Table of Contents

20. Description of the pTXB expression vectors:

21. Description of the pKYB1 expression vector:

22. Description of the pCYB expression vectors:

23. Do the pTYB, pTXB, pTWIN or pCYB vectors carry any marker for the selection of the recombinant plasmids?

24. Which *E.coli* host strains can be used with the IMPACT vectors?

25. How can the target gene be transferred from the pCYB vectors into the pTYB vectors?

26. Can single-stranded DNA be prepared from the pTYB, pTXB, pTWIN and pCYB vectors?

27. In the case of the C-terminal fusion vectors pTYB1,2,3 and 4, which residues at the C-terminal of the target protein may inhibit cleavage or cause *in vivo* cleavage?

28. In the case of the N-terminal fusion vectors, pTYB21, pTYB22, pTYB11 and 12, which residues at the N-terminus of the target protein may inhibit cleavage or cause *in vivo* cleavage?

29. In the case of the C-terminal fusion vectors pTXB1 and 3, which residues at the C-terminal of the target protein may inhibit cleavage or cause in vivo cleavage?

CLONING

30. Should I clone my target gene into the C-terminal fusion vectors or N terminal fusion vectors?

31. What can I do if my gene of interest has a SapI and I want to make an exact fusion in pTXB1 or pTYB1?

32. What are the features of other *E. coli* IMPACT vectors?

33. How to design PCR Primers for cloning into C-terminal fusion vectors pTXB, pTYB (pTYB 1,2,3 and 4), or pCYB vectors?

34. Cloning into pTWIN Vectors

35. Primer Design for pTWIN vectors

36. Which residues or sequence should be inserted at the cleavage site to improve controllable cleavage in C-terminal pTYB vectors (pTYB1,2,3 and

4)?

37. How should primers with a SapI site be designed for cloning the gene of interest into pTXB1/pTYB1 or pTXB3/pTYB3 vectors?

38. How should a cloning strategy for C-terminal fusions (pTXB1,3 and pTYB1,2,3,4) be designed if the gene of interest contains both NdeI and NcoI sites?

39. Can XmaI be used for cloning the gene of interest into pTYB2 or pTYB4 vector?

40. How can the recombinant plasmid (pTYB, pTXB or pCYB vector with an insert) be analyzed by restriction digestion when the SapI or SmaI site is used for cloning?

41. In the case of C-terminal fusion vectors, if the EcoRI site is used as the 3´ cloning site, which vector should be chosen?

Expression

42. Recommended strains for the IMPACT vectors?

43. Which *E. coli* host strains can be used with the pCYB vectors?

44. Should DTT be added to the SDS Sample Buffer for SDS PAGE analysis?

45. How many positive clones should be checked for expression and purification?

46. What factors contribute to the low expression of some proteins?

47. How can I optimize expression?

48. What causes lower yields for eucaryotic proteins using the IMPACT system in *E.coli*?

49. What are some suggestions for expressing a toxic gene?

50. Should a pTXB or a pTYB vector be used as a control for expression?

51. Can the intein-CBD enhance solubility?

52. What if my fusion protein is insoluble?

53. IMPACT-TWIN Renaturation Protocol

54. Is there an alternative way to break cells other than sonication?

55. Are there antibodies available for western blot detection?

56. What are the common problems associated with western blot analysis?

57. What does this mean if Western blot analysis with anti-CBD serum detects a 28 kDa product (when pTXB1,3 are used) or a 55 kDa product (when the C terminal TYB fusion vectors are used) and not the fusion precursor?

58. Is *in vivo* cleavage an all or nothing phenomenon?

59. General Troubleshooting

PURIFICATION AND ON-COLUMN CLEAVAGE

60. What is the binding capacity of the resin?

61. How can the binding efficiency be improved?

62. The protein to be purified requires DTT or 2-mercaptoethanol in the Column Buffer. What effect will this have on cleavage?

63. Which reducing reagents can be used to stabilize a thiol-dependent enzyme?

64. Will protease inhibitors interfere with the intein-mediated cleavage reaction?

65. Is the 0.1% Triton X-100 in the Column Buffer necessary or could it be replaced by a comparable concentration of another detergent?

66. If the fusion protein forms inclusion bodies, can I try to purify protein under denaturing conditions?

67. If the target protein is sensitive to DTT, are there alternative means to induce on-column cleavage?

68. How can an on-column cleavage reaction be carried out if the target protein has optimal activity at low salt?

69. How can *E.coli* GroEL be eliminated from the elution sample?

70. The target protein is not eluted after DTT treatment, but is present on the chitin column.

71. What are the solutions to reduce co-purification of CBD fusion proteins?

72. If the on-column cleavage is inefficient and the fusion protein is the major product on the chitin beads, what should I do?

73. How can cleavage of the fusion protein be slowed down during the purification process?

74. How do I remove DTT after cleavage?

75. When I use the N-terminal fusion vectors (pTYB21, pTYB22, pTYB11 and pTYB12), how do I remove the small peptide co-eluted with the target protein?

76. Are there known ways to accelerate the hydrolysis of DTT from the C terminus of the target protein in a controlled manner?

77. When using the N- and C-terminal fusion construct (pTWIN system) 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?

78. Can the chitin beads be regenerated?

APPLICATIONS/LIGATION(IPL) AND TWIN

79. What is Intein-mediated Protein Ligation (IPL)?

80. What has IPL been used for?

81. Can IPL be used to generate substrates for enzyme analysis?

82. What vectors are suitable for IPL and cyclization?

83. What cleavage reagent should be used for IPL?

84. Which products should be ordered for IPL?

85. What are the column buffers recommended for chitin chromatography and on-column cleavage?

86. What is the protocol for ligation?

87. Which residues at the C-terminus of the target protein may inhibit cleavage or cause in vivo cleavage when the pTXB vectors are used?

88. What are the sequencing primers for the pTXB vectors?

89. References for Intein-mediated Protein Ligation(IPL):

90. Can the IMPACT System be used in other expression systems?

91. How can the C-terminal fusion vectors be used to label the C-terminus of the target protein?

92. Can the IMPACT System be used to pull out proteins from *E. coli* or mammalian cell extracts which react with the target protein?

93. Can the IMPACT System be used to purify peptides?

94. What is your recommended protocol for scaling-up the chitin column?

CHITIN-BINDING DOMAIN (CBD)

95. What is chitin?

96. Where does the chitin binding domain (CBD) in the IMPACT system originate from?

97. Does the *Bacillus circulans* CBD have an affinity for low molecular weight chitin (2 to 6 GlcNAc residues)?

98. Is there any method to elute the complete fusion protein from the chitin column? For instance, by eluting with high concentrations of GlcNAc or low pH?

99. Can the chitin beads be regenerated?

100. Can you provide additional information about the chitin beads?

101. What is the Km of the chitin binding domain?

102. What is the specificity of the CBD to chitin, cellulose, and polyglucans?

FAQs for IMPACT Protein Purification System

GENERAL QUESTIONS

1. What is IMPACT?

IMPACT, Intein-Mediated Purification with an Affinity Chitin-binding Tag, is a novel protein purification system that allows recombinant proteins to be purified without affinity tag in a single chromatographic step. This method was developed at New England Biolabs (NEB) from studies of the mechanism of protein splicing. It distinguishes itself from all other purification systems by its ability to purify a recombinant protein with its native sequence by a single affinity column, without the use of a protease. The IMPACT system utilizes the inducible self-cleavage activity of an engineered protein splicing element (intein) to separate the target protein from the affinity tag. The target protein is fused to a tag consisting of the intein and the chitin binding domain (CBD) which allows affinity purification of the fusion precursor on the chitin column. The intein undergoes specific cleavage by a thiol reagent or pH and temperature shift which releases the target protein from the chitin-bound tag resulting in a single column purification of the target protein.

The IMPACT system includes a series of *E.coli* expression vectors, which utilize engineered inteins of 134-454 amino acid residues. These vectors are designed for protein expression and purification in *E.coli* as well as protein manipulations such as protein labeling, ligation and cyclization.

The IMPACT Kit, as well as vectors sold separately, is available to meet your research goal. The compatibility of the multiple cloning sites of the vectors allows insertion of the same target gene fragment into different vectors for optimal expression and purification.

The gene encoding the target protein is inserted into the multiple cloning site of the IMPACT expression vector, to create an in-frame fusion between the target gene and the affinity tag consisting of the intein and chitin binding domain (CBD, 52 amino acid residues) When crude extracts of induced *E.coli* cells are passed over a chitin column, the fusion protein of the target protein intein-CBD binds to the chitin beads while all other contaminants are washed off the column. On-column cleavage is induced at 4°C by addition of a reducing agent (such as DTT) or temperature and pH shift (pTWIN vectors). The target protein is released while the intein-CBD tag remains bound to the column, resulting in a single-column purification of the target protein.

The IMPACT Kit (#E6901) contains expression vectors, which allow the fusion of the cleavable intein tag to either the C-terminus (pTXB1) or N-terminus (pTYB21) of the target protein. This flexibility in fusion protein construction maximizes the probability of successful expression and purification of a target protein.

The pTWIN vectors (#N6951 and #N6952) offer many advantages: (1) the facile isolation of native proteins without the use of a thiol reagent - using pH and

temperature shift (Intein 1) (2) the isolation of proteins with an N-terminal cysteine [for intein mediated ligation (IPL)] or residue other than methionine without the use of exogenous proteases which can be costly and non-specific (3) the purification of proteins with a C-terminal thioester for use in IPL reactions which can insert non-coded amino acids into a protein or label a bacterially expressed protein (4) the generation of circular protein species.

All vectors use a T7 promoter and the lacI gene to provide stringent control of the fusion gene expression. Binding of the lac repressor to the lac operator sequence immediately downstream of the T7 promoter suppresses basal expression of the fusion gene in the absence of IPTG induction. All vectors carry the Amp^r gene marker (the bla gene), which conveys ampicillin resistance to the host strain, except for pKYB1 (#N6706), which carries the kanamycin resistance gene.

2. What are the advantages of the IMPACT System?

- Single column purification - no additional steps to remove the affinity tag
- Yields native amino acid sequence
- Fusion of a target protein to either the C-terminus or N-terminus of the intein tag
- Proteases are NOT required to remove the affinity tag from the target protein
- Use of either thiol reagent or pH and temperature shift to induce on-column cleavage
- Isolation of proteins with or without an N-terminal methionine residue
- Production of proteins possessing an N-terminal cysteine and/or C-terminal thioester for use in protein labeling, ligation and cyclization
- T7 Promoter for high-level expression and tight regulation of transcription
- Protein semisynthesis - The ability to create a thioester at the C-terminus for ligation with a N-terminal containing cysteine peptide or tag
- The ability to label the C-terminus of the target protein
- IMPACT vectors pTYB21(#N6709) and pTYB22(#N6710) contain a MCS that is compatible with other NEB expression vectors.

3. What is protein splicing?

Protein splicing is defined as a post-translational processing event in which an internal protein segment, the intein, is excised from a protein precursor and the two external polypeptide sequences, the exteins, are then ligated together, resulting in the production of two distinct proteins. Many inteins, have been shown to mediate splicing without other protein cofactors. These self-splicing inteins have been engineered to undergo peptide bond cleavage at their N- and/or C- terminal junction. For more information on inteins please go to Inbase, the Intein Database and Registry.

4. Which kit should I use? / Table: IMPACT Vectors and Applications

The Summary of IMPACT vectors table describes some important considerations when choosing a vector: for protein purification:

1.) Site of target protein fusion -
   - May determine if target protein is active as a fusion
   - May increase or decrease level of expression
   - May affect cleavage reactions

2.) Cleavage site residue preference: actual cleavage efficiency is dependent on the

adjacent residues as well as the folding of the fusion protein.
3.) Thiol requirement - If your target protein is sensitive to thiol reagents, use of the TWIN vectors is warranted.

Summary of IMPACT vectors

Vectors	Site of target protein fusion	Intein Tag (kDa)	Recommeded cloning sites	Preferred residues at cleavage sitea	Method of cleavageb,c	Applications
pTXB1 pTXB3	C-terminus	Mxe GyrA intein (28)	NdeI-SapI (or SpeI)	M, Y, F, LEM (Unfavorable residues- S, P, E, D)	DTT (or MESNA) pH 8.0-8.5, 4°C	Purification; C-terminal thioester for ligation and modification
pTYB21 pTYB22 pTYB11 pTYB12	N-terminus	Sce VMA1 intein (56)	SapI-PstI BsmI (or NdeI)-PstI SapI-PstI BsmI (or NdeI)-NotI	A, Q, M, G, L, N, W, F, Y (Unfavorable residues- P, S, C, T, R)	DTT c pH 8-8.5, 25° C	Purification
pTYB1 pTYB2 pTYB3 pTYB4	C-terminus	Sce VMA1 intein (56)	NdeI- SapI NdeI-SmaI (or XhoI) NcoI-SapI NcoI-SmaI (or XhoI)	G, LEG (Unfavorable residues- P, C, N, D, R)	DTT (or MESNA) pH 8-8.5 4°C	Purification and ligation
pTWIN1	C-terminus (Intein 2)	Mth RIR1 Intein (28)	NdeI-SapI (or SpeI)	M, Y, F, LEM (Unfavorable residues- S, P, E, D)		Purification; C-terminal thioester for ligation and modification
pTWIN2	C-terminus (Intein 2)	Mth RIR1 intein (22)	NdeI-SapI (or SpeI)	G, A, LEG (Unfavorable residues- P, E, D)		
pTWIN1d	N-terminus (Intein 1)	Ssp DnaB mini-intein (27)	SapI-SapI SapI-PstI (or BamHI) BsrGI-PstI (or BamHI)	C, S, A, G, M, T, CRAM (Unfavorable residues- P)	pH 6-7 25° C	Purification; Defined N-terminus (e.g. Cys); Ligation
pTWIN1d	N-terminus (Intein 1) & C-terminus (Intein 2)	Ssp DnaB mini-intein (27) Mxe GyrA intein (28)	SapI-SapI	C, S, A, G, M, T, CRAM (Unfavorable residues- P) M, Y, F, LEM (Unfavorable residues- S, P, E, D)	Step 1: pH 6-7 25° C Step 2: DTT (or MESNA) pH 8-8.5 4° C	Purification; Defined N-terminus (e.g. Cys); Ligation and cyclizatione

pTWIN2d	N-terminus (Intein 1) & C-terminus (Intein 2)	Ssp DnaB mini-intein (27) Mth RIR1 Intein (22)	SapI-SapI	C, S, A, G, M, T, CRAM (Unfavorable residues- P) G, A, LEG (Unfavorable residues- P, E D)		

a. Actual cleavage efficiency is dependent on the adjacent residues as well as the folding of the fusion protein.
b. Dithiothreitol (DTT) is used only for protein purification. 2-mercaptoethanesulfonic acid (MESNA) is used for isolation of proteins possessing a C-terminal thioester for ligation, labeling and cyclization.
c. Cysteine can be used in the place of DTT.
d. pTWIN vectors allow for induction of Ssp DnaB (intein 1) intein cleavage by a pH/temperature shift.
e. When creating circular proteins it is typical that the linear form will co-purified, requiring a further step to separate the two protein species.

5. What vectors are included in the IMPACT kit?

The IMPACT kit includes three *E. coli* expression vectors, which allows for the fusion of the cleavable intein tag to either the C-terminus (pTXB1) or N-terminus (pTYB21) of the target protein.
The pMXB10 vector serves as a positive control for expression and purification and can also be used as a cloning vector.
pTXB1 (#N6707) contains a mini-intein from the *Mycobacterium xenopi* GyrA gene (Mxe GyrA intein; 198 amino acid residues) that has been modified to undergo thiol-induced cleavage at its N-terminus (Evans, T.C., et al. (1998) Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci*. 7, 2256–2264; Southworth, M. W., et al.(1999) BioTechniques 27, 110–120). The vector allows for the purification of a target protein without any extra amino acids by cloning into the NdeI and SapI sites. The target protein is fused at its C-terminus to a self-cleavable intein tag (~28 kD) that contains the chitin binding domain (CBD, 6 kDa) allowing for affinity purification of the fusion precursor on a chitin column.
The pTYB21 (#N6709) vector utilizes an intein from the *Saccharomyces cerevisiae* VMA1 gene (Sce VMA1 intein; 454 amino acids)(Chong, S. et al. (1998) Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. *Nucl. Acids Res*. 26, 5109–5115; Chong, S., et al. (1998) Modulation of protein splicing of the *Saccharomyces cerevisiae* vacuolar membrane ATPase intein. *J. Biol. Chem*. 273, 10567–77). The target protein is fused at its N-terminus to a self-cleavable VMA1 intein-CBD tag (56 kD); the tag allows for the affinity purification of the fusion precursor on a chitin column. The vector is designed to allow for purification of a target protein without any extra amino acids, or without an N-terminal methionine residue, by cloning its 5’ end into the SapI site. If the SapI site is not used extra amino acids will be added to the N-terminus of the

target protein. pTYB21(#N6709) and pTYB22(#N6710) contain a MCS compatible with other NEB expression vectors.
The control vector, pMXB10 (#N6903), derived from pTXB1 carries the control target protein, maltose-binding protein (MBP), already inserted upstream of the Mxe GyrA intein-CBD. Induction yields the MBP-GyrA intein-CBD fusion which, when cleaved, results in the elution of MBP. The polylinker regions flanking the coding region for MBP can conveniently be used to clone a gene of interest. However, after intein cleavage the target protein will contain additional amino acids at its C-terminus, including (LEY), which has had a high rate of successful cleavage.
The IMPACT vectors utilize a T7 promoter to provide stringent control of expression of the fusion gene in *E. coli* (Dubendorff, J. W. and Studier, F. W. (1991) Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *J. Mol. Biol*. 219, 45–59.). The IMPACT vectors carry the Amp^r gene marker (the *bla* gene), which conveys ampicillin resistance to the host strain; one vector, pKYB1 (#N6706), is kanamycin resistant (available separately).

6. What are the features of other *E. coli* IMPACT vectors?

The C-terminal fusion vectors, pTYB1(#N6701) and pTYB2 (#N6702), utilize an intein from the *Saccharomyces cerevisiae* VMA1 gene [Sce VMA intein; Kane, et al., (1990) *Science*,250, 651-657]. The intein has been modified to undergo a self-cleavage reaction at its N-terminal junction at 4°C when induced by thiol reagents such as DTT. pTYB1 and pTYB2 use ATG of the NdeI site in the multiple cloning region for translation initiation. The gene encoding the target protein is inserted into the NdeI/SapI site for pTYB1 and the NdeI/SmaI site for pTYB2 to create a fusion between the C-terminus of the target gene and the N-terminus of the gene encoding the intein. DNA encoding a chitin binding domain (CBD) from *Bacillus circulans* [Watanabe, T., et al. (1994) The role of the C-terminal domain and type III domains of chitinase A1 from *Bacillus circulans* WL-12 in chitin degradation *J. Bacteriol*. 176, 4465–4472] has been added to the C-terminus of the intein to allow affinity purification of the 3-part fusion.
pTYB11(#N6701) and pTYB12 (#N6902) are N-terminal fusion vectors in which the N-terminus of the target protein is fused to the C-terminus of the intein tag consisting of the intein and CBD. In the pTYB11 vector the SapI site must be used to clone the 5' end of the gene.This vector allows the use of BsmI or NdeI for cloning the 5′ end of the target gene (instead of the SapI site in pTYB11). The CBD is inserted in a loop region of the engineered Sce VMA intein such that the cleavage activity of the intein is not affected. The engineered intein can undergo cleavage at its C-terminus triggered by thiol-induced cleavage at its N-terminal junction.
pTXB1, pTYB1, pTYB21 and pTYB11 contain a SapI cloning site, which allows the target gene to be cloned adjacent to the cleavage site of the intein tag; this results in the purification of a target protein without any vector-derived, non-native residues attached to its terminus. For pTXB1 and pTYB1 only the SapI site should be used to clone the 3′ end of the target gene and for pTYB11 only the SapI site should be used to clone the 5′ end of the target gene . This strategy will result in the fusion of the target gene adjacent to the intein tag (and the cleavage site). The target protein can be purified without any extra non-native residues. The use of SapI site also allows for the addition of amino acid residues favorable for cleavage (by engineering them into the coding region of the primers).
Furthermore, pTYB1 and pTXB1 contain the same or compatible restriction sites in the

multiple cloning region. This flexibility in fusion protein construction maximizes the probability of successful expression and purification of a target protein.

Use of pTYB2, pTYB22 or pTYB12 yields a target protein with extra residue(s) added to its C-terminus or N-terminus, respectively, after the cleavage of the intein tag. For instance, cloning the 3´ end of a target gene using the SmaI site in pTYB2 adds an extra glycine residue to the C-terminus of the target protein. Likewise, cloning the 5´ end of a target gene using the NdeI site in pTYB12 and pTYB22 adds four extra residues (Ala-Gly-His-Met) to the N-terminus of the target protein. These extra residues have been shown to successfully cleave. pTYB21 and pTYB22 contain a MCS compatible with other NEB expression vectors.

The IMPACT vectors use a T7/lac promoter and the lacI gene to provide stringent control of the fusion gene expression. Binding of the lac repressor to the lac operator sequence immediately downstream of the T7 promoter suppresses basal expression of the fusion gene in the absence of IPTG induction. The vectors also contain the origin of DNA replication from bacteriophage M13, which allows for the production of single-stranded DNA by helper phage (M13KO7 helper phage, #N0315) superinfection of cells bearing the plasmid. Except in the cases of pTYB11 and pTYB12, the background transcription is further reduced by the placement of five tandem transcription terminators (*rrnB* T1) upstream of the T7 promoter sequence. All IMPACT vectors carry the Amp^r gene marker (the *bla* gene), which conveys ampicillin resistance to the host strain except for pKYB1 (#N6706), which carries the kanamycin resistance gene.

There are three TWIN *E. coli* expression vectors - pTWIN1(#N6951), pTWIN2 (#N6952) and pTWIN-MBP1(#N6953). The pTWIN vectors allow for the fusion of a mini-intein tag to the C-terminus, N-terminus, or both the C- and N-termini of a target protein, depending on the cloning sites chosen. Fusion of the intein tag to the C-terminus of the target protein allows thiol-dependent purification of the target protein from the chitin resin. This thiol-dependent cleavage permits the purification of proteins with a C-terminal thioester for use in intein-mediated protein ligation (IPL) reactions. IPL reactions can be used to incorporate fluorescent or biotinylated tags and/or non-coded amino acids into the C-terminus of a bacterially-expressed protein.

Alternatively, fusion of the intein tag (Intein1 or Ssp DnaB intein) to the N-terminus of the target protein permits the release of the target protein from the chitin resin by a pH (from 8.5 to 6) and temperature (from 4 °C to 20 - 25 °C) shift. This is advantageous for the purification of proteins that are sensitive to reducing agents such as DTT or 2-mercaptoethanol. Also, the C-terminal fusions can be used to purify proteins with an N-terminal amino acid other than Met. Finally, insertion of the target gene between two self-cleaving intein tags allows the generation of cyclic protein species which posses a peptide bond at the site of ligation.

Both pTWIN1 and pTWIN2 contain SapI 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. The pTWIN1 and pTWIN2 vectors both use a modified Ssp DnaB intein (154 amino acid residues) as intein1 and differ only in the identity of intein 2. pTWIN1 uses a modified Mxe GyrA intein (198 amino acid residues, same as in pTXB1) while pTWIN2 uses a modified Mth RIR1 intein (134 amino acid residues). However, both pTWIN1 and pTWIN 2 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 NcoI to SacI 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 42 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 [Evans, T.C. Jr., Benner, J., and Xu, M.-Q. (1999) The cyclization and polymerization of bacterially-expressed proteins using modified self-splicing inteins. *J. Biol. Chem*.274, 18359-18363]. Cloning into the NcoI to XhoI 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.
Two more pTYB C-terminal fusion vectors (pTYB3 NEB #N6703 and pTYB4 NEB #N6704) are available for cloning a target gene in which the C-terminus of a target protein is fused in-frame to the N-terminus of the Sce VMA intein-CBD tag. pTYB3 and pTYB4 contain an NcoI site overlapping the initiating methionine codon in place of the NdeI site in pTYB1 and pTYB2, respectively. Digestion of the insert with BspHI, BspLU11I, and AflIII can also generate NcoI- compatible overhangs.
TXB3 (NEB #N6708) is a IMPACT C-terminal fusion vector that is identical to pTXB1, except it utilizes the NcoI site overlapping the initiating methionine codon in place of NdeI.

7. What are the components of the IMPACT Kit?

All components of the IMPACT System are sold separately.
Components of the IMPACT Kit are:

- pTXB1 (#N6707)
- pTYB21 (#N6909)
- pMXB10 (#N6903) - Control plasmid
- *E. coli* Strain ER2566 – not competent
- Chitin Beads (#S6651)
- 1,4-Dithiothreitol (DTT), 1M
- Anti-Chitin Binding Domain Serum (anti-CBD) (#S6654)
- 3X SDS-PAGE Sample Buffer
- Blue Loading Buffer
- Instruction Manual

A competent version of ER2566, T7 Express Competent *E. coli* (NEB #C2566), can be purchased separately.
A Special Offer is provided for the IMPACT Kit with T7 Express Cells (#E0543).

8. Which inteins are used in the IMPACT vectors?

The pTXB vectors use an 198 amino acid residue intein from the gyrA gene of Mycobacterium xenopi. The intein has been engineered for thiol-inducible cleavage at its N-terminus. [Evans, T.C., et al. (1998) Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci*.7, 2256-2264; Telenti, A., et al. (1997) The *Mycobacterium xenopi* GyrA protein splicing element: Characterization of a minimal intein. *J. Bacteriol*.179, 6378-6382; Southworth, M.W., et al. (1999) Purification of proteins fused to either the amino or carboxy terminus of the *Mycobacterium xenopi* Gyrase A intein. *BioTechniques* 27, 110-120.]

The pTYB vectors utilize the intein (454 amino acid residues) from the VMA1 gene of

Saccharomyces cerevisiae (Sce VMA intein). The intein used in the C-terminal fusion vectors (pTYB1-4) is a splicing deficient mutant (carrying Asn454Ala substitution) that only undergoes N-terminal junction cleavage in the presence of a thiol compound or hydroxylamine. The C-terminal cleavage and splicing activities have been abolished [Chong et al.,(1996) *J. Biol. Chem*.271,22159-22168; Chong et al.,(1997) Gene 192,271-281]. In the N-terminal fusion vectors (pTYB21, pTYB22, pTYB11 and pTYB12) the N-terminus of a target protein is fused to the C-terminus (Asn454) of the intein. The CBD (56 amino acids) is inserted in a loop region of the intein without affecting its cleavage activity. A sequence consisting of the first 10 residues of the maltose-binding protein is used as the N-extein sequence to provide a favorable translational start for the fusion protein. The intein contains a single substitution, which changes its penultimate histidine residue (His453) to a glutamine. These substitutions allow for the inducible cleavage at both termini of the intein - the thiol-induced cleavage at the N-terminus of the intein/CBD tag (510 amino acids) triggers the cleavage at the C-terminus tag. [Chong et al.,(1998) *J. Biol. Chem*.273,10567-10577; Chong et al.,(1998) *Nucl Acids Res*.26,5109-5115]

The pTWIN vectors utilize inteins from the Ssp dnaB gene (Ssp DnaB intein, 154 aa), the Mxe gyrA gene (Mxe GyrA intein, 198 aa), and the Mth rir1 gene (Mth RIR1 intein, 134 aa).

Intein 1 in the pTWIN1, pTWIN2 and pTWIN-MBP1 vectors is a mini-intein of 154 amino acids derived from the *Synechocystis sp dnaB* gene [Wu, H., et al. (1998) *Biochem Biophys Acta* 1387, 422-432] engineered to undergo pH (from pH 8.5 to 6) and temperature (from 4 °C to 20 - 25 °C) dependent cleavage at its C-terminus. 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. [See Applications: Protein ligation and labeling, below].

In the pTWIN vectors intein2 is either a mini-intein from the *Mycobacterium xenopi* GyrA gene (pTWIN1) [Telenti, A. et. al. (1997) *J. Bacteriol*.179. 6378-6382] or from the *Methanobacterium thermoautotrophicum* rir1 gene (pTWIN2) [Smith, D. R., et al. (1997) *J. Bacteriol*.179(22), 7135-7155]. These inteins have been modified to undergo thiol-induced cleavage at their N-terminus. The use of thiol reagents, such as 2-mercaptoethanesulfonic acid (MESNA), releases a reactive thioester at the C-terminus of the target protein for use in IPL. Following cleavage the target protein is eluted from the chitin resin while the intein-CBD tag remains bound to the chitin resin.

9. What causes the cleavage?

In the C-terminal fusion vectors (pTYB1,2,3,4 and pTXB1,2,) the engineered intein can undergo a N-S acyl shift to move the upstream polypeptide (the target protein) to the side chain of the intein N-terminal cysteine to create a thioester linkage between the cysteine and the preceding residue (the C-terminal residue of a target protein). The thioester bond is susceptible to cleavage by thiol compounds such as DTT. Thus the equilibrium of the N-S acyl shift can be drastically shifted in the presence of DTT, resulting in release of the target protein. [Chong et al.,(1998) *Gene* 192, 271-281; Evans et al., (1998*) Protein Sci*.7, 2256-2264]

In the N-terminal fusion vectors (pTYB21, pTYB22, pTYB11 and pTYB12) the

engineered Sce VMA intein contains a substitution, which changes its penultimate histidine residue (His453) to a glutamine. This substitution attenuates the succinimide formation by the adjacent C-terminal residue (Asn454), which, in conjunction with a substitution at the first C-extein residue, allows inducible cleavage at both termini of the intein. The thiol-induced cleavage at the N-terminus of the intein/CBD tag triggers the cleavage at the C-terminus. [Chong et al.,(1998) *J. Biol. Chem*.273,10567-10577; Chong et al.,(1998) *Nucl. Acids Res*.26,5109-5115]

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 (MESNA). 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.

For cleavage without the use of a thiol regent, the N-terminus of a target protein can be fused to the C-terminus (an Asn) of Intein1 of the pTWIN vectors (the Ssp DnaB intein). A CBD, 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 ring with the concomitant breakage of the peptide bond.

10. If my target protein is sensitive to DTT , which vector(s) should I use?

You may use the pTWIN1 or pTWIN2 vector. Fusion of an intein tag (Intein 1) to the N-terminus of a target protein allows one-column protein purification with a pH (from pH 8.5 to 6) and temperature (from 4 °C to 20 - 25 °C) shift. No thiol reagents are required for cleavage.

The N-terminus of a target protein can be fused to the C-terminus (an Asn) of Intein1 of the pTWIN vectors (the Ssp DnaB intein). A CBD, 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 ring with the concomitant breakage of the peptide bond. Intein 1 fusions are used to generate proteins with an N-terminal cysteine.

11. What is the success rate of procaryotic protein purification?

80% of the prokaryotic proteins tested with the IMPACT System were expressed and purified with yields ranging from 0.5-30 mg/liter culture. 50% of these proteins were purified with their native sequences (no vector derived residues).

12. What size range of proteins have been purified?

The size range of proteins we have purified from *E.coli* host cells is up to 124 kDa.

13. IMPACT References:

pTYB1-4, pCYB1-4

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). Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* 192, 271-281. PubMed ID: 9224900

pTYB11,12

Chong, S., Montello, G.E., Zhang, A., Cantor, E.J., Liao, W., Xu, M-Q., and Benner, J. (1998) Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. *Nucleic Acids Res*. 26, 5109-5115. PubMed ID: 9801307

Chong, S., Williams, K. S., Wotkowicz, C., and Xu, M. Q. (1998). Modulation of protein splicing of the Saccharomyces cerevisiae vacuolar membrane ATPase intein. *J. Biol. Chem.* 273:10567-77. PubMed ID: 9553117

pTXB1,3

Evans, T.C. Jr., Benner, J., and Xu, M.-Q. (1998) Semisynthesis of cytotoxic proteins using a modified protein splicing element.*Protein Sci*. 7, 2256-2264. PubMed ID: 9827992

Southworth, M.W., Amaya, K., Evans, T.C., Xu, M.-Q., and Perler, F.B. (1999) Purification of proteins fused to either the amino or carboxy terminus of the Mycobacterium xenopi Gyrase A intein. *BioTechniques* 27, 110-120. PubMed ID: 10407673

pTWIN1,2

Evans, T.C. Jr., Benner, J., and Xu, M.-Q.(1999) The cyclization and polymerization of bacterially-expressed proteins using modified self-splicing inteins. *J. Biol. Chem.* 274, 18359-18363. PubMed ID: 10373440

Evans, T.C. Jr., Benner, J., and Xu, M.-Q. (1999) The *in vitro* ligation of bacterially expressed protein using an intein from*Methanobacterium thermoautotrophicum*. *J. Biol. Chem.* 274, 3923-3926. PubMed ID: 9933578

Mathys, S., Evans, T.C. Jr., Chute, I.C., Wu, H., Chong, S., Benner, J., Liu, X.-Q., Xu, M.-Q. (1999) Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene*, 231:1-13. PubMed ID: 10231563

Reviews

Xu, M.Q., Paulus, H. and Chong, S. (2000) Fusions to self-splicing inteins for protein purification. *Methods Enzymol.* 326:376-418. PubMed ID: 11036654

Evans T. C. & Xu, M.-Q, (1999) Intein-mediated protein ligation: harnessing nature's escape artists. *Biopolymers* 51, 333-42. PubMed ID: 10685044

For additional references see InBase, the Intein registry Web site.

VECTORS

14. The IMPACT kit vectors:

The IMPACT Kit (NEB#E6901) contains two expression vectors, which allow for the fusion of the cleavable intein tag to either the C-terminus (pTXB1, C-terminal fusion) or N-terminus (pTYB21, N-terminal fusion) of a target protein. This flexibility in fusion

protein construction maximizes the probability of successful expression and purification of a target protein. The use of the SapI site in pTXB1 and pTYB21 vectors allow the cloning of a target gene immediately adjacent to the intein cleavage site resulting in the purification of a native target protein without any vector-derived extra residues after the cleavage. The SapI site must be used to clone the 3' end of the gene in to pTXB1.

pTXB1 (#N6707) contains a mini-intein from the *Mycobacterium xenopi* GyrA gene (Mxe GyrA intein; 198 amino acid residues) that has been modified to undergo thiol-induced cleavage at its N-terminus [Evans, T.C., et al. (1998) Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci* . 7, 2256–2264; Southworth, M. W., et al.(1999) *BioTechniques* 27, 110–120]. The vector allows for the purification of a target protein without any extra amino acids by cloning into the NdeI and SapI sites. The target protein is fused at its C-terminus to a self-cleavable intein tag (~28 kD) that contains the chitin binding domain (CBD, 6 kDa) allowing for affinity purification of the fusion precursor on a chitin column.

The pTYB21 (#N6709) vector utilizes an intein from the *Saccharomyces cerevisiae* VMA1 gene (Sce VMA1 intein; 454 amino acids) [Chong, S. et al. (1998) Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. *Nucl. Acids Res*. 26, 5109–5115; Chong, S., et al. (1998) Modulation of protein splicing of the Saccharomyces cerevisiae vacuolar membrane ATPase intein. *J. Biol. Chem*. 273, 10567–77]. The target protein is fused at its N-terminus to a self-cleavable VMA1 intein-CBD tag (56 kD); the tag allows for the affinity purification of the fusion precursor on a chitin column. The vector is designed to allow for purification of a target protein without any extra amino acids, or without an N-terminal methionine residue, by cloning its 5’ end into the SapI site. If the NdeI site is used extra amino acids (GRAM) will be added to the N-terminuspTYB21 and pTYB22 contain a MCS compatible with other NEB expression vectors.

The control vector, pMXB10 (#N6903) , derived from pTXB1 carries the control target protein, maltose-binding protein (MBP), already inserted upstream of the Mxe GyrA intein-CBD. Induction yields the MBP-GyrA intein-CBD fusion which, when cleaved, results in the elution of MBP. The polylinker regions flanking the coding region for MBP can conveniently be used to clone a gene of interest. However, after intein cleavage the target protein will contain additional amino acids at its C-terminus, including (LEY), which has had a high rate of successful cleavage.

The IMPACT vectors utilize a T7 promoter to provide stringent control of expression of the fusion gene in *E. coli* (Dubendorff, J. W. and Studier, F. W. (1991) Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *J. Mol. Biol*. 219, 45–59.). The IMPACT vectors carry the Amp^r gene marker (the *bla* gene), which conveys ampicillin resistance to the host strain; one vector, pKYB1 (#N6706), is kanamycin resistant (available separately).

15. What are the polylinkers of the IMPACT vectors?

See the polylinkers of the various IMPACT vectors on the DNA Sequences and Maps tool.

16. Description of the pTYB expression vectors:

pTYB vectors (pTYB1,2,3,4,21,22,11, and 12) contain the Sce VMA1 intein and are

designed for expressing and purifying a protein produced from a cloned gene in *E. coli*.

pTYB1 (and 3) and 2 (and 4) are medium copy number *E. coli* plasmids, which are 7477, and 7474 base pairs in length, respectively. These vectors allow the fusion of the cleavable intein tag to the C- terminus of a target protein. The pTYB1 (NEB #N6701S) and pTYB2 (NEB #N6702) contain a NdeI site which is used to clone the 5' end of the target gene. pTYB3 (NEB#N6703) and pTYB4 (NEB#N6704) vectors contain an NcoI site, overlapping the initiating methionine codon, in place of the NdeI site in pTYB1 and pTYB2. Digestion of the insert with BspHI, BspLU11I and AflIII can also generate NcoI-compatible overhangs.

pTYB21 and pTYB22 are medium copy *E.coli* plasmids which are 7412 and 7415, base pairs in length, respectively. These vectors allow the fusion of the cleavable intein tag to the N- terminus of a target protein. Using the SapI site in pTYB21 allows for purification of a targe protein without any extra amino acids. If another site is used extra amino acids will be added on to the protein.

pTYB11 and 12 are medium copy number *E. coli* plasmids, which are 7413, and 7416 base pairs in length, respectively. These vectors allow the fusion of the cleavable intein tag to the N-terminus of a target protein. The SapI site can be used to clone the 5' end of the target gene in pTYB11 while the BsmI or NdeI is used to clone the 5' end of the target gene in pTYB12.

pMYB5, a control plasmid for IMPACT expression and purification, may also be used for cloning a gene of interest as a C-terminal fusion. Replacing the *malE* gene sequence (1.1kb) between the NdeI and XhoI sites by an open reading frame leaves three extra amino acid residues (LeuGluGly) at the C-terminus of the target protein following cleavage.

These vectors contain translation initiation signals (Shine Dalgarno sequence) for the strongly expressed T7 gene 10 protein (G10) and also contain either an NdeI (pTYB1 and 2) or NcoI (pTYB3 and 4) overlapping the initiating methionine codon. The polylinkers allow in-frame insertion of the target gene in frame to the intein. Ideally, PCR primers with the appropriate restriction enzyme sites should be used to clone the gene. In the case of pTYB1 and pTYB11by cloning in to the SapI site a fusion protein is produced that after cleavage yields a protein with its native sequence (i.e. no additional amino acids). Alternatively, other sites are available to clone in the target gene but additional amino acids will be present at the N-terminus or C-terminus of the target protein after cleavage. For both pTYB1 and pTYB11 vectors only the SapI site should be used to clone the 3´ and 5´ end respectively, of the target gene. This strategy will result in the fusion of the target gene adjacent to the intein tag (and the cleavage site). The target protein can be purified without any extra non-native residues. The use of SapI site allows for the addition of amino acids favorable for cleavage (by engineering them into the primers).

All IMPACT plasmids contain the ampicillin resistance gene (except for pKYB1), the M13 origin, the ColE1 origin of replication and the lacI gene. The origin of DNA replication from bacteriophage M13, allows the production of single-stranded DNA by helper phage (M13KO7 helper phage, NEB #N0315) superinfection of cells bearing the plasmid. The expression of the fusion gene in a pTYB vector is under the control of the T7 promoter, which is inducible by IPTG. Production of the fusion protein occurs after expression of T7 RNA polymerase. ER2566 is provided in the kit as an *E. coli* host strain for expression of the fusion protein from a pTYB vector. ER2566 cells or T7

Express (NEB #C2566) carry a chromosomal copy of the T7 RNA polymerase gene inserted into the lac gene, and under the control of the lac promoter. In the absence of IPTG induction, expression of the T7 RNA polymerase gene is suppressed by the binding of the *lac* repressor to the *lac* promoter. In addition, the presence of the lac operator immediately downstream of the T7 promoter results in the lowest basal level expression from a pTYB vector due to the binding of lac repressor to the lac operator. Furthermore, the five tandem copies of transcription terminators (rrnB T1) placed upstream of the T7 promoter minimize background transcription (not present in pTYB11,12). The T7 transcription terminator downstream of the CBD prevents continued transcription.

17. Comparison between the C-terminal fusion vectors (pTXB1, pTYB1) and N terminal fusion vectors (pTYB21).

The IMPACT™ system allows the fusion of a self-cleavable intein tag to either the C-terminus or N-terminus of a target protein. Different target proteins, due to certain structural constraints, may prefer either fusion to allow proper folding of the precursor protein and a higher level of protein expression. Our studies show that some target proteins results in higher yields when expressed in N-terminal fusion vectors, whereas others give higher yields in the C-terminal fusion vectors. There are also cases in which both fusions produce similar levels of protein expression and final yield. See the table that lists some differences in usage between the C-terminal and N-terminal fusion vectors.

	pTYB1, pTXB1	**pTYB21, pTYB22, pTYB11, pTYB12**
Cloning	The C-terminus of the target protein is fused to the intein tag	The N-terminus of the target protein is fused to the intein tag
N-terminus of the target protein	Met	Not restricted to Met; Pro blocks cleavage; Cys, Ser and Thr result in protein splicing
C-terminus of the target protein	See FAQ 27 and 29	Any amino acid; less *in vivo* cleavage
Expression	Influenced by the target protein	Less affected by the target protein
Purification		
Induction of on-column cleavage	4-23°C for 16 hr; efficient	At 4°C, less efficient; may require 16°C or higher for a longer time

Sensitivity to pH	pH 6.0-9.0, the cleavage favors basic pH	pH 6.0 inhibits cleavage; recommended pH 8.0 or above
Use of Cys for cleavage	Cys is covalently atttached to the target protein	Cys is not attached to the target protein
Elution	Normally elute after an overnight incubation	May require a longer incubation time before elution
Applications for labeling and peptide ligation	Yes	No

18. Description of the pTWIN expression vectors:

The pTWIN vectors allow a target protein to be sandwiched between two self-cleaving inteins. Chitin binding domains present on both inteins allow the affinity purification of the precursor protein on a chitin resin. Intein1 is a mini-intein derived from the Synechocytis spp dnaB gene engineered to undergo pH and temperature dependent cleavage at its C-terminus. 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 intein mediated protein ligation (IPL) reactions. Intein2 is either a mini-intein from the *Mycobacterium xenopi* GyrA gene (pTWIN1) or from the Methanobacterium thermoautotrophicum rir1 gene (pTWIN2). These inteins have been modified to undergo thiol-induced cleavage at their N-terminus. The use of thiol reagents such as 2-mercaptoethanesulfonic acid (MESNA) releases a reactive thioester at the C-terminus of the target protein for use in IPL. Following cleavage of the intein the target protein is eluted from the chitin resin while the intein remains bound through the chitin binding domain.

pTWIN-MBP1 can be used both as a control vector and a cloning vector. Cloning of a target gene into the NcoI to SacI 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 [Evans, T.C. Jr., Benner, J., and Xu, M.-Q. (1999) The cyclization and polymerization of bacterially-expressed proteins using modified self-splicing inteins. *J. Biol. Chem*.274, 18359-18363]. Cloning into the NcoI to XhoI 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 TWIN system offers many advantages:

1. Isolation of a native protein without the use of a thiol reagent (Intein1); cleavage is induced by a pH shift (from 8.5 to 6) and temperature shift (from 4 °C to 20 - 25 °C)
2. The facile isolation of native proteins without an affinity tag without the use of exogenous proteases which can be costly and non-specific

3. The isolation of proteins with an N-terminal cysteine (for IPL) or residue other than methionine
4. 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.
5. The generation of circular protein species. Furthermore, chitin, an extremely abundant organic substance, is stable, inexpensive and reusable.

19. Comparison between the pTWIN vectors

	pTWIN1a	**pTWIN1 or pTWIN2**	**pTWIN1**b
Fusion	C-Terminal	N-Terminal	N- and C- terminal
Cloning recommended.	NdeI to SapI	SapI to PstI the target	SapI to SapI
N-terminus of the target protein	Met	Cys-Arg, Gly-Arg, and Ser-Arg are recommended. Pro is not. Other amino acids depend on the influence of the target protein.	Cys-Arg, Gly-Arg, and Ser-Arg are recommended. Pro is not recommended. Other amino acids depend on the influence of the target protein.
C-terminus of the target protein	Met, Tyr and Phe are recommended. Asp, Pro and Gly are not recommended. Other amino acid residues depend on the influence of the target protein. (see 2.12)	Any amino acid	Met, Tyr and Phe are recommended. Asp, Pro and Gly are not recommended. Other amino acid residues depend on the influence of the target protein. (see FAQ 29)
Expression	Influenced by the target protein	Less affected by the target protein	Less affected by the 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 at room temperature	Shift from pH 8.5 buffer at 4°C to pH7.0 at room temperature followed by addition of thiol reagents 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 intein thiolinduced 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
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 NdeI to SapI 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.

b. Cloning into the pTWIN2 SapI sites results in the Ssp DnaB intein and the Mth RIR1 intein (intein 2 n pTWIN2) fused to the N-and C-terminus of the target protein, respectively. The Ssp DnaB intein in pTWIN2 is the same as in pTWIN1 (intein 1 n pTWIN1 and pTWIN2). 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.

20. Description of the pTXB expression vectors:

The pTXB1(NEB#N6707) and pTXB3(NEB#N6708) vectors are *E. coli* IMPACT C-terminal expression vectors sold separately. The pTXB vectors are designed for:
Purification of recombinant proteins
Isolation of proteins for protein ligation [Evans et al.,(1998) *Protein Sci*.7,2256-2264].
These plasmids are identical to the C-terminal fusion pTYB plasmids, except for the replacement of the yeast SceVMA intein with the 198 amino acid residue intein from the gyrA gene of Mycobacterium xenopi (Mxe GyrA intein). A target gene is inserted into the polylinker region of the pTXB vectors so that it is in-frame with the Mxe GyrA intein/chitin binding domain (CBD) coding region. Expression of the fusion gene is under control of the IPTG-inducible T7 promoter.
The Mxe GyrA intein (198 amino acid residues) has been engineered for thiol reagent-inducible cleavage between a protein of interest and the N-terminus of the intein, which results in a thioester at the C-terminus of the target protein [Evans et al.,(1998) *Protein Sci*.7,2256-2264]. The intein was engineered by substituting the intein C-terminal asparagine residue with an alanine to block splicing and C-terminal cleavage. These mini-intein vectors may result in higher expression levels. It may be advantageous to express a protein of interest using both pTXB and pTYB vectors (cloning the insert using the same restriction sites) and examine both constructs for

expression and purification.
The polylinker region of pTXB1 (NEB#N6707) is identical to pTYB1(NEB#6701) while the polylinker region of pTXB3 (#N6708) is identical to pTYB3(#N6703). pTYB and pTXB differ in the identities of the intein.
The double stranded plasmids, pTXB1 and pTXB3, are 6706 base pairs in length.
Please note that the SalI in the polylinker is not unique in pTXB1 and pTXB3. The SapI site must be used for cloning the 3´ end of the insert (see data card and FAQ's for details).
Based on mutagenesis studies using paramyosin as the target protein, Ser, Pro or Asp at the -1 position (the amino acid preceding the intein) blocks the cleavage When the C-terminus of a target protein contains an unfavorable residue, additional residue(s) such as Met or Tyr (the native residue preceding the Mxe GyrA intein), may be inserted between the target protein and the N-terminus of the intein to improve controlled cleavage. However, the cleavage efficiency may vary when a different target protein is fused to the intein.
For more details please see FAQ titled "In the case of the C-terminal fusion vectors pTXB1 and 3, which residues at the C-terminal of the target protein may inhibit cleavage or cause in vivo cleavage?"

21. Description of the pKYB1 expression vector:

The pKYB1 vector (NEB#N6706) is an *E.coli* C-terminal expression vector for use with the IMPACT System. pKYB1 carries the kanamycin resistance gene (Kn) from Tn 903. It was constructed by replacing the ampicillin gene in pTYB1 with the kanamycin resistance gene, therefore all other features are identical to pTYB1 (MCS, T7 promoter, etc.). This double stranded vector is 8196 bp in length. Both T7 Express (NEB #C2566) and BL21(DE3) (NEB #C2527) or a T7 derivative can be used as a host with this vector.

22. Description of the pCYB expression vectors:

The pCYB series of C-terminal fusion vectors were previously supplied with the IMPACT I system, and are still available upon special request. The four pCYB vectors (pCYB1, 2, 3 and 4) allow for expression and purification of a protein produced from a cloned gene in *E. coli*. Expression of the fusion gene in a pCYB vector is under the control of the tac promoter and is inducible by IPTG. pCYB1, 2, 3 and 4 are medium copy number *E. coli* plasmids which are 6846, 6843, 6846, 6843 base pairs in length, respectively. All of the pCYB vectors contain the gene for ampicillin resistance, the M13 origin, the ColE1 origin of replication and the *lacIq* gene. The vectors are identical in sequence except for the polylinker regions. The polylinkers allow in-frame insertion of the target gene upstream of the intein and contain either an NdeI (pCYB1 and 2) or NcoI (pCYB3 and 4) overlapping the initiating methionine codon. pCYB1,2,3 and 4 are similar to pTYB1,2,3 and 4 except the former contain the *tac* promoter and the latter the T7 promoter.

23. Do the pTYB, pTXB, pTWIN or pCYB vectors carry any marker for the selection of the recombinant plasmids?

No. However, to reduce the background from self-ligation of the vector, the ligation sample can be digested prior to transformation with an enzyme whose site is deleted from the polylinker during cloning and is also absent in the insert. This will linearize

any remaining parental vectors without the insert.

24. Which *E.coli* host strains can be used with the IMPACT vectors?

The fusion constructs should initially be established in a non-expression strain. Vector DNA supplied with the IMPACT system is prepared from a restriction-deficient *E. coli* strain (r-m-). When introduced into a strain with wild type EcoK (hsd+) DNA will be restricted. Thus plasmid DNA or ligated DNA should be introduced into a restriction-deficient *E. coli* strain (r-m- or r-m+).

The IMPACT kit includes the non-competent NEB T7 expression strain ER2566. A competent version of ER2566, T7 Express Competent *E. coli* (NEB #C2566), can be purchased separately from New England Biolabs, Inc. or is available as a Special Offer with the IMPACT Kit (NEB#E0543).

Several commercially available strains have also been tested, all of which allowed for expression and purification with the IMPACT system. Note: the culture should be inoculated with a freshly grown colony. For optimal expression of each fusion protein different strains should be tested.

For *E.coli* strains to be used, go to FAQ titled "Recommended strains for the IMPACT vectors?" under Expression.

25. How can the target gene be transferred from the pCYB vectors into the pTYB vectors?

If the gene of interest has previously been inserted into one of the pCYB vectors, the gene of interest, along with part or all of the intein-CBD coding region, can be transferred into the C-terminal fusion pTYB vectors (pTYB 1,2,3 and 4) by using the compatible unique restriction sites that exist between the two vectors (See Compatible pTYB Sites chart). Be sure that the sites used are not present in the gene of interest. Furthermore, pMYB5, a control plasmid may be used for cloning a gene of interest. Replacing the malE gene sequence (1.1Kb) between the NdeI and XhoI sites by an open reading frame leaves three extra amino acid residues (LeuGluGly) at the C-terminus of the target protein following cleavage.

pCYB1-4	**Compatible pTYB Sites**	
NdeI-SapI (pCYB1) NdeI-SmaI (pCYB2)	NdeI-Acc65I NdeI-KpnI NdeI-BsrGI NdeI-HindIII NdeI-BgIIINdeI-BstBI	NdeI-PmlI NdeI-SacII NdeI-NsiI NdeI-AgeI NdeI-MfeI NdeI-PstI

NcoI-SapI (pCYB3) NcoI-SmaI (pCYB4)	NcoI-Acc65I NcoI-KpnI NcoI-BsrGI NcoI-HindIII NcoI-BgIII NcoI-BstBI	NcoI-PmlI NcoI-SacII NcoI-NsiI NcoI-AgeI NcoI-MfeI NcoI-PstI
NdeI-XhoI (pCYB2)	NdeI-XhoI (pTYB2)*	
NcoI-XhoI (pCYB4)	NcoI-XhoI (pTYB4)*	

*Note: Other restriction sites listed above may also be used.

26. Can single-stranded DNA be prepared from the pTYB, pTXB, pTWIN and pCYB vectors?

Single stranded DNA can be prepared from these vectors. The antisense strand is rescued, so the Kunkel mutagenesis primers should be complementary and annealed to the antisense strand of the gene of interest. The CJ236 strain with M13K07 helper phage (NEB#N0315) has been tested. The IMPACT vectors contain a M13 origin, which should be compatible with other rescue systems such as CJ236 with VCSm13 helper. Other F' hosts can be used to rescue single stranded DNA.

27. In the case of the C-terminal fusion vectors pTYB1,2,3 and 4, which residues at the C-terminal of the target protein may inhibit cleavage or cause *in vivo* cleavage?

The data in the C-terminal fusion vectors table is based on analysis of cleavage reactions using the *E. coli* maltose-binding protein (MBP) as the target protein (MYB) with amino acid substitutions at the position (-1) immediately upstream of the cleavage site in the sequence (L-3E-2 X-1/C1). In summary, the data indicate that Pro, Cys and Asn inhibit thiol-induced cleavage, while Asp, Arg, His, Glu and Thr cause *in vivo* cleavage of the fusion proteins. Therefore if the target protein contains one of these "unfavorable residues" at its C-terminus, the inclusion of a favorable residue or sequence immediately adjacent to the cleavage site may be necessary for controllable cleavage reaction. In the case of the MBP fusion, when a glycine is added between the proline and the intein N-terminal cysteine, inducible cleavage is greatly improved. However, this is not to say that all proteins will behave in the same manner when fused to the intein-chitin binding domain. In each case, the effect of different C-terminal residues on the cleavage reaction has to be evaluated.

C-terminal Residue of the Target Protein	*In vivo* cleavage	*In vitro* Cleavage with DTT (40 mM)	
		4°C	16°C
Gly	-	+ + +	+++
Ala	-	+ + +	+++
Ile*	-	+	+
Leu*	-	+	+++
Met*	-	+++	+++
Phe*	-	+++	+++
Val*	-	+	++
Gln*	-	+++	+++
Ser	-	++	+++
Trp*	-	+++	+++
Tyr*	-	+++	+++
Lys*	-	+++	+++
Thr*	25%	++	+++
Glu*	50%	++	+++
His*	50%	++	++

Arg*	75%	not determined	not determined
Asp	100%	not determined	not determined
Asn	-	-	-
Cys	-	-	-
Pro	-	-	-

(–) = Less than 10% cleavage; (+) = 30%–49% cleavage;
(++) = 50%–74% cleavage; (+++) = 75%–100% cleavage.

** Boiling in DTT-containing SDS-PAGE sample buffer may cause partial or complete cleavage with these amino acids at the -1 position. If substantial in vivo cleavage is observed, the cell extract should be evaluated in a SDS Sample Buffer containing no DTT.*

28. In the case of the N-terminal fusion vectors, pTYB21, pTYB22, pTYB11 and 12, which residues at the N-terminus of the target protein may inhibit cleavage or cause *in vivo* cleavage?

The data in the N-terminal fusion vectors table is based on analysis of cleavage reactions using the MYT4 fusion system in which the Sce VMA intein was fused between *E. coli* maltose-binding protein (MBP; as the N-extein) and phage T4 DNA ligase (as the C-extein). To investigate the effect of the N-terminal residue of a target protein on cleavage in the C-terminal fusion vectors, amino acid substitutions at the first C-extein residue Cys455 were made. In summary, the data indicate that Pro, Cys and Ser are not favorable for thiol-induced cleavage. However, this is not to say that all proteins will behave in the same manner when fused to the intein-chitin binding domain. In each case, the effect of different N-terminal residues on the cleavage reaction has to be evaluated.

N-terminal residue of the target protein	% cleavage after 16 hr*			% cleavage after 40 hr*		
	4°C	16°C	23°C	4°C	16°C	23°C
Met Ala Gln	40-60	>80	>95	60-90	>90	>95
Gly Leu Asn Trp Phe Tyr	10-40	50-80	75-95	40-60	>90	>90
Val Ile Asp Glu Lys Arg His	<10	30-50	50-80	10-20	70-90	70-95
Pro	<10	<10	<10	<10	<10	<10
Thr	7	40	80	20	80	>90
Ser	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Cys	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

* Cleavage reactions were conducted in a model system (the MYT4 fusion system) [Chong, S., Montello, G.E., Zhang, A., Cantor, E.J., Liao, W., Xu, M-Q., and Benner, J. (1998) Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. *Nucleic Acids Res*. 26, 5109-5115] in which T4 DNA ligase is the target protein. The purified fusion protein was treated with 40 mM DTT in 30 mM Hepes, pH 8.0, 0.5 M NaCl

for 16 or 40 hr at 4°C, 16°C and 23°C. The percentage of cleavage (% cleavage) was determined by comparing the fusion precursors from the DTT-treated samples with those from the samples without the DTT-treatment in scanned images of Coomassie Blue stained SDS-PAGE gels. n.d., not determined.

29. In the case of the C-terminal fusion vectors pTXB1 and 3, which residues at the C-terminal of the target protein may inhibit cleavage or cause in vivo cleavage?

See Table 1A in the product manual for the effect of the C-terminal residue of a target protein on DTT-induced cleavage with pTXB1. The C-terminal amino acid of the target protein, paramyosin, was mutated immediately upstream of the intein cleavage site. Cleavage was induced with 40 mM DTT in 30 mM Tris, pH 8.5, 0.5M NaCl. Percent cleavage was determined by Coomassie stained SDS-PAGE analysis of chitin beads before and after DTT cleavage.

Note: Boiling in SDS Sample Buffer containing DTT can cause partial or complete cleavage, resulting in an overestimation of in vivo cleavage. If substantial in vivo cleavage is observed, the cell extract should be evaluated in a SDS Sample Buffer containing no DTT.

[1] Leu showed ~50% *in vivo* cleavage when induced at 15°C; at 37°C *in vivo* cleavage was less than 5%.

[2] Asp showed ~50% *in vivo* cleavage when expression was induced at 15°C and 37°C.

C-Terminal Residue of the Target Protein	% Cleavage After 16 Hours*		% Cleavage After 40 Hours*	
	4°C	23°C	4°C	23°C
Tyr Phe Gln Asn Thr Lys Ala His Leu1 Met	65-80	80-95	75-90	85-95
Ile Arg Glu Trp Cys	30-55	60-85	50-70	70-95
Val	30	70	60	90

Gly	10	40	20	60
Asp2	10	20	20	30
Ser Pro	5-15	5-15	5-15	5-20

Table 1A: Effect of the C-terminal residue of a target protein on DTT-induced cleavage with pTXB1. The C-terminal amino acid of the target protein, paramyosin, was mutated immediately upstream of the intein cleavage site. Cleavage was induced with 40 mM DTT in 30 mM Tris, pH 8.5, 0.5M NaCl. Percent cleavage was determined by Coomassie stained SDS-PAGE analysis of chitin beads before and after DTT cleavage.

Note: Boiling in SDS Sample Buffer containing DTT can cause partial or complete cleavage, resulting in an overestimation of in vivo cleavage. If substantial in vivo cleavage is observed, the cell extract should be evaluated in a SDS Sample Buffer containing no DTT.

1 Leu showed ~50% *in vivo* cleavage when induced at 15°C; at 37°C *in vivo* cleavage was less than 5%.

2 Asp showed ~50% in vivo cleavage when expression was induced at 15°C and 37°C.

CLONING

30. Should I clone my target gene into the C-terminal fusion vectors or N terminal fusion vectors?

It is conceivable that different target proteins, due to certain structural constraints, may prefer either C-terminal or N-terminal fusion to allow proper folding of the fusion precursor and a high level of protein expression. One may express the target gene in both N- and C-terminal fusion vectors and determine which type of fusion results in a better expression and yield. Use of pTXB1, pTYB1, pTYB21 (SapI site is used), pTYB11, pTWIN1 or pTWIN2 allows expression and purification of a target protein without any extra vector-derived residues. One should also take into consideration the differences between the C-terminal and N-terminal fusions with respect to amino acid at the fusion junction and if the purified target protein is also intended for C-terminal labeling or peptide ligation, pTXB1, pTWIN, pTYB1 or pTYB2 should be used. If the target protein requires a defined N-terminal residue (other than Met) or has an unfavorable residue (such as Pro) as the C-terminal residue, pTWIN, pTYB21, pTYB22, pTYB11, or pTYB12 should be used or extra amino acids should be added to the target protein.
For pTXB1 and pTYB1 only the SapI site should be used to clone the 3´ end of the target gene and for pTYB11 only the SapI site should be used to clone the 5´end of

the gene. This strategy will result in the fusion of the target gene adjacent to the intein tag (and the cleavage site). The target protein can be purified without any extra non-native residues. The use of the SapI site also allows for the addition of amino acids favorable for cleavage (by engineering their codons into the primers).

31. What can I do if my gene of interest has a SapI and I want to make an exact fusion in pTXB1 or pTYB1?

For cloning in to pTXB1 you can use the SpeI site in the Mxe intein. Your 3' primer will be designed as the reverse complement to: 15-18 bp target gene – TGC....MxeI intein ACTAGT NNNNNN where NNNNNN is any 6 nt to allow for efficient cleavage of your PCR product. The TGC is the first cysteine of the Mxe intein and ACTAGT is the SpeI site. You can cut the vector and insert with NdeI/SpeI and you will regenerate the intein sequence getting an exact fusion.

There could be a two-step procedure for cloning a gene containing a single internal SapI site into pTYB1 vector. First, the 5' fragment of the target gene (NdeI in the forward primer - internal SapI site sequence plus Not or XhoI in the reverse primer) is cloned into pTYB1. This results in a NdeI-SapI-Xho-SapI-intein-CBD fusion. Please remember that you can use SapI for directional cloning in most cases where the three-nt overhangs generated at the two SapI sites are not the same. The second PCR uses one forward primer containing the internal SapI site and a reverse primer containing a SapI site generating an overhang compatible with the SapI generated end of pTYB1.

Alternatively, one may amplify the full length gene (with the reverse primer containing a SapI site required for pTXB1, pTYB1 vectors). The PCR fragment is cloned into a T-vector or blunt end cloning vector. The recombinant plasmid can then be treated by SapI to obtain a partial digestion pattern (vary enzyme concentration and time of digest).

32. What are the features of other *E. coli* IMPACT vectors?

pTYB1 (NEB #N6701), pTYB2 (NEB#N6702), pTYB3 (NEB #N6703) and pTYB4 (NEB #N6704), are available for cloning a target gene in which the C-terminus of the target protein is fused to the intein-CBD tag. pTYB3 and pTYB4 contain an NcoI site, overlapping the initiating methionine codon, in place of the NdeI site in pTYB1 and pTYB2. Digestion of the insert with BspHI, BspLU11I and AflIII can also generate NcoI-compatible overhangs.

pTXB3 (NEB #N6708) and pTXB1 (NEB #N6707) are IMPACT C-terminal fusion vectors using an engineered mini-intein (198 residues) from the gyrA gene of Mycobacterium xenopi. These vectors are designed for fusing the C-terminus of a target protein to the Mxe intein-CBD tag for protein purification and generation of C-terminal thioester tagged proteins for labeling and protein ligation [See Evans, T.C., Benner, J., and Xu, M.-Q.(1998) Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci*.7,2256-2264]. [See Part 6: Applications: Protein ligation and labeling, below] Use of pTXB vectors may result in higher expression and reduce in vivo cleavage. The polylinker region of pTXB1 is identical to pTYB1 and the polylinker region of pTXB3 is identical to pTYB3.

pTYB21(NEB#N6709), pTYB22(NEB#N6710), pTYB11 (NEB#N6901) and pTYB12 (NEB#N6902) are available for cloning a target gene in which the N-terminus of the target protein is fused to the intein-CBD tag. If a N-terminus other than Met is

required these vectors can be used. For pTYB11 only the SapI site should be used to clone the 5′ end of the target gene. Cloning the 5′ end of a target gene using the NdeI site in pTYB21, pTYB22, or pTYB12 adds extra residues (Ala-Gly-His) to the N-terminus of the target protein.
The pTWIN vectors allow for cloning of the target gene to the N- and/or C-terminus of the target protein. The inteins in these vectors are mini-inteins, less than 30 kDa. When Intein 1 (Ssp DnaB intein) is used the cleavage is induced by pH shift from 8.5 to 6 and/or temperature from 4°C to room temperature. When Intein 2 (Mxe GyrA intein in pTWIN1 and Mth RIR1 intein) in pTWIN2 is used cleavage is induced with a thiol reagent. More details on the pTWIN vectors can be found under FAQ titled "Description of the pTWIN expression vectors".

33. How to design PCR Primers for cloning into C-terminal fusion vectors pTXB, pTYB (pTYB 1,2,3 and 4), or pCYB vectors?

The following examples demonstrate how to design primers used for PCR reactions to generate a target gene fragment with (1) an NdeI, NcoI, or BspHI site at its 5′ end, (2) an XhoI (pTYB2) or SapI (or BspQI) site at its 3′ end, or (3) blunt-end(s). The sequence corresponding to the sequence of the putative target gene is in lower case. Restriction enzyme recognition sequences are underlined. The addition of extra bases (NNN NNN) 5′ to the restriction site are required for efficient cleavage by the corresponding enzyme. Although the length of such a sequence may vary according to a specific enzyme, 6 bases are generally recommended. Please note the SapI isoschizomer BspQI (NEB #R0712) can be used instead of SapI.

Forward Primers:

- Forward primer #1 (NdeI) for insertion into the NdeI site:

5′ - NNN NNN CAT atg ggt aat ctg tct caa acc ca-3′

- Forward primer #2 (NcoI) for insertion into the NcoI site:

5′ - NN NNN NCC atg ggt aat ctg tct caa acc cac-3′

- Forward primer #3 (BspHI) for insertion in the NcoI site:

5′ - NN NNN NTC atg aat aat ctg tct caa acc cac-3′

- Forward primer #4 for insertion into the filled-in NcoI site or NruI site:

5′ - ggt aat ctg tct caa acc cac-3′

Reverse Primers:

- Reverse primer #5 (SapI) for insertion into the SapI site:

5′ -NN NNN NGC TCT TCCGCA ttc ctt cct cct taa tct ttc tt-3′

- Reverse primer #6 (SapI) for insertion into the SapI site (for producing protein with a Gly residue at the C-terminus):

5′ -NN NNN NGC TCT TCCGCA ACC ttc ctt cct cct taa tct ttc tt-3′

- Reverse primer #7 for insertion into the SmaI site (with a Gly tag) or filled-in SapI site:

5′ -ttc ctt cct cct taa tct ttc tt -3′

- Reverse primer #8 for insertion into the SmaI site (with a LeuGluGly tag):

5′ -CTC GAG ttc ctt cct cct taa tct ttc tt -3′

- Reverse primer #9 (XhoI) for insertion into the XhoI site of pTYB2 or pTYB4 (leaving a LeuGluProGly tag):

5′ - NNN NNN CTC GAG ttc ctt cct cct taa tct ttc tt -3′

34. Cloning into pTWIN Vectors

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. Both pTWIN1 and pTWIN2 contain SapI 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. The pTWIN1 and pTWIN2 vectors both use a modified Ssp DnaB intein as Intein1 and differ only in the identity of intein 2. 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.

TWIN-MBP1 can be used both as a control vector and a cloning vector. Cloning of a target gene into the NcoI to SacI 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 NcoI to XhoI 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.

35. Primer Design for pTWIN vectors

Normally, a target gene is amplified by PCR before it is inserted in-frame into the polylinker 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 SapI sites in the pTWIN1 or pTWIN2 vector. It should be noted that the SapI digested DNA does not undergo self-ligation because DNA restriction at both SapI sites results in non-complimentary overhangs. Please note the SapI isoschizomer BspQI (NEB #R0712) can be used instead of SapI.

To clone your insert you may do any of the following:

1. Use SapI sites (pTWIN1, pTWIN2
   (SapI site is not regenerated after cloning)
2. Use NcoI and XhoI (pTWIN-MBP1)
3. Use NcoI and SacI (pTWIN-MBP1)

Examples of primer design for the insertion of a target gene between two inteins:

Cloning vectors: pTWIN1 or pTWIN2
5′ cloning site primer:
SapI 5′ -GGT GGT T GC TCT TCC AACNNN…-3′
3′ cloning site primer:

SapI 5′ -GGT GGT T GC TCT TCC GCA NNN…-3′

Cloning vectors: pTWIN1, pTWIN2 or pTWIN-MBP1
5′ cloning site primer:
NcoI* 5′ -GG TGG TCC ATG GNN N…-3′

Cloning vector: pTWIN-MBP1
5′ cloning site primer:
NcoI* 5′ -GGT GGT CC ATG GNN N…-3′
3′ cloning site primer:
SacI* 5′ -GGT GGT TT G AGC TCN NN…-3′

Cloning vector: pTWIN-MBP1
5′ cloning site primer:
NcoI* 5′ -GGT GGT CC ATG GNN N…-3′
X hoI* 5′ -GGT GGT CTC GAG NNN…-3′

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 the restriction site is close to the 5′ end of the primer.

SapI site is lost after cloning.

A stop codon should NOT be included in the reverse primer when the C-terminus of the target protein is fused to the N-terminus of a intein tag.

* Do NOT use in conjunction with SapI cloning.

36. Which residues or sequence should be inserted at the cleavage site to improve controllable cleavage in C-terminal pTYB vectors (pTYB1,2,3 and 4)?

Gly, LeuGluGly, LeuGluProGly or GlyThrLeuGluGly have been found to improve controllable cleavage in the C-terminal pTYB vectors, although these may not be the optimal sequences for every protein. The desired sequence can be included in the reverse PCR primer, and either the SapI or SmaI sites may be used for cloning. For both pTYB1 and pTYB11 vectors only the SapI site should be used to clone the 3′ and 5′ end respectively, of the target gene. This strategy will result in the fusion of the target gene adjacent to the intein tag (and the cleavage site). The target protein can be purified without any extra non-native residues. The use of SapI site allows for the addition of amino acids favorable for cleavage(by engineering them into the primers).

37. How should primers with a SapI site be designed for cloning the gene of interest into pTXB1/pTYB1 or pTXB3/pTYB3 vectors?

NEB sells BspQI (NEB#R0712), an isoschizomer of SapI (NEB#R0569), which can be used instead of SapI.

Some researchers are not familiar with the SapI restriction enzyme [5′ -GCTCTTC(1/4) -3′ , NEB #R0569], which is used in the cloning of target genes into the IMPACT expression vectors. Below is information regarding primer design for PCR amplification of a target gene. The putative gene sequence is in lower case. The SapI-recognition sequence [5′ -GCTCTTC-3′] is incorporated in the reverse primer as illustrated in the following examples. The SapI-recognition sequence [5′ -GCTCTTC-3′] is inverted in the PCR product. SapI digestion creates a 5′ GCA overhang for ligation with an IMPACT vector with a 5′ TGC overhang generated by SapI digestion.

Example 1:

This reverse primer is designed for insertion at the SapI site. The resulting fusion construct is for generating protein with no vector-derived amino acid residue at the C-terminus.
5′ -CT CGA GGC TCT TCC GCA ttc ctt cct cct taa tct ttc tt-3′
The 3′ end of the PCR product (the codon for the intein N-terminal cysteine is indicated):
...aa gaa aga tta agg agg aag gaa **TGC** GGA AGA GCC TCG AG 3′
...tt ctt tct aat tcc tcc ttc ctt **ACG** CCT TCT CGG AGC TC 5′
The 3′ end of the PCR product after *Sap*I digestion:
...aa gaa aga tta agg agg aag gaa 3′
...tt ctt tct aat tcc tcc ttc ctt **ACG** 5′

Example 2:

reverse primer is designed for insertion at the SapI site. The resulting fusion should generate protein with a glycine residue at the C-terminus.
5′ -CT CGA GGC TCT TCC GCA ACC ttc ctt cct cct taa tct ttc tt-3′
The 3′ end of the PCR product (The added glycine codon and the codon for the intein N-terminal cysteine are indicated)
...aa gaa aga tta agg agg aag gaa **GGT TGC** GGA AGA GCC TCG AG 3′
...tt ctt tct aat tcc tcc ttc ctt **CCA ACG** CCT TCT CGG AGC TC 5′
The 3′ end of the PCR product after *Sap*I digestion:
...aa gaa aga tta agg agg aag gaa **GGT** 3′
...tt ctt tct aat tcc tcc ttc ctt **CCA** ACG 5′

38. How should a cloning strategy for C-terminal fusions (pTXB1,3 and pTYB1,2,3,4) be designed if the gene of interest contains both NdeI and NcoI sites?

Several other restriction enzymes are compatible with NcoI (c/cATGg) and thus may be used for subcloning a target gene or ORF into pTXB3, pTYB3 or pTYB4. The ATG within the recognition sequences of these enzymes can be used for translation initiation. For example, a target gene can be amplified by PCR using a sense primer, which has a BspHI (t/cATGa) or a BspLU11I (a/cATGt) tag (depending on the restriction map of the target gene and the second codon of the ORF). Another enzyme AflIII (a/cRYGt) may also be used for cloning. Second, the NcoI site in the vector (pTYB3 or pTYB4) can be filled-in and used for ligation with a 5′ blunt end fragment starting from the second codon of the ORF. Third, other sites downstream of NdeI or

NcoI sites may be chosen as 5´ cloning site. However, this would result in addition of vector-derived amino acid residues to the N-terminus of the protein of interest. Finally, an alternative strategy is to use the XbaI site, upstream of the polylinker region, as a 5´ cloning site. The insert in this case should contain an optimal Shine-Dalgarno sequence and translation initiation codon.

39. Can XmaI be used for cloning the gene of interest into pTYB2 or pTYB4 vector?

It is not recommended to use XmaI for cloning a gene or ORF into a pTYB N-terminal fusion vector. The recognition sequence (CCCGGG) of SmaI and XmaI in the polylinker of pTYB2 or pTYB4 vector encodes for Pro-Gly adjacent to the intein sequence. When the XmaI site is used for cloning, ligation of XmaI-digested insert and a XmaI-digested pTYB vector results in the regeneration of the XmaI site (forming an ORF-CCCGGG/intein-CBD fusion). Thus, the purified protein contains Pro-Gly at the carboxyl terminus. Instead, SmaI should be used to prepare the pTYB vector so that the purified protein product contains only a single extra amino acid residue, glycine. SmaI digestion of a pTYB vector leaves a blunt-end (5´ GGG/intein...) for ligation with the blunt 3´ end of the insert (forming an ORF-GGG/intein-CBD fusion). Thus, it is not correct to design a XmaI/SmaI site in the reverse PCR primer for cloning unless proline is the native carboxyl amino acid residue.

40. How can the recombinant plasmid (pTYB, pTXB or pCYB vector with an insert) be analyzed by restriction digestion when the SapI or SmaI site is used for cloning?

If SapI or SmaI sites are used for cloning, the recognition sequence may not be regenerated. The KpnI site, which is located 12 nucleotides downstream of the N-terminus of the SceVMA1 intein in pTYB1-4 (or pCYB1-4 vectors), can be used for restriction analysis. The SpeI site present 17 nucleotides downstream of the N-terminus in the Mxe GyrA intein present in the pTXB vectors, can be used for restriction analysis. The NdeI (or NcoI) site may be used to cleave the 5´ end of the insert. Other combinations of double digestions (e.g. NdeI and PstI, NcoI and PstI, XbaI and KpnI) may also be used to determine whether the clones contain the desired insert. Analysis of an insert in pTYB21, pTYB22, pTYB11 and pTYB12 can be achieved using the SacII (unique in the intein) and PstI sites.

41. In the case of C-terminal fusion vectors, if the EcoRI site is used as the 3´ cloning site, which vector should be chosen?

The residues (Ser Ser) encoded by the polylinker in pTXB1 or pTYB1 and pTXB3 or pTYB3 preceding the intein results in poor cleavage. Please note that the SapI site can be used to introduce residues favorable for cleavage. If a 3´ cloning site other than the SapI site is chosen, either the pTYB2 or pTYB4 vectors should be used since glycine is the native residue at the N-terminal splice junction of the yeast intein (Sce VMA), thus optimizing cleavage reaction.

Expression

42. Recommended strains for the IMPACT vectors?

Host Strain	Genotype	Availability
ER2566	*F- lamda- fhuA2 [lon] ompT lacZ:: T7 gene1 gal sulA11* D*(mcrC-mrr)114::IS10 R(mcr-73::miniTn10—TetS)2 R(zgb-210::Tn10) (TetS) endA1 [dcm]*	New England Biolabs
BL21 (DE3)	F- *ompT hsdSB (rB-mB-) gal dcm*(DE3)	Invitrogen (cat# C6000-03)
BL21 (DE3)pLysS	F- *ompT hsdSB* (rB-mB-) *gal dcm*(DE3) pLysS (camR)	Invitrogen (cat# C6060-10)

A competent version of ER2566, T7 Express Competent *E. coli* (NEB #C2566H), can be purchased separately from New England Biolabs, Inc. Non-competent strains from New England Biolabs are available upon request at no extra charge with an order or for the cost of shipping if ordered separately. New England Biolabs, Inc. 1-800-632-5227.

43. Which *E. coli* host strains can be used with the pCYB vectors?

pCYB plasmids (previously supplied with the IMPACT I system) 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. Thus plasmid DNA or ligated DNA should be introduced into a restriction-deficient *E. coli* strain (r-m- or r-m+). The table that lists the expression strains, which have been tested with the pCYB vectors in NEB. Note: the culture should be inoculated with a freshly grown colony. For optimal expression of each fusion protein different strains should be tested.

Host Strain	Genotype	Availability
TB1	F-*ara* D*(lac-proAB) rpsL* (Strr) [*phi80*d*lac*D*(lacZ)M15*] *thi hsdR* (rk-mk+)	New England Biolabs (#E4122S)
ER2267	e14-(McrA-) *endA1 supE44 thi-1 relA1? rfbD1? spoT1?* D*(mcrC-mrr)114::IS10* D*(argF-lac)U169 recA1[F' proA+B+ lac1q* D*(lacZ)M15 zzf::*mini-Tn*10*(Kanr)	New England Biolabs (#E4103S)
JM109	*e14-(McrA -) recA1 endA1 gyrA96 thi-1 hsdR17 (rk-mk+) supE44 relA1* D*(lacproAB) [F' traD36 proAB lacIqZ*D*M15]*	New England Biolabs (#E4107S)

Ordering Information:
Non-competent strains from New England Biolabs are available upon request at no extra charge with an order or for the cost of shipping if ordered separately. New England Biolabs, Inc. 1-800-632-5227.

44. Should DTT be added to the SDS Sample Buffer for SDS PAGE analysis?

DTT should be added to an aliquot of 3X SDS Sample Buffer to a final concentration of 40 mM. For analysis of IMPACT purification you may use SDS Sample Buffer containing DTT. However, boiling in DTT-containing SDS Sample Buffer can cause partial or complete cleavage of the fusion protein, resulting in an overestimation of in vivo cleavage. If substantial in vivo cleavage is observed, the cell extract should be evaluated in a SDS Sample Buffer without DTT.
TCEP [tris-(2-carobxyethyl) phosphine] or TCCP [tris-(2-cyanoethyl) phosphine] (1-5 mM) may also be used as a reducing agent in SDS-PAGE Sample Buffer in place of DTT. TCEP is easier than TCCP to dissolve.
Dissolve TCCP or TCEP in the buffer at the concentration recommended and adjust the pH to 8.5.
TCEP:
We used the chemical from PIERCE Catalog #20490.
SIGMA Product Number- 93284
Invitrogen: Catalog Number T-2556

45. How many positive clones should be checked for expression and purification?

It is recommended that several positive clones, previously identified by restriction digestion, should be analyzed for expression of the fusion protein by stained SDS-PAGE and if not detectable by Coomassie staining then western blot analysis. Mini cultures (2-10 mls) may be done to check for expression. For optimizing expression 50-100 ml cultures should be used. PCR may generate stop codons and other mutations, which may affect the expression of the fusion protein or the activity of the target protein. If a purified protein exhibits low activity, other clones should be used for purification and activity assays. The clones should eventually be sequenced. The ends of the insert can be sequenced using the vector primers provided in the kit.

46. What factors contribute to the low expression of some proteins?

Successful expression and purification of a target protein in an IMPACT vector depend on the following factors: (a) *E. coli* strain; (b) culture conditions (e.g., temperature, aeration, cell density); (c) induction and protein expression conditions (temperature, time, IPTG concentration); (d) expression level of the fusion protein; (e) solubility of the fusion protein; (f) efficiency of thiol-induced cleavage; (g) solubility of the target protein after the cleavage.
The expression of the fusion protein in the C-terminal fusion vectors (pTXB1,3,pTYB1,2,3 and 4) may be primarily determined by the expression level of the N-terminal target protein. In the case of the N-terminal fusion vectors (pTYB21, pTYB22, pTYB11 and 12) the expression level is less dependent on the target protein. It is conceivable that different target proteins, due to certain structural constraints, may prefer either C-terminal or N-terminal fusion to allow proper folding of the fusion precursor and a high level of protein expression. Some target proteins tested (the

majority of them eukaryotic proteins) did not express well in E. coli. The use of a eukaryotic expression system may ultimately help to increase the yield and obtain properly modified eukaryotic proteins. The presence of RNA structure that sequesters the translation initiation sequence may decrease the efficiency of translation. Poor codon usage, mRNA degradation or proteolysis 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 as well as proteolysis. Protease deficient hosts should be tested to minimize proteolysis. Improved codon usage or translational coupling may be effective in increasing the expression level. Another possibility for poor expression could be that the clone is not stable due to the toxicity of the target protein to the host cells. One should inoculate the medium with a freshly grown colony and induce the expression at lower temperatures with lower IPTG concentrations (0.01 - 0.3 mM). Some target proteins become insoluble after on-column cleavage and therefore are only eluted after incubation of the resin with0.3 N NaOH. In this case, increasing the salt concentration (0.5-2 M NaCl) or adding a nonionic detergent to the Cleavage Buffer may improve the solubility of the target protein. (see "What if my fusion protein is insoluble?")

47. How can I optimize expression?

Successful expression and purification of a target protein in an IMPACT vector depends on the following factors: (a) *E. coli* strain - use fresh colony (b) culture conditions (e.g., temperature, aeration, cell density); (c) induction and protein expression conditions (temperature, time, IPTG concentration); (d) expression level of the fusion protein; (e) solubility of the fusion protein; (f) efficiency of thiol-induced cleavage; (g) solubility of the target protein after intein-tag cleavage.

Regarding induction of the fusion protein here are a few suggestions - The most important thing is to start with a fresh colony EVERY time for expression. If you do a transformation you should use the plate the following day to start your culture from one of the colonies. If you leave your plate at 4°C for a couple of days either restreak some colonies or perform a new transformation and start your culture the following day with a fresh colony.

Regarding the temperature of induction - We usually grow our cells at 37°C until an OD600 subscript of 0.5 - 0.7 (0.8 is too high - may not see induction if cells are overgrown) is reached. We then induce with 0.1 - 0.5 mM IPTG (need to titrate for optimal induction) and leave overnight in a 12 -15°C shaker (at 250 rpm). IF performing a 37°C induction the cells should be induced at an OD600 of 0.5 and can be induced for 2-4 hours at 37°C.

We would like to suggest the following protocol (essentially the same as we recommend in the manual)

1. Transform or streak from old culture or colony onto an LB+Amp plate. Incubate the plate at 30°C or 37°C overnight.
2. Inoculate a freshly grown colony in 1 L LB+Amp media Grow the culture to an OD600 of 0.5~0.6. Transfer 2 ml to a sterile tube as an uninduced control.
3. You may try inducing with 0.3 mM IPTG at 37°C or 30°C for 2-3 hours (induce at OD600=0.5); for optimal solubility and folding a lower temperature induction at 15°C overnight (induce at OD600=0.6.

4. Mix 40 µl culture with 20 µl 3X SDS-PAGE sample buffer (+DTT) or 20 µl SDS-PAGE sample buffer (-DTT). Samples: uninduced total cells (+DTT) induced total cells (+DTT in sample buffer)
5. Boil for 10 min. Load 15 µl on SDS-PAGE. Check for expression.
6. For purification, spin down cells (and store at -20°C). Break cells (by sonication at 4°C) in 100 ml cold column buffer/L culture. Centrifuge at 20,000 x g for 30 min. Mix a 40 µl aliquot with 20 µl SDS-PAGE sample buffer. Load the clarified extract onto a 10-20 ml (bed volume) chitin column and follow the protocol in the manual for purification. To check for soluble expression check the clarified extract for the presence of the induced protein.

For optimization of expression 50 - 100 mls cultures should be used.
To reduce the chance of unwanted proteolysis a protease inhibitor may be added. The presence of PMSF (20 µM) or the protease inhibitor cocktail tablets (Sigma, Roche) in the buffers have had no effect on the efficiency of the intein-tag self-cleavage reaction. Therefore, the protease inhibitors may be added to the column buffer for cell lysis. Performing all the steps including induction and purification in the cold may reduce degradation.
NOTE: We offer many different vectors with different inteins because each fusion protein will express and cleave differently. We are unable to predict which intein will work best.

48. What causes lower yields for eucaryotic proteins using the IMPACT system in *E.coli*?

40% of the eukaryotic proteins tested were expressed and purified using the IMPACT System using the C-terminal fusion vectors. The final yields ranged between 0.5-1 mg/liter. The lower success rate for eukaryotic protein purification using the IMPACT System can be attributed to the fact that many eukaryotic proteins express poorly and/or as inclusion bodies in E. coli. Since the target gene is cloned into the polylinker in-frame with the N-terminus of the intein, the expression level of the fusion protein is primarily dependent on the expression of the N-terminal target protein. Thus, we would expect more variation in the yields of both prokaryotic and eukaryotic proteins than when using a C-terminal fusion system (with a highly expressed N-terminal tag). Furthermore, tests at NEB and customer feedback indicate that the use of a eukaryotic expression system may increase expression levels and solubility of eukaryotic proteins.
Novagen sells Rosetta(DE3) cells while Stratagene sells BL21-CodonPlus™ (DE3) Competent Cells. These cells are engineered to contain extra copies of the genes that encode tRNAs for codons in E. coli that are rarely used. Expression in this strain may result in a higher level of expression of your fusion protein. Use a T7 expression strain.

49. What are some suggestions for expressing a toxic gene?

Establish the initial construct in a non expression strain with no T7 RNA polymerase gene (like NEB 10-beta NEB #C3019).
Incubate the plates and inoculate the culture (from a fresh colony) at 20-30°C (not at 37°C)
When you are ready to express, try using a strain which is designed for the expression

of extremely toxic genes. Use of BL21(DE3)/pLys may result in tighter control of expression. Stratagene has a specific system called the Lambda CE6 induction kit, which is based on BL21 cells and bacteriophage CE6. In this case T7 RNA polymerase is introduced via CE6. BL21 cells alone do not carry the T7 RNA polymerase gene.
In the case of solubility problems:
Induce expression at low temperatures to reduce solubility problems (15°C overnight)
Following sonication, check both crude cell extracts and clarified cell extracts by SDS-PAGE and Western analysis to figure out if there is a solubility problem
Use the intein-mediated protein ligation [See Part 6: Applications: Protein ligation and labeling] technique to express and purify a truncated, inactive protein and ligate to the missing expressed sequence to restore activity [Evans et. al.,(1998) *Protein Sci*.7,2256-2264].

50. Should a pTXB or a pTYB vector be used as a control for expression?

pTXB, pTYB vectors are not good controls for expression since the intact polylinker region apparently causes poor expression. The control plasmids pMXB10 (NEB #N6903), pMYB5 (NEB# N6906) or pTWIN-MBP1 (NEB#N6953), which express maltose binding protein (MBP) as the target protein, can be used as a control for expression and purification.

51. Can the intein-CBD enhance solubility?

Based on the tests with the IMPACT system, there appears to be no direct correlation between enhanced solubility and the intein-CBD fusion construct. However, there have been a few examples of proteins that have been solubilized when fused to the intein-CBD fusion. For example, a plant protein was solubilized after being expressed as a three part fusion with the intein and CBD. In another case, the introduction of an extra Gly residue (which is the native residue upstream of the intein) between the protein of interest and the Sce VMA1 intein has also been shown to improve the solubility of the fusion protein. Furthermore, the addition of a nonionic detergent to the column buffer (0.1-0.5% Triton X-100 and 0.1% Tween 20), or higher salt concentration (up to 2 M NaCl) may increase the solubility. Also, since lower temperatures should lessen aggregation, it is recommended to try an induction at 15°C overnight. We now offer many IMPACT vectors containing different inteins and allowing for both N- and C-terminal fusions.

52. What if my fusion protein is insoluble?

If you have problems with solubility one of the first things to try is varying the temperature of induction (15 °C or lower; some customers have used 8 -12 °C induction temperatures) and/or the concentration of IPTG.
There are some steps which effect solubilization of the protein -
A fresh colony from an overnight plate is necessary for optimal expression
Induction conditions - Overnight at 15ºC with 0.3 mM IPTG or 30ºC for 3-6 hours is recommended.
You may also try lowering the IPTG concentration to 0.01-0.1 mM.
Resuspend the cell pellet in at least 100 ml column buffer/L culture.
Sonication - It is crucial to handle the sonication cautiously. If the protein solution gets warm or the solution foams there is a possibility that proteins are being denatured. Use an ice-chilled water bath to keep the cell suspension cool.

If you wish to use a detergent in your buffer, we recommend 0.1 - 0.5% Triton X-100 or 0.1 - 0.2% Tween-20. Too much detergent will impair binding to the chitin column.
When purifying a protein that may be insoluble, several factors should be considered:

1. The binding efficiency of the intein-tag to the chitin resin is lower at 4 M urea or higher
2. The intein-mediated cleavage reaction may be carried out at 0 - 2 M urea
3. The higher the urea concentration, the better the chance to solublize a target protein. However, the cleavage reaction should be performed in 0 - 2 M urea

The following are suggestions for methods to try to keep the fusion protein soluble during cleavage and elution as well as methods to resolubilize after cleavage. Because there are a lot of variables to try, small scale trials are recommended.

1. Buffer : Use the suggested Column Buffer of 20mM HEPES/Tris (pH 8.5), 0.5M NaCl. To this buffer, you could add:
    a. a non-ionic detergent like 0.1% to 0.5% Triton X-100 or 0.1-0.2% Tween-20
    b. Urea: 1-2M Urea will not affect cleavage or binding
2. After Cleavage : if you find varying the buffer conditions did not help
    a. wash the column after the DTT cleavage step with an increasing gradient of Urea, 0-8M. After 6M Urea you will probably see some of the CBD-intein fusion also washing off
    b. pH gradient: perhaps your protein will solubilize at a different pH
3. For renaturation of your protein:
    a. For refolding, solubilize in urea and then dialyse it out. OR
    b. Add PDI (protein disulphide isomerase) after denaturing with 140 mM DTT and 6 M urea - use renaturation buffer of Tris and glutathione (both reduced and oxidized form) and then add PDI OR
    c. Just add PDI

Cleavage of the intein depends on the proper folding of the intein so refolding after denaturation is a crucial step.
We have successfully refolded inteins but each fusion protein is different.
The following protocol was used for refolding the Ssp DnaB intein (Intein1 in the pTWIN vectors):

53. IMPACT-TWIN Renaturation Protocol

The following protocol has been successfully applied to the recovery of proteins fused to intein1 or intein2 of the pTWIN1 vector from inclusion bodies. If the intein-tag is fused to the C-terminus of the target protein then DTT should not be included in the solutions.

1. Resuspend the cell pellet from 1 L of *E. coli* culture in 100 mL Cell Lysis Buffer.
2. Break cells by sonication.
3. Spin down cell debris containing the inclusion bodies at 15,000 x g and 4°C for 30 min.
4. Pour out supernatant and resuspend pellet in 100 mL Breaking Buffer.
5. Stir solution for 1 h at 4°C.
6. Spin remaining cell debris down at 15,000 x g and 4°C for 30 min.
7. Load supernatant into dialysis bag and dialyze against Renaturation Buffer

A, B, C, D, and 2 times E. Each step is against 1 L of the renaturation buffer and dialysis should take at least 3 h at 4°C. During dialysis the buffer should be stirred by a stir bar.

8. Centrifuge the dialyzed solution containing the renatured protein at 15,000g and 4°C for 30 min to remove any remaining impurities or incorrectly folded protein which is again aggregated.
9. Use a standard protocol for chitin chromatography and use the cleavage conditions recommended for the specific intein-tag. Elute the protein product and analyze both the eluate and chitin beads for cleavage extent and protein solubility.
10. Solutions:
    Cell Lysis Buffer: 20 mM Tris-HCl, pH 8.5 and 0.5 M NaCl.
    Breaking Buffer: 20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 7M Guanidine-HCl, 10 mM DTT.
    Renaturation Buffer A: 20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 8 M urea,10 mM DTT.
    Renaturation Buffer B: 20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 6 M urea, 1 mM DTT.
    Renaturation Buffer C: 20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 4 M urea, 1 mM DTT.
    Renaturation Buffer D: 20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 2 M urea, 0.1 mM oxidized glutathione, 1 mM reduced glutathione.
    Renaturation Buffer E: 20 mM Tris-HCl, pH 8.5, 0.5 M NaCl, 0.1 mM oxidized glutathione, 1 mM reduced glutathione.

We have also used the following conditions to solubilize a protein - try solubilizing it in 2, 4, 6 or 8 M urea or fully denature in 6M guanidine-HCl in column buffer and then dialyze in 8 M urea + buffer. Then dialyze or dilute to 4 M and 2 M urea for binding to chitin followed by an extensive wash step. Conduct cleavage at 2 M urea.
You may initially solubilize the fusion protein in 4M Urea and/or 0.5% Triton and then carry out cleavage in 2M Urea and/or 0.1% Triton.
References in:

- Lyles MM., Gilbert, HF. 1991. Catalysis of the oxidative folding of ribonuclease A by protein disulfide isomerase: Dependence of the rate on the composition of the redox buffer. *Biochemistry*. 30: 613-619
- Evans, T.C. Jr., Benner, J., and Xu, M.-Q. (1998) Semisynthesis of cytotoxic proteins using a modified protein splicing element. Protein Sci. 7, 2256-2264.

54. Is there an alternative way to break cells other than sonication?

Cells can also be broken with a French press or by cell lysis reagents (B-PER etc.). Since egg white lysozyme is known to bind and digest chitin, it is not recommended to use lysozyme for cell lysis. Cells can also be broken with a sonicator or French press or by the addition of T4 lysozyme. If a sonicator or a French press or T4 lysozyme 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 MgCl2 (to 5 mM final concentration) to reduce viscosity before the clarified extract can

be passed through the chitin column. Cell lysis products (B-PER etc.) can also be used.

55. Are there antibodies available for western blot detection?

50 µl of rabbit serum (#S6654) raised against the *Bacillus circulans* chitin binding domain (CBD) is provided with the IMPACT kit. A 1:5000 dilution of anti-chitin binding domain serum (Polyclonal anti-CBD) may be used to analyze crude extracts or purified samples. A western blot is carried out when the fusion protein is not detectable by Coomassie staining. NEB also sells murine anti-CBD monoclonal antibody (#E8034S), which can be used at 1:1000. The advantage of the monoclonal antibody is that it has higher specificity and results in lower background compared to the polyclonal antibody. Rabbit serum raised against the S. *cerevisiae* VMA1 intein is available upon special request.

56. What are the common problems associated with western blot analysis?

One of the most common problems in western blot analysis occurs when SDS-polyacrylamide gels are overloaded with protein sample. This results in protein bands, which appear smeared or aggregated, particularly in the absence of DTT in the loading sample buffer. Usually, 1-10 µl of induced culture should be loaded. If cell pellet from 1 liter culture is resuspended in 100 ml column buffer, 1-2 µl of each sample should be loaded. We also load 1:10 and 1:100 dilutions of the cell extract. TCEP [tris-(2-carboxyethyl)phosphine] (PIERCE) can be used as a reducing agent in SDS-PAGE Sample Buffer in place of DTT to improve the quality of SDS-PAGE. Unlike DTT, TCEP should not cause cleavage of the fusion protein. Alternatively, samples may be prepared with DTT-containing SDS-PAGE Sample Buffer as long as a control sample (induced cell extract) prepared with a DTT-free sample buffer is included for comparison during western blot analysis or stained SDS-PAGE. This will prevent one from overestimating the amount of in vivo cleavage.

If non-specific bands, about the size of the target fusion protein, are detected on Western Blot analysis then the murine anti-CBD monoclonal antibody (#E8034) should be used.

57. What does this mean if Western blot analysis with anti-CBD serum detects a 28 kDa product (when pTXB1,3 are used) or a 55 kDa product (when the C terminal TYB fusion vectors are used) and not the fusion precursor?

This data indicates that either the target protein is degraded or cleaved *in vivo*. The protein samples should be prepared in SDS-PAGE sample buffer without DTT (or 2-mercaptoethanol) since boiling in DTT-containing sample buffer may cause cleavage of the fusion protein. If possible, perform a Western blot with target protein antiserum to differentiate between proteolysis versus intein-mediated cleavage. If proteolysis is evident, try different hosts and/or change the induction temperature. It is believed that in vivo cleavage is caused by hydrolysis of the thioester linkage between the target protein and the intein. In vivo cleavage may be reduced by inducing cells at different temperatures (for example, 12-15°C overnight or 30°C room temperature for 3-6 hours). Another possibility is to include a favorable residue or sequence immediately adjacent to the cleavage site, which may improve controllable cleavage. Use of another C-terminal fusion vector (different intein) vector may reduce in vivo

cleavage and increase the final yield. Furthermore, expression using a N-terminal fusion vector (pTYB21, pTYB22, pTYB11 or 12) may circumvent the problem caused by an unfavorable C-terminal residue(s).

58. Is *in vivo* cleavage an all or nothing phenomenon?

No. In many cases *in vivo* cleavage is significantly less when cells are induced at a different temperature. Also, a shorter induction time should be tested. Furthermore, even with some in vivo cleavage enough of the fusion protein may be present to continue with purification.

59. General Troubleshooting

1. No expression: Try different induction conditions or a different host. Make sure you start the culture with a fresh colony.
2. Inclusion body: Induction at lower temperature and/or solubilization with urea (see "What if my fusion protein is insoluble?").
3. Premature cleavage in vivo: *In vivo* cleavage may be significantly less when cells are induced at different temperatures (12-18°C). You may also clone the target gene using a different cloning sites and add favorable C- or N-terminal sequence. You may also try using another intein (Mxe, Sce, Ssp, Mth inteins).
4. No or poor cleavage: Try adding a favorable C-terminal or N-terminal sequence to your protein. Use a different intein or fusion system. Make sure pH is at least 8.5 (except for Intein 1 in the pTWIN vectors where cleavage is with a pH 6 buffer).
5. The target protein is insoluble after cleavage and is only eluted by 0.3 M NaOH solution: Use high salt and/or a detergent in the buffer to reduce this effect.

PURIFICATION AND ON-COLUMN CLEAVAGE

60. What is the binding capacity of the resin?

Based on the binding assays with the maltose-binding protein (MBP) fusion, 2 mg of MBP protein was released from a 1 ml bed of chitin beads, after cleavage of the MBP fusion protein. The kit includes 20 ml of chitin beads, which are supplied as a 40 ml slurry in 20% ethanol. Chitin beads are also sold separately, in both 20 ml (NEB#S6651S) and 100 ml (NEB#S6651L) sizes. Store Chitin Beads at 4°C. Temporary storage at -20°C will not affect the binding capacity.

61. How can the binding efficiency be improved?

The binding efficiency can be improved by increasing the size of the chitin column (10-20ml of resin/1L culture), and/or by diluting the clarified cell extracts (100ml column buffer/1L culture) prior to loading. Loading should be conducted slowly (0.5-1 ml/min).

62. The protein to be purified requires DTT or 2-mercaptoethanol in the Column Buffer. What effect will this have on cleavage?

The presence of low concentrations of DTT (1 mM) or 2-mercaptoethanol (5 mM) may not cause significant cleavage in a few hours. If DTT and 2-mercaptoethanol are absolutely essential for the target protein, try to keep the extract at 4°C and perform column chromatography promptly. (See the next question for the use of other reducing agents).

63. Which reducing reagents can be used to stabilize a thiol-dependent enzyme?
TCEP [tris-(2-carboxyethyl)phosphine] (PIERCE) can be used at 0.1-1 mM final concentration in column buffer to stabilize oxidation-sensitive proteins during purification [Burns, J.A., et. al., (1991) *J. Org. Chem*. 56, 2648-2650]. These compounds specifically reduce disulfide bonds without affecting intein-mediated cleavage reaction and thus it can be used to stabilize proteins with essential thiols. Furthermore, TCEP or TCCP (1-5 mM) may also be used as a reducing agent in SDS-PAGE Sample Buffer in place of DTT, which can cause partial or complete cleavage of the fusion protein when certain amino acids are at the position preceding the cleavage site (C-terminal fusion vectors).

64. Will protease inhibitors interfere with the intein-mediated cleavage reaction?
The presence of PMSF (20 µM) or the protease inhibitor cocktail tablets (Sigma, Roche) in the buffers have had no effect on the efficiency of the cleavage reaction. These protease inhibitors can be added to the column buffer for cell lysis.

65. Is the 0.1% Triton X-100 in the Column Buffer necessary or could it be replaced by a comparable concentration of another detergent?
0.1% Triton X-100 is not necessary in the column buffer. High salt concentrations (up to 2 M NaCl tested), Triton X-100 (up to 0.5%), Tween 20 (up to 0.2% tested), and 50 mM EDTA had little effect on chitin binding or intein cleavage. However comparable concentrations of other detergents have not been tested yet. We recommend that other detergents be tested using the MBP fusion protein which can be expressed from the control plasmids pMXB10 (NEB #N6903), pMYB5 (NEB#N6906) or pTWIN-MBP1 (NEB#N6953).

66. If the fusion protein forms inclusion bodies, can I try to purify protein under denaturing conditions?
Purification and cleavage of the C-terminal fusion (MBP-Sce VMA intein-CBD fusion) protein has been tested under various urea concentrations. 1-4 M urea allowed 80% binding efficiency while 5-8 M urea showed 30-50% binding efficiency. The efficiency of cleavage with 30 mM DTT in 2 M urea is at least 50%, in comparison to reactions under native conditions. Based on these data, chitin chromatography and on-column cleavage may be carried out in 1-2 M urea. Since DTT may cause rapid cleavage of the fusion protein under denaturing conditions, following the wash step the column flow should be stopped and the DTT stock solution should be added directly to the column.

67. If the target protein is sensitive to DTT, are there alternative means to induce on-column cleavage?
If the activity of the target protein is affected by high concentrations of DTT or 2-mercaptoethanol, lower concentrations of DTT or 2-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. Alternatively, 50 mM of freshly prepared hydroxylamine (for pTYB1 and pTYB2) or cysteine solution (at pH 8-9) can be used to induce cleavage at 4-25°C. Be aware that when hydroxylamine or cysteine is used with pTXB1,3, pTYB1,2,3 or 4, they form a

stable covalent bond with the C-terminus of the target protein. One should determine whether a C-terminal hydroxylmate or cysteine affects the activity of the target protein. When cysteine is used for cleavage with pTYB21, pTYB22, pTYB11 or pTYB12, the cysteine is not attached to the target protein. If hydroxylamine or cystein is used the C-terminal thioester is not generated.

If the recombinant protein is to be used in a ligation reaction the use of 2-mercaptoethanesulfonic acid (MESNA) is recommended [Evans et al., (1998) *Protein Sci*. 7, 2256-2264].

68. How can an on-column cleavage reaction be carried out if the target protein has optimal activity at low salt?

The composition of the Cleavage Buffer is not critical for the cleavage reaction and can be made similar to the final storage buffer for the target protein. Wash the column extensively (at least 20 column volumes) with high salt Column Buffer (this removes proteins that may bind to chitin by ionic or other types of nonspecific interactions). Wash the column with at least 3 volumes of the low salt Cleavage Buffer (without DTT or 2-mercaptoethanol) and then flush with 3-5 volumes of DTT-containing Cleavage Buffer. This ensures that the target protein is eluted in the specified buffer.

69. How can *E.coli* GroEL be eliminated from the elution sample?

Proteins can be eluted from the chitin column by virtue of their affinity for a target protein. *E. coli* host chaperone protein GroEL (with an apparent molecular weight about 57-60 kDa on SDS-PAGE) has been co-purified with the target protein in some cases. This may be indicative of (partial) misfolding of the target protein. Try different temperature inductions to aid in folding (12-37°C). Washing the column with different column buffers prior to the cleavage step may reduce the chances of co-purification of such proteins. High salt concentration (1-2 M), nonionic detergents, and ligand or co-factors (such as ATP or GTP) may be used in the column buffer for different proteins. Washing the column extensively (20 column volumes) with different column buffer prior to the cleavage step may reduce the chances of co-purification of such proteins. High salt concentration (1-2 M), nonionic detergents, and ligand or co-factors (such as ATP or GTP) may be used in the column buffer for different proteins. You may use one of the following protocols:

1. Wash the column with 5-10 volumes of regular buffer.
   Then wash the column with at least 10 column volumes of wash buffer at room temperature containing 10 mM MgCl2 and 5 mM ATP (fresh, pH 8.5). You may increase the salt concentration to 1M. Make sure the buffer is well mixed in - you may add this buffer, shake the column so that the buffer is thoroughly mixed in and then wash it thoroughly. Then return the column to the cold room and wash with 5 column volumes of regular buffer. You may try different combinations of ATP and/or high salt etc.
2. Another researcher found that doing a few quick 5 minute washes at room temperature with ATP and $MgCl_2$ eliminates these protein-protein interactions. The procedure was to load and wash the chitin resin as indicated in the IMPACT manual, then place the column at room temperature and then remove the remaining buffer. Add two column

volumes of 5 mM ATP and 10 mM $MgCl_2$ in column buffer and mix buffer and resin by rocking for 5 minutes at room temperature. Collect the wash and repeat 3 to 5 times. Then move the column back to the fridge if that is where you are doing your cleavage and move on to the cleavage step following the IMPACT manual. Some chaperones like DnaK seem to come off easily in the first room temperature wash, while others like GroEL take longer to come off.

3. Mr. Li Qi from the New York State Department of Health and School of Public Health, State University of New York at Albany has used the IMPACT system for purification of a protein of interest, which is co-purified with E.coli GroEL (based on protein sequence analysis). The following simple protocol was used to get rid of GroEL. Following the wash step, the chitin column was washed with 5 column volumes of high salt buffer (50 mM Tris-HCl, pH 8/1M NaCl/1mM EDTA).

70. The target protein is not eluted after DTT treatment, but is present on the chitin column.

If the target protein along with the intein-CBD fusion partner is present on the chitin column after elution of the target protein, this suggests that the target protein became insoluble after cleavage. Increasing the salt concentration (0.5-2 M NaCl) or the addition of a nonionic detergent in the Cleavage Buffer may improve the solubility of the target protein. A number of nonionic detergents (0.1-0.5% Triton X-100 or 0.1-0.2% Tween 20) examined had little effect on binding or cleavage. If urea is used to elute the column, some intein-CBD molecules may also elute with the target protein. The target protein may need to be further purified and then refolded.

71. What are the solutions to reduce co-purification of CBD fusion proteins?

In some experiments, uncleaved fusion protein and intein-CBD species may be present in various amounts in the elution samples. This leakage may occur, for example, when the chitin column is saturated. Furthermore, when cleavage is inefficient, the fusion protein may appear as a dominant contamination species. Several approaches may be used to improve the purity of the eluted samples. First, an increase in the chitin column volume may help to reduce potential overloading problems. Also make sure that the load is dilute (resuspend pellet from 1L of cells in at least 100 mls Column buffer). Second, use of column buffer without detergents may help to reduce the background in some cases. Third, wask the column with at least 20 column volumes of buffer. Finally, the CBD fusion proteins may be reduced by passage of the elution sample slowly through a second chitin column. To improve cleavage efficiency a different fusion construct, for example, with additional favorable residue(s) adjacent to the cleavage site, may help to solve this problem.

72. If the on-column cleavage is inefficient and the fusion protein is the major product on the chitin beads, what should I do?

Longer incubation time (>24 hours at 4°C), or higher temperatures (10-25°C) may help to facilitate the cleavage reaction. Higher pH (8-9) and higher DTT concentration (100 mM) may also improve cleavage efficiency. Try inducing cleavage with other thiol compounds such as 2-mercaptoethanol or hydroxylamine. The low cleavage efficiency may be caused by an unfavorable residue or sequence of the target protein,

adjacent to the intein cleavage site. The inclusion of favorable residue(s) between the target protein and the intein may be necessary to improve the efficiency of the cleavage reaction.
For cleavage of intein1 in the pTWIN vectors 0.2% Tween in the buffer at pH 6 helped cleavage.

73. How can cleavage of the fusion protein be slowed down during the purification process?

Purification should generally be carried out promptly and in the cold in order to reduce the risk of in vitro cleavage. Solutions and samples should be kept at 4°C. Although cell extracts should be loaded slowly onto the chitin column, the flow rate during the wash step can be increased. Some fusion proteins may cleave in vivo and during purification due to hydrolysis of the thioester linkage. If a significant portion of a fusion protein is cleaved in vivo, cleavage may occur during purification, resulting in lower final yield of the product. Rarely does a significant amount of cleavage occur during purification; since higher pH (>pH8) increases the rate of hydrolysis. If *in vitro* cleavage is a problem, extracts may be loaded and washed at pH 6. On-column cleavage is then induced by equilibrating the chitin column using DTT-containing column buffer at pH 8.5.

74. How do I remove DTT after cleavage?

Dialysis will remove the free DTT from solution, and the DTT moiety will be released from the carboxyl terminus of the target protein following cleavage.

75. When I use the N-terminal fusion vectors (pTYB21, pTYB22, pTYB11 and pTYB12), how do I remove the small peptide co-eluted with the target protein?

Dialysis will remove the small peptide (1.2 kDa) that is co-eluted with the target protein. Due to its small molecular weight, the N-extein cleaved peptide cannot be detected on a regular SDS-PAGE gel, and the DTT moiety will be released from the carboxyl terminus of the target protein following cleavage.

76. Are there known ways to accelerate the hydrolysis of DTT from the C terminus of the target protein in a controlled manner?

Storage of the eluted protein samples at higher pH - such as 8 or 9 will increase the rate of hydrolysis. Also storage at higher temperature, such as room temperature, will allow hydrolysis to proceed faster than at 4ºC.

77. When using the N- and C-terminal fusion construct (pTWIN system) 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 [Iwai, H., and Pluckthun, A. (1999) *FEBS Lett*. 459, 166-172].

78. Can the chitin beads be regenerated?

The chitin beads can be regenerated at least 5 times. After stripping the chitin column with 3 volumes of 0.3 M NaOH (Stripping solution). Allow the beads to soak in the Stripping Solution for 30 minutes, wash the resin with additional 7 column volumes of 0.3 M NaOH. Then wash the resin with 20 bed volumes of water followed by 5 column volumes of Column Buffer. For long term storage, add 0.02% sodium azide to the Column Buffer. The beads should be stored at 4°C. Temporary storage at -20°C will not affect the binding capacity.

APPLICATIONS/LIGATION(IPL) AND TWIN

79. What is Intein-mediated Protein Ligation (IPL)?

The IPL reaction allows the ligation of a synthetic peptide or a protein with an N-terminal cysteine residue to the thioester on the C-terminus of an expressed protein through a native peptide bond [Evans et al., (1998) *Protein Sci*.7, 2256-2264]. The IPL protocol employs the IMPACT C-terminal fusion vectors to express and purify a protein of interest and to generate a thioester at its C-terminus. For IPL we recommend the use of pTXB1 when possible. The IPL reaction applies the chemistry described for "native chemical ligation" which fuses two synthetic peptides when the N-terminal cysteine of one peptide attacks a C-terminal thioester of another peptide [Dawson et al.,(1994) Science 266, 776-779; Tam et al., (1995) *Proc. Natl. Acad. Sci*. USA 92, 12485-12489]. Initially, a new thioester bond is formed by transthioesterification involving attack by the sulfhydryl group of the N-terminal cysteine residue on the C-terminal thioester. The transitory ligation product then undergoes a spontaneous S-N acyl rearrangement from a thioester to a stable peptide bond. This technique has also been described as "expressed protein ligation" [Muir et al.,(1998) Proc. Natl. Acad. Sci. USA 95,6705-6710; Severinov and Muir(1998) *J. Biol. Chem*. 273, 16205-16209].

80. What has IPL been used for?

Because IPL allows the fusion of synthetic peptides, as well as bacterially expressed proteins, with an N-terminal cysteine, to a protein expressed in the IMPACT system, IPL has been used in a variety of ways including:

- The expression of cytotoxic proteins [Evans et al., (1998) *Protein Sci*.7: 2256-2264].
- The labeling of proteins with radioactive compounds as well as with synthetic peptides containing biotin or fluorescein [Chong et al., (1997) Gene 192: 277-281; Evans et al., (1998) *Protein Sci*.7: 2256-2264, Muir et al.,(1998*) Proc. Natl. Acad. Sci*. USA 95:6705-6710]
- The study of protein-protein interactions [Severinov et al., (1998) *J. Biol. Chem*. 273:16205-16209]
- The generation of kinase substrates by varying the kinase recognition site at the protein level instead of at the DNA level [Ghosh, I. et al., (2004) *J. of Imm. Methods*. 293:85-95; Xu, J., Sun, L., Ghosh, I., and Xu, M.-Q. (2004) Western blot analysis of Src kinase assays using peptide substrates ligated to a carrier protein. *Biotechniques*. 36:976-998.]
- The generation of phosphatase substrates [Kochinyan, S. et. al. (2007)

Use of intein-mediated phosphoprotein arrays to study substrate specificity of protein phosphatases. *Biotechniques* 42(1): 63-9]

- The isotopic labeling of proteins for NMR analysis [Xu et al., (1999)*Proc. Natl. Acad. Sci. USA* 96,388-393]
- Generation of substrates for protein arrays [Sun, L., et al., (2004) *Biotechniques*. 37:430-443.]
- Site specifically incorporating lipid moieties into a protein [Rak et al., (2003) *Science* 302, 646-650]

Please refer to "What Is IMPACT?" for more references on IPL.

81. Can IPL be used to generate substrates for enzyme analysis?

Please refer to the Peptide Carrier Kit (NEB#E6600) and the following references for more information.

1. Sun, L., Ghosh, I., Barshevsky, T., Kochinyan, S. and Xu, M.Q. (2007) Design and Use of Ligated Phosphoproteins: Study of Protein Phosphatases by Dot Blot Array, ELISA and Western Blot Analysis. *Methods*. 42, 220-226.
2. Kochinyan, S., Sun, L., Ghosh, I., Barshevsky, T., Xu, J. and Xu , M.Q. (2007) Use of Intein-Mediated Phosphprotein Arrays to Study Substrate Specificity of Protein Phosphatases. *Biotechniques* 42: 63-69.
3. Xu, M.Q., Ghosh, I., Kochinyan, S. and Sun, L. (2006) Intein-mediated Peptide Arrays for Epitope Mapping and Kinase/Phosphatase Assays. Methods in Molecular Biology, vol., Microarrays: Methods and Protocols Edited by J.B. Rampal. Humana Press Inc., NY: 313-338.
4. Sun, L., Rush, J., Ghosh, I., Maunus, J.R. and Xu, M.-Q. (2004) Producing peptide arrays for epitope mapping by intein-mediated protein ligation. *Biotechniques*. 37: 430-443.
5. Ghosh, I, Sun, L., Evans, T.C. Jr., and Xu, M.-Q. (2004) An improved method for utilization of peptide substrates for antibody characterization and enzymatic assays. *J. of Imm. Methods*. 293:85-95
6. Xu, J., Sun, L., Ghosh, I., and Xu, M.-Q. (2004) Western blot analysis of Src kinase assays using peptide substrates ligated to a carrier protein. *Biotechniques*. 36:976-998.

82. What vectors are suitable for IPL and cyclization?

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 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 (fusion with intein 1), 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.

Studies by NEB scientists have indicated that pTXB vectors are more suitable for IPL than the pTYB vectors because they cleave more proficiently with thiol reagents that

are best for ligation, such as 2-mercaptoethanesulfonic acid (MESNA) and thiophenol. Cleavage of fusion proteins expressed from a pTYB vector may require a higher incubation temperature with these thiol reagents. However, it may be advantageous to express a protein of interest using both pTXB1,3 and pTYB1,2,3,4 vectors (cloning the insert using the same restriction sites) and examining both constructs for expression and purification.

83. What cleavage reagent should be used for IPL?

The thiol reagent used to induce cleavage during the standard IMPACT purification, typically 1,4-dithiothreitol (DTT), lowers the proficiency of the ligation reaction [Evans et al.,(1998) *Protein Sci*.7,2256-2264]. Use of 2-mercaptoethanesulfonic acid (MESNA) or thiophenol results in greater than 90% ligation proficiencies. We recommend the use of 2-mercaptoethanesulfonic acid, sodium salt (MESNA) (SIGMA Chemical Co., Cat.No. M-1511) because it is an odorless compound. We recommend the use of pTXB1 or pTXB3 for IPL as the Sce VMA intein in pTYB vectors does not cleave well with MESNA, resulting in lower yields of the target protein.

84. Which products should be ordered for IPL?

The cloning, expression and purification procedures for IPL are essentially the same as the standard IMPACT protocol. You need to order a pTXB or pTWIN vector, primers for sequencing the insert and MESNA. You may order the standard IMPACT Kit (NEB #E6901) in addition to the primer or individual components (TXB vector, chitin beads, anti-CBD antibody, primers).

85. What are the column buffers recommended for chitin chromatography and on-column cleavage?

A column buffer of 20 mM Tris-HCl, pH 8.5 containing 500 mM NaCl is recommended for the IMPACT purification step. Cleavage buffer should be freshly prepared using the column buffer supplemented with 50 mM 2-mercaptoethanesulfonic acid at 4°C. Other possible buffer conditions: 10-100 mM Tris-HCl, pH 8.0-9.0, 100-2000 mM NaCl.

86. What is the protocol for ligation?

A typical lPL reaction can be carried out by mixing two reactants at 4°C – 25°C at pH 8 - 9. One of the components should have a final concentration of at least 0.5 - 1 mM. Regarding ligation of a peptide to a protein, for greater extent of ligation, we use the protein at 1 -10 uM with 0.5 - 1 mM peptide; we add the two components in the presence of 0.1 M Tris pH 8.5 and 10mM MESNA and let the reaction go overnight at 4°C or one hour at room temperature.

The following protocol illustrates a typical labeling experiment. Mix 4 µL of L-[35S]-cysteine (11.0 mCi/mL, Perkin Elmer, cat# NEG-022T) or 4 µL of peptide with an N-terminal cysteine (10 mM) with a 36 µL aliquot of protein freshly isolated using the IMPACT system. Incubate the reaction solution at 4°C overnight. To examine the labeled or ligated protein, add 20 µL of 3X SDS-PAGE sample buffer (with DTT) to the protein sample and boil for 10 minutes. Analysis of the sample can then be performed using SDS-PAGE.

The following references may be useful:

Reviews

- Xu, M.-Q., and Evans, T. C. Jr. Intein-mediated ligation and cyclization of

Expressed Proteins. *Methods* 24, 257-277.
- Xu, M.Q., Paulus, H. and Chong, S. (2000) Fusions to self-splicing inteins for protein purification. *Methods Enzymol*. 326:376-418.
- Evans T. C. & Xu, M.-Q., (1999) Intein-mediated protein ligation: harnessing nature's escape artists. *Biopolymers* 51, 333-42.

pTXB1,3
- Evans, T.C. Jr., Benner, J., and Xu, M.-Q. (1998) Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci*. 7, 2256-2264.

pTWIN1,2
- Evans, T.C. Jr., Benner, J., and Xu, M.-Q.(1999) The cyclization and polymerization of bacterially-expressed proteins using modified self-splicing inteins. *J. Biol. Chem*. 274, 18359-18363.
- Evans, T.C. Jr., Benner, J., and Xu, M.-Q. (1999) The in vitro ligation of bacterially expressed protein using an intein from Methanobacterium thermoautotrophicum. *J. Biol. Chem*. 274, 3923-3926.
- Mathys, S., Evans, T.C. Jr., Chute, I.C., Wu, H., Chong, S., Benner, J., Liu, X.-Q., Xu, M.-Q. (1999) Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. *Gene*, 231:1-13.

Applications
- Kochinyan, S.; Sun, L.; Ghosh, I.; Barshevsky, T.; Xu, J.; Xu, M. Q. (2007) Use of intein-mediated phosphoprotein arrays to study substrate specificity of protein phosphatases. *Biotechniques* 42(1): 63-9.
- Sun, L., Rush, J., Ghosh, I., Maunus, J.R. and Xu, M.-Q. (2004) Producing peptide arrays for epitope mapping by intein-mediated protein ligation. *Biotechniques*. 37:430-443.
- Ghosh, I, Sun, L., Evans, T.C. Jr., and Xu, M.-Q. (2004) An improved method for utilization of peptide substrates for antibody characterization and enzymatic assays. *J. of Imm. Methods*. 293:85-95.
- Xu, J., Sun, L., Ghosh, I., and Xu, M.-Q. (2004) Western blot analysis of Src kinase assays using peptide substrates ligated to a carrier protein. *Biotechniques*. 36:976 -981.

87. Which residues at the C-terminus of the target protein may inhibit cleavage or cause in vivo cleavage when the pTXB vectors are used?

Based on mutagenesis studies using paramyosin as the target protein, Ser, Pro or Asp at the -1 position (the amino acid preceding the intein) blocks the cleavage When the C-terminus of a target protein contains an unfavorable residue, additional residue(s) such as Tyr (the native residue preceding the Mxe GyrA intein), may be inserted between the target protein and the N-terminus of the intein to improve controlled cleavage. However, the cleavage efficiency may vary when a different target protein is fused to the intein.

88. What are the sequencing primers for the pTXB vectors?

The insert in a pTXB vector can be sequenced by the T7 Universal Primer (NEB#S1248), and Mxe intein reverse II primer (NEB#S1285, 5´

GATTGCCATGCCGGTCAAGG -3´), which anneals to the intein sequence 107 base pairs from the cleavage site and sequences in the orientation opposite to transcription.

89. References for Intein-mediated Protein Ligation(IPL):

1. Evans, T.C. Jr., Benner, J., and Xu, M.-Q. (1999) The cyclization and polymerization of bacterially-expressed proteins using modified self-splicing inteins. *J. Biol. Chem*.274, 18359-18363.
2. Evans, T.C., Benner, J., and Xu, M.-Q.(1998) Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci*. 7, 2256-2264.
3. Evans, T.C., Benner, J., and Xu, M.-Q.(1999) The in vitro Ligation of Bacterially Expressed protein Using an Intein from Methanobacterium thermoautotrophicum. *J. Biol. Chem*.274. 3923-3926.
4. Muir, T.W., Sondhi, D., and Cole, P.A.(1998) Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci*. USA 95, 6705-6710.
5. Severinov, K., and Muir, T.W.(1998) Expressed protein ligation, a novel method for studying protein-protein interactions in transcription. *J. Biol. Chem*. 273, 16205-16209.
6. Xu, R., Ayers, B., Cowburn, D., Muir, T. W.(1999) Chemical ligation of folded recombinant proteins: Segmental isotopic labeling of domains for NMR studies. *Proc. Natl. Acad. Sci. USA* 96, 388-393.
7. Gimble, F.S.(1998) Putting protein splicing to work. Chem. and Biol. 5, R251-R256.
8. Holdford, M., Muir, T. W.(1998) Adding 'splice' to protein engineering.*Structure* 6,945-949.

90. Can the IMPACT System be used in other expression systems?

The IMPACT system has been successfully used in baculovirus and yeast. The IMPACT vectors for expression in baculovirus cells have been constructed and are available for ß-testing. It has been shown that a eukaryotic expression system may improve expression and solubility of a eukaryotic protein. Proteins in the range of 20-200 kDa have been purified.

91. How can the C-terminal fusion vectors be used to label the C-terminus of the target protein?

The C-terminus of the target protein can be covalently labeled by using L-[35S]-cysteine or by using a biotinylated synthetic peptide with an N-terminal cysteine (see below), immediately following thiol-induced on-column cleavage and elution. The thiol-tagged protein is mixed with the label and incubated at 4°C overnight. The sulfhydryl group of the cysteine can initiate the attack at the thioester bond present at the C-terminus of the target protein. The cysteine residue then forms a stable peptide bond with the C-terminus of the target protein by a spontaneous S-N shift. The L-[35S]-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. The choice of thiol for cleavage and labeling depends on the desired outcome. DTT causes efficient on-column cleavage of the precursor protein and leads to maximum protein yields. However, the percentage of total protein

labeled is lower than another thiol, 2-mercaptoethanesulfonic acid (MESNA). The use of MESNA as the thiol to induce on column cleavage has been found to increase the percentage of the total protein that is labeled. However, MESNA is not as efficient at causing on-column cleavage as DTT and so the total yield (labeled and unlabeled) of protein may be lower. In the cases in which a high specific activity or high biotin incorporation is required the use of MESNA is recommended. The incorporation of a biotinylated peptide onto the C-terminus of a protein can be accomplished using freshly isolated target protein and a biotinylated peptide with an N-terminal cysteine. NEB supplies peptides to specifically label the C-terminus of a thioester tagged protein with either a fluorescein (NEB #P6606) or a biotin (NEB #P6607) conjugated molecule.

92. Can the IMPACT System be used to pull out proteins from *E. coli* or mammalian cell extracts which react with the target protein?

Since proteins are purified in a single chromatographic step under native conditions, it is possible to use the IMPACT System to perform the proposed experiment. The protein complex may be released by induction of the intein-mediated cleavage reaction. Otherwise the uncleaved fusion protein and the bound protein may be stripped by 0.3 M NaOH.

93. Can the IMPACT System be used to purify peptides?

The IMPACT System provides an attractive alternative method for peptide synthesis and purification. In each case, expression, stability and inducible cleavage of the fusion have to be optimized. It is also possible to use this system to study peptide interactions with other proteins or ligands.

We believe that it is possible to purify peptides using the IMPACT system but you need to consider two major factors. The first one is that certain residues or sequences may not exhibit efficient cleavage activity; the cleavage efficiency of different amino acid residues flanking the cleavage site varies depending on the target and intein sequence. The second factor is that the yield for peptides could be low due to their small mass. Furthermore, the cleavage efficiency of each intein-target protein fusion may vary. For peptide expression using the IMPACT system please refer to the following reference

> Evans, T.C. Jr., Benner, J., and Xu, M.-Q. 1999. The cyclization and polymerization of bacterially-expressed proteins using modified self-splicing inteins. *J. Biol. Chem*. 274, 18359-18363.

We have created cyclic peptides as small as 9 amino acids and as large as >350 amino acids (Evans, et al (1999) *J. Biol. Chem*. vol 274, pp 18359-18363). In these cases the expression of the fusion proteins were good (see figures 2A and 4A of the previously mentioned reference). This led to good yields of cyclic MBP (ca. 380 amino acids) of ca. 5-10 mg per liter. However, the 9, 10, and 14 amino acid peptides had a much lower yield by mass. This was due to the fact that even though the moles of peptide purified was probably comparable to the moles of MBP, the significantly smaller MW of the peptides resulted in much less mass of product. We got in the microgram range. The bottom line is that peptides can be produced using the IMPACT system, however the yields per liter will probably be in the 50-1000 microgram range.

Here are some more references you may be interested in:

Pezza JA, Allen KN, Tolan DR. (2004) Intein-mediated purification of a recombinantly expressed peptide. *Chem Commun* (Camb). (21):2412-3. (pTWIN1, pTYB1) PMID: 15514791
Morassutti, C., DeAmicis, F., Skerlavaj, B., Zanetti, M., and Marchetti, S. (2002) Production of a recombinant antimicrobial peptide in transgenic plants using a modified VMA intein expression system. *FEBS Letters*. 519: 141-146. PubMedID: 12023033
Cottingham, I.R., Millar, A., Emslie, E., Colman, A., Schnieke, A.E. and McKee, C.(2001) A method for the amidation of recombinant peptides expressed as intein fusion proteins in Escherichia coli. *Nat. Biotechnol*. 10:974-977. (pTYB1, pCYB1) PubMedID: 11581666

94. What is your recommended protocol for scaling-up the chitin column?

First, we recommend estimating the mass of fusion protein to be loaded to the column. Values from the pilot runs should be sufficient in estimating the amount of fusion protein produced, as a percentage of total soluble protein. As a guideline, one can assume a yield of approx. 4-5g of total soluble protein from 100g of cells. Assuming that approximately 10% of that is fusion protein, then typically about 500mg of fusion will be produced in a 100g cell paste (or the equivalent of approximately 20L of culture). Using a value of 2 mg/ml of chitin resin, a 250ml column bed volume would be required in this example. The clarified extract is loaded to the column at a relatively slow flow rate to ensure complete binding, i.e.1-5ml/min. The column is washed with 20 column volumes at a faster rate. The use of a closed column (i.e. with a top flow adaptor) is recommended to ensure complete flushing of the unbound fraction of the cell extract during the wash.

To ensure homogeneous equilibration of the column with DTT, the elution is performed by cycling 2 column volumes of DTT-containing buffer at a concentration of 50 µg/ml through the column. Set up the DTT buffer in a flask on a magnetic stir plate and connect both the inlet and outlet of the column to this flask. Gently stirring, start the flow at the same rate as was used to wash the column, and run the equilibration for approximately 3-5 column volumes. Let the column stand overnight to complete the cleavage. Depending on the rate of cleavage for your fusion, some cleavage of the bound protein may occur during the cycling, but most of the cleavage will continue overnight.

The next day, some of the protein will be in the stirring flask and some will be in the void volume of the column. Set up the outlet of the column to a fraction collector and flush with an additional 2 column volumes of DTT buffer and continue collecting until the protein reading from the column reaches baseline. The protein-containing fractions can then be pooled with the protein remaining in the flask.

An alternate method of loading the column with the clarified cell extract is to do a "batch" absorption. Simply add the required amount of resin to the clarified extract and gently rock the slurry on an oscillating table in a cold room for 1-2 hours. Then pack the absorbed slurry into the column and perform the wash and remaining steps as outlined above.

For HPLC the following protocol has been carried out:

Use of a 300 ml Pharmacia XK 50 column. Chitin beads packed at 5 ml/min. and run at 2 ml/min. If the pressure exceeds .5 MPa the column could burst. For cleavage use a final conc. of 50 mM DTT at 25 degrees for 16 hours. Note: The purified protein was stable at room temperature.

CHITIN-BINDING DOMAIN (CBD)

95. What is chitin?

Chitin is a homopolysaccharide of ß-(1-4) linked N-acetyl glucosamine (GlcNAc). It is a structural material and is one of the most abundant biopolymers. It can be found in most fungi, some algae and it is the major structural material in exoskeletons of many arthropods and mollusks.

96. Where does the chitin binding domain (CBD) in the IMPACT system originate from?

The 52 amino acids of the chitin binding domain are from the carboxyl terminus of the chitinase A1 gene of *Bacillus circulans* WL-12 [Watanabe et al.,(1994) J. Bacteriol.176,4465-4472; Chong et al.,(1997)*Gene* 192, 277-281].

97. Does the *Bacillus circulans* CBD have an affinity for low molecular weight chitin (2 to 6 GlcNAc residues)?

No, two different tests have been performed which illustrate that the CBD does not bind to these low molecular weight chitin molecules. One was to try eluting the CBD fusions bound to chitin with low MW sugars of chitin, 1 to 6 GlcNAc residues. The fusion remained bound to the chitin when washed with each of these six different length chitin sugars, showing that at least the CBD has a higher affinity for the high MW chitin. In another test CBD fusions were loaded onto columns that had single GlcNAc residues or triacetylchitotriose residues linked to agarose beads. No binding was detected in either case.

98. Is there any method to elute the complete fusion protein from the chitin column? For instance, by eluting with high concentrations of GlcNAc or low pH?

The only way to remove the complete fusion from the column at this time is by denaturation with either 0.3 N NaOH, 0.5-1.0% SDS or 6M Guanidinium. 5-8M Urea can only cause partial elution of the fusion protein. The chitin binding domain (CBD) remains bound to chitin at pH4 (in the range pH 3-11). The elution of a CBD-target protein can be performed by 6 M GnCl (pH4) Otherwise, CBD can be stripped at pH 12 or pH 2 - 2.5.

The following conditions were tested for elution of CBD fusions from chitin with little or small effect on binding: 50 mM-2M NaCl, 0.1 - 0.5% Triton X-100, 0.1 - 0.2% Tween 20, 1 - 4 M Urea, pH 2.5-9 and 10 - 25% glycerol. Conditions that seemed to make the fusions probably precipitate and disappear both from wash and final elution with SDS were 10% Acetic Acid, 100% Methanol, 25% Ammonium Sulfate, and 80% Isopropanol.

99. Can the chitin beads be regenerated?

The chitin beads can be regenerated at least 5 times. After stripping the chitin column with 3 volumes of 0.3 M NaOH (Stripping solution). Allow the beads to soak for 30 minutes, wash the resin with additional 7 column volumes of 0.3 M NaOH. Then wash the resin with 20 bed volumes of water followed by 5 column volumes of Column Buffer. For long term storage, add 0.02% sodium azide to the Column Buffer. The beads should be stored at 4°C. Temporary storage at -20°C will not affect the binding capacity.

100. Can you provide additional information about the chitin beads?

The chitin beads are approximately 50-100 µm in size and are made by acetylation of chitosan, with a MW in the range of 750,000. For more information on the preparation of chitin beads, please refer to the Methods section in Fig.3 of the paper by Chong et al.,(1997)Gene,192, 277-281.

101. What is the Km of the chitin binding domain?

The *Bacillus circulans* CBD binds only high MW chitin that is insoluble, making binding kinetics difficult to determine. At this time binding studies have not been conducted, and the only known way to elute the CBD fusion is by denaturation, for example, with 0.3 M NaOH, 1% SDS or 6M Guanidinium hydrochloride.

102. What is the specificity of the CBD to chitin, cellulose, and polyglucans?

The *Bacillus circulans* CBD seems to be very specific for chitin. It does not bind cellulose, xylans, or amylose.