IMPACT™-TWIN

Purification, Ligation and Cyclization of Recombinant Proteins Using Self-Cleavable Affinity Tags

For additional information, including vector sequences and frequently asked questions, see the NEB website: www.neb.com

Instruction Manual

Catalog #E6950S
Store at –20°C
Note: Store chitin beads at 4°C

Version 1.5
7/08

Table of Contents:

Kit Components 2
Introduction 3
System Components 4
pTWIN Vector Map 6
Protein Purification Strategies 8
N-terminal Fusion 9
C-terminal Fusion 15
N- and C-terminal Intein Fusions 21
Cloning into a pTWIN Vector 27
Affinity Purification and On-column Cleavage 30
Simplified Purification Protocol for Test Experiment 34
Media and Solutions 38
Frequently Asked Questions 40
References 46
Appendices
Appendix I: Comparison Between Intein-Tag Fusions 48
Appendix II: Protein Labeling and Intein-mediated Protein Ligation 50
Appendix III: Sequencing Primers 52
Appendix IV: The Chemical Mechanism of Protein Splicing 54
Appendix V: The Chemical Mechanism of Thiol-inducible and pH-inducible Cleavage Reactions 56
Appendix VI: Regeneration of Chitin Resin 58
Appendix VII: Research Use Assurance Statement 59
Kit Components Sold Separately/Companion Products 60

Notice to Buyer/User: The buyer and user have a non-exclusive license to use this system or any components thereof for RESEARCH PURPOSES ONLY. See RESEARCH USE ASSURANCE STATEMENT attached hereto as Appendix VII for detail on terms of the license granted hereunder (U.S. Patent Nos. 5,496,714 and 5,834,247).

Information presented herein is accurate and reliable to the best of our knowledge and belief, but is not guaranteed to be so. Nothing herein is to be construed as recommending any practice or any product in violation of any patent or violation of any law or regulation. It is the user's responsibility to determine for himself or herself the suitability of any material and/or procedure for a specific purpose and to adopt such safety precautions as may be necessary.

Kit Components:

IMPACT™-TWIN System components are described in detail on page 4. A list of components sold separately and companion products can be found on page 60.

- **Vector DNA (3 vectors)** 10 µg of each (50 µl)
- **Sequencing Primers (3 primers)** 200 picomoles of each
- ***E. coli* Strain ER2566** 0.2 ml stock cells
- **Anti-Chitin Binding Domain Serum** (rabbit) 50 µl
- **Chitin Beads** (store at 4°C) 20 ml
- **1, 4-Dithiothreitol (DTT), 1M** 5 ml
- **Blue Loading Buffer** 1 ml
- **Instruction Manual**

Introduction:

The IMPACT-TWIN (Intein Mediated Purification with an Affinity Chitin-binding Tag-Two Intein) system is a novel protein purification system which utilizes the inducible self-cleavage activity of protein splicing elements (termed inteins) to separate the target protein from the affinity tag. Unlike other purification systems, IMPACT-TWIN is able to isolate native recombinant proteins possessing an N-terminal cysteine and/or possessing a reactive C-terminal thioester in a single chromatographic step without the use of exogenous proteases (1,2). These reactive groups can be used in Intein-mediated Protein Ligation (IPL, Appendix II) to specifically attach proteins, peptides, or labels to the N-and/or C-terminus of a target protein (1–11) using chemistry described previously (12,13). Furthermore, an intramolecular reaction makes it possible to generate circular protein species.

The IMPACT-TWIN system and its companion product, the IMPACT-CN protein purification system, were the direct result of New England Biolabs investigation into the mechanism of protein splicing (Appendix IV, Refs. 14,15). The IMPACT-TWIN system allows a target protein to be sandwiched between two self-cleaving inteins (Figure 9, Ref. 3). Chitin binding domains (16) present on both inteins allow the affinity purification of the precursor protein on a chitin resin. Intein1 is a mini-intein derived from the *Synechocystis* sp *dnaB* gene (17) engineered to undergo pH and temperature dependent cleavage at its C-terminus (18). Cleavage of this intein can liberate an N-terminal amino acid residue other than Met on a target protein. A protein with an N-terminal cysteine residue can be used in IPL reactions. Intein2 is either a mini-intein from the *Mycobacterium xenopi gyrA* gene (19) (pTWIN1) or from the *Methanobacterium thermoauto-trophicum rir1* gene (20) (pTWIN2). These inteins have been modified to undergo thiol-induced cleavage at their N-terminus (1,4,21). The use of thiol reagents, such as 2-mercaptoethanesulfo-nic acid (MESNA), releases a reactive thioester at the C-terminus of the target protein for use in IPL. Following cleavage of both inteins the target protein is eluted from the chitin resin while the inteins remain bound through the chitin binding domains.

This system offers many advantages: (i) the facile isolation of native proteins without an affinity tag that may alter its characteristics. (ii) the isolation of proteins with an N-terminal cysteine or residue other than methionine without the use of exogenous proteases which can be costly and non-specific (the intein has not been observed to cleave unexpected sites). (iii) the purification of

proteins with a C-terminal thioester for use in IPL reactions to insert non-coded amino acids into or label a bacterially expressed protein. NEB supplies peptides to specifically label the C-terminus of a protein with either a fluorescein (NEB #P6606S) or a biotin (NEB #P6607S) molecule. (iv) the generation of circular protein species (Figure 9).

System Components:

Cloning Vector DNA

The pTWIN vectors are used for the cloning and expression of recombinant proteins in *E. coli*. The exact vector and cloning strategy that should be employed depends both on the desired outcome and the properties of the target protein (see Protein Purification Strategies). Both pTWIN1 (7375 bp) and pTWIN2 (7192 bp) contain Sap I sites which allow the gene of interest to be cloned between the intein-tags without the addition of any vector derived residues at either termini of the target gene. Both the pTWIN1 and pTWIN2 vectors use a modified *Ssp* DnaB intein as Intein1 and differ only in the identity of intein2 (see Figure 9 for the relative positions of Intein1 and Intein2). pTWIN1 uses a modified *Mxe* GyrA intein while pTWIN2 uses a modified *Mth* RIR1 intein. However, both pTWIN1 and pTWIN2 contain the same multiple cloning sites which simplifies the insertion of a target gene into both vectors to determine the optimal expression plasmid.

pTWIN-MBP1 can be used both as a control vector and a cloning vector. Cloning of a target gene into the Nco I to Sac I sites in pTWIN-MBP1 adds 3 amino acids to the protein's N-terminus and 23 amino acids to its C-terminus. When additional amino acids will not alter the behavior of the target protein this linker may increase the yields of circular species. In the case of the 43 kDa *E. coli* maltose binding protein (MBP) these extra amino acids were found to permit cyclization whereas without these linker sequences no circular MBP was detected. Cloning into the Nco I to Xho I sites in pTWIN-MBP1 can be used if a smaller linker is desired. This results in 3 amino acids attached to the protein's N-terminus and 3 amino acids to its C-terminus.

The pTWIN vectors utilize a T7 promoter to provide stringent control of the fusion gene expression (22). In the absence of IPTG, basal expression of the fusion gene is suppressed by binding of the lac repressor to the lac operator sequence immediately downstream of the T7 promoter. The presence of the *lac*I gene, which encodes lac repressor, in the pTWIN vectors permits

tight transcriptional control even in *E. coli* strains lacking an endogenous *lac*I gene. Background transcription is further reduced by the placement of five tandem transcription terminators (rrnB T1) upstream of the T7 promoter sequence. The origin of DNA replication from bacteriophage M13 permits the production of single stranded DNA using helper phage (M13K07 Helper Phage, NEB #N0315S) superinfection of plasmid bearing cells. The pTWIN vectors carry the Amp^r gene marker (the *bla* gene), which conveys ampicillin resistance to the host strain.

Sequencing Primers

Three sequencing primers (200 pmoles, lyophilized) are included for sequencing genes inserted into the pTWIN vectors (Appendix III). The *Ssp* DnaB Intein Forward Primer (NEB #S1269S, 5´-ACTGGGACTCCATCGTTTCT-3´) is complementary to the *Ssp* DnaB intein sequence 87–107 nucleotides upstream of the *Ssp* DnaB intein C-terminus. The *Mxe* Intein Reverse II Primer (NEB #S1285S, 5´-GATTGCCATGCCGGTCAAGG-3´) is complementary to the *Mxe* GyrA intein sequence 85 nucleotides downstream of the *Mxe* GyrA intein N-terminus. The *Ssp* DnaB Intein Forward Primer and the *Mxe* Intein Reverse Primer are used to sequence a target gene inserted between the inteins in pTWIN1. The *Mth* RIR1 Intein Reverse II Primer (NEB #S1270S, 5´-GGTGAACGGTTTGCCCTC-3´) is complementary to the *Mth* RIR1 intein gene sequence 60–77 nucleotides downstream of the *Mth* RIR1 intein N-terminus. The *Mth* RIR1 Intein Reverse Primer is for use in conjunction with the *Ssp* DnaB Intein Forward Primer to sequence a target gene inserted between the inteins in pTWIN2.

There are 2 additional sequencing primers that can be purchased for sequencing genes inserted into the N- or C-terminal fusion constructs (Appendix III) of the pTWIN vectors. The T7 Universal Primer (NEB #S1248S, 5´-TAATACGACTCACTATAGGG-3´) is complementary to the T7 promoter and yields sequencing data in the direction of transcription. The T7 Universal Primer is used to sequence a target gene inserted in place of the CBD-Intein1 in either pTWIN1 or pTWIN2 (Figure 6). The T7 Terminator Reverse Primer (NEB #S1271S, 5´-TATGCTAGTTATTGCTCAG-3´) is complementary to a site within the T7 transcription terminator region 48–66 nucleotides downstream of the BamH I site. The T7 Terminator Reverse Primer is used to sequence a target gene inserted in place of the Intein2-CBD in either pTWIN1 or pTWIN2 (Figure 3).

pTWIN1
7375 bp

Sap I 6437
PaeR7 I - Tli I - Xho I 6430
EcoR I 6424
Not I 6416
Nco I 6410
Sap I 6404
Stu I 7062
Age I 7048
SgrA I 6957
Sac II 6798
Bsm I 6769
Aat II - Zra I 6688
BsiW I 6628
Spe I 6462
Pst I 7228
BamH I 7239
Blp I 7291
EcoO109 I 7318
Sca I 444
Pvu I 555
Fsp I 702
Swa I 1096
Psi I 1194
Dra III 1319
AlwN I 1896
Pci I 2310
BstZ17 I 2480
PflF I - Tth111 I 2504
Afe I 3194
Kas I - Nar I - Sfo I 3459
Hpa I 3594
EcoR V 3650
BssH II 3687
Apa I - PspOM I 3891
BstE II 3916
Nru I 6377
BseR I 6243
Bmt I - Nhe I 5971
Nde I 5722
Xba I 5683
BsaB I 5618
Pme I 5606
ori
rop
lacI

pTWIN1 Polylinker

```
5´...AC TGG GAC TCC ATC GTT TCT ATT ACG GAG ACT GGA GTC GAA GAG GTT TTT
      Ssp DnaB Intein Forward Primer →
                                                      ← Intein
                   ...Ssp DnaB Intein... Val Ala Asn Asp Ile Ile Val His Asn▼
GAT TTG ACT GTG CCA GGA CCA CAT AAC TTT GTC GCG AAT GAC ATC ATT GTA CAC AAC
                                          Nru I
                                              Intein →
Gly Arg Ala Met Gly Gly Arg Glu Phe Leu Glu Gly Ser Ser▼Cys Ile Thr Gly
GGA AGA GCC ATG GGC GGC CGC GAA TTC CTC GAG GGC TCT TCC TGC ATC ACG GGA GAT
   Sap I   Nco I    Not I    EcoR I   Xho I     Sap I
...Mxe GyrA Intein...
GCA CTA GTT GCC CTA CCC GAG...................................3´
    Spe I                    49nt   ← Mxe Intein Reverse II Primer
```

Figure 1: *pTWIN1 map with the unique sites. The polylinker for pTWIN1 is also shown.*

pTWIN2
7192 bp

Sma I - Xma I 6859
Ale I 6841
Bsm I 6814
Sac II 6677
Pml I 6602
Bsu36 I 6554
Rsr II 6480
Spe I 6472
Xho I - Tli I - PaeR7 I 6430
EcoR I 6424
Not I 6416
Nco I 6410
Nru I 6377
BseR I 6243
Nde I 5722
Xba I 5683
BsaB I 5618
Pme I 5606
Age I 6865
Stu I 6879
Pst I 7045
BamH I 7056
Blp I 7108
Eco O109 I 7135
Sca I 444
Pvu I 555
Fsp I 702
Swa I 1096
Psi I 1194
Dra III 1319
AlwN I 1896
Pci I 2310
Acc I - BstZ17 I 2480
PflF I - Tth111 I 2504
Afe I 3194
Kas I - Nar I - Sfo I 3459
Hpa I 3594
EcoR V 3650
BssH II 3687
Apa I - PspOM I 3891
BstE II 3916
Mlu I 4098
EcoN I 4562

pTWIN2 Polylinker

```
5´...AC TGG GAC TCC ATC GTT TCT ATT ACG GAG ACT GGA GTC GAA GAG GTT TTT
      Ssp DnaB Intein Forward Primer →
                                                      ← Intein
                   ...Ssp DnaB Intein... Val Ala Asn Asp Ile Ile Val His Asn▼
GAT TTG ACT GTG CCA GGA CCA CAT AAC TTT GTC GCG AAT GAC ATC ATT GTA CAC AAC
                                          Nru I
                                              Intein →
Gly Arg Ala Met Gly Gly Arg Glu Phe Leu Glu Gly Ser Ser▼Cys Val Ser Gly Asp Thr
GGA AGA GCC ATG GGC GGC CGC GAA TTC CTC GAG GGC TCT TCC TGC GTA TCC GGT GAC ACC ATT
   Sap I   Nco I    Not I    EcoR I   Xho I     Sap I
...Mth RIR1 Intein...
GTA ATG ACT AGT GGC GGT CCG CGC ACT GTG GCT GAA CTG GAG GGC AAA CCG TTC ACC...3´
        Spe I                               ← Mth RIR1 Intein Reverse Primer
```

Figure 2: *pTWIN2 map with the unique sites. The polylinker for pTWIN2 is also shown.*

Protein Purification Strategies:

There are 3 possible intein-tag fusions when using the pTWIN vectors; fusion of the intein-tag to the N-terminus, the C-terminus or to both the N-and C-termini of the target protein. The choice of which fusion to use depends on the intended purpose. For example, purification or protein engineering and the nature of the target protein.

Considerations for choosing a particular cloning strategy are outlined below.

N-terminal Fusion

- Protein Purification
  -using one-column step
  -of proteins with an N-terminal amino acid other than Met
  -does not require thiol reagents
  -of proteins with an N-terminal Cys for use in IPL reactions (see Appendix II)

C-terminal Fusion

- Protein Purification
  -using one-column step
  -of proteins with a C-terminal thioester for use in IPL reactions (see Appendix II)

N- and C-terminal Fusion

- Protein Purification
  -using one-column step
  -of circular protein species
  -of proteins with an N-terminal Cys for use in IPL reactions (see Appendix II)

N-terminal Fusion:

Fusion of the *Ssp* DnaB (intein 1) self-cleavable intein-tag to the N-terminus of a target protein allows one-column protein purification with a pH and temperature shift. No thiol reagents are required for cleavage, which is beneficial for isolating proteins that are sensitive to reducing agents such as DTT. Furthermore, the N-terminal fusion permits the isolation of a protein without an N-terminal methionine. This is useful for IPL in which a protein with an N-terminal cysteine is required as one of the reactants.

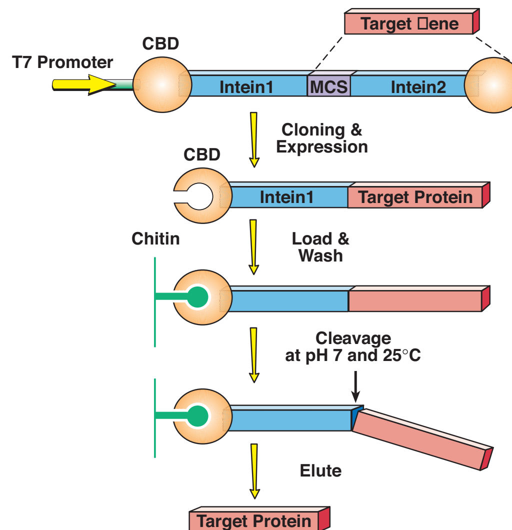


Figure 3: *Protein Purification by the pTWIN1 Vector. Fusion of the intein-tag to the N-terminus of the target protein and release of the target by pH and temperature induced cleavage.*

Overview of protein purification using the N-terminal fusion construct.

1. Clone the target gene into the appropriate pTWIN vector.
2. Transform the plasmid bearing the target gene into an applicable *E. coli* host strain (i.e. T7 Express Competent *E. coli* (NEB #C2566)).
3. Grow the cells at 37°C until an OD_{600} of 0.5–0.7 is reached.
4. Induce protein expression.
5. Equilibrate a chitin column in Buffer B1. (Buffer compositions are described in Media and Solutions on Page 38).
6. Lyse cells in Buffer B1 and slowly apply the clarified cell extract to the chitin column.
7. Wash the column with Buffer B1 to remove the unbound proteins.
8. Induce on-column cleavage of the intein-tag by flushing with 3 column volumes and equilibrating the chitin resin in Buffer B2.
9. Allow cleavage reaction to proceed overnight at room temperature.
10. Elute the target protein.
11. Dialyze the target protein into an appropriate storage buffer.
12. Determine intein-tag cleavage efficiency by taking an aliquot of the chitin beads or by SDS elution of the chitin beads.

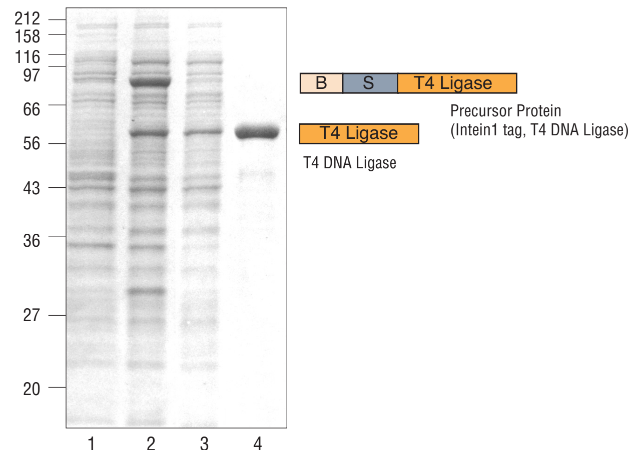


Figure 4: *N-terminal fusion. The intein-tag (Ssp DnaB intein) is fused to the N-terminus of the target protein (T4 DNA Ligase). Lane 1: Uninduced crude cell extract. Lane 2: Crude cell extract following induction of protein expression. The precursor protein and T4 DNA Ligase, released due to some in vivo cleavage of the intein, are visible. Lane 3: Crude cell extract after passage over chitin beads. The precursor binds to the resin through the chitin binding domain. Lane 4: A fraction eluted from the chitin beads after inducing cleavage of the intein-tag.*

Cloning the Target Gene (N-terminal fusion):

Choice of Vectors

An important consideration when using the pTWIN vectors for N-terminal fusion is whether or not to use extra amino acid residues at the target proteins N-terminus. Appendix I describes recommended amino acids for use with the *Ssp* DnaB intein. However, studies have shown that amino acid residues other than the target protein N-terminal residue can have a significant effect on the cleavage behavior of the *Ssp* DnaB intein. When possible, it has been found that adding the amino acids CRA or GRA prior to the N-terminal methionine of the target protein increases the controlled cleavage reaction.

Cloning a target gene into the Sap I to Pst I sites in either pTWIN1 or pTWIN2 results in no extra vector-derived residues added to the N- or C-terminus of the target protein. In this case the Sap I site directly 3´ of the *Ssp dnaB* intein gene should be used. **Please note that the other sites in the polylinker region should not be used in conjunction with the Sap I site.** The cloning strategy is illustrated in Figure 5.

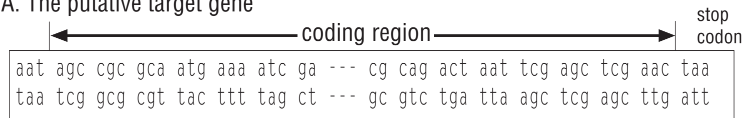
Cloning into the Nco I to Pst I sites in pTWIN1 or pTWIN-MBP1 also results in the fusion of the *Ssp* DnaB intein to the N-terminus of the target protein. However, three vector derived amino acid residues, GRA for pTWIN1 and CRA for pTWIN-MBP1, are added to the N-terminus of the target protein. These residues have been found to permit controlled cleavage of the *Ssp* DnaB intein.

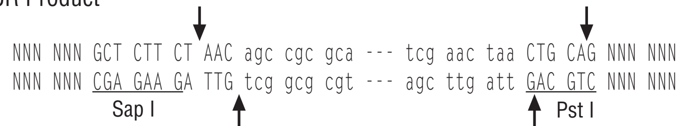
Primer Design

Normally, a target gene is amplified by PCR before it is inserted in-frame into the cloning site of one of the pTWIN vectors. Appropriate restriction sites, absent in the target gene, are incorporated into the forward and reverse primers. The choice of the restriction sites in the primers determines whether any extra vector derived amino acid residues will be attached to the termini of the target protein after the cleavage of the intein-tag.

For example, to obtain a target protein with three extra vector derived residues (GRA), which facilitate controlled cleavage, the target gene is cloned into the Nco I to Pst I sites in the pTWIN1 vector. Alternatively, cloning the target gene into the Sap I to Pst I sites in the pTWIN1 vector results in a fusion with no extra vector derived amino acid residues on the target protein. Table 1, Table 2 and Figure 5 illustrate primer design and cloning strategies for the use of the *Ssp* DnaB intein (Intein 1) as a self-cleaving affinity tag for purification.

Cloning into the Sap I to Pst I sites in pTWIN1. This allows the utilization of intein1 (the *Ssp* DnaB intein) as an N-terminal fusion for protein purification without thiol-reagents.





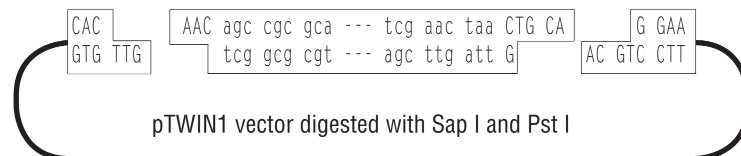


Figure 5: *Cloning of a PCR product into the Sap I site 3´ of the Ssp dnaB intein gene and the Pst I site of pTWIN1 or pTWIN2. The putative target gene is in lower case letters. The addition of extra bases (NNN NNN) to the 5´ and 3´ ends of the PCR product is required for efficient restriction enzyme activity. Following digestion with Sap I and Pst I, the insert and vector can be ligated. The Pst I site is regenerated, however the Sap I site is lost. In this example, the first codon of the target gene (agc) is ligated next to the C-terminal asparagine of the Ssp DnaB intein (AAC). Following intein cleavage the target protein will have an N-terminal serine residue. A stop codon (taa) should be included in the PCR product as illustrated.*

Table 1: Examples of forward primer design for use of the Ssp DnaB intein (Intein 1) as a self-cleavable affinity tag.

Restriction Site Used in the Forward Primer	Forward Primer†	Cloning Vector
Sap I*	5´- GGT GGT TGC TCT TCC AAC NNN...-3´	pTWIN1
Nco I	5´- GGT GGT CC ATG GNN N...-3´	pTWIN1 or pTWIN-MBP1

† The target gene sequence is represented by "NNN...". Restriction sites are underlined. The "GGT GGT" sequence at the 5´ end of the primer is to ensure efficient DNA cleavage by the restriction enzyme.

* Sap I site is lost after cloning.

Table 2: Examples of reverse primer design for use of the Ssp DnaB intein (Intein 1) as a self-cleavable affinity tag.

Restriction Site Used in the Reverse Primer	Reverse Primer†	Cloning Vector
Pst I	5´- GGT GGT CTG CAG TTA NNN...-3´	pTWIN1 or pTWIN-MBP1
BamH I	5´- GGT GGT GGA TCC TTA NNN...-3´	pTWIN1 or pTWIN-MBP1

† The target gene sequence is represented by "NNN...". Restriction sites are underlined. The "GGT GGT" sequence at the 5´ end of the primer is to ensure efficient DNA cleavage by the restriction enzyme. When using the Pst I or BamH I sites a stop codon should be included in the reverse primer.

C-terminal Fusion:

Protein purification with the intein-tag fused to the C-terminus of the target protein allows the isolation of a protein using the thiol-inducible cleavage activity of either the *Mxe* GyrA (intein 2 in pTWIN1) or *Mth* RIR1 intein (intein 2 in pTWIN2). A reactive C-terminal thioester can be generated on the target protein when 2-mercaptoethanesulfonic acid (MESNA) is used as the thiol-reagent to induce cleavage of the intein. The C-terminal thioester can be used in IPL to ligate bacterially-expressed proteins or synthetic peptides to the target protein (Appendix II).

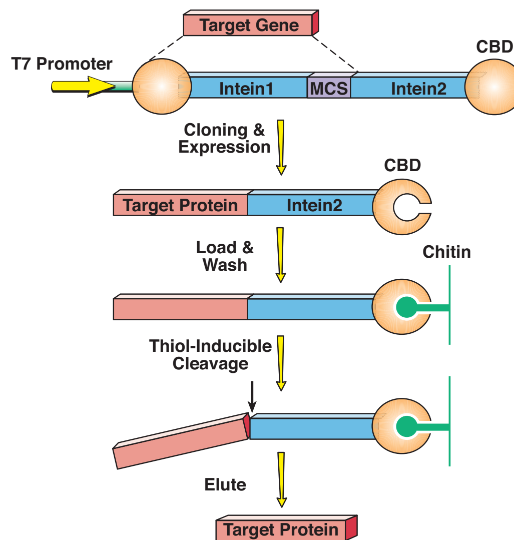


Figure 6: *Protein Purification by a pTWIN Vector. Fusion of Intein2 to the C-terminus of the target protein and release of the target by thiol-induced cleavage.*

Overview of protein purification using the C-terminal fusion construct.

1. Clone the target gene into the appropriate pTWIN vector.
2. Transform the plasmid bearing the target gene into an applicable *E. coli* host strain (i.e. T7 Express Competent *E. coli* (NEB #2566)).
3. Grow the cells at 37°C until an OD_{600} of 0.5–0.7 is obtained.
4. Induce protein expression.
5. Equilibrate a column containing chitin beads in Buffer B1. (Buffer compositions are described in Media and Solutions on Page 38).
6. Lyse cells in Buffer B1 and slowly apply the clarified cell extract to the chitin column.
7. Wash the column with Buffer B1 to remove unbound proteins.
8. On-column cleavage of the intein-tag is induced by equilibrating the chitin beads in Buffer B3. If a reactive thioester is desired for subsequent IPL reactions then Buffer B4 should be used in place of Buffer B3.
9. Allow the on-column cleavage reaction to proceed overnight.
10. Elute the target protein with the buffer used for cleavage (Buffer B3 or B4).
11. If Buffer B4 was used as the cleavage buffer for IPL reactions, then the freshly eluted thioester-tagged target protein should be used immediately. If not performing an IPL reaction go to step 12.
12. Dialyze the target protein into an appropriate storage buffer.
13. Determine intein-tag cleavage efficiency by taking an aliquot of chitin beads or by SDS elution of the chitin beads.

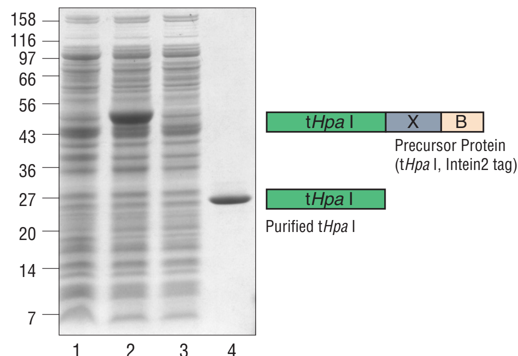


Figure 7: *C-terminal fusion. The intein-tag (Mxe GyrA intein) is fused to the C-terminus of the target protein tHpa I, a truncated form of the restriction enzyme Hpa I. Lane 1: Uninduced crude cell extract. Lane 2: Crude cell extract following induction of protein expression. The precursor protein is visible. Lane 3: Crude cell extract after passage over chitin beads. The precursor binds to the resin through the chitin binding domain. Lane 4: Fractions eluted from the chitin beads contain the purified tHpa I.*

Cloning the Target Gene (C-terminal Fusion):

Choice of Vectors

When using the pTWIN vectors to fuse an intein-tag to the C-terminus of a target protein it is necessary to decide which intein to use, either the *Mxe* GyrA or *Mth* RIR1 intein, and whether or not to add extra amino acids to the C-terminus of the target protein. The *Mxe* GyrA intein is recommended for purification because it has been found to have predictable cleavage behavior with a wide range of target protein fusions. However, the *Mxe* GyrA intein did not cleave well when a test protein, maltose binding protein, had amino acid residues with small side chains, such as glycine and alanine, at its C-terminus (see FAQ #2.12 at www.neb.com). This resulted in

lower than expected yields of purified maltose binding protein. The *Mth* RIR1 intein may perform better in this case as it undergoes proficient thiol-induced cleavage with glycine at the C-terminus of the target protein. Because intein cleavage efficiency will vary depending on the target protein sequence, it is advisable to test both the *Mxe* GyrA and the *Mth* RIR1 inteins and compare the yields of protein.

Appendix I describes amino acid residues that are recommended at the C-terminus of the target protein when using the *Mxe* GyrA or *Mth* RIR1 intein. However, studies have shown that amino acid residues other than the target protein C-terminal residue can have a significant effect on an inteins cleavage behavior. When possible, it has been found that adding the amino acids MRM to the C-terminus of the target protein increases the controlled cleavage reaction with the *Mxe* GyrA intein.

Cloning into the Nde I to Sap I sites in pTWIN1 or pTWIN2 results in the fusion of the *Mxe* GyrA intein or the *Mth* RIR1 intein, respectively, to the C-terminus of the target protein. The cloning strategy using pTWIN1 or pTWIN2 as a C-terminal intein fusion vector is illustrated in Figure 8.

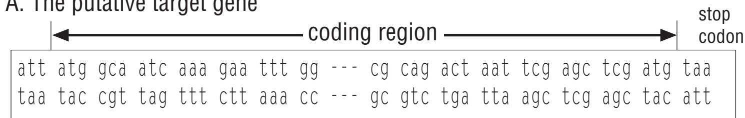
Primer Design

Normally, a target gene is amplified by PCR before it is inserted in-frame into the multiple cloning site (MCS) of one of the pTWIN vectors. Appropriate restriction sites, absent in the target gene, are incorporated into the forward and reverse primers. The choice of the restriction sites in the primers determines whether any extra vector derived amino acid residues will be attached to the termini of the target protein after the cleavage of the intein-tag.

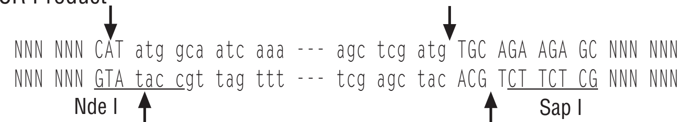
For example, to obtain a target protein with no extra vector derived amino acid residues, the target gene is cloned into the Nde I to Sap I sites in the pTWIN1 or pTWIN2 vector. Table 3 and Figure 8 illustrate primer design and cloning strategies for the use of the *Mxe* GyrA or *Mth* RIR1 intein as a self-cleaving affinity tag for purification.

Cloning into the Nde I to Sap I sites in pTWIN1 or pTWIN2. This allows the utilization of intein2 (the *Mxe* GyrA *or Mth* RIR1 intein) as a C-terminal fusion for protein purification.

A. The putative target gene



B. PCR Product



C. Ligation

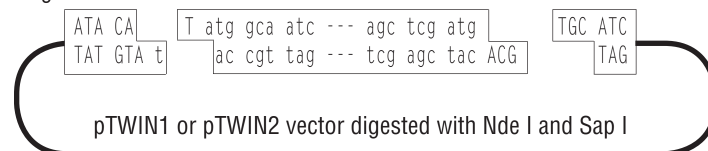


Figure 8: *Cloning of a PCR product into the Nde I site and the Sap I site directly 5´ of intein2 of pTWIN1 or pTWIN2. The putative target gene is in lower case letters. The addition of extra bases (NNN NNN) to the 5´ and 3´ ends of the PCR product is required for efficient restriction enzyme activity. Following digestion with Nde I and Sap I and ligation of the appropriate fragments the Nde I site is regenerated but the Sap I site is lost. With this cloning strategy, no extra vector derived amino acid residues are present following cleavage of intein2.* ***A stop codon should not be included at the 3´ end of the target gene.***

Table 3: Examples of primer design for use of the Mxe GyrA or Mth RIR1 intein as a self-cleavable affinity tag.

Restriction Site Used in the Primer	Forward Primer†	Cloning Vector
Nde I	5´- GGT GGT CAT ATG NNN...-3´	pTWIN1 or pTWIN2
Restriction Site Used in the Primer	**Reverse Primer†**	**Cloning Vector**
Sap I*	5´- GGT GGT TGC TCT TCC GCA NNN...-3´	pTWIN1 or pTWIN2

† *The target gene sequence is represented by "NNN...". Restriction sites are underlined. The "GGT GGT" sequence at the 5´ end of the primer is to ensure efficient DNA cleavage by the restriction enzyme.*

* *Sap I site is lost after cloning.*

*A stop codon should **not** be included in the reverse primer.*

N- and C-terminal Intein Fusions:

Fusing an intein to both the N- and C-termini of a target protein allows: (i) the isolation of a protein with an N-terminal residue other than methionine, (ii) the purification of a protein with an N-terminal residue other than methionine and a C-terminal thioester. In both cases a cysteine is often desired as the N-terminal residue for use in IPL reactions (Appendix II). In particular, the generation of an N-terminal cysteine and a C-terminal thioester on the same protein allows the creation of circular protein species (Figures 9 and 10).

The use of two inteins to isolate a protein with an N-terminal residue other than methionine can be particularly advantageous when the *Ssp* DnaB intein, Intein1, undergoes uncontrollable *in vivo* cleavage with the target protein. This makes it unusable as a single N-terminal affinity tag fusion. In this case, if the target protein is fused between two inteins then Intein1 can cleave *in vivo* and the *Mxe* GyrA (intein 2 in pTWIN1) or the *Mth* RIR1 intein (intein2 in pTWIN2) can be used as the self-cleavable affinity tag for protein purification.

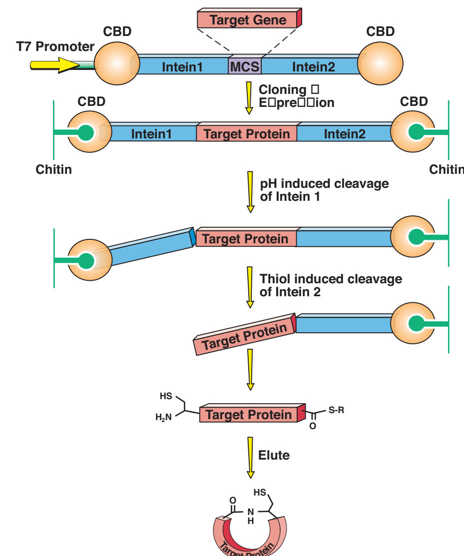
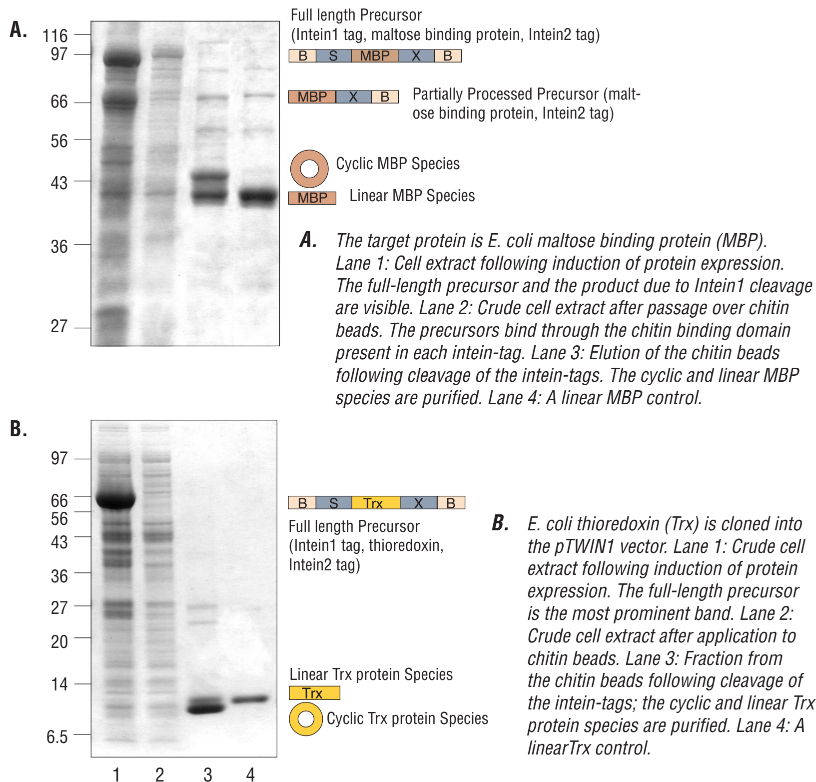


Figure 9: *Protein Purification by a pTWIN Vector. Fusion of the target protein between Intein1 and Intein2 and release of the target with an N-terminal cysteine and a C-terminal thioester.*

Figure 10: *N- and C-terminal fusion. The target protein is sandwiched between two intein-tags.*



A. *The target protein is E. coli maltose binding protein (MBP). Lane 1: Cell extract following induction of protein expression. The full-length precursor and the product due to Intein1 cleavage are visible. Lane 2: Crude cell extract after passage over chitin beads. The precursors bind through the chitin binding domain present in each intein-tag. Lane 3: Elution of the chitin beads following cleavage of the intein-tags. The cyclic and linear MBP species are purified. Lane 4: A linear MBP control.*

B. *E. coli thioredoxin (Trx) is cloned into the pTWIN1 vector. Lane 1: Crude cell extract following induction of protein expression. The full-length precursor is the most prominent band. Lane 2: Crude cell extract after application to chitin beads. Lane 3: Fraction from the chitin beads following cleavage of the intein-tags; the cyclic and linear Trx protein species are purified. Lane 4: A linearTrx control.*

Overview of the isolation of a protein when fused between two inteins.

1. Clone the target gene into the appropriate pTWIN vector.
2. Transform the plasmid bearing the target gene into an applicable *E. coli* host strain (i.e. ER2566 contained in kit).
3. Grow the cells at 37°C until an OD_{600} of 0.5–0.7 is reached.
4. Induce protein expression.
5. Equilibrate chitin beads in Buffer B1.
6. Lyse the cells in Buffer B1 and slowly apply the clarified cell extract to the chitin column.
7. Wash the column with Buffer B1 to remove unbound proteins.
8. Induce on-column cleavage of Intein1 by equilibrating the chitin beads in Buffer B2.
9. Allow the cleavage reaction to proceed overnight at room temperature.
10. Wash the column with Buffer B2 and then B1 to remove any unbound protein species.
11. On-column cleavage of Intein2 is induced by equilibrating the chitin beads in Buffer B3. If a reactive thioester is desired for subsequent IPL reactions, such as cyclization, then Buffer B4 should be used in place of Buffer B3.
12. Incubate the chitin column overnight.
13. Elute the target protein with the buffer used for cleavage (Buffer B3 or B4).
14. Determine cleavage efficiency taking an aliquot of chitin beads or by SDS elution of the chitin beads.

Cloning the Target Gene (N- and C-terminal fusion):

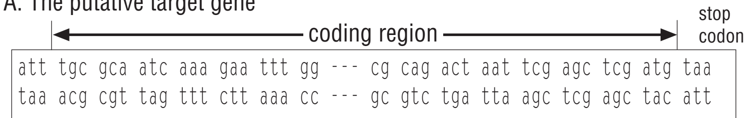
Choice of Vectors

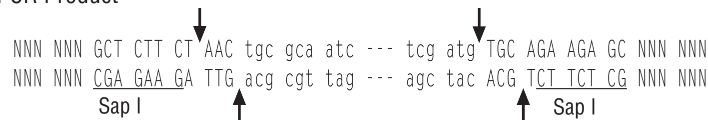
When fusing a target gene between two inteins using the pTWIN vectors it is necessary to decide which intein to fuse to the C-terminus of the target protein, either the *Mxe* GyrA or *Mth* RIR1 intein. The intein fused to the N-terminus of the target protein is the *Ssp* DnaB intein in either case. It is also necessary to decide whether or not to add extra amino acids to either the N-terminus, the C-terminus or both termini of the target protein. The considerations of whether to use the *Mxe* GyrA or the *Mth* RIR1 intein is the same as described in the section on C-terminal Fusion (see page 15). The *Mxe* GyrA intein is recommended for purification because it has been found to have predictable cleavage behavior with a wide range of target protein fusions. However, the *Mxe* GyrA intein did not cleave well when a test protein, maltose binding protein, had amino acid residues with small side chains, such as glycine and alanine, at its C-terminus. This resulted in lower than expected yields of purified maltose binding protein. The *Mth* RIR1 intein may perform better in this case as it undergoes proficient thiol-induced cleavage with glycine at the C-terminus of the target protein. As a general rule, because the target protein sequence can influence the cleavage behavior of the intein, it is advisable to test both the *Mxe* GyrA and the *Mth* RIR1 inteins and compare the final yields of protein.

Appendix I describes the recommended amino acids at the N- and C-termini of the target protein with the various inteins. However, studies have shown that amino acid residues other than the target protein C-terminal residue can have a significant effect on an inteins cleavage behavior. When possible, it has been found that adding the amino acids MRM to the C-terminus of the target protein increases the controlled cleavage reaction with the *Mxe* GyrA intein.

Cloning into the Sap I sites in pTWIN1 or pTWIN2 results in the fusion of the *Mxe* GyrA intein or the *Mth* RIR1 intein, respectively, to the C-terminus of the target protein. The cloning strategy using pTWIN1 or pTWIN2 is illustrated in Figure 11.

Cloning into the Sap I to Sap I sites in pTWIN1 or pTWIN2. This allows the sandwiching of the target gene between intein1 (the *Ssp* DnaB intein) and intein2 (the *Mxe* GyrA or *Mth* RIR1 intein).





C. Ligation

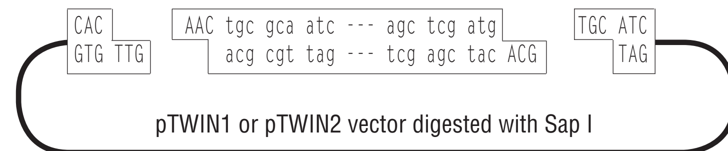


Figure 11: *Cloning of a PCR product into the Sap I sites of pTWIN1 or pTWIN2 results in the fusion of intein-tags to the N- and C-termini of the target protein. The putative target gene is in lower case letters. The addition of extra bases (NNN NNN) to the 5´ and 3´ ends of the PCR product is required for efficient restriction enzyme activity. Following digestion with Sap I and ligation of the appropriate fragments both of the Sap I sites are lost. With this cloning strategy, no extra vector derived amino acid residues are present following cleavage of the intein-tags. A stop codon should **not** be included at the 3´ end of the target gene. Other restriction enzyme sites in the pTWIN1 or pTWIN2 MCS should **not** be used in conjunction with the Sap I sites.*

Primer Design

Normally, a target gene is amplified by PCR before it is inserted in-frame into the multiple cloning site (MCS) of one of the pTWIN vectors. Appropriate restriction sites, absent in the target gene, are incorporated into the forward and reverse primers. The choice of the restriction sites in the primers determines whether extra amino acid residues, if any, will be attached to the termini of the target protein after the cleavage of the intein-tag.

For example, to obtain a target protein with no extra vector derived amino acid residues, the target gene is cloned into the two Sap I sites in the pTWIN1 or pTWIN2 vector. It should be noted that the Sap I digested DNA does not undergo self-ligation because DNA restriction at both Sap I sites results in non-complimentary overhangs. Table 4 and Figure 11 illustrate primer design and cloning strategies to insert a target gene between two inteins using the pTWIN vectors.

Table 4: *Examples of primer design for the insertion of a target gene between two inteins.*

Restriction Site Used in the Forward Primer	Forward Primer†	Cloning Vector
Sap I††	5´- GGT GGT TGC TCT TCC AAC NNN...-3´	pTWIN1 or pTWIN2
Nco I*	5´- GGT GGT CC ATG GNN N...-3´	pTWIN1, pTWIN2 or pTWIN-MBP1
	Reverse Primer†	
Sap I††	5´- GGT GGT TGC TCT TCC GCA NNN...-3´	pTWIN1 or pTWIN2
Sac I*	5´- GGT GGT TGA GCT CNN NN...-3´	pTWIN-MBP1
Xho I*	5´- GGT GGT CTC GAG NNN...-3´	pTWIN-MBP1

† *The target gene sequence is represented by "NNN...". Restriction sites are underlined. The "GGT GGT" sequence at the 5´ end of the primer is to ensure efficient DNA cleavage by the restriction enzyme.*

* *Do **not** use in conjunction with Sap I cloning.*

†† *Sap I site is lost after cloning.*

*A stop codon should **not** be included in the reverse primer.*

Cloning into a pTWIN vector:

The following is a protocol to clone an amplified target gene fragment using restriction enzymes that create non-compatible sticky ends. For blunt-end or single site cloning, the vector may need to be treated with an alkaline phosphatase (Antarctic Phosphatase, NEB #M0289 or CIP, NEB #M0290). Also more units of T4 DNA Ligase (NEB #M0202) or the Quick Ligation Kit (NEB #M2200) should be used. This is only a suggested cloning strategy and users may vary the protocol or follow their own protocols.

1. **Purification of the Amplified Gene Fragment**: The reaction mixture containing the amplified target gene fragment (generally 100 µl) is directly loaded onto a 1% low melting-agarose gel. The correct fragment is excised from the gel using a razor blade. The gel slice (100–150 µl, the smaller the volume the better) is melted at 65°C for 7 minutes. Following an incubation at 42°C for 7 minutes to cool the gel slice, 2 µl of β-Agarase I (NEB #M0392) per 100 µl gel slice is added and the agarase is allowed to react at 42°C for 1 hour. The melted gel slice containing the target gene is then used for restriction enzyme digestion; no DNA precipitation is necessary.

2. **Restriction Digestion and Ligation**: The above gel-purified gene fragment (use 80 µl) is double-digested with the appropriate restriction enzymes in a 100 µl reaction mixture (Note use the suggested 10X NEBuffer for double digestion). At the same time, a pTWIN vector (0.5 µg) is digested with the same enzymes in a 50 µl reaction mixture. Following a 2 to 4 hour digestion, both reaction mixtures are loaded onto a 1% low melting-agarose gel. The gel slices (100–150 µL) containing the digested gene fragment and digested pTWIN vector are combined in the same eppendorf tube and melted at 65°C for 10 minutes. After cooling at 42°C for 7 minutes, 2 µl of β-Agarase I (NEB #M0392) per 100 µl of gel slice is added and allowed to react for 1 hour at 42°C. The ligation of the DNA fragments is conducted in the same mixture. 2 µl of T4 DNA Ligase (NEB #M0202) and an appropriate volume

of 10X T4 DNA Ligase Buffer are added to the mixture and the ligation reaction incubated overnight at 16°C.

3. **Transformation**: The above ligation mixture (15 µl) is used to transform 150 µl of competent cells. To reduce the background from vector self-ligation, the ligation sample can be digested prior to transformation with a restriction enzyme whose recognition site is deleted from the polylinker during cloning and is also absent from the insert. This linearizes any remaining parental vectors. The vectors supplied in the kit are prepared from a restriction-deficient *E. coli* strain (r^-m^-). When introduced into a strain with wild type EcoK (hsd^+), DNA will be restricted, therefore, fusion constructs should initially be established in a non-restricting (r^-m^- or r^-m^+), non-expression host.

4. **Screening for Inserts**: Plasmids are isolated from each colony and those containing the correct target gene insert are identified by digesting the plasmid DNA with the same restriction enzymes as were used for cloning the target gene fragment. An exception to this is when the Sap I restriction enzyme is used because the Sap I site is lost after ligation of the target gene into the pTWIN vector. Alternatively, it is possible to use other restriction sites present in the plasmid, as long as they are unique and are not found in the target gene. Colony PCR or colony hybridization can be used to screen a large number of transformants for the presence of the inserted target gene. Also, immunodetection with Anti-Chitin Binding Domain Serum is able to detect the full-length fusion precursor in total cell lysates. Clones should be further confirmed by DNA sequencing before proceeding to the cell culture and protein expression steps.

Cell Culture and Fusion Protein Expression

The expression of the fusion protein from a pTWIN vector may be affected by the following factors: (i) *E. coli* strain; (ii) cell culture conditions (e.g., temperature, aeration, cell density, etc.); (iii) protein expression induction conditions (temperature, duration, IPTG concentration, etc.). *E. coli* strain ER2566 is supplied in the kit as a host for fusion protein expression from a pTWIN vector. It should be noted that the supplied *E. coli* strain ER2566 is not competent. These cells can be made competent using standard procedures (23).Other commercially available strains (e.g. BL21(DE3) and derivatives can also be used as hosts for the pTWIN vectors. More stringent control of protein expression is possible by using hosts carrying the pLysS or the pLysE plasmid encoding T7 lysozyme. Different bacterial strains may be tested to achieve optimal expression of each fusion protein. Expression of a toxic protein may require lowering the culture temperature to maintain the vector stability; induction of protein expression at 15–30°C can often help the folding and solubility of the fusion protein and increase the cleavage efficiency of the intein. The following protocol is provided as a guideline.

Cell Culture

1 liter of LB medium containing 100 µg/ml ampicillin is inoculated with a freshly grown colony or 10 ml of freshly grown culture. The culture is incubated in an air shaker at 37°C until the OD_{600} reaches 0.5–0.7.

Induction of Protein Expression

IPTG is added to a final concentration of 0.3–0.5 mM before the culture is transferred to a 15°C air shaker (one can achieve a 15°C incubation by installing a normal 37°C shaker in a 4°C cold room and adjusting the heating of the shaker). The time required for the sample to cool from 37°C to 15°C allows the accumulation of sufficient T7 RNA polymerase for protein expression. Induction at 15°C should be conducted overnight. Other induction conditions such as 37°C for 2 hours, 30°C for 6 hours, or 20–25°C overnight may also be tested.

Cell Harvest

The cells from the IPTG-induced culture are spun down at 5000 x g for 15 minutes at 4°C and the supernatant is discarded. At this stage the cell pellet can be stored at –20°C.

Affinity Purification and On-column Cleavage:

Preparation of Chitin Column

The chitin-binding domain(s) (CBD) present in the intein-tag(s) allows affinity purification of the fusion protein. The CBD has a high affinity for chitin beads, which allows efficient and facile recovery of the fusion protein from the crude cell extract. In addition, stringent wash conditions (*e.g.* high salt concentration or the use of nonionic detergents) can be employed to reduce nonspecific binding, thus increasing purity. Generally, a column packed with 10 ml of chitin beads (bed volume) should be used for a one-liter culture. The chitin column should be equilibrated at 4°C with 10 column volumes of the appropriate buffer for the intein construct being used prior to the loading of the crude cell extract.

Preparation of Crude Cell Extract

The cell pellet from a one-liter culture is resuspended in at least 100 ml of the appropriate ice-cold cell lysis buffer (See the section corresponding to the intein fusion used). The inclusion of nonionic detergents in the cell lysis buffer can reduce nonspecific protein binding to the chitin resin during the affinity column step. Nonionic detergents should only be used if they do not inactivate the target protein. We recommend using either 0.1% Triton X-100 or 0.1% Tween 20 in the cell lysis buffer. Furthermore, protein degradation by the action of proteases can be minimized by the inclusion of protease inhibitors such as Phenylmethylsulfonyl fluoride (PMSF). Oxidation sensitive proteins can be stabilized during purification by using the reducing agents TCEP [tris-(2-carboxyethyl)phosphine] or TCCP [tris-(2-cyanoethyl)phosphine] (0.1 mM) in the cell lysis buffer. Following resuspension of the cell pellet, the cells can be lysed by sonication or by the use of a French Press. Egg white lysozyme is not recommended for cell lysis because it is known to bind and degrade the chitin beads. However, if a sonicator or a French press is not available, a low level of lysozyme (10–20 µg/ml) can be used (incubate at 4°C for 1 hour). An increase in viscosity indicates when the cells are broken. If the mixture becomes extremely viscous, it may be necessary to dilute the cell lysate and/or add 10 µg/ml of protease-free DNase plus $MgCl_2$ (to 5 mM final concentration) to reduce viscosity.

The supernatant from the centrifugation of the lysed cells at 19,000 x g for 30 minutes is saved and referred to as the clarified cell extract. Samples (40 µl) are taken from the total cell extract (before centrifugation) and the clarified cell extract (supernatant after centrifugation), mixed with 20 µl 3X SDS Sample Buffer and analyzed by SDS-PAGE and/or Western blot. If the intein-tag target protein fusion precursor is detected in the total cell extract but not in the clarified cell extract, this may suggest that the fusion protein is expressed but in a form that is not soluble.

Loading the Clarified Cell Extract

The clarified cell extract is slowly loaded onto the chitin column at a flow rate no faster than 0.5–1 ml/min. A sample (40 µl) from the flow-through (undiluted) is mixed with 20 µl 3X SDS Sample Buffer and analyzed by SDS-PAGE. Comparison of the flow-through and the clarified cell extract samples indicates the binding efficiency of the fusion precursor to the chitin column.

Washing the Chitin Column

Due to the high affinity of the chitin binding domain for the chitin beads, a higher flow rate (e.g., 2 ml/min) and stringent wash conditions (high salt concentration (1M NaCl) and/or nonionic detergents) can be used. Loading and washing at high salt with nonionic detergents may reduce nonspecific binding of other *E. coli* proteins. Normally, no less than 20 bed volumes of the appropriate buffer is required to thoroughly wash the column.

Induction of On-column Cleavage

The conditions necessary to induce on column cleavage of the intein-tag will depend on the intein or inteins present in the precursor fusion protein and the desired protein product (e.g. thioester-tagged, circular, etc). The specific conditions for on-column cleavage are discussed in the sections describing the intein fusion of interest.

Elution of the Target Protein

Following on-column cleavage of the intein or inteins, the target protein is released from the intein-tag and can be eluted from the column using the appropriate buffer or a specific storage buffer. Fraction sizes of about 1/3 the column bed volume (for example 3 ml fractions for a 10 ml bed volume column) typically results in the target protein being eluted in the first few fractions.

The protein concentration in each fraction can be determined by the Bradford dye binding assay (23) and the eluted fractions should be analyzed by SDS-PAGE. If the protein concentration is too low for SDS-PAGE, acetone can be used to precipitate the protein before electrophoretic analysis (add 0.6 ml acetone to a 0.2 ml aliquot of protein sample, mix and store at –20°C for 30 minutes. Centrifuge the solution in a microfuge and carefully pour off the supernatant. The pellet (white precipitate) is resuspended with 40 µl of 1X SDS Sample Buffer and boiled before electrophoresis). At this point, some chitin resin can be removed to determine the cleavage efficiency (see Stripping the chitin resin below). If a large amount of the precursor still remains uncleaved, the incubation of the column should continue for an additional 12–24 hours before a second elution is conducted.

Stripping the Chitin Resin

The uncleaved fusion precursor protein and the intein-tag remain bound to the chitin resin during the target protein elution and can be stripped from the resin by 1% SDS. A 200 µl resin slurry is gently removed from the chitin column and mixed with 100 µl 3X SDS Sample Buffer. After boiling for 5 minutes, the resin is spun down. The supernatant (3–10 µl) is directly used for SDS-PAGE analysis. Alternatively, the bound proteins can be eluted by passing 3 bed volumes of the Stripping Buffer, containing 1% SDS, through the column. The elution should be conducted at room temperature to prevent the precipitation of the SDS. Since protein concentrations cannot be determined by the Bradford dye binding assay when SDS is present, the samples should be examined by SDS-PAGE.

Regeneration of the Chitin Resin

The chitin resin can be regenerated 4–5 times by the following protocol. Wash the chitin resin with 3 bed volumes of 0.3 M NaOH (stripping solution). Allow the resin to soak for 30 minutes and then wash with an additional 7 bed volumes of stripping solution. Rinse with 20 bed volumes of water followed by 5 bed volumes of column buffer. The resin can be stored at 4°C. For long term storage 0.02% sodium azide should be added to the column buffer.

Simplified Purification Protocol for Test Experiment:

The following is a protocol for a pilot experiment which can serve as a guideline for testing optimal conditions for the expression and purification of a target protein. Samples taken throughout the experiment can be analyzed by SDS-PAGE.

1. Culture Preparation: Inoculate 1 liter of LB media supplemented with ampicillin (100 µg/ml final concentration) with a freshly grown colony or 10 ml of freshly grown culture. Incubate the 1 liter culture at 37°C until an OD_{600} of 0.5–0.7 is reached.

2. Induction: The 1 liter culture should be split into 5 samples (200 ml each in 1 liter flasks) to test for optimal expression conditions. One sample should be used as a control for uninduced cells (no IPTG).

   Note: The optimal incubation temperature and time for induction will vary depending on the target protein. We recommend testing several conditions such as 37°C for 2 hours, 30°C for 3 hours, 22–25°C for 6 hours and 15°C overnight using 0.3 mM IPTG.

3. Cell Harvest: Spin down the 200 ml cultures at 5000 x g for 10 minutes at 4°C. Discard the supernatant. If desired, the pellet can be stored at –20°C for future use.

   Note: Steps 4–9 should be performed at 4°C, unless otherwise stated.

4. Crude Cell Extracts: Resuspend each cell pellet in 20 ml of ice-cold Buffer B1. Lyse the cells by sonication on ice. Remove 40 µl of the crude cell extract and mix with 20 µl of 3X SDS Sample Buffer (Sample 1: crude extract/uninduced; Sample 2: crude extract/induced cells). Prepare clarified cell extracts by centrifugation at 19,000 x g for 30 minutes. Save the pellet at –20°C for future analysis. If the target protein is absent from the clarified cell extract, this may indicate a solubility problem in which case the pellet should be tested for the presence of insoluble fusion protein. Remove 40 µl of the clarified cell extract (the supernatant) and mix with 20 µl of 3X SDS Sample Buffer (Sample 3: clarified cell extract). The samples saved from each step can be analyzed by SDS-PAGE and/or by Western blot with the Anti-Chitin Binding Domain Serum to detect the fusion protein.

5. Chitin Column Equilibration: Aliquot the chitin bead slurry into separate columns (6 ml bed volume), 1 column for each clarified cell extract sample. Equilibrate the columns by passing 60 ml of the appropriate buffer (the same buffer used in step 4) through the column.

6. Loading: Slowly load the clarified cell extracts onto the chitin columns (at a flow rate of no more than 0.5–1.0 ml/min). Remove 40 µl of the flow-through and mix with 20 µl 3X SDS Sample Buffer (Sample 4). This sample, when compared with the clarified cell extract sample, illustrates the binding efficiency of the construct.

7. Washing the Column: Wash the column with 100 ml of Buffer B1 at a flow rate of 2 ml/min. Be sure that all traces of crude cell extract have been washed off the sides of the column.

8. On-column Cleavage Reaction:

   Note: The conditions for on-column cleavage depend on whether the intein-tag is fused to the N-terminus, the C-terminus, or the N-and C-termini of the target protein. Choose the description below that corresponds to the intein fusion being used.

8a. On-column Cleavage Reaction for an Intein-tag fused to the N-terminus of the target protein: Flush the column quickly with 18 ml (or about 3 column volumes) of Buffer B2 [20 mM HEPES or Tris-HCl, pH 6.5 containing 500 mM NaCl, and 1 mM EDTA]. Remove 40 µl of the flow through from the quick flush and mix with 20 µl 3X SDS Sample Buffer (Sample 5a). Stop the column flow, and leave at room temperature overnight.

8b. On-column Cleavage Reaction for an Intein-tag fused to the C-terminus of the target protein: Flush the column quickly with 18 ml (or about 3 column volumes) of Buffer B3 [20 mM HEPES or Tris-HCl, pH 8.5 containing 40 mM DTT, 500 mM NaCl, and 1 mM EDTA]. If a thioester tagged target protein is desired, 50 mM MESNA should be substituted for the DTT in Buffer B3 (see Buffer B4 in Media and Solutions). Remove 40 µl of the flow through from the quick flush and mix with 20 µl 3X SDS Sample Buffer (Sample 5b). Stop the column flow, and leave at 4°C overnight.

8c. On-column Cleavage Reaction for a target protein fused between 2 intein-tags: Flush the column quickly with 18 ml (or about 3 column volumes) of Buffer B2 [20 mM HEPES or

Tris-HCl, pH 6.5 containing 500 mM NaCl, and 1 mM EDTA]. Remove 40 µl of the flow through from the quick flush and mix with 20 µl 3X SDS Sample Buffer (Sample 5c). Stop the column flow, and leave at room temperature overnight. Following the overnight incubation, wash the column with 60 ml (or 10 column bed volumes) of Buffer B2. Stop the column flow and flush the column quickly with 18 ml (or about 3 column volumes) of Buffer B3 [20 mM HEPES or Tris-HCl, pH 8.5 containing 40 mM DTT, 500 mM NaCl, and 1 mM EDTA]. For the formation of a circular or thioester-tagged protein for use in IPL reactions, 50 mM MESNA should be substituted for the 40 mM DTT in Buffer B3 (See Buffer B4 in Media and Solutions). Remove 40 µl of the flow through from the quick flush and mix with 20 µl 3X SDS Sample Buffer (Sample 5d). Stop the column flow, and leave at 4°C overnight.

9. Target protein elution: Elute the target protein using additional cleavage buffer (see above) and collect 2 ml-fractions for 3 column volumes. Monitor fractions by the Bradford dye binding assay to determine the concentration of the target protein in each fraction. Remove 40 µl from the fractions with the highest protein concentrations and mix with 20 µl 3X SDS Sample Buffer (Sample 6). If the protein concentration is too low for SDS-PAGE, add 0.6 ml acetone to a 0.2 ml aliquot, mix and store at –20°C for 30 minutes. Centrifuge and carefully pour off the supernatant. Then add 40 µl of 1X SDS Sample Buffer to the pellet (white precipitate) and boil 5 minutes before electrophoresis.

10. Fusion Precursor Elution: The remaining intein-tag as well as uncleaved fusion precursor can be stripped from the column with 1% SDS. To assess cleavage efficiency, remove 200 µl chitin resin and mix with 100 µl 3X SDS Sample Buffer. After boiling for 5 minutes, the supernatant is analyzed on SDS-PAGE to determine the cleavage efficiency.

Media and Solutions:

The following are suggested media for cell culture, cell lysis and protein purification. They can be modified according to the specific properties of the target protein.

LB broth (per liter)
10 g tryptone
5 g yeast extract
10 g NaCl
Adjust pH to 7.0 with NaOH

Buffer B1*
20 mM Na-HEPES (alternatively, Tris-HCl or Na-Phosphate) (pH 8.5)
500 mM NaCl (or 50–1000 mM NaCl)
1 mM EDTA

**Note: When Buffer B1 is being used for cell lysis a nonionic detergent, e.g., 0.1–0.2% Tween 20 and protease inhibitors, e.g., PMSF (20 µM) can be included. For a target protein sensitive to oxidation, a low concentration (1 mM) of DTT or β-mercaptoethanol, or 0.1 mM TCEP [tris-(2-carboxyethyl)phosphine] or TCCP [tris-(2-cyanoethyl)phosphine] may be used.*

Buffer B2*
20 mM Na-HEPES (alternatively, Tris-HCl or Na-Phosphate), pH 6.5
500 mM NaCl (or 50–1000 mM NaCl)
1 mM EDTA

**Note: The pH of Buffer B2 can be varied from 6.0–7.0. When Buffer B2 is being used for cell lysis a nonionic detergent, e.g., 0.1-0.2% Tween 20 and protease inhibitors, e.g., PMSF (20 µM) can be included. For a target protein sensitive to oxidation, a low concentration (1 mM) of DTT or β-mercaptoethanol, or 0.1 mM of TCEP [tris-(2-carboxyethyl)phosphine] or TCCP [tris-(2-cyanoethyl)phosphine] may be used.*

Buffer B3
20 mM HEPES (or Tris-HCl), pH 8.5
500 mM NaCl (or 50–1000 mM NaCl)
40 mM DTT
1 mM EDTA

Buffer B4 (to generate a thioester-tagged protein for IPL)
20 mM HEPES (or Tris-HCl), pH 8.5
500 mM NaCl (or 50–1000 mM NaCl)
50 mM 2-mercaptoethanesulfonic acid
1 mM EDTA

Stripping Solution
0.3 M NaOH

Frequently Asked Questions (FAQs):

Cloning

- *Should I fuse an intein-tag to the C-terminus, N-terminus or to both of the termini of my target protein?*

  This depends on the desired outcome. If the system is being used strictly for protein purification, then an N- or C-terminal fusion is usually sufficient. It is advisable to try both the N- and C-terminal fusions because it is conceivable that different target proteins, due to certain structural constraints, may prefer one construct over the other to allow the proper folding of the fusion precursor and a high level of protein expression. Our studies show that some target proteins have higher yields when expressed as an N-terminal fusion, whereas others give higher yields as a C-terminal fusion.

  However, if the target protein requires a defined N-terminus, such as an N-terminal cysteine, then either the fusion of an intein-tag to the N-terminus or to the N- and C-termini of the target protein should be used. If a reactive C-terminal thioester is desired then the C-terminal or the N- and C-terminal fusions should be used. For the generation of a thioester-tagged protein MESNA should be used in place of DTT.

  If a protein species with an N-terminal cysteine and a C-terminal thioester is desired for reactions such as the generation of cyclic protein species then intein-tags should be fused to both the N-and C-termini of the target protein. MESNA should be used in place of DTT for these reactions.

- *The target protein contains an unfavorable residue such as aspartate at its C-terminus. Can I still use the intein-tag fused to the target protein's C-terminus?*

  The data in Appendix I are based on our studies using the maltose-binding protein (MBP) as the target protein. It is possible that an aspartate residue at the C-terminus may not affect the *in vivo* or thiol-induced cleavage of the intein in the context of a different target protein. However, it is advisable to also add an extra amino acid residue that is more favorable for controlled cleavage [e.g. a Met or Tyr for the *Mxe* GyrA intein (pTWIN1) or a Gly for the *Mth* RIR1 intein (pTWIN2)]. See FAQ # 2.12 at www.neb.com.

Expression

- *What factors contribute to poor expression and low yields of some proteins?*

  The yield of a target protein is generally determined by (i) the expression level of the fusion precursor; (ii) the solubility of the fusion protein; (iii) the controlled cleavage of the intein-tag; and/or (iv) the solubility of the target protein after cleavage. The expression level of the fusion protein is mostly influenced by the target protein. Poor codon usage, mRNA degradation or proteolysis due to protein misfolding may all contribute to poor expression. Different growth and induction conditions should be tested to optimize the expression of the fusion protein. Induction at lower temperatures may reduce the formation of inclusion bodies (i.e., improve the solubility of the fusion protein) as well as reduce the level of proteolysis. Protease deficient hosts should also be tested to minimize proteolysis. Optimizing codon usage may be helpful in increasing the expression level. Another possibility for poor expression could be that the clone is not stable due to toxicity of the target protein to host cells. Accordingly, one should inoculate the medium with a freshly grown colony and induce the expression at lower temperatures with lower IPTG concentrations. Some target proteins become insoluble after on-column cleavage and therefore are only eluted after incubation of the resin with SDS. In this case, one may try to increase the salt concentration (0.5–1 M NaCl) or add a nonionic detergent to the buffer used during cleavage to improve the solubility of the target protein.

- *What does it mean if the western blot analysis with Anti-Chitin Binding Domain Serum detects a product corresponding to the intein and CBD tag but not the full length fusion protein in the crude cell extract?*

  The data indicate that either the target protein is degraded or that the intein cleaved *in vivo*. One should make sure that the samples were prepared in SDS Sample Buffer without DTT or β-mercaptoethanol since boiling in DTT-containing sample buffer may cause cleavage of the fusion protein. Perform a Western blot with anti-target protein serum to differentiate between proteolysis and intein-mediated cleavage. If proteolysis is evident, try different hosts. *In vivo* cleavage may be reduced by inducing cells at lower temperatures (for example, 12–15°C overnight or room temperature for 3–6 hours). You may also try adding residues that are optimal for controlled cleavage to the protein (for example GRA or CRA at the N-terminus of the target protein for use with the *Ssp* DnaB intein, M or MRM at the C-terminus for use with the *Mxe* GyrA intein, or G for use with the *Mth* RIR1 intein).

Purification

- *Is there an alternative way to break cells other than sonication?*

  Cells can also be broken with a French press, or by the addition of egg white lysozyme. However, since egg white lysozyme is known to bind and digest chitin, it's use is not recommended. If a sonicator or a French press is not available, try a low level of lysozyme (10–20 µg/ml) and incubate at 4°C for 1 hour. The increase in viscosity indicates when the cells are broken. If the mixture becomes extremely viscous, it may be necessary to add 10 µg/ml of protease-free DNase plus $MgCl_2$ (to 5 mM final concentration) to reduce viscosity before the clarified extract can be passed through the chitin column.

- *Can I use a thiol reducing agent during purification to protect my protein from oxidation?*

  Yes, a low concentration (1 mM) of dithiothreitol (DTT) or β-mercaptoethanol (β-ME) can be used throughout the purification procedures as long as the purification (before the on-column cleavage step) can be finished in a short period of time. Alternatively, TCEP [tris-(2-carboxyethyl)phosphine] or TCCP [tris-(2-cyanoethyl)phosphine] can be used at 0.1 mM final concentration in the cell lysis and column buffers to stabilize oxidation-sensitive proteins during purification. These compounds specifically reduce disulfide bonds without affecting the intein-mediated cleavage reaction and thus can be used to stabilize proteins with essential thiols (Burns, J.A., *et. al. J. Org. Chem.* 56, 2648–2650). Furthermore, TCEP or TCCP (5 mM) may also be used as a reducing agent in the SDS Sample Buffer.

- *When using the N- and C-terminal fusion construct to generate cyclic proteins both the linear and circular forms are purified from the chitin resin. Is there a way to get only the circular species?*

  Currently, it is not possible to elute only the circular protein species from the chitin resin while having the linear form remain bound. The ratio of linear to circular protein will vary dramatically depending on the ease with which the N- and C-termini of the target protein can come into contact and react. One potential way to separate the linear and circular protein forms is to use a His tag and an exo-protease as described previously (11).

Cleavage

- *What should I do if the fusion precursor is the major product from the SDS elution of the chitin resin (after induction of on-column cleavage)?*

  This means that the induced on-column cleavage is not efficient, which invariably leads to a low yield of the target protein. One may try the following options to increase the cleavage efficiency: (i) increase the temperature of the on-column cleavage; An 8 hour incubation at 16–23°C may result in more released target protein than a 16 or 40 hour incubation at 4°C; (ii) for intein1 cleavage, lower the pH value of the cleavage buffer from pH 7.0 to pH 6.0 or pH 6.5; (iii) increase the duration of the on-column cleavage; If the target protein is sensitive to high temperatures, a longer incubation time at a lower temperature may be needed; (iv) change the residue(s) adjacent to the intein cleavage site. Often the low cleavage efficiency is caused by an unfavorable terminal residue of the target protein which is placed adjacent to the intein cleavage site when Sap I is used for cloning. Appendix I may serve as a guideline for placing a suitable residue adjacent to the cleavage site. Sometimes, it is necessary to add more than one residue to the terminus of a target protein in order to achieve a high cleavage efficiency.

- *What does it mean if the target protein is not eluted after on-column cleavage, but is present in the SDS elution fractions?*

  If both the target protein and the intein-tag (or tags) are present in the SDS elution fractions, it suggests that the target protein becomes insoluble after induced on-column cleavage. One may try to increase the salt concentration (0.5–1 M NaCl) or add a nonionic detergent to the cleavage buffer to improve the solubility of the target protein. A number of nonionic detergents examined (0.1–0.5% Triton X-100 or 0.1–0.2% Tween 20) had little effect on binding or cleavage. If urea is used to elute the column, some intein-tag may also co-elute with the target protein. In this case one may need to repurify and refold the target protein.

- *If my target protein is sensitive to DTT, are there alternative means to induce the on-column cleavage of the C-terminal intein-tag fusions?*

  The first choice would be to use the *Ssp* DnaB intein as an N-terminal intein-tag (see the N-terminal Fusion section under purification strategies) instead of a C-terminal intein-tag

fusion. Cleavage of the *Ssp* DnaB intein-tag is induced by pH and temperature shifts and does not require the use of DTT. Alternatively, if a C-terminal fusion is preferred, lower concentrations of DTT or β-mercaptoethanol (5–10 mM) may be used for on-column cleavage. However, longer incubation time or higher temperatures (up to room temperature) may be required for efficient cleavage. Also, 50 mM freshly prepared hydroxylamine can be used to induce cleavage at 4–25°C of the C-terminal intein-tag. Be aware that when hydroxylamine is used, it forms a stable covalent bond with the C-terminus of the target protein. One should determine whether a C-terminal hydroxamate affects the activity of the target protein.

- *The target protein has optimal activity at low salt concentrations. How do I perform on-column cleavage?*

  Following on-column cleavage and target protein elution from the chitin resin, the target protein buffer can be changed by dialysis. Alternatively, the cleavage buffer can be made to match the optimal salt concentration for the enzyme, however salt concentrations of < 50 mM are not recommended. To carry out cleavage in a low salt buffer the column should first be washed extensively (10 volumes) with high salt buffer (this removes proteins that nonspecifically bind to chitin). The chitin resin is then flushed quickly with 3 volumes of the low salt cleavage buffer. This ensures that the target protein is eluted in the specified buffer.

- *After the on-column cleavage, the SDS-PAGE analysis shows that the target protein is contaminated with other bands. What could these proteins be?*

  These may be *E. coli* proteins that nonspecifically bind to chitin beads. Generally, washing the column extensively with buffer containing high salt (0.5–1 M NaCl) and nonionic detergent will effectively reduce the background. If lysozyme is used for breaking cells, it may be present in the eluant since lysozyme can bind and digest chitin. It is also possible that the proteins have been eluted by virtue of their affinity for the target protein. For example, the *E. coli* host chaperone protein GroEL (~60 kDa) has been seen to co-purify with some eucaryotic proteins or mutant proteins which fold poorly in *E. coli* host cells. Addition of ATP to the wash buffer may allow separation of the GroEL protein from the target protein. The target protein may be heterogeneous owing to proteolytic degradation. Finally, in some cases, small amounts of the fusion protein or the intein-tag may appear in the eluate following the on-column cleavage reaction. The latter two possibilities can be checked by Western blot analysis using antibodies against the target protein and the chitin binding domain. Use of protease inhibitors, protease-deficient hosts and induction at low temperatures (15°C for 16 hours) may reduce proteolytic degradation.

- *How do I remove DTT after cleavage?*

  After elution of the target protein, free DTT can be removed from the sample by dialysis.

More FAQs can be found at www.neb.com--technical reference--gene expression and protein purification--FAQs for IMPACT Kit (www.neb.com/nebecomm/tech_reference/gene_expression/IMPACTFaq.asp)

References:

1. Xu, M.-Q. and Evans Jr., T. C. (2003). *E. coli* Gene Expression Protocols. In P. E. Vaillancourt (Ed.), *Methods in Mol. Biol.* Vol. 205 (pp. 43–68). Totowa, N.J.:Human Press Inc.
2. Xu, M.-Q. and Evans Jr., T. C. (2001) *Methods.* 24, 257–277.
3. Evans, J., T. C., Benner, J. and Xu, M.-Q. (1998) *Protein Sci.* 7, 2256–2264.
4. Evans, J., T. C., Benner, J. and Xu, M.-Q. (1999) *J. Biol. Chem.* 274, 3923–3926.
5. Evans, J., T. C., Benner, J. and Xu, M.-Q. (1999) *J. Biol. Chem.* 274, 18359–18363.
6. Camarero, J. A. and Muir, T. W. (1999) *J. Amer. Chem. Soc.* 121, 5597–5598.
7. Cotton, G. J., Ayers, B., Xu, R. and Muir, T. W. (1999) *J. Am. Chem. Soc.* 121, 1100–1101.
8. Muir, T. W., Sondhi, D. and Cole, P. A. (1998) *Proc. Natl. Acad. Sci. USA* 95, 6705–6710.
9. Severinov, K. and Muir, T. W. (1998) *J. Biol. Chem.* 273, 16205–16209.
10. Xu, R., Ayers, B., Cowburn, D. and Muir, T. W. (1999) *Proc. Natl. Acad. Sci. USA* 96(2), 388–393.
11. Iwai, H. and Pluckthun, A. (1999) *FEBS Lett* 459, 166–172.
12. Dawson, P. E., Muir, T. W., Clark-Lewis, I. and Kent, S. B. (1994) *Science* 266, 776–779.
13. Tam, J. P., Lu, Y.-A., Liu, C.-F. and Shao, J. (1995) *Proc. Natl. Acad. Sci. USA* 92, 12485–12489.
14. Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F. B. and Xu, M.-Q. (1996) *J. Biol. Chem.* 271, 22159–22168.
15. Xu, M.-Q. and Perler, F. B. (1996) *EMBO J.* 15(19), 5146–5153.
16. Watanabe, T., Ito, Y., Yamada, T., Hashimoto, M., Sekine, S. and Tanaka, H. (1994) *J. Bacteriol.* 176, 4465–4472.
17. Wu, H., Xu, M.-Q. and Liu, X.-Q. (1998) *Biochim Biophys Acta* 1387, 422–432.
18. Mathys, S., Evans, J., T. C., Chute, I. C., Wu, H., Chong, S., Benner, J., Liu, X.-Q. and Xu, M.-Q. (1999) *Gene* 231, 1–13.
19. Telenti, A., Southworth, M., Alcaide, F., Daugelat, S., Jacobs, J., William, R. and Perler, F. B. (1997) *J. Bacteriol.* 179, 6378–6382.
20. Smith, D. R., Doucette-Stamm, L. A., Deloughery, C., Lee, H., Dubois, J., Aldredge, T., Bashirzadeh, R., Blakely, D., Cook, R., Gilbert, K., Harrison, D., Hoang, L., Keagle, P., Lumm, W., Pothier, B., Qiu, D., Spadafora, R., Vicaire, R., Wang, Y., Wierzbowski, J., Gibson, R., Jiwani, N., Caruso, A., Bush, D., Safer, H., Patwell, D., Prabhakar, S., McDougall, S., Shimer, G., Goyal, A., Pietrokovski, S., Church, G. M., Daniels, C. J., Mao, J.-I., Rice, P., Nolling, J. and Reeve, J. N. (1997) *J. Bacteriol.* 179(22), 7135–7155.
21. Southworth, M. W., Amaya, K., Evans, J., T. C., Xu, M.-Q. and Perler, F. B. (1999) *Biotechniques* 27, 110–120.
22. Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W. (1990) *Methods Enzymol.* 185, 60–89.
23. Sambrook, J. and Russell D. (2001). *Molecular Cloning: A Laboratory Manual*, (2nd ed.), (pp. 1.105–1.122). Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
24. Perler, F. B., Davis, E. O., Dean, G. E., Gimble, F. S., Jack, W. E., Neff, N., Noren, C. J., Thorner, J. and Belfort, M. (1994) *Nucl. Acids Res.* 22, 1125–1127.
25. Chong, S., Mersha, F. B., Comb, D. G., Scott, M. E., Landry, D., Vence, L. M., Perler, F. B., Benner, J., Kucera, R. B., Hirvonen, C. A., Pelletier, J. J., Paulus, H. and Xu, M.-Q. (1997) *Gene* 192(2), 271–281.
26. Chong, S., Montello, G. E., Zhang, A., Cantor, E. J., Liao, W., Xu, M.-Q. and Benner, J. (1998) *Nucl. Acids Res.* 26(22), 5109–5115.
27. Wood, D. W., Wu, W., Belfort, G., Derbyshire, V. and Belfort, M. (1999) *Nature Biotechnology* 17(9), 889–892.
28. Bradford, M. M. (1976) *Anal. Biochem.* 72, 248–254.

Appendix I: Comparison Between Intein-Tag Fusions

N-and C-terminal Fusions

The IMPACT-TWIN™ system allows the fusion of a self-cleavable mini intein-tag to the C-terminus, the N-terminus or both the N-and C-termini of a target protein by using the appropriate cloning sites. Different target proteins, due to certain structural constraints, may prefer a particular fusion strategy to allow proper folding of the precursor protein and a higher level of protein expression. The following table lists some differences and uses of the various fusion constructs.

	pTWIN1[a]	pTWIN1 or pTWIN2	pTWIN1[b]
Fusion	C-terminal	N-terminal	N- and C-termini
Cloning	Nde I to Sap I	Sap I to Pst I	Sap I to Sap I
N-terminus[c] of the target protein	Met	Cys-Arg, Gly-Arg, and Ser-Arg are recommended; Pro is not recommended; other amino acids depend on the influence of target protein	Cys-Arg, Gly-Arg, and Ser-Arg are recommended; Pro is not recommended; other amino acids depend on the influence of target protein
C-terminus[c] of the target protein	Met, Tyr and Phe are recommended; Asp, Pro, Gly and Ala are not recommended; other amino acid residues depend on the influence of the target protein	Any amino Acid	Met, Tyr and Phe are recommended Asp, Pro, Gly and Ala are not recommended; other amino acid residues depend on the influence of the target protein
Expression	Influenced by target protein	Less affected by target protein	Less affected by target protein
Purification			
Induction of on-column cleavage	Thiol reagents such as DTT and MESNA	Shift from pH 8.5 buffer at 4°C to pH 7.0 buffer at room temperature	Shift from pH 8.5 buffer at 4°C to pH 7.0 buffer at room temperature followed by addition of a thiol reagent such as DTT or MESNA
Sensitivity to pH	pH 7–9, cleavage favors basic pH	pH affects cleavage activity	pH affects cleavage activity of the *Ssp* DnaB intein; the *Mxe* GyrA thiol-induced cleavage favors basic pH
Incubation	Overnight at 4–25°C	Overnight at 20–25°C; incubation can also be at 4°C but will require a longer time	Overnight at 20–25°C for the *Ssp* DnaB intein followed by overnight at 4–25°C for *Mxe* GyrA intein
Application	Purification; C-terminal thioester generation for IPL	Purification; isolation of a protein with an N-terminal amino acid other than Met; generation of a protein with an N-terminal Cys for IPL	Purification; isolation of a protein with an N-terminal amino acid other than Met; producing cyclic protein species

[a] cloning into the pTWIN2 Nde I to Sap I sites results in the Mth RIR1 intein being fused to the C-terminus of the target protein. It is recommended that a Gly or Ala be present at the C-terminus of the target protein. Asp and Pro are not recommended and all other amino acids will depend on the influence of the target protein. The other conditions are as described for pTWIN1.

[b] cloning into the pTWIN2 Sap I sites results in the Ssp DnaB intein and the Mth RIR1 intein fused to the N-and C-termini of the target protein, respectively. The Ssp DnaB intein characteristics in pTWIN2 are the same as in pTWIN1. The amino acids Asp and Pro are not recommended at the C-terminus of the target protein when using the Mth RIR1 intein, while Gly and Ala are recommended. Other amino acids will depend on the influence of the target protein. The other conditions are as described for pTWIN1.

[c] for references on the cleavage activity of these inteins see References 1–5, 18, and 21.

Appendix II: Protein Labeling and Intein-mediated Protein Ligation

Labeling of the Target Protein

The C-terminus of the target protein can be covalently labeled using L[^{35}S]-cysteine or a biotinylated synthetic peptide with an N-terminal cysteine immediately following thiol-induced on-column cleavage and elution. Currently NEB has available 2 peptides for protein labeling. One contains a fluorescein group (NEB #P6606S) and the other a biotin (NEB #P6607S). Typically 2-mercaptoethanesulfonic acid (MESNA) is used as the thiol reagent to induce intein-mediated cleavage. This produces a C-terminal thioester on the target protein. The following protocol illustrates a typical labeling experiment. Mix 4 µl of L[^{35}S]-cysteine (11.0 mCi/ml, 0.0102 µmol/ml, available through NEN) or 4 µl of biotinylated peptide (10 mM, NH_2-CDPEK*DS-COOH, in which the biotin was incorporated as biotinylated lysine during synthesis) with a 36 µl aliquot of the freshly purified protein sample and incubate the reaction mixture (pH 8.0) at 4°C overnight. To examine the labeled protein, add 20 µl of 3X SDS Sample Buffer (with DTT) to the protein sample, boil for 5 minutes and run on an SDS-PAGE gel. The L-[^{35}S]-cysteine sample can be analyzed by SDS-PAGE and autoradiography. The biotinylated protein sample can be analyzed by SDS-PAGE followed by a Western blot with anti-biotin antibody (Cell Signaling Technology #7055 or #7075). Any peptide or protein sample can be substituted for the biotinylated peptide as long as it contains an N-terminal cysteine residue.

Intein-mediated Protein Ligation (IPL)

The IPL reaction, also referred to as expressed protein ligation, allows the ligation of a bacterially expressed protein or a synthetic peptide with an N-terminal cysteine residue to a bacterially expressed protein or synthetic peptide with a C-terminal thioester through a native peptide bond (1-10,18). The IPL protocol applies a previously described chemoselective ligation strategy in which an N-terminal cysteine of a peptide attacks a C-terminal thioester present on another peptide (12,13). This results in a peptide bond at the site of ligation. The pTWIN vectors are designed to generate reactants for IPL, however the performance of the IPL reaction itself will vary significantly depending on the protein or peptide reactants used. For intermolecular IPL reactions, ones that do not involve cyclization, it is necessary to have the reactants as

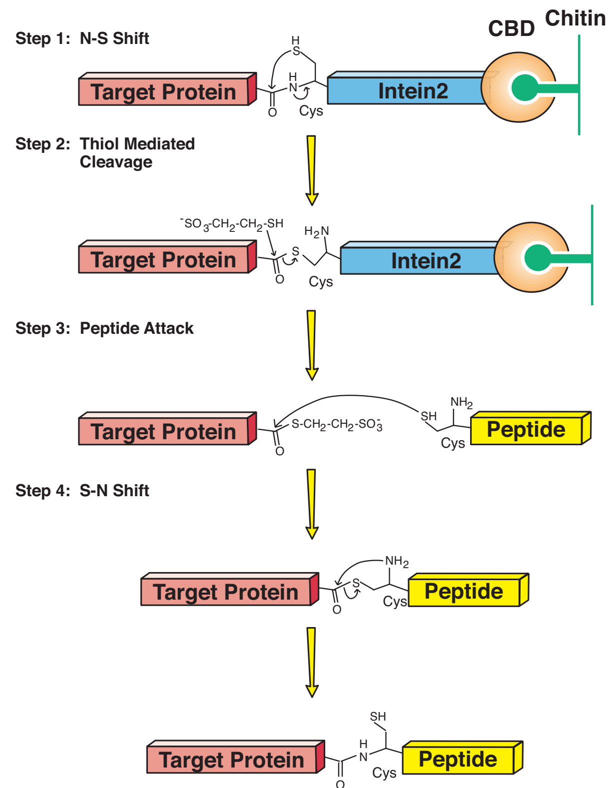


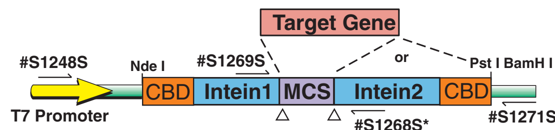
Figure 12: *The Mechanism of Intein-mediated Protein Ligation (IPL).*

concentrated as reasonably possible. In an ideal case, at least one of the reactants should be close to a concentration of 1 mM. Because the cyclization is an intramolecular reaction, the reactant concentration is not critical. By cloning a gene into the appropriate pTWIN restriction sites it is possible to isolate a target protein with an N-terminal cysteine, a C-terminal thioester, or both. The presence of both an N-terminal cysteine and a C-terminal thioester allows the *in vitro* cyclization of a target protein with a peptide bond at the site of fusion (5,6,11).

Appendix III: Sequencing Primers

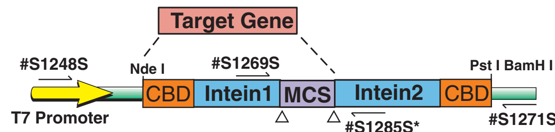
There are a variety of sequencing primers available for checking the target gene following insertion into a pTWIN vector. Three sequencing primers are included with the IMPACT-TWIN kit. These are the *Ssp* DnaB Intein Forward Primer (NEB #S1269S), the *Mxe* Intein Reverse II Primer (NEB #S1285S), and the *Mth* RIR1 Intein Reverse Primer (NEB #S1270S). Two other sequencing primers are sold separately, the T7 Universal Primer (NEB #S1248S) and the T7 Terminator Reverse Primer (NEB #S1271S). The annealing position and direction of sequencing for each primer are shown on the accompanying figure.

Fusion of the N-terminus of a target protein to the C-terminus of Intein1



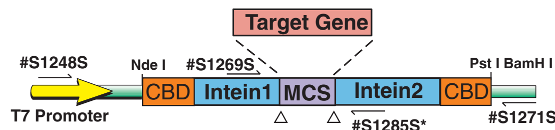
Note: The target gene fragment should contain a translation termination codon.

Fusion of the C-terminus of a target protein to the N-terminus of Intein2



Note: The target gene should not carry a translation termination codon.

Fusion of both N- and C-termini of a target protein to the inteins



Note: The target gene should not carry a translation termination codon.

***Cloning into pTWIN2 requires use of sequencing primer #S1270S in place of #S1285S.**

Figure 13: *The different cloning strategies using pTWIN1 or pTWIN2 and relevant sequencing primers.*

Appendix IV: The Chemical Mechanism of Protein Splicing

Protein splicing is a posttranslational processing event involving the excision of an internal protein segment, the intein, from a precursor protein with the concomitant ligation of the flanking N- and C-terminal regions (the exteins) (14,15,24). Sequence alignment reveals that there are highly conserved residues at the two splice junctions: a cysteine or serine residue at the N-terminus of the intein, His-Asn at the C-terminus of the intein, and Cys, Ser or Thr as the first residue of the C-terminal extein. These conserved splice junction residues are directly involved in the catalysis of peptide bond cleavage and formation during protein splicing reactions.

The chemical mechanism of protein splicing with an intein which has cysteine residues at its N-terminus and adjacent to its C-terminus is shown in Figure 14:

Step 1. Formation of a linear thioester intermediate by an N-S acyl rearrangement of Cys1 at the N-terminus of the intein.

Step 2. Formation of a branched intermediate by transthioesterification involving attack by the Cys immediately following the C-terminus of the intein on the thioester formed in Step 1.

Step 3. Excision of the intein by peptide bond cleavage coupled to succinimide formation involving the intein C-terminal Asn residue.

Step 4. Spontaneous S-N acyl rearrangement of the transitory ligation product from a thioester to a stable amide bond.

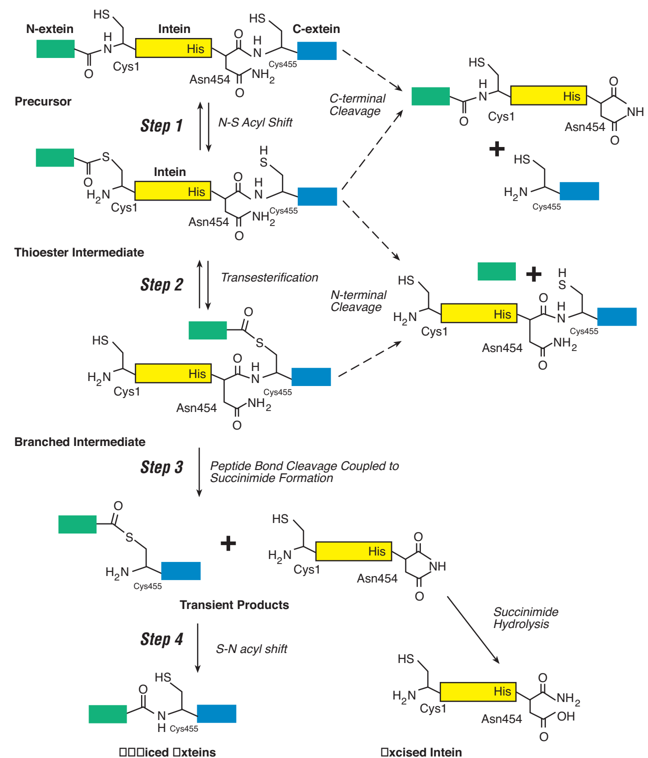


Figure 14: *The chemical mechanism of protein splicing involving the intein from the Saccharomyces cerevisiae VMA1 gene.*

Appendix V: The Chemical Mechanism of the Thiol-inducible and pH-inducible Cleavage Reactions

The understanding of the mechanism of protein splicing has enabled the modulation of protein splicing elements so that there is efficient, controllable peptide bond cleavage at either termini of the intein (1,4,18,21,24–27). The pTWIN system utilizes inteins from the *Ssp dnaB* gene (*Ssp* DnaB intein), the *Mxe gyrA* gene (*Mxe* GyrA intein), and the *Mth rir1* gene (*Mth* RIR1 intein).

The pTWIN vectors allow the fusion of Intein2 (either the *Mxe* GyrA or *Mth* RIR1 intein) to the C-terminus of a target protein. The C-terminal residue (an Asn) of the intein has been mutated to an alanine. This blocks the splicing reaction but still allows an N-S acyl rearrangement to occur at the intein N-terminus (Cys1) resulting in the formation of a thioester linkage between the target protein and the intein. Cleavage of the thioester bond can be induced by thiol reagents, such as 1,4-dithiothreitol (DTT) or 2-mercaptoethanesulfonic acid (Figure 15A). Use of 2-mercaptoethanesulfonic acid results in the formation of a reactive thioester at the C-terminus of the target protein. This thioester can be used in subsequent IPL reactions.

Alternatively, the N-terminus of a target protein can be fused to the C-terminus (an Asn) of Intein1 (the *Ssp* DnaB intein). A chitin binding domain, present at the N-terminus of the *Ssp* DnaB intein, facilitates purification using a chitin resin. The N-terminal cysteine (Cys1) of the intein has been changed to an alanine to block the splicing reaction. The *Ssp* DnaB intein with this mutation undergoes a temperature and pH dependent cleavage of the peptide bond between the C-terminus of the intein and the downstream amino acid. This occurs by the cyclization of the C-terminal Asn side chain to form a succinimide group with the concomitant breakage of the peptide bond (Figure 15B).

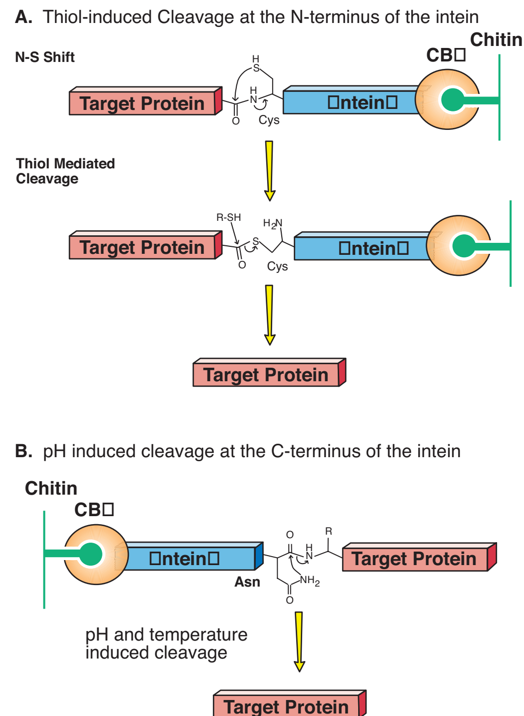


Figure 15: The proposed chemical mechanisms of the inducible peptide bond cleavage reactions for inteins found in the pTWIN vectors. (A) Thiol-induced cleavage of Intein2, which is used for C-terminal fusions (B) pH and temperature induced cleavage of Intein1, which is used for N-terminal fusions.

Appendix VI: Regeneration of Chitin Resin

The chitin resin can be regenerated 4-5 times by the following protocol:
Wash the chitin resin with 3 bed volumes of 0.3 M NaOH (stripping solution). Allow the resin to soak for 30 minutes and then wash with an additional 7 bed volumes of stripping solution. Rinse with 20 bed volumes of water followed by 5 bed volumes of column buffer. The resin can be stored at 4°C. For long term storage 0.02% sodium azide should be added to the column buffer.

Appendix VII: Research Use Assurance Statement

The buyer and user have a non-exclusive sublicense to use this system or any component thereof for **RESEARCH PURPOSES ONLY**, based upon agreement to the following assurances.

Transfer of the host cells that contain the cloned copy of the T7 gene 1 to third parties is explicitly prohibited. This limitation applies to *E. coli* strain *ER2566* which is provided in combination with the IMPACT™-TWIN system or in combination with appropriate vectors for said system.

A license to use this system or any components thereof for commercial purposes may be obtained from New England Biolabs, Inc.

Commercial Laboratory Buyer and User: Use of the host cells that contain the cloned copy of T7 gene 1, the gene for T7 RNA polymerase for any purpose other than in combination with the IMPACT™-TWIN system is explicitly prohibited.

Use of the host cells that contain the copy of the T7 gene 1, the gene for T7 RNA polymerase with any other vector(s) containing a T7 promoter to direct the production of RNA or protein requires a license from Brookhaven National Laboratory. Information about research-use or commercial-use license agreements may be obtained from the Office of Technology Transfer, Brookhaven National Laboratory, Building 475D, P.O. Box 5000, Upton, New York, 11973-5000; telephone: 631-344-7134, fax: 631-344-3729.

You may refuse this non-exclusive research license agreement by returning the enclosed materials unused. By keeping or using the enclosed materials, you agree to be bound by the terms of this sub-license.

Kit Components Sold Separately:

pTWIN1 Vector
#N6951S 10 µg

pTWIN2 Vector
#N6952S 10 µg

pTWIN-MBP1 Vector
#N6953S 10 µg

Ssp DnaB Intein Forward Primer (20-mer)
#S1269S 0.5 A_{260} units

Mxe Intein Reverse II Primer (18-mer)
#S1285S 0.5 A_{260} units

Mth RIR1 Intein Reverse Primer (18-mer)
#S1270S 0.5 A_{260} units

Chitin Beads
#S6651S 20 ml
#S6651L 100 ml

Anti-Chitin Binding Domain Serum
#S6654S 0.05 ml

Blue Loading Buffer Pack
#B7703S

Companion Products:

Flu-P1
#P6606S 0.29 mg

Bio-P1
#P6607S 0.29 mg

IMPACT Kit
#E6901S

pTYB3 Vector
#N6703S 10 µg

pTYB4 Vector
#N6704S 10 µg

pTXB1 Vector
#N6707S 10 µg

pTXB3 Vector
#N6708S 10 µg

pKYB1 Vector
#N6706S 10 µg

T7 Universal Primer (20-mer)
#S1248S 0.5 A_{260} units

T7 Terminator Reverse Primer (19-mer)
#S1271S 0.5 A_{260} units

For additional components and information, see the NEB website: www.neb.com

New England Biolabs, Inc.
240 County Road
Ipswich, MA 01938
USA
Telephone (978) 927-5054
Toll Free (USA Orders)1-800-632-5227
Toll Free (USA Tech)1-800-632-7799
Fax (978) 921-1350
e-mail: info@neb.com
www.neb.com

New England Biolabs, Ltd.
1815 Ironstone Manor, Unit 6
Pickering, Ontario, L1W 3W9
Canada
Telephone (905) 837-2234
Toll Free 1-800-387-1095
Fax (905) 837-2994
Fax Toll Free 1-800-563-3789
e-mail: info@ca.neb.com

New England Biolabs GmbH
Brüningstrasse 50, Geb.G 810
65926 Frankfurt am Main
Germany
Telephone +49/(0)69/305 23140
Free Call 0800/246 5227 (Germany)
Fax +49/(0)69/305 23149
Free Fax 0800/246 5229 (Germany)
e-mail: info@de.neb.com

New England Biolabs (Beijing), Ltd.
Wangzhuang Lu No. 1
Tsinghua Tongfang High-Tech Plaza
Building B, 6th Floor B
Haidian District, Beijing 100083
China, People's Republic
Telephone 010-82378266
Fax 010-82378262
e-mail: beijing@neb-china.com

New England Biolabs (UK) Ltd.
75/77 Knowl Piece, Wilbury Way
Hitchin, Hertfordshire SG4 0TY
England, UK
Telephone (01462) 420616
Call Free 0800 318486
Fax (01462) 421057
Fax Free 0800 435682
e-mail: info@uk.neb.com

NEW ENGLAND
BioLabs® Inc.
the leader in enzyme technology