

Enzyme Free Cloning for high throughput gene cloning and expression

Rob N. de Jong · Mark A. Daniëls ·
Rob Kaptein · Gert E. Folkers

Received: 21 August 2006 / Accepted: 14 December 2006 / Published online: 13 February 2007
© Springer Science+Business Media B.V. 2007

Abstract Structural and functional genomics initiatives significantly improved cloning methods over the past few years. Although recombinational cloning is highly efficient, its costs urged us to search for an alternative high throughput (HTP) cloning method. We implemented a modified Enzyme Free Cloning (EFC) procedure, a PCR-only method that eliminates all variables other than PCR efficiency by circumventing enzymatic treatments. We compared the cloning efficiency of EFC with that of Ligation Independent Cloning (LIC). Both methods are well suited for HTP cloning, but EFC yields three times more transformants and a cloning efficiency of 91%, comparable with recombinational cloning methods and significantly better than LIC (79%). EFC requires only nanogram amounts of both vector and insert, does not require highly competent cells and is, in contrast to LIC, largely insensitive to variations in PCR product concentration. Automated protein expression screening of expression strains directly transformed with EFC reactions showed, that the traditional preceding step via a cloning strain can be circumvented. EFC proves an efficient and robust HTP cloning method, that is compatible with existing Ligation Independent Cloning vectors, and highly suitable for automation.

Keywords Ligation independent cloning · Enzyme free cloning · High throughput · Functional genomics · Recombinant protein expression · PCR-only cloning

R. N. de Jong · M. A. Daniëls · R. Kaptein ·
G. E. Folkers (✉)
Bijvoet Center for Biomolecular Research, Utrecht
University, Padualaan 8, 3584 CH, Utrecht,
The Netherlands
e-mail: g.folkers@chem.uu.nl

Introduction

The overwhelming amount of genome sequences available and the large proportion of proteins without reliable functional annotation, have led to a genome wide functional analysis, generally referred to as functional genomics or proteomics [1, 2]. In spite of the lower accuracy of such high throughput (HTP) methods, a wealth of information on gene function and functional relationships between gene products has accumulated [3–6]. Recently, methods for the production of smaller amounts of recombinant protein for whole proteome analysis have become available [7]. The demand for large amounts of soluble protein for structural and functional genomics boosted technological developments in recombinant protein production in *Escherichia coli* over the past few years [8–10]. Most steps from gene to soluble protein can at present be performed using a liquid handling station, which dramatically increased the throughput and reduced the costs through miniaturization [11]. These small-scale expression-screening experiments predict sufficiently well the production in larger quantities generally required for downstream applications [12]. In an ongoing effort to automate small-scale expression screening, we searched for a cheap and efficient cloning method that is robust enough to directly transform expression strains with cloning reactions, thereby circumventing the intermediate isolation of plasmid DNA from an *E. coli* cloning strain.

Traditionally, molecular cloning employs restriction enzymes (RE) to digest DNA, which is enzymatically joined using T4 DNA ligase. Although RE mediated cloning is rather time consuming and needs a lot of manual interference, it can be used for HTP cloning

[13]. Recombinational cloning methods have been successfully implemented in the HTP cloning of PCR fragments and are particularly advantageous if a specific gene has to be cloned in many different vectors [14]. Disadvantageous can be the presence of a recombination site as part of the open reading frame, which can affect protein function. Recombination sequences are not translated, when placed outside the open reading frame, but this requires large PCR primers with Shine–Dalgarno or Kozak sequences between the recombination site and the genespecific termini, and it reduces the flexibility. Also, the need to construct an entry clone prior to preparation of an expression construct and the costs associated with these methods are drawbacks when applied to HTP protein expression in an academic environment.

A good alternative to recombinational cloning is Ligation Independent Cloning (LIC) [15, 16]. This procedure uses the exonuclease activity of T4 DNA polymerase to generate 5' extended complementary cohesive ends in both the vector and the PCR product. A disadvantage of this method is the strict requirement for correct T4 treatment of both insert and vector. Furthermore, it is not entirely cloning vector and sequence independent, because it relies on terminal cloning sequences of 12–15 nucleotides lacking one nucleotide type to generate complementary ends.

The hetero-stagger PCR method circumvents enzymatic treatments of PCR products by generating two PCR products that each contain a cohesive end for a specific restriction site on either the 5' or the 3' end of the PCR product [17, 18]. This method was successfully applied in the HTP screening of soluble recombinant proteins [19]. Using hetero-stagger PCR cloning with longer cohesive tails, Tillett and Neilan reported an Enzyme Free Cloning (EFC) method that does not depend on restriction sites [20, 21]. The absence of enzymatic manipulations after PCR amplification suggested that EFC could be useful in automated HTP applications. Two disadvantages explain why EFC is not a commonly used method. It is not straightforward to prepare high quality vector DNA by PCR around the vector on the large scale needed for HTP applications, and the need for two sets of gene-specific primers creates additional costs. We can overcome the first disadvantage by combining LIC-treated vectors with EFC-created PCR products. Parallel oligonucleotide synthesis led to a significant cost reduction, making this modified EFC method now a potentially attractive HTP cloning method.

To evaluate the utility of EFC in HTP cloning in an unbiased pairwise comparison with LIC, we applied both methods to clone 24 different PCR products into

identical LIC compatible vectors. All PCR products can be cloned reliably by both methods, but with higher efficiency by EFC. We show that this cloning method is fast, does not require gelpurification of PCR products and requires only nanogram amounts of both insert and vector DNA in combination with house-made chemically competent cells. Furthermore, it can be completely performed using a liquid handling station or multichannel pipettes, demonstrating the feasibility of EFC as a HTP cloning method.

Material and methods

Plasmids

The plasmid pLICHIS was obtained by cloning in the NcoI and BamHI sites of pET15b two annealed oligonucleotides with the sequences: 5'CATGGGCAGCAGCCATCATCATCATCATCACAGCAGCCATA TGCTGGTGCCGCGCGGCAGCCGCGGGGTTCT TCTTGAG, and 5'GATCCTCAAGAAGAACCCCG CGGCTGCCGCGCGGCACCAGCATATGGCTGC TGTGATGATGATGATGATGGCTGCTGC. GB1, Trx and GST fusions were constructed by inserting the PCR fragments of the respective fusion tags amplified from respectively pTrx-Fus (Invitrogen), pGEX2T (GE-Healthcare), and pTH3 [22] in the NcoI-NdeI sites of pLICHIS. Dual tag plasmids HIS-GST, HIS-Trx and HIS-GB1 were made by insertion of a ds oligonucleotide with NcoI overhang in the NcoI digested fusions. The S-tag fusion was constructed by inserting two annealed oligonucleotides with the sequences 5'TATGAAAGAAACCGCTGCTGCTAAATTTCGA ACGCCAGCACATGGACAGCCTGG TGCCGCGC GGCAGCCGC, and 5'GGCTGCCGCGCGGCAC CAGGCTGTCCATGTGCTGGCGTTTCGAATTTA GCAGCAGCGTTTCTTTCA in the NdeI and SacII digested pLICHIS. All constructs were sequence verified. Details on construction of these vectors are available upon request.

EFC

For PCR amplification, two sets of primers were designed, (Eurogentec, genomic primers, at 10 or 50 nmol without further purification) where the gene-specific sequence is extended either with GCCGCGCGGCAGCCTG-gene-specific (LICFW) or TG-gene-specific (FW) as forward primers and CA-AGAAGAACCCCTCA-gene-specific (LICRV) or TCA-gene-specific (RV) as reverse primer respectively. PCR products were amplified in two separate

tubes, one with LICFW combined with RV, and one with FW combined with LICRV, using cDNA prepared as described before [12]. When plasmid DNA was used as PCR template, it was digested with the methylation dependent restriction enzyme DpnI (MBI Fermentas) after PCR amplification. To remove primer dimers and non-consumed primers, PCR products were purified using the AMPure PCR purification method (Agencourt) according to instructions of the company using either multichannel pipettes or a Hamilton Star liquid handling station. About 1–100 ng of the two PCR products were mixed and incubated for 5 min at 95°C and slowly cooled to 25°C in at least 5 min (PCR block or water bath). This hybridized PCR product was mixed with 2 ng of T4 treated pLICHIS (as described under Ligation Independent Cloning) in a total volume of 5 μ l, incubated at room temperature for 5 min, and transferred to an ice bath. A total of 25 μ l of DH5 α competent cells (CaCl₂ method [23], 1·10⁵–3·10⁶ colonies/ μ g plasmid DNA) was added to the reaction mixture, incubated on ice for 30 min, heat shocked for 60 s at 42°C, and recovered for 30 min at 37°C in 75 μ l LB. A total of 60 μ l of the transformation mixture was transferred to 12-well LB plates containing 50 μ g/ml ampicillin and distributed equally by twisting the plates, followed by 15 min drying at 37°C without lid. During transformations in the expression strain BL21(DE3) Rosetta2-LysS (Novagen), the transformation mixture was plated on LB amp plates containing 35 μ g/ml chloramphenicol. Protein expression and detection was performed as described before [12].

Ligation independent cloning

pLICHIS was digested with SacII (NEB) or Cfr42I (MBI Fermentas) for 4–16 h and dephosphorylated using CIAP (MBI Fermentas), and gelpurified using the QiaexII gel extraction kit (Qiagen). A total of 500 ng of vector was treated with 0.3 units of T4 DNA polymerase (NEB) in the presence of 2.5 mM dTTP in T4 polymerase buffer (33 mM TrisAc pH 7.9, 10 mM MgAc, 66 mM KAc, 0.1 mg/ml BSA, 0.5 mM DTT), incubated for 25 min at 20°C and subsequently heat inactivated for 20 min at 70°C. PCR products were PCR amplified using the LICFW and LICRV primer and purified as described above. A total of 100 ng of PCR product was treated with 0.3 unit of T4 DNA polymerase (NEB) as described for the vector but with 2.5 mM dATP instead of dTTP, in T4 polymerase buffer. About 1–100 ng of T4 treated insert and 1–3 ng of T4 treated vector were mixed together and transformed to DH5 α as described above.

Results

Experimental setup

An ideal cloning method should require small amounts of insert and vector DNA, and be cheap, effective, easily automatable and insensitive to variations in DNA concentrations [24]. Ligation independent cloning [15] meets most of these criteria, but we observed substantial variation in cloning efficiency. Parameters influencing efficiency are the amount of PCR product, purity of the PCR product, exact temperature of T4 DNA polymerase treatment and T4 DNA polymerase activity differences, caused by enzyme batch variations and activity loss over time (unpublished results). As a result, PCR products can be under- or over-treated by the exonuclease activity of T4 DNA polymerase, which can significantly influence the cloning efficiency.

The Enzyme Free Cloning method we implemented here, deviates from the EFC published before [21]. We use a LIC reaction [15] to generate the vector overhangs instead of a vector-long PCR reaction, which enabled us to make much larger preparations of EFC compatible vector as needed for constant HTP cloning quality. We use the hetero-stagger PCR method for gene amplification [17] to create cohesive termini complementary to the T4 treated vector and circumvent the T4 treatment of PCR product needed for LIC. This combination of methods effectively should eliminate all variables during cloning except for the PCR efficiency. The EFC method requires two sets of gene-specific PCR primers, since one forward and one reverse primer are extended with a sequence complementary to the termini of the T4 treated vector. Two separate PCR reactions are performed, one with the extended forward primer (1) and a normal reverse primer (3), and one with a normal forward (2) and an extended reverse primer (4). The two PCR products are purified, mixed, denatured and renatured, creating cohesive termini in which 25% of the possible combinations are compatible with the T4 treated vector (Fig. 1d).

In this setup, we can use an identical vector for both EFC and LIC, which allows us to do a pairwise comparison of the two cloning methods that is unbiased by the chosen vector. We constructed a LIC compatible vector pLICHIS (Fig. 1b) and derived various fusion proteins by inserting GST, Trx or GB1 in between the HIS-tag and the LIC site. All plasmids have a pET15b plasmid backbone, 14 basepairs of the thrombin cleavage site that are a LIC compatible sequence because they lack thymidines, followed by a SacII site and a 13 bp reverse LIC site lacking an adenosine. A digestion with SacII and exonuclease treatment using

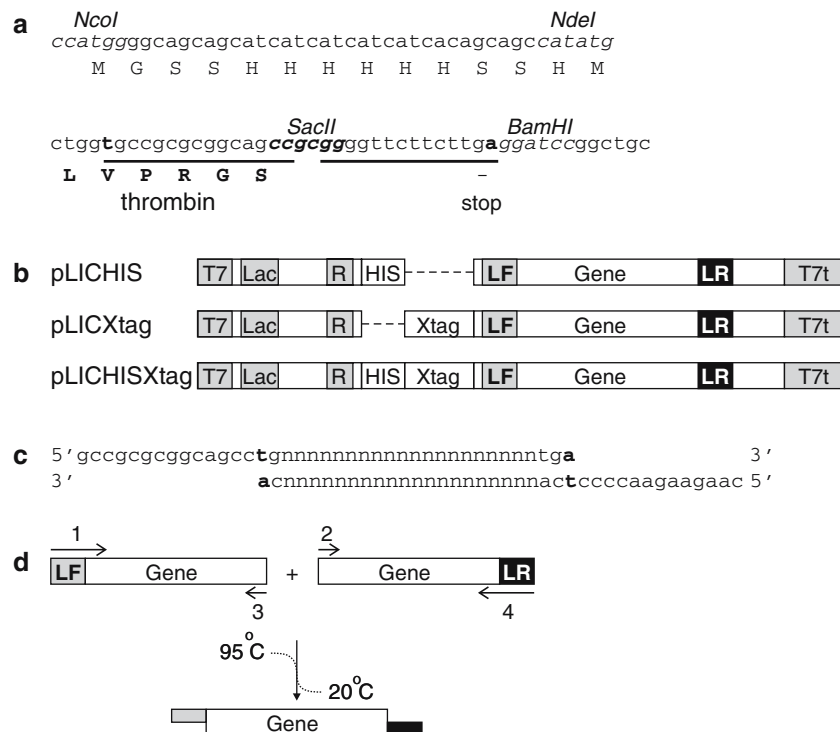


Fig. 1 Plasmid and cloning procedure. **(a)** Nucleotide sequence of cloning site of pLICHIS and restriction sites used for construction of the various fusion tags. The plasmid sequence outside the *NcoI* and *BamHI* fragment is identical to pET15b. Translated protein is indicated below, with the thrombin cleavage site in bold. The LIC extensions are underlined and the dTTP nucleotide or its complement incorporated by T4 DNA polymerase is indicated in bold. The *SacII* site used for linearization prior to T4 treatment of the vector is indicated. **(b)** Schematic representation of the plasmids constructed for small scale expression screening. X-tag refers to either a GST-tag (pLICGST), Thioredoxin (pLICTrx) or Gb1 (pLICGb1) tag, or a combination of the HIS tag with a GST tag (pLICHISGST), a Thioredoxin tag (pLICHISTrx), a Gb1 tag (pLICHISGb1), or an S-tag (pLICHISS). Indicated are the T7 promoter (T7), T7

terminator (T7t), lac operon (Lac), ribosomal binding site (R), LIC forward site (LF) and LIC reverse site (LR). Gene refers to the inserted gene sequence, while the dashed lines indicate the absence of sequences present in the other vectors. **(c)** Overhang of the PCR amplified target protein with the dATP nucleotide and its complement that will be incorporated by T4 DNA polymerase indicated in bold. **(d)** For EFC, a LIC compatible overhang is created using two different primer sets as described in detail in the Materials and Methods section. The LIC sequence extended primers 1 and 4 and the non-extended primers 2 and 3 are used for PCR amplification as schematically depicted. After mixing, denaturing and reannealing of the two PCR products, 25% of the PCR products contain a LIC-compatible overhang

T4 DNA polymerase in the presence of dTTP generates a non-self complementary overhang, identical for all vectors, that permits efficient cloning of T4 treated PCR products in vectors with different fusion tags (Fig. 1).

Enzyme free cloning is a robust cloning method

Various amounts of PCR products were either treated with T4 DNA polymerase for the LIC method, or mixed for EFC and hybridized with increasing amounts of the same T4 treated vector resulting in a molar insert/vector ratio ranging from 2 to 2000. A vector concentration dependent increase in the amount of colonies was observed for both methods (Fig. 2a, b). While for EFC increasing amounts of both vector and insert stimulated colony formation, remarkably, we

observed that LIC displayed a significant drop in cloning efficiency at higher insert concentrations. The insert concentration dependence for LIC is probably caused by an incorrect ratio of T4 polymerase and PCR product during exonuclease treatment. To test this hypothesis, a T4 treated PCR product was diluted, and various amounts were hybridized with different amounts of vector. Under these circumstances, LIC cloning efficiency increased with insert concentration similar to EFC (Fig. 2c), showing that the efficiency of T4 treatment is significantly influenced by the PCR product concentration. Under all conditions EFC resulted in more colonies than LIC, with an average increase of 2.6 ± 1.3 fold more colonies per ng of PCR product. The overall number of correct cloning events was also larger for EFC than for LIC (see below, Fig. 3b).

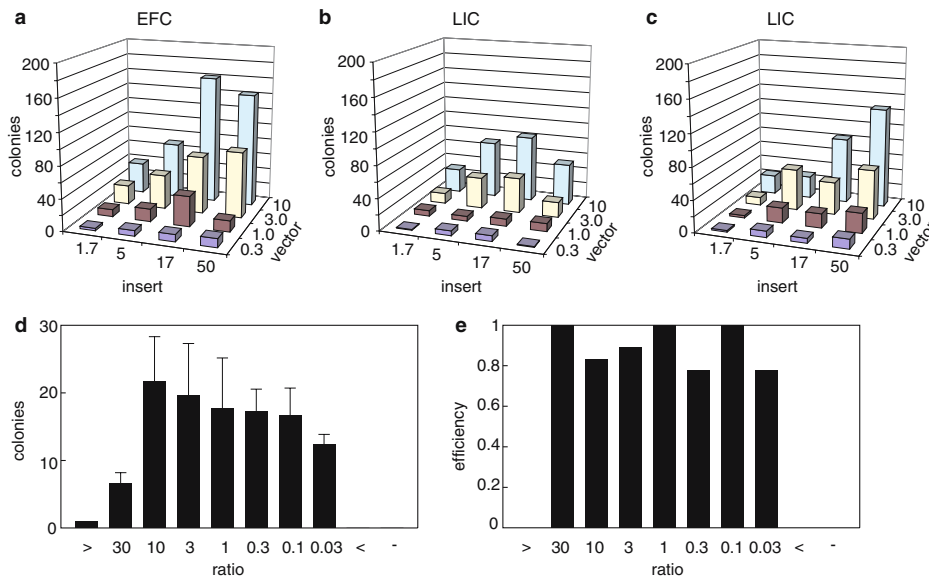


Fig. 2 Cloning efficiency for LIC and EFC. **(a–c)** Number of colonies obtained after transformation of EFC or LIC reactions in DH5 α with the indicated amount of insert and vector, using the same insert prepared by **(a)** EFC reaction or by LIC reaction with T4 treatment using the same amount of T4 DNA polymerase either **(b)** after dilution or **(c)** before dilution of the PCR products. Representative experiments are shown. **(d, e)** Influence of the ratio between the two different PCR products on the EFC efficiency. **(d)** The number of colonies obtained after transformation of EFC reactions in DH5 α , where the two PCR

products were mixed in different ratios ranging from 30:1 to 1:0.03. The controls included were without (-) or with only one of the two PCR products present in the EFC reaction (> and <). The average \pm standard deviation of three independent experiments is shown. **(e)** The fraction of correct cloning reactions (efficiency) as determined by the fragment sizes obtained after PCR amplification of the insert of the plasmid, using purified DNA of the transformants under the experimental conditions described for **d**

The EFC method relies on the formation of the hybrid between two PCR products with different extensions. Although PCR is quite robust, it can be anticipated that in HTP applications the two PCR reactions of each target will not be equally efficient. To test the effect of such variations on the cloning efficiency, molar ratios of the two PCR products to be hybridized were varied, ranging from 30:1 to 0.03:1. As shown in Fig. 2d, in the absence of PCR product no colonies were obtained, while when only one PCR product was present, we observed either no or one colony that appeared to be an empty vector (data not shown). Colonies were obtained at all molar ratios, which appeared to contain the correct insert in most of the cases (89%, $N=56$). We only observed a significant decrease in the number of colonies at the most skewed ratio's, while under all other conditions a comparable number of colonies was obtained.

Since DNA concentration variations of both insert and vector and the relative ratio of the complementary PCR products all have only a minor effect on the cloning efficiency, we conclude that EFC should be robust enough for HTP cloning.

EFC is a highly efficient cloning method

To test the suitability of EFC for HTP cloning, 24 different targets were selected (Table 1) and PCR amplified according to both the EFC protocol and the LIC protocol. All targets except clone 21 were successfully amplified from both reverse transcribed cDNA (Fig. 3) and plasmid DNA (data not shown), and all amplified targets could be successfully cloned using both methods. Two distinct bands were obtained for clone six, the larger of the anticipated size, the smaller most probably a PCR artifact; the majority of the plasmids contained the smaller PCR artifact. The absence of such artifacts from the PCR using LIC primers is not a general observation and probably due to experimental variation. Initially PCR amplification of clone 21 failed (Fig. 3a) and, as expected, we were unable to clone this PCR product both using EFC and LIC. After reamplification of both clone 21 and 6 using a new batch of cDNA and addition of 5% glycerol and 5% DMSO, products of the expected size were obtained and successfully cloned using both methods (Fig. 3b, c and data not shown).

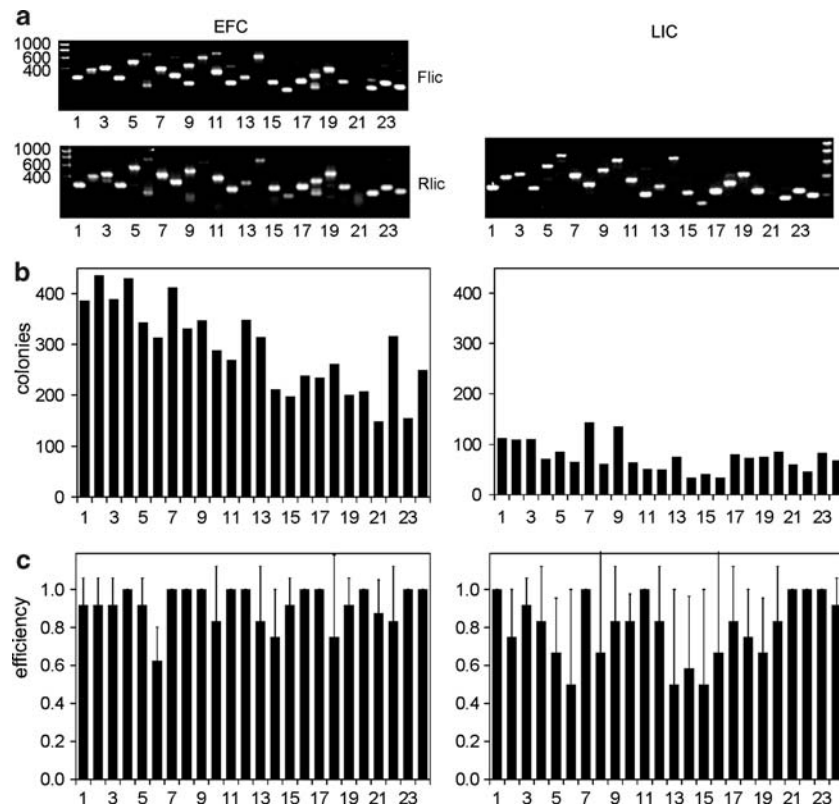


Fig. 3 Pair wise comparison between EFC and LIC for HTP cloning. The left panel shows the results obtained for EFC, while the right panel shows the data for LIC. **(a)** The RT-PCR amplification of the 24 different targets after PCR product purification. Note that for EFC target 6 a PCR product of incorrect size was obtained for both primer sets. For target 21 the first RT-PCR reaction for both EFC and LIC failed. Amplification of these failed reactions using a new batch of cDNA and addition of 5% glycerol and 5% DMSO resulted in the correct fragments. For further analysis the data obtained for the

reamplified product for clone 21 for both EFC and LIC were used. **(b)** The number of colonies obtained after transformation of 10–100 ng PCR product (depending on the PCR efficiency) hybridized with 3 ng vector DNA in DH5 α . **(c)** The fraction of correct cloning reactions for the different targets as determined by the fragment sizes obtained after PCR amplification of the insert of the plasmid, using colony PCR or purified DNA. Data are the average \pm standard deviation of three independent experiments using in total 10 colonies for every target using both methods

Cloning of a comparable amount of either the LIC treated PCR product or the two mixed PCR products, using the same T4 treated vector in the same competent cells, resulted in 2.4 ± 1.4 fold more transformants using the EFC approach. Four independent experiments were performed with two different batches of PCR products and two different vectors using two different competent cell preparations (Fig. 3b, and data not shown), showing a reproducibly better result for EFC than for LIC in agreement with the results presented in Fig. 2. For each target, 10 colonies from three independent experiments were analyzed by PCR using plasmid specific primers, showing a success rate of 91 and 79% for EFC and LIC respectively (Fig. 3c, and data not shown). Using EFC, only a small fraction (14%) of the incorrect constructs ($N = 21$) had no insert. Few contained primer dimers (24%), while the remainder of the failures were incorrectly sized PCR

artifacts (33%), or contained deletions not attributed to incorrect PCR amplification (29%). For LIC a comparable error distribution was found (data not shown).

Since we used standard *Thermus aquaticus* DNA polymerase, lacking 3'–5' exonuclease (proofreading) activity, PCR products with 3' extensions are formed (“extendase activity”), which might be incorporated into the plasmid. DNA sequencing of the vectors never revealed frameshifts due to this extendase activity, neither with EFC, nor with LIC (data not shown). Over 60 plasmids were sequenced; all but one construct were correctly inserted. The single failure was most probably due to incorrect priming of a primer with the plasmid template (data not shown).

As observed before (Fig. 2b), the EFC method was not sensitive to differences in the PCR amplification efficiency of the two products (e.g. 10, 13, 14, 16). The length of the fragments only marginally influenced

Table 1 Targets used for comparison of the EFC and LIC cloning methods. A SwissProt accession number is given if available, otherwise a GenBank accession number is shown (#). Most common gene name, first and last amino acid and size (base pairs) of the PCR amplified fragment are shown

SwissProt	Gene	First residue	Last residue	PCR-size
Q64252	eIF3S6	327	411	288
Q9Y5B5	usp15	1	120	393
Q92541	KIAA0252	345	476	429
Q9C0C8	KIAA1735	408	490	282
Q9Y4E8	HRPT2	358	526	540
Q8TEY7	Usp33	713	942	723
P78344	I4G2	540	669	423
P78332	NY-LU12	1027	1123	324
Q86VG0	VHLAK	1	151	486
Q96GC6	hfb101	148	347	633
Q86UW6	B3BP	1657	1770	375
Q8IWZ8	SF4	187	253	234
Q9NVV9	THA1	1	90	303
O75312	ZPR1	212	423	669
Q92926	BAF60C	252	321	243
Q15424	SAF-B	26	69	165
O75398	DEAF2	200	274	258
O14640	Dishevelled1	387	475	300
O60381	HMGBP1	208	338	426
Q9Y6K1	DNMT3A	289	362	255
4104419#	HNF6	278	384	354
665918#	RO52	285	338	195
Q96L91	p400	765	835	246
Q9P2D1	CHD7	1802	1860	210

cloning efficiency; we have successfully cloned fragments up to 2000 basepairs. Finally, also if PCR amplification was less efficient, cloning could be performed, albeit with a lower success rate (Fig. 2a, data not shown). Overall we conclude that, provided primer dimers are removed, all correctly amplified PCR products can be successfully cloned by both LIC and EFC, EFC being significantly more efficient than LIC.

EFC is suitable for small-scale expression screening

The high cloning efficiency suggests that EFC can be used in HTP small-scale expression screening as an alternative to LIC. We transformed the expression strain BL21(DE3) Rosetta2-LysS either with EFC cloned plasmids isolated from DH5 α , or directly with EFC reactions and evaluated protein expression. All direct transformations were successful as judged by the number of colonies obtained after transformation (ranging from 2 to 30 colonies with a transformation efficiency of 2×10^5 colonies/ μ g of plasmid DNA). Transformation using sequencing verified plasmids, and direct transformation of EFC reactions resulted in largely comparable protein expression (Fig. 4a). Differences in the protein expression levels were probably due to experimental variation and are

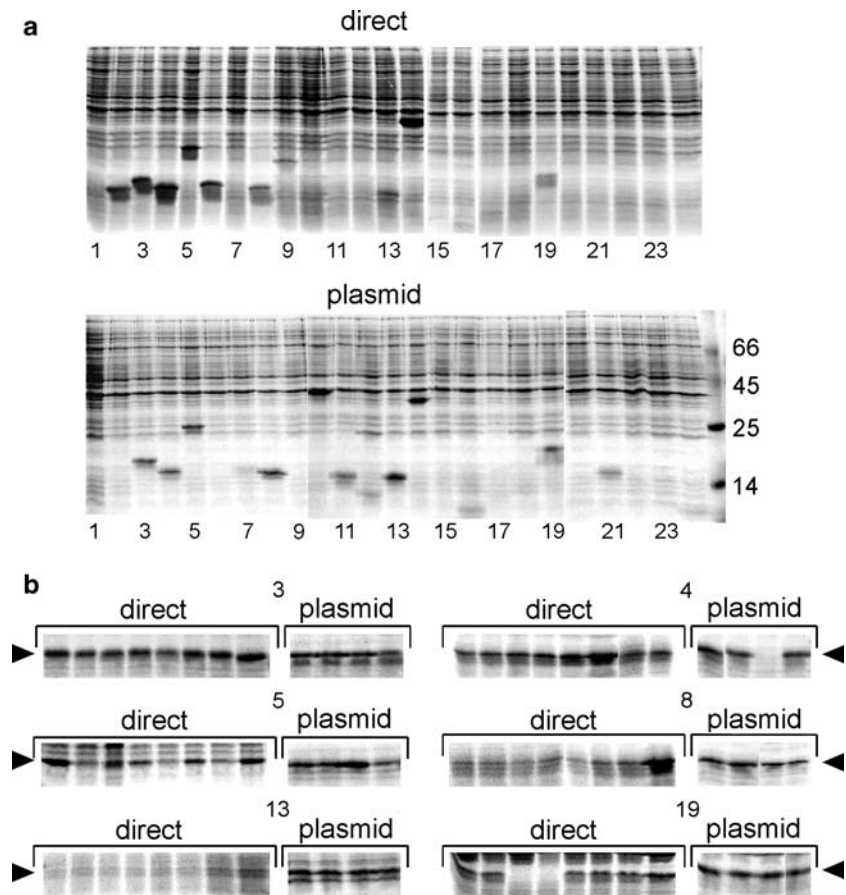
frequently observed in small-scale expression screening (unpublished observations). In contrast to verified plasmid transformation, the direct transformation of the EFC reaction of clone 21 did not yield detectable protein expression, most likely caused by the poor quality of the PCR amplification (Fig. 2a). Furthermore, for clone 6 the smaller PCR product was cloned and expressed.

To test the reproducibility of the direct transformation of EFC reactions, we analyzed the protein expression of eight individual colonies for each clone that expressed significant amounts of protein in Fig. 4a. As a control, expression of 4 colonies from the transformation of a verified plasmid was performed. Figure 4b shows that, like for plasmid transformation, essentially all EFC reactions showed expression of the expected protein. The overall success rate for direct transformation of EFC reactions, 94% (n=80), is comparable with the results obtained above (Fig. 2, 3). Most importantly, both qualitatively (Fig. 4a) and quantitatively (Fig. 4b), protein expression by direct transformation of EFC reactions was only slightly less efficient than expression by transformation of verified plasmids. We conclude that EFC reactions can be used for small-scale expression screening of PCR products without the time consuming intermediary cloning step in a general purpose *E.coli* cloning strain.

EFC proves reliable in High Throughput Cloning

To verify the utility of EFC for HTP cloning, we tested its efficiency in the cloning and expression of 222 targets ranging from 4 kD to ~200 kD unrelated to the previous test group (Table 2). We amplified 210/222 targets successfully from cDNA, while after EFC and transformation to home-made competent DH5 α cells, colonies were obtained for over 98% of all reactions. One colony was selected for each reaction, and PCR analysis revealed a success rate of 65%. For failed reactions, three more colonies were analyzed, of which 52% appeared correct. In total, we screened 450 colonies to obtain 180 constructs, being 86% of all successfully amplified PCR products (data not shown). Failures were generally caused by either poor PCR amplification or the presence of PCR products of unexpected size. Since we performed cDNA amplifications, such incorrectly sized PCR products could also result from alternatively spliced mRNA's. Small scale protein expression screening of EFC cloned plasmids revealed that over 80% of all successfully cloned constructs expressed a protein of the expected size as judged by SDS-PAGE (data not shown). We conclude that EFC can be used as a HTP cloning method.

Fig. 4 EFC for HTP small-scale expression screening. **(a)** Total protein expression in BL21 Rosetta-LysS that were transformed by either EFC cloned plasmids isolated from DH5 α (lower panel) or directly using EFC reactions (upper panel). For direct transformation, 10–100 ng of PCR product (Fig. 3a) was hybridized with 3 ng T4 treated pLICHIS. A total of 15 μ l of IPTG induced cultures grown under identical conditions were analyzed by SDS-PAGE. **(b)** After direct transformation of EFC cloning reactions (direct) or verified plasmids (plasmid), respectively, 8 and 4 colonies were tested in small-scale expression screening. Total protein expression for the various targets is shown. Arrowheads indicate the overexpressed proteins



Discussion

We have compared Ligation Independent Cloning with a modified Enzyme Free Cloning method and show that EFC yields a higher total number of colonies (Fig. 2b), that EFC yields a higher fraction of correct colonies (Fig. 2c) and that EFC is more tolerant to low amounts of vector and insert (Fig. 2a). We show that every successfully amplified PCR product can be cloned using EFC, while the cloning efficiency is only marginally influenced by the ratios of the two PCR products (Fig. 2b). We established the usefulness of this method in HTP applications by cloning 222 targets, and show that after successful PCR amplification, provided that sufficient amount of vector (1–3 ng) is used, the success rate is only marginally influenced by the PCR amplification efficiency (Fig. 2, 3, and data not shown) or fragment size within the tested range from 150 bp to approximately 2 kb (Fig. 3 and data not shown). Evaluation of all EFC experiments revealed an average efficiency of 2.9 colonies per ng PCR product using 1–3 ng of T4 treated vector and competent cells with an efficiency of $10^6/\mu\text{g}$ DNA. Therefore, we advise to use at most 2 ng of vector and

preferably three or more ng of PCR product in high throughput applications. Additionally, we observed that EFC is also an efficient method for site directed mutagenesis and multiple fragment cloning (not shown), in agreement with published data [25].

Cloning efficiencies reported here for EFC are comparable to the Invitrogen Gateway and Clontech In-Fusion system determined in a pairwise comparison of these two systems [24] for the size of fragments cloned here (< 1 kb). Although LIC is used by many structural genomics groups [26, 27], its use is not as widespread as recombinational cloning methods [7, 8, 14, 28, 29]. Dieckmann et al. have shown that LIC can be used for prokaryotic expression vector preparation, but they did not report on efficiencies of the cloning reactions [26]. Given the large percentage of expressed proteins reported by this group, using LIC, efficiency should approach 90%. The commercially available LIC vectors of Novagen claim to have recombination efficiencies greater than 95%, although we are not aware of data supporting these efficiencies. In our hands we obtain a slightly lower cloning efficiency for LIC reactions using all in house made reagents, which is probably due to the presence of PCR artifacts in the

Table 2 Summary of the EFC cloning of 222 targets. All gene fragments were amplified from a home-made cDNA library, cloned using EFC in DH5 α . Plasmids were isolated and protein was expressed in Rosetta2 -pLysS. At each protein production step, both the successrate per step and overall successrates are

indicated. Poor quality PCR: PCR reactions containing very faint bands, additional bands of unexpected height, smears or primer/dimer artifacts. Good quality PCR: PCR reactions with a clear band of correct size

	Success per step (%)	Overall fraction (%)
Targets selected ($N = 222$)	100	100
detectable PCR amplification	95	95
(of which: poor quality)		(8)
(of which: good quality)		(86)
Cloned	86	81
expressed	80	65

cloning reactions such as primer dimers and inappropriate PCR products. These can easily be removed by gel purification of PCR products, although this is not ideal in a HTP cloning protocol. Results from several groups [26, 27] including our unpublished experience and experiments presented here (Fig. 2, 3), show that LIC is a reliable cloning method albeit less efficient than EFC.

Crucial for efficient EFC was the removal of primer dimers. Different PCR cleanup methods that fail to completely remove PCR artifacts produce undesirable numbers of false positives (data not shown). With the currently used PCR cleanup method only 24% of the failed cloning attempts, being 2% of all analyzed ($n > 300$), gave primer dimers. Empty vectors were infrequently found (14% of the failed reactions, 1% of all reactions). The cloning of PCR artifacts or undefined rearrangements between vector and insert caused the remainder of the failures (62%). Importantly, the type of errors and the relative frequency of occurrence for EFC and LIC are comparable. Furthermore, although differences in ratio's of the two PCR products are well tolerated (Fig. 2B), correct PCR amplification of both fragments is preferred for high efficiency cloning as illustrated by the high frequency of failures obtained when no or a smeary PCR product (e.g. clone 21) or an additional PCR fragment was present (e.g. clone 6, 18).

The most important disadvantage of the EFC method are its additional costs. The necessity for two sets of primers, requiring on average 40 additional nucleotides per cloning reaction in comparison with LIC, and the double PCR reaction and purification makes this method more expensive than LIC. These additional costs associated with the EFC method could be decreased using two universal adaptor primers for all PCR products that encode the vector compatible overhang and recognize a standardized 5' end of gene specific primers. Advantageous is the higher success rate, which reduces the number of transformants that need to be analyzed. We further show that the

transformation in a cloning strain and subsequent plasmid isolation can be omitted with EFC for small-scale expression screening (Fig. 4). The reproducibility of the direct transformation EFC method was underscored by the small percentage of colonies (~10%) that failed to express the protein (Fig. 4b). This increases the throughput, and reduces the costs associated with the normally used protocol. Altogether, we feel that the limited additional costs of 40 nucleotides for an additional primer set are well balanced by the absence of enzymatic treatments after PCR and the increased efficiency and throughput.

The EFC method as proposed by Tillett and Neilan (20) uses a vector that is PCR amplified also using EFC, making cloning completely independent of vector sequences. We were unable to prepare vectors consistently in the quantities needed for HTP applications (unpublished results). By combining a T4 treated vector and a hetero-staggered PCR product, we extend the versatility of EFC applications. Moreover, existing LIC vectors are perfectly compatible with EFC. Extrapolating the EFC cloning efficiency and using commercially available chemical competent cells (10^8 /ug), we should expect approximately 10^5 colonies when 10 ng of vector and 50 ng PCR product are used. Such efficiencies should make it possible to efficiently perform applications like e.g. library construction for putatively interacting proteins, domain swapping, antibiotics resistance gene swapping or promoter swapping. We currently evaluate some of these applications.

We expect that through the implementation of this cloning method, we will be able to completely automate small-scale expression screening. The reported EFC efficiencies, the ability to directly clone PCR products in an expression strain without a subcloning step, and the ability to automate these experiments, will greatly facilitate small-scale expression screening, and holds the promise to increase the throughput of protein production for functional and structural genomics.

Acknowledgements This work was supported by the EU SPINE (Structural Proteomics in Europe) grant QLG2-CT-2002–00988. M.A.D. was supported by a NWO Horizon Breakthrough Project grant 050-71-629. The authors would like to thank Darren Hart for stimulating discussions, Martin Hammarström for plasmids, and H. van Aken for technical support.

References

1. Eisenberg D, Marcotte EM, Xenarios I, Yeates TO (2000) *Nature* 405:823
2. Vukmirovic OG, Tilghman SM (2000) *Nature* 405:820
3. Bader GD, Heilbut A, Andrews B, Tyers M, Hughes T, Boone C (2003) *Trends Cell Biol* 13:344
4. Bork P, Serrano L (2005) *Cell* 121:507
5. Ge H, Walhout AJ, Vidal M (2003) *Trends Genet* 19:551
6. Grant BD, Wilkinson HA (2003) *Current Opinion in Cell Biology* 15:206
7. Braun P, LaBaer J (2003) *Trends Biotechnol* 21:383
8. Pearlberg J, LaBaer J (2004) *Curr Opin Chem Biol* 8:98
9. Rual JF, Hirozane-Kishikawa T, Hao T, Bertin N, Li S, Dricot A, Li N, Rosenberg J, Lamesch P, Vidalain PO, Clingingsmith TR, Hartley JL, Esposito D, Cheo D, Moore T, Simmons B, Sequerra R, Bosak S, Doucette-Stamm L, Le Peuch C, Vandenhoute J, Cusick ME, Albala JS, Hill DE, Vidal M (2004) *Genome Res* 14:2128
10. Yokoyama S (2003) *Curr Opin Chem Biol* 7:39
11. Hunt I (2005) *Protein Expr Purif* 40:1
12. Folkers GE, van Buuren BN, Kaptein R (2004) *J Struct Funct Genomics* 5:119
13. Klock HE, White A, Koesema E, Lesley SA (2005) *J Struct Funct Genomics* 6:89
14. Walhout AJ, Temple GF, Brasch MA, Hartley JL, Lorson MA, van den Heuvel S, Vidal M (2000) *Methods Enzymol* 328:575
15. Aslanidis C, de Jong PJ (1990) *Nucleic Acids Res* 18:6069
16. Haun RS, Serventi IM, Moss J (1992) *Biotechniques* 13:515
17. Liu Z (1996) *Nucleic Acids Res* 24:2458
18. Zeng G (1998) *Biotechniques* 25:206
19. Shih YP, Kung WM, Chen JC, Yeh CH, Wang AH, Wang TF (2002) *Protein Sci* 11:1714
20. Neilan BA, Tillett D (2002) *Methods Mol Biol* 192:125
21. Tillett D, Neilan BA (1999) *Nucleic Acids Res* 27:e26
22. Hammarstrom M, Hellgren N, van den Berg S, Berglund H, Hard T (2002) *Protein Sci* 11:313
23. Sambrook EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York
24. Marsischky G, LaBaer J (2004) *Genome Res* 14:2020
25. Chiu J, March PE, Lee R, Tillett D (2004) *Nucleic Acids Res* 32:e174
26. Dieckman L, Gu M, Stols L, Donnelly MI, Collart FR (2002) *Protein Expr Purif* 25:1
27. Doyle SA (2005) *Methods Mol Biol* 310:107
28. Braun P, Hu Y, Shen B, Halleck A, Koundinya M, Harlow E, LaBaer J (2002) *Proc Natl Acad Sci U S A* 99:2654
29. Reboul J, Vaglio P, Rual JF, Lamesch P, Martinez M, Armstrong CM, Li S, Jacotot L, Bertin N, Janky R, Moore T, Hudson JR, Jr, Hartley JL, Brasch MA, Vandenhoute J, Boulton S, Endress GA, Jenna S, Chevet E, Papasotiropoulos V, Tolia PP, Ptacek J, Snyder M, Huang R, Chance MR, Lee H, Doucette-Stamm L, Hill DE, Vidal M (2003) *Nat Genet* 34:35