[10]   Construction of Protein Fragment Complementation Libraries Using Incremental Truncation

By DAVID E. PASCHON and MARC OSTERMEIER

Introduction

Many proteins can have their peptide backbone cut by proteolytic or genetic means, yet the two fragments can associate to make an active heterodimer. This ''monomer-to-heterodimer conversion'' is referred to as protein fragment complementation (PFC). Such complementation is the reverse of evolutionary processes in which domains are recruited and fused at the genetic level.[1] Classic examples of protein fragment complementation include ribonuclease S[2] and β-galactosidase.[3] Protein fragment complementation can be used to examine theories of protein evolution,[4–6] protein folding,[7] macromolecular assembly,[8] structure–function relationships,[9,10] and mapping contacts in membrane proteins.[11] Sites for successful protein bisection for protein fragment complementation are quite varied. Bisection sites need not fall between well-defined domains or structural units and fall within conserved and nonconserved regions, as well as within secondary structure elements such as α helixes.[9,12] Overlapping sequences at the bisections point are often tolerated and, in some cases, even required.

For most locations, bisection does not lead to protein fragment complementation, presumably due to inefficient assembly or improper folding of the fragments. For some sites, this can be overcome by fusion of the fragments to dimerization domains to facilitate correct assembly. This is known as assisted protein reassembly (APR). A bisection point for APR

[1] M. Ostermeier and S. J. Benkovic, *Adv. Protein Chem.* **55,** 29 (2000).

[2] I. Kato and C. B. Anfinsen, *J. Biol. Chem.* **244,** 1004 (1969).

[3] A. Ullmann, F. Jacob, and J. Monod, *J. Mol. Biol.* **24,** 339 (1967).

[4] B. L. Bertolaet and J. R. Knowles, *Biochemistry* **34,** 5736 (1995).

[5] G. de Prat-Gay, *Protein Eng.* **9,** 843 (1996).

[6] B. Hocker, S. Beismann-Driemeyer, S. Hettwer, A. Lustig, and R. Sterner, *Nature Struct. Biol.* **8,** 32 (2001).

[7] A. G. Ladurner, L. S. Itzhaki, G. P. Gray, and A. R. Fersht, *J. Mol. Biol.* **273,** 317 (1997).

[8] M. L. Tasayco and J. Carey, *Science* **255,** 594 (1992).

[9] K. Shiba and P. Schimmel, *Proc. Natl. Acad. Sci. USA* **89,** 1880 (1992).

[10] J.-M. Betton and M. Hofnung, *EMBO J.* **13,** 1226 (1994).

[11] H. Yu, M. Kono, T. D. McKee, and D. D. Oprian, *Biochemistry* **34,** 14963 (1995).

[12] M. Ostermeier, A. E. Nixon, J. H. Shim, and S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **96,** 3562 (1999).

must meet two requirements: (a) the respective protein fragments do not assemble efficiently into an active protein in the absence of a dimerization domain and (b) when these fragments are each fused to one half of a dimer (a dimerization domain), the activity of the protein is restored. In the case where the bisected protein is an enzyme, assembly of the dimerization domains can be detected as the reconstitution of enzyme activity. Such enzymatic two-hybrid systems [also referred to as ''protein fragment complementation assays'' (PCAs)] have a number of biotechnological applications.[13] Enzymes that have been used as PCAs include dihydrofolate reductase,[14–16] glycinamide ribonucleotide transformylase,[17] green fluorescent protein,[18] ubiquitin,[19,20] β-galactosidase,[21,22] β-lactamase,[23,24] and aminoglycoside and hygromycin B phosphotransferases.[17] The primary use of these PCAs has been as an enzymatic two-hybrid system to evaluate protein–protein interactions[16] and to select for interacting proteins,[25,26] including antibody/antigen pairs.[27]

Incremental Truncation

Incremental truncation,[12] a method for rapidly generating a DNA library of every 1 base pair deletion of a gene or gene fragment (Fig. 1), can be used to evaluate ''protein fragment complementation space.'' Separate,

[13] S. W. Michnick, *Chem. Biol.* **7,** R217 (2000).

[14] J. N. Pelletier, F. X. Campbell-Valois, and S. W. Michnick, *Proc. Natl. Acad. Sci. USA* **95,** 12141 (1998).

[15] I. Remy and S. W. Michnick, *Proc. Natl. Acad. Sci. USA* **96,** 5394 (1999).

[16] I. Remy, I. A. Wilson, and S. W. Michnick, *Science* **283,** 990 (1999).

[17] S. W. Michnick, I. Remy, F. X. Campbell-Valois, A. Vallee-Belisle, and J. N. Pelletier, *Methods Enzymol.* **328,** 208 (2000).

[18] I. Ghosh, A. D. Hamilton, and L. Regan, *J. Am. Chem. Soc.* **122,** 5658 (2000).

[19] N. Johnson and A. Varshavsky, *Proc. Natl. Acad. Sci. USA* **91,** 10340 (1994).

[20] G. Karimova, J. Pidoux, A. Ullmann, and D. Ladant, *Proc. Natl. Acad. Sci. USA* **95,** 5752 (1998).

[21] F. Rossi, C. A. Charlton, and H. M. Blau, *Proc. Natl. Acad. Sci. USA* **94,** 8405 (1997).

[22] F. M. Rossi, B. T. Blakely, C. A. Charlton, and H. M. Blau, *Methods Enzymol.* **328,** 231 (2000).

[23] A. Galarneau, M. Primeau, L. E. Trudeau, and S. W. Michnick, *Nature Biotechnol.* **20,** 619 (2002).

[24] T. Wehrman, B. Kleaveland, J. H. Her, R. F. Balint, and H. M. Blau, *Proc. Natl. Acad. Sci. USA* **99,** 3469 (2002).

[25] K. M. Arndt, J. N. Pelletier, K. M. Muller, T. Alber, S. W. Michnick, and A. Pluckthun, *J. Mol. Biol.* **295,** 627 (2000).

[26] J. N. Pelletier, K. M. Arndt, A. Pluckthun, and S. W. Michnick, *Nature Biotechnol.* **17,** 683 (1999).

[27] E. Mossner, H. Koch, and A. Pluckthun, *J. Mol. Biol.* **308,** 115 (2001).
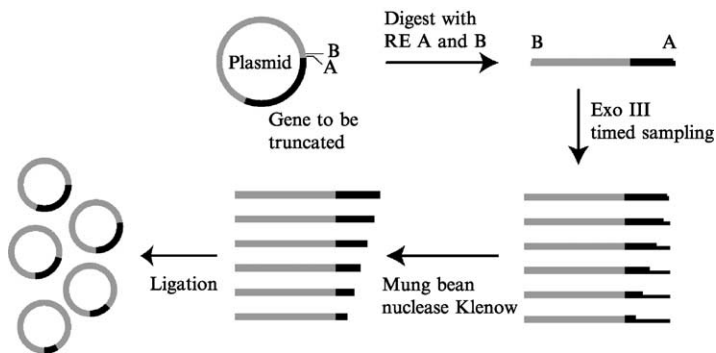
FIG. 1. Incremental truncation. The gene or gene fragment is cloned into a plasmid designed for incremental truncation. The vector is then digested with restriction enzymes A and B. Restriction enzyme A leaves a 3′ recessed end, which is susceptible to Exo III digestion, whereas restriction enzyme B leaves a 3′ overhang, which is resistant to Exo III digestion. Following the addition of Exo III, small samples are taken after discrete time intervals. Single-stranded tails left from Exo III digestion on the plasmids are removed by incubating with mung bean nuclease and the ends are blunted by Klenow fragment. The plasmids are then recircularized using ligase under dilute conditions and are thus ready for transformation into the desired host.

comprehensive libraries of N-terminal and C-terminal truncations of any gene can be constructed in compatible vectors and cotransformed (Fig. 2). Provided that a sufficient selection or screen is available for the activity of the bisected protein, all possible protein fragment pairs can be evaluated in a single experiment. Sites for PFC or APR can be thus identified.

There are two basic methods for constructing incremental truncation libraries: time-dependent truncation (described here) and a method involving the polymerase-catalyzed incorporation of α-phosphothioate nucleotides (THIO truncation).[28,29] Although the latter has the advantage of experimental convenience and has its advantages for the construction of ITCHY libraries,[30] for purposes of creating libraries for PFC or APR, it suffers from a bias toward short truncations and limited control over the truncation range. Time-dependent truncation is preferred, as it produces a much more even distribution of truncation products.[30]

A schematic depiction of time-dependent truncation is shown in Fig. 1. Plasmid DNA containing the gene to be truncated is prepared by digesting closed circular plasmid DNA with restriction enzymes such that (1) a 3′

[28] S. Lutz, M. Ostermeier, and S. J. Benkovic, *Nucleic Acids Res.* **29,** e16 (2001).
[29] M. Ostermeier and S. Lutz, *Methods Mol. Biol.* **239,** 129 (2003).
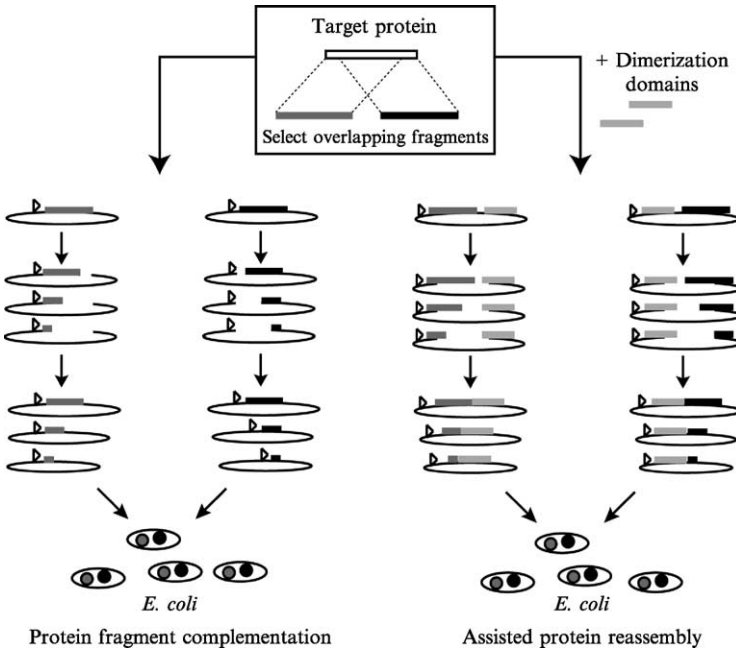[30] M. Ostermeier, *Biotechnol. Bioeng.* **82,** 564 (2003).

Fig. 2. Use of incremental truncation to identify fragments for PFC and APR. Large, overlapping (and ideally, inactive) N-terminal and C-terminal fragments are selected from the gene of the target protein. For PFC, these fragments are cloned into vectors designed for performing incremental truncation. The N-terminal fragment is truncated from its C terminus, and the C-terminal fragment is truncated from its N terminus. N-terminal fragments are fused to a series of stop codons in all three frames, and C-terminal fragments are fused to an ATG start codon. Individual N-terminal and C-terminal truncation libraries are first constructed and then cotransformed into *E. coli* cells. For APR, the same procedure is followed except that the vectors contain dimerization domains, which are fused to the truncated genes after truncation has been performed.

recessed end is left at the end of the DNA to be truncated and (2) a 3′ overhang is left on the opposite end of the vector to protect it from degradation (the end not to be truncated can also be protected by the incorporation of α-phosphothioate nucleotides[31]). Exonuclease III (Exo III) is used to digest linear double-stranded DNA. Exo III is a 3′–5′ exonuclease that can digest 3′ recessed or blunt ends efficiently, but cannot digest from a 3′ overhang of four or more bases.[32] The incremental

[31] S. D. Putney, S. J. Benkovic, and P. R. Schimmel, *Proc. Natl. Acad. Aci. USA* **78,** 7350 (1981).
[32] J. D. Hoheisel, *Anal. Biochem.* **209,** 238 (1993).

truncation library is created by time-dependent sampling during Exo III digestion. The rate of Exo III digestion is controlled carefully by a combination of low temperature and salt. During digestion, small aliquots are removed and the reaction is quenched in a low-pH, high-salt buffer. The resulting single-stranded DNA tails are removed by digestion with mung bean nuclease and the ends are then blunted using the Klenow fragment. The blunt ends of the linear plasmid are then ligated, and the resulting library is transformed into the desired host organism.

Aside from constructing libraries to identify fragments for PFC and APR, the following protocols can also be used to optimize peptide linkers in gene fusions and to construct libraries of promoter variants that vary the distance between important sites within the promoter. In addition, the truncation of a protein can result in altered properties. Removal of the first 70 N-terminal amino acids of a phospholipase resulted in the change of substrate specificity from a phospholipase to a lipase.[33] Incremental truncation can also be used to create fusion libraries between genes (called ITCHY libraries). Detailed protocols for the construction of ITCHY libraries have been published elsewhere[29] and are not described here.

Materials and Methods

Vectors for Creating PFC or APR Libraries

The vectors used for incremental truncation are shown in Fig. 3; other vectors are suitable as well, provided that they meet certain criteria outlined later. Plasmid pDIM-N2 is used for making libraries of N-terminal fragments, whereas pDIM-C8 is used for making libraries of C-terminal fragments. Each plasmid contains a different antibiotic resistance marker: ampicillin for pDIM-N2 and chloramphenicol for pDIM-C8. Each also has a different origin of replication so that both plasmids may be stably maintained in the same bacterial cell. The f1 origin of replication allows for the possibility of packing DNA into phage as another method for transformation into *Escherichia coli*. IPTG (Isopropyl-BD-Thiogalactoside) inducible *lac*-based promoters allow for leaky expression of the resulting protein fragments under noninducing conditions, with higher expression under inducing conditions.

The gene to be truncated is cloned between the *Nde*I and the *Bam*HI (or *Spe*I) sites in pDIM-N2 and between the *Bgl*II and the *Spe*I sites in pDIM-C8. For use in creating libraries for PFC, pDIM-N2 contains stop codons in all three frames between the *Nsi*I and the *Kpn*I restriction sites. This ensures that all N-terminal protein fragments will have a stop codon

[33] J. K. Song, B. Chung, Y. H. Oh, and J. S. Rhee, *Appl. Environ. Microbiol.* **68,** 6146 (2002).
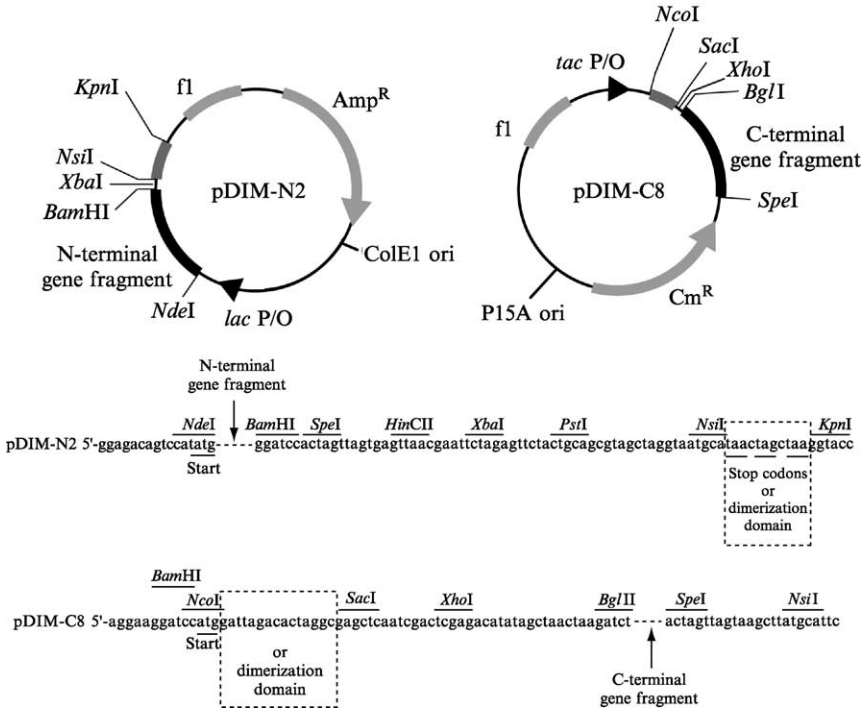
FIG. 3. Circular maps and sequences of cloning and truncation regions of incremental truncation vectors. All vectors have the filamentous phage origin of replication (f1), as well as plasmid origins of replication (ColE1 ori and p15A ori). In pDIM-N2, the N-terminal gene to be truncated is cloned between the *Nde*I and the *Bam*HI sites downstream from an IPTG-inducible lac promoter (*lac* P/O). The vector also has an antibiotic resistance gene (ampicillin, Amp[R]). In pDIM-C8, the C-terminal gene to be truncated is cloned between the *Bgl*II and the *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[R]). For PFC, pDIM-N2 contains a series of stop codons between the *Nsi*I and the *Kpn*I sites. For APR, dimerization domains are located between the *Nsi*I and the *Kpn*I sites of pDIM-N2 and the *Nco*I and the *Sac*I sites of pDIM-C8.

proximal to the truncated gene and contain C-terminal extensions of no more than three residues not in the original protein (one-third of all N-terminal protein fragments will contain no extensions). Plasmid pDIM-C8 contains an *Nco*I site, which, upon digestion and blunt end formation, leaves an ATG start codon that is fused to the truncated C-terminal gene fragment. One-third of this library will have this start codon in-frame with the truncated gene, the other two-thirds will not produce truncation fragments except in the case of spurious internal translation initiation.

For constructing libraries for APR, the dimerization domains must be positioned properly in the vectors. This is accomplished by cloning the gene for one subunit of the dimer between *Nsi*I and *Kpn*I in pDIM-N2. If the first codon of the dimerization domain begins with a T, then this can be the last T in the *Nsi*I site. Alternatively, the last T in the *Nsi*I site can serve as the wobble base for genes in which truncation stopped after removal of the third base of a codon; however, this will necessarily result in some truncations having a nonsynonymous mutation at the fusion point. The gene for the other half of the dimer is cloned between the *Nco*I and the *Sac*I sites in pDIM-C8. In most cases the last codon of the dimerization domain either already ends with a G or can be mutated silently to a G, which can be designed to be the first G of the *Sac*I site.

Preparation of Plasmid DNA for Truncation

Plasmid DNA to be truncated must be prepared to ensure that the majority of plasmid molecules do not contain a nick. 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. It is also important to limit the amount of restriction enzyme used to avoid nicking of the DNA for the same reasons. We routinely isolate plasmid DNA of pDIM-N2 and pDIM-C8 vectors from *E. coli* strain DH5α using commercial plasmid prep kits, although for pDIM-C8 vectors, we have found that the fraction of nicked molecules and the total yield of plasmid DNA are highly dependent on growth conditions.[34] Note that *Sac*I and *Xho*I have difficulty digesting supercoiled DNA, thus necessitating higher amounts of enzyme during digestions to ensure the complete digestion of the pDIM-C8 vector.

Digestion mixtures are as follows. pDIM-N2 consists of 10 μg of pDIM-N2, 10 μl of 10× NEB buffer 3, 1 μl of 100× bovine serum albumin (BSA), 1.5 μl of *Pst*I (30 units), 2 μl of *Nsi*I (20 units), and water to 100 μl. pDIM-C8 consists of 10 μg of pDIM-C8, 15 μl of 10× NEB buffer 1, 1 μl of 100× BSA, 7.5 μl of *Sac*I (150 units), 10 μl of *Xho*I (200 units), and water to 150 μl.

1. Incubate both digestions at 37° for 1.5–2 h.
2. Incubate at 65° for 20 min to inactivate enzymes.
3. Add 5× volume of digestion of QIAquick buffer PB.

[34] M. Ostermeier, S. Lutz, and S. J. Benkovic, *in* ''Protein–Protein Interactions: A Molecular Cloning Manual'' (E. Golemis, ed.). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2002.

    4. Follow QIAquick protocol.
    5. Elute DNA from the column with 50 μl of QIAquick buffer EB.

Construction of Individual Truncation Libraries

The following protocol makes repeated use of the QIAquick DNA purification kit from QIAGEN. We have found that QIAquick buffer PB is a suitable buffer for inactivating Exo III.

    1. Equilibrate 250 μl QIAquick buffer PB at room temperature in a 1.5-ml tube (tube A).
    2. Into a 0.5-ml tube (tube B) add 2 μg DNA, 6 μl of 10× Exo III buffer, the desired amount of 1 M NaCl, and water to 60 μl. At 22°, the rate of Exo III digestion can be estimated as a function of NaCl concentration using the equation:

$$\text{Rate(bp/min)} = 48 \times 10^{-0.00644[\text{NaCl}]}$$

where the concentration of NaCl is in mM.
    3. Equilibrate tube B in a minifridge or thermocycler at 22°.
    4. At time = 0, add 200 units of Exo III to tube B and mix immediately.
    5. Begin removing 1-μl samples from tube B every 20–30 s and add to tube A. Mix tube A well. Because the rate of Exo III digestion is highly dependent on temperature, leave tube B open during sampling to avoid warming the tube by repeated handling. QIAquick buffer PB is slightly viscous; therefore, some liquid will stick to the side of the pipette tip when adding the truncation samples to tube A. Thus the final volume will be lower than the starting 250 μl of buffer PB plus the 60 μl of truncation samples.
    6. After all samples are taken, estimate the volume remaining in tube A and add 5× volume of QIAquick buffer PB.
    7. Follow the QIAquick protocol.
    8. In the final step of the QIAquick protocol, elute DNA from the column with 44 μl of QIAquick buffer EB.
    9. Add 5 μl of 10× mung bean nuclease buffer and 3 μl (3 units) of mung bean nuclease and incubate tube at 30° for 30 min. Occasionally, the amount of mung bean nuclease added has to be adjusted to ensure the formation of a high-quality library. The amount of mung bean nuclease to add can be determined experimentally using a single truncation time point. For example, digest 10 μg of DNA with Exo III for a sufficient amount of time to digest 300 bp. This single time point is then divided into five tubes with 2 μg of DNA per tube. Add 5 μl of 10× mung bean nuclease buffer to each tube and then add 5 μl of various amounts of mung bean nuclease (say 0, 1, 3, 6, and 12 units) to each sample. A portion of the DNA is then

run on a 0.8% agarose gel. DNA not digested by mung bean nuclease will run as a slightly diffuse band larger than the expected size due to the large single-stranded overhangs. As the level of mung bean is increased, the DNA will start to smear between the size obtained with no mung bean nuclease and the expected size. When the level of mung bean nuclease is optimum, the DNA will run as a focused band at the expected size. When there is too much mung bean nuclease, the DNA will smear to smaller sizes as the mung bean nuclease begins to make double-stranded breaks.

10. Add 250 μl of QIAquick buffer PB.

11. Follow the QIAquick protocol.

12. Elute truncated DNA from the QIAquick column with 82 μl of QIAquick buffer EB except for constructing pDIM-C8 PFC libraries, in which case elution is with 90 μl of QIAquick buffer EB.

For pDIM-N2 (both PFC and APR libraries) or for pDIM-C8 in the case of constructing APR libraries skip to step 17. Only perform steps 13–16 for pDIM-C8 in the case of constructing PFC libraries. These steps are necessary in preparing the ATG start codon for fusion to the truncated gene.

13. For pDIM-C8 only, add 10 μl of 10× NEB buffer 1, 1 μl of 100× BSA, 18 units of *Nco*I, and water to 100 μl. Incubate at 37° for 2 h.

14. Add 0.5 ml of QIAquick buffer PB.

15. Follow the QIAquick protocol to purify digested pDIM-C8.

16. Elute digested pDIM-C8 from the column with 82 μl of QIAquick buffer EB.

17. Equilibrate pDIM-N2 and pDIM-C8 at 37°.

18. For pDIM-N2, add 10 μl of Klenow mix [20 mM Tris–HCl, pH 8.0, 100 mM MgCl$_2$, 0.25 units/μl Klenow (exo$^+$)], incubate at 37° for 3 min, add 10 μl of dNTPs (0.125 mM each), and incubate at 37° for 5 min. For pDIM-C8, add 10 μl of dNTP, 10 μl of Klenow mix, and incubate at 37° for 5 min.

19. Inactivate Klenow by incubating at 72° for 20 min.

20. Cool to room temperature and add 0.4 ml of ligase mix (320 μl of water, 40 μl of 10× T4 DNA ligase buffer, 40 μl of 50% PEG 8000, and 18 Weiss units of T4 DNA ligase).

21. Incubate at room temperature ≥12 h.

22. Concentrate by ethanol precipitation. To 250 μl of ligation mixture add 125 μl of 7.5 M ammonium acetate, 750 μl of 100% ethanol (at −20°), and 2 μl Pellet Paint (Novagen) and incubate on ice for 30 min. Centrifuge at 10,000g at 4° for 10 min. Wash the DNA pellet with 750 μl of 70% ethanol (at −20°). After an additional 2-min spin and thorough removal of

all liquid by pipetting, air dry the DNA pellet for 10 min. Resuspend the DNA pellet in 15 μl water.

23. Electroporate $\leq$5 μl of DNA into 50 μl of electrocompetent DH5α-E *E. coli* cells (e.g., Bio-Rad Gene Pulser II set to 25 μF capacitance, 200 Ω resistance, and 2.5 kV). Allow cells to recover for 1 h in 1 ml of SOC media while shaking at 200 rpm at 37°.

24. Plate 1 μl of the cells on a small plate with growth media containing either 100 μg/ml ampicillin for pDIM-N2 or 50 μg/ml chloramphenicol for pDIM-C8. Plate the remaining electroporation mixture ($\sim$1 ml) on a large LB plate (e.g., 245 $\times$ 245-mm bioassay dish; Nalgene-Nunc) with the appropriate antibiotics. Colonies that grow on the small plate are used to determine the number of transformants (total transformants = number of colonies on the small plate $\times$ 1000) and will also be used for evaluating the truncation libraries.

25. Incubate at 37° overnight.

Evaluating the Individual Truncation Libraries

After creating a truncation library for pDIM-N2 or pDIM-C8, verify that the library has a random distribution in the expected size range. This can be performed by mini-prepping the DNA from individual transformants and performing restriction enzyme digests followed by analysis using agarose gel electrophoresis. It is more convenient to perform PCR on individual transformants using primers flanking the insertion site as described next.

1. Combine 70 μl of 10$\times$ *Taq* buffer (containing 15 mM MgCl$_2$, Promega), 56 μl of dNTPs (2.5 mM each), 42 μl of 10 μM forward primer, 42 μl of 10 μM reverse primer, 4 μl *Taq* DNA polymerase (20 units, Promega), and 486 μl water to create the PCR mixture.

2. Add 25 μl of PCR mixture per tube.

3. Transfer one colony per tube from the 1-μl plate using a sterile pipette tip.

4. Perform PCR for 30 cycles of 1 min at 94°, 1 min at 56°, and 1 min at 72°. Also do a control PCR on the original pDIM-N2 and pDIM-C8 (plasmid before truncation) using the same primers as described earlier. This will give an upper limit on the size of fragments obtained from the PCR.

5. Analyze the PCR reactions by agarose gel electrophoresis (Fig. 4 shows a sample result).

If results show that the library is biased toward too little truncation, reconstruct the library with a lower concentration of NaCl. If the library is
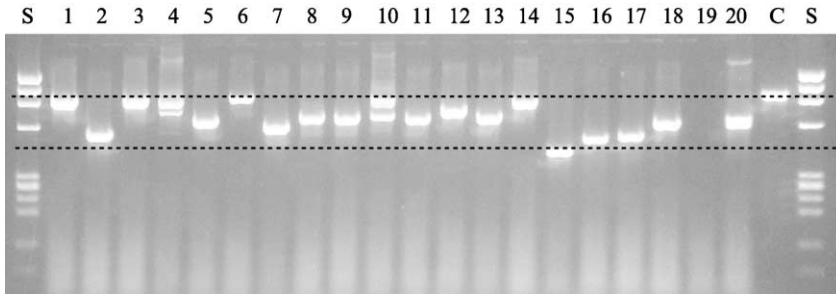
FIG. 4. PCR analysis of truncation-length diversity. Individual C-terminal library members were amplified by colony PCR using primers flanking the gene fragment to be truncated. The two standard lanes (S) are the ϕX174-*Hae*III-digested DNA marker (NEB). The control lane (C) is a PCR amplification of the vector before truncation has been performed. This untruncated fragment is the largest size that should appear in the library. Dotted horizontal lines give the desired size range for this particular truncation library. Note that lanes 4 and 10 each have two bands. Two bands can result from a single cell receiving two plasmids with different truncation lengths.

biased toward too much truncation, reconstruct the library with a higher concentration of NaCl. Note that too much truncation can also be caused by overdigestion with mung bean nuclease.

Recovering and Storing the Individual Truncation Libraries

Libraries are recovered from the large plate. One portion of the library is saved as frozen cell stocks and from another portion the library is isolated in plasmid form.

1. Recover cells from a 245 × 245-mm bioassay dish by adding 2 × 15 ml storage media [18 ml LB, 9 ml 50% glycerol, 3 ml 20% (w/v) glucose] to the top of the plate, scrape cells from media using a cell spreader, and then pipette cells into a 50-ml polypropylene centrifuge tube.

2. Spin cells in a centrifuge at 5000 rpm (maximum allowable speed for polypropylene centrifuge tubes) at 4° for 10 min.

3. Decant supernatant and add 2 ml of storage media.

4. Resuspend the pelleted cells by gentle shaking.

5. Store 4 × 200-μl aliquots in 1.5-ml tubes at −80°.

6. Pellet the remaining cells in a centrifuge at 5000 rpm at 4° and follow the protocol for the QIAGEN HiSpeed plasmid midiprep kit (or other midiprep kit) to recover the library in plasmid form.

Cotransformation of the Individual Truncation Libraries

The individual N-terminal and C-terminal truncation libraries are now ready to be cotransformed into *E. coli* strain DH5α–E. It should be noted that using too much DNA for the transformation could lead to problems when screening the final library. If too much DNA is used, multiple library members of either pDIM-N2 or pDIM-C8 can transform a single bacterial cell, thus complicating the analysis and identification of positives. Using 20 ng of DNA from each library for the cotransformation appears to minimize the occurrence of multiple copies of library members in the same cell while still resulting in cotransformed libraries of significant size.

1. Combine 20 ng of DNA from each library in a 0.5-ml tube and chill on ice. Add 50 μl of electrocompetent DH5α-E *E. coli* cells and electroporate (e.g., Bio-Rad Gene Pulser II set to 25 μF capacitance, 200 Ω resistance, and 2.5 kV). Allow cells to recover for 1 h in 1 ml of SOC media shaking at 200 rpm at 37°.

2. Plate 1 μl of the cells on each of three small LB plates (one with Cm, one with Amp, and one with Cm and Amp). Plate the remaining electroporation mixture (~1 ml) on a large LB plate (e.g., 245 × 245-mm bioassay dish; Nalgene-Nunc) with Cm and Amp. Colonies that grow on the small plate with both antibiotics are used to determine the number of transformants (total transformants = number of colonies on the small plate × 1000). Small plates with only one antibiotic are used to determine which plasmid is limiting the number of double transformants, should the number of Amp^{R}Cm^R transformants be small.

3. Recover the library from the large plate as before and store as a frozen cell stock.

Confirmation of True Positives

The library of heterodimers can now be screened or selected for the desired function. However, once positives are identified, these must be analyzed further to ensure that recombination has not occurred. It is important to separate the plasmids and retransform them together for each possible positive to ensure that (1) recombination producing a full-length gene has not occurred, (2) one of the protein fragments alone is not functional, and (3) the function is exhibited only when both halves of the heterodimer are present.

Recombination is the most frequent cause of false positives. Because the individual truncation libraries originated from large overlapping fragments, there is a high probability of finding large overlaps in the N-terminal and C-terminal genes in individual bacterial cells. Presumably these

overlapping DNA sequences sometimes lead to homologous recombination between the pDIM-N2 and the pDIM-C8 plasmids to reconstitute the original, full-length gene. This can occur at a significant frequency (about 3 in 10^4) even in $recA^-$ strains of *E. coli*. Therefore, positives must be screened further to ensure that recombination has not occurred. Although PCR (with forward and reverse primers for the starting gene), minipreps (most recombinants result in larger or abnormal size plasmids), and restriction digests can be used as a quick screen to eliminate recombinants, the most definitive test to ensure that a positive is a true heterodimeric positive is to isolate the individual plasmids, retransform each plasmid individually into fresh cells, and confirm the absence of the function. After it is confirmed that the individual plasmids do not produce a protein with the desired function, these plasmids are cotransformed together again and the presence of the function is confirmed. The truncated genes residing in the plasmids are then sequenced as a final confirmation.

1. Miniprep the plasmid DNA from 5 ml of an overnight inoculum of a positive clone.

2. Prepare a 1:100–1:1000 dilution of the miniprep. The miniprep needs to be diluted so that there is a high probability of transforming only one of the plasmids into a given cell.

3. Electroporate 5 μl of this diluted DNA into DH5α-E (e.g., Bio-Rad Gene Pulser II set to 25 μF capacitance, 200 Ω resistance, and 2.5 kV). Allow cells to recover for 1 h in 1 ml of SOC media shaking at 200 rpm at 37$°$.

4. Spread 5 μl of the electroporation mixture on a plate containing 100 μg/ml Amp and 50 μl of the electroporation mixture on a plate containing 50 μg/ml Cm. pDIM-C8 has a lower copy number than pDIM-N2, resulting in a lower frequency of cells transformed with pDIM-C8. Therefore, a greater volume needs to be plated to obtain a similar number of colonies.

5. Incubate plates at 37$°$ overnight.

6. Pick individual colonies from the Amp^R transformant plate and streak first onto a Cm plate and then immediately onto an Amp plate. If the colony contains cells with only the pDIM-N2 plasmid, nothing should grow when streaked onto the Cm plate. If there is growth on the Cm plate, either the cells contain both plasmids or a recombination has occurred (recombinants often contain a large plasmid with both antibiotic markers). Generally, when transforming this level of dilution, virtually 100% of Amp^R transformants will contain only pDIM-N2.

7. Perform the analogous streak test with colonies selected from the Cm^R transformant plate, except now streak first onto an Amp plate. If

there is growth on the Amp plate, either the cells contain both plasmids or a recombination has occurred. In general, greater than 90% of the Cm^R transformants will contain only pDIM-C8. This frequency is lower than the frequency of transformants containing only pDIM-N2, as pDIM-C8 is the lower copy number plasmid.

8. Pick a transformant that only has one antibiotic resistance and test for the desired function of the heterodimer. If the desired function is present, either the protein fragment maintains the function by itself or recombination has occurred such that a full-length gene is present on an individual plasmid.

9. If the function is not present for both separated plasmids, the final step is to retransform them together to confirm that the function requires the presence of both plasmids. Grow an inoculum and miniprep the plasmid DNA from about 5 ml of culture. Mix together 2 μl each of DNA from pDIM-N2 and pDIM-C8 minipreps, electroporate into the desired strain, and plate onto plates containing 100 μg/ml Amp and 50 μg/ml Cm.

10. Test these cotransformants for the desired function.