr>
</tbody>
</table>
AVLQS_rev  aagacggcgCGCGCTGCCATGATGATG
ep_caspase_forw (5' phosphorylated)  AGATATACATATGCAACCA
ep_caspase_rev  gcatccgatATCTTATTA
vector_forw (5' phosphorylated)  CCTTCTTTACTTAAACTAATAT
vector_rev  TAATAAAGATtaggatgc

Construct sequences

Active Npro-cp ATCase
Nucleotide sequence: pyrI
ATGGGCACAC ACGATAATAA ATTGCAGGTT GAAGCTATTA AACGCGGCAC GGTAATTGAC
60
CATATCCCCG CCCAGATCGG TTTTAAGCTG TTGAGTCTGT TCAAGCTGAC CGAAACGGAT
120
CAGCGCATCA CCATCGGTCT GAACTTACCT TCTGGCGAGA TGGGCCGCAA AGATCGTAC
180
CAGCCACGG TTACCGGAT CGACAATACT GAAGTTGTTG GATTAATCGG CCCAAGTCTG
240
CCGGAGGCAG TAAGCAATGT GCTGTTCTGC CCCACACAGA ACTCTATACG CATCGCCGA
300
CCGGTTTCTA CAGCCTTTCG CGTGCAAGAA CGCAGCAATT ACATCGGCT CAAATGCAAA
360
TACTGTTAAG AAGAAGTTTG CCATAATGTG GTGCTGGCCA ATAA
420

Protein sequence: PyrI
MGTHDNKLQV EA1KRTGVID HIPA1Q1GFLK LSLFLKTETD QRT1G1L1P1 50
SGEMRKRDLI KI1TFL1SED QVDQ1L1YAP QATVNR1DNY EVEVGKSRPSL 100
PERIDNVLC PNSNC1S1AE PVS1SSF1AVRK RAND1ALKY YCE1KEFSHV 150
VLAN 154

Nucleotide sequence: Npro-pyrB
ATGGAACTCA ATCATTTCGA ACTGCTCTAC AAAACTAGCA AGCAAAAACC TGTTGGCGTT
60
GAAGAGCCGG TCTACGATAC TGCAGGTCGT CCTCTTTTTG GGAATCCGTC CGAAGTGCAC
120
CCCCAGTCAA CCCTCAAGCT TCCCCATGAC CGCGGACGCG GTGACATTCG TACAACGCTG
180
CGCGATCTGC CTCGTAAAGG CGATTGTCGC TCTGGAAACC ACC
240
TAGGTCC GGTGTCGGGC
260
ATTTAATTTA AACCAGGTCC CGTCTATTAC CAAGACTACA CTGGTCCGGT TTACCATCGT
320
GCACCTCTGG AATTCTTTGA TGAAGCTCAA TTTTGCGAAG TGACTAAACG TATTGGCCGT
360
GTAACCGCTT GCAACGAGCC AACTCCCTGC ACCCTGGAAC TCTAGGTCAT TTGACTACTAC
420
GCACCTCTGG AATTCTTTGA TGAAGCTCAA TTTTGCGAAG TGACTAAACG TATTGGCCGT
460

Protein Sequence: Npro-pyrB
MELNHFELLY KTSKQKPVGV EEPVYDTAGR PLFGNPSEVH PQSTLKLPHD
50
RGRGDIRTTL RDLPRGKDCR SGNHLGPVSG IYIKPGPVYY QDYTGPVYHR
100
APLLEFFDEAQ FCEVTKR1GR VTSGDGKLYH IYVCDVGCLIL LKLAKRGTPR
150
TLKWRNFTN CPLWVTSCMT RVQKRLDFS EYANVKAPFV LRASDLHNAR
200
ANMKVLHLP L RPDEIATDV KTPHAWYFQQ AGNGIFARQA LLALVLNRA
250
Inactive \textsc{N}^{pro}\textsc{-}cp ATCase

\textbf{Protein sequence: iN^{pro}\textsc{-pyrB}}

\begin{verbatim}
MELNHFELLY KTSI KQPKPGVGG EEPVYDTAGR PLFGNPSEVH PQSTLKLPHD 50
RGRDGRDQTL RLDRGDGCR SGNHLGFPVS GQYIKGPVPYY QDYTPVYHPH 100
APELLEDFEAQ PEVVTKRIGR VTSGDGKLYH IYCVDGCHIL LKLABRTGTPR 150
TLKWRNFTN CPLAVTSCMT RQVKELLRLPS EYANVKQFVQ LRSQDLHAL 200
ANMKVLHPLF RVDEIATDVF KTPHAWHFQQ AGNGIFARQA LLALVNLKAN 250
PLYQKHIISI NDLSRDDLNL VLATAAKLKA NPQPELLKHK VIASCFFEAS 300
TRTRLSFETS MHRLGAASVG FDSSANTSLG KKGETLADTI SVISTYGVRL 350
VMRHPQEGGAA RLATEFSGNV PVLNAGDGSN QHPTQTLDDL FTIQETQGRL 400
DLNHLVANVGD LKYGRVTHSL TQALAKFDGN RFYFIAPDAL AMPQYILDML 450
\end{verbatim}

\textbf{cp ATCase c227 PyrI D73E}

\textbf{MCS I Protein Sequence:}

\begin{verbatim}
MGTHDNKLQV EAIKRGTVDD HPIAPOIQFKL LSLFKLTETD QRITILGLNL 50
SGEMGRDLKI SFGSLSLEDG QVEDLALYAP QATVNRIDNY EVVGAQGSPSL 100
PERIDNLVLC PNSNCISHAE PVSSSSAVKR RANDLAKCK YEKEFSNV 150
\end{verbatim}

\textbf{MCS II Protein Sequence:}

\begin{verbatim}
MTRVQKERLD PSEYANVKAQ FVLRAEHDLHN AKANMNVLHP LPRVEIADTH 50
VDKTHOWYF QAGNIGFAR QALLALVLNR ANPLYQKHII SINDLRSDDL 100
NLVLATAXKL KANPQPELKK HKVIAECFEE ASRTRLSEF TSMHLRASAV 150
VGFSDSANTS LKGGKGETLAD TISVISTYVD AVMRHPQEG AARLATEFSG 200
NVPVLNAGDG SNQHPTQTLDDL FTIQETQGRL LDNLHVVMD YLQKGRYHV 250
\end{verbatim}

\textbf{cp caspase-2}

\textbf{Nucleotide Sequence:}

\begin{verbatim}
ATGCCATCATC ATCACCATCA TGCGCAAAAT CATGCAGGTA GTCCGGGTTG TGAAGAAAGC 60
GCAGCACGTA AAGAAAAACT GCGCGAAATG CGTCTGCCCA CCCCSTACCGA TATGATTGTG 120
GGTTTAGCAT GTCTGAAGGG CACCCGCGGA ATCGTGATATA CCAAGCTTGG TAGCTGTATAT 180
ATGGACACGT TGGCAGAGGT TTTTACGCAGG GTGTGATGAT ATATGATGAT TGCGATATTG 240
CGTGTGTAATG TAAGCGGCCCT GATTTAAGAG CTGAGGCTTG TGGCGCAGGG TACAAGATT 300
CATGCGGGTA AAGAAGTGG CGAGTAGTTGT AGACCCCTGT GTGGCTATCT GTACCTGTTT 360
CCGGGCCATC CTCCGCGGGTG ATCCGGTCGG CTTGGTGCGT AGGTATAACC GTTGTACCGG 420
GAATTTTATC AGACATTTT TCAACGTCGCA ATACGTGTAC ATACGTGTAC ATACGTGTAC 480
GCACTGCTGC TATACGTGATA CTACGTGATA TCGAAGACGG ACCTGCTGCA CGTGGCTGCT 540
GGTTGAGTGG ATACGTTAGT CTTGTTAACT GTGCTGATCT AATGACGTCG TGGCAGGAGT 600
GCTCTGTTGG ATACGACGCG ACAAGAGAAT CAAAGGAATG TCGAAGACGG ACCTGCTGCA 660
CTCCCAGATC TGTGCTGACG TACAGTCTAT GTGTGATGAT TGCGATATTG TGGCGCAGGG 720
CAATTTATTG GTGTTGAGTG CACACGCGCC GACACGGGCG TGGCAGGAGT TGGCGCAGGG 780
AATGGCAAAAT CTTCCGGCTGT TGCGAAGATG TCGACGCTTG TGGCGCAGGG TGGCGCAGGG 840
GGTGAGTGG ACGTCTCCTG GTGGCGACGG CAGGTATACT ATGGTGTTTG AA 882
\end{verbatim}

\textbf{Protein sequence:}

\begin{verbatim}
MHHHHIGKN HAGSPGCGES AAGKEKLPKM RLPLTRSDMIC GYACLKGTAA 50
MRNTKRSWY IELAQVFVEF RADCMHVADM LVKNVALIKD REGAYPAFEP 100
HRCKEMSYPE STLCRLHLYL GFHPPTGSGP VCLPQVFCTTO EFXQTHFQLA 150
YRLGSRPRGL ALVLSVHVT GEKEFERSG GDVGHSLTVL LFKLGLYVHD 200
VLCQTAQEM QEKQLQFQAQL PAVRVTSDLI VALSUHVGVEG AIVGVDKGLL 250
QLQEVFQLFD NANCPSLQNK PAMIFQACR GDETHQGDQ QD 292
\end{verbatim}
S9 (cp caspase-2 E105V)

**Nucleotide sequence:**

```
ATGCACCATC ATCACCATCA TGGCAAAAAT CATGCAGGTA GTCCGGGTTG TGAAGAAAGC

GCAGCAGGTA AAGAAAAACT GCCGAAAATG CGTCTGCCGA CCCGTAGCGA TATGATTTGT

GGTTATGCAT GTCTGAAAGG CACCGCAGCA ATGCGTAATA CCAAACGTGG TAGCTGGATG

GCAATTTATG ATGTGGATATG CAAACTGCTG ATGCGTTACG GAATTTTACC

CCGGGTCATC CTCCGACCGG ATCCGGTCCG GTTTGTCTGC AGGTTAAACC GTGTACACCG

GCAATTTATG ATGTGGATATG CAAACTGCTG ATGCGTTACG GAATTTTACC

CCGGGTCATC CTCCGACCGG ATCCGGTCCG GTTTGTCTGC AGGTTAAACC GTGTACACCG
```

**Protein sequence:**

```
MHHHHHHGKN HAGSPGCEES AAGKEKLPKM RLPTRSDMIC G

YACLKGTAA

MRNTKRGSWY IEALAQVFSE RACDMHVADM LVKVNALIKD REGYAPGTEF

HRCKVMSEYC STLCRHLYLF PGHPPTGSGP VCLQVKPCTP EFYQTHFQLA

YRLQSRPRGL ALVLSNVHFT GEKELEFRSG GDVDHSTLVT LFKLLGYDVH

VLCDQTAQEM QEKLQNFAQL PAHRVTDSCI VALLSHGVEG AIYVGDKL

QLQEVFGQLD NANCPSLQNK FKMFFIACR QDTEDRGVDQ 292
```

**SARS-CoV-2 MPro**

**Nucleotide sequence:**

```
ATGAGCGGTTTTGCTTTAAATGGCCTTTGCCGCTTAAAGTTGTAAGTTGTAAGTTTCAGGTTACCTG

TGCGGACCACTCAGCAGCTGAGTGATGTGTTTTATTATGTGAAATGGCAAACATCCGCTGTCG

ATGGTTGCTGTCGTTGAGAGAATTTTGAAGAGACGTTGCTGATCCAGAAGACTGCTGTTTCAACAT

ATTAGGCAGGTCTTCTGTAATGTTAGCTGTGGTTACGTTGTTTTCAACATTGATTATGATTGCTTGA

GCTTCTGGAACGCTACATCAGAGTCCAGCTGAGTTGCTGATCCAGCAGCTGATGGAAGGTAACG

TTTATGCTGGGTTGGTGATGCTGACTGCAACAGCTGACAGCTGCTGATCTGCTGATCTGCTGATCTG

TTGTTCTGCTGCACTGAGATGCCATACCATGCATGTTGATGCTGATGCTGATGCTGATGCTGATGCTG

TGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
```

**Protein sequence:**

```
MSGFRKMAFPSGKVEGCMVQVGTCTTTLNGLWLDDVYVCPRHIVICTSEDMLNPYEDLLIRKNSHNFL

VQAGNVQLRVIGHSMQVCNLKLVDTAPNKTPKFYKVFQRIQPGTFSVLACYNPSGSPVYQCAMRPNF

IKGSFLNGSIVGSGFNDVSPCYMHHMLPTGVHAGDLEGNFYPFVQRDQTQAAGTDITTVN

VLAWLYAIAVINGDRWFLNFRFFTLLNDNLFMLVAMKYNEPQDOLHDVILGPLSAQTXIAVLDMCALS

LQNGMNGRTILGSALLEDELPFEFDVQRQCSGVTFQ
```
Table S5 Composition of used buffers and media

<table>
<thead>
<tr>
<th>Buffer/Media</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBS (Phosphate buffered saline, pH 7.4)</td>
<td>137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄</td>
</tr>
<tr>
<td>TY (Tryptone yeast) medium</td>
<td>1 % (w/v) peptone, 0.7 % (w/v) yeast extract, 0.25 % (w/v) NaCl</td>
</tr>
<tr>
<td>SOC (Super optimal broth with catabolite repression)</td>
<td>20 g/L tryptone, 5 g/L yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 20 mM glucose</td>
</tr>
<tr>
<td>Supplemented M9 minimal medium</td>
<td>50 mM Na₂HPO₄, 20 mM KH₂PO₄, 10 mM NaCl, 1 mM MgSO₄, 0.1 mM CaCl₂, 0.4 % (w/v) glucose, 0.5 % (w/v) casamino acids, 20 mM NH₄Cl, 10 µg/ml FeSO₄, vitamins (0.001 mg/ml of each biotin, thiamine, riboflavin, pyridoxine, niacinamide), 0.025–1 mM IPTG</td>
</tr>
<tr>
<td>Used antibiotics</td>
<td>Kanamycin (50 µg/ml)</td>
</tr>
<tr>
<td></td>
<td>Ampicillin (100 µg/ml)</td>
</tr>
<tr>
<td></td>
<td>Chloramphenicol (34 µg/ml)</td>
</tr>
</tbody>
</table>