28 Generation of Protein Fragment Libraries by Incremental Truncation

Marc Ostermeier,[1] Stefan Lutz,[2] and Stephen J. Benkovic[2]

[1]*Department of Chemical Engineering, Johns Hopkins University, Baltimore, Maryland 21218;*
[2]*Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802*

INTRODUCTION

Protein fragmentation and domain swapping are valuable methods for the study of inter- and intra-domain and subdomain interactions in proteins. Building on classic examples of protein fragment complementation (Ullman et al. 1967; Kato and Afinsen 1969), protein fragmentation has been used in biophysical studies, particularly in the elucidation of protein folding mechanisms (Tasayco and Carey 1992; Ladurner et al. 1997). More recently, protein engineering using the recombination of protein fragments has gained popularity (for reviews, see Lutz and Benkovic 2000; Ostermeier and Benkovic 2000).

Various methodologies for protein fragmentation have been successfully implemented. On the protein level, target enzymes are usually hydrolyzed by chemical and proteolytic means.

Protein–Protein Interactions: A Molecular Cloning *Manual*, © 2002 by Cold Spring Harbor Laboratory Press, Chapter 28.

Although these fragmentation methods are relatively simple, the resulting complex product mixtures require extensive purification and characterization of the various products. Although this problem is solved by limiting the product "libraries" to a dozen or fewer variants, observed biases toward more accessible regions of a target structure such as surface loops cannot easily be overcome.

Alternatively, protein fragments can be generated by genetic approaches. Using rational design, specific PCR primers can amplify predetermined gene fragments that, upon cloning and transformation into a suitable expression host, allow the isolation of a desired protein segment. Although very efficient on an individual basis, the approach rapidly becomes laborious for multiple fragments. Furthermore, it depends heavily on the availability of structural information to choose the fragments to study. In contrast, incremental truncation is a rapid method for generating a comprehensive gene library that encodes all possible fragment lengths of a target protein and does so in the absence of detailed structural information (Ostermeier et al. 1999c). One important caveat of this method is that a genetic selection system or a high-throughput screen may be needed to analyze the library if the number of possible constructs is large.

The incremental truncation methodology can be applied in a number of different ways (Ostermeier et al. 1999a). First, a library of amino-terminal or carboxy-terminal truncations of a single gene can be generated. Such protein fragments are useful in defining relationships between sequence and structure and function (Jasin et al. 1983), including the detection of regions responsible for protein–protein interactions. Furthermore, the random pairing of members of such libraries can be used to identify protein fragments that can non-covalently associate to form functional proteins (Ostermeier et al. 1999c). This method for protein fragment complementation can also be viewed as a strategy for the conversion of a monomeric protein into a heterodimer. Finally, incremental truncation can be used to create comprehensive fusion libraries between fragments of two genes (Ostermeier et al. 1999b). This method has application in protein engineering, particularly in the generation of combinatorial fusion libraries between genes with low levels of homology.

OUTLINE OF PROCEDURE

Basic Concepts

Incremental truncation is a method for creating a combinatorial library containing every one-base-pair deletion of a gene or gene fragment. As shown in Figure 1, the substrate for incremental truncation is linear double-stranded DNA. The substrate is prepared by digesting closed circular plasmid DNA with restriction enzymes such that a blunt end or 5′ overhang lies near the DNA end to be truncated, whereas a 4-base 3′ overhang on the opposite end protects the remaining vector from degradation. Exonuclease III (Exo III), a 3′-to-5′ exonuclease used to create the truncations, uses 3′ recessed ends or blunt ends as a substrate but cannot digest certain 4-base 3′ overhangs (see Table 1). Alternatively, the protected end can be prepared by filling a 5′ overhang with α-phosphothioate nucleotides (Putney et al. 1981).

The key step in the creation of an incremental truncation library is the time-dependent sampling during Exo III digestion. Over the course of the DNA digestion, small aliquots are removed frequently and the reaction is quenched in a low-pH, high-salt buffer. Temperature and buffer conditions are chosen such that the rate of hydrolysis is balanced with the rate of sampling. For example, if the rate of Exo III is controlled to 10 bases per minute and small aliquots are removed every 20 seconds, the average amount of truncation will increase by 3.3 bases for every sample. Because Exo III digests DNA at a substantially uniform and synchronous rate (Wu et al. 1976) but with a standard deviation of truncation length of 0.22 times the average number of bases truncated (Hoheisel 1993), a population of DNA is created that contains every possible integer trun-

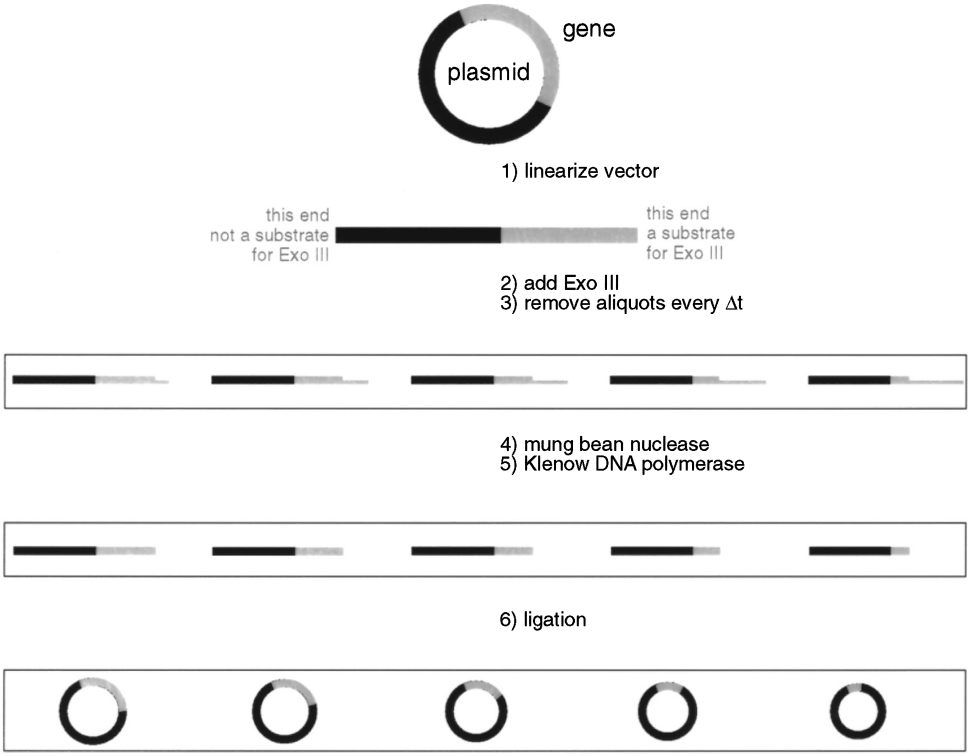

FIGURE 1. General incremental truncation schematic. Incremental truncation is performed on a linear piece of DNA (containing the segment of DNA to be truncated, shown in gray) that has one end protected from digestion and the other end susceptible to digestion. This is easily accomplished, for example, by digestion of plasmid DNA with two restriction enzymes: one that produces a 3´ overhang (which is not a substrate for Exo III) and the other which produces a 5´ overhang (which is a substrate for Exo III). Digestion with Exo III proceeds when the digestion rate is slow enough that the removal of aliquots at frequent intervals results in a DNA library with every one-base deletion. The ends of the DNA can be blunted by treatment with mung bean nuclease and Klenow DNA polymerase so that unimolecular ligation results in the desired incremental truncation library. For some applications, additional DNA manipulations are required before recircularizing the vector.

TABLE 1. Resistance to Exo III Digestion of 4-base 3´ Overhangs Created by Commercially Available Restriction Enzymes

Restriction enzyme	Is 3´ overhang resistant to Exo III?			Successfully used to protect ends for incremental truncation?
	Ref. 1[a]	Ref. 2[b]	Ref. 3[c]	
*Aat*II	yes	yes	yes	–
*Apa*I	no	no	yes	–
*Ban*II[d]	no	yes	yes	–
*Bgl*I[d]	–	yes	yes	–
*Bsp1286*I[d]	no	-	–	–
*BstX*I[d]	–	yes	yes	–
*Hae*II[d]	no	yes	yes	–
*Kpn*I	no	–	yes	–
*Nsi*I	yes	yes	yes	yes
*Pst*I	yes	no	yes	yes
*Sac*I	–	yes	yes	yes
*Sph*I	yes	yes	yes	yes
*Sst*I	yes	–	–	–

[a]Hoheisel (1993).
[b]New England BioLabs (1998).
[a]Promega (1995).
[d]These enzymes contain ambiguous bases in their cut sites. Their resistance may depend on the exact sequence.

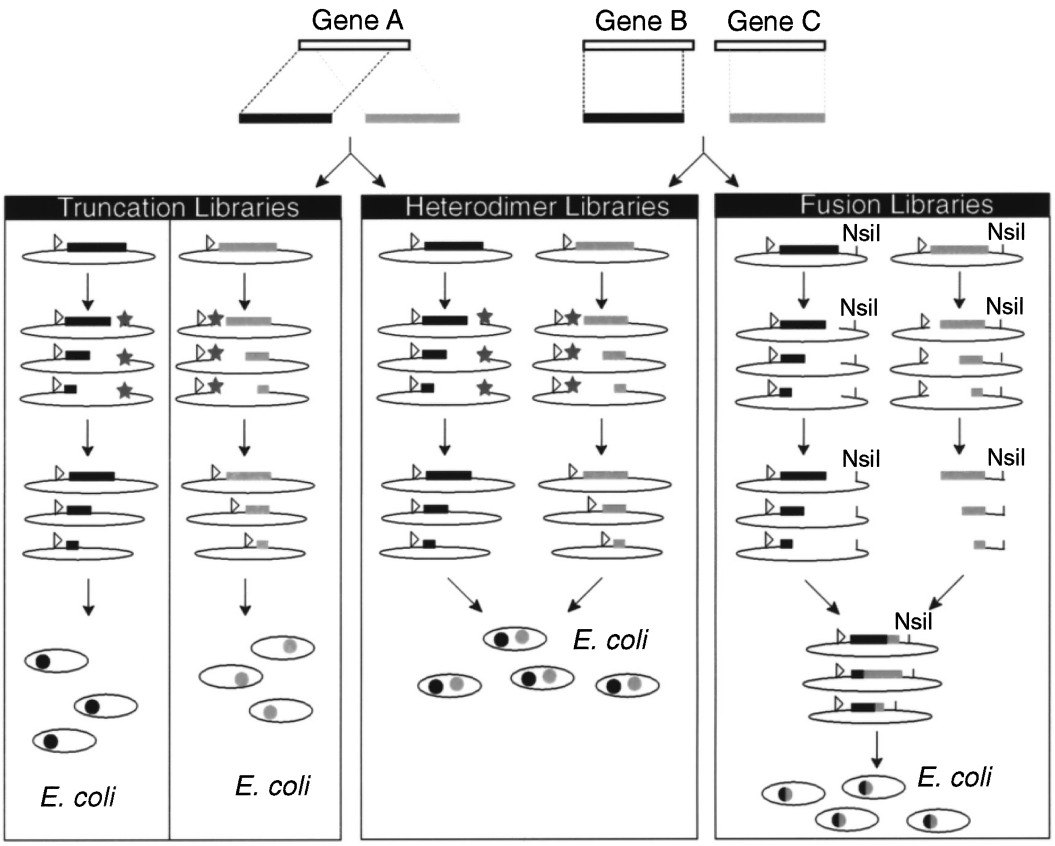

FIGURE 2. Combining incremental truncation libraries. In the left panel, two incremental truncation libraries derived from the 5′ and 3′ fragments of the same gene are prepared separately. Members of the individual library are ligated, introducing the necessary start or stop codons as indicated by the stars. In the middle panel, the two libraries, resulting from a single gene A or two different genes B and C, are randomly paired by cotransformation into *E. coli* cells. Each *E. coli* cell expresses a 5′ fragment and a 3′ fragment, allowing the possible detection fragment complementation. In the right panel, the two truncation libraries are fused on the DNA level such that each *E. coli* cell receives a fusion of a randomly paired 5′ fragment of gene B and a 3′ fragment of gene C.

cation of the starting DNA. Following the time-dependent truncation with Exo III, the single-stranded DNA tails are removed by digestion with a single-strand nuclease, preferably mung bean nuclease, and the ends are blunted using a DNA polymerase. Subsequent steps depend on the intended use of the library.

Three possible uses of libraries are presented in Figure 2. When applied to a single gene (Fig. 2, left), vectors that contain 3′-truncation libraries must carry stop codons in all three reading frames that, upon ligation, align with the library to terminate translation of the protein fragment. Vectors containing 5′-truncation libraries, on the other hand, require a start codon, fused to the 5′ end of the library upon ligation. As shown in Figure 2 (center), truncations can be performed in opposite directions on two overlapping fragments of the same gene A or for two different genes B and C. The two libraries are kept on separate plasmids and can be transformed into the same cells to randomly pair members of the two truncation libraries together. Finally, fusion libraries of the 5′ and 3′ fragments, termed ITCHY libraries, can be prepared from genes B and C as shown in the right-hand panel of Figure 2.

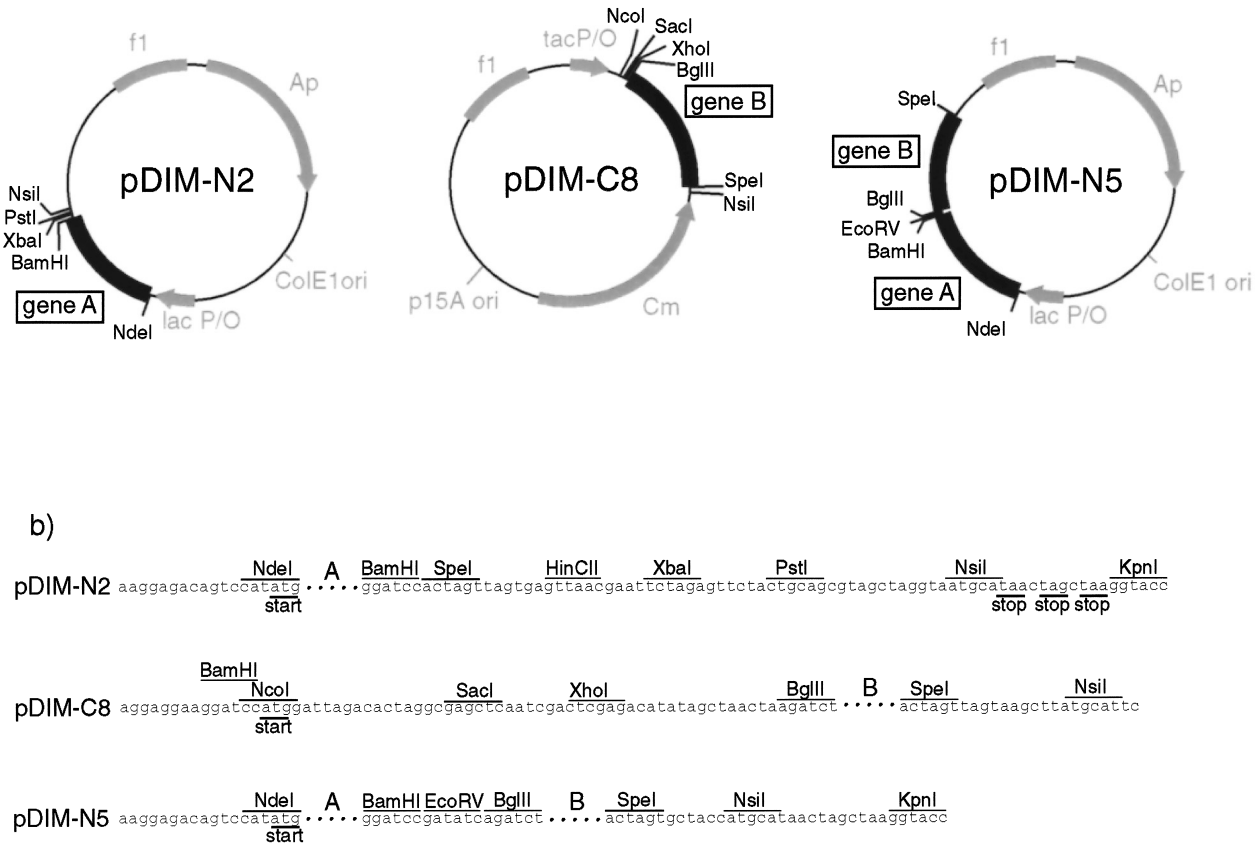

FIGURE 3. Vectors for incremental truncation. Circular maps (*a*) and sequences of cloning (*b*) and truncation regions of incremental truncation vectors. All vectors have the filamentous phage origin of replication (f1) as well as plasmid origins of replication (ColE1 ori and p15A ori). In pDIM-N2 (Ostermeier et al. 1999a), the amino-terminal gene to be truncated is cloned between the *Nde*I and *BamH*I sites downstream from an IPTG-inducible *lac* promoter (lac P/O). The vector also has an antibiotic-resistance gene (ampicillin, Ap). In pDIM-C8 (Ostermeier et al. 1999b), the carboxy-terminal gene to be truncated is cloned between the *Bgl*II and *Spe*I sites downstream from an IPTG-inducible trp/lac hybrid promoter (tac P/O). The vector also has an antibiotic-resistance gene (chloramphenicol, Cm). In pDIM-N5, both gene fragments to be truncated are cloned downstream from an IPTG-inducible *lac* promoter (lac P/O) between the indicated restriction enzyme sites (*Nde*I/*Bam*HI and *Bgl*II/*Spe*I). Between the two gene fragments is located a unique restriction site that produces blunt ends (*Eco*RV).

Description of Vectors for Incremental Truncation

Vectors designed for performing incremental truncation are shown in Figure 3. Single-gene, heterodimer, or fusion libraries can be prepared using vectors pDIM-N2 and pDIM-C8 for the amino- and carboxy-terminal fragments, respectively. Vector pDIM-N5 is used solely for making fusion libraries. The salient features of these vectors include (1) different antibiotic-resistance genes and origins of replication in pDIM-N2 and pDIM-C8, (2) an f1 phage origin of replication, allowing the possibility of packaging into phage for efficient transfer of vector DNA into *E. coli*, (3) unique restriction enzyme sites for cloning the DNA to be truncated, (4) unique restriction enzyme sites for preparing the vectors for incremental truncation, (5) unique restriction enzyme sites for creating incremental truncation fusion libraries (ITCHY libraries), and (6) *lac*-based promoters to express fusion proteins.

Protocol 1

Incremental Truncation

This protocol is divided into the following stages: Preparation of Plasmid DNA, Preparation of DNA for Truncation, and Incremental Truncation. Detailed comments are included at the beginning of each stage.

MATERIALS

CAUTION: See Appendix for appropriate handling of materials marked with <!>.

Buffers and Solutions

Ammonium acetate (7.5 M) <!>

Store at room temperature.

100x Bovine serum albumin (10 mg/ml) (BSA)

Store at –20°C.

Buffer EB (Qiagen, proprietary)

10 mM Tris-HCl (pH 8.5)

Store at room temperature.

Buffer PB

Store at room temperature.

Buffer PE (Qiagen, proprietary)

Store at room temperature.

dNTP mix

0.125 mM dATP
0.125 mM dCTP
0.125 mM dGTP
0.125 mM dTTP

Store at –20°C.

Ethanol (100%) <!>

Store at –20°C.

10x Exo III buffer

660 mM Tris-HCl (pH 8.0)
6.6 mM $MgCl_2$ <!>

Store at 4°C.

7.4x Exo III stop buffer

0.3 M potassium acetate (pH 4.6)
2.5 M NaCl
10 mM $ZnSO_4$
50% glycerol

Concentrations are approximate because this buffer can be difficult to prepare. The following protocol should be used. Mix the following together: 5.5 ml of 3 M potassium acetate, pH 4.6 (prepared by dissolving 2.94 g of potassium acetate in 2 ml of H_2O and adjusting the pH to 4.6 with 4–5 ml of glacial

acetic acid <!> followed by adjusting the volume to 10 ml), 25 ml of 5 M NaCl, 0.27 g of $ZnSO_4 \cdot 7H_2O$, and 25 ml of 100% glycerol.

Store at –20°C.

1× Klenow buffer
- 20 mM Tris-HCl (pH 8.0)
- 100 mM $MgCl_2$ <!>

Store at 4°C.

LB medium (per liter)
- 10 g of tryptone
- 5 g of yeast extract
- 10 g of NaCl

10× Ligase buffer
- 500 mM Tris-HCl (pH 7.5)
- 100 mM $MgCl_2$ <!>
- 100 mM dithiothreitol <!>
- 10 mM ATP
- 250 µg/ml bovine serum albumin

Store at –20°C.

10× Mung bean buffer
- 500 mM sodium acetate (pH 5.0)
- 300 mM NaCl
- 10 mM $ZnSO_4$

Store at 4°C.

NaCl (1 M)

Store at room temperature.

10× NEB1 buffer
- 100 mM Bis Tris propane-HCl (pH 7.0)
- 100 mM $MgCl_2$ <!>
- 10 mM dithiothreitol <!>

Store at 4˚C.

10× *Nsi*I buffer
- 100 mM Tris-HCl (pH 8.4)
- 1 M NaCl
- 10 mM dithiothreitol

Store at 4°C.

PEG (50%) <!>

Store at room temperature.

2x TY (per liter)
- 16 g of tryptone
- 10 g of yeast extract
- 5 g of NaCl

Enzymes

Exonuclease III (~200 units/µl) (Promega)

Store at –20°C.

Klenow mix
- 20 µl of 1× Klenow buffer
- 1 µl of Klenow (5 units) (Promega)

Store at –20°C. Prepare fresh daily.

Ligase mix
320 µl of H_2O
40 µl of 10x ligase buffer
40 µl of 50% PEG
18 Weiss units of T4 DNA Ligase (Promega)
Store at –20°C. Prepare fresh daily.
Mung bean nuclease (~10 units/µl)
*Nco*I (~10 units/µl)
*Nsi*I (~10 units/µl)
Store at –20°C.

Additional Equipment and Reagents

Ampicillin
Agarose gel electrophoresis system (TAE buffer)
Chloramphenicol <!>
Heat blocks at 15°C, 22°C, 30°C, 37°C, and 72°C
Reagent kit (Qiagen Midi Prep kit)

METHOD

Stage 1: Preparation of Plasmid DNA

For incremental truncation, it is important to prepare plasmid DNA free of single-stranded nicks. Exo III can digest from single-stranded nicks in double-stranded DNA, leaving single-stranded gaps. These single-stranded gaps will be digested by mung bean nuclease and result in the undesired effect of random deletions throughout the entire plasmid.

We routinely isolate plasmid DNA of pDIM-N vectors from *E. coli* strain DH5α using commercial plasmid-prep kits. For pDIM-C vectors, we have found that the fraction of nicked molecules and total yield of plasmid DNA are highly dependent on growth conditions. However, if the following protocol for growth is followed, we routinely prepare pDIM-N and pDIM-C vectors with <10% nicked DNA. For high yield of supercoiled DNA, use the following growth conditions:

1. Grow an overnight inoculum of DH5α containing vectors in 10 ml of LB medium, 0.2% glucose, and the appropriate antibiotic (100 µg/ml ampicillin for pDIM-N2 and pDIM-N5; 50 µg/ml chloramphenicol for pDIM-C8) in a 25-ml test tube at 37°C.
2. For pDIM-N2, inoculate 100 ml of 2x TY medium, 2.0% glucose, and 100 µg/ml ampicillin with 2 ml of overnight culture and grow in 500-ml shake flask overnight at 37°C. For pDIM-C8 inoculate 500 ml of LB and 50 µg/ml chloramphenicol with 10 ml of overnight culture and grow overnight in a 2-liter baffled shake flask at 37°C.
3. Isolate plasmid DNA using a standard reagent kit (Qiagen Midi Prep kit). A typical yield of plasmid DNA is ~1 µg/ml culture for pDIM-N2 and pDIM-N5 and 0.2 µg/ml culture for pDIM-C8.

   Should production of sufficiently pure non-nicked DNA prove difficult, methods to purify non-nicked from nicked plasmid DNA include CsCl-ethidium bromide gradients (Sambrook et al. 1989), acid-phenol extraction (Zasloff et al. 1978), or removal of the nicked DNA by enzymatic digestion (Gaubatz and Flores 1990). Alternatively, treatment of nicked DNA with T4 DNA ligase should presumably repair the nicks.

Stage 2: Preparation of DNA for Truncation

Another important factor to consider in preparing non-nicked DNA for incremental truncation is the restriction enzyme digestion to linearize the DNA in preparation for truncation. Restriction enzymes from suppliers may have nuclease contamination. In addition, restriction enzymes may have single-stranded nicking activity at a high enzyme-to-DNA ratio. For this reason, we recommend digesting the DNA with the minimum amount of restriction enzyme necessary to digest the DNA fully and avoid conditions contributing to star activity (relaxed or altered specificity) in restriction enzymes (see manufacturer's product specifications).

We digest closed circular pDIM-N vectors with 1.5 units of *Xba*I and 1.5 units of *Nsi*I or *Pst*I per microgram of DNA in the manufacturer's suggested buffer at 37°C for 1.5–2 hours. pDIM-C vectors are digested with 15 units of *Sac*I and 20 units of *Xho*I per microgram of DNA in the manufacturer's suggested buffer for 1.5–2 hours at 37°C. *Sac*I and *Xho*I have difficulty digesting supercoiled DNA (New England Biolabs 2000), accounting for the higher level of enzyme activity required.

It is important to note that some 4-base 3′overhangs are not completely resistant to Exo III digestion. Table 1 summarizes commercially available enzymes that create 4-base 3′ overhangs and lists whether the 3′ overhangs are resistant to Exo III. Discrepancies between the sources as to whether a 3′ end is resistant to Exo III result from the criteria for designating an end as resistant (most 4-base 3′ overhangs are very weak substrates of Exo III) and, in the case of sites with ambiguous bases, the exact sequence of the site used for testing. An alternative way to protect DNA from digestion is to fill in a 5′ overhang with α-phosphothioate nucleotides (Putney et al. 1981).

1. Digest 10 µg of plasmid DNA with the appropriate restriction enzymes in a total volume of 100 µl at 37°C.
2. After 1.5–2 hours, heat-inactivate the enzymes at the manufacturer's recommended conditions and purify the DNA by ethanol precipitation with 50 µl of ammonium acetate and 0.3 ml of ethanol.
3. Redissolve the DNA pellet in 100 µl of buffer EB. It can be stored for several months at –20°C.

   Alternatively, quench the restriction digestion with 500 µl of buffer PB and purify the DNA by following the QIAquick protocol.

Stage 3: Incremental Truncation

Incremental truncation is performed using temperature and NaCl concentration to control the rate of Exo III digestion. We found the rate of truncation to vary with NaCl at 22°C by the following equation: rate (bp/min)= $47.9 \times 10^{(-0.00644 \times N)}$ where N = concentration of NaCl in mM (0–150 mM). Using this equation, the rate of Exo III digestion is ~10 bases/minute at 22°C in the presence of 100 mM salt. This rate expression is valid for a DNA concentration (1 µg/30 µl) and ratio of Exo III to DNA (100 units/µg DNA).

In contrast to the truncation of pDIM-N2/ -C8 where one of the 3′ ends is protected from Exo III hydrolysis because it is a 3′-overhang (see above), the truncation of pDIM-N5 for making fusion libraries proceeds simultaneously on both ends of the linearized vector. Therefore, twice the amount of Exo III is required to keep the ratio of units of enzyme per moles of 3′ ends constant.

The dependence of Exo III digestion rate on temperature in the absence of NaCl (Henikoff 1987; Hoheisel 1993) and NaCl concentration at higher temperatures (Tomb and Barcak 1989; Hoheisel 1993) has been determined by other investigators.

Upon addition of Exo III to the reaction mixture, aliquots are taken at constant time intervals, usually every 20–60 seconds, depending on rate and overall length of truncation. After quenching the final aliquot, the single-stranded tails are removed by treatment with a single-stranded nuclease. Although we have successfully used S1 nuclease to hydrolyze the single-stranded overhang in the past, recent experiments using mung bean nuclease have produced better results, including larger, more evenly distributed libraries (Ostermeier and Benkovic 2001). Although mung bean nuclease requires a buffer change, this is conveniently accomplished using DNA purification spin columns such as Qiagen's QIAquick column. The use of spin columns greatly improves the yield and quality of truncated DNA compared to ethanol precipitation.

Subsequent processing of the DNA depends on the desired use. Accordingly, this section is divided into descriptions of the construction of heterodimer libraries and ITCHY fusion libraries: (1) single gene truncations and heterodimer libraries, (2) fusion libraries using pDIM-N2 and pDIM-C8, and (3) fusion libraries using pDIM-N5.

This general protocol is applicable for making incremental truncation libraries of a single gene, for making heterodimer libraries, or for making fusion libraries. It is written for truncating over a 300-bp range at a rate of 10 bp/minute. The rate of truncation can be changed by altering the NaCl concentration in step 2. The sample size and frequency of sampling in step 5 may be changed accordingly to truncate for different lengths of time.

1. Equilibrate 180 µl of 1× Exo III stop buffer on ice in a 1.5-ml tube (tube A).
2. To a second 0.65-ml tube (tube B) add:

DNA	2 µg
10× Exo III buffer	6 µl
1 M NaCl	6 µl

   Add H_2O to 60 µl.
3. Equilibrate tube B at 22°C.
4. At time = 0, add 200 units of Exo III to tube B and mix immediately.
5. Beginning at 30 seconds, remove 1-µl samples every 30 seconds and add to tube A. Mix tube A well. Note that all time points are removed to tube A, which is kept on ice.

   The rate of Exo III is very temperature dependent. It is preferable to leave tube B open during the sampling to avoid warming the tube by repeated handling.
6. After all samples are taken, add 1.2 ml of QIAquick buffer PB.
7. Follow the QIAquick protocol.
8. Elute DNA from column with 47 µl of buffer EB.
9. Add:

   5 µl of 10× mung bean nuclease buffer
   1 µl of mung bean nuclease (10 units)
10. Incubate for 30 minutes at 30°C.
11. Add 250 µl of QIAquick buffer PB.
12. Follow the QIAquick protocol to purify truncated DNA.
13. Follow Method A, B, or C (step 14 on) below depending on the type of library.

Single Gene Truncation and Heterodimer Libraries

When truncating 5′ and 3′ gene fragments individually, start and stop codons must be introduced in a separate step. For the 3′-truncated fragment, a stop codon must be fused to the library. Because the reading frame at the end of the truncated gene is unknown, a series of stop codons in all three frames is used. For two of the frames, this results in the addition of one to three carboxy-terminal residues absent in the original protein. However, for the third frame, no extra residues are added. Similarly, it is unknown what reading frame will provide the correct start codon for the 5′ truncated fragment. Thus, one-third of the library will be in-frame and therefore produce meaningful carboxy-terminal protein fragments.

As shown in Figure 4, the stop codon triplet and start codon can be exposed upon restriction enzyme digestion. Digestion of pDIM-N2 with *Nsi*I followed by removal of the 3′ single-stranded overhang exposes the sequence 5′-TAACTAGCTAA-3′ containing stop codons in all three frames for fusion to the truncated gene. In this case, the 3′ overhang is removed using the 3′ to 5′ exonuclease activity of Klenow DNA polymerase in the absence of dNTPs. Subsequently, dNTPs are added to fill in any truncations past the blunt end of the former *Nsi*I site and to blunt the end of the DNA that was truncated. Digestion of pDIM-C8 with *Nco*I followed by a fill-in reaction with Klenow DNA polymerase and dNTPs on the 5′ single-stranded overhang exposes an ATG start codon for fusion to the truncate gene. Both libraries are then circularized by ligation under dilute conditions so as to favor intramolecular ligation and prepared for transformation into *E. coli*.

Method A: Making an Incremental Truncation Library of a Single Gene or for Heterodimers

14. Elute the truncated DNA from the QIAquick column with 90 µl of buffer EB.
15. For restriction enzyme digestions: For pDIM-N2, add 10 µl of 10× *Nsi*I buffer, 1 µl of 100× BSA, and 15 units of *Nsi*I. For pDIM-C8, add 10 µl of 10× NEB1 buffer, 1 µl of 100× BSA, and 18 units of *Nco*I.
16. Incubate for 2 hours at 37°C.
17. Add 0.5 ml of buffer PB.
18. Follow the QIAquick protocol.
19. Elute DNA from the column with 82 µl of buffer EB.
20. Equilibrate the tube at 37°C.
21. Klenow treatment: For pDIM-N2, add 10 µl of Klenow buffer, incubate for 3 minutes at 37°C, add 10 µl of dNTP mix, and incubate for 5 minutes at 37°C. For pDIM-C8, add 10 µl of dNTP mix and 10 µl of Klenow buffer and incubate for 5 minutes at 37°C.
22. Inactivate the Klenow buffer by incubating for 20 minutes at 72°C.
23. Cool to room temperature and add 0.4 ml of ligase buffer.
24. Incubate for ≥12 hours at room temperature.
25. Concentrate by ethanol precipitation with 250 µl of ammonium acetate and 1.5 ml of ethanol.
26. Transform into the desired host.

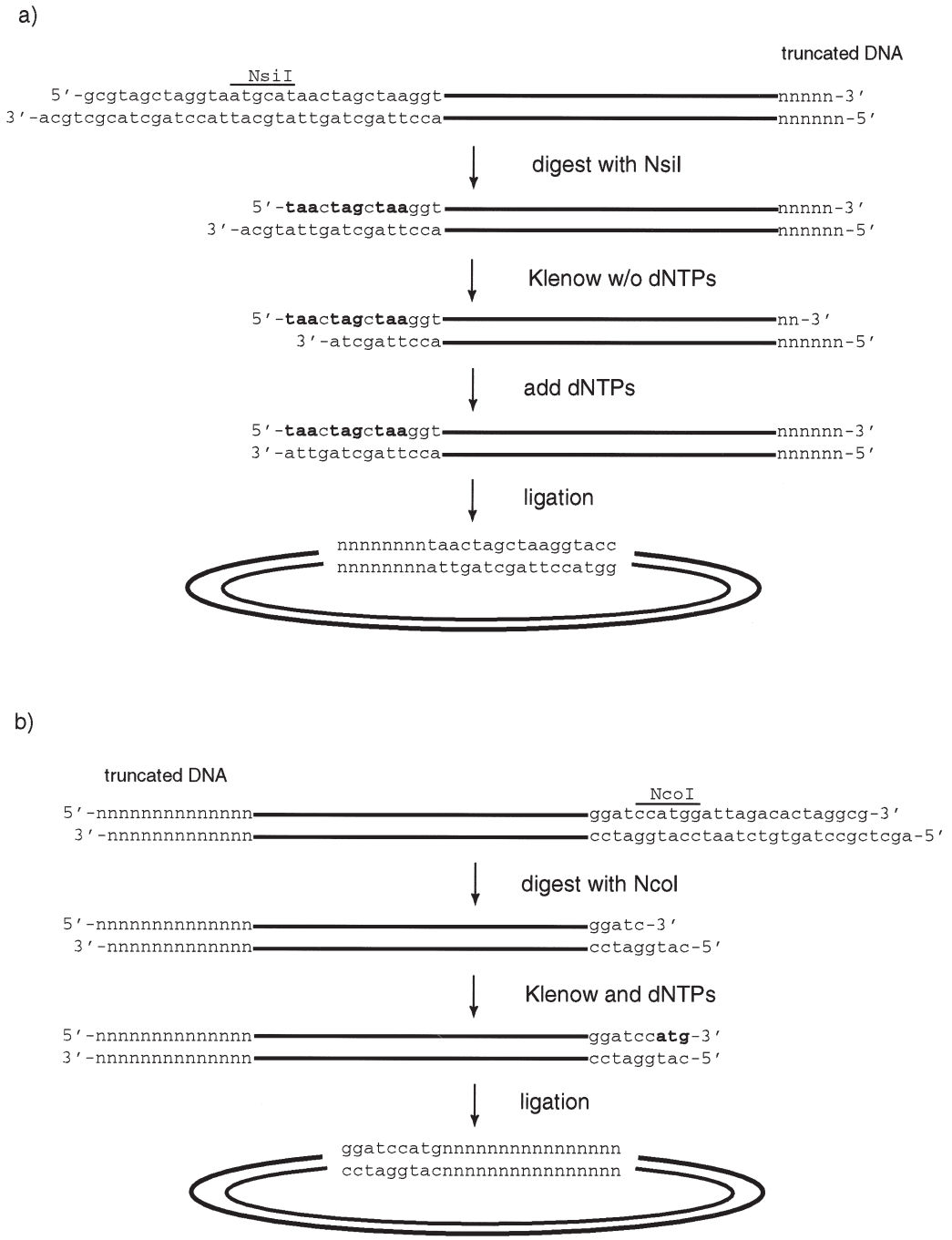

FIGURE 4. Fusion of truncation libraries to stop or start codons. (*a*) Truncations from the 3´ end of a gene need to be fused to a stop codon for the appropriate fragment to be expressed. Digestion with *Nsi*I exposes a series of three stop codons (shown in bold) in all three frames. The 3´-to-5´ exonuclease activity of Klenow is used to remove the 3´ overhang next to the first stop codon so that fusion to the truncated gene can be done in a seamless manner. Thus, if the gene fragment is in frame with this first stop codon, no additional carboxy-terminal residues will be added to the truncated protein. The other two frames result in one to three additional carboxy-terminal residues. (*b*) Truncations from the 5´ end of a gene need to be fused to a start codon in order for the appropriate fragment to be expressed. This is conveniently achieved by digestion with *Nco*I followed by a fill-in reaction with Klenow and dNTPs. The ATG start codon (shown in bold) is fused seamlessly to the truncated gene fragment, one-third of which will be in frame.

Method B: Making Fusion Libraries Using pDIM-N2 and pDIM-C8

After the mung bean nuclease treatment and purification using the QIAquick column according to the manufactuer's instructions, the DNA is blunt-ended with Klenow fragment DNA polymerase and dNTPs. Following enzyme inactivation by heat denaturation, the DNA is digested with *Nsi*I to prepare pDIM-N2 to receive the truncated library constructed in pDIM-C8. The *Nsi*I-digested DNA is subjected to agarose gel electrophoresis, and the desired range of DNA sizes for pDIM-N2 with its truncation library and the truncation library from pDIMC-8 is isolated. We obtain our largest libraries when we use electroelution to recover the DNA from the agarose, followed by ethanol precipitation and ligation in a volume of 20 µl or less.

14. Elute the truncated DNA from the QIAquick column with 72 µl of buffer EB.
15. Equilibrate the tube at 37°C.
16. Klenow treatment: add 10 µl of dNTP mix and 10 µl of Klenow buffer. Incubate for 5 minutes at 37°C.
17. Inactivate the Klenow buffer by incubating for 20 minutes at 72°C.
18. For restriction enzyme digestions: add 10 µl of 10× *Nsi*I buffer, 1 µl of 100× BSA, and 15 units of *Nsi*I.
19. Incubate for 2 hours at 37°C.
20. Incubate for 20 minutes at 72°C.
21. Isolate the large fragment from the pDIM-N2 digestion and the small fragment library from pDIM-C8 digestion by agarose gel electrophoresis.

    We have had the most success in creating large libraries by electroeluting the DNA from the gel slice using the S&S Elutrap electro-separation system (Schleicher & Schuell, Keene, New Hampshire 03431) followed by concentration by ethanol precipitation with ammonium acetate into 17 µl of H_2O.
22. Ligation:

    17 µl of DNA
    2 µl of ligase buffer
    6 Weiss units of T4 DNA ligase
23. Incubate for ≥12 hours at 15°C.
24. Transform into the desired host.

Method C: Making Fusion Libraries Using pDIM-N5

In this variation, after the mung bean nuclease step and purification using the QIAquick column, the DNA is treated with Klenow fragment DNA polymerase and dNTPs to fully blunt the truncated gene. After enzyme inactivation by heat denaturation, the DNA is digested with *Nsi*I to uncouple the synchronized truncation libraries. No gel purification is required, and ligation will randomly fuse members of the two libraries creating the ITCHY library.

The avoidance of the gel purification step saves time and increases the yield of DNA for subsequent ligation. However, this comes at the expense of being unable to select only those truncations that occur in the desired size range, which (1) may be important in some applications and (2) eliminates the low level of abhorrent truncation products that appear due to the small fraction of nicked DNA in the starting material and for other reasons (Henikoff 1984).

14. Elute the truncated DNA from the QIAquick column with 72 µl of buffer EB.
15. Equilibrate the tube at 37°C.
16. Klenow treatment: add 10 µl of dNTP mix and 10 µl of Klenow buffer and incubate for 5 minutes at 37°C.
17. Inactivate the Klenow buffer by incubating for 20 minutes at 72°C.
18. For restriction enzyme digestions: add 10 µl of 10× *Nsi*I buffer, 1 µl of 100× BSA, and 15 units of *Nsi*I.
19. Incubate for 2 hours at 37°C.
20. Incubate for 20 minutes at 72°C.
21. Ethanol-precipitate into 17 µl of H_2O.
22. Ligation:

    17 µl of DNA
    2 µl of ligase buffer
    6 Weiss units of T4 DNA ligase
23. Incubate for ≥12 hours at 15°C.
24. Transform into the desired host.

Controls for Incremental Truncation

Simple control experiments can provide valuable information concerning the performance of Exo III during the truncation reaction. By doubling the scale of the digestion reaction but removing only the normal sample volume over the course of the digestion, half of the volume will remain after the last time point. This second half is then quenched in a separate tube and processed as normal through the mung bean nuclease digestion and subsequent purification. The size of this DNA, which corresponds to DNA that has undergone maximum truncation, can then be analyzed by agarose gel electrophoresis to determine whether truncation has proceeded normally and over the desired length. To make size determination more accurate, or in cases where the length of truncation is small in comparison to the size of the plasmid, it may be advantageous to digest this DNA with a restriction enzyme prior to electrophoresis. It is important to remember that the truncated DNA will run as a smear on the gel because the standard deviation in truncation length varies at 0.22 times the truncation length.

Verification of Library Diversity

Following transformation into a suitable host strain, the easiest and quickest method to verify the diversity of the incremental truncation library is to perform PCR on randomly selected colonies using primers that are complementary to regions outside the desired range of truncation. The size of the amplification product is determined by agarose gel electrophoresis. Alternatively, restriction analysis can be performed on plasmid DNA isolated from randomly selected library members.

TABLE 2. Potential Applications of the Incremental Truncation Technology

Amino-terminal or carboxy-terminal truncations of a single gene
Identify subdomains or motifs responsible for dimerization or heterologous interactions
Define minimal functional units
Assign subdomains or motifs responsible for various protein functions
Create proteins with altered properties (most likely loss of function)
Identify protein fragments suitable for crystallization experiments
Combining amino-terminal and carboxy-terminal truncations of a single gene (heterodimers)
Identify interacting fragments of the protein
Distinguish independent folding units within a protein
Identify possible ancestral fusion points
Fusing amino-terminal and carboxy-terminal fusions of a single gene
Identify subdomains or motifs responsible for dimerization or heterologous interactions
Generate functional proteins with internal deletions or tandem duplications
Define minimal functional units
Assign function to subdomains or motifs within the protein
Map out flexible segments of the protein that can accommodate changes
Create proteins with altered properties
Fusing amino-terminal and carboxy-terminal fusions of two different genes
Force independent proteins or domains in proximity to one another
Create hybrid protein libraries
Optimize linker regions between proteins to be fused

CONCLUSION

The methodology outlined above provides multiple opportunities to study protein interactions. In addition, it allows the modification and alteration of existing protein properties, as well as the generation of novel activities in the context of combinatorial libraries. A list of potential applications of the incremental truncation technology is given in Table 2.

REFERENCES

Gaubatz J.W. and Flores S.C. 1990. Purification of eukaryotic extrachromosomal circular DNAs using exonuclease III. *Anal. Biochem.* **184:** 305–310.

Henikoff S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28:** 351–359.

———. 1987. Unidirectional digestion with exonuclease III in DNA sequence analysis. *Methods Enzymol.* **155:** 156–165.

Hoheisel J.D. 1993. On the activities of *Escherichia coli* exonuclease III. *Anal. Biochem.* **209:** 238–246.

Jasin M., Regan L., and Schimmel P. 1983. Modular arrangement of functional domains along the sequence of an aminoacyl tRNA synthetase. *Nature* **306:** 441–447.

Kato I. and Anfinsen C.B. 1969. On the stabilization of ribonuclease S-protein by ribonuclease S-peptide. *J. Biol. Chem.* **244:** 1004–1007.

Ladurner A.G., Itzhaki L.S., de Prat Gay G., and Fersht A.R. 1997. Complementation of peptide fragments of the single domain protein chymotrypsin inhibitor 2. *J. Mol. Biol.* **273:** 317–329.

Lutz S. and Benkovic S.J. 2000 Homology-independent protein engineering. *Curr. Opin. Biotechnol.* **11:** 319–324.

New England BioLabs 1998. *Exo-size deletion kit instruction manual*. New England BioLabs, Beverly, Massachusetts.

———. 2000. *New England BioLabs catalog and technical reference*. New England BioLabs, Beverly, Massachusetts.

Ostermeier M. and Benkovic S.J. 2000. Evolution of protein function by domain swapping. *Adv. Protein Chem.* **55:** 29–77.

———. 2001. Construction of hybrid gene libraries involving the circular permutation of DNA. *Biotech. Lett.* **23:** 303–310.

Ostermeier M., Nixon A.E., and Benkvoic S.J. 1999a. Incremental truncation as a strategy in the engineering of novel catalysts. *Bioorg. Med. Chem.* **7:** 2139–2144.

Ostermeier M., Shim J.H., and Benkovic S.J. 1999b. A combinatorial approach to hybrid enzymes independent of DNA homology. *Nat. Biotechnol.* **17:** 1205–1209.

Ostermeier M., Nixon A.E., Shim J.H., and Benkovic S.J. 1999c. Combinatorial protein engineering by incremental truncation. *Proc. Natl. Acad. Sci.* **96:** 3562–3567.

Promega. 1995. *Erase-a-base system technical manual,* #TM006. Promega, Madison, Wisconsin.

Putney S.D., Benkovic S.J., and Schimmel P.R. 1981. A DNA fragment with an α-phosphorothioate nucleotide at one end is asymmetrically blocked from digestion by exonuclease III and can be replicated in vivo. *Proc. Natl. Acad. Sci.* **78:** 7350–7354.

Sambrook J., Fritsch E.F., and Maniatis T. 1989. *Molecular cloning: A laboratory manual,* 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Tasayco M.L. and Carey J. 1992. Ordered self-assembly of polypeptide fragments to form nativelike dimeric trp repressor. *Science* **255:** 594–597.

Tomb J.-F. and Barcak G.J. 1989. Regulating the 3′-5′ activity of exonuclease III by varying the sodium chloride concentration. *BioTechniques* **7:** 932–933.

Ullmann A., Jacob F., and Monod J. 1967. Characterization by in vitro complementation of a peptide corresponding to an operator-proximal segment of the beta-galactosidase structural gene of *Escherichia coli. J. Mol. Biol.* **24:** 339–343.

Wu R., Rube G., Siegel B., Jay E., Spielman P., and Tu C.D. 1976. Synchronous digestion of SV40 DNA by exonuclease III. *Biochemistry* **15:** 734–740.

Zasloff M., Ginder G.D., and Felsenfeld G. 1978. A new method for the purification and identification of covalently closed circular DNA molecules. *Nucleic Acids Res.* **5:** 1139–1152.