A versatile bacterial expression vector designed for single-step cloning of multiple DNA fragments using homologous recombination

Mats A. Holmberg, Naveen Kumar Chandappa Gowda, Claes Andréasson

Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, S-10691 Stockholm, Sweden. Tel: +46 8 164202. Fax: +46 8 15 98 37. Email: claes.andreasson@su.se.

Abbreviations:
LIC, ligation independent cloning; YHRC, yeast homologous recombination cloning
Abstract

Production of recombinant proteins is the starting point for biochemical and biophysical analyses and requires methodology to efficiently proceed from gene sequence to purified protein. While optimized strategies for the efficient cloning of single-gene fragments for bacterial expression is available, efficient multiple DNA fragment cloning still presents a challenge. To facilitate this step, we have developed an efficient cloning strategy based on yeast homologous recombination cloning (YHRC) into the new pET-based bacterial expression vector pSUMO-YHRC. The vector supports cloning for untagged expression as well as fusions to His6-SUMO or His6 tags. We demonstrate that YHRC from single PCR products of 6 independent genes into the vector results in virtually no background. Importantly, in a quantitative assay for functional expression we find that single-step YHRC of 7 DNA fragments can be performed with very high cloning efficiencies. The method and reagents described in this paper significantly simplifies the construction of expression plasmids from multiple DNA fragments, including complex gene fusions, chimeric genes and polycistronic constructs.

Protein expression/Protein purification/cloning/recombination cloning
SUMO/nanobody/
Introduction

Straightforward cloning and production of biologically active proteins is central for studies in biochemistry and biophysics. During recent years, several improvements in recombinant protein expression and purification techniques have been developed. Yet many laboratories still depend on unreliable cloning techniques based on ligation of endonuclease restriction products. Especially, when cloning multiple DNA fragments into one expression vector traditional methodology requires time-consuming reiteration of the process and the cloning step frequently becomes limiting in the protein production process.

Current cloning methodology relies on PCR amplification of the relevant gene using two primers designed to facilitate the subsequent cloning step. Traditionally, unique restriction sites in the flanking primer regions are digested and the PCR product is ligated into a vector containing compatible cohesive ends. Similarly, in TA-cloning a single adenosine residue added at each end of the PCR product by Taq DNA polymerase facilitates ligation into a restricted vector carrying complementary thymidine residues [1]. An alternative to these methods is ligation independent cloning (LIC). In LIC flanking sequence extensions are added to the PCR product and are by enzymatic modification converted to single stranded overhangs that are annealed to a vector carrying complementary ends [2].

Cloning of multiple DNA fragments into one expression vector usually relies on the tedious process of subcloning, i.e. reiterative restriction and ligation cloning. In seamless cloning the introduction of extra DNA sequences with restriction sites between the gene fragments is avoided [3]. The most wide-spread methodology for the seamless assembly of multiple DNA fragments is fusion PCR, also called overlap extension PCR [4]. Briefly, PCR products are generated that by flanking sequence homology can prime each other in a consecutive PCR reaction and produce a fused PCR product that can be cloned by any method of choice. The challenge of the task and work investment increase rapidly with the number of fragments that have to be fused and the size of the product. A recent development in multiple DNA fragment cloning is the Gibson one-step isothermal in vitro recombination assembly method [5]. The method is based on incubation of the DNA fragments with T5 exonuclease, Phusion DNA polymerase and Taq DNA ligase and has been commercialized by New England Biolabs. However, efficiency of assembly decreases as the number or length of fragments increases.

Homologous recombination cloning offers an alternative for the seamless assembly of multiple DNA fragments. The methodology is versatile and in principle any stretch of sequence homology of sufficient length between the DNA fragments and the vector will suffice for homologous recombination. In E. coli the in vivo expressed bacteriophage-based recombination systems recET and lambda Red system can be used to recombine PCR products to a plasmid vector during transformation [6, 7]. In practice the limited recombination frequency (5x10⁻⁴) often requires either direct selection for the insert (antibiotic marker) or counter-selection of the vector (toxic counterselection marker). Simple
restriction of the vector at the site of recombination could in principle suffice as counter-selection but since transformation of *E. coli* greatly favors supercoiled plasmid over linear DNA fragments any trace of undigested vector will result in false positive clones making this approach inefficient.

In budding yeast (*Saccharomyces cerevisiae*) plasmid construction by yeast homologous recombination cloning is straightforward with only restriction linearization of the vector as the means of its counter selection. Simply transforming yeast cells using a vector, restricted at the site for recombination, together with a PCR product that gaps the restriction point with vector sequence homology at each end results in highly efficient cloning [8-11]. Moreover, multiple PCR fragments or even oligonucleotides or genomic DNA can efficiently be cloned by YHRC in a single step involving several recombination events [12, 13]. Given that the vector employed is a yeast shuttle vector it can readily be purified from the yeast transformants using standard silica-based plasmid minipreps and introduced in *E. coli* to facilitate analysis and protein expression [14]. *E. coli* plasmids can be converted into single copy yeast shuttle vectors simply by introducing a selection marker together with a short autonomous replication sequence (*ARS*) [15, 16] and a centromere sequence (*CEN*) [17]. Therefore YHRC potentially could offer a robust methodology for the assembly *E. coli* expression vectors.

Current approaches for recombinant protein purification frequently rely on affinity tags such as the hexahistidine tag (His<sub>6</sub>), Glutathione-S-transferase or Maltose Binding Protein [18]. Certain tags have the added benefit of enhancing the soluble expression of its fusion partner [19]. For example, an N-terminal fusion of SUMO (yeast Smt3) dramatically enhances soluble expression of its fusion partner [20-22]. If properly cloned, SUMO will be fused directly to the native N-terminal residue and all SUMO residues are fully removed after cleavage with SUMO protease (yeast Ulp1). Although the SUMO tag does not itself offer an affinity feature it can readily be combined with other tags, for example a His<sub>6</sub> tag.

In this study we present an efficient strategy for the cloning of multiple gene fragments in a single-step using YHRC. Our approach is based on the novel pET-derived bacterial expression vector pSUMO-YHRC that replicates in yeast and confers drug resistance in both yeast and *E. coli*. 
Material and Methods

Reagents, strains and media

Oligonucleotide and gene synthesis as well as sequencing were performed by Eurofins MWG Operon (Ebersberg, Germany). Restriction enzymes and Phusion DNA polymerase were obtained from Thermo Fisher Scientific Inc. (Waltham, MA). Miniprep kits were obtained from Zymo Research (Irvine, CA). Yeast lytic enzyme (lysozyme) was purchased from MP Biomedicals (Santa Ana, CA). Agar, peptone, tryptone and yeast extract were obtained from Formedium Ltd. (Hunstanton, England, UK). SUMO protease (Ulp1-His6) was expressed and purified as described previously [23]. Monoclonal mouse anti-GFP antibodies were obtained from Roche Applied Science (Roche Diagnostics Scandinavia AB, Sweden). Other reagents were purchased from Sigma-Aldrich (Pittsburgh, PA) if not otherwise stated.

Standard media used to propagate yeast strain CAY1015 was YPD (1% yeast extract, 2% peptone, 2% glucose). Yeast strain CAY1015 (MATa his3Δ1 leu2Δ0 ura3Δ0) was obtained a meiotic segregant of the S288C strain BY4743 [24]. For yeast antibiotic selection 200 mg/l G418 was included in the media solidified by 2% agar.

LB media was used for propagating the E. coli DH10B pyrF derivative strain DH10B-U [25]. The protein expression strain BL21-SI (Invitrogen, Life Technologies Ltd, Paisley, UK) that carries a chromosomal salt-inducible gene encoding T7 RNA polymerase was propagated using the low salt medium 2YTON (1.6% tryptone, 1% yeast extract) supplemented with 1 mM MgSO4. When plasmid selection was required 50 mg/l kanamycin and 25 mg/l chloramphenicol were added to the media.

Construction of SUMO fusion expression vector

YHRC SUMO fusion expression vector pSUMO-YHRC is a derivative of the already described pET24d-derived SUMO fusion vector pCA528 [23] and was constructed in a single step of homologous recombination between three DNA fragments. Briefly, 100 ng of Clal restricted pCA528, 500 ng DpnI restricted pUG6 [26] and 500 ng of a 586 bp CEN6/ARS209 PCR product amplified from pRS316 [27] using primers 1 and 2 (Table 1) were used to cotransform strain CAY1015 to G418 resistance. Plasmids were rescued from the pooled transformants and used to transform DH10B-U to kanamycin resistance. One clone was validated by DNA sequencing and saved as pSUMO-YHRC.

Yeast homologous recombination cloning (YHRC)

(i) PCR of genes: PCR products to be used for homologous recombination cloning was obtained with Phusion High-Fidelity DNA Polymerase according to the manufacturers instructions. Briefly, a 100 µl PCR reaction was setup in HF Buffer (200 µM dNTPs, 60 ng template DNA, 0.2 µM of each primer, 5% DMSO, 2 U Phusion High-Fidelity DNA Polymerase) and cycled in a LifePro Thermal Cycler.
(Bioer Technology, China) (1× 97°C for 2 min; 30× 97°C for 45 s, 55°C for 20 s, 72°C for 1 min). To remove any plasmid PCR template 1× FastDigest Buffer and 2 FDU DpnI was added to the completed PCR reactions followed by 2 h incubation at 37°C.

(ii) Preparation of the vector: 1 µg of pSUMO-YHRC was restricted for 2 h at 37°C in a total reaction buffer of 30 µl containing FastDigest Buffer and 1 FDU of BamHI, HindIII and Xhol.

(iii) Yeast transformation: Yeast was transformed by an adapted protocol described by [28]. Specifically, 10 ml CAY1015 was grown over night in YPD medium at 30°C. In the morning the strain was diluted in fresh YPD medium (5 ml per transformation) and grown at 30°C from starting OD600 = 0.1. After ~5 h, the cultures had reached OD600 = 1.0 and 3 ml of cell culture for each transformation was harvested by centrifugation (1.5 ml Eppendorf tubes) and washed by resuspension and centrifugation in 100 mM LiOAc (1 ml per transformation). The cell pellet was resuspended in 50 µl carrier DNA (2 mg/ml heat-denatured salmon testis DNA (Sigma D1626) in 100 mM LiOAc 1 mM EDTA) together with 5 µl vector digest (above) and 20 µl crude PCR reaction (above). The suspension was briefly mixed by vortexing followed by the addition of 300 µl 39% PEG 3350 100 mM LiOAc and intense vortexing for 15 s. After heat shock in a water bath at 42°C for 40 min cells were pelleted by centrifugation, resuspended in 150 µl Milli-Q water and plated on solid YPD medium supplemented with 200 mg/l G418. Colonies were formed after 2 days of incubation at 30°C.

(iv) Plasmid rescue from yeast: Fresh colonies (>500 cfu per plate, 10⁴ cfu/µg restricted pSUMO-YHRC DNA) were harvested and pooled by washing the cells of the agar surface using 2 ml of TE. 20 µl of the cell suspension was transferred to an Eppendorf tube and mixed with 200 µl TE and 5 µl of lyticase solution (5 U/µl lyticase in TE, 50% glycerol, 5 mM β-mercaptoethanol). After incubation at 37°C for >1 h to remove the cell wall by enzymatic digestion, a standard commercial silica column miniprep kit was used (Zippy™ Plasmid Miniprep Kit, Zymo Research Corporation) according to the manufacturers instructions with the modification of starting with the addition of the 7x Lysis Buffer that contains SDS and NaOH. Silica columns were eluted using 30 µl of water and 2 µl of miniprep DNA was used to transform DH10B-U to kanamycin resistance by electroporation.

Cloning of expression plasmids

Plasmids pCA856 (His₆-SUMO-NanobodyGFPI), pCA861 (His₆-SUMO-Rpo41T⁹²⁰-L₁₂¹⁷-HA), pCA862 (His₆-SUMO-DHFRmutD-FLAG), pCA863 (His₆-SUMO-3HA-stGnd1), pCA892 (His₆-SUMO-Sis1) and pCA893 (His₆-SUMO-Ydj1) were constructed by YHRC of pSUMO-YHRC after PCR amplification of the genes with primer pairs 3/4, 5/6, 7/8, 9/10, 11/12 and 13/14 (Table 1). Plasmids or yeast genomic DNA were used as PCR templates (C.A. collection) and in the case of NanobodyGFPI [29] the 354 bp E. coli codon-optimized gene was synthesized
including a 5’ in-frame SUMO homology sequence (5’-GATATTATTGAGGCTCACAGAAGATTGGTG-3’). The generated plasmids were sequenced using primers 15 (anneals in SUMO, coding strand) and 16 (anneals downstream of the insert, non-coding strand).

**Expression and purification of His6-SUMO-NanobodyGFP**

BL21-SI cells were transformed with pCA856 and pMJS10 (ARp-Erv1-DsbC, p15A, Cam^R^) [30] and colonies were selected on YTON+Kan+Cam. An overnight culture in 2YTON+Kan+Cam was setup by inoculating with 20 pooled colonies. The next morning the overnight culture was diluted 1:100 in fresh medium and expanded over day to early stationary phase (OD$_{600}$ ∼1.0-1.5). Pre-expression of the sulfhydryl oxidase and the disulfide isomerase was induced by the addition of 0.2% (w/v) L-arabinose and the culture was allowed to cool to 20°C for 1 h before His$_6$-SUMO-NanobodyGFP expression was induced by the addition of 0.5 mM IPTG (isopropyl beta-D-thiogalactopyranoside) and 0.3 M NaCl. Cells were harvested the following morning by centrifugation.

Cells from 1 L of culture were chemically lysed in 10 ml of 40 mM Tris-HCl pH 7.5, 400 mM NaCl by the addition of 1 g/L lysozyme, 0.5% (w/v) SB3-14, 0.5% (v/v) Triton X-100 in the presence of a few crystals of DNase I and 1 mM PMSF. After preclearing the lysate in a JA17 rotor (Beckman Coulter) at 14000 rpm for 30 min the supernatant was applied to 1 g Protino Ni-IDA resin (Machery-Nagel GmbH & Co. KG, Düren, Germany) and was incubated for 30 minutes. After extensive washing with 40 mM Tris-HCl pH 7.5, 400 mM NaCl bound protein was eluted in fractions using 40 mM Tris-HCl, 400 mM NaCl, 250 mM imidazole pH 7.5. Protein peak fractions were pooled and the His$_6$-SUMO tag was severed from the Nanobody$^{GFP}$ by digestion with 1% (w/w total protein) Ulp1-His$_6$ SUMO protease during 16 dialysis against 40 mM Tris-HCl pH 7.5, 400 mM NaCl. The protein solution was depleted of His$_6$-SUMO and Ulp1-His$_6$ by passing it over a gravity flow column containing 1 g Protino Ni-IDA. Following extensive dialysis against 10 mM Hepes pH 7.4, 50 mM NaCl, 5% glycerol, the nanobody solution was filtered through 50 kDa MWCO Vivaspin ultrafilter (GE Healthcare UK Ltd, UK) and protein concentration was determined by A$_{280}$ (Extinction coefficient 27055 M$^{-1}$ cm$^{-1}$). All steps were performed at 4°C.

**Nanobody$^{GFP}$ immobilization and GFP immunoprecipitation**

6 mg of purified Nanobody$^{GFP}$ was immobilized onto 0.29 g NHS-Sepharose slurry according to the manufacturers instructions (GE Healthcare). Briefly, purified nanobody at a concentration of 1.1 mg/mL was diluted 1:1 with standard coupling buffer (0.2 M NaHCO$_3$, 0.5 M NaCl, pH 8.3) and was allowed to react with the matrix before it was washed extensively with Tris-based buffer.

For GFP immunoprecipitation uracil prototrophic transformants of yeast strain CAY1015 carrying either EGFP-expressing pUG35 [31] or empty vector pRS316 [27] were expanded at 30°C in 500 ml SC medium lacking methionine. Cells were harvested at OD$_{600}$ 0.8 by centrifugation and were resuspended in 500 µl 10 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5 % NP40, 1 mM PMSF, 1x concentration. After gravity flow column containing 1 g Protino Ni-IDA resin (Machery-Nagel GmbH & Co. KG, Düren, Germany) and was incubated for 30 minutes. After extensive washing with 40 mM Tris-HCl pH 7.5, 400 mM NaCl bound protein was eluted in fractions using 40 mM Tris-HCl, 400 mM NaCl, 250 mM imidazole pH 7.5. Protein peak fractions were pooled and the His$_6$-SUMO tag was severed from the Nanobody$^{GFP}$ by digestion with 1% (w/w total protein) Ulp1-His$_6$ SUMO protease during 16 dialysis against 40 mM Tris-HCl pH 7.5, 400 mM NaCl. The protein solution was depleted of His$_6$-SUMO and Ulp1-His$_6$ by passing it over a gravity flow column containing 1 g Protino Ni-IDA. Following extensive dialysis against 10 mM Hepes pH 7.4, 50 mM NaCl, 5% glycerol, the nanobody solution was filtered through 50 kDa MWCO Vivaspin ultrafilter (GE Healthcare UK Ltd, UK) and protein concentration was determined by A$_{280}$ (Extinction coefficient 27055 M$^{-1}$ cm$^{-1}$). All steps were performed at 4°C.

**Nanobody$^{GFP}$ immobilization and GFP immunoprecipitation**

6 mg of purified Nanobody$^{GFP}$ was immobilized onto 0.29 g NHS-Sepharose slurry according to the manufacturers instructions (GE Healthcare). Briefly, purified nanobody at a concentration of 1.1 mg/mL was diluted 1:1 with standard coupling buffer (0.2 M NaHCO$_3$, 0.5 M NaCl, pH 8.3) and was allowed to react with the matrix before it was washed extensively with Tris-based buffer.
Complete protease inhibitors (Roche Diagnostics Scandinavia AB, Sweden) and 1 mg/ml each of DNAse and RNAse followed by glass bead lysis using a FastPrep-24 (MP Biomedicals), 6.5 M/s for 30 s. After preclearing by centrifugation the lysates were incubated first with 100 µl of Sepharose bead slurry and then with 100 µl of Sepharose-NHS-Nanobody\textsuperscript{GFP} slurry for 1 h each. Sepharose-NHS-Nanobody\textsuperscript{GFP} beads were washed 3 times with 0.5 mL 10 mM Tris-HCl pH 7.5, 300 mM NaCl, 0.5 mM EDTA, 0.5 % NP40, 1 mM PMSF and 1x Complete protease inhibitors and bound protein was released by boiling in SDS-Page sample buffer.

Assay for multiple fragment assembly of LacZ

YHRC into pSUMO-YHRC was performed as described above using lacZ PCR fragments individually amplified from the plasmid YCpAGP1-LacZ [32]. The following lacZ PCR products were amplified: 1 fragments of 3146 bp (Primers 17/18); 2 fragments of 1714 bp (primers 17/25) and 1482 bp (primers 26/18); 3 fragments of 890 bp (primers 17/21), 1296 bp (primers 22/27) and 1060 bp (primers 28/18); and 7 fragments of 515 bp (primers 17/19), 428 bp (primers 20/21), 442 bp (primers 22/23), 482 bp (primers 24/25), 472 bp (primers 26/27), 468 bp (primers 28/29) and 611 bp (primers 30/18). The DNA extracted plasmid DNA from the G418 resistant yeast cells was used to transform the lacZ T7 RNA polymerase expressing strain T7 Express LysY (New England Biolabs). Kanamycin resistant transformants were selected on LB medium supplemented with 40 mg/L X-gal (5-bromo-4-chloro-3-indolyl-\(\beta\)-D-galactopyranoside) and 10 µM IPTG and colonies were scored for expression of \(\beta\)-galactosidase activity by blue color development.

Results and discussion

Vector pSUMO-YHRC

We set out to construct a vector for protein expression in E. coli that is compatible with YHRC and based on the widely used inducible T7 expression system [33]. Starting with pCA528 [23], a pET24a derivative expressing His\textsubscript{6}-SUMO under the control of Lacl-repressed T7 promoter [34], we introduced sequences for replication, mitotic segregation and selection in yeast. As genetic elements for stable single-copy replication and mitotic segregation we used a 515 bp cassette (CEN6/ARS209) from the pRS-series of plasmids [27]. To obtain a compact selection marker functional in both yeast and E. coli we decided to replace the bacterial promoter of the transposon Tn903 kanamycin resistance gene (\textit{kan}) on pCA528/pET24a with a bifunctional promoter. In yeast, expression of aminoglycoside 3'-phosphotransferase from \textit{kan} in the drug resistance marker cassette kanMX results in resistance to G418 [35]. We had previously noted that also \textit{E. coli} cells transformed with plasmids that carried kanMX were resistant to kanamycin, indicating that cryptic \textit{E. coli} promoters are present in the promoter of kanMX. We searched the yeast TEF promoter of the kanMX cassette for Sigma70 promoters and found putative -10 (5'-ATGTATAAT-3') and -35 (5'-TTTAGC-3') boxes starting 247 bp and 265 bp upstream of the initiating codon of \textit{kan}, respectively. Encouraged by this finding we used YHRC to transfer the entire 378 bp TEF promoter from kanMX to pCA528 together with
CEN6/ARS4 from pRS316 and selected for G418-resistant yeast transformants (see Material and Methods for details). The resulting plasmid pSUMO-YHRC (Fig. 1A) was rescued to E. coli and conferred robust growth on LB supplemented with kanamycin (Fig. 1B).

Next, we performed growth experiments to make sure that the level of antibiotic resistance conferred by the TEF promoted kanamycin resistance gene in pSUMO-YHRC is sufficient for robust growth in selective media. Untransformed E. coli or transformants carrying pET24a, pCA528 or pSUMO-YHRC were plated onto LB medium supplemented with increasing levels of kanamycin. The transformants grew similarly at the standard kanamycin concentration of 50 mg/L. However, at tenfold higher concentration of kanamycin (500 mg/L), transformants carrying pSUMO-YHRC exhibited more robust growth than transformants of pET24a or its derivative pCA528. The difference in growth when increasing the kanamycin concentration ten-fold above standard laboratory concentrations suggests that the TEF promoter drives somewhat higher expression of aminoglycoside 3’-phosphotransferase than the promoter originally present in the Tn903 kanamycin resistance gene of pET24a. We conclude that the bifunctional TEF promoted kanamycin resistance gene in pSUMO-YHRC is functional in both yeast and E. coli.

**YHRC of single fragments in pSUMO-YHRC**

We directly assessed the efficiency of YHRC into pSUMO-YHRC by cloning six different genes encoding: NanobodyGFP [36, 37], Rpo41T920-L1217-HA, DHFRmutD-FLAG, [38], 3HA-stGnd1 [39], Sis1 [40] and Ydj1 [41]. First, we defined DNA sequences in pSUMO-YHRC suitable for homologous recombination (Fig. 2A, 5’ homology region and 3’ homology region). These 36 bp and 38 bp sequences, respectively, are of sufficient length to support efficient homologous recombination in yeast but yet small enough to be included as 5’ overhangs in PCR primers. Next, the six genes were PCR amplified with primers flanked by the defined homology regions, mixed with BamHI, HindIII and Xhol restricted pSUMO-YHRC and used to transform yeast to G418 resistance (Fig. 2B, for details see Materials and Methods). In this step in vivo homologous recombination between the restricted pSUMO-YHRC and the PCR product circularizes the DNA, thereby allowing it to replicate in yeast. While restricted vector alone resulted in a transformation efficiency of $10^3$ cfu/µg, inclusion of insert yielded $10^4$ cfu/µg (Fig. 2C). The low level of vector background was likely caused by residual circular vector and rare recombination events. The resulting yeast transformants were pooled and the rescued library of plasmid DNA was used to transform E. coli to kanamycin resistance (Fig. 2B). Restriction analysis of 29 independent plasmid clones from the six clonings indicated that the overall efficiency of the method was very high with a total of 97% correct clones (Table 2). As anticipated, sequencing of one clone from each cloning demonstrated that YHRC produced perfect recombination.

**Expression and purification of functional NanobodyGFP**
We decided to test the functionality of pSUMO-YHRC in protein expression and purification setups by focusing on the pSUMO-YHRC derived vector pCA856 that expresses His6-SUMO-Nanobody\(^{GFP}\) (above). Nanobody\(^{GFP}\) is a single domain antibody fragment (V\(\text{H}1\)) derived from a llama antibody that binds to GFP with nanomolar affinity and has been described under several different names including GFP-binding protein 1 (GBP1), Enhancer and GFP-nanobody [29, 36, 37, 42]. Solubility tags similar to His6-SUMO are known to facilitate high-level expression of nanobodies in the cytoplasm of \(E.\ coli\) [43]. Since oxidation of the single cysteine pair in the structure of the nanobody cannot normally take place in the reducing cytoplasm, we coexpressed sulfhydryl oxidase Erv1 and a disulfide isomerase DsbC from plasmid pMJS10 [30] together with His6-SUMO-Nanobody\(^{GFP}\) in BL21-SI cells.

After induction of the T7 RNA polymerase-dependent expression of His6-SUMO-Nanobody\(^{GFP}\) overnight in 1 L of culture, cells were harvested and protein extract was prepared (Fig. 3A). Protein was subjected to Ni-IDA affinity chromatography and 72 mg of protein was obtained. Following proteolytic severing of the tag with SUMO protease overnight, His6-SUMO was depleted from the protein solution by passing it over Ni-IDA matrix resulting in 8.3 mg of purified Nanobody\(^{GFP}\).

We tested if the purified Nanobody\(^{GFP}\) expressed from pSUMO-YHRC was properly folded and functional by a GFP binding experiment. Briefly, Nanobody\(^{GFP}\) was immobilized on NHS-Sepharose beads and the resulting matrix was used to probe for EGFP binding. Protein extracts were prepared from yeast cells that expressed EGFP from the \(MET25\) promoter on a single-copy plasmid [31] and then subjected to Nanobody\(^{GFP}\) immunoprecipitation (Fig. 3B). Nanobody\(^{GFP}\) quantitatively bound EGFP from the extracts. The identity of EGFP was confirmed by both immunoblot analysis and a control purification from lysates that did not express EGFP. We noted that the Nanobody\(^{GFP}\) matrix also was bound by a yet unidentified yeast protein migrating right below the 51 kDa marker. In conclusion pSUMO-YHRC expresses soluble and functional Nanobody\(^{GFP}\).

\textit{YHRC of multiple fragments into pSUMO-YHRC}

We asked if YHRC into pSUMO-YHRC was compatible with efficient seamless multiple DNA fragment assembly in a single step. The 3.1 kb \(E.\ coli\) gene \(lacZ\) that encodes \(\beta\)-galactosidase was used as a model and was amplified as a single PCR product or split into 2, 3 and 7 PCR products (Fig. 4A and B). Primers were designed so that the ends of \(lacZ\) carried the standard homology to pSUMO-YHRC (see above) and so that each fragment carried 50 bp of homology to the neighboring fragments. Following transformation of yeast to G418 resistance with the restricted pSUMO-YHRC and the \(lacZ\) fragments (in all cases \(10^4\) cfu/\(\mu\)g), plasmid DNA was extracted and used to transform the \(\DeltalacZ\) \(E.\ coli\) strain T7 Express LysY to score for functional expression of His6-SUMO-\(\beta\)-galactosidase. While the vector control did not result in any Lac\(^+\) colonies, functional expression was observed for YHRC using the 1, 2, 3 or 7 PCR products of \(lacZ\) (Fig. 4C). Importantly, cloning efficiency was not severely negatively
impacted by increasing the number of gene fragments, e.g., 2 fragments resulted in 90.7% Lac+ colonies while the corresponding number for 3 fragments was 88.3% and for 7 fragments 85.7%. Paradoxically, single-fragment recombination gave reproducibly the lowest frequency of Lac+ colonies (74.7%). However, this is explained by the fact that the single 3.1 kb lacZ fragment was not as well amplified as the smaller fragments (Fig. 4B). Taken the data together, multiple-fragment YHRC into pSUMO-YHRC is efficient even with as many as 7 gene fragments.

Tag versatility of pSUMO-YHRC

We mainly use pSUMO-YHRC for N-terminal His6-SUMO fusions, however by simply exchanging the primer homology regions, or by omitting the termination codon in the primers during YHRC several other options are available (Table 3). For example, genes can be cloned to express no N-terminal tag or a His6 tag at either the N- or C-terminus. Similarly, since the YHRC methodology supports high-efficiency cloning involving multiple gene fragments, other tags (e.g. GST and MBP) can readily be introduced in a single cloning step. The vector and YHRC methodology are potentially useful when in a single step creating expression cassettes for the simultaneous expression of multiple genes, both in polycistronic and multiple monocistronic designs. In summary pSUMO-YHRC can therefore replace all commonly used pET vectors.

Time and efficiency considerations

YHRC is a highly efficient strategy for the cloning of expression plasmids in bacteria. The high efficiency stems from three sequential selections when processing pSUMO-YHRC (Fig. 2B). First, the transformation of yeast selects for functional genes that express G418 resistance. Second, the plasmid rescue selects for plasmids that can replicate in yeast. Third, transformation of E. coli selects for plasmids that can replicate and confers kanamycin resistance. To this end we have scored a close to 100% YHRC efficiency for single genes by restriction analysis (Table 2). Analysis of functional expression after multiple-fragment YHRC yields a >85% cloning efficiency (Fig. 4). This is a value that also takes inactivating PCR and primer mutations into account.

YHRC has a low work load and saves time due to its high efficiency and single step cloning of multiple gene fragments. This makes the methodology useful also for high-throughput applications, e.g., testing different fusion partners. Unlike most other methods, there is no need for agarose gel purification of the DNA and the crude PCR reaction is directly mixed with the restricted pSUMO-YHRC during yeast transformation. By treating the pool of yeast transformants as carriers of a library of plasmids, processing time in yeast is limited to two days and bacterial transformation and growth requires an additional day (Fig. 2B). We note that for many laboratories growth of the yeast transformants over the weekend might be a rational way to optimize working time. Similarly, frozen competent yeast cells that instantly are ready for transformation minimizes work [44].
In comparison to other available methods for multiple DNA fragments assembly, particularly the Gibson one-step isothermal *in vitro* recombination assembly method, our approach has three advantages. Firstly, YHRC retains high efficiency even at large numbers of DNA fragments. Further, its does not rely on commercial kits. Lastly, the method is not sensitive to the presence buffer components or enzymes and therefore does not require purification of the DNA.

**Acknowledgements**

We thank Dr. Lloyd W. Ruddock, University of Oulu, Finland for supplying pMJS10. This work was funded by the Swedish Science Council (C.A.) and Carl Tryggers Stiftelse (C.A. and M.H.).
References


sensor of external amino acids and F-Box protein Grr1p are required for transcriptional induction of the AGP1 gene, which encodes a broad-specificity amino acid permease. Molecular and cellular biology 19 (1999) 989-1001.


**Figure legends**

**Fig. 1**
pSUMO-YHRC is a vector for the expression of His$_6$-SUMO fusions. A. Map of the pSUMO-YHRC and its cloning site. The vector is a derivative of pET24a and contains a gene encoding His$_6$-SUMO transcribed by a lac-operated T7 RNA polymerase promoter (P$_{lacT7}$). A transcription terminator from bacteriophage T7 (T$_{T7}$) is positioned downstream of the cloning site that is flanked by restriction sites (BamHI, HindIII and XhoI). The lacI gene ensures efficient promoter repression under non-inducing conditions. The plasmid replicates in E. coli by means of a pBR322 origin of replication (ori) and copy number is controlled by rop. Single-copy replication in yeast is mediated by a centromeric and an autonomously replicating sequence (CEN6/ARS209). The TEF promoter (P$_{TEF}$) drives expression of kan that confers resistance to kanamycin in E. coli and G418 in yeast. The lower panel shows the cloning site with relevant restriction sites and 5' and 3' homology regions suitable for YHRC. B. Growth assay of strain DH10B-U either untransformed or transformed to kanamycin resistance using pET24a, pCA528 or pSUMO-YHRC. 10-fold serially diluted cell suspensions were applied onto solid LB medium with the indicated concentrations of kanamycin and incubated at 37°C overnight.

**Fig. 2**
Cloning into pSUMO-YHRC by homologous recombination in yeast. A. Design of PCR primers with 5' and 3' homology regions. B. Schematic representation of homologous recombination in competent yeast between restricted pSUMO-YHRC (BamHI, HindIII and XhoI) and a crude PCR product carrying flanking homology sequences as in A. Yeast transformants form colonies on G418 containing medium after 2 days. Plasmids are rescued and used to transform E. coli to kanamycin resistance from which the cloned expression vector can be isolated. C. Restricted pSUMO-YHRC and a PCR product of SIS1 (as in B) were used to transform yeast to G418 resistance. DNA was analyzed on a 0.8% agarose gel (left panel). The resulting plates with yeast transformants were photographed after two days incubation (right panel).

**Fig. 3**
pSUMO-YHRC functions as a vector for the expression and purification of functional Nanobody$_{GFP}$. A. SDS-Page analysis (Coomassie Brilliant Blue staining) of protein fractions from the expression and purification of Nanobody$_{GFP}$ expressed from pCA856 in BL21-SI cells. Expression was analyzed before and after induction of the T7 promoter with IPTG and NaCl. Soluble protein in the precleared fraction was applied to Ni$_2$-IDA matrix and bound protein was eluted with imidazole. His$_6$-SUMO-Nanobody$_{GFP}$ was cleaved with SUMO protease and the solution was depleted from His$_6$-SUMO by incubation with Ni-IDA matrix. B. Immobilized Nanobody$_{GFP}$ immunoprecipitates EGFP from yeast protein extracts. Protein extracts from yeast cells that either expressed EGFP from pUG35 or that carried an empty vector control (EV) were prepared by bead beating and centrifugation. The precleared lysates were incubated with Sepharose beads and the flow-through fractions (Sepharose FT) were applied to Sepharose beads with conjugated Nanobody$_{GFP}$. Flow-through fractions (Nanobody$_{GFP}$ FT) were
collected. Bound protein was washed and was eluted by boiling in SDS-Page sample buffer. Each fraction was analyzed by SDS-Page and Coomassie Brilliant Blue staining (Upper panel) as well as by immunoblot analysis with GFP specific antibodies (Lower panel). The position of migration of an unidentified yeast protein that binds to Nanobody\textsuperscript{GFP} is marked (*). For immunoblot analysis 10% of the fraction volumes used for the Coomassie Brilliant Blue stained gel was loaded with the exception of the eluted fraction where 1% was loaded.

Fig. 4
YHRC of multiple fragments into pSUMO-YHRC. A. Schematic representation of amplified \textit{lacZ} PCR products with the size of each PCR product indicated (bp). Lines marked 1, 2, 3 and 7 show how the corresponding number of PCR fragments can be cloned by YHRC into pSUMO-YHRC based on 50 bp of homology shared between adjacent fragments. B. Crude PCR reaction mixtures (5 \mu l) from the amplification of \textit{lacZ} products described in A analyzed on a 1% agarose gel. C. Frequency of \textit{E. coli} transformants that display functional expression of \textbeta-galactosidase (Lac+) after YHRC described in A. Data is from 3 experiments and error bars indicate standard deviation.
### Table 1 PCR primers

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGGGGTCCTTTTCATCACGTGC</td>
</tr>
<tr>
<td>2</td>
<td>CGCTGGCCGGTGACCCGGGGGACGAGGCAACGGATCGCTTGGCTCTGCTAACCTAC</td>
</tr>
<tr>
<td>3</td>
<td>GAATCCCATATTATGGCTCAC</td>
</tr>
<tr>
<td>4</td>
<td>CCAACTACGCTTCCTTCTCGGGCTTTGTAGCAGGCTGCAGCTCACAGAAGAAGATTTACGTCG</td>
</tr>
<tr>
<td>5</td>
<td>GATATTATGGCTCACAGAAGAAGATTTACGTCG</td>
</tr>
<tr>
<td>6</td>
<td>GCAGCCGAGCTCACTAGGTTGTTGTTGTTGCTCTGAGCTATTAAAGATTTTTCATCTATAGCC</td>
</tr>
<tr>
<td>7</td>
<td>GATATTATGGCTCACAGAAGAAGATTTACGTCG</td>
</tr>
<tr>
<td>8</td>
<td>GCAGCCGAGCTCACTAGGTTGTTGTTGCTCTGAGCTATTAAAGATTTTTCATCTATAGCC</td>
</tr>
<tr>
<td>9</td>
<td>GATATTATGGCTCACAGAAGAAGATTTACGTCG</td>
</tr>
<tr>
<td>10</td>
<td>GCAGCCGAGCTCACTAGGTTGTTGTTGCTCTGAGCTATTAAAGATTTTTCATCTATAGCC</td>
</tr>
<tr>
<td>11</td>
<td>GATATTATGGCTCACAGAAGAAGATTTACGTCG</td>
</tr>
<tr>
<td>12</td>
<td>GCAGCCGAGCTCACTAGGTTGTTGTTGCTCTGAGCTATTAAAGATTTTTCATCTATAGCC</td>
</tr>
<tr>
<td>13</td>
<td>GATATTATGGCTCACAGAAGAAGATTTACGTCG</td>
</tr>
<tr>
<td>14</td>
<td>GCAGCCGAGCTCACTAGGTTGTTGTTGCTCTGAGCTATTAAAGATTTTTCATCTATAGCC</td>
</tr>
<tr>
<td>15</td>
<td>CTTAAGATTTCTTGTACGAC</td>
</tr>
<tr>
<td>16</td>
<td>CTAGTTATGGCTACGGG</td>
</tr>
<tr>
<td>17</td>
<td>GATATTATGGCTCACAGAAGAAGATTTTTCATCTATAGCC</td>
</tr>
<tr>
<td>18</td>
<td>GCAGCCGAGCTCACTAGGTTGTTGTTGCTCTGAGCTATTAAAGATTTTTCATCTATAGCC</td>
</tr>
<tr>
<td>19</td>
<td>ACGGCAGCCGGTGTGAC</td>
</tr>
<tr>
<td>20</td>
<td>TTGTAGGGCTTAAAACTCAG</td>
</tr>
<tr>
<td>21</td>
<td>TACACCAAGCAGCTCATCG</td>
</tr>
<tr>
<td>22</td>
<td>ACGGCAGCCGGTGTGAC</td>
</tr>
<tr>
<td>23</td>
<td>CAACTTTGCTTACATCAC</td>
</tr>
<tr>
<td>24</td>
<td>CAGCTGTTACGGATTCAC</td>
</tr>
<tr>
<td>25</td>
<td>GATACTGAGAAAGCCTG</td>
</tr>
<tr>
<td>26</td>
<td>TACACATTTGCGGCTTAC</td>
</tr>
<tr>
<td>27</td>
<td>TACGCCTACGCTGAGCAG</td>
</tr>
<tr>
<td>28</td>
<td>AACACAGCGACGGAGAGAG</td>
</tr>
<tr>
<td>29</td>
<td>GACGGTGCTTAAATCACAG</td>
</tr>
<tr>
<td>30</td>
<td>TGCACTTGACGGCAGTAC</td>
</tr>
</tbody>
</table>

### Table 2 Single gene YHRC efficiency

<table>
<thead>
<tr>
<th>Resulting plasmid</th>
<th>Insert expressing</th>
<th>Correct clonesa</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCA856</td>
<td>NanobodyGFP</td>
<td>5/6</td>
</tr>
<tr>
<td>pCA861</td>
<td>Rpo41T920-L1217-HA</td>
<td>4/4</td>
</tr>
<tr>
<td>pCA862</td>
<td>DHFRmutD-FLAG</td>
<td>4/4</td>
</tr>
<tr>
<td>pCA863</td>
<td>3HA-stGnd1</td>
<td>4/4</td>
</tr>
<tr>
<td>pCA892</td>
<td>Sis1</td>
<td>6/6</td>
</tr>
<tr>
<td>pCA893</td>
<td>Ydj1</td>
<td>6/6</td>
</tr>
</tbody>
</table>

**Overall efficiency:** 28/29 (97%)  

*a Number of correct clones of total number of clones analyzed by restriction digests.
Table 3 Primer design for expression with alternative N- and C-termini using pSUMO-YHRC

<table>
<thead>
<tr>
<th>N-terminus</th>
<th>5’ homology sequence in the PCR primer (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>His6-SUMO-</td>
<td>GATATTATTGAGGCTCACAGAGAACAGATTGGTGGG</td>
</tr>
<tr>
<td>Met-Gly-His6-</td>
<td>GAGATATACATATGTTGCATCCATCCATCCATCCACAT</td>
</tr>
<tr>
<td>Untagged</td>
<td>GAAAATAATTTGTTTTAATTTAAGGAAGCATATACAT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C-terminus</th>
<th>3’ homology sequence in the PCR primer (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untagged</td>
<td>GCAGCCGGATCTCAAGTTGGGCGTGGTGTTGGTGCTCGAG</td>
</tr>
<tr>
<td>-His6</td>
<td>GCAGCCGGATCTCAAGTTGGGCGTGGTGTTGGTG</td>
</tr>
</tbody>
</table>
A

5' homology + Forward primer
5' - GATATTATTGAGGCTCACAGAGAACAGATTGGTGGG - 3'

PCR product

GAGCTCGTGGTGGTGGTGGTGACTCTAGGCCGACG

3' homology + Reverse primer

CTCGAGCACCACCACCACCACCACTGAGATCCGGCTGC

B

Crude PCR

BamHI Xhol HindIII

pSUMO-YHRC

Competent Yeast

E. coli

Kan

Expression vector

Yeast

G418

C

<table>
<thead>
<tr>
<th>kB</th>
<th>PCR</th>
<th>pSUMO-YHRC</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- PCR

+ PCR

G418
Before induction
After induction
Precleared
Eluted
Cleaved
Depleted

His<sub>6</sub>-SUMO-
Nanobody<sup>GFP</sup>

A

B

Sepharose Nanobody<sup>GFP</sup>

<table>
<thead>
<tr>
<th>Precleared</th>
<th>FT</th>
<th>FT</th>
<th>Eluted</th>
</tr>
</thead>
<tbody>
<tr>
<td>EV EGFP</td>
<td>EV EGFP</td>
<td>EV EGFP</td>
<td>EV EGFP</td>
</tr>
</tbody>
</table>

kDa

28
39
51
64

His<sub>6</sub>-SUMO
Nanobody<sup>GFP</sup>

EGFP

* 

α-GFP

28

<sup>GFP</sup>
A His6-SUMO lacZ 5' UTR
1 3146 1714 1482
2 890 1296 1060
3 512 428 482 472 468 611
7

B

C

1.0 3.0 1.0

1 2 3 7

L 3146 1214 628 632 890 1060 502 428 482 472 468 611

# lacZ PCR fragments

Lac+ (%)