

**ACEMBL Expression System Series**  
***MultiColi***  
**Multi-Protein Expression for E.coli**

**User Manual**

*Vers. 1.0*  
*July 2010*

This manual is based on the original ACEMBL manual (November 2009) written by Yan Nie, Christoph Bieniossek and Imre Berger but has been revised, updated and, wherever necessary, modified to meet customer demands.

ACEMBL was developed at the European Molecular Biology Laboratory, EMBL Grenoble Outstation, Grenoble, France.

## Table of Contents

A. The ACEMBL System Kit: Contents and Storage	3
B. Synopsis	4
C. The ACEMBL System	5
C.1. ACEMBL vectors	5
C.2. The multiple integration element (MIE)	6
C.3. Tags, promoters, terminators	7
C.4. Complex Expression	8
D. Procedures	10
D.1. Cloning into ACEMBL vectors	10
D.1.1. Single gene insertion into the MIE by SLIC	10
D.1.2. Polycistron assembly in MIE by SLIC	15
D.1.3. Gene insertion by restriction/ligation	20
D.1.4. Multiplication by using the HE and BstXI sites	23
D.2. Cre-LoxP reaction of Acceptors and Donors	26
D.2.1. Cre-LoxP fusion of Acceptors and Donors	28
D.2.2. Deconstruction of fusion vectors by Cre	31
D.3. Co-expression by Co-transformation	33
E. ACEMBL multi-gene combination: Examples	35
E.1. SLIC cloning into ACEMBL vectors: human TFIIIF	35
E.2. The Homing endonuclease/BstXI module: yeast RES complex	36
F. Appendix	37
F.1. DNA sequence of MIE	37
F.2. DNA sequences of ACEMBL vectors	38
F.2.1. pACE1	38
F.2.2. pACE2	39
F.2.3. pDC	40
F.2.4. pDK	41
F.2.5. pDS	42
F.2.6. ACEMBL plasmid maps	43

## A. The ACEMBL System Kit: Contents and Storage

Reagents to be supplied in ACEMBL system kit:

Acceptor vectors: pACE1, pACE2 – approx. 2 µg plasmid DNA per vial

Donor vectors: pDC, pDK, pDS – approx. 2 µg plasmid DNA per vial

keep at 4°C for short-term storage and in a freezer at -20°C or lower for medium- and long-term storage (take care to avoid repeated freeze-thaw cycles, e.g. by aliquotting DNA prior to freezing)

*E. coli* transformed with Acceptor and Donor vectors, provided as agar stabs; for plating bacteria as a starting point for plasmid preparations

keep agar stabs at 4°C or room temperature

pirLC, pirHC cells<sup>†</sup>

<sup>†</sup> *E. coli* strains expressing the *pir* gene for propagation of donor vectors (any other strain with *pir*<sup>+</sup> background can be used as well). LC: low copy number propagation, HC: high copy number propagation of plasmids with R6Kγ origin.

### Additionally required reagents:

Antibiotics: ampicillin, chloramphenicol, kanamycin, spectinomycin, tetracycline

Enzymes: Cre recombinase

T4 DNA polymerase (for recombination insertion of genes)

Phusion polymerase (for PCR amplification of DNA)

Restriction enzymes and T4 DNA ligase (for conventional cloning)

Regular laboratory *E. coli* strain for cloning (TOP10, HB101, DH5α)

Expression strain(s) of choice

## B. Synopsis

ACEMBL *MultiColi* is a 3<sup>rd</sup> generation multi-gene expression system for complex production in *E. coli*, created at the European Molecular Biology Laboratory EMBL, at Grenoble. ACEMBL can be applied both manually and also in an automated setup by using a liquid handling workstation. ACEMBL applies tandem recombination steps for rapidly assembling many genes into multi-gene expression cassettes. These can be single or polycistronic expression modules, or a combination of these elements. ACEMBL also offers the option to employ conventional approaches involving restriction enzymes and ligases if desired, which may be the methods of choice in laboratories not familiar with recombination approaches.

The following strategies for multi-gene assembly and expression are provided for in the ACEMBL system and detailed in Sections C and D:

- (1) Single gene insertions into vectors (recombination or restriction/ligation)
- (2) Multi-gene assembly into a polycistron (recombination or restriction/ligation)
- (3) Multi-gene assembly using homing endonucleases
- (4) Multi-gene plasmid fusion by Cre-LoxP reaction
- (5) Multi-gene expression by cotransformation

These strategies can be used individually or in conjunction, depending on the project and user.

In Section D, step-by-step protocols are provided for each of the methods for multi-gene cassette assembly that can be applied in the ACEMBL *MultiColi* system. Each procedure is illustrated by corresponding complex expression experiments in Section D of this Supplement.

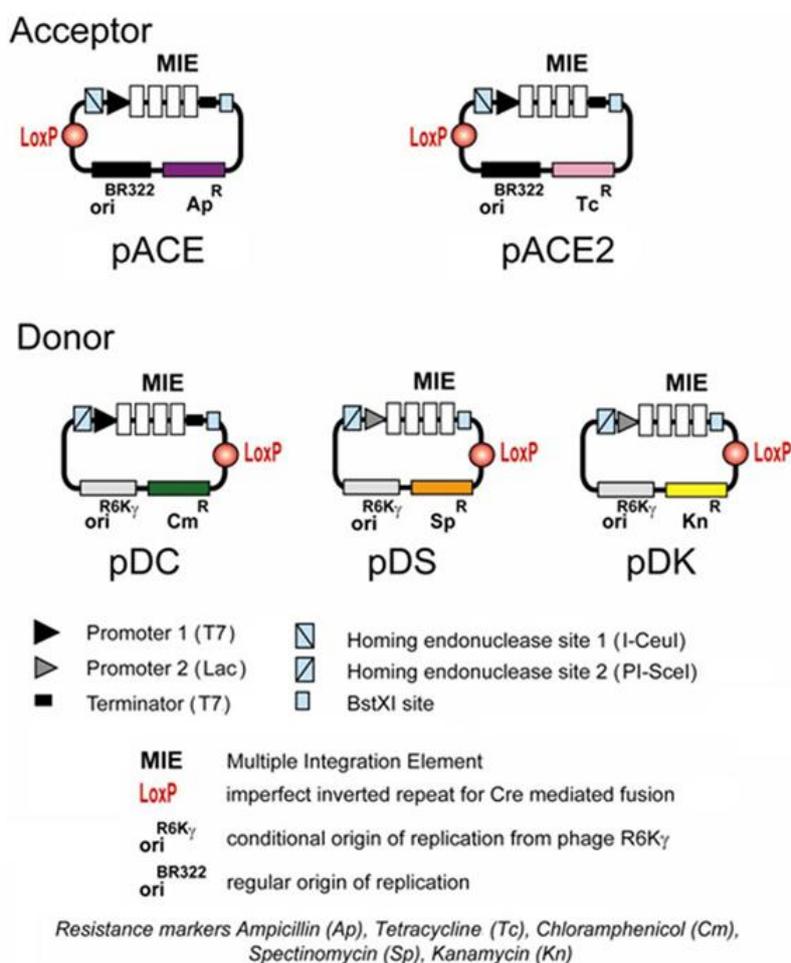
DNA sequences of ACEMBL vectors are provided in the Appendix and can be copied from there for further use.

## C. The ACEMBL System

### C.1. ACEMBL vectors

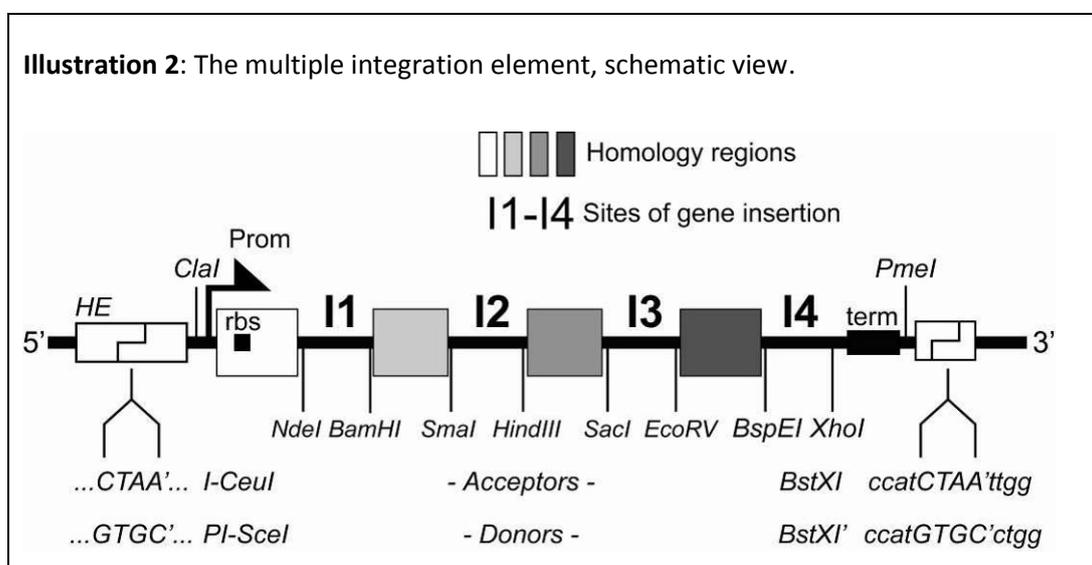
At the core of the technology are five small *de novo* designed vectors which are called “Acceptor” and “Donor” vectors (Illustration 1). Acceptor vectors (pACE1, pACE2) contain origins of replication derived from ColE1 and resistance markers (ampicillin or tetracycline). Donor vectors contain conditional origins of replication (derived from R6K $\gamma$ ), which make their propagation dependent on hosts expressing the *pir* gene. Donor vectors contain resistance markers kanamycin, chloramphenicol, or spectinomycin. Up to three Donor vectors can be used in conjunction with one Acceptor vector.

**Illustration 1:** ACEMBL system for multi-protein complex production.



All Donor and Acceptor vectors contain a loxP imperfect inverted repeat and in addition, a multiple integration element (MIE). This MIE consists of an expression cassette with a promoter of choice (prokaryotic, mammalian, insect cell specific or a combination thereof) and a terminator (prokaryotic, mammalian, insect cell specific or a combination thereof). In between is a DNA segment which contains a number of restriction sites that can be used for conventional cloning approaches or also for generating double-strand breaks for the integration of expression elements of choice (further promoters, ribosomal binding sites, terminators and genes). The MIE is completed by a homing endonuclease site and a specifically designed restriction enzyme site (BstXI) flanking the promoter and the terminator (see C.2.) Vector DNA sequences are provided in the Appendix. Maps of all vectors are shown at the end of this manual.

## C.2. The multiple integration element (MIE)



The MIE was derived from a polylinker<sup>1</sup> and allows several approaches for multi-gene assembly (Section D). Multiple genes can be inserted into the MIE of any one of the vectors by a variety of methods, for example BD-In-Fusion recombination<sup>2</sup> or SLIC (sequence and ligation independent cloning)<sup>3</sup>. For this, the vector needs to be linearized, which can also be carried out efficiently by PCR reaction with appropriate primers, since the vectors are all small (2 to 3 kb). Use of ultrahigh-fidelity

<sup>1</sup> Tan, S. et al. *Protein Expr. Purif.* **40**, 385 (2005)

<sup>2</sup> ClonTech TaKaRa Bio Europe, [www.clontech.com](http://www.clontech.com)

<sup>3</sup> Li, M. and Elledge, S., *Nat. Methods* **4**, 251 (2007)

polymerases such as Phusion<sup>4</sup> is recommended. Alternatively, if more conventional approaches are preferred, i.e. in a regular wet lab setting without robotics, the vectors can also be linearized by restriction digestion, and a gene of interest can be integrated by restriction / ligation (Section D). The DNA sequence of the MIE is shown in the Appendix.

### C.3. Tags, promoters, terminators

Current vectors of the ACEMBL system contain the default promoters T7 and Lac, as well as the T7 terminator element (Illustr.1, 10). The T7 system is currently most commonly used; it requires bacterial strains which contain a T7 polymerase gene in the *E. coli* genome. The Lac promoter is a strong endogenous promoter which can be utilized in most strains. All ACEMBL vectors contain the lac operator element for repression of heterologous expression.

Evidently, all promoters and terminators present in ACEMBL Donor and Acceptor vectors, and in fact the entire multiple integration element (MIE) can be exchanged with a favored expression cassette by using restriction/ligation cloning with appropriate enzymes (for example ClaI/PmeI, Illustration 2) or insertion into linearized ACEMBL vectors where the MIE was removed by sequence and ligation independent approaches such as SLIC. In pDC, the T7 promoter was substituted with a trc promoter (pDC<sup>trc</sup>) and the T7 promoter in pACE with an arabinose promoter (pACE<sup>ara</sup>), and the resulting vectors were used successfully in co-expression experiments by inducing with arabinose and IPTG.

Currently, the ACEMBL system vectors do not contain DNA sequences encoding for affinity tags that facilitate purification or solubilization of the protein(s) of interest. We typically use C- or N-terminal oligohistidine tags, with or without protease sites for tag removal. We introduce these by means of the respective PCR primers used for amplification of the genes of interest prior to SLIC mediated insertion. We recommend outfitting Donors or Acceptors of choice by the array of custom tags that are favored in individual user laboratories prior to inserting

---

<sup>4</sup> Finnzymes/New England BioLabs, [www.neb.com](http://www.neb.com)

recombinant genes of interest. This is best done by a design which will, after tag insertion, still be compatible with the recombination based principles of ACEMBL system usage.

#### **C.4. Complex Expression**

For expression in *E.coli*, the ACEMBL multi-gene expression vector fusions with appropriate promoters or terminators are transformed into the appropriate expression host of choice. In the current version (T7 and lac promoter elements), most of the wide array of currently available expression strains can be utilized. If particular expression strains already contain helper plasmids with DNA encoding for chaperones, lysozyme or else, the design of the multi-gene fusion should ideally be such that the ACEMBL vector containing the resistance marker that is also present on the helper plasmid is not included in multi-gene vector construction (although this is probably not essential).

Alternatively, the issue can be resolved by creating new versions of the ACEMBL vectors containing resistance markers that circumvent the conflict. This can be easily performed by PCR amplifying the vectors minus the resistance marker, and combine the resulting fragments with a PCR amplified resistance marker by recombination (SLIC) or blunt-end ligation (using 5'phosphorylated primers). Note that resistance markers can also be exchanged in between ACEMBL vectors by restriction digestion with AlwNI and ClaI (for Donors) and AlwNI and PmeI (for Acceptors).

Donor vectors depend on the *pir* gene product expressed by the host, due to the R6Ky conditional origin of replication. In regular expression strains, they rely on fusion with an Acceptor for productive replication. Donors or Donor-Donor fusions can nonetheless be used even for expression when not fused with an Acceptor, by using expression strains carrying a genomic insertion of the *pir* gene. Such strains have recently become available (e.g. from Novagen Inc., Madison WI, USA).

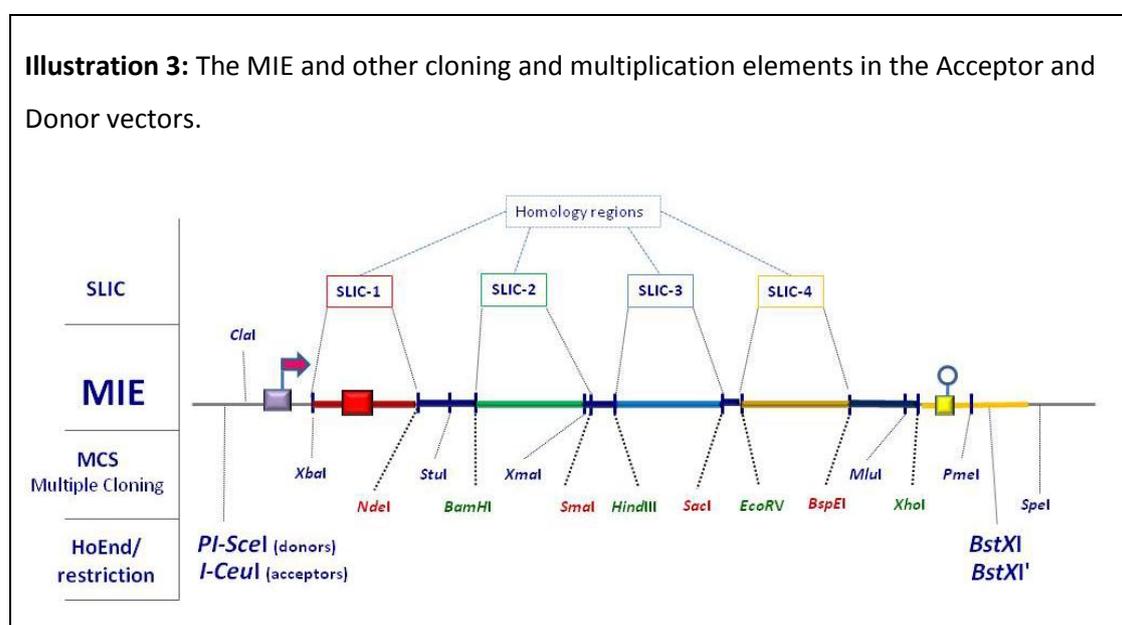
Co-transformation of two plasmids can also lead to successful protein complex expression. The ACEMBL system contains two Acceptor vectors, pACE and

pACE2, which are identical except for the resistance marker (Illustration 1). Therefore, genes present on pACE1 or pACE2, respectively, can be expressed by co-transformation of the two plasmids and subsequent simultaneous exposure to tetracyclin and ampicillin. In fact, entire Acceptor-Donor fusions containing several genes, based on pACE1 or pACE2 as Acceptors, can in principle be co-transformed for multi-expression, if needed.

## D. Procedures

### D.1. Cloning into ACEMBL vectors

All Donors and Acceptors contain an identical MIE with exception of the homing endonuclease site / BstXI tandem encompassing the MIE (Illustrations 1 and 3, plasmid maps in the appendix). The MIE is tailored for sequence and ligation independent gene insertion methods. In addition, the MIE also contains a series of unique restriction sites, and therefore can be used as a classical polylinker for conventional gene insertion by restriction/ligation. We suggest to choose the methods a user lab is most familiar with. For automated applications, restriction/ligation is essentially ruled out. In this case, recombination approaches can be used efficiently for gene insertion (SLIC).



#### D.1.1. Single gene insertion into the MIE by SLIC

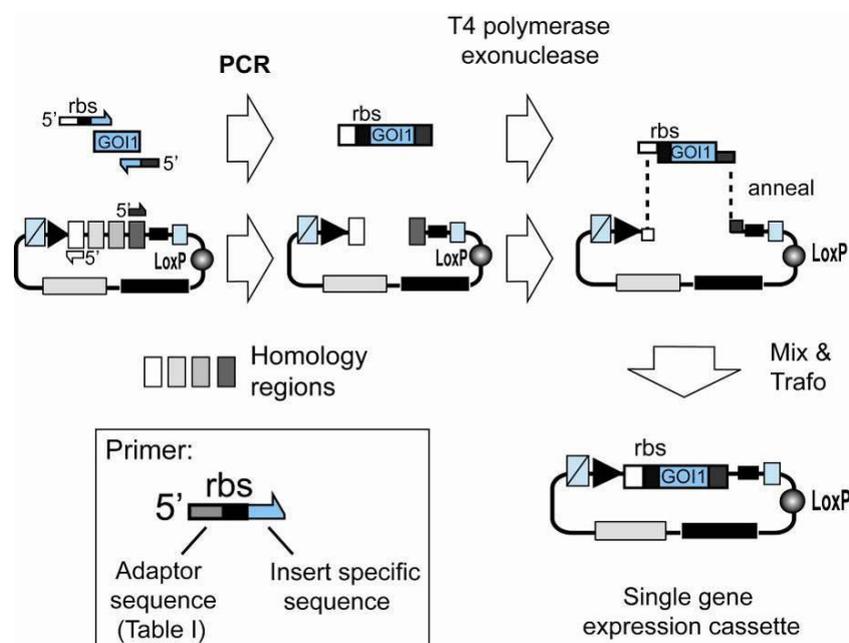
Several procedures for restriction/ligation independent insertion of genes into vectors have been published or commercialized (Novagen LIC, Becton-Dickinson BD In-Fusion and others), each with its own merit. All of these systems rely on the exonuclease activity of DNA polymerases. In the absence of dNTPs, 5' extensions are created from blunt ends or overhangs by digestion from the 3' end. If two DNA fragments contain the same approx. 20 bp sequence at their termini at opposite

ends, this results in overhangs that share complementary sequences capable of annealing. This can be exploited for ligation independent combination of two or several DNA fragments containing homologous sequences.

If T4 DNA polymerase is used, this can be carried out in a manner that is independent of the sequences of the homology regions (Sequence and Ligation Independent Cloning, SLIC) and detailed protocols have become available. In the context of multi-protein expression, this is particularly useful, as the presence of unique restriction sites, or their creation by mutagenesis, in the ensemble of encoding DNAs ceases to be an issue.

We adapted SLIC for the insertion of encoding DNAs amplified by Phusion polymerase into the ACEMBL Acceptor and Donor vectors according to the published protocols. This not only allows seamless integration of genes into the expression cassettes, but also concatamerization of expression cassettes into multi-gene constructs via a simple and repetitive routine that can be readily automated.

**Illustration 4: Single gene insertion by SLIC.** A gene of interest (GOI 1) is PCR amplified with specific primers and integrated into a vector (Acceptor, Donor) linearized by PCR with complementary primers (complementary regions are shaded in light gray or dark grey, respectively). Resulting PCR fragments contain homology regions at their ends. T4 DNA polymerase acts as an exonuclease in the absence of dNTP and produces long sticky overhangs. Mixing (optionally annealing) of T4DNA polymerase exonuclease treated insert and vector is followed by transformation, yielding a single gene expression cassette.



We use an improved protocol for SLIC which was modified from the original publication<sup>5</sup>. This protocol, as applied manually, is detailed below (Protocol 1). If other systems are used (BD-InFusion etc.), follow the manufacturer's recommendations. For robotics applications, modifications of the protocol may be necessary and will be detailed elsewhere<sup>6</sup>.

**Protocol 1:** Single gene insertion by SLIC.

Reagents required:

Phusion Polymerase  
5x HF Buffer for Phusion Polymerase  
dNTP mix (10 mM)  
T4 DNA polymerase (and 10x Buffer)  
DpnI enzyme  
*E. coli* competent cells  
100mM DTT, 2M Urea, 500 mM EDTA  
Antibiotics

**Step 1:** Primer design

Primers for the SLIC procedure are designed to provide the regions of homology that result in long sticky ends after treatment with T4 DNA polymerase in the absence of dNTP:

Primers for the insert contain a DNA sequence corresponding to this region of homology ("Adaptor sequence" in Illustration 4, inset), followed by a sequence stretch that specifically anneals to the insert to be amplified (Illustration 4, inset). Useful adaptor sequences for SLIC are listed below (Table I).

If the gene of interest (GOI) is amplified from a vector already containing expression elements (e.g. the pET vector series), this "insert specific sequence" can be located upstream of a ribosome binding site (rbs). Otherwise, the forward primer needs to be designed such that a ribosome binding site is also provided in the final construct (Illustration 4, inset).

<sup>5</sup> Li, M. and Elledge, S., *Nat. Methods* **4**, 251 (2007)

<sup>6</sup> Bieniossek, C., Nie, Y. et al., *Nat. Methods* **6**, 447 (2009).

Primers for PCR linearization of the vector backbone are simply complementary to the two adaptor sequences present in the primer pair chosen for insert amplification (Illustration 4).

### Step 2: PCR amplification of insert and vector

Identical reactions are prepared in 100- $\mu$ l volume for DNA insert to be cloned and vector to be linearized by PCR:

ddH <sub>2</sub> O	75 $\mu$ l
5 $\times$ Phusion HF Reaction buffer	20 $\mu$ l
dNTPs (10 mM stock)	2 $\mu$ l
Template DNA (100 ng/ $\mu$ l)	1 $\mu$ l
5' SLIC primer (100 $\mu$ M stock)	1 $\mu$ l
3' SLIC primer (100 $\mu$ M stock)	1 $\mu$ l
Phusion polymerase (2 U/ $\mu$ l)	0.5 $\mu$ l

PCR reactions are then carried out with a standard PCR program (unless very long DNAs are amplified, in which case you should the double extension time):

1 x 98° C for 2 min  
 30 x [98° C for 20 sec. -> 50° C for 30 sec. -> 72° C for 3 min]  
 Hold at 10° C

Analysis of the PCR reactions by agarose gel electrophoresis and ethidium bromide staining is recommended.

### Step 3: DpnI treatment of PCR products (optional)

PCR reactions are then supplied with 1  $\mu$ l DpnI enzyme which cleaves parental plasmids (that are methylated). For insert PCR reactions, DpnI treatment is not required if the resistance marker of the template plasmid differs from the destination vector.

Reactions are carried out as follows:

Incubation: 37° C for 1-4h  
 Inactivation: 80° C for 20 min

### Step 4: Purification of PCR products

**! PCR products must be purified to remove residual dNTPs!**

Otherwise, the T4 DNA polymerase reaction (Step 5) may be compromised.

Product purification is best performed by using commercial PCR Purification Kits or NucleoSpin Kits (Qiagen, MacheryNagel or others). It is recommended to perform elution in the minimal possible volume indicated by the manufacturer.

### Step 5: T4 DNA polymerase exonuclease treatment

Identical reactions are prepared in 20- $\mu$ l volume for insert and for vector (eluted in Step 4):

10x T4 DNA polymerase buffer	2 $\mu$ l
100mM DTT	1 $\mu$ l
2M Urea	2 $\mu$ l
DNA eluate from Step 3 (vector or insert)	14 $\mu$ l
T4 DNA polymerase	1 $\mu$ l

Reactions are then carried out as follows:

Incubation:	23°C for 20 min
Arrest:	Addition of 1 $\mu$ l 500 mM EDTA
Inactivation:	75°C for 20 min

### Step 6: Mixing and Annealing

Exonuclease treated insert and vector are then mixed, followed by an (optional) annealing step which was found to enhance efficiency<sup>7</sup>:

T4 DNA pol treated insert:	10 $\mu$ l
T4 DNA pol treated vector:	10 $\mu$ l
Annealing:	65°C for 10 min
Cooling:	Slowly (in heat block) to RT

### Step 7: Transformation

Mixtures are next transformed into competent cells following standard transformation procedures.

Reactions for pACE1 and pACE2 derivatives are transformed into standard *E. coli* cells for cloning (such as TOP10, DH5 $\alpha$ , HB101) and after recovery (2-4h) plated on agar containing ampicillin (100  $\mu$ g/ml) or tetracycline (25  $\mu$ g/ml), respectively.

Reactions for Donor derivatives are transformed into *E. coli* cells expressing the *pir* gene (such as BW23473, BW23474, or PIR1 and PIR2, Invitrogen) and plated on agar containing chloramphenicol (25  $\mu$ g/ml, pDC), kanamycin (50  $\mu$ g/ml, pDK), and spectinomycin (50  $\mu$ g/ml, pDS).

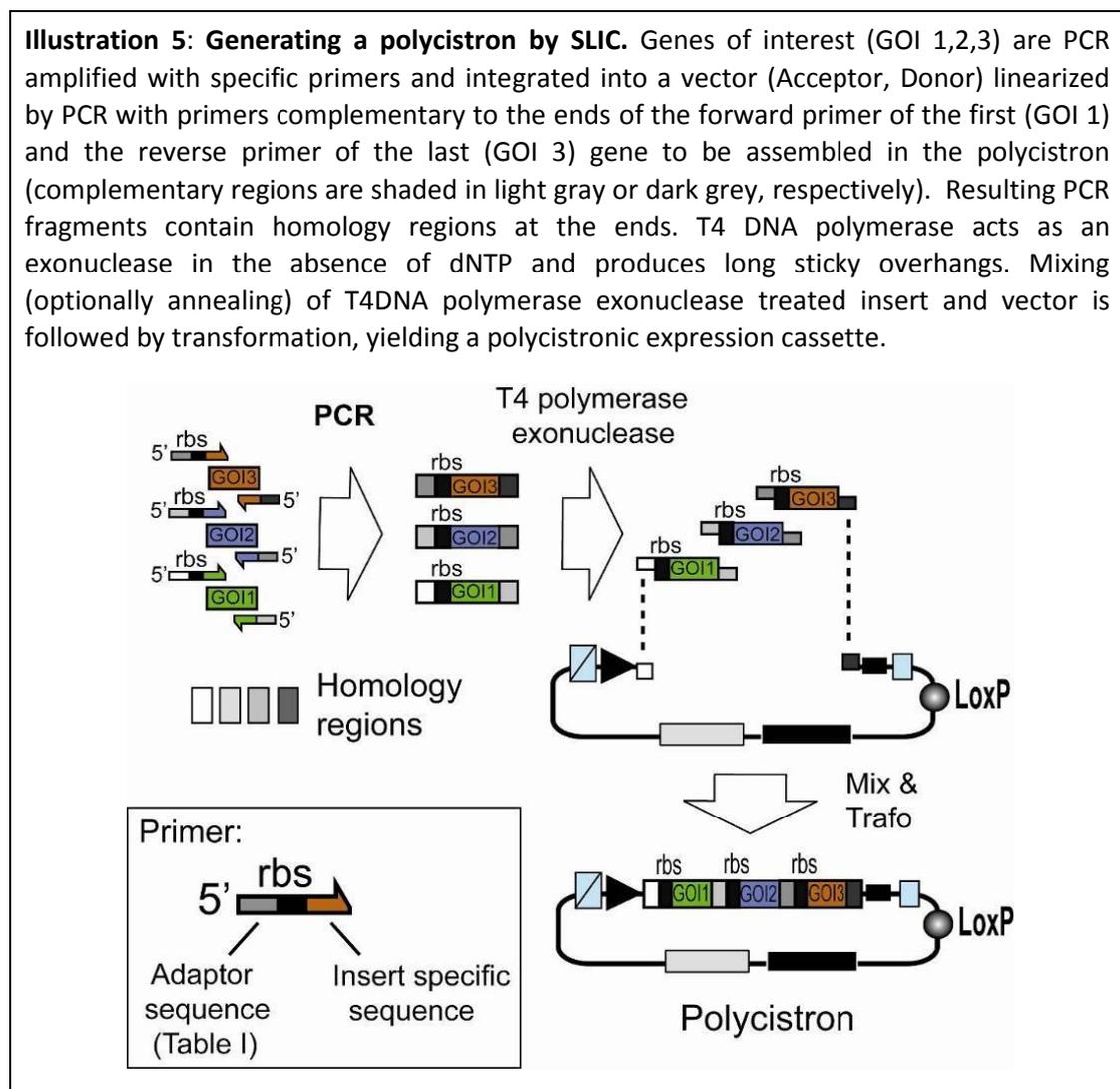
### Step 8: Plasmid analysis

Plasmids are cultured in small-scale in media containing the corresponding antibiotic, and analyzed by sequencing and (optionally) restriction mapping with an appropriate restriction enzyme.

<sup>7</sup> Dr. Rolf Jaussi, PSI Villigen, personal communication

## D.1.2. Polycistron assembly in MIE by SLIC

The multiple integration element can also be used to integrate genes of interest by using multi-fragment SLIC recombination as shown in Illustration 4. Genes preceded by ribosome binding sites (rbs) can be assembled into polycistrons in this way.

**Protocol 2.** Polycistron assembly by SLIC.

Reagents required:

Phusion Polymerase

5x HF Buffer for Phusion Polymerase

dNTP mix (10 mM)

T4 DNA polymerase (and 10x Buffer)

*E. coli* competent cells

100mM DTT, 2M Urea, 500 mM EDTA

Antibiotics

### Step 1: Primer design

The MIE element is composed of tried-and-tested primer sequences. These constitute the “Adaptor” sequences that can be used for inserting single genes or multi-gene constructs. Recommended adaptor sequences are listed below (Table I).

Adaptor sequences form the 5' segments of the primers used to amplify DNA fragments to be inserted into the MIE. Insert specific sequences are added at 3', DNA encoding a ribosome binding site can be inserted optionally if not already present on the PCR template

### Step 2: PCR amplification of insert and primer

Identical reactions are prepared in 100- $\mu$ l volume for all DNA insert (GOI 1,2,3) to be cloned and the vector to be linearized by PCR:

ddH <sub>2</sub> O	75 $\mu$ l
5 $\times$ Phusion HF Reaction buffer	20 $\mu$ l
dNTPs (10 mM stock)	2 $\mu$ l
Template DNA (100 ng/ $\mu$ l)	1 $\mu$ l
5' SLIC primer (100 $\mu$ M stock)	1 $\mu$ l
3' SLIC primer (100 $\mu$ M stock)	1 $\mu$ l
Phusion polymerase (2 U/ $\mu$ l)	0.5 $\mu$ l

PCR reactions are then carried out with a standard PCR program (unless very long DNAs are amplified, in which case you should double the extension time):

1 x 98° C for 2 min

30 x [98° C for 20 sec. -> 50°C for 30 sec. -> 72°C for 3 min]

Hold at 10°C

Analysis of the PCR reactions by agarose gel electrophoresis and ethidium bromide staining is recommended.

### Step 3: DpnI treatment of PCR products (optional)

PCR reactions are then supplied with 1  $\mu$ l DpnI enzyme which cleaves parental plasmids (that are methylated). For insert PCR reactions, DpnI treatment is not required if the resistance marker of the template plasmids differs from the destination vector.

Reactions are then carried out as follows:

Incubation: 37°C for 1-4h

Inactivation: 80°C for 20 min

#### Step 4: Purification of PCR products

**! PCR products must be purified to remove residual dNTPs !**

Otherwise, the T4 DNA polymerase reaction (Step 5) may be compromised.

Product purification is best performed by using commercial PCR Purification Kits or NucleoSpin Kits (Qiagen, MacheryNagel or others). It is recommended to perform elution in the minimal possible volume indicated by the manufacturer.

#### Step 5: T4 DNA polymerase exonuclease treatment

Identical reactions are prepared in 20- $\mu$ l volume for each insert (GOI 1,2,3) and for the vector (eluted in Step 4):

10x T4 DNA polymerase buffer	2 $\mu$ l
100mM DTT	1 $\mu$ l
2M Urea	2 $\mu$ l
DNA eluate from Step 3 (vector or insert)	14 $\mu$ l
T4 DNA polymerase	1 $\mu$ l

Reactions are then carried out as follows:

Incubation: 23°C for 20 min

Arrest: Addition of 1  $\mu$ l 500 mM EDTA

Inactivation: 75°C for 20 min

#### Step 6: Mixing and Annealing

T4 DNA polymerase exonuclease treated insert and vector are then mixed, followed by an (optional) annealing step which was found to enhance efficiency<sup>8</sup>:

T4 DNA pol treated insert 1 (GOI 1):	5 $\mu$ l
T4 DNA pol treated insert 2 (GOI 2):	5 $\mu$ l
T4 DNA pol treated insert 3 (GOI 3):	5 $\mu$ l
T4 DNA pol treated vector:	5 $\mu$ l
Annealing:	65°C for 10 min
Cooling:	Slowly (in heat block) to RT

#### Step 7: Transformation

Mixtures are next transformed into competent cells following standard transformation procedures.

---

<sup>8</sup> Dr. Rolf Jaussi, PSI Villigen, personal communication.

Reactions for pACE1 and pACE2 derivatives are transformed into standard *E. coli* cells for cloning (such as TOP10, DH5 $\alpha$ , HB101) and after recovery plated on agar containing ampicillin (100  $\mu\text{g/ml}$ ) or tetracycline (25  $\mu\text{g/ml}$ ), respectively.

Reactions for Donor derivatives are transformed into *E. coli* cells expressing the *pir* gene (such as BW23473, BW23474, or PIR1 and PIR2, Invitrogen) and plated on agar containing chloramphenicol (25  $\mu\text{g/ml}$ , pDC), kanamycin (50  $\mu\text{g/ml}$ , pDK), and spectinomycin (50  $\mu\text{g/ml}$ , pDS).

### Step 8: Plasmid analysis

Plasmids are cultured and correct clones are selected based on specific restriction digestion and DNA sequencing of the inserts.

**Table I.** Adaptor sequences

For single gene or multi-gene insertions into ACEMBL vectors by SLIC.

Adaptor <sup>1</sup>	Sequence	Description
T7InsFor	TAGGTATCGATAATAC GACTCACTATAGGG	Forward primer for <u>insert</u> amplification, if gene of interest (GOI) is present in a T7 system vector (i.e. pET series). No further extension (rbs, insert specific overlap) required.
T7InsRev	CCTCAAGACCCGTTTA GAGGCCCAAGGGGT TATGCTAG	Reverse primer for <u>insert</u> amplification, if GOI is present in a T7 system vector (i.e. pET series). No further extension (stop codon, insert specific overlap) required.
T7VecFor	CTAGCATAACCCCTTG GGGCCTCTAACGGGT CTTGAGG	Forward primer for <u>vector</u> amplification, reverse complement of T7InsRev. No further extension required.
T7VecRev	CCCTATAGTGAGTCGT ATTATCGATACCTA	Reverse primer for <u>vector</u> amplification, reverse complement of T7InsFor. No further extension required.
NdeInsFor	GTTTAACTTTAAGAAG GAGATATACATATG	Forward primer for <u>insert</u> amplification for insertion into MIE site I1 (Illustration 2). Further extension at 3' (insert specific overlap) required. Can be used with adaptor XhoInsRev in case of single fragment SLIC (Illustr. 3).
XhoInsRev	pACE1,pACE2,pDC (T7): CTTTGTTAGCAGCCGG ATCTCTCGAG pDK,pDS (lac): GGGTTTAAACGGAAGT AGTCTCGAG	Reverse primer for <u>insert</u> amplification for insertion into MIE site I4 (Illustr. 2). Further extension at 3' (stop codon, insert specific overlap) required. Can be used with adaptor NdeInsFor in case of single fragment SLIC (Illustr. 3).
XhoVecFor	pACE1,pACE2,pDC (T7):	Forward primer for <u>vector</u> amplification, reverse complement of XhoInsRev

	CTCGAGAGATCCGGCT GCTAACAAAG pDK,pDS (lac): CTCGAGACTAGTTCCGT TTAAACCC	No further extension required.
NdeVecRev	CATATGTATATCTCCTT CTTAAAGTTAAAC	Reverse primer for <u>vector</u> amplification, reverse complement of NdeInsFor. No further extension required.
SmaBam	GAATTCAGTGGCCGTC GTTTTACAGGATCC	Reverse primer for <u>insert</u> amplification (GOI1) for insertion into MIE site I1 (Illustr. 2). Further extension at 3' (stop codon, insert specific overlap) required. Use with adaptor NdeInsFor.
BamSma	GGATCCTGTAACACGA CGGCCAGTGAATTC	Forward primer for <u>insert</u> amplification (GOI2) for insertion into site I2 (Illustr. 2,4). Further extension at 3' (rbs, insert specific overlap) required. Use with adaptor SacHind (multifragment SLIC, Illustr. 4)
SacHind	GCTCGACTGGGAAA ACCCTGGCGAAGCT T	Reverse primer for <u>insert</u> amplification (GOI2) insertion into MIE site I2 (Illustr. 2, 4). Further extension at 3' (stop codon, insert specific overlap) required. Use with adaptor BamSma (multifragment SLIC, Illustr. 4)
HindSac	AAGCTTCGCCAGGG TTTTCCCAGTCGAGC	Forward primer for <u>insert</u> amplification (GOI3) for insertion into site I3 (Illustr. 2,4). Further extension at 3' (rbs, insert specific overlap) required. Use with adaptor BspEco (multifragment SLIC, Illustr. 4)
BspEco5	GATCCGGATGTGAA ATTGTTATCCGCTGG TACC	Reverse primer for <u>insert</u> amplification (GOI3) insertion into MIE site I3 (Illustr. 2, 4). Further extension at 3' (stop codon, insert specific overlap) required. Use with adaptor HindSac.(multifragment SLIC, Illustr. 4)
Eco5Bsp	GGTACCAGCGGATA ACAATTTACATCCG GATC	Forward primer for <u>insert</u> amplification (GOI4) for insertion into site I4 (Illustr. 2,4). Further extension at 3' (rbs, insert specific overlap) required. Use with adaptor XhoInsRev (multifragment SLIC, Illustr. 4)

<sup>1</sup> All Adaptor primers (without extension) can be used as sequencing primers for genes of interest that were inserted into the MIE.

### D.1.3. Gene insertion by restriction/ligation

The MIE can also be interpreted as a simple multiple cloning site with a series of unique restriction sites. The MIE is preceded by a promoter and a ribosome binding site, and followed by a terminator. Therefore, cloning into the MIE by classical restriction/ligation also yields functional expression cassettes.

Genes of interest can be subcloned by using standard cloning procedures into the multiple integration element (MIE) (see Appendix) of ACEMBL vectors (the MIE is identical in all vectors).

#### **Protocol 3.** Restriction/ligation cloning into the MIE.

##### Reagents required:

- Phusion Polymerase
- 5x HF Buffer for Phusion Polymerase
- dNTP mix (10 mM)
- 10 mM BSA
- Restriction endonucleases (and 10x Buffer)
- T4 DNA ligase (and 10x Buffer)
- Calf or Shrimp intestinal alkaline phosphatase
- E. coli* competent cells
- Antibiotics

##### **Step 1:** Primer design

For conventional cloning, PCR primers are designed containing chosen restriction sites, preceded by appropriate overhangs for efficient cutting (c.f. New England Biolabs catalog), and followed by  $\geq 20$  nucleotides overlapping with the gene of interest that is to be inserted.

All MIEs are identical in the ACEMBL vectors. They contain a ribosome binding site preceding the NdeI site. For single gene insertions, therefore, a rbs does not need be included in the primer.

If multi-gene insertions are planned (for example into insertion sites I1-I4 of the MIE), primers need to be designed such that a rbs preceding the gene and a stop codon at its end are provided.

In particular for polycistron cloning by restriction/ligation, it is recommended to construct templates by custom gene synthesis. In the process, the restriction sites present in the MIE can be eliminated from the encoding DNAs.

**Step 2: Insert preparation**

PCR of insert(s):

Identical PCR reactions are prepared in 100 µl volume for genes of interest to be inserted into the MIE:

ddH <sub>2</sub> O	75 µl
5× Phusion HF Reaction buffer	20 µl
dNTPs (10 mM stock)	2 µl
Template DNA (100 ng/µl)	1 µl
5' primer (100 µM stock)	1 µl
3' primer (100 µM stock)	1 µl
Phusion polymerase (2 U/µl)	0.5 µl

PCR reactions are then carried out with a standard PCR program (unless very long DNAs are amplified, then double extension time):

1 x 98° C for 2 min

30 x [98° C for 20 sec. -> 50°C for 30 sec. -> 72°C for 3 min]

Hold at 10°C

Analysis of the PCR reactions by agarose gel electrophoresis and ethidium bromide staining is recommended.

Product purification is best performed by using commercial PCR Purification Kits or NucleoSpin Kits (Qiagen, MacheryNagel or others). It is recommended to perform elution in the minimal possible volume indicated by the manufacturer.

Restriction digestion of insert(s):

Restriction reactions are carried out in 40 µl reaction volumes, using the specific restriction enzymes as specified by the manufacturer's recommendations (c.f. New England Biolabs catalog and others).

PCR Kit eluate (≥1 µg)	30 µl
10x Restriction enzyme buffer	4 µl
10 mM BSA	2 µl
Restriction enzyme for 5'	2 µl
Restriction enzyme for 3'	2 µl (in case of double digestion, else ddH <sub>2</sub> O)

Restriction digestions are performed in a single reaction with both enzymes (double digestion) or sequentially (two single digestions) if the buffer conditions required are incompatible.

Gel extraction of insert(s):

Processed insert is then purified by agarose gel extraction using commercial kits (Qiagen, MachereyNagel etc). It is recommended to elute the extracted DNA in the minimal volume defined by the manufacturer.

### Step 3: Vector preparation

Restriction digestion of ACEMBL plasmid(s):

Restriction reactions are carried out in 40  $\mu$ l reaction volumes, using specific restriction enzymes as specified by manufacturer's recommendations (c.f. New England Biolabs catalog and others).

ACEMBL plasmid ( $\geq 0.5$ $\mu$ g) in ddH <sub>2</sub> O	30 $\mu$ l
10x Restriction enzyme buffer	4 $\mu$ l
10 mM BSA	2 $\mu$ l
Restriction enzyme for 5'	2 $\mu$ l
Restriction enzyme for 3'	2 $\mu$ l (for double digestions, else ddH <sub>2</sub> O)

Restriction digestions are performed in a single reaction with both enzymes (double digestion) or sequentially (two single digestions) if the buffer conditions required are incompatible.

Gel extraction of vector(s):

Processed vector is then purified by agarose gel extraction using commercial kits (Qiagen, MachereyNagel etc). It is recommended to elute the extracted DNA in the minimal volume defined by the manufacturer.

### Step 4: Ligation

Ligation reactions are carried out in 20  $\mu$ l reaction volumes according to the recommendations of the supplier of T4 DNA ligase:

ACEMBL plasmid (gel extracted)	8 $\mu$ l
Insert (gel extracted)	10 $\mu$ l
10x T4 DNA Ligase buffer	2 $\mu$ l
T4 DNA Ligase	0.5 $\mu$ l

Ligation reactions are performed at 25°C (sticky end) for 1h or at 16°C (blunt end) overnight.

### Step 5: Transformation

Mixtures are next transformed into competent cells following standard transformation procedures.

Reactions for pACE1 and pACE2 derivatives are transformed into standard *E. coli* cells for cloning (such as TOP10, DH5 $\alpha$ , HB101) and after recovery plated on agar containing ampicillin (100  $\mu$ g/ml) or tetracycline (25  $\mu$ g/ml), respectively.

Reactions for Donor derivatives are transformed into *E. coli* cells expressing the *pir* gene (such as BW23473, BW23474, or PIR1 and PIR2, Invitrogen) and plated

on agar containing chloramphenicol (25 µg/ml, pDC), kanamycin (50 µg/ml, pDK), and spectinomycin (50 µg/ml, pDS).

#### **Step 6:** Plasmid analysis

Plasmids are cultured and correct clones are selected based on specific restriction digestion and DNA sequencing of the inserts.

#### D.1.4. Multiplication by using the HE and BstXI sites

All ACEMBL system vectors contain a homing endonuclease (HE) site and a designed BstXI site that envelop the multiple integration element (MIE). The homing endonuclease site can be used to insert entire expression cassettes, containing single genes or polycistrons, into a vector already containing one gene or several genes of interest. Homing endonucleases have long recognition sites (20-30 base pairs or more). Although not all equally stringent, homing endonuclease sites are most probably unique in the context of even large plasmids, or, in fact, entire genomes.

In the ACEMBL system, Donor vectors contain a recognition site for homing endonuclease PI-SceI (Illustr. 2). This HE site yields upon cleavage a 3' overhang with the sequence -GTGC. Acceptor vectors contain the homing endonuclease site I-CeuI, which upon cleavage will result in a 3' overhang of -CTAA. On Acceptors and Donors, the respective HE site is preceding the MIE. The 3' end of the MIE contains a specifically designed BstXI site, which upon cleavage will generate a matching overhang. The basis of this is the specificity of cleavage by BstXI. The recognition sequence of BstXI is defined as CCANNNNN'NTGG (apostrophe marks position of phosphodiester link cleavage). The residues denoted as N can be chosen freely. Donor vectors thus contain a BstXI recognition site of the sequence CCATGTGC'CTGG, and Acceptor vectors contain CCATCTAA'TTGG. The overhangs generated by BstXI cleavage in each case will match the overhangs generated by HE cleavage. Note that Acceptors and Donors have different HE sites.

The recognition sites are not symmetric. Therefore, ligation of a HE/BstXI digested fragment into a HE site of an ACEMBL vector will be (1) directional and (2) result in a hybrid DNA sequence where a HE halfsite is combined with a BstXI halfsite. This site will be cut by neither HE nor BstXI. Therefore, in a construct that

had been digested with a HE, insertion by ligation of HE/BstXI digested DNA fragment containing an expression cassette with one or several genes will result in a construct which contains all heterologous genes of interest, enveloped by an intact HE site in front, and a BstXI site at the end. Therefore, the process of integrating entire expression cassettes by means of HE/BstXI digestion and ligation into a HE site can be repeated iteratively.

**Protocol 4.** Multiplication by using homing endonuclease/BstXI.

Reagents required:

- Homing endonucleases PI-SceI, I-CeuI
- 10x Buffers for homing endonucleases
- Restriction enzyme BstXI (and 10x Buffer)
- T4 DNA ligase (and 10x Buffer)
- E. coli* competent cells
- Antibiotics

**Step 2:** Insert preparation

Restriction reactions are carried out in 40 µl reaction volumes, using homing endonucleases PI-SceI (Donors) or I-CeuI (Acceptors) as recommended by the supplier (c.f. New England Biolabs catalog and others).

ACEMBL plasmid ( ≥ 0.5 µg) in ddH <sub>2</sub> O	32 µl
10x Restriction enzyme buffer	4 µl
10 mM BSA	2 µl
PI-SceI (Donors) or I-CeuI (acceptors)	2 µl

Reactions are then purified by PCR extraction kit or acidic ethanol precipitation, and next digested by BstXI according to the recommendations of the supplier.

HE digested DNA in ddH <sub>2</sub> O	32 µl
10x Restriction enzyme buffer	4 µl
10 mM BSA	2 µl
BstXI	2 µl

Gel extraction of insert(s):

Processed insert is then purified by agarose gel extraction using commercial kits (Qiagen, MachereyNagel etc). It is recommended to elute the extracted DNA in the minimal volume defined by the manufacturer.

**Step 3: Vector preparation**

Restriction reactions are carried out in 40  $\mu$ l reaction volumes, using homing endonucleases PI-SceI (Donors) or I-CeuI (Acceptors) as recommended by the supplier (c.f. New England Biolabs catalog and others).

ACEMBL plasmid ( $\geq 0.5 \mu\text{g}$ ) in ddH <sub>2</sub> O	33 $\mu$ l
10x Restriction enzyme buffer	4 $\mu$ l
10 mM BSA	2 $\mu$ l
PI-SceI (Donors) or I-CeuI (acceptors)	1 $\mu$ l

Reactions are then purified by PCR extraction kit or acidic ethanol precipitation, and next treated with intestinal alkaline phosphatase according to the recommendations of the supplier.

HE digested DNA in ddH <sub>2</sub> O	17 $\mu$ l
10x Alkaline phosphatase buffer	2 $\mu$ l
Alkaline phosphatase	1 $\mu$ l

Gel extraction of vector:

Processed vector is then purified by agarose gel extraction using commercial kits (Qiagen, MachereyNagel etc). It is recommended to elute the extracted DNA in the minimal volume defined by the manufacturer.

**Step 4: Ligation**

Ligation reactions are carried out in 20  $\mu$ l reaction volumes:

HE/Phosphatase treated vector (gel extracted)	4 $\mu$ l
HE/BstXI treated insert (gel extracted)	14 $\mu$ l
10x T4 DNA Ligase buffer	2 $\mu$ l
T4 DNA Ligase	0.5 $\mu$ l

Ligation reactions are performed at 25°C for 1h or at 16°C overnight.

**Step 5: Transformation**

Mixtures are next transformed into competent cells following standard transformation procedures.

Reactions for pACE and pACE2 derivatives are transformed into standard *E. coli* cells for cloning (such as TOP10, DH5 $\alpha$ , HB101) and after recovery plated on agar containing ampicillin (100  $\mu\text{g}/\text{ml}$ ) or tetracycline (25  $\mu\text{g}/\text{ml}$ ), respectively.

Reactions for Donor derivatives are transformed into *E. coli* cells expressing the *pir* gene (such as BW23473, BW23474, or PIR1 and PIR2, Invitrogen) and plated on agar containing chloramphenicol (25  $\mu\text{g}/\text{ml}$ , pDC), kanamycin (50  $\mu\text{g}/\text{ml}$ , pDK), and spectinomycin (50  $\mu\text{g}/\text{ml}$ , pDS).

**Step 6: Plasmid analysis**

Plasmids are cultured and correct clones selected based on specific restriction digestion and DNA sequencing of the inserts.

**Note:** Integration can likewise be performed by sequence and ligation independent cloning. It is recommended to carry out linearization of the vector by digestion with HE, if heterologous genes are already present, to avoid PCR amplifications over encoding regions. The fragment to be inserted is generated by PCR amplification resulting in a PCR fragment containing a 20-25 base pair stretch at its 5' end that is identical to the corresponding DNA sequence present at the HE site counted from the site of cleavage towards 5' (site of cleavage is position -4). At the 3' end of the PCR fragment, the homology region is 20-25 base pairs counted from the site of cleavage towards 3'.

## D.2. Cre-LoxP reaction of Acceptors and Donors

Cre recombinase is a member of the integrase family (Type I topoisomerase from bacteriophage P1). It recombines a 34 bp loxP site in the absence of accessory protein or auxiliary DNA sequence. The loxP site is comprised of two 13 bp recombinase-binding elements arranged as inverted repeats which flank an 8 bp central region where cleavage and ligation reaction occur.

The site-specific recombination mediated by Cre recombinase involves the formation of a Holliday junction (HJ). The recombination events catalyzed by Cre recombinase are dependent on the location and relative orientation of the loxP sites. Two DNA molecules, for example an Acceptor and a Donor plasmid, containing single loxP sites will be fused. Furthermore, the Cre recombination is an equilibrium reaction with 20-30% efficiency in recombination. This provides useful options for multi-gene combinations for multi-protein complex expression.

### Illustration 6: LoxP imperfect inverted repeat

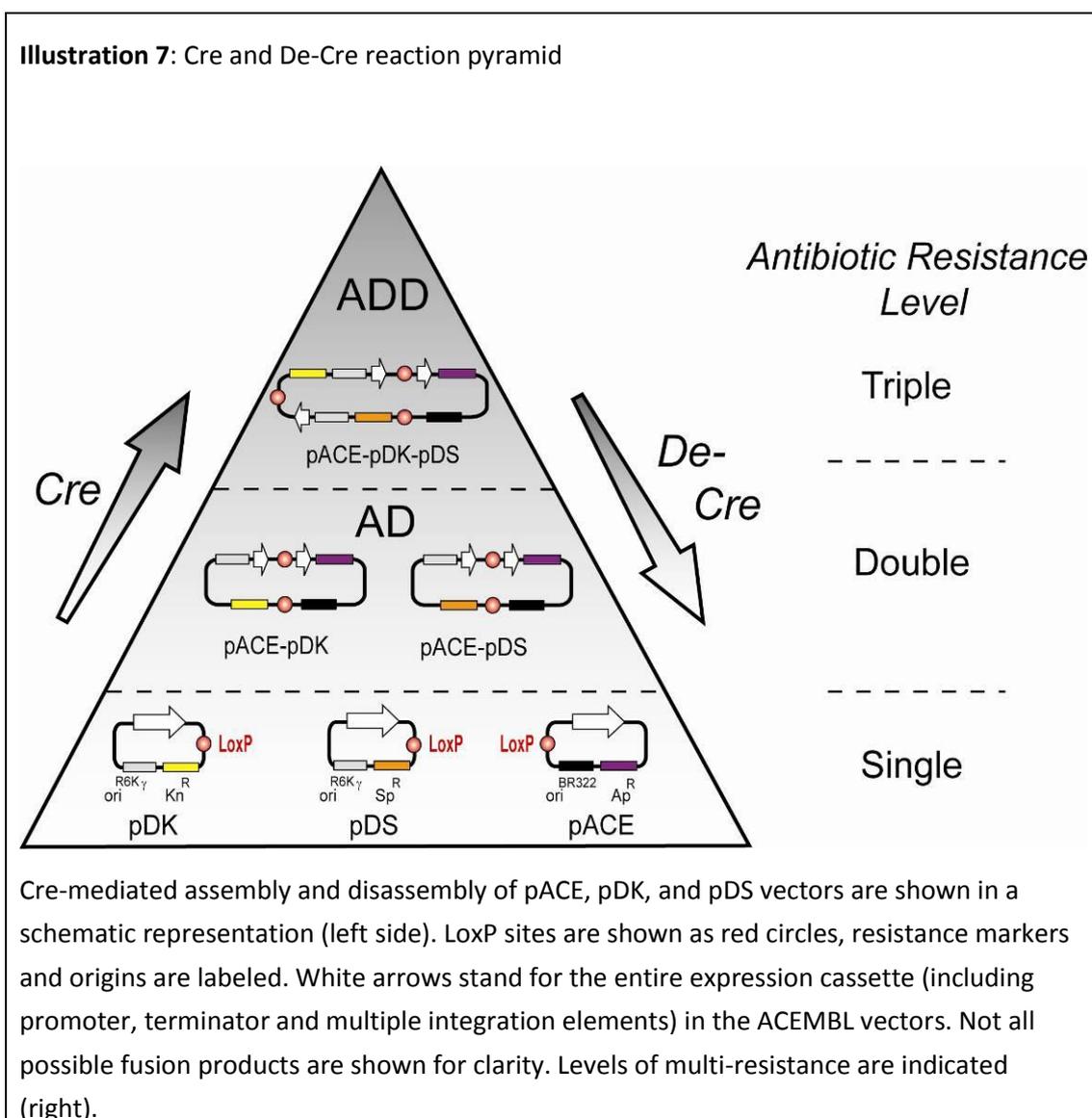
13bp	8bp	13bp
5'...ATAACTTCGTATA	GCATACAT	TATACGAAGTTAT...3'
3'...TATTGAAGCATAT	CGTATGTA	ATATGCTTCAATA...5'
inverted repeat	Spacer	inverted repeat

In a reaction where several DNA molecules such as Donors and Acceptors are incubated with Cre recombinase, the fusion/excision activity of the enzyme will result in an equilibrium state where single vectors (educt vectors) and all possible fusions coexist. Donor vectors can be used with Acceptors and/or Donors, likewise

for Acceptor vectors. Higher order fusions are also generated where more than two vectors are fused. This is shown schematically in Illustration 6.

The fact that Donors contain a conditional origin of replication that depends on a *pir*<sup>+</sup> (*pir* positive) background now allows for selecting out from this reaction mix all desired Acceptor-Donor(s) combinations. For this, the reaction mix is used to transform to *pir* negative strains (TOP10, DH5 $\alpha$ , HB101 or other common laboratory cloning strains). Then, Donor vectors will act as suicide vectors when plated out on agar containing the antibiotic corresponding to the Donor encoded resistance marker, unless fused with an Acceptor. By using agar with the appropriate combinations of antibiotics, all desired Acceptor-Donor fusions can be selected for.

Using this approach, fusion vectors of 25 kb and larger have generated. In stability tests (serial passaging for more than 60 generations), even such large plasmids proved to be stable as checked by restriction mapping, even if only one of the antibiotics corresponding to the encoded resistance markers was provided in the growth medium.

**Illustration 7:** Cre and De-Cre reaction pyramid

### D.2.1. Cre-LoxP fusion of Acceptors and Donors

This protocol is designed for generating multi-gene fusions from Donors and Acceptors by Cre-LoxP reaction.

#### Reagents:

Cre recombinase (from NEB or self made)

Standard *E. coli* competent cells (*pir<sup>-</sup>* strain)

Antibiotics

96well microtiter plates

12 well tissue-culture plates (or petri dishes) w. agar/antibiotics

## LB medium

1. For a 20µl Cre reaction, mix 1-2 µg of each educt in approximately equal amounts. Add ddH<sub>2</sub>O to adjust the total volume to 16-17 µl, then add 2 µl 10x Cre buffer and 1-2µl Cre recombinase.
2. Incubate Cre reaction at 37°C (or 30°C) for 1 hour.
3. Optional: load 2-5 µl of Cre reaction on an analytical agarose gel for examination.

*Heat inactivation at 70°C for 10 minutes before the gel loading is strongly recommended.*

4. For chemical transformation, mix 10-15µl Cre reaction with 200 µl chemical competent cells. Incubate the mixture on ice for 15-30 minutes. Then perform heat shock at 42°C for 45-60 s.

*Up to 20 µl Cre reaction (0.1 volumes of the chemical competent cell suspension) can be directly transformed into 200 µl chemical competent cells.*

For electrotransformation, up to 2 µl Cre reaction could be directly mixed with 100 µl electrocompetent cells, and transformed by using an electroporator (e.g. BIORAD E. coli Pulser) at 1.8-2.0 kV.

*Larger volume of Cre reaction must be desalted by ethanol precipitation or PCR purification column before electrotransformation. The desalted Cre reaction mix should not exceed 0.1 volumes of the electrocompetent cell suspension.*

*The cell/DNA mixture could be immediately used for electrotransformation without prolonged incubation on ice.*

5. Add up to 400 µl of LB media (or SOC media) per 100 µl of cell/DNA suspension immediately after the transformation (heat shock or electroporation).
6. Incubate the suspension in a 37°C shaking incubator overnight or for at least 4 hours (recovery period).
7. Plate out the recovered cell suspension on agar containing the desired combination of antibiotics. Incubate at 37°C overnight.
8. Clones from colonies present after overnight incubation can be verified by restriction digestion at this stage (refer to steps 12-16).

*Especially in the case that only one multifusion plasmid is desired.*

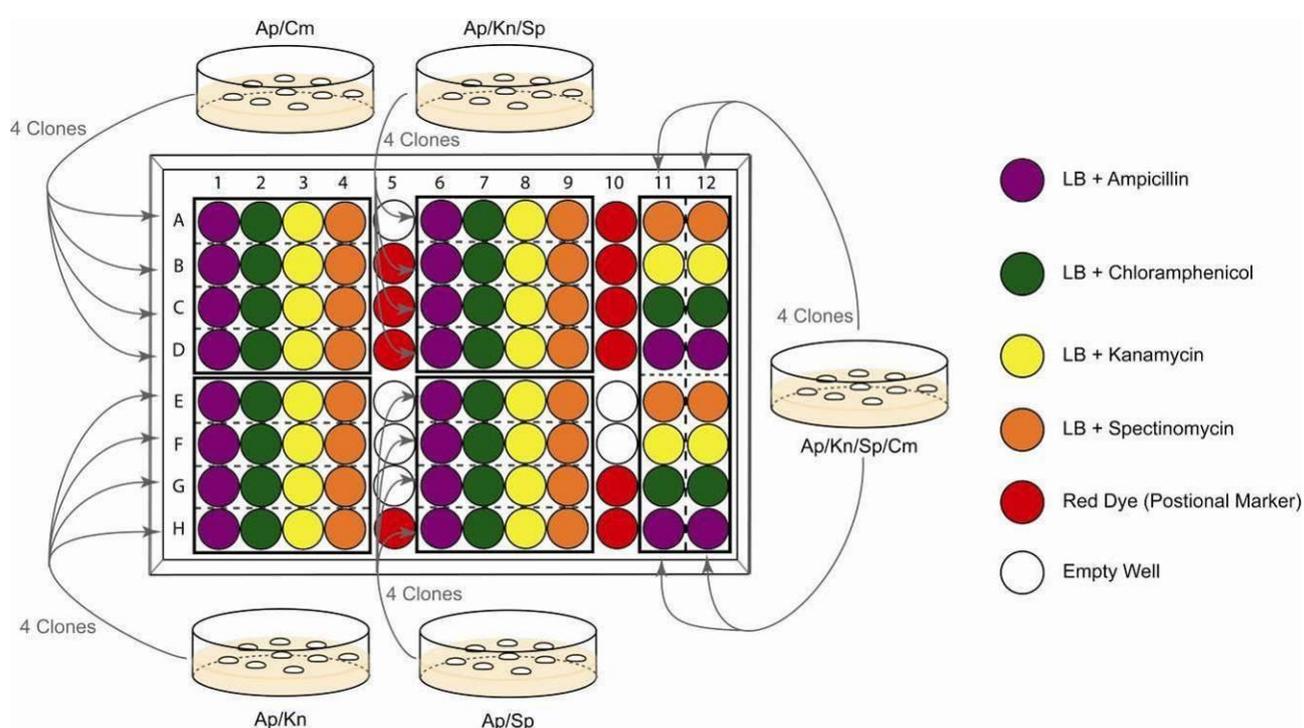
For further selection by single antibiotic challenges on a 96 well microtiter plate, continue to step 9.

Several to many different multifusion plasmid combinations can be processed and selected on one 96 well microtiter plate in parallel.

9. For 96 well antibiotic tests, inoculate four colonies from each agar plate with different antibiotic combination into ~500  $\mu$ l LB media without antibiotics. Incubate the cell cultures in a 37°C shaking incubator for 1-2 hours.
10. During the incubation of colonies, fill a 96 well microtiter plate with 150  $\mu$ l antibiotic-containing LB media (following Illustration 7). It is recommended to add coloured dye (positional marker) in the wells indicated.

A typical arrangement of the solutions, which is used for parallel selections of multifusion plasmids, is shown in Illustration 7. The concept behind the 96 well plate experiment is that every cell suspension from single colonies needs to be challenged by all four single antibiotics for unambiguous interpretation.

**Illustration 8:** 96 well analysis of Cre assembly



11. Add 1  $\mu$ l aliquots of pre-incubated cell culture (Step 9) to the corresponding wells. Then incubate the inoculated 96 well microtiter plate in a 37°C shaking incubator overnight at 180-200 rpm.

*Recommended: use parafilm to wrap the plate to avoid drying out.*

*The remainder of the pre-incubated cell cultures could be kept at 4°C for further inoculations if necessary.*

12. Select transformants containing desired multifusion plasmids based on antibiotic resistance, according to the combination of dense (positive) and clear (no growth) cell microcultures from each colony. Inoculate 10-20  $\mu$ l cell

culture into 10 ml LB media with corresponding antibiotics. Incubate in a 37°C shaking incubator overnight.

13. Centrifuge the overnight cell cultures at 4000g for 5-10 minutes. Purify plasmid from the resulting cell pellets with common plasmid miniprep kits, according to manufacturers' recommendation.
14. Determine the concentrations of purified plasmid solutions by using UV absorption spectroscopy (e.g. by using a NanoDrop™ 1000 machine).
15. Digest approx. 0.5-1 µg of the purified plasmid solution in a 20 µl restriction digestion with appropriate endonuclease(s). Incubate under recommended reaction condition for ~2 hours.
16. Use 5-10 µl of the digestion for analytical agarose (0.8-1.2%) gel electrophoresis. Verify plasmid integrity by comparing the experimental restriction pattern to a restriction pattern predicted *in silico* (e.g. by using program VectorNTI from Invitrogen or similar programs).

#### D.2.2. Deconstruction of fusion vectors by Cre

The protocol is suitable for releasing any single educt from multifusion constructs (deconstruction). This is achieved by Cre-LoxP reaction, transformation and plating on agar with appropriately reduced antibiotic resistance level (c.f. Illustration 6). In the liberated educt entity, encoding genes can be modified and diversified. Then, the diversified construct is resupplied by Cre-LoxP reaction (C.2.1.).

#### Reagents:

Cre recombinase (and 10x Buffer)

*E. coli* competent cells

(*pir*<sup>+</sup> strains, *pir*<sup>-</sup> strains could be used only when partially deconstructed Acceptor-Donor fusions are desired).

Antibiotics

1. Incubate approx. 1 µg multifusion plasmid with 2 µl 10x Cre buffer, 1-2 µl Cre recombinase, add ddH<sub>2</sub>O to adjust the total reaction volume to 20 µl.
2. Incubate this Cre deconstruction reaction mixture at 30°C for 1-4 hour.
3. Optional: load 2-5 µl of the reaction on an analytical agarose gel for examination.

*Heat inactivation at 70°C for 10 minutes before the gel loading is strongly recommended.*

4. For chemical transformation, mix 10-15µl De-Cre reaction with 200 µl chemical competent cells. Incubate the mixture on ice for 15-30 minutes. Then perform heat shock at 42°C for 45-60 s.

*Up to 20 µl De-Cre reaction (0.1 volumes of the chemical competent cell suspension) can be directly transformed into 200 µl chemical competent cells.*

For electrotransformation, up to 2 µl De-Cre reaction can be directly mixed with 100 µl electrocompetent cells, and transformed by using an electroporator (e.g. BIORAD *E. coli* Pulser) at 1.8-2.0 kV.

*Larger volume of De-Cre reaction must be desalted by ethanol precipitation or PCR purification column before electrotransformation. The desalted De-Cre reaction mix should not exceed 0.1 volumes of the electrocompetent cell suspension.*

*The cell/DNA mixture could be immediately used for electrotransformation without prior incubation on ice.*

5. Add up to 400 µl of LB media (or SOC media) per 100 µl of cell/DNA suspension immediately after the transformation (heat shock or electroporation).
6. Incubate the suspension in a 37°C shaking incubator (recovery).

*For recovery of partially deconstructed double/triple fusions, incubate the suspension in a 37°C shaking incubator for 1 to 2 hours.*

*For recovery of individual educts such as single ACEMBL vectors from pACKS plasmid, incubate the suspension in a 37°C shaking incubator overnight or for at least 4 hours.*

7. Plate out the recovered cell suspension on agar containing the desired (combination of) antibiotic(s). Incubate at 37°C overnight.
8. Colonies after overnight incubation might be verified directly by restriction digestion at this stage (refer to steps 12-16).

*Especially recommended in the case that only one single educt or partially deconstructed multifusion plasmid is desired.*

For further selection by single antibiotic challenge on a 96 well microtiter plate, continue with step 9.

*Several different single educts/partially deconstructed multifusion plasmids can be processed and selected on one 96 well microtiter plate in parallel.*

9. For 96 well analysis, inoculate four colonies each from agar plates containing a defined set of antibiotics into ~500 µl LB media without antibiotics. Incubate the cell cultures in a 37°C shaking incubator for 1-2 hours.
10. During the incubation of colonies, fill a 96 well microtiter plate with 150 µl antibiotic-containing LB media or coloured dye (positional marker) in the corresponding wells.

*Refer to Illustrations 7 and 12 for the arrangement of the solutions in the wells, which are used for parallel selection of single educts or partially deconstructed multifusion plasmids. The concept is that every cell suspension from a single colony needs to be challenged by all four antibiotics separately for unambiguous interpretation.*

11. Add 1  $\mu$ l aliquots from the pre-incubated cell cultures (Step 9) into the corresponding wells. Then incubate the 96 well microtiter plate in a 37°C shaking incubator overnight at 180-200 rpm.

*Recommended: use parafilm to wrap the plate to prevent desiccation.*

*The remainder of the pre-incubated cell cultures can be kept in 4°C fridge for further inoculations if necessary.*

12. Select transformants containing desired single educts or partially deconstructed multifusion plasmids according to the combination of dense (growth) and clear (no growth) cell cultures from each colony. Inoculate 10-20  $\mu$ l cell cultures into 10 ml LB media with corresponding antibiotic(s). Incubate in a 37°C shaking incubator overnight.
13. Centrifuge the overnight cell cultures at 4000g for 5-10 minutes. Purify plasmid from cell pellets with common plasmid miniprep kits, according to manufacturers' information.
14. Determine the concentrations of purified plasmid solutions by using UV absorption spectroscopy (e.g. NanoDrop™ 1000).
15. Digest approx. 0.5-1  $\mu$ g of the purified plasmid solution in a 20  $\mu$ l restriction digestion (with 5-10 unit endonuclease). Incubate under recommended reaction condition for ~2 hours.
16. Use 5-10  $\mu$ l of the digestion for analytical agarose gel (0.8-1.2%) electrophoresis. Verify the plasmid integrity by comparing the actual restriction pattern to predicted restriction pattern *in silico* (e.g. by using VectorNTI, Invitrogen, or any other similar program).
17. Optional: Possibly, a deconstruction reaction is not complete but yields partially deconstructed fusions which still retain entities to be eliminated. In this case, we recommend to pick these partially deconstructed fusions containing and perform a second round of Cre deconstruction reaction (repeat steps 1-8) by using this construct as starting material.

*In our hands, two sequential deconstruction reactions were always sufficient to recover all individual modules, for instance all four single ACEMBL vectors from a pACKS plasmid. Liberation of single educts from double/triple fusions were found to be often more efficient than from quadruples such as the pACKS plasmid of the system kit (Section E).*

### **D.3. Co-expression by Co-transformation**

Protein complexes can be expressed also from two separate vectors co-transformed in expression strains. The co-transformed vectors can have the same or different origins of replication, however, they must encode different resistance markers.

Plasmids pACE1 (ampicillin resistance marker) and pACE2 (tetracycline resistance marker) have both a ColE1 derived replicon and can therefore be used with all common expression strains. pACE1 and pACE2 derivatives (also including fused Donors if needed) can be co-transformed into expression strains, and double transformants selected for by plating on agar plates containing both ampicillin and tetracyclin antibiotics.

Transformations are carried out by using standard transformation protocols.

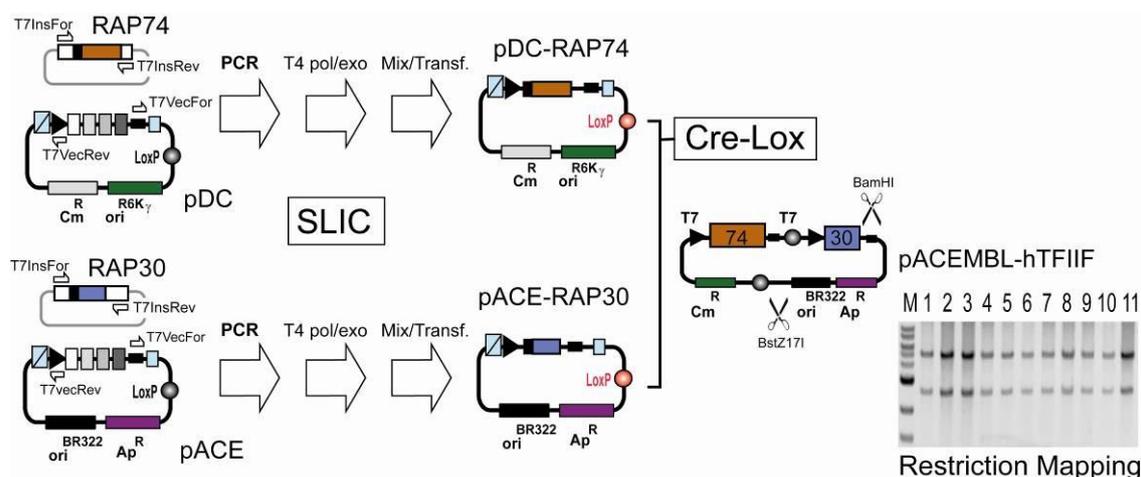
## E. ACEMBL multi-gene combination: Examples

Examples of multi-protein expressions by ACEMBL illustrating the gene combination procedures detailed in Section C are shown in the following. Reactions presented were carried out manually following the protocols provided, and also on a Tecan Freedom Evoll 200 robot with adapted protocols. These and other examples have been published in Bieniossek et al., *Nat. Methods* **6**, 447 (2009) plus supplementary material.

### E.1. SLIC cloning into ACEMBL vectors: human TFIIIF

Genes encoding full-length human RAP74 with a C-terminal oligo-histidine tag and full-length human RAP30 were amplified from pET-based plasmid template<sup>9</sup> following the protocols described above. Linearized vector backbones were generated by PCR amplification from pACE1 and pDC. SLIC following Protocol 1 (Section D), yielded pACE1-RAP30 and pDC-RAP74his (Fig 8). These plasmids were fused by Cre-LoxP reaction (Section C). Results from restriction mapping by BstZ171/BamHI double digestion of 11 double resistant (Cm, Ap) colonies are shown by a gel section from 1% E-gel electrophoresis (M: NEB 1kb DNA marker). All clones tested showed the expected pattern (5.0 + 2.8 kb). One clone was transformed in BL21(DE3) cells. Expression and purification by Ni<sup>2+</sup>-capture and S200 chromatography resulted in human TFIIIF complex.

**Illustration 9:** ACEMBLing TFIIIF (Image courtesy of Dr. I. Berger, EMBL Grenoble).



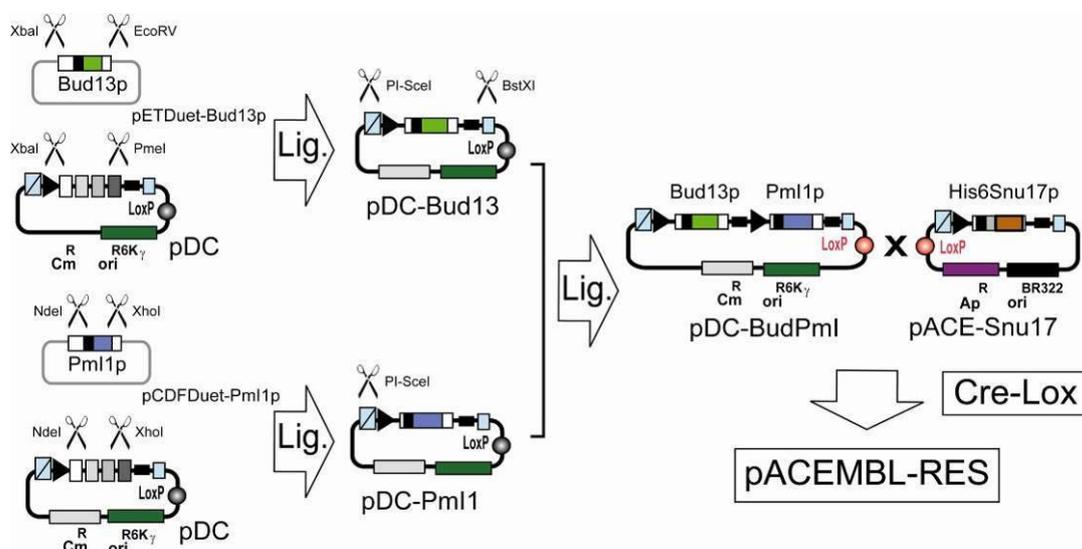
<sup>9</sup> Gaiser, F., Tan, S. and Richmond, T.J. *J. Mol. Biol.* **302**, 1119-1127 (2000).

## E.2. The Homing endonuclease/BstXI module: yeast RES complex

Plasmids pCDFDuet-Pml1p, pRSFDuet-Snu17p-NHis and pETDuet-Bud13p, encoding for yeast proteins (all full-length) Pml1p, Snu17p and Bud13p, respectively, were a kind gift from Dr. Simon Trowitzsch and Dr. Markus Wahl (MPI Göttingen). Snu17p contains a six-histidine tag fused to its N-terminus. The gene encoding for His6-tagged Snu17p was excised from pRSFDuet-Snu17p-NHis by NcoI/XhoI restriction, and ligated into a NcoI/XhoI digested pACE1 construct (containing an unrelated gene between the NcoI and XhoI sites) resulting in pACE1-Snu17. The gene encoding for Bud13p was liberated from pETDuet-Bud13p by restriction digestion with XbaI and EcoRV, and placed into XbaI/PmeI digested pDC resulting in pDC-Bud13. The gene encoding Pml1p was liberated from pCDFDuet-Pml1p by restriction digestion with NdeI and XhoI, and placed into NdeI/XhoI digested pDC resulting in pDC-Pml1. Next, the expression cassette for Bud13p was liberated from pDC-Bud13 by digestion with PI-SceI and BstXI. The liberated fragment was inserted into PI-SceI digested and alkaline phosphatase treated pDC-Pml1p resulting in pDC-Bud13p-Pml1p.

pACE-Snu17 and pDC-BudPml were then fused by Cre-LoxP reaction and selected for by plating on agar plates containing ampicillin and chloramphenicol. Fusion plasmids were transformed into BL21(DE3) cells. Expression and purification by Ni<sup>2+</sup>-capture and S200 size exclusion chromatography resulted in the trimeric RES complex.

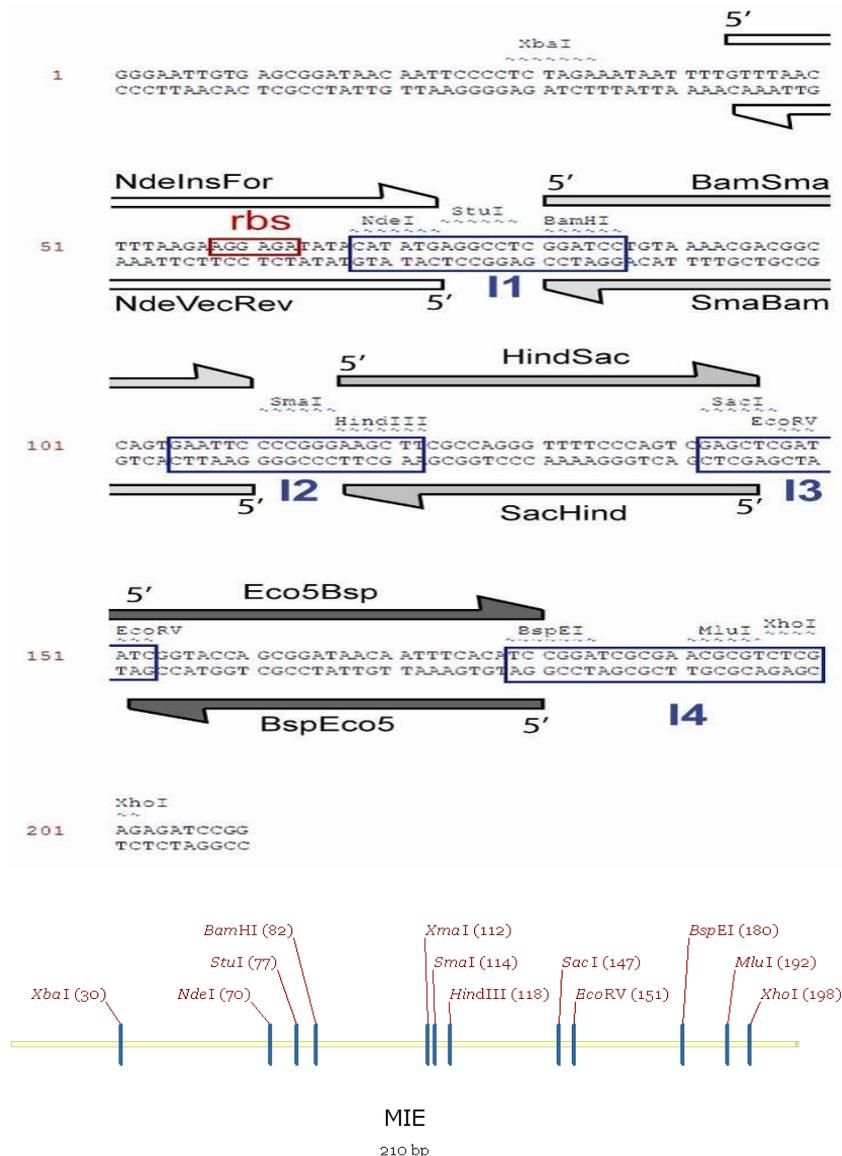
**Illustration 10:** The HE/BstXI multiplication module.



## F. Appendix

### F.1. DNA sequence of MIE

Below are the sequence and map of the MIE fragment between T7/lac promoter and T7 terminator in ACEMBL vectors. Forward and reverse primers for sequencing can be standard vector primers for T7 and lac. Adaptor primer sequences (c.f. Table I) are indicated. DNA sequences in these homology regions contain tried-and-tested sequencing primers<sup>10</sup>. Sites of insertion (I1-I4) are shown.. The adaptor sequences, and probably any sequence in the homology regions, can be used as adaptors for multifragment insertions. The ribosome binding site present in the MIE (rbs) is boxed in red.



<sup>10</sup> Tan S. et al. *Protein Expr. Purif.* **40**, 385, (2005).

## F.2. DNA sequences of ACEMBL vectors

### F.2.1. pACE1

```

1   GGTACCGCGG CCGCGTAGAG GATCTGTTGA TCAGCAGTTC AACCTGTTGA
51  TAGTACTTCG TTAATACAGA TGTAGGTGTT GGCACCATGC ATAACTATAA
101 CGGTCCCTAAG GTAGCGACCT AGGTATCGAT AATACGACTC ACTATAGGGG
151 AATTGTGAGC GGATAACAAT TCCCCTCTAG AAATAATTTT GTTTAACTTT
201 AAGAAGGAGA TATACATATG AGGCCTCGGA TCCTGTAAAA CGACGGCCAG
251 TGAATTCCCC GGAAGCTTC GCCAGGGTTT TCCCAGTCGA GCTCGATATC
301 GGTACCAGCG GATAACAATT TCACATCCGG ATCGCGAACG CGTCTCGAGA
351 GATCCGGCTG CTAACAAAGC CCGAAAGGAA GCTGAGTTGG CTGCTGCCAC
401 CGCTGAGCAA TAACTAGCAT AACCCCTTGG GGCTCTAAA CGGCTCTTGA
451 GGGGTTTTTT GGTTTAAACC CATCTAATTG GACTAGTAGC CCGCCTAATG
501 AGCGGGCTTT TTTTAAATTC CCCTATTTGT TTATTTTCT AAATACATTC
551 AAATATGTAT CCGCTCATGA GACAATAACC CTGATAAATG CTTCAATAAT
601 ATTGAAAAAG GAAGAGTATG AGTATTCAAC ATTTCCGTGT CGCCCTTATT
651 CCCTTTTTTG CGGCATTTTG CCTTCCTGTT TTTGCTCACC CAGAAACGCT
701 CGTGAAAGTA AAAGACGCAG AGGACCAATT GGGGGCACGA GTGGGATACA
751 TAGAACTGGA CTTGAATAGC GGTAAAATCC TTGAGAGTTT TCGCCCTGAA
801 GAGCGTTTTT CAATGATGAG CACTTTCAAA GTTCTGCTAT GTGGAGCAGT
851 ATTATCCCGT GTAGATGCGG GGCAAGAGCA ACTCGGACGA CGAATACACT
901 ATTCGCAGAA TGACTTGGTT GAATACTCCC CAGTGACAGA AAAGCACCTT
951 ACGGACGGAA TGACGGTAAG AGAATTATGT AGTGCCGCCA TAACGATGAG
1001 TGATAACACT GCGGCGAACT TACTTCTGAC AACCATCGGT GGACCGAAGG
1051 AATTAACCGC TTTTTTGCAC AATATGGGAG ACCATGTAAC TCGCCTTGAC
1101 CGTTGGGAAC CAGAACTGAA TGAAAGCCATA CCAAACGACG AGCGAGACAC
1151 CACAATGCCT GCGGCAATGG CAACAACAT ACGCAAACTA TTAAC TGGCG
1201 AACTACTTAC TCTGGCTTCA CGGCAACAAT TAATAGACTG GCTTGAAGCG
1251 GATAAAGTTG CAGGACCACT ACTGCGTTCG GCACTTCCTC CTGGCTGGTT
1301 TATTGCTGAT AAATCTGGGG CAGGAGAGCG TGGTTCACGG GGTATCATTG
1351 CCGCACTTGG ACCAGATGGT AAGCCTTCCC GTATCGTAGT TATCTACACG
1401 ACGGGTAGTC AGGCAACTAT GGACGAACGA AATAGACAGA TTGCTGAAAT
1451 AGGGGCTTCA CTGATTAAGC ATTGGTAAAC CGATACAATT AAAGGCTCCT
1501 TTTGGAGCCT TTTTTTTTGG ACGGACCGGT AGAAAAGATC AAAGGATCTT
1551 CTTGAGATCC TTTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA
1601 CCACCGCTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT
1651 TTTTCCGAAG GTAAC TGGCT TCAGCAGAGC GCAGATACCA AATACTGTCC
1701 TTCTAGTGTA GCCGTAGTTA GGCCACCACT TCAAGAACTC TGTAGCACCG
1751 CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG
1801 CGATAAGTCG TGTCTTACCG GGTGGACTC AAGACGATAG TTACCGGATA
1851 AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGCACACA GCCCAGCTTG
1901 GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA
1951 AAGCGCCACG CTTCCCGAAG GGAGAAAAGC GGACAGGTAT CCGGTAAGCG
2001 GCAGGGTCGG AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC
2051 TGGTATCTTT ATAGTCCTGT CGGGTTTCGC CACCTCTGAC TTGAGCGTCG
2101 ATTTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGAAA AACGCCAGCA
2151 ACGCGGCCTT TTTACGGTTC CTGGCCTTTT GCTGGCCTTT TGCTCACATG
2201 TTCTTTCCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCTT
2251 TGAGTGAGCT GATACCGCTC GCCGCAGCCG AACGACCGAG CGCAGCGAGT
2301 CAGTGAGCGA GGAAGCGGAA GAGCGCCTGA TCGGGTATTT TCTCCTTACG
2351 CATCTGTGCG GTATTTTACA CCGCAATGGT GCACTCTCAG TACAATCTGC
2401 TCTGATGCCG CATAGTTAAG CCAGTATACA CTCCGCTATC GCTACGTGAC
2451 TGGGT CATGG CTGCGCCCCG ACACCCGCCA ACACCCGCTG ACGCGCCCTG
2501 ACGGGCTTGT CTGCTCCCGG CATCCGCTTA CAGACAAGCT GTGACCGTCT
2551 CCGGGAGCTG CATGTGTCAG AGGTTTTTAC CGTCATCACC GAAACGCGCG
2601 AGGCAGGGGG AATTCCAGAT AACTTCGTAT AATGTATGCT ATACGAAGTT
2651 AT

```

## F.2.2. pACE2

```

1   ATGAAATCTA ACAATGCGCT CATCGTCATC CTCGGCACCG TCACCTGGA
51  TGCTGTAGGC ATAGGCTTGG TTATGCCGGT ACTGCCGGGC CTCTTGCGGG
101 ATATCGTCCA TTCCGACAGC ATCGCCAGTC ACTATGGCGT GCTGTAGCG
151 CTATATGCGT TGATGCAATT TCTATGCGCA CCCGTTCTCG GAGCACTGTC
201 CGACCGCTTT GGCCGCCGCC CAGTCCTGCT CGCTTCGCTA CTTGGAGCCA
251 CTATCGACTA CGCGATCATG GCGACCACAC CCGTCCTGTG GATTCTCTAC
301 GCCGGACGCA TCGTGGCCGG CATCACCGGC GCCACAGGTG CGGTTGCTGG
351 GCCTATATC  GCCGACATCA CCGATGGGGA AGATCGGGCT CGCCACTTCG
401 GGCATGAG  CGCTTGTTTC GCGTGGGTA TGGTGGCAG  CCCCGTGGCC
451 GGGGGACTGT TGGGCGCCAT CTCCTTACAT GCACCATTCC TTGCGGCGGC
501 GGTGCTCAAC GGCTCAACC  TACTACTGGG CTGCTTCCTA ATGCAGGAGT
551 CGCATAAGGG AGAGCGCCGA CCCATGCCCT TGAGAGCCTT CAACCCAGTC
601 AGCTCCTTCC GGTGGGCGCG GGGCATGACT ATCGTCGCCG CACTTATGAC
651 TGTCTTCTTT ATCATGCAAC TCGTAGGACA GGTGCCGGCA GCGCTCTGGG
701 TCATTTTCGG CGAGGACCGC TTTGCTGGA GCGGACGAT GATCGGCCCTG
751 TCGCTTGCGG TATTCGGAAT CTTGCACGCC CTCGCTCAAG CCTTCGTCAC
801 TGGTCCC GCCAACGTT TCGGCGAGAA GCAGGCCATT ATCGCCGGCA
851 TGGCGGCCGA CGCGCTGGGC TACGTCTTGC TGGCGTTCGC GACGCGAGGC
901 TGGATGGCCT TCCCCATTAT GATTCTTCTC GCTTCCGGCG GCATCGGGAT
951 GCCCGCGTTG CAGGCCATGC TGTCCAGGCA GGTAGATGAC GACCATCAGG
1001 GACAGCTTCA AGGATCGCTC GCGGCTCTTA CCAGCCTAAC TTCGATCATT
1051 GGACCGCTGA TCGTCACGGC GATTTATGCC GCCTCGGCCGA GCACATGGAA
1101 CGGGTTGGCA TGGATTGTAG GCGCCGCCCT ATACCTTGTC TGCCTCCCCG
1151 CGTTGCGTCG CGGTGCATGG AGCCGGGCCA CCTCGACCTG AACCGATAACA
1201 ATTAAAGGCT CTTTTTGAG CTTTTTTTTT TGGACGGACC GGTAGAAAAAG
1251 ATCAAAGGAT CTTCTTGAGA TCTTTTTTTT CTGCGCGTAA TCTGCTGCTT
1301 GCAAACAAAA AAACCACCGC TACCAGCGGT GTTTTGTTT GCGGATCAAG
1351 AGTACCAAC TCTTTTCCG AAGGTAAC TGCTCAGCAG AGCGCAGATA
1401 CCAAATACTG TCCTTCTAGT GTAGCCGTAG TTAGGCCACC ACTTCAAGAA
1451 CTCTGTAGCA CCGCCTACAT ACCTCGCTCT GCTAATCCTG TTACCAGTGG
1501 CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA CTCAAGACGA
1551 TAGTTACCGG ATAAGGCGCA GCGGTCGGGC TGAACGGGGG GTTCGTGCAC
1601 ACAGCCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC
1651 GTGAGCTATG AGAAAGCGCC ACGCTTCCG AAGGGAGAAA GCGGACAGG
1701 TATCCGGTAA GCGGCAGGGT CGGAACAGGA GAGCGCACGA GGGAGCTTCC
1751 AGGGGGAAAC GCCTGGTATC TTTATAGTCC TGTCGGGTTT CGCCACCTCT
1801 GACTTGAGCG TCGATTTTTG TGATGCTCGT CAGGGGGGGCG GAGCCTATGG
1851 AAAAACGCCA GCAACGCGGC CTTTTTACGG TTCTGGCCCT TTTGCTGGCC
1901 TTTTGCTCAC ATGTTCTTTC CTGCGTTATC CCCTGATTCT GTGGATAACC
1951 GTATTACCGC CTTTGAGTGA GCTGATAACG CTCGCCGAG CCGAACGACC
2001 GAGCGCAGCG AGTCAGTGAG CGAGGAAGCG GAAGAGCGCC TGATGCGGTA
2051 TTTTCTCCTT ACGCATCTGT GCGGTATTTT ACACCGCAAT GGTGCACTCT
2101 CAGTACAATC TGCTCTGATG CCGCATAGTT AAGCCAGTAT ACACTCCGCT
2151 ATCGCTACGT GACTGGGTCA TGGCTGCGCC CCGACACCCG CCAACACCCG
2201 CTGACGCGCC CTGACGGGCT TGTCGCTCC CGGCATCCGC TTACAGACAA
2251 GCTGTGACCG TCTCCGGGAG CTGCATGTGT CAGAGGTTTT CACCGTCATC
2301 ACCGAAACGC GCGAGGCAGG GGGAAATCCA GATAACTTCG TATAATGTAT
2351 GCTATACGAA GTTATGGTAC CGCGGCCCGG TAGAGGATCT GTTGATCAGC
2401 AGTTCAACCT GTTGATAGTA CTTGTTAAT ACAGATGTAG GTGTTGGCAC
2451 CATGCATAAC TATAACGGTC CTAAGGTAGC GACCTAGGTA TCGATAATAC
2501 GACTCACTAT AGGGGAATTG TGAGCGGATA ACAATTCCCC TCTAGAAATA
2551 ATTTTGTTTA ACTTTAAGAA GGAGATATAC ATATGAGGCC TCGGATCCTG
2601 TAAAACGACG GCCAGTGAAT TCCCCGGGAA GCTTCGCCAG GTTTTTCCCA
2651 GTCGAGCTCG ATATCGGTAC CAGCGGATAA CAATTCACA TCCGATCGC
2701 GAACGCGTCT CGAGAGATCC GGCTGCTAAC AAAGCCCGAA AGGAAGCTGA
2751 GTTGGCTGCT GCCACCGCTG AGCAATAACT AGCATAACCC CTTGGGGCCT
2801 CTAAACGGGT CTTGAGGGGT TTTTGGTTT AAACCCATCT AATTGGACTA
2851 GTAGCCC GCC TAATGAGCGG GCTTTTTTTT AATTCCTTA TTTGTTTATT
2901 TTTCTAAATA CATTCAAATA TGTATCCGCT CATGAGACAA TAACCTGTAT
2951 AAATGCTTCA ATAATATTGA AAAAGGAAGA GT

```

## F.2.3. pDC

```

1      ATCAACGTCT  CATTTCGCC  AAAAGTTGGC  CCAGATCTAT  GTCGGGTGCG
51     GAGAAAGAGG  TAATGAAATG  GCACCTAGGT  ATCGATAATA  CGACTCACTA
101    TAGGGGAATT  GTGAGCGGAT  AACCAATCCC  CTCTAGAAAT  AATTTTGTTT
151    AACTTTAAGA  AGGAGATATA  CATATGAGGC  CTCGGATCCT  GTAAAACGAC
201    GGCCAGTGAA  TTCCCCGGA  AGCTTCGCCA  GGGTTTTCCC  AGTCGAGCTC
251    GATATCGGTA  CCAGCGGATA  ACAATTTTAC  ATCCGGATCG  CGAACCGGTC
301    TCGAGAGATC  CGGCTGCTAA  CAAAGCCCGA  AAGGAAGCTG  AGTTGGCTGC
351    TGCCACCGCT  GAGCAATAAC  TAGCATAACC  CCTTGGGGCC  TCTAAACGGG
401    TCTTGAGGGG  TTTTTTGTT  TAAACCCATG  TGCCTGGCAG  ATAACCTCGT
451    ATAATGTATG  CTATACGAAG  TTATGGTACC  GCGGCCGCGT  AGAGGATCTG
501    TTGATCAGCA  GTTCAACCTG  TTGATAGTAC  GTACTAAGCT  CTCATGTTTC
551    ACGTACTAAG  CTCTCATGTT  TAACGTAATA  AGCTCTCATG  TTTAACGAAC
601    TAAACCCTCA  TGGCTAACGT  ACTAAGCTCT  CATGGCTAAC  GTACTAAGCT
651    CTCATGTTTC  ACGTACTAAG  CTCTCATGTT  TGAACAATAA  AATTAATATA
701    AATCAGCAAC  TTAATAGCC  TCTAAGGTTT  TAAGTTTTAT  AAGAAAAAAA
751    AGAATATATA  AGGCTTTTAA  AGCTTTTAA  GTTTAACGGT  TGTGGACAAC
801    AAGCCAGGGA  TGTAACGCAC  TGAGAAGCCC  TTAGAGCCTC  TCAAAGCAAT
851    TTTGAGTGAC  ACAGGAACAC  TTAACGGCTG  ACAGAATTAG  CTTACGCTG
901    CCGCAAGCAC  TCAGGGCGCA  AGGGCTGCTA  AAGGAAGCGG  AACACGTAGA
951    AAGCCAGTCC  GCAGAAACGG  TGCTGACCCC  GGATGAATGT  CAGCTGGGAG
1001   GCAGAATAAA  TGATCATATC  GTCAATTATT  ACCTCCACGG  GGAGAGCCTG
1051   AGCAAACCTG  CCTCAGGCAT  TTGAGAAAGCA  CACGGTCACA  CTGCTTCCGG
1101   TAGTCAATAA  ACCGGTAAAC  CAGCAATAGA  CATAAGCGGC  TATTTAACGA
1151   CCCTGCCCTG  AACCGACGAC  CGGGTCGAAT  TTGCTTTCGA  ATTTCTGCCA
1201   TTCATCCGCT  TATTATCACT  TATTCAGGCG  TAGCAACCAG  GCGTTTAAGG
1251   GCACCAATAA  CTGCCTTAAA  AAAATTACGC  CCCGCCCTGC  CACTCATCGC
1301   AGTACTGTTG  TAATTCATTA  AGCATTCTGC  CGACATGGAA  GCCATCACAA
1351   ACGGCATGAT  GAACCTGAAT  CGCCAGCGGC  ATCAGCACCT  TGTGCTTGTG
1401   CGTATAATAT  TTGCCCATGG  TGAAAACGGG  GGCGAAGAAG  TTGTCCATAT
1451   TGGCCACGTT  TAAATCAAAA  CTGGTGAAAC  TCACCCAGGG  ATTGGCTGAG
1501   ACGAAAAACA  TATTCTCAAT  AAACCTTTA  GGGAAATAGG  CCAGGTTTTT
1551   ACCGTAACAC  GCCACATCTT  GCGAATATAT  GTGTAGAAAC  TGCCGGAAT
1601   CGTCGTGGTA  TTCCTCCAG  AGCGATGAAA  ACGTTTCAGT  TTGCTCATGG
1651   AAAACGGTGT  AACAAGGGTG  AACACTATCC  CATATCACCA  CATATCACCA
1701   TTTCAATTGCC  ATACGGAATT  CCGATGAGC  ATTCATCAGG  CGGGCAAGAA
1751   TGTGAATAAA  GGCCGGATAA  AACTTGTGCT  TATTTTTCTT  TACGGTCTTT
1801   AAAAAGGCCG  TAATATCCAG  CTGAACGGTC  TGGTTATAGG  TACATTGAGC
1851   AACTGACTGA  AATGCCTCAA  AATGTTCTTT  ACGATGCCAT  TGGGATATAT
1901   CAACGGTGGT  ATATCCAGTG  ATTTTTTTCT  CCATTTTAGC  TTCCTTAGCT
1951   CCTGAAAATC  TCGATAACTC  AAAAAATACG  CCCGGTAGTG  ATCTTATTTT
2001   ATTATGGTGA  AAGTTGACC  CTCTTACGTG  CCGATCAACG  TCTCATTTTT
2051   GCCAAAAGTT  GGCCAG

```

## F.2.4. pDK

```

1 CTATGTCGGG TGCGGAGAAA GAGGTAATGA AATGGCACCT AGGTATCGAT
51 GGCTTTACAC TTTATGCTTC CGGCTCGTAT GTTGTGTGGA ATTGTGAGCG
101 GATAACAATT TCACACAGGA AACAGCTATG ACCATGATTA CGAATTTCTA
151 GAAATAAATTT TGTTTAACTT TAAGAAGGAG ATATACATAT GAGGCCCTCGG
201 ATCCTGTAAA ACGACGGCCA GTGAATFCCC CGGGAAGCTT CGCCAGGGTT
251 TTCCCAGTCG AGCTCGATAT CGGTACCAGC GGATAACAAT TTCACATCCG
301 GATCGCGAAC GCGTCTCGAG ACTAGTTCCG TTTAAACCCA TGTGCCCTGGC
351 AGATAACTTC GTATAATGTA TGCTATACGA AGTTATGGTA CGTACTAAGC
401 TCTCATGTTT CACGTACTAA GCTCTCATGT TTAACGTACT AAGCTCTCAT
451 GTTTAAACGAA CTAAACCCCTC ATGGCTAACG TACTAAGCTC TCATGGCTAA
501 CGTACTAAGC TCTCATGTTT CACGTACTAA GCTCTCATGT TTGAACAATA
551 AAATTAATAT AAATCAGCAA CTTAAATAGC CTCTAAGGTT TTAAGTTTTTA
601 TAAGAAAAAAA AAGAATATAT AAGGCTTTTAA AAGCTTTTAA GGTTTAACGG
651 TTGTGGACAA CAAGCCAGGG ATGTAACGCA CTGAGAAGCC CTTAGAGCCT
701 CTCAAAGCAA TTTTCAGTGA CACAGGAACA CTTAACGGCT GACAGAATTA
751 GCTTCACGCT GCCGCAAGCA CTCAGGGCGC AAGGGCTGCT AAAGGAAGCG
801 GAACACGTAG AAAGCCAGTC CGCAGAAACG GTGCTGACCC CGGATGAATG
851 TCAGCTACTG GGCTATCTGG ACAAGGGAAA ACGCAAGCGC AAAGAGAAAAG
901 CAGGTAGCTT GCAGTGGGCT TACATGGCGA TAGCTAGACT GGGCGGTTTTT
951 ATGGACAGCA AGCGAACCGG AATTGCCAGC TGGGGCGCCC TCTGGTAAGG
1001 TTGGGAAGCC CTGCAAAGTA AACTGGATGG CTTTCTTGCC GCCAAGGATC
1051 TGATGGCGCA GGGGATCAAG ATCTGATCAA GAGACAGGAT GAGGATCGTT
1101 TCGCATGATT GAACAAGATG GATTGCACGC AGGTTCCTCCG GCCGCTTGGG
1151 TGGAGAGGCT ATTCGGCTAT GACTGGGCAC AACAGACAAT CGGCTGCCTC
1201 GATGCCGCCG TGTTCCGGCT GTCAGCGCAG GGGCGCCCGG TTCTTTTTTGT
1251 CAAGACCGAC CTGTCCGGTG CCCTGAATGA ACTGCAGGAC GAGGCAGCGC
1301 GGCTATCGTG GCTGGCCACG ACGGGCGTTC CTTGCGCAGC TGTGCTCGAC
1351 GTTGTCACTG AAGCGGGAAG GGACTGGCTG CTATTGGGCG AAGTGCCGGG
1401 GCAGGATCTC CTGTCACTC ACCTTGCTCC TGCCGAGAAA GTATCCATCA
1451 TGGCTGATGC AATGCGGCGG CTGCATACGC TTGATCCGGC TACCTGCCCA
1501 TTGACCACC AAGCGAAAACA TCGCATCGAG CGAGCACGTA CTCGGATGGA
1551 AGCCGGTCTT GTCGATCAGG ATGATCTGGA CGAAGAGCAT CAGGGGCTCG
1601 CGCCAGCCGA ACTGTTCCGC AGGCTCAAGG CGCGCATGCC CGACGGCGAG
1651 GATCTCGTCG TGACACATGG CGATGCCTGC TTGCCGAATA TCATGGTGGA
1701 AAATGGCCGC TTTTCTGGAT TCATCGACTG TGGCCGGCTG GGTGTGGCGG
1751 ACCGCTATCA GGACATAGCG TTGGCTACCC GTGATATTGC TGAAGAGCTT
1801 GGCGGCGAAT GGGCTGACCG CTTCTCGTG CTTTACGGTA TCGCCGCTCC
1851 CGATTTCGAG CGCATCGCCT TCTATCGCCT TCTTGACGAG TTCTTCTGAG
1901 CGGGACTCTG GGGTTCGAAA TGACCGACCA AGCGACGCC AACCTGCCAT
1951 CACGAGATTT CGATTCCACC GCCGCCTTCT ATGAAAGGTT GGGCTTCGGA
2001 ATCGTTTTTC GGGACGCCGG CTGGATGATC CTCCAGCGCG GGGATCTCAT
2051 GCTGGAGTTC TTCGCCACC CCGGGAT

```

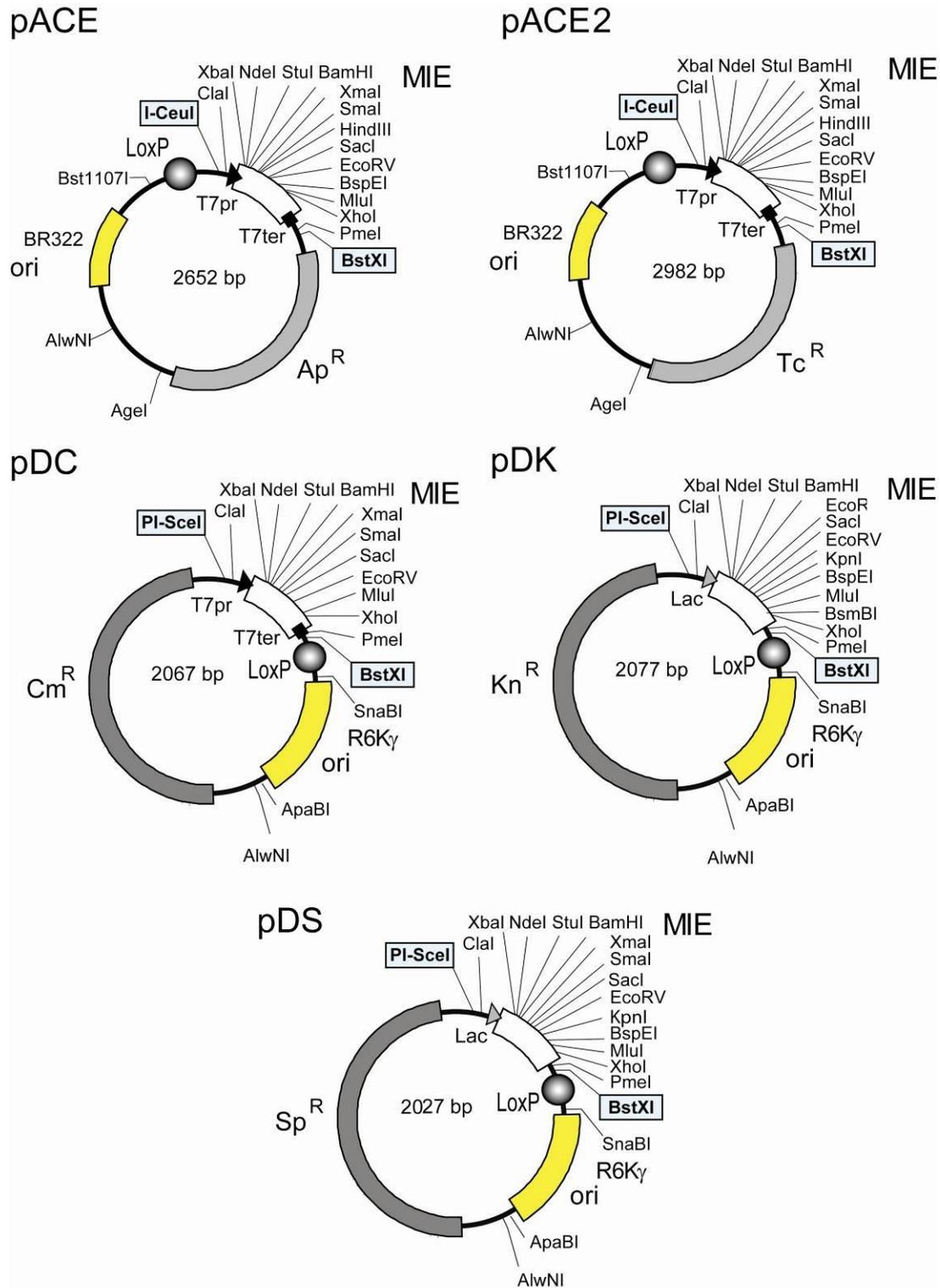
## F.2.5. pDS

```

1      CTATGTCGGG  TCGGAGAAAA  GAGGTAATGA  AATGGCACCT  AGGTATCGAT
51     GGCTTTACAC  TTTATGCTTC  CGGCTCGTAT  GTTGTGTGGA  ATTGTGAGCG
101    GATAACAATT  TCACACAGGA  AACAGCTATG  ACCATGATTA  CGAATTTCTA
151    GAAATAATTT  TGTTTAACTT  TAAGAAGGAG  ATATACATAT  GAGGCCTCGG
201    ATCCTGTAAA  ACGACGGCCA  GTGAATTCCC  CGGGAAGCTT  CGCCAGGGTT
251    TTCCCAGTCG  AGCTCGATAT  CGGTACCAGC  GGATAACAAT  TTCACATCCG
301    GATCGCGAAC  GCGTCTCGAG  ACTAGTTCG  TTTAAACCCA  TGTGCCTGGC
351    AGATAACTTC  GTATAATGTA  TGCTATACGA  AGTTATGGTA  CGTACTAAGC
401    TCTCATGTTT  CACGTACTAA  GCTCTCATGT  TTAACGTACT  AAGCTCTCAT
451    GTTTAACGAA  CTAAACCCTC  ATGGCTAACG  TACTAAGCTC  TCATGGCTAA
501    CGTACTAAGC  TCTCATGTTT  CACGTACTAA  GCTCTCATGT  TTGAACAATA
551    AAATTAATAT  AAATCAGCAA  CTTAAATAGC  CTCTAAGGTT  TTAAGTTTTA
601    TAAGAAAAAA  AAGAATATAT  AAGGCTTTTA  AAGCTTTTAA  GGTTTAACGG
651    TTGTGGACAA  CAAGCCAGGG  ATGTAACGCA  CTGAGAAGCC  CTTAGAGCCT
701    CTCAAAGCAA  TTTTGAGTGA  CACAGGAACA  CTTAACGGCT  GACATAATTC
751    GGTTCACGCG  TGCCGCAAG  ACTCAGGGCG  CAAGGGCTGC  TAAAGGAAGC
801    GGAACACGTA  GAAAGCCAGT  CCGCAGAAAC  GGTGCTGACC  CCGGATGAAT
851    GTCAGCTGGG  AGGCAGAATA  AATGATCATA  TCGTCAATTA  TTACCTCCAC
901    GGGGAGAGCC  TGAGCAAAC  GGCCTCAGGC  ATTTGAGAAG  CACACGGTCA
951    CACTGCTTCC  GGTAGTCAAT  AAACCGGTAA  GTAGCGTATG  CGCTCACGCA
1001   ACTGGTCCAG  AACCTTGACC  GAACGCAGCG  GTGGTAACGG  CGCAGTGGCG
1051   GTTTTCATGG  CTTGTTATGA  CTGTTTTTTT  GGGGTACAGT  CTATGCCTCG
1101   GGCATCCAAG  CAGCAAGCGC  GTTACGCCGT  GGGTCGATGT  TTGATGTTAT
1151   GGAGCAGCAA  CGATGTTACG  CAGCAGGGCA  GTCGCCCTAA  AACAAAGTTA
1201   AACATCATGA  GGGAAAGCGT  GATCGCCGAA  GTATCGACTC  AACTATCAGA
1251   GGTAGTTGGC  GTCATCGAGC  GCCATCTCGA  ACCGACGTTG  CTGGCCGTAC
1301   ATTTGTACGG  CTCCGCAGTG  GATGGCGGCC  TGAAGCCACA  CAGTGATATT
1351   GATTTGCTGG  TTACGGTGAC  CGTAAGGCTT  GATGAAACAA  CGCGGCGAGC
1401   TTTGATCAAC  GACCTTTTGG  AAACCTTCGG  TTCCCCTGGA  GAGAGCGAGA
1451   TTCTCCGCGC  TGTAGAAGTC  ACCATTGTTG  TGCACGACGA  CATCATTCGG
1501   TGGCGTTATC  CAGCTAAGCG  CGAACTGCAA  TTTGGAGAAT  GGCAGCGCAA
1551   TGACATTCTT  GCAGGTATCT  TCGAGCCAGC  CACGATCGAC  ATTGATCTGG
1601   CTATCTTGCT  GACAAAAGCA  AGAGAACATA  GCGTTGCCTT  GGTAGGTCCA
1651   GCGGCGGAGG  AACTCTTTGA  TCCGGTTCCT  GAACAGGATC  TATTTGAGGC
1701   GCTAAATGAA  ACCTTAACGC  TATGGAACCT  GCCGCCCGAC  TGGGCTGGCG
1751   ATGAGCGAAA  TGTAGTGCTT  ACGTTGTCCC  GCATTTGGTA  CAGCGCAGTA
1801   ACCGGCAAAA  TCGCGCCGAA  GGATGTCGCT  GCCGACTGGG  CAATGGAGCG
1851   CCTGCCGGCC  CAGTATCAGC  CCGTCATACT  TGAAGCTAGA  CAGGCTTATC
1901   TTGGACAAGA  AGAAGATCGC  TTGGCCTCGC  GCGCAGATCA  GTTGGAAAGAA
1951   TTTGTCCACT  ACGTGAAAGG  CGAGATCACC  AAGGTAGTCG  GCAAATAATG
2001   TCTACAATT  CGTTCAAGCC  GACGGAT

```

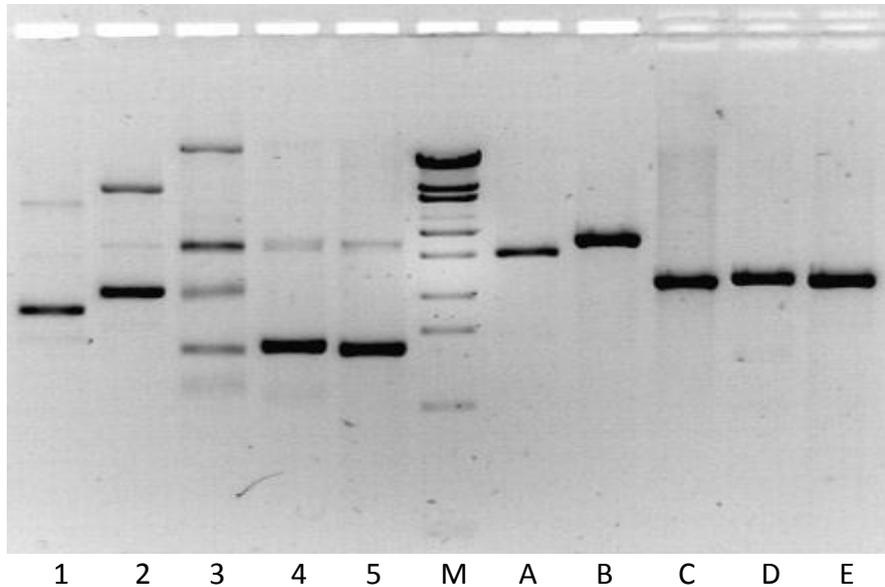
## F.2.6. ACEMBL plasmid maps



Acceptor vectors pACE1 (pACE) and pACE2, containing a T7 promoter and terminator, are shown. Donor vectors pDK, pDS and pDC contain conditional origins of replication. pDS and pDK have a lac promoter. pDC has a T7 promoter. Resistance markers are shown in gray, origins of replication in yellow. LoxP imperfect inverted repeat sequences are shown as spheres. Homing endonuclease sites and corresponding BstXI sites are boxed. The restriction enzyme sites in the multiple integration element (MIE) are indicated. All MIEs

have the same DNA sequence between *Cl*I and *P*meI. Differences in unique restriction site composition stem from differences in the plasmid backbone sequences.

All ACEMBL vectors were analyzed by *B*amHI restriction digestion. The undigested and digested ACEMBL vectors are shown below:



**Restriction mapping of ACEMBL vectors.** Both undigested as well as *B*amHI-digested Acceptor and Donor vectors are. All restriction reactions yield the expected sizes. Lane 1-5 show uncut pACE1, pACE2, pDC, pDK, and pDS vectors; lane M shows  $\lambda$  Styl marker; lane A-E show *B*amHI-digested pACE1, pACE2, pDC, pDK, and pDS vectors.