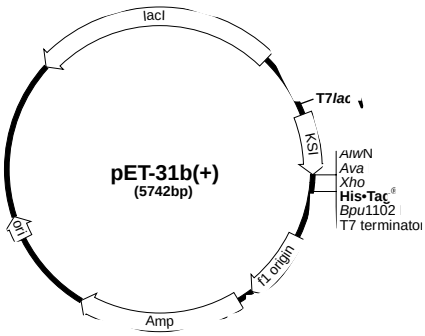
pET Peptide Expression System 31

Table of Contents

I. Description

pET Peptide Expression System 31	69953-3
pET Peptide Expression System 31 plus Competent Cells	69958-3
pET-31b(+) DNA, *Alw*N I digested, dephosphorylated	69954-3



The pET Peptide Expression System 31 is designed for the production of peptides in *E. coli*. Based on the pET-31b(+) vector, the system has several features that facilitate the expression and purification of target peptides (1). Peptide coding sequences are placed downstream of a 125aa ketosteroid isomerase (KSI) gene and upstream of a His•Tag sequence. Unique *Alw*N I and *Xho* I cloning sites are available. Use of the *Alw*N I site allows unidirectional insertion of coding sequences immediately adjacent to a methionine residue. For high-yield production, small (10-25aa) peptide coding sequences can be cloned as tandem repeats interspersed with single methionine residues. Long peptides (25-75aa) can be expressed in high yield from constructs containing single copies of coding sequences. The KSI-peptide-His•Tag fusion protein is expressed at high levels in the presence of a source of T7 RNA polymerase in the *recA*$^-$, protease-deficient BLR(DE3)pLysS host. The fusion protein is purified in one step under denaturing conditions using Novagen's His•Bind metal chelation resin and then cleaved with CNBr to release monomer peptide units each terminating with a homoserine lactone, along with insoluble KSI and the His•Tag peptide. The HS-lactone can easily be converted to an amide or homoserine, or reacted with other compounds such as analogs of fluorescein and biotin to produce C-terminal derivatized peptides. The system has been used to make biologically active peptides of 14-67aa at yields of >50mg purified peptide per liter of culture (1).

The figure on p. 2 shows the cloning strategy for insertion of tandem repeats of peptide coding regions into the *Alw*N I site of pET-31b(+). Complementary oligonucleotides corresponding to the coding region are designed such that they lack internal methionine codons and produce the three-base 3' overhangs shown in the annealed molecule. The oligonucleotides are 5' phosphorylated to allow the formation of multimeric copies upon ligation. Products corresponding to the desired number of tandem copies (usually 3-6) are isolated by agarose gel electrophoresis, and then ligated into the *Alw*N I digested, dephosphorylated vector.

The resulting construct carries several tandem copies of the target peptide coding sequence interspersed with and flanked by single methionine codons. The initial vector + insert ligation is transformed into a non-expression host such as NovaBlue to allow the establishment of stable plasmid recombinants. The plasmid is then transferred into the BLR(DE3)pLysS expression host for target protein production upon the addition of IPTG to the culture.

The KSI moiety of the fusion protein is extremely insoluble and causes the formation of large peptide-containing inclusion bodies in the host cell. The inclusion bodies are isolated by centrifugation, washed, and the fusion protein is purified by His•Bind® metal chelation chromatography under denaturing conditions. Elution is achieved by raising the imidazole concentration in the buffer. The subsequent steps vary somewhat depending on the characteristics of the target peptide. In general, the purified fusion protein is cleaved with CNBr under standard conditions, followed by differential precipitation of the extremely hydrophobic KSI protein to yield pure peptide. Two alternative protocols, suitable for purification of hydrophobic peptides and hydrophilic peptides, are provided in Section V.

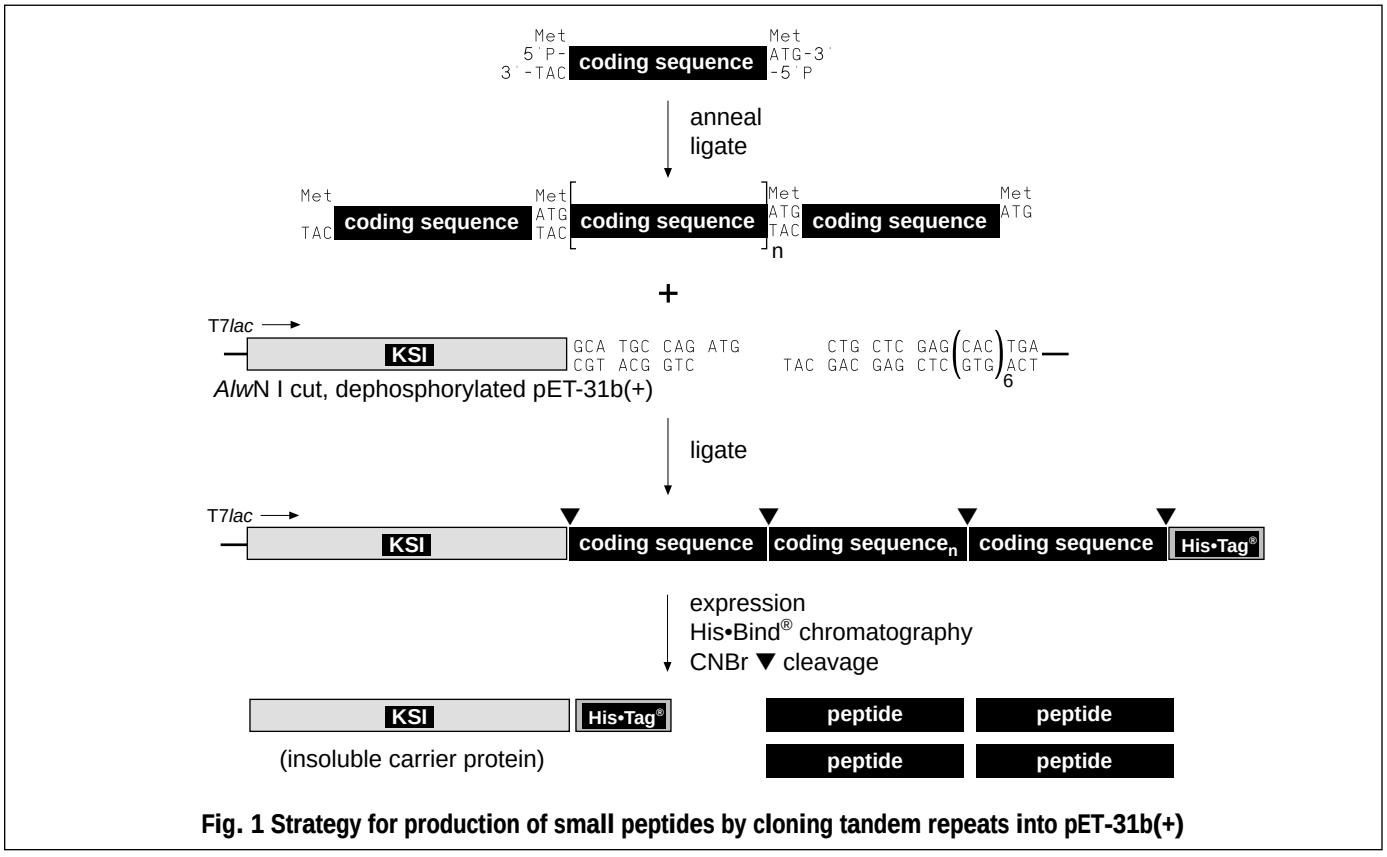


Fig. 1 Strategy for production of small peptides by cloning tandem repeats into pET-31b(+)

II. System Components

The **pET Peptide Expression System 31** is available as a choice of kits with individual components available separately. The basic system contains:

- pET-31b(+) DNA, 10µg
- Host bacterial strains, BLR and BLR(DE3)pLysS, glycerol stocks
- Induction control clone H, glycerol stock
- His•Bind® Resin, 10ml
- 4 polypropylene chromatography columns
- Vector map, pET Manual and pET Peptide Expression System 31 protocols
- Novagen Vector Diskette containing all Novagen plasmid sequences (Macintosh or DOS format)

The **pET Peptide Expression System 31 plus Competent Cells** contains two host strains ready for high-efficiency transformation. One 0.2ml aliquot each of the initial cloning host NovaBlue and the expression host BLR(DE3)pLysS, plus SOC medium, are included.

A kit containing 2µg ***Alw*N I digested dephosphorylated pET-31b(+) DNA** and a Control Insert is also available (Cat. No. 69954-1).

Other Supplemental Products

• pET-31b(+) DNA (uncut)	10µg	69952-3
• BLR(DE3)pLysS Competent Cells	0.4ml	69956-3
	1ml	69956-4
• His•Bind Resin	10ml	69670-3
	50ml	69670-4
• His•Bind Buffer Kit		69755-3

III. Construction of Target Peptide$_n$-His•Tag Genes

The procedure used to create the tandem repeat of the target peptide fused to the C-terminus of the 125 codons for the bacterial ketosteroid isomerase (KSI) gene is shown in Figure 1. The sense and anti-sense oligonucleotide strands encoding the target peptide should be designed using the preferred codon usage of *E. coli* (2). In general , the tandem repeat strategy is recommended for peptides up to 25 amino acids (75bp). Larger peptides may be produced in suitable yield from monomer constructs. Internal methionine residues must be mutated to other conserved residues in order to avoid cyanogen bromide cleavage at non-junctional methionines. In several cases, we have found that mutation of these internal methionines to isoleucine has no measurable effect on bioactivity (1). Tandem peptide construction requires that the 3' end of the coding strand include a 3-base ATG extension, and the 3' end of the non-coding strand include a 3-base TAC extension to create suitable overhangs in the duplex for unidirectional end-to-end self ligation.

The following protocols call for phosphorylation/ligation of the oligos on a relatively large scale. Sufficient quantities (25µg, or about 1 A_{260} unit) of highly pure oligonucleotides help to ensure that the ligations proceed efficiently and that sufficient material will be recovered from gel fractionation. HPLC or gel purified oligonucleotides are suitable. A procedure for gel purification is provided in Appendix A. Note that it is also necessary to use 5' phosphorylated oligonucleotides to allow tandem ligation, and for ligation into the *Alw*N I digested dephosphorylated pET-31b(+) vector available from Novagen.

Phosphorylation

1. If the purified oligos have already been phosphorylated proceed to the annealing reaction below. If they are not phosphorylated assemble two reactions, each containing one oligo strand, as follows. We recommend Novagen's Polynucleotide Kinase (Cat. No. 69248-1, which includes 10X buffer and 10mM ATP).

   25µg (~1500-2000pmol) purified oligo in TE buffer (10mM Tris-HCl pH 8.0, 1mM EDTA)

10µl 10X Kinase buffer
10µl 10mM ATP
125U T4 polynucleotide kinase
xµl Nuclease-free water
100µl Total volume

2. Add the enzyme last, stir gently to mix and incubate at 37°C for 30 minutes.

Annealing

1. Combine the two phosphorylated oligos into one tube (total volume 200µl). Annealing can be performed in the phosphorylation reaction buffer. If pre-phosphorylated oligos are used, dissolve them in TE buffer or water at ~100pmol/µl and add equimolar amounts (1500-2000pmol) of each to one tube. Add 20µl of 10X annealing buffer (10X = 400mM Tris-HCl pH 8.0, 100mM $MgCl_2$, 500mM NaCl) and water to a total volume of 200µl.
2. Heat the combined oligos at 99°C for 10 minutes, then cool to 30°C over a 15 minute period. This procedure is conveniently performed in a thermal cycler.
3. Extract the reaction once with an equal volume of TE-buffered phenol:CIAA (1:1; CIAA is 24 parts chloroform and 1 part isoamyl alcohol). Extract the aqueous phase with 1 volume of CIAA. Transfer the final aqueous phase to a fresh tube. Add $MgCl_2$ to 10mM, 0.1 volume 3M Na acetate pH ~6 and 2 volumes of 100% ethanol. Vortex and place at -20°C for > 1 hour. Centrifuge at 12,000 × g for 5 minutes, drain, rinse the pellet with 100% ethanol and air dry. Resuspend the annealed oligos in TE buffer at 100pmol/µl. Remember that 1pmol of each strand when combined equals 1pmol annealed oligo!

Ligation and purification of tandem repeats

1. Assemble the ligation reaction as follows (qualified ligation components are included in the DNA Ligation Kit, Cat. No. 69838-3):

   2.5µl phosphorylated, annealed oligos (250pmol)
   2.5µl 10X Ligase buffer (200mM Tris-HCl pH 7.6, 50mM $MgCl_2$)
   2.5µl 100mM DTT
   0.5µl 10mM ATP
   0.5µl T4 DNA ligase (2 Weiss units)
   16.5µl Nuclease-free water
   25µl Total volume

2. Add the enzyme last and mix gently by stirring with the pipet tip. Incubate the reaction at 16°C for 2 hours. The extent of ligation can be checked by gel analysis of 2µl (~0.5µg) sample of the reaction.
3. Add concentrated gel loading buffer to 1X. Load the entire sample in a 1cm × 1.5mm well of a 3% low melting point agarose gel containing 0.5µg/ml ethidium bromide. Run appropriate size markers (e.g. Novagen's PCR Markers, Cat. No. 69278-1) in an adjacent lane. Run the gel at 70V for 3 hr at 4°C (voltage and time given for standard 10cm long minigel apparatus).
4. Visualize the DNA bands under UV illumination (~300-360nm). Cut out bands containing the desired multimer species (several repeat sizes can be tested) and recover the DNA by standard melting/phenol extraction procedures (e.g. Sambrook et al. (1989) in "Molecular Cloning, A Laboratory Manual", 2nd edition, pp. 6.30-6.31).
5. Resuspend the final purified DNA in ~10µl TE buffer and determine its concentration spectrophotometrically or by gel analysis of a sample vs. known amounts of marker DNA. The DNA should now be ready for ligation with the pET-31b(+) vector. Calculate the molar concentration of the DNA as follows:

   A_{260} × dilution factor × 50µg/ml = x µg/ml
   x µg/ml ÷ (660µg/µmol/bp × #bp) = y µmol/ml
   y µmol/ml × 1/1000 = z pmol/µl

Note: If difficulties are encountered in producing the desired tandem species (as monitored in Step 2) or in recovering sufficient material for cloning, use more annealed oligos (up to 50µg, or 2000pmol) and more ligase (15U) in the 25µl

ligation reaction. In addition, increase the incubation time to 16 hours and reduce the temperature to 8°C. The extent of ligation may be further increased by adding more ligase (10U) and ATP (2µl 10mM) at this point and allowing the reaction to proceed for another 24 hr at 8°C.

Ligation with *Alw*N I digested dephosphorylated pET-31b(+)

The pET-31b(+) vector is prepared by *Alw*N I digestion and dephosphorylation. Qualified vector ready for ligation of tandem inserts and a positive control insert are available as a kit (Cat. No. 69954-1). Each different size tandem DNA multimer is ligated with the vector in separate reactions to create an array of pET-31(target peptide)$_n$His•Tag® constructs.

1. Assemble the following ligation reaction:

   2µl (0.026pmol) *Alw*N I digested, dephosphorylated pET-31b(+)
   1µl 10X Ligase buffer (200mM Tris-HCl pH 7.6, 50mM $MgCl_2$)
   1µl 100mM DTT
   0.5µl 10mM ATP
   xµl (0.13-0.26pmol) purified tandem DNA insert
   0.5µl (2 Weiss units) T4 DNA ligase
   yµl Nuclease-free water
   10µl Total volume

   In addition, set up a negative control reaction containing all of the same components except for the target insert (add an equivalent amount of TE buffer instead). This reaction will allow the establishment of vector background. A positive control reaction may also be set up, which should contain a known amount of target DNA that has been successfully used for cloning (see below for information on the Positive Control insert supplied with the *Alw*N I digested dephosphoylated vector).

2. Add the enzyme last and mix by stirring with the pipet tip. Incubate at 16°C overnight.
3. Transform competent NovaBlue cells and screen colonies for inserts.

Positive Control

A Positive Control insert is provided with the *Alw*N I digested, dephosphorylated vector to verify the cloning steps. The insert encodes a 15aa peptide corresponding to the S•Tag™ peptide. It is provided as a 5' phosphorylated monomer with appropriate overhangs for *Alw*N I site ligation. Use 2µl (0.2pmol) of the insert for ligation. Note that this insert is provided for testing purposes only and may not be used for peptide production or distributed outside your laboratory.

Protocols for transformation and colony screening are found in the accompanying pET System Manual. Screen potential recombinants by colony PCR with the KSI primer (Cat. No. 69947-3) and T7 terminator primer (Cat. No. 69337-3) combination. The presence of inserts and number of target peptide encoding units can be determined by agarose gel electrophoresis of the PCR products. The vector without an insert produces a 158bp amplification product with the indicated primers. Candidate recombinants should also be sequenced to verify the construction. The appropriate constructs are then transformed into BLR(DE3)pLysS for protein overproduction.

The ligations containing the target oligo and Positive Control insert should produce many more colonies than the negative control ligation. If the target insert ligation produces nearly the same number of colonies as the negative control, cloning efficiency is low and screening will likely reveal a large percentage of clones containing only the religated vector. Instead of preparing more insert, it may be possible to obtain the desired clone by digesting the remainder of the ligation with *Alw*N I prior to transformation. This treatment will reduce the background of religated vector and should therefore increase the likelihood of finding the desired recombinant.

IV. Expression of KSI-Target Peptide$_n$His•Tag® Fusion

A detailed discussion of induction strategies, culture maintenance and storage is found in the pET System Manual. The protease deficient strain BLR(DE3)pLysS should be used for protein production. The BLR strain is a *recA*$^-$ derivative of BL21 and is potentially superior for stabilizing tandem repeats.

1. The day before induction, inoculate 25-50ml of LB media supplemented with 50µg/ml carbenicillin (Carb) or ampicillin and 34µg/ml chloramphenicol (Cam) in a 250ml flask with a single colony of the recombinant (grown on LB agar plus antibiotics). Shake at 250rpm at 37°C overnight.
2. The next morning, dilute the starter culture 40-fold into 500ml of fresh LB/Carb/Cam in a 2.8 liter Fernbach flask prewarmed to 37°C, using as many flasks as needed to suit the scale of the production.
3. Shake at 250rpm at 37°C until the OD_{600} reaches 0.3-0.5. Induce the production of the target protein by adding 1mM IPTG. Harvest the cells (by centrifugation) at an OD ~ 2.0, which is generally reached 3-6 hr after induction. The drained cell pellet (~2.5g/l) can be stored at -20°C for at least one week without any loss in yield of fusion protein.
4. Resuspend the cell pellet in 0.1X the culture volume of 1X Binding Buffer (5mM imidazole, 40mM Tris-HCl pH 7.9, 500mM NaCl), sonicate on ice until no longer viscous, and centrifuge at 12,000 × g for 10 min at 4°C. If desired, remove a small sample prior to centrifugation for gel analysis of total cell protein.
5. Carefully remove the supernatant (soluble) fraction to a fresh tube. If desired, the pellet (insoluble) fraction can be washed by brief sonication in 1X Binding Buffer, recentrifuged as above, and drained. Resuspend the pellet in 0.1 culture volume of 1X Binding Buffer containing 6M guanidine-HCl. After the pellet has dissolved completely, centrifuge again at 12,000 × g for 10 min at 4°C to remove any residual particulate matter.
6. Since the KSI fusion partner is very insoluble, it is likely that the target fusion protein will be found in the insoluble fraction, which is the pellet from the centrifugation in Step 4. The supernatant, pellet and pre-centrifugation sample (if saved) can be quickly analyzed for the target protein by SDS-gel electrophoresis and staining with Coomassie blue. The following formula can be applied to determine an appropriate amount of sample to load on a standard mini-gel apparatus. Note that the presence of salt and guanidine may affect the mobility of the target protein; if this is crucial follow the protocol in the pET Manual for processing a portion of the culture in a low-salt buffer.

   (30/concentration factor) × (3/OD_{600} at harvest) = x

   Add xµl of sample to an equal volume of 2X SDS sample buffer.

   Use the same amount of soluble, insoluble and total samples described above.

V. Purification by His•Bind® Chromatography and CNBr Cleavage

The following protocols are based on processing 200ml of *E. coli* culture. Scale all steps in proportion to the amount of culture used. All steps are performed at room temperature unless otherwise noted. **Use appropriate safety precautions when handling CNBr and other hazardous chemicals during this procedure.**

Protocol 1: Production and purification of hydrophobic peptides

The insoluble cellular protein pellet will generally contain the majority of the fusion protein as dense inclusion bodies. If gel analysis of the soluble and insoluble fractions shows that the target protein is in the redissolved inclusion bodies, His•Bind chromatography is carried out under denaturing conditions.

1. Load the resolubilized inclusion bodies from Step 5 above onto a 2.5ml His•Bind column which has been charged with $NiSO_4$ and equilibrated with 1X Binding Buffer (1X = 20mM Tris-HCl pH 7.9, 0.5M NaCl, 5mM imidazole) plus 6M guanidine-HCl, at a flow rate of 0.25ml/min.

2. Wash the column with 15ml of 1X Binding Buffer plus 6M guanidine-HCl and then with 20ml of the same buffer containing 16mM imidazole instead of 5mM. An 8X stock of this solution (without guanidine-HCl) can be prepared by adding 4 parts 8X Binding Buffer plus 1 part 8X Wash Buffer supplied in the His•Bind Buffer Kit. Collect and save the flow-through and wash fractions.
3. Elute the fusion protein with 15ml of 300mM imidazole in the same buffer. (A 4X stock of this elution buffer without guanidine-HCl can be prepared from His•Bind Buffer kit components as follows: first dilute the kit's 8X Binding Buffer with an equal volume of water to bring it to 4X, then add 2.35 parts of 4X Binding Buffer to 1 part of the kit's 4X Elute Buffer.) Collect 2ml fractions. Analyze samples of each fraction by SDS-polyacrylamide gel electrophoresis. Combine the peak fractions (ca. 10ml), dialyze overnight at 4°C against 2×1 liter of H_2O in 12-14 kDa cut-off dialysis bags. The majority of the protein should form a white precipitate which can be pelleted by centrifugation at $2000 \times g$ for 10 min at 4°C.
4. Working in a fume hood, dissolve the dense white pellet in 6ml of 80% formic acid. Transfer to a 50ml round-bottomed flask and add 0.2g CNBr. Bubble in nitrogen, wrap the flask in aluminum foil and stir for 18-22 hr. Attach to a rotary evaporator and evaporate to dryness at 28°C. Resuspend the resultant transparent proteinaceous gel in a minimal volume of 40% CH_3CN/60% H_2O/0.1 % TFA (ca. 2-4ml) and stir for 1 hr. Centrifuge this suspension at $12,000 \times g$ for 10 min at 4°C. The supernatant can be further clarified by passage through a 0.22µm filter and the purity of the recombinant target peptide analyzed by reverse-phase HPLC and UV spectroscopy.

Protocol 2: Production and purification of hydrophilic peptides

The purification and CNBr cleavage of the target peptide from the carrier fusion protein is the same as in Protocol 1 until the CNBr cleavage step. Here the two protocols diverge due to the differing solubility between the hydrophilic and the hydrophobic target peptides. After the CNBr cleaved peptide mixture is dried by rotary evaporation, the gelatinous material is resuspended in 20mM KH2PO4, 100mM NaCl, and the pH adjusted to 7.4 with 1M $NaHCO_3$. This mixture is stirred overnight under N_2 and wrapped in aluminum foil. The suspension is centrifuged at $5000 \times g$ for 15 min at 4°C and the peptide in the supernatant analyzed by UV spectroscopy. The hydrophilic peptide in the supernatant can be either concentrated by rotary evaporation or dried by lyophilization.

VI. Preparation of Peptide Conjugates and Affinity Supports

CNBr cleavage of the tandem peptides at the junctional methionines will generate a C-terminal homoserine lactone. In most cases, the bioactivity should be unaffected by the presence of this extra amino acid. One may wish to take advantage of this uniquely reactive residue and couple a wide variety of primary amine-containing compounds such as biotin, fluorescein, and other fluorescent reagents to the C-terminus of the target peptide, or prepare C-terminally coupled peptide affinity supports.

Preparation of peptide-homoserine conjugates

In the simplest case, the lactone ring can be completely and immediately converted to the homoserine (HS) free acid form by dropwise addition of 0.2N NaOH, until a final pH of 12.5–13 is reached and then neutralized with phosphate-buffered saline, pH 7.4. For each of the following reactions, 0.3-0.8mg of peptide-lactone are placed in 1.7ml polypropylene tubes and completely lactonized by addition of 20µl of 100% trifluoroacetic acid (TFA) and immediately dried in a Speed Vac (Savant). The dried pellets are dissolved in 50µl anhydrous DMF delivered by a gas-tight syringe and 8µl of Et_3N are added. All reaction tubes are equipped with a micro-stir bar.

Synthesis of peptide-HS-amide: 50µl of anhydrous DMF saturated with NH_3 gas are added directly to the dried lactonized pellet prior to the addition of Et_3N. This reaction is stirred at 44°C and amidation is complete in 30 min.

Synthesis of peptide-HS-biotin: 100 equivalents of biotin ethylenediamine (Molecular Probes) are added to the lactonized peptide solution and the reaction allowed to proceed for 5 hr at 48°C.

Synthesis of peptide-HS-fluorescein: 28 equivalents of fluoresceinyl glycine amide (Molecular Probes) are dissolved in an additional 100µl DMF and added to the lactonized peptide solution and the reaction is allowed to proceed for 8 hr at 48°C.

Synthesis of peptide-HS-dansyl: 37 equivalents of dansyl ethylenediamine (Molecular Probes) are added to the lactonized peptide solution and the reaction is allowed to proceed for 7 hr at 48°C.

Synthesis of peptide-HS-ANS: 100 equivalents of 5((2-aminoethyl)amino)naphthalene-1-sulfonic acid (Molecular Probes) are dissolved in an additional 200µl DMF plus an additional 32µl Et_3N and added to the lactonized peptide solution and the reaction is allowed to proceed for 6 hr at 45°C.

Preparation of C-terminally coupled peptide-affinity supports

Four mg of lyophilized peptide-HS-lactone are placed into a 1.7ml polypropylene tube and solubilized with 765µl 0.88 M NaOAc, and the pH is adjusted to 6 with NaOH. 0.5ml of aminoalkyl Sepharose (Affigel 102, Bio Rad) is added and the mixture is gently rocked for 170 hr at 37°C. The extent of the coupling is monitored by removing small aliquots at various time points and determining the amount of peptide remaining in the supernatant by the Bradford Method.

VII. References

1. Kuliopulos, A. and Walsh, C.T. (1994) *J. Am. Chem. Soc.* **116**, 4599-4607.

2, Boer, H. A. d., and Kastelein, R. A. (W. Reznikoff and L. Gold, Ed.) in "Maximizing Gene Expression", Butterworths: Stoneham, MA, 1986; pp 225-285.

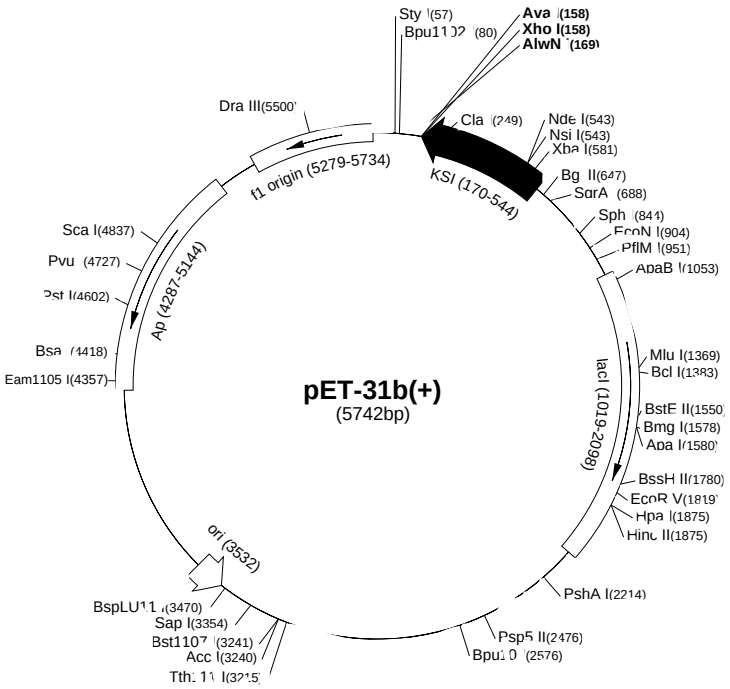
Appendix A: Polyacrylamide Gel Purification of Oligonucleotides

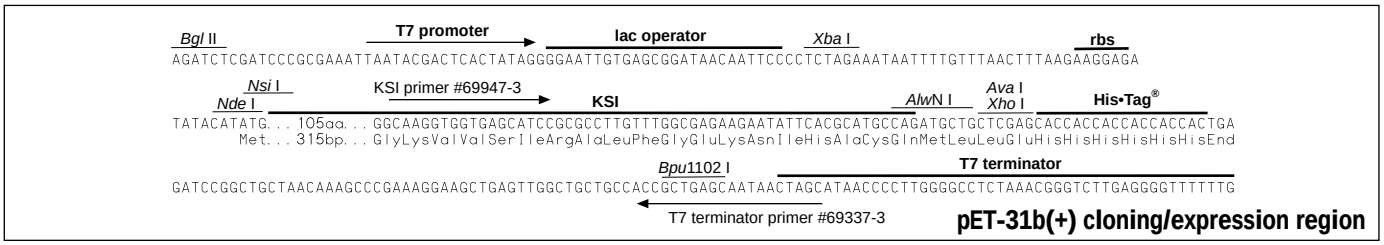
1. Synthesize oligonucleotide, determine concentration by reading the absorbance at 260nm. Use 20µg/ml = 1 A_{260}
2. Pour 12% polyacrylamide minigel:

   4.2g urea
   1.0ml 10X TBE (Tris-Borate-EDTA)
   1.8ml water
   4.0ml 30% acrylamide/0.8% bisacrylamide Protogel (National Diagnostics) solution
   100µl 10% ammonium persulfate
   5µl TEMED
   10ml total volume

   Degas gel, pour 0.75 mm gel, overlay with tert-amyl alcohol, let gel polymerize.
3. Rinse away urea prior to loading oligo. Load 200µg of oligo in 1X DNA loading buffer with bromophenol blue dye, run at 200 volts × ca. 30 min in minigel apparatus. Take apart gel, place on Saran wrap.
4. Using hand-held short-wave UV lamp with fluorescent-indicator TLC plate, excise dark band of interest (UV absorbing) with a clean razor blade. Using forceps, place gel slice into 15ml Falcon tube and mash-up gel slice with a glass rod. Add 10ml 100mM NH_4Ac, 1 mM EDTA, pH 6-6.5 and rock overnight at room temperature.
5. Pellet gel at 3000 × g 1 min. Transfer supernatant to fresh Falcon tube.
6. Pre-equilibrate Waters C-18 Sep-Pak with 10ml MeOH, then 20ml H_2O. Slowly load 10ml of oligo-containing supernatant, avoid pushing bubbles through Sep-Pak at all times. Wash with 10ml H_2O. Elute very slowly with 2 × 1ml 50% MeOH/50% H_2O into two 1.7ml polypropylene tubes labeled E1 and E2.
7. Take absorbance readingsat 260nm of 50% MeOH/50% H_2O eluates E1 and E2, blanked against same solvent. Use 30µg/ml = 1 A_{260}
8. Dry in Speed Vac.

Appendix B: pET-31b(+) Maps and Sequence





pET-31b(+) cloning/expression region

Unique sites are shown. Note that the sequence is numbered by the pBR322 convention, so the T7 expression region is reversed (*i.e.* the sequence below shows the reverse complement of the expressed strand; the above diagram shows the expressed strand). The f1 origin in pET-31b(+) is oriented such that the single stranded DNA produced upon helper phage infection will anneal with the T7 terminator primer.

pET-31b(+) Sequence

```
   1  ATCCGGATAT AGTTCCTCCT TTCAGCAAAA AACCCCTCAA GACCCGTTTA
  51  GAGGCCCCAA GGGGTTATGC TAGTTATTGC TCAGCGGTGG CAGCAGCCAA
 101  CTCAGCTTCC TTTCGGGCTT TGTTAGCAGC CGGATCTCAG TGGTGGTGGT
 151  GGTGGTGCTC GAGCAGCATC TGGCATGCGT GAATATTCTT CTCGCCAAAC
 201  AAGGCGCGGA TGCTCACCAC CTTGCCGGCG CCATTGAAGC GAAAGTGATC
 251  GATGGGCGCA ACTACGGTCT TGCGGCCCTG ATACTCGAAG CTGACGGTGA
 301  AAGCGAAGGC CGCTTCGTTG GCGACCGCGC GTACCTCCTG CGTCAGCTCC
 351  ACCGCCAAAG GCAGTTTGAG CGAGTTGGCG TAAAACTCAC GAATCGCAGC
 401  CGTACCGGAC CTGGGCTCGG AACCCACGGG GTCTTCCACC GTGGCGTCAT
 451  CGGCAAACAG CGCGACGATG CCGTCCAGAT CGCCGGCATT GAGCGCAGCC
 501  ACAAAGCGCT GTACCACGGC GGTGATGTGT TCTGGGGTAT GCATATGTAT
 551  ATCTCCTTCT TAAAGTTAAA CAAAATTATT TCTAGAGGGG AATTGTTATC
 601  CGCTCACAAT TCCCCTATAG TGAGTCGTAT TAATTTCGCG GGATCGAGAT
 651  CTCGATCCTC TACGCCGGAC GCATCGTGGC CGGCATCACC GGCGCCACAG
 701  GTGCGGTTGC TGGCGCCTAT ATCGCCGACA TCACCGATGG GGAAGATCGG
 751  GCTCGCCACT TCGGGCTCAT GAGCGCTTGT TTCGGCGTGG GTATGGTGGC
```

```
 801  AGGCCCCGTG GCCGGGGGAC TGTTGGGCGC CATCTCCTTG CATGCACCAT
 851  TCCTTGCGGC GGCGGTGCTC AACGGCCTCA ACCTACTACT GGGCTGCTTC
 901  CTAATGCAGG AGTCGCATAA GGGAGAGCGT CGAGATCCCG GACACCATCG
 951  AATGGCGCAA AACCTTTCGC GGTATGGCAT GATAGCGCCC GGAAGAGAGT
1001  CAATTCAGGG TGGTGAATGT GAAACCAGTA ACGTTATACG ATGTCGCAGA
1051  GTATGCCGGT GTCTCTTATC AGACCGTTTC CCGCGTGGTG AACCAGGCCA
1101  GCCACGTTTC TGCGAAAACG CGGGAAAAAG TGGAAGCGGC GATGGCGGAG
1151  CTGAATTACA TTCCCAACCG CGTGGCACAA CAACTGGCGG GCAAACAGTC
1201  GTTGCTGATT GGCGTTGCCA CCTCCAGTCT GGCCCTGCAC GCGCCGTCGC
1251  AAATTGTCGC GGCGATTAAA TCTCGCGCCG ATCAACTGGG TGCCAGCGTG
1301  GTGGTGTCGA TGGTAGAACG AAGCGGCGTC GAAGCCTGTA AAGCGGCGGT
1351  GCACAATCTT CTCGCGCAAC GCGTCAGTGG GCTGATCATT AACTATCCGC
1401  TGGATGACCA GGATGCCATT GCTGTGGAAG CTGCCTGCAC TAATGTTCCG
1451  GCGTTATTTC TTGATGTCTC TGACCAGACA CCCATCAACA GTATTATTTT
1501  CTCCCATGAA GACGGTACGC GACTGGGCGT GGAGCATCTG GTCGCATTGG
1551  GTCACCAGCA AATCGCGCTG TTAGCGGGCC CATTAAGTTC TGTCTCGGCG
1601  CGTCTGCGTC TGGCTGGCTG GCATAAATAT CTCACTCGCA ATCAAATTCA
1651  GCCGATAGCG GAACGGGAAG GCGACTGGAG TGCCATGTCC GGTTTTCAAC
1701  AAACCATGCA AATGCTGAAT GAGGGCATCG TTCCCACTGC GATGCTGGTT
1751  GCCAACGATC AGATGGCGCT GGGCGCAATG CGCGCCATTA CCGAGTCCGG
1801  GCTGCGCGTT GGTGCGGATA TCTCGGTAGT GGGATACGAC GATACCGAAG
1851  ACAGCTCATG TTATATCCCG CCGTTAACCA CCATCAAACA GGATTTTCGC
1901  CTGCTGGGGC AAACCAGCGT GGACCGCTTG CTGCAACTCT CTCAGGGCCA
1951  GGCGGTGAAG GGCAATCAGC TGTTGCCCGT CTCACTGGTG AAAAGAAAAA
2001  CCACCCTGGC GCCCAATACG CAAACCGCCT CTCCCCGCGC GTTGGCCGAT
2051  TCATTAATGC AGCTGGCACG ACAGGTTTCC CGACTGGAAA GCGGGCAGTG
2101  AGCGCAACGC AATTAATGTA AGTTAGCTCA CTCATTAGGC ACCGGGATCT
2151  CGACCGATGC CCTTGAGAGC CTTCAACCCA GTCAGCTCCT TCCGGTGGGC
2201  GCGGGGCATG ACTATCGTCG CCGCACTTAT GACTGTCTTC TTTATCATGC
2251  AACTCGTAGG ACAGGTGCCG GCAGCGCTCT GGGTCATTTT CGGCGAGGAC
2301  CGCTTTCGCT GGAGCGCGAC GATGATCGGC CTGTCGCTTG CGGTATTCGG
2351  AATCTTGCAC GCCCTCGCTC AAGCCTTCGT CACTGGTCCC GCCACCAAAC
2401  GTTTCGGCGA GAAGCAGGCC ATTATCGCCG GCATGGCGGC CCCACGGGTG
2451  CGCATGATCG TGCTCCTGTC GTTGAGGACC CGGCTAGGCT GGCGGGGTTG
2501  CCTTACTGGT TAGCAGAATG AATCACCGAT ACGCGAGCGA ACGTGAAGCG
2551  ACTGCTGCTG CAAAACGTCT GCGACCTGAG CAACAACATG AATGGTCTTC
2601  GGTTTCCGTG TTTCGTAAAG TCTGGAAACG CGGAAGTCAG CGCCCTGCAC
2651  CATTATGTTC CGGATCTGCA TCGCAGGATG CTGCTGGCTA CCCTGTGGAA
2701  CACCTACATC TGTATTAACG AAGCGCTGGC ATTGACCCTG AGTGATTTTT
2751  CTCTGGTCCC GCCGCATCCA TACCGCCAGT TGTTTACCCT CACAACGTTC
2801  CAGTAACCGG GCATGTTCAT CATCAGTAAC CCGTATCGTG AGCATCCTCT
2851  CTCGTTTCAT CGGTATCATT ACCCCCATGA ACAGAAATCC CCCTTACACG
2901  GAGGCATCAG TGACCAAACA GGAAAAAACC GCCCTTAACA TGGCCCGCTT
2951  TATCAGAAGC CAGACATTAA CGCTTCTGGA GAAACTCAAC GAGCTGGACG
3001  CGGATGAACA GGCAGACATC TGTGAATCGC TTCACGACCA CGCTGATGAG
3051  CTTTACCGCA GCTGCCTCGC GCGTTTCGGT GATGACGGTG AAAACCTCTG
3101  ACACATGCAG CTCCCGGAGA CGGTCACAGC TTGTCTGTAA GCGGATGCCG
3151  GGAGCAGACA AGCCCGTCAG GGCGCGTCAG CGGGTGTTGG CGGGTGTCGG
3201  GGCGCAGCCA TGACCCAGTC ACGTAGCGAT AGCGGAGTGT ATACTGGCTT
```

```
3251  AACTATGCGG CATCAGAGCA GATTGTACTG AGAGTGCACC ATATATGCGG
3301  TGTGAAATAC CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCTC
3351  TTCCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGGTCGTT CGGCTGCGGC
3401  GAGCGGTATC AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA
3451  GGGGATAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG
3501  GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC
3551  CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG
3601  ACAGGACTAT AAAGATACCA GGCGTTTCCC CCTGGAAGCT CCCTCGTGCG
3651  CTCTCCTGTT CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC
3701  CTTCGGGAAG CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTATCTCAGT
3751  TCGGTGTAGG TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT
3801  TCAGCCCGAC CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC
3851  CGGTAAGACA CGACTTATCG CCACTGGCAG CTGGTAACAG GATTAGCAGA
3901  GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA
3951  CGGCTACACT AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG
4001  TTACCTTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC
4051  GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG CAGCAGATTA CGCGCAGAAA
4101  AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC
4151  AGTGGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG ATTATCAAAA
4201  AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT
4251  CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA
4301  GTGAGGCACC TATCTCAGCG ATCTGTCTAT TTCGTTCATC CATAGTTGCC
4351  TGACTCCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT TACCATCTGG
4401  CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCTCCAGATT
4451  TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCCT
4501  GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAGCTAG
4551  AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTG
4601  CAGGCATCGT GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC
4651  GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCCATGT TGTGCAAAAA
4701  AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT AAGTTGGCCG
4751  CAGTGTTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC TCTTACTGTC
4801  ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC
4851  ATTCTGAGAA TAGTGTATGC GGCGACCGAG TTGCTCTTGC CCGGCGTCAA
4901  TACGGGATAA TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT
4951  GGAAAACGTT CTTCGGGGCG AAAACTCTCA AGGATCTTAC CGCTGTTGAG
5001  ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT
5051  TTACTTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC
5101  GCAAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT
5151  CCTTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG
5201  GATACATATT TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCGCGC
5251  ACATTTCCCC GAAAAGTGCC ACCTGAAATT GTAAACGTTA ATATTTTGTT
5301  AAAATTCGCG TTAAATTTTT GTTAAATCAG CTCATTTTTT AACCAATAGG
5351  CCGAAATCGG CAAAATCCCT TATAAATCAA AAGAATAGAC CGAGATAGGG
5401  TTGAGTGTTG TTCCAGTTTG GAACAAGAGT CCACTATTAA AGAACGTGGA
5451  CTCCAACGTC AAAGGGCGAA AAACCGTCTA TCAGGGCGAT GGCCCACTAC
5501  GTGAACCATC ACCCTAATCA AGTTTTTTGG GGTCGAGGTG CCGTAAAGCA
5551  CTAAATCGGA ACCCTAAAGG GAGCCCCCGA TTTAGAGCTT GACGGGGAAA
5601  GCCGGCGAAC GTGGCGAGAA AGGAAGGGAA GAAAGCGAAA GGAGCGGGCG
```

```
5651  CTAGGGCGCT GGCAAGTGTA GCGGTCACGC TGCGCGTAAC CACCACACCC
5701  GCCGCGCTTA ATGCGCCGCT ACAGGGCGCG TCCCATTCGC CA
```

pET-31b(+) sequence landmarks

T7 promoter*	616-632
T7 transcription start*	615
lacI coding sequence	1019-2098
KSI coding sequence*	170-544
*Alw*N I site	169
His•Tag® coding seq.*	140-157
T7 terminator*	26-72
pBR322 origin	3532
bla coding sequence*	4287-5144
f1 origin*	5279-5734

* indicates sequence shown is reverse complement

pET-31b(+) Restriction Map

Enzyme	Sequence	# Sites	Locations							
AccI	GT mk_AC	1	3240							
AciI	C CGC	84								
AflIII	A CryG_T	2	1369	3470						
AluI	AG CT	26								
AlwI	GGATCnnnn n_	14								
Alw21I	G_wGCw C	8	159	869	1353	2464	3288	3788	4943	5028
AlwNI	CAG_nnn CTG	1	169							
ApaI	G_GGCC C	1	1580							
ApaBI	GCA_nnnnn TGC	1	1053							
ApaLI	G TGCA_C	4	1349	3284	3784	5024				
ApoI	r AATT_y	3	1644	5302	5313					
AvaI	C yCGr_G	1	158							
AvaII	G GwC_C	8	407	1921	2297	2385	2476	2755	4495	4717
BanI	G GyrC_C	10								
BanII	G_rGCy C	5	417	753	767	1580	5575			
BbsI	GAAGACnn nnnn_	5	424	1515	1854	2228	2588			
BbvI	GCAGC(n)8 (n)4n_	28								
BccI	CCATC	14								
Bce83I	CTTGAG(n)14_nn	7	21	2183	2353	3561	3859	4094	4962	
BcefI	ACGGC(n)12 n_	8	385	456	532	888	1229	1856	3966	5526
BcgI	CGA(n)6TGC(n)10_nn	8	1661	1695	2195	2229	3047	3081	4862	4896
BclI	T GATC_A	1	1383							
BfaI	C TA_G	7	70	582	2484	3959	4212	4547	5651	
BglI	GCCn_nnn nGGC	2	2433	4477						
BglII	A GATC_T	1	647							
BmgI	GkGCCC	1	1578							
BpmI	CTGGAGn(14)_nn	5	1207	1696	2330	2997	4427			
Bpu10I	CC TnA_GC	1	2576							
Bpu1102I	GC TnA_GC	1	80							
BsaI	GGTCTCn nnnn_	1	4418							
BsaAI	yAC GTr	2	3222	5500						
BsaBI	GATnn nnATC	3	646	652	2667					
BsaHI	Gr CG_yC	8	228	444	692	713	827	1326	2009	4894
BsaJI	C CnnG_G	10								
BsaWI	w CCGG_w	8	2	404	1688	2191	2659	3676	3823	4648
BsaXI	ACnnnnnCTCC	2	2028	5448						
BsbI	CAACAC	2	3186	5407						
BscGI	CCCGT	14								
BsiI	C TCGT_G	2	3643	5021						
BsiEI	CG_ry CG	6	325	2154	3386	3810	4727	4876		
BslI	CCnn_nnn nnGG	22								
BsmAI	GTCTCn nnnn_	7	1066	1471	1597	1984	3111	4418	5194	
BsmBI	CGTCTCn nnnn_	2	1984	3111						
BsmFI	GGGACn(10) nnnn_	4	830	2371	2741	5715				
BsoFI	GC n_GC	48								
Bsp24I	GACn(6)TGGn(10)_nn	10								
Bsp1286I	G_dGCh C	13								
BspEI	T CCGG_A	2	2	2659						
BspGI	CTGGAC	2	475	2996						
BspLU11I	A CATG_T	1	3470							
BsrI	ACTG_Gn	24								
BsrBI	GAG CGG	4	602	3403	5198	5644				
BsrDI	GCAATG_nn	4	1416	1782	4418	4592				
BsrFI	r CCGG_y	9	224	482	679	688	1055	2267	2427	4437	5601
BssHII	G CGCG_C	1	1780							
Bst1107I	GTA TAC	1	3241							
BstEII	G GTnAC_C	1	1550							
BstXI	CCAn_nnnn nTGG	3	1171	1300	1423					
BstYI	r GATC_y	11								
Cac8I	GCn nGC	40								

Enzyme	Sequence	# Sites	Locations								
CjeI	ACn(6)TGGn(9) n(6)_	28									
CjePI	CCAn(7)TCn(7) n(6)_	20									
ClaI	AT CG_AT	1	249								
CviJI	rG Cy	89									
CviRI	TG CA	25									
DdeI	C TnA_G	11									
DpnI	GA TC	28									
DraI	TTT AAA	3	4223	4242	4934						
DraIII	CAC_nnn GTG	1	5500								
DrdI	GACnn_nn nnGTC	4	470	3163	3578	5455					
DrdII	GAACCA	2	1092	5505							
DsaI	C CryG_G	5	424	439	514	806	2442				
EaeI	y GGCC_r	4	677	809	2043	4745					
Eam1105I	GACnn_n nnGTC	1	4357								
EarI	CTCTTCn nnn_	3	987	3354	5152						
EciI	TCCGCC	4	1146	3544	3690	4512					
Eco47III	AGC GCT	4	507	774	2275	2724					
Eco57I	CTGAAGn(14)_nn	2	4012	5024							
EcoNI	CCTnn n_nnAGG	1	904								
Eco0109I	rG GnC_Cy	3	53	802	2476						
EcoRII	=CCwGG_	8	409	1092	1407	1947	2004	3496	3617	3620	
EcoRV	GAT ATC	1	1819								
FauI	CCCGCnnnn nn_	17									
FokI	GGATGnnnnnnnnn nnnn_	11									
FspI	TGC GCA	2	2451	4579							
GdiII	y GGCCG	4	677	809	2043	4745					
HaeI	wGG CCw	5	1097	2418	3485	3496	3942				
HaeII	r_GCGC y	16									
HaeIII	GG CC	25									
HgaI	GACGCnnnnn nnnnn_	14									
HgiEII	ACCnnnnnnGGT	2	967	4050							
HhaI	G_CG C	52									
Hin4I	GAbnnnnnvTC	4	412	1268	4356	4430					
HincII	GTy rAC	1	1875								
HinfI	G AnT_C	15									
HpaI	GTT AAC	1	1875								
HphI	GGTGAnnnnnnn_n	19									
MaeII	A CG_T	15									
MaeIII	=GTnAC_	18									
MboII	GAAGAnnnnnnn_n	16									
MluI	A CGCG_T	1	1369								
MmeI	TCCrACn(18)_nn	3	3685	3869	5477						
MnlI	CCTCnnnnnn_n	27									
MseI	T TA_A	28									
MslI	CAynnnnrTG	10									
MspI	C CG_G	34									
MspA1I	CmG CkG	10									
MwoI	GCnn_nnn nnGC	46									
NarI	GG CG_CC	5	228	692	713	827	2009				
NciI	CC s_GG	12									
NdeI	CA TA_TG	1	543								
NgoAIV	G CCGG_C	6	224	482	679	2267	2427	5601			
NlaIII	_CATG	25									
NlaIV	GGn nCC	26									
NsiI	A_TGCA T	1	543								
NspI	r_CATG y	5	177	844	2815	3107	3474				
Pfl1108I	TCGTAG	1	951								
PflMI	CCAn_nnn nTGG	2	395	569							
PleI	GAGTCnnnn n_	9	630	918	1005	1801	3364	3849	4346	5435	5443
PshAI	GACnn nnGTC	1	2214								
Psp5II	rG GwC_Cy	1	2476								
Psp1406I	AA CG_TT	6	1031	2399	2795	4583	4956	5285			
PstI	C_TGCA G	1	4602								
PvuI	CG_AT CG	1	4727								
PvuII	CAG CTG	4	1969	2062	3061	3880					
RcaI	T CATG_A	3	767	4184	5192						
RsaI	GT AC	6	332	403	512	1516	3276	4837			
SapI	GCTCTTCn nnn_	1	3354								
Sau96I	G GnC_C	20									
Sau3AI	=GATC_	28									
ScaI	AGT ACT	1	4837								
ScrFI	CC n_GG	20									
SfaNI	GCATCnnnnn nnnn_	23									
SfcI	C TryA_G	5	615	3735	3920	4598	5719				
SgrAI	Cr CCGG_yG	1	688								
SphI	G_CATG C	2	177	844							
SspI	AAT ATT	3	184	5161	5292						
StyI	C CwwG_G	1	57								
TaqI	T CG_A	13									
TaqII	GACCGAnnnnnnnnn_nn	9	1277	1495	2168	3372	4705	4890	5043	5060	5404
TfiI	G AwT_C	6	391	2048	2350	2520	3024	3445			
ThaI	CG CG	38									
TseI	CG CG	28									
Tsp45I	=GTsAC_	8	1550	2378	2909	3122	3217	4613	4824	5673	
Tsp509I	=AATT_	15									
Tth111I	GACn n_GTC	1	3215								
Tth111II	CAArCAnnnnnnnnn_nn	8	212	470	1208	1901	2931	4054	4061	4093	
UbaJI	GCsGC	20									
VspI	AT TA_AT	4	630	2054	2113	4529					
XbaI	T CTAG_A	1	58								
XcmI	CCAnnnn_n nnnnTGG	3	1225	1741	1759						
XhoI	C TCGA_G	1	158								
XmnI	GAAnn nnTTC	2	3028	4956							

Enzymes that do not cut pET-31b(+):

AatII	AflII	AgeI	AscI	AvrII	BaeI	BamHI	BseRI
BsmI	BspMI	BsrGI	Bsu36I	EagI	EcoRI	FseI	HindIII
KpnI	MunI	NcoI	NheI	NotI	NruI	NspV	PacI
PmeI	PmlI	RleAI	RsrII	SacI	SacII	SalI	SexAI
SfiI	SgfI	SmaI	SnaBI	SpeI	SrfI	Sse8387I	StuI
SunI	SwaI						