Published in final edited form as:
Methods Mol Biol. 2012 ; 794: 187–197. doi:10.1007/978-1-61779-331-8_11.

Experimental methods for scanning unnatural amino acid mutagenesis

Jia Liu and **T. Ashton Cropp***
Department of Chemistry and Biochemistry, University of Maryland, College Park, MD, 20742, USA

Summary

Site-specific incorporation of unnatural amino acids into proteins *in vivo* relies on the genetic reassignment of nonsense or quadruplet codons. Here, we describe a general procedure for the random introduction of these codons into open reading frames resulting in protein libraries that are scanned with unnatural amino acid residues. These libraries can enable large-scale mutagenesis experiments aimed at understanding and improving protein function.

Keywords

Scanning mutagenesis; tRNA; Mu transposon; Codon; Unnatural amino acids

1. Introduction

More than 50 different unnatural amino acids with distinctive properties have been added to the genetic codes of bacteria, yeasts and mammalian cells (1). Central to this methodology is the reassignment of nonsense or quadruplet codons using orthogonal aminoacyl-tRNA synthetase/tRNA pairs. These chemical tools allow one to control protein functions with light (2), detect transient protein-protein interactions (3), perform protein bio-conjugation reactions (4, 5), among other tasks. While the unnatural amino acid mutagenesis methodology is generally efficient, large-scale mutagenesis experiments for scanning unnatural amino acids is hindered by the generation of genes containing random nonsense or quadruplet codons. Indeed some unnatural amino acids that have been genetically encoded would be even more useful if one could randomly distribute them throughout protein sequence space.

We and others have described new approaches to generating collections of proteins that contain single amino acid mutations located in random positions (6–8). These libraries are rationally diversified and the mutation in question can be any genetically encoded amino acid. Here, we provide detailed experimental methods for scanning unnatural amino acid mutagenesis that can be applied to any protein of interest. This is a sequential method that consists of creating a small library of open reading frames in which each member contains a random, single and in-frame amber stop codon TAG (or another nonsense codon). The

*To whom correspondence should be addressed. Tel: (301)405-1734; Fax: (301)314-9121; acropp@gmail.com.

mutants can be separated and expressed individually with desired unnatural amino acids. Alternatively, the mutant library can be expressed with unnatural amino acid as a mixed population, which can be subsequently used for functional screening or genetic selection. The protocol from start to finish can be expected to take approximately two weeks to complete.

2. Materials

1. Intein targeting plasmid (pIT) (6).
2. Entranceposon (M1-Cam^R) (Finnzymes, Espoo, Finland) for PCR template (*see* Figure 1 and Note 1).
3. DNA containing gene of interest (*see* Note 2).
4. pInSALect vector (9).
5. Appropriate expression vector (*see* Note 3).
6. MuA transposase and 10× transposon reaction buffer (Finnzymes, Espoo, Finland), stored at −20 °C.
7. 10 mM Tris-HCl buffer, pH 8.5 at 25 °C.
8. TE buffer, 10 mM Tris-HCl buffer, 0.1 mM EDTA, pH 8.5 at 25 °C.
9. Primers (*Mly*I restriction sites underlined and reverse complement of TAG codon in italic).
    - Forward (*see* Note 4) and reverse (*see* Note 5) primers for cloning the gene of interest into pIT vector.
    - Forward and reverse primers for cloning the gene of interest from the pIT vector into a desired expression vector.
    - *Mly*I transposon 5'-gcttagatctgactcggcgcacgaaaaacgcgaaag-3’.
    - Linker FWD 5'Phos-ggatcgactctcctgggtattcgcaataatcttaatactgag-3’.
    - Linker REV 5'Phos-*cta*gatctgactcaattaccaatgcttaatcagtgaggcacct-3’ (*see* Note 6).

[1]In this protocol, the *Mly* I transposon DNA is generated by PCR from Entranceposon (M1-Cam^R) DNA. Once constructed, this DNA can be inserted into a cloning vector such that it can be released from the vector by *Bgl* II digestion and contain a 5'-GATC overhang that is important for the transposon reaction (11). This will ensure correctly digested DNA is available for future use (Figure 1B).
[2]The *BamH* I and *Sal* I sites in the genes of interest, if any, should be destroyed by site-directed mutagenesis for downstream manipulations.
[3]The expression vector will be double transformed into an expression cell strain containing a plasmid for incorporation of unnatural amino acids (10). These plasmids contain a chloramphenicol selection marker and p15A origin of replication. Therefore any expression vector used for the protein of interest should have an alternative selection marker and origin of replication.
[4]The forward primer should include a *BamH* I site before the start codon of the gene of interest. The number of oligonucleotides between the start codon and *BamH* I site should be adjusted such that the gene of interest will be in the same reading frame of the N-terminal fusion peptide from the pIT vector (Figure 1A).
[5]The reverse primer should contain a *Sal* I site after the stop codon.
[6]It is recommended to PAGE-purify this primer to eliminate truncated products. We have also found that including a phosphate at the 5'- end can increase the efficiency of linker ligation.

Dissolve primer stocks in TE buffer at a final concentration of 100 μM and store at −20 °C. Dilute to 10 μM prior to use.

10. dNTP stock solution, 10 mM total, 2.5 mM each, stored as 20 μL aliquots at −20 °C.
11. Phusion DNA polymerase (New England Biolabs, Ipswich, MA), stored at −20 °C.
12. Restriction enzymes *Sal* I, *Bam*H I, *Bgl* II, FastDigest *Mly* I and appropriate buffers (Fermentas, Glen Burnie, MD). Store at −20 °C.
13. T4 DNA ligase (Fermentas), stored at −20 °C.
14. 0.7% and 1% agarose gels.
15. QIAquick PCR purification kit (Qiagen, Valencia, CA).
16. QIAEX II gel extraction kit (Qiagen, Valencia, CA).
17. Chemically or electro-competent *Escherichia coli* cells for cloning, such as DH10B. Store at −80 °C.
18. Chemically or electro-competent *Escherichia coli* cells for expression, such as BL21 (DE3). Store at −80 °C.
19. 1000× Antibiotic stock solutions: ampicillin (50 mg/mL), kanamycin (50 mg/mL) and chloramphenicol (35 mg/mL).
20. 1000× reagents for protein induction, such as 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) or 20% arabinose).
21. Unnatural amino acid, sufficient to make media at 2 mM final concentration.
22. Luria Bertani (LB) medium (1 L): 10 g tryptone, 5 g yeast extract and 10 g sodium chloride.
23. SOC medium (1 L): 20 g tryptone, 5 g yeast extract, 10 mM sodium chloride, 25 mM potassium chloride. Add magnesium chloride to 10 mM and glucose to 20 mM after autoclaving.
24. 2×YT medium (1 L): 16 g tryptone, 10 g yeast extract and 5 g sodium chloride, pH 7.2.
25. LB agar plates with appropriate antibiotics.

3. Methods

3.1. Generating a pIT-X vector containing gene of interest

1. Digest 2 μg of the pIT vector (6) using *BamH* I and *Sal* I. Gel-purify the 3.2 Kbp fragment using QIAEX II gel extraction kit (*see* Note 7).

[7]When gel slices are incubated with solubilization buffer (Qiagen), any temperature above 50 °C is not recommended. Higher temperatures denature double-stranded DNA and reduce the yield of gel-purification.

2. Amplify the gene of interest using PCR to introduce a *BamH* I and *Sal* I at the 5' and 3'-ends, respectively.
3. Purify the PCR product using with the QIAquick PCR purification kit.
4. Digest the PCR products using *BamH* I and *Sal* I and gel-purify the digestion products using the QIAEX II gel extraction kit.
5. Ligate the purified digest of pIT vector and PCR product using T4 DNA ligase. Name the recombinant plasmid as pIT-X (*see* Note 8).

3.2. Transposition reaction

1. Prepare the *Mly* I transposon DNA by PCR as follows: 10 ng Entranceposon (M1-Cam^R) template, 0.2 mM dNTPs, 0.5 μM of primer *Mly*I transposon (which serves as the forward and reverse primer), 1× Phusion HF buffer (NEB) and 0.5 unit of Phusion DNA polymerase (NEB) in a 50 μL solution.
2. Incubate the reaction with the following cycle conditions: initial denaturation at 98 °C for 30 s, 30 cycles of 98 °C for 10 s, 60 °C for 30 s and 72 °C for 1 min and final extension at 72 °C for 10 min.
3. Purify the PCR product using with the QIAquick PCR purification kit.
4. Digest the *Mly* I transposon DNA as follows: 2 μg DNA, 10 U of *Bgl* II and 1×buffer O (Fermentas) in a 50 μL solution (*see* Note 9).
5. Load the digestion products on a 0.7% agarose gel and purify the 1.3 kbp fragment.
6. Load 1 or 2 μL of the purified transposon DNA on a 1% agarose gel and determine the concentration by comparison with a DNA standard.
7. Store the purified transposon DNA at −20 °C prior to use.
8. Perform a 20 μL transposition reaction with the following components: 400 ng of pIT-X plasmid DNA, 1.3 molar excess of *Mly* I transposon DNA (*see* Note 10), 1× reaction buffer (Finnzymes) and 1 U of MuA transpoase (Finnzymes).
9. Incubate the reaction at 30 °C for 4 hours.
10. Stop the reaction by adding SDS to a final concentration of 0.1% and heating at 75 °C for 10 min. Cool the reaction solution on ice.
11. Transform each 1 μL of products into 50 μL of electro-competent *E. coli* cells (*see* Note 11) and recover in 500 μL SOC at 37 °C for 1 hour.

---

[8]When inserted into pIT vector, the gene of interest is in-frame with an N-terminal fusion peptide (Figure 1A) which contains a Tat signal sequence and an N-terminal region of the VMA *cis*-splicing intein (VMA-N) (9).

[9]It is not recommended to use excess DNA (e.g. > 2 μg) in *Bgl* II digestion. In our experience, excess DNA can result in incorrect or incomplete digestion products which will reduce the efficiency of transposon integrations.

[10]Depending on the size of the target gene (and pIT-X), the amount of transposon DNA should be adjusted. It is important to maintain the molar excess below 2.0 to avoid formation of unstable complexes containing multiple transposon insertions.

[11]When using competent cells with a transformation efficiency of 5×10^7 colonies/μg DNA, 20 μL of transposition products typically results in >80,000 colonies. Unused products can be stored at −20 °C for at least one month without any noticeable decrease of transformation yield.

12. Plate transformation cells on LB agar supplemented with 50 μg/mL kanamycin and 10 μg/mL chloramphenicol. Grow at 37 °C overnight.

13. For a gene of *L* bps, collect 9×(*L*+1,500) colonies from the transposition reaction. (*see* Note 12). Mix well the collected colonies in LB medium supplemented with 50 μg/mL kanamycin and 10 μg/mL chloramphenicol.

14. Save 5 tubes of 1 mL stock cells with an OD_{600} of >1.0 in 15% glycerol, stored at −80 °C (*see* Note 13). Extract the plasmid DNA from the remaining cells to build the pIT-X-transposon library.

3.3. Isolation of transposon insertions located in the gene of interest

1. Digest 2 μg of the transposon library DNA using *Sal* I and *BamH* I.

2. Load the digestion product on a 0.7% agarose gel and gel purify the two fragments using the QIAEX II gel extraction kit corresponding to the pIT vector backbone and the gene of interest with transposon insertions (*see* Note 14).

3. Ligate the two purified DNA fragments with T4 DNA ligase. Incubate at 16 °C for 16 hours. Inactivate the ligation product by heating at 70 °C for 10 min.

4. Transform the ligation products into electro-competent *E. coli* cells and plate cells on LB agar supplemented with 50 μg/mL kanamycin and 10 μg/mL chloramphenicol.

5. Collect 9×*L* colonies (*see* Note 15) and save cell stocks as above. Extract the plasmid DNA from the remaining cells to obtain the purified transposon library.

3.4. Creation of random triplet nucleotide deletions

1. Digest the purified transposon library DNA with *Mly* I as follows: 1 μg of DNA, 1 unit of FastDigest *Mly* I (Fermentas) (*see* Note 16) and 1×FastDigest buffer in 50 μL solution.

2. Incubate at 37 °C for 1 hour (*see* Note 17).

3. Load the digestion products on a 0.7% agarose gel and purify the larger fragment using the QIAEX II gel extraction kit. This corresponds to a linearized pIT-X

---

[12]To calculate the number of colonies required for full coverage, the number of allowed sites and site preference of transposon insertions must be considered. The number of insertion sites possible in the pIT vector backbone is 1,500 bps (2,900 of total length minus 600 bps of replication origin and 800 bps of kanamycin resistance gene). Therefore, there are (*L*+1,500) possible insertion sites for the pIT vector carrying a gene of *L* bps. Three-fold degeneracy 3×(*L*+1,500) of colonies are required to cover 95% of all the possible insertion events, assuming even distribution of insertions (12). To compensate for the site preference of transposon insertions, we assume that additional 3-fold degeneracy is sufficient to accommodate any intrinsic site preference of Mu transposase (13). Therefore, 3×3×(*L*+1,500) = 9×(*L*+1,500) colonies are sufficient to give a 95% coverage for a gene of *L* bps.

[13]If needed, the whole tube (1 mL) of stock cells should be used for inoculation to maintain the library diversity.

[14]As the transposon can insert both inside and outside of the gene of interest, a *BamH* I/*Sal* I digestion of the transposon library will result in four DNA fragments: i) vector backbone with transposon, ii) vector backbone, iii) gene of interest with transposon insertions and iv) gene of interest.

[15]9L colonies is required for maintaining library diversity because transposon insertions outside of the gene of interest have been purged.

[16]Under our experimental conditions, Fermentas FastDigest *Mly* I with FastDigest buffer cleaves DNA more precisely than NEB *Mly* I with NEBuffer 4 does, as judged by sequencing results of obtained mutants.

[17]Longer incubation time is not recommended as overdigestion with *Mly* I can remove extra bases.

library in which each DNA copy contains a random triplet nucleotide deletion in the gene of interest.

4. Load 1 or 2 μL of the purified deletion library DNA on a 0.7% agarose gel and determine the DNA concentration by comparison with DNA standard.

5. Store the purified DNA at −20 °C prior to use.

3.5. Linker ligation and reading-frame selection

1. Amplify the reading-frame selection linker by PCR as follows: 10 ng pInSALect template, 0.2 mM dNTPs, 0.5 μM of each primer LinkerFWD and LinkerFWD, 1× Phusion HF buffer (NEB) and 0.5 unit of Phusion DNA polymerase (NEB) in a total 50 μL solution.

2. Incubate the reaction with the following cycle conditions: initial denaturation at 98 °C for 30 s, 30 cycles of 98 °C for 10 s, 60 °C for 30 s and 72 °C for 1 min and final extension at 72 °C for 10 min.

3. Purify the PCR products with QIAquick PCR purification kit. Store at −20 °C prior to use (*see* Note 18).

4. Load 1 or 2 μL of the purified linker DNA on a 1% agarose gel and determine the concentration by comparison with a DNA standard.

5. Perform linker ligation in 100 μL total volume containing: 1 μg linker DNA, 500 ng of linearized pIT-X deletion library DNA, 1 μL of T4 DNA ligase and 1× T4 ligase buffer. This will produce a molar linker:library ratio of approximately 5:1. This can be adjusted appropriately given the size of pIT-X.

6. Incubate reaction at 16 °C for 16 hours. Heat-inactivate the ligation reaction at 70 °C for 10 min.

7. Pool all of the ligation product, ethanol-precipitate the DNA and resuspend in 10 μL of sterilized water.

8. Transform 1 μL of the concentrated ligation product into 50 μL electro-competent cells and recover in 500 μL SOC at 37 °C for 1 hour.

9. Plate cells on LB agar supplemented with 50 μg/mL kanamycin and 40 μg/mL ampicillin (*see* Note 19). Grow plates at 30 °C overnight (*see* Note 20).

10. Collect 9×*L* colonies and save stock cells as stated above. Extract plasmid DNA from remaining cells to obtain pIT-X-linker library. Store at −20 °C prior to use.

---

[18]The PCR products of reading-frame selection linker contain the C-terminal region of VMA *cis*-splicing intein (VMA-C) and the β-lactamase gene (BLA), flanked by a *Mly* I site at each end and a TAG codon at the 3' end (Figure 1C).

[19]When the linker DNA is inserted in-frame with the gene of interest, VMA-N from the pIT vector will be in the same reading-frame of VMA-C from the linker sequence. The VMA intein will self-splice and assemble the Tat signal peptide and BLA, which will subsequently confer cells ampicillin resistance.

[20]Incubation at a reduced temperature of 30 °C is critical for correct intein-mediated splicing (9).

3.6. Generation of in-frame TAG mutations

1. Digest the linker-containing library DNA as follows: 1 μg of DNA, 1 unit of FastDigest *Mly* I (*see* Note 16) and 1×FastDigest buffer in 50 μL solution (*see* Note 21).
2. Incubate at 37 °C for 1 hour (*see* Note 17).
3. Load the digestion product on a 0.7% agarose gel and purify the larger fragment using QIAEX II gel extraction kit. This fragment corresponds to a linearized pIT-X library in which each DNA copy contains a random TAG codon mutation in the gene of interest.
4. Determine the concentration of purified TAG mutation library by comparison with a DNA standard on an agarose gel.
5. Perform a 10 μL intra-molecular ligation reaction as follows: ~20–30 ng of linearized TAG mutation library DNA, 1 μL T4 DNA ligase and 1× T4 ligase buffer.
6. Incubate reaction at 16 °C for 16 hours. Heat-inactivate ligase at 70 °C for 10 min.
7. Transform each 1 μL of the ligation products into electro-competent cells and recover at 37 °C for 1 hour.
8. Plate cells on LB agar supplemented with 50 μg/mL kanamycin. Grow at 37 °C overnight.
9. Collect 9×*L* colonies and save stock cells as stated above. Extract plasmid DNA from remaining cells to obtain pIT-X(TAG) mutation library. Store DNA at −20 °C.
10. Select individual colonies to verify mutations by DNA sequencing.

3.7. Cloning the TAG mutation library into an expression vector and production of mutant proteins

1. PCR amplify the TAG mutation library with primers appropriate for your expression vector and purify the PCR products with QIAquick PCR purification kit.
2. Digest the purified PCR products with desired restriction enzymes and gel-purify the digestion products using QIAEX II gel extraction kit.
3. Digest the desired expression vector using the same restriction enzymes and gel-purify the digestion products using QIAEX II gel extraction kit.
4. Ligate the purified PCR products and vector DNA using T4 DNA ligase as described above.
5. Incubate at 16 °C for 16 hours. Heat-inactivate the ligase at 70 °C for 10 min.

[21]This *Mly* I digestion will release the linker sequence from the pIT-X library, leaving an in-frame TAG codon to fill the triplet nucleotide scar.

6. Transform the ligation products into electro-competent *E. coli* cells.
7. Plate cells on LB agar supplemented with the antibiotic corresponding to the selection marker in the expression vector.
8. Collect the 9×*L* colonies (*see* Note 15) and save stock cells as stated above. Extract the plasmid DNA from remaining cells (*see* Note 22).
9. Transform the collected library DNA into competent cells that are suitable for protein expression, such as those containing a pSUP plasmid for incorporation of unnatural amino acids (10) (*see* Notes 23 and [24]).
10. Express the mutant proteins in liquid culture in the presence of desired unnatural amino acid at a final concentration of 2 mM. Depending on the ultimate use of the protein(s), screens or selections can be performed on individual clonal isolates or mixed populations.

Acknowledgments

The authors thank the National Institutes of Health (GM084396) for financial support. We are also grateful to Prof. Stefan Lutz (Emory University) for the plasmid pInSALect and Prof. Peter G. Schultz for the plasmid pSUP.

References

1. Cropp TA, Schultz PG. An expanding genetic code. Trends Genet. 2004; 20:625–630. [PubMed: 15522458]
2. Wu N, et al. A genetically encoded photocaged amino acid. J Am Chem Soc. 2004; 126:14306–14307. [PubMed: 15521721]
3. Chin JW, et al. Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. Proc Natl Acad Sci U S A. 2002; 99:11020–11024. [PubMed: 12154230]
4. Wang L, et al. Addition of the keto functional group to the genetic code of Escherichia coli. Proc Natl Acad Sci U S A. 2003; 100:56–61. [PubMed: 12518054]
5. Deiters A, et al. Site-specific PEGylation of proteins containing unnatural amino acids. Bioorg Med Chem Lett. 2004; 14:5743–5745. [PubMed: 15501033]
6. Daggett KA, Layer M, Cropp TA. A general method for scanning unnatural amino acid mutagenesis. ACS Chem Biol. 2009; 4:109–113. [PubMed: 19199564]
7. Baldwin AJ, et al. Expanded molecular diversity generation during directed evolution by trinucleotide exchange (TriNEx). Nucleic Acids Res. 2008; 36:e77. [PubMed: 18559360]
8. Baldwin AJ, et al. Expanded chemical diversity sampling through whole protein evolution. Mol Biosyst. 2009; 5:764–766. [PubMed: 19562116]
9. Gerth ML, Patrick WM, Lutz S. A second-generation system for unbiased reading frame selection. Protein Eng Des Sel. 2004; 17:595–602. [PubMed: 15331775]
10. Ryu Y, Schultz PG. Efficient incorporation of unnatural amino acids into proteins in *Escherichia coli*. Nat Methods. 2006; 3:263–265. [PubMed: 16554830]

[22]In the pIT-X plasmid, the N-terminal fusion peptide (Tat signal peptide and VMA-N) usually disrupts the native function of the protein of interest. Therefore, activity assay may not be directly performed under the context of pIT vector.
[23]We have obtained the best results by creation of competent cells that already contain a plasmid for expressing a orthogonal aminoacyl-tRNA synthetase/tRNA pair. For example, one can transform BL21(DE3) with a pSUP plasmid and prepare electrocompetent cells using chloramphenicol selection. The TAG mutant library (in the expression vector) can then be transformed into these cells and selected for both plasmids. This approach results in higher protein expression yields and transformation efficiency.
[24]We have found acceptable expression results using the pSUP series of plasmids(10) that express orthogonal aminoacyl-tRNA synthetase/tRNA pairs. Recently alternative plasmid systems(14) have been described that may result in improved yields depending on the protein being expressed. The choice of plasmid does not impact the methodology described here for TAG mutant library construction.

11. Haapa S, et al. An efficient and accurate integration of mini-Mu transposons in vitro: a general methodology for functional genetic analysis and molecular biology applications. Nucleic Acids Res. 1999; 27:2777–2784. [PubMed: 10373596]

12. Patrick WM, Firth AE, Blackburn JM. User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries. Protein Eng. 2003; 16:451–457. [PubMed: 12874379]

13. Poussu E, et al. Probing the alpha-complementing domain of *E. coli* beta-galactosidase with use of an insertional pentapeptide mutagenesis strategy based on Mu in vitro DNA transposition. Proteins. 2004; 54:681–692. [PubMed: 14997564]

14. Young TS, et al. An enhanced system for unnatural amino acid mutagenesis in *E. coli*. J Mol Biol. 2010; 395:361–374. [PubMed: 19852970]

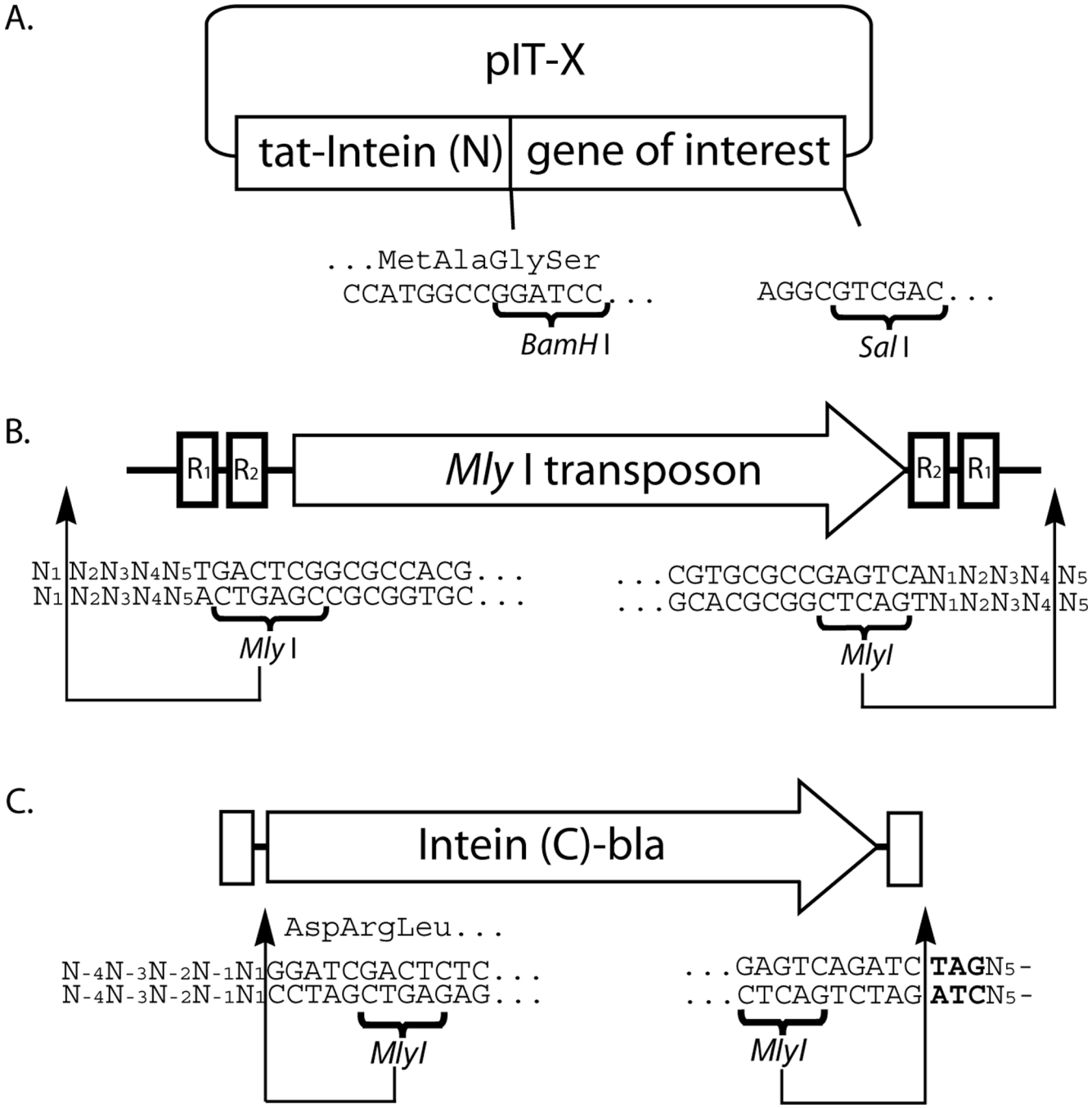

Figure 1.

Schematic diagram showing the DNA components described in this protocol with nucleotides numbered relative to original transposon insertion site. A) The cloning site of pIT-X indicating correct reading frame. Oligonucleotides used for protein of interest must match the correct reading frame. B) *Mly* I transposon showing positioning of restriction sites as it would insert into a segment of DNA. C) Reading-frame selectable linker showing restriction sites, frame, and amber stop codon scar in bold.