

Exonuclease III-Based High Throughput Construction of DNA Templates for *In Vitro* Expression

Stephanie Uder, Brian Ward, Kristen Opper, and Keming Song
 Research and Development, Sigma-Aldrich Biotechnology Division
 P.O. Box 14508, St. Louis, MO 63178



Abstract

High throughput gene expression is critical in genome-wide analysis of protein-protein interaction (PPI), functional domain mapping, mutagenesis, and protein array research. Traditional methods involve laborious gene cloning and *in vivo* expression requiring weeks of experimentation while *in vitro* methods of expression offer a rapid, automatable approach. Although a variety of *in vitro* transcription and translation systems are widely used, preparation of expression ready DNA templates (ERTs) has been a major bottleneck for high throughput applications. Here we present an exonuclease III-based ligation-mediated PCR assembly method (Exo-LMPCR) for high throughput construction of ERTs. The method involves generation of gene-specific PCR products using a specially formulated thionucleotide mix followed by controlled Exonuclease III digestion to produce cohesive ends. The first round PCR products are directionally ligated to adaptors containing sequence elements such as promoters, ribosomal binding sites, and affinity tags. Finally, second round PCR amplification of the assembled construct is conducted to generate ERTs for direct use in *in vitro* transcription and translation. Exo-LMPCR was characterized using FLAG[®] and c-Myc affinity tags and over 96 mammalian genes ranging in size from 239 to 2207 bases. Exo-LMPCR ERTs expressed in commercial *in vitro* expression systems (rabbit reticulocyte and wheat germ) were functionally active as determined by protein-protein interaction and functional domain mapping studies. Furthermore, this technology was integrated with a 96-well plate-based affinity capture and detection system for analysis of PPIs. The entire system enables investigators to rapidly generate ERTs, express interacting proteins *in vitro*, and capture and detect interactive partners using ELISA. Exo-LMPCR surpasses existing methods due to the combined properties of being universal and directional without reliance on base recognition enzymes, being streamlined for superior flexibility in control of 3' and 5' end adaptors, and because the system is extremely robust due to specific assembly and efficient ligation.

Introduction

Standard methods of generating protein for function studies involve traditional cloning of open reading frames (ORFs) into vectors for expression in appropriate hosts. *In vitro* expression in current systems such as rabbit reticulocyte lysate, wheat germ extract, and *Escherichia coli* lysate may overcome many of the drawbacks associated with *in vivo* expression. *In vitro* expression may be accomplished in hours versus the days and weeks required for *in vivo* methods. Most genes can be expressed using one promoter and at levels adequate for protein function studies, thus facilitating expression of a large number of genes in high throughput studies. *In vitro* expression may also overcome difficulties often encountered in *in vivo* systems where the over-expressed protein is toxic to the cell, the protein is insoluble or forms inclusion bodies, or the protein undergoes rapid proteolytic degradation due to intracellular proteases. Due to its speed, simplicity, and universality, *in vitro* methods are increasingly attractive tools for protein function studies.

Exo-LMPCR offers a solution to the bottleneck of generating expression ready DNA templates or ERTs for *in vitro* expression in a high throughput format. Templates for coupled or uncoupled *in vitro* transcription/translation systems need only to be linear segments of DNA containing a T7 promoter for transcription and a ribosomal binding site such as Kozak or Shine-Dalgarno 5' to the ORF for initiation of protein translation. An affinity tag such as FLAG[®] or c-Myc is usually desired for convenient purification and detection of the recombinant protein. There are a number of ways to create linear ERTs containing all required elements through recombinatorial PCR methods. Drawbacks of current methods include long primers associated with PCR assembly and use of restriction enzymes and other base recognition enzymes to create cohesive ends for ligation-mediated PCR assembly. Exo-LMPCR utilizes technology based on Exonuclease III. ORFs are amplified in the presence of a mixture of dNTPs containing thiouridines, dATP_S and dGTP_S. Exonuclease III, a 3'-5' phosphoric diester hydrolase, sequentially degrades linear duplex DNA from the 3' termini. Because Exonuclease III cleaves phosphorothioates extremely slowly, 5' termini are produced that are compatible with cohesive ends produced by restriction enzymes. This property of Exonuclease III is exploited here to directionally ligate PCR products to 5' and 3' adaptors containing elements for *in vitro* expression. Exo-LMPCR completely eliminates concern of internal restriction sites and provides total freedom in choosing sequences for cohesive ends providing a high throughput system of generating ERTs for *in vitro* protein expression.

Results

Overview of Exo-LMPCR

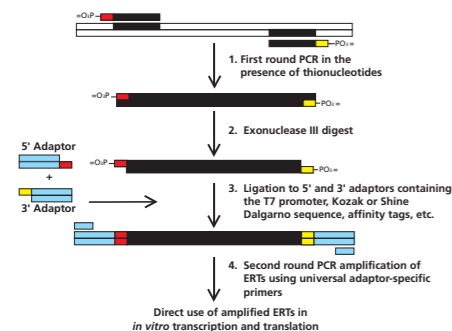


Figure 1. Schematic diagram of exonuclease III-based assembly of DNA templates for use in *in vitro* transcription/translation. (1) ORFs undergo 1st round PCR amplification in the presence of a dNTP mix containing dATP_S and dGTP_S using gene specific primers containing a 5' AGCTT overhang on the sense primer and a 5' GATCT overhang on the anti-sense primer. These overhangs correspond to the cohesive sequences of *Hind* III and *Bgl* II respectively. (2) Amplification products immediately undergo a 10 minute reaction with Exonuclease III, which sequentially removes bases from 3' termini, in order to reveal the *Hind* III and *Bgl* II cohesive sites introduced by the primer. (3) Exonuclease III digested products are then ligated with selected 5' and 3' adaptors for 10 minutes. (4) The assembled template is then amplified by 2nd round PCR using universal adaptor-specific primers for direct use in transcription/translation reactions.

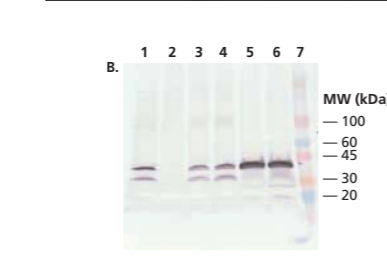
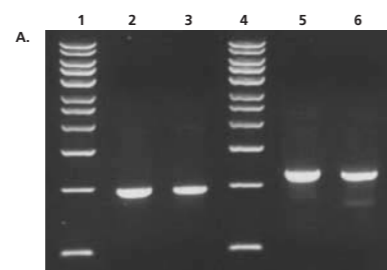


Figure 2. Assembly of FLAG-p53 and FLAG-Ik-B-α DNA templates (A) for *in vitro* expression (B). (A) Human p53 and Ik-B-α were amplified via 1st round PCR (lanes 2 and 3) in a specially formulated dNTP mix containing dGTP_S and dATP_S using gene specific primers containing overhangs complementary to *Hind* III and *Bgl* II cohesive ends. After digestion with Exonuclease III, the first round product was ligated to a 5' adaptor containing a T7 promoter, Kozak sequence, and FLAG affinity tag and to a 3' adaptor containing a stop codon and the CMV-24 priming sequence. The assembled DNA template was then amplified in second round PCR (lanes 5 and 6) using universal adaptor-specific primers. Lanes 1 and 4 contain 1kb ladder. The second round PCR products were used directly in a rabbit reticulocyte coupled transcription and translation system. (B) Translated samples were analyzed by SDS-PAGE/Western analysis, colorimetric detection using monoclonal M2 anti-FLAG[®]-AP and NBT/BCIP substrate. Lane 1 corresponds to a positive control which consists of a second round FLAG-p53 template cloned into a pETBlue-2a vector. Amplicons from this vector, using the universal adaptor-specific primers, were used as a control to represent a reaction from one species versus a potential mixture. Lane 2 is a negative control consisting of rabbit reticulocyte lysate minus template. Lanes 3 and 4 are *in vitro* expressed FLAG-p53 and lanes 5 and 6 are *in vitro* expressed FLAG-Ik-B-α. Lane 7 contains the ColorBurst[™] protein molecular weight markers (C4105).

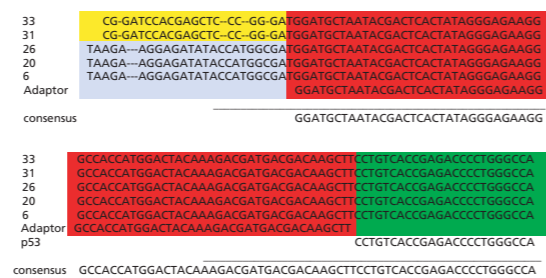


Figure 3. Sequence analysis of adaptor/gene junctions to demonstrate specificity of assembly. As indicated in figure 2, second round PCR amplicon FLAG-p53 was cloned into a pETBlue-2a blunt vector. Recombinant clones were selected by colony PCR using internal p53 primers. Clones 33, 31, 26, 20, and 6 (only known to contain 300 bases of p53) were sequenced using vector specific sequencing primers. Sequencing through the 5' and 3' adaptor/gene junctions revealed all sequences were intact and correctly assembled. Above, the red blocks represent the 5' T7-FLAG adaptor region while the blue represents the start of the p53 gene for the six clones analyzed. The yellow and green blocks represent the pETBlue-2a vector (nondirectional cloning).

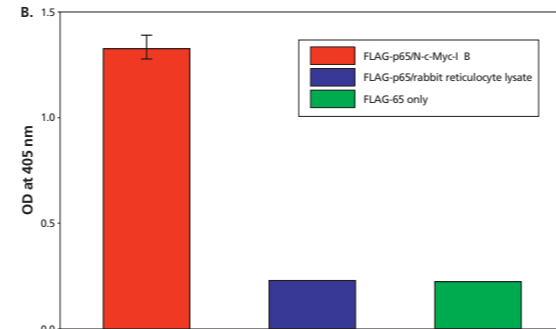
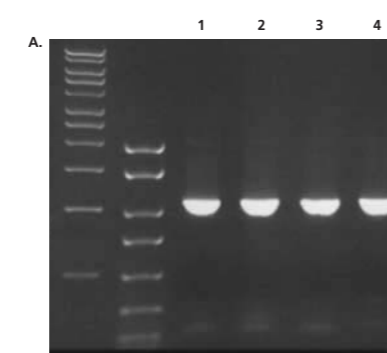
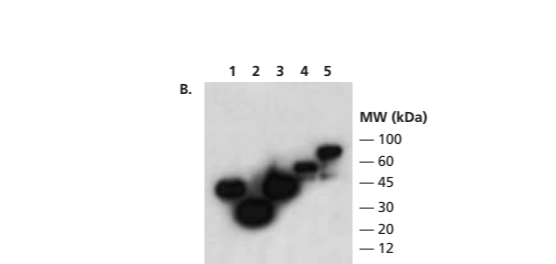
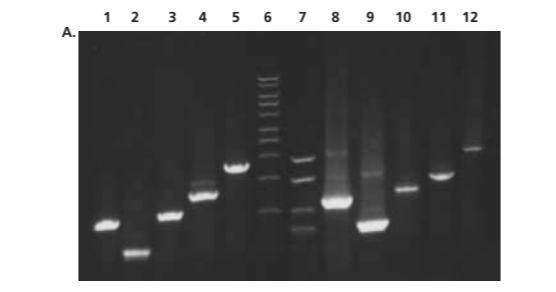


Figure 4. Assembly of N- and C-terminal tagged templates (A) and functional assay of *in vitro* translated N-c-Myc-Ik-B-α template (B). Human Ik-B-α was amplified in a dNTP mix containing dGTP_S and dATP_S using gene specific primers with *Hind* III and *Bgl* II overhangs. After Exonuclease III digestion of the first round product, the Exo-treated Ik-B-α was ligated to four different adaptor combinations (5' T7-Kozak-FLAG-*Hind* III/3' *Bgl* II-Stop-CMV-24, 5' T7-Kozak-c-Myc-*Hind* III/3' *Bgl* II-Stop-CMV-24, 5' T7-Kozak-*Hind* III/3' *Bgl* II-FLAG-CMV-24, 5' T7-Kozak-*Hind* III/3' *Bgl* II-c-Myc-CMV-24). The assembled templates were amplified by second round PCR using universal adaptor primers as seen in (A) lanes 1-4 which correspond to N-FLAG-Ik-B-α, N-c-Myc-Ik-B-α, Ik-B-α-C-FLAG and Ik-B-α-C-Myc. (B) The N-c-Myc-Ik-B-α ERT from second round PCR was used directly in coupled transcription/translation in a rabbit reticulocyte system. Cos-7 cells were transfected with FLAG-p65-CMV-2 plasmid. FLAG-p65 from the lysate was then incubated with the captured FLAG-p65. Protein-protein interaction was measured by detection of N-c-Myc-Ik-B-α using monoclonal Anti-c-Myc-AP in ELISA format.



NCBI Accession Number	HeLa Gene	Size (bp)
X66365	PLSTIRE for serine/threonine protein kinase	980
X16937	phosphoinositide-dependent kinase type II beta subunit	647
Z11695	40 kDa protein kinase related to rat ERK2	1046
L31951	human protein kinase (JNK 2)	1274
L19559	protein kinase HSTPK13	1811

Figure 5. High throughput modeled assembly (A) and translation (B) of N-FLAG HeLa genes (C). Human kinase genes of varying size were amplified from HeLa cDNA using gene-specific primers containing *Hind* III and *Bgl* II overhangs in the presence of a dNTP mixture containing dGTP_S and dATP_S. First round amplicons are shown in (A) lanes 1-5: 1) X66365 2) X16937 3) Z11695 4) L31951 5) L19559. Exonuclease III treated first round PCR products (5 μl, varying amounts of DNA) were ligated to 2 ng each adaptor of 5' T7-Kozak-FLAG-*Hind* III/3' *Bgl* II-Stop-CMV-24 adaptor set in a 20 μl reaction using the Quick-Link[™] Ligation Kit (LIG-2). Second round amplification using universal adaptor-specific primers was performed to amplify assembled templates in the ligation reaction (1 μl ligation mix in 50 μl reaction). Second round PCR amplicons of assembled ERTs are shown in (A) lanes 8-12: 8) X66365 9) X16937 10) Z11695 11) L31951 12) L19559. 5 μl of each amplified ERT were then directly used in coupled transcription/translation in rabbit reticulocyte lysate. Translated FLAG fusion-proteins were captured on and eluted from an Anti-FLAG M2 96-well affinity capture plate followed by western detection with Anti-FLAG M2-HRP shown in (B) lanes 1-5: 1) X66365 2) X16937 3) Z11695 4) L31951 5) L19559.

High Throughput Modeled Assembly of 93 Mammalian ORFs

Targets ranging from 239 to 2207 base pairs

Sample Set	Template	Number Correct Size	Percent Correct Size
First Round PCR	HeLa cDNA	81/93	87%
Second Round PCR	Assembly using first round product	67/81	83%

Table 1. High throughput assembly of mammalian ERTs. High throughput modeled assembly was performed for 93 HeLa genes as indicated in Figure 5. A band of the correct size was obtained in second round PCR for 83% of assembly reactions where a band of the correct size was obtained in first round amplification from cDNA. The number reflects 72% of the original target number. The targets ranged in size from 239 to 2207 bases.

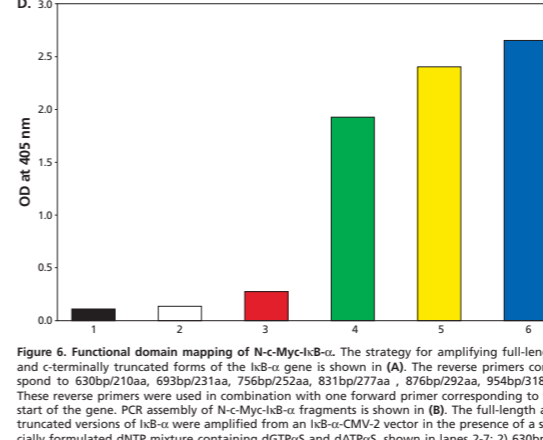
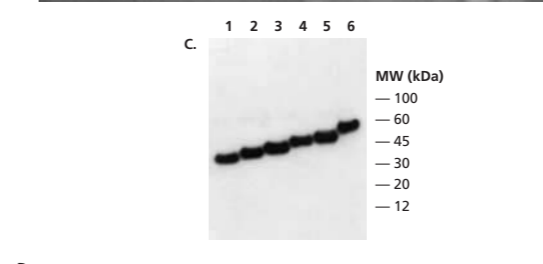
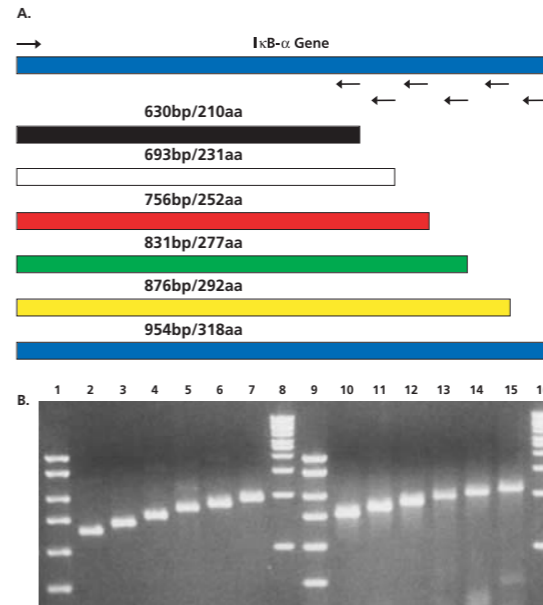


Figure 6. Functional domain mapping of N-c-Myc-Ik-B-α. The strategy for amplifying full-length and C-terminally truncated forms of the Ik-B-α gene is shown in (A). The reverse primers correspond to 630bp/210aa, 693bp/231aa, 756bp/252aa, 831bp/277aa, 876bp/292aa, 954bp/318aa. These reverse primers were used in combination with one forward primer corresponding to the start of the gene. PCR assembly of N-c-Myc-Ik-B-α fragments is shown in (B). The full-length and truncated versions of Ik-B-α were amplified from an Ik-B-α-CMV-2 vector in the presence of a specially formulated dNTP mixture containing dGTP_S and dATP_S shown in lanes 2-7: 2) 630bp 3)

693bp 4) 756bp 5) 831bp 6) 876bp 7) 954bp. After treatment with Exonuclease III, the fragments were ligated to the N-c-Myc adaptor set (5' T7-Kozak-c-Myc-*Hind* III/3' *Bgl* II-Stop-CMV-24). Second round amplification of assembled templates with universal primers is shown in lanes 10-15 (bases correspond only to Ik-B-α portion of amplicon, additional bases added by adaptor set): 10) 630bp 11) 693bp 12) 756bp 13) 831bp 14) 876bp 15) 954bp. Full length and truncated N-c-Myc-Ik-B-α fragments were then *in vitro* transcribed/translated in a coupled rabbit reticulocyte lysate system, purified on an Anti-c-Myc 96-well affinity capture plate, eluted and detected by western blot with monoclonal Anti-c-Myc/rabbit anti-mouse-HRP as shown in (C): 1) 210aa 2) 231aa 3) 252aa 4) 277aa 5) 292aa 6) 318aa. The functional domain-mapping experiment (D) was performed by capturing FLAG-p65 expressed in Cos-7 cells on an Anti-FLAG M2 96-well affinity capture plate. N-c-Myc-Ik-B-α fragments were then incubated with the anchored FLAG-p65. Protein-protein interaction was detected in ELISA format using monoclonal Anti-c-Myc-AP to detect captured N-c-Myc-Ik-B-α fragments as shown in (D): 1) 210aa 2) 231aa 3) 252aa 4) 277aa 5) 292aa 6) 318aa. The necessary region for interaction with p65 appears to map to the region between 252aa and 277aa, thus quickly narrowing the region for future studies to a 25 amino acid region. The entire domain-mapping study from cDNA to ELISA results was completed in 1.5 days.

Conclusions

- Exo-LMPCR provides a unique universal system to directionally insert multiple ORFs into adaptor sets containing all elements required for *in vitro* transcription/translation.
- The use of thionucleotides and Exonuclease III to create cohesive ends in Exo-LMPCR eliminates concern of internal restriction sites that are problematic in other directional methods.
- Exo-LMPCR adaptors offer flexibility and allow one to generate an array of tagged fusion proteins (n-terminal, c-terminal, or both) needed for downstream applications.
- Use of the system allows researchers to go from cDNA to expressed protein in only 1 day.
- The approach is streamlined and automatable allowing one to generate protein for many genes in a high throughput format in a fraction of the time required for standard methods.
- The system has been shown effective for a large number of genes as well as compatible with different *in vitro* systems such as coupled, uncoupled, and linked transcription/translation, rabbit reticulocyte lysate and wheat germ extract.
- Exo-LMPCR assembly is extremely robust and accurate. One assembly reaction can provide enough template for >100 second round amplifications which in turn provides enough ERT for >1000 transcription/translation reactions. Sequence analysis thus far has shown 100% accuracy in the adaptor/gene junctions.
- The Exo-LMPCR system overcomes the bottleneck of being able to generate DNA templates for *in vitro* transcription/translation in a high throughput format. This method, combined with *in vitro* expression and multi-well affinity capture, makes it feasible to express a large number of genes and to study their function at a speed unachievable by existing technologies.

Acknowledgements

We would like to acknowledge Ron Hernan, Tom Juehne, Chuan Li, and Henry George for their help in this work.

References

1. Zhang, L., S. Uder, T. Juehne, B. Brizzard, and K. Song. 2002. Nonradioactive assay of FLAG-tagged MAPK using anti-FLAG antibody coated multiwell plates. *Biotechniques* 32: 442-447.
2. Putney, S.D., S.J. Benkovic and P.R. Schimmel. 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.
3. Li, C. and R.M. Evans. 1997. Ligation independent cloning irrespective of restriction site compatibility. *Nucleic Acids Res.* 25: 4165-4166.
4. Kaluz, S., K. Kolbe and K.B.M. Reid. 1992. Directional cloning of PCR products using exonuclease III. *Nucleic Acids Res.* 20: 4369-4370.
5. Ohuchi, S., H. Nakano and T. Yamane. 1998. *In vitro* method for the generation of protein libraries using PCR amplification of a single DNA molecule and coupled transcription/translation. *Nucleic Acids Res.* 26: 4339-4346.

Related Sigma Products

Catalog #	Product Description
D1313	JumpStart [™] REDAccuTag [™] LA DNA Polymerase Mix
E1131	Exonuclease III from <i>Escherichia coli</i> BE25 /psGR3
RDC-1	Director Universal Cloning System
LIG-2	Quick-Link [™] DNA Ligation Kit
P2983	ANTI-FLAG [®] High Sensitivity, M2 coated 96-well plates
P2241	Anti-c-Myc Coated HS 96-well Plates

Corresponding Author

Stephanie Uder
 R&D Scientist
 Sigma-Aldrich Biotechnology
 P.O. Box 14508
 St. Louis, MO 63178
 E-mail: suder@sial.com
 Phone: (314) 289-8496 ext. 1458
 Fax: (314) 286-7645