

# Selection of a T7 promoter mutant with enhanced *in vitro* activity by a novel multi-copy bead display approach for *in vitro* evolution

Siddhartha Paul<sup>1</sup>, Alexander Stang<sup>1</sup>, Klaus Lennartz<sup>2</sup>, Matthias Tenbusch<sup>1</sup> and Klaus Überla<sup>1,\*</sup>

<sup>1</sup>Department of Molecular and Medical Virology, Ruhr-University Bochum, D-44780 Bochum and <sup>2</sup>Department of Molecular Genetics, Institute of Cell Biology (Cancer Research), West German Cancer Center Essen, University of Duisburg-Essen Medical School, 45122 Essen, Germany

Received April 27, 2012; Revised September 5, 2012; Accepted September 18, 2012

## ABSTRACT

***In vitro* evolution of nucleic acids and proteins is a powerful strategy to optimize their biological and physical properties. To select proteins with the desired phenotype from large gene libraries, the proteins need to be linked to the gene they are encoded by. To facilitate selection of the desired phenotype and isolation of the encoding DNA, a novel bead display approach was developed, in which each member of a library of beads is first linked to multiple copies of a clonal gene variant by emulsion polymerase chain reaction. Beads are transferred to a second emulsion for an *in vitro* transcription-translation reaction, in which the protein encoded by each bead's amplicon covalently binds to the bead present in the same picoliter reactor. The beads then contain multiple copies of a clonal gene variant and multiple molecules of the protein encoded by the bead's gene variant and serve as the unit of selection. As a proof of concept, we screened a randomized library of the T7 promoter for high expression levels by flow cytometry and identified a T7 promoter variant with an ~10-fold higher *in vitro* transcriptional activity, confirming that the multi-copy bead display approach can be efficiently applied to *in vitro* evolution.**

## INTRODUCTION

Generation of a large number of variants and high-throughput selection of the best variants during re-iterative rounds has proven to be a successful strategy to improve biological and physical properties of nucleic

acids and proteins (1–7). This approach generally depends on the maintenance of a stable linkage between the genotype and the phenotype during the selection procedure and on efficient determination of the genotype encoding the selected phenotype. Classical examples are phage display, bacterial surface display and yeast surface display, all of which are dependent on cellular expression pathways and the replication capacity of the respective units of selection (e.g. phage, bacteria or yeast cells) carrying the genetic information for variation in the phenotype (8–12). These approaches require the transformation of living cells, limiting the number of independent variants that can be screened. In addition, the requirement for living cells or infectious virions restricts the conditions applicable during the selection step. Other approaches have, therefore, been developed that are performed exclusively *in vitro* (13–17). Some of them can only be used for a rather narrow range of applications. The SELEX approach, for example, is used to optimize the binding of only DNA or RNA molecules to various ligands and to some extent the enzymatic activity of DNA and RNA (18–20). In ribosome and mRNA display strategies, a linkage is formed between variants of the mRNA and the protein encoded by the RNA (15,16,21). This allows selection of proteins with particular binding and enzymatic activities. However, the poor stability of the RNA and the RNA-protein complexes severely restricts the experimental screening conditions (22,23). To improve the stability of the phenotype-genotype linkage, DNA display methods were developed, in which single molecules of DNA were transcribed and translated *in vitro* in picolitre reactors generated by water-in-oil emulsions (17,22–26). The proteins encoded by each of the single DNA molecules contain a constant binding domain for the encoding DNA. Using the O-6-alkylguaninyl-transferase (SNAP) domain and DNA labelled with the SNAP substrate benzylguanine (BG) (27) Stein *et al.* (23)

\*To whom correspondence should be addressed. Tel: +49 234 322 31 89; Fax: +49 234 321 43 52; Email: klaus.ueberla@rub.de

The authors wish to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

were even able to covalently link the encoded protein with its encoding DNA. In addition to the constant DNA binding domain, the proteins encoded by the DNA contain a variable domain that is under selection. After breaking up the water-in-oil emulsion, the complex of a single DNA molecule with the encoded protein is then selected by the affinity of the variable domain to a defined ligand. Amplification of the co-selected DNA allows subsequent rounds of selection and, finally, determination of the genotype. The stoichiometry of this *in vitro* evolution procedure suggests limitations in the stringency that can be used during the affinity-based selection process. Each picolitre reactor is spiked by a single DNA molecule, which is transcribed into multiple copies of RNA, each of which is subsequently translated. Therefore, a vast excess of the protein molecules synthesized in each picolitre reactor is not coupled to the encoding DNA. The single DNA-coupled protein molecule of a picolitre reactor is forced to compete for binding of its variable domain to the ligand with an excess of non-coupled protein molecules of the same picolitre reactor and also protein molecules with lower affinity to the ligand from other picolitre reactors. Thus, a balance has to be found between quantitative recovery of the high-affinity binders and the stringency of the selection conditions.

Instead of breaking up the picolitre reactors, they can also be used as cell-like microcapsules by, for example, applying microfluidic selection strategies (28–30). Although this opens novel ways to screen for enzymatic activities, the experimental conditions of the screening reaction have to be compatible with the *in vitro* transcription–translation (IVTT) reaction. Another practical difficulty is that the genotype responsible for the selected phenotype has to be determined from single DNA molecules (4,7,14,17,22,25,31,32).

Therefore, the aim of our study was to develop a robust and versatile *in vitro* evolution platform, in which the screening reaction is independent from the IVTT, and in which the unit of selection contains multiple copies of the

DNA stably linked to multiple molecules of the encoded protein. In this manuscript, we now describe a novel multi-copy bead display approach, which is based on microbeads as solid units of selection and two consecutive compartmentalization steps. With this bead display approach, we selected a T7 promoter variant, which has a 10-fold higher *in vitro* transcriptional activity compared with the wild-type T7 promoter, and it results in a 2-fold higher protein expression levels in IVTT reactions.

## MATERIALS AND METHODS

### Preparation of templates for emulsion polymerase chain reaction

T7 promoter expression cassettes were amplified with primers T7-s2 and T7Term-a by conventional polymerase chain reaction (PCR). The PCR products were purified by agarose gel electrophoresis, quantified photometrically and used as templates for emulsion PCR. The T7 promoter library was constructed by PCR amplification of a cloned open reading frame encoding a green fluorescent protein (GFP)–SNAP fusion protein with primers T7-wobble-s and T7Term-a (sequences of all primers are shown in Table 1). T7-wobble-s starts with the sequence of the T7-s2 primer followed by the T7 promoter sequence, in which 10 nucleotides flanking the transcription initiation site are completely randomized (Figure 2B) followed by a spacer and primer binding site on the open reading frame (ORF) of GFP–SNAP fusion protein. Thus, the PCR product is a library of randomized T7 promoter variants followed by the GFP and SNAP coding regions. The wild-type T7 promoter construct was also produced using primers T7-consensus (without any randomization in the T7 promoter) and T7Term-a, thereby generating a PCR product identical in sequence to the library, with the exception of the randomized region. The randomization of the 10 nucleotides was confirmed by sequencing the PCR product.

**Table 1.** List of primers

Name	Sequence	Binding Site <sup>a</sup>
T7-s2+TG spacer	5'-Amino-C12-TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGAGCAACCGCACCTGTGG-3'	a
T7-s2	5'-AGCAACCGCACCTGTGG-3'	b
T7-consensus-s	5'-AGCAACCGCACCTGTGGTAATAACGACTCACTATAGGGAGATTGTGAGCGGATAACAATTCCCC-3'	c
T7-wobble-s	5'-AGCAACCGCACCTGTGGTAATAACGACTCACNNNNNNNNNNNTTGTGAGCGGATAACAATTCCCC-3'	d
Luc_ov_Rev	5'-ATGTTTTTGGCGTCTTCCATGGTATATCTCCTTCTAAAGTTAAACAAA-3'	e
T7pro_RP_seq2	5'-GCCGTAGGTCAGGGTGGT-3'	f
T7Term-a	5'-GGATATAGTTCCTCCTTTCAGC-3'	g
T7Term-a (BG) <sup>b</sup>	5'-BG-C6-GGATATAGTTCCTCCTTTCAGC-3'	h
Ovrlprim_NNB_T7	5'-GACCACCCTGACCTACGGC-3'	
Luc_ov_Fwd	5'-ATGGAAGACGCCAAAAACAT-3'	
Luciferase_RP	5'-CATTTAGGTGACACTATAGAATAGGG-3'	
Luc-seq	5'-AGCTTCTGCCAACCGAAC-3'	
Luc-RT-FP1	5'-GGAAGTCGGGGAAGCG-3'	
Luc-RT-RP1	5'-TCTCACACACAGTTCGCCTC-3'	
Anti T7bs2 (BG) <sup>b</sup>	5'-BG-C6-CCACAGGTGCGGTTG-3'	

<sup>a</sup>For primer binding sites, see Figure 2.

<sup>b</sup>These primers are covalently coupled with BG-substrate through a thiol group with C6 spacer in the 5'-end (33); N = A, G, C or T.

### Covalent coupling of oligonucleotides to microbeads or BG

The 5'-amino-modified forward primers T7-s2 or T7-s2+TG spacer were coupled to magnetic beads (Dynabeads Myone carboxylic acid, Invitrogen) according to the manufacturer's guidelines and a previously published report (26) with minor modifications. In all,  $7\text{--}12 \times 10^8$  beads (100  $\mu\text{l}$  of the bead solution) were transferred into 1.5 ml Eppendorf tubes and were washed three times with 100  $\mu\text{l}$  of 0.01 N NaOH and three times with 100  $\mu\text{l}$  deionized water. Beads were then resuspended in 50  $\mu\text{l}$  of 25 mM 2-(*N*-morpholino) ethane sulphonic acid buffer, pH 6.0. Twenty microlitres of a 100  $\mu\text{M}$  5'-amino modified primer solution was added to this solution and incubated on an end-over-end rotator for 30 min. After 30 min, 3 mg of EDC (3-dimethylaminopropyl carbodiimide hydrochloride) dissolved in 30  $\mu\text{l}$  of 25 mM 2-(*N*-morpholino) ethane sulphonic acid buffer (pH 6.0) were added and incubated overnight on a end-over-end rotator at 4°C. Beads were then washed five times with 100  $\mu\text{l}$  TE (pH 8.0) and were stored in 100  $\mu\text{l}$  of TE buffer at 4°C. The reverse primer (T7Term-a) was chemically coupled to BG substrate (New England Biolabs) using a protocol previously published (23) through a custom service of the Sigma Aldrich company (Germany).

### BG binding assay

GFP-SNAP fusion protein was expressed in a cell-free expression system, using the *in vitro* coupled transcription-translation kit TNT T7 Quick for PCR DNA (Promega) according to the manufacturer's guidelines. Eight hundred nanograms of purified PCR product was used for cell-free expression of proteins. Ten microlitres of the translation reaction was incubated with 10  $\mu\text{l}$  of BG-agarose beads (Covalys) for 25 min at room temperature in the dark. For competition assay, 10  $\mu\text{l}$  of the translation reaction was first incubated with 10  $\mu\text{l}$  BG-NH<sub>2</sub> (2 mg/ml) (New England Biolabs) for 25 min at room temperature in the dark. This reaction mixture was then incubated with 10  $\mu\text{l}$  of BG-agarose beads for 25 min at room temperature in the dark. All beads were then analysed under a fluorescence microscope.

### Water-in-oil emulsion PCR

The aqueous phase is composed of 100  $\mu\text{l}$  PCR mix containing the following reagents:  $3 \times 10^9$  copies of linear template DNA,  $7\text{--}12 \times 10^7$  beads coupled with forward primer (T7-s2+TG spacer), 0.01  $\mu\text{M}$  of soluble forward primer (T7-s2), 3  $\mu\text{M}$  of BG coupled reverse primer, 0.5 mM of dNTPs, 15 U of DreamTaq DNA polymerase (Fermentas) and 10  $\mu\text{l}$  of bovine serum albumin (10 mg/ml). The mineral oil mixture containing 2% Abil-90 EM (Evonik) and 0.05% Triton X-100 (Applichem) was prepared as published previously (34). In a 2 ml cryovial with a flat bottom, the aqueous phase was gradually added drop wise to 500  $\mu\text{l}$  of the oil phase for 5 min while stirring constantly on ice at 2000 r.p.m. Stirring was continued for another 3 min to obtain a homogenous emulsion. The emulsion was then dispensed in 50  $\mu\text{l}$  aliquots into ten

200  $\mu\text{l}$  thin-walled PCR tubes. PCR was then performed in a conventional thermocycler with the following temperature profile: 95°C for 3 min, 45 cycles of 95°C for 30 s, 58°C for 30 s, 72°C for 90 s followed by a final extension at 72°C for 10 min. After PCR, all samples were collected in a 1.5 ml Eppendorf tube and were centrifuged at 13 000 r.p.m. for 5 min. The oil was removed from the top, and the beads with yet intact emulsion remained at the bottom. The emulsion was disrupted by adding 100  $\mu\text{l}$  of Bind and Wash buffer [10.0 mM of Tris-HCl (pH 7.5), 1.0 mM of ethylenediaminetetraacetic acid (EDTA), 2.0 M of NaCl] and 1 ml of hexane and vortexing for 10 s. The disrupted oil phase (top) in hexane was removed and discarded, and the beads settled down in the aqueous layer. To completely remove the oil from the bead suspension, hexane extraction was repeated three more times. Residual hexane was removed by centrifugation under vacuum at room temperature for 5 min. The beads were then washed five times with TE pH 8.0 and resuspended in 9  $\mu\text{l}$  of nuclease-free water.

### Cell-free expression of proteins in water-in-oil emulsion

The TNT T7 Quick for PCR DNA system (Promega) was used to perform IVTT reaction in emulsions. Fifty microlitres of reaction mixes were assembled on ice by combining 40  $\mu\text{l}$  of TNT T7 Quick for PCR DNA mix supplemented with 1  $\mu\text{l}$  of methionine (1 mM) and 9  $\mu\text{l}$  of beads suspended in nuclease-free water. The reaction mixture was added to the oil mixture as described previously for the emulsion PCR to form the emulsion IVTT. The emulsion was then incubated at 30°C for 5–120 min for the expression of proteins. The emulsion was kept on ice for 5 min before breaking up the emulsion as previously described (35). One hundred microlitres of breaking buffer (phosphate buffered saline containing 1 mM dithiothreitol and 10 pM BG-NH<sub>2</sub>) was also added during the recovery of beads. The beads were then washed twice with breaking buffer and resuspended in 100  $\mu\text{l}$  of FACS buffer (phosphate buffered saline + 0.5% bovine serum albumin).

### Staining of beads and flow cytometry

One hundred microlitres of a 1:100 dilution of anti-GFP rabbit serum (Invitrogen) in FACS buffer was added to the beads, mixed well and incubated at room temperature for 1 h with intermittent mixing at every 20 min. The beads were then washed twice with 100  $\mu\text{l}$  of FACS buffer and stained with 100  $\mu\text{l}$  of 1:100 diluted anti-rabbit IgG (Invitrogen) for 1 h in the dark at room temperature with intermittent mixing at every 20 min. The beads were then washed two times with 100  $\mu\text{l}$  of FACS buffer, resuspended in 500  $\mu\text{l}$  FACS buffer and analysed in a FACS Calibur (Bectin Dickenson).

For the screening, beads displaying the library were stained with antibodies as mentioned previously and were sorted in a FACS Di VA cell sorter (Bectin Dickenson) after gating on single beads by forward and side scatter. Each round of sorting protocol involved two sorts. Beads were first yield sorted with a speed of 5000–10 000 beads/s. After the yield sort, sorted beads

underwent purity sort, which was performed at a speed below 2000/s. The purity sort beads were directly collected in 50  $\mu$ l nuclease-free water for subsequent PCR amplification.

#### **Generation of beads containing GFP–SNAP and MS2–SNAP**

Beads carrying expression cassettes encoding for GFP–SNAP were used in an emulsion IVTT, which was carried out in the presence of increasing amounts of additional soluble expression cassettes encoding for bacteriophage coat protein MS2 (MS2)–SNAP fusion protein (0,  $1 \times 10^{11}$ ,  $4 \times 10^{11}$  and  $1 \times 10^{12}$  copies). Beads were fluorescently stained for GFP and were analysed by flow cytometry as described previously.

#### **Amplification of DNA from beads**

The quantification of the number of amplicons bound to the beads after the emulsion PCR was determined by quantitative real time PCR with the primers T7s2 and T7pro\_RP\_seq2 using the Quantitect SYBR Green PCR Kit (Qiagen). Approximately 5000 beads were added to the PCR in a final volume of 20  $\mu$ l. The temperature profile was as follows: 94°C for 15 min, 40 cycles of 94°C for 15 s, 60°C for 30 s, 72°C for 30 s and acquiring SYBR Green fluorescence at 75°C were performed followed by standard melting curve analysis of products. Serial dilutions of GFP–SNAP DNA with known copy numbers were run in parallel and were allowed to calculate the number of amplicons bound to the beads.

In the T7 promoter selection experiment, a short fragment containing the T7 promoter variant was amplified directly from beads using the primers T7-s2 and T7pro\_RP\_seq2 with Faststart Taq polymerase (Roche) with limited number of PCR cycles (25 cycles) with the following temperature profile: 94°C for 4 min, 25 cycles of 94°C for 15 s, 62°C for 30 s, 72°C for 30 s and analysed in a 1.5% agarose gel. The longer fragment containing GFP and SNAP-tag was amplified by PCR using the primers ovr1prim\_NNB\_T7 and T7Term-a from a plasmid encoding for the same with Faststart Taq polymerase (Roche). The temperature profile for this PCR was 94°C for 3 min, 35 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 90 s followed by final extension at 72°C for 7 min. The entire cassette for expression was rebuilt by using the smaller fragment and the longer fragment mixed in equimolar concentrations through overlap extension PCR using the primers T7-s2 and T7Term-a with Dreamtaq polymerase (Fermentas). The temperature profile for this PCR was 94°C for 3 min, 35 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 90 s followed by final extension at 72°C for 10 min. The resulting fragment was gel purified and was used for another round of bead display and sorting. After the final round, single beads were sorted directly in a 96-well PCR plate, each containing 10  $\mu$ l of nuclease-free water. The DNA was again amplified from these beads using the primers T7-s2 and T7pro\_RP\_seq2 as described previously and sequenced.

#### **Determination of the T7 promoter activity**

The mutant T7 promoter sequences were PCR amplified using the primers T7-s2 and Luc\_ov\_Rev to confer each of these variants with a segment of the luciferase gene for subsequent overlap extension PCR. The PCR was performed using Faststart Taq polymerase (Roche) with the following temperature profile: 94°C for 4 min, 25 cycles of 94°C for 15 s, 60°C for 30 s, 72°C for 30 s followed by final extension at 72°C for 7 min. The luciferase gene without any promoter was amplified from pCR-Luc plasmid with the primers Luc\_ov\_Fwd and Luciferase\_RP using Faststart Taq polymerase (Roche) with the following temperature profile: 94°C for 4 min, 25 cycles of 94°C for 15 s, 60°C for 30 s, 72°C for 90 s followed by final extension at 72°C for 7 min. Both of these fragments were then mixed in equimolar concentration to perform the overlap extension PCR to produce the entire expression cassettes for the luciferase gene under control of different T7 promoter variants. This PCR was performed using Dreamtaq polymerase (Fermentas) with the following temperature profile: 94°C for 3 min, 35 cycles of 94°C for 30 s, 62°C for 30 s, 72°C for 90 s followed by final extension at 72°C for 10 min. Products were gel purified, and the concentration was determined photometrically.

The luciferase gene under the control of different promoter variants was expressed using the TNT T7 Quick for PCR DNA kit (Promega). One-fifth of one reaction was used for analysis of each template. Two hundred nanograms of DNA was mixed with 8  $\mu$ l of master mix containing 1 mM of methionine in 0.5 ml tubes. Expression of proteins was done at 30°C for 15–120 min. After expression, the reactions were snap chilled in ice for 5 min before depleting the reaction of RNA by adding 0.5  $\mu$ l of RNase IF (New England Biolabs) and incubating the reaction at 37°C for 30 min. Expression levels of luciferase were analysed in a luminometer, using Bright Glo luciferase substrate (Promega) as described by the manufacturer.

#### **Determination of transcriptional and translational activity of mutant and wild-type T7 promoters**

DNA containing the T7 promoter upstream of the luciferase gene was transcribed into RNA by the TNT T7 Quick for PCR DNA kit (Promega) at 30°C for 15–120 min. The reaction was stopped by immediately adding lysis buffer from the QIAamp Blood DNA mini kit (Qiagen) and incubating at 56°C for 10 min. This lysis buffer contains proteinase K, which should completely destroy RNA polymerase and thereby terminate the transcription of DNA. RNA was purified using the same Blood DNA purification kit, which is known to co-purify RNA efficiently. The purified RNA was further treated with TURBO DNA-free™ (Ambion) to completely remove residual DNA. The amount of RNA transcripts was then determined by a quantitative reverse transcriptase (RT)-PCR using primers Luc\_FP1 and Luc\_RP1. RT-PCR was performed using the Quantitect SYBR Green RT-PCR Kit (Qiagen) in a volume of 20  $\mu$ l with 5  $\mu$ l (1:1000 diluted) purified RNA and each primer in a final concentration of 0.5  $\mu$ M. All quantitative PCRs

were performed on a Rotor-Gene 3000 (Corbett Research). The temperature profile was as follows: 50°C for 20 min for reverse transcription, 94°C for 15 min, 40 cycles of 94°C for 15 s, 60°C for 30 s, 72°C for 30 s and acquiring SYBR Green fluorescence at 81°C were performed, followed by standard melting curve analysis of products. Appropriate no RT and no template control samples were used for each experiment. A standard curve was used to calculate copy numbers of the transcripts.

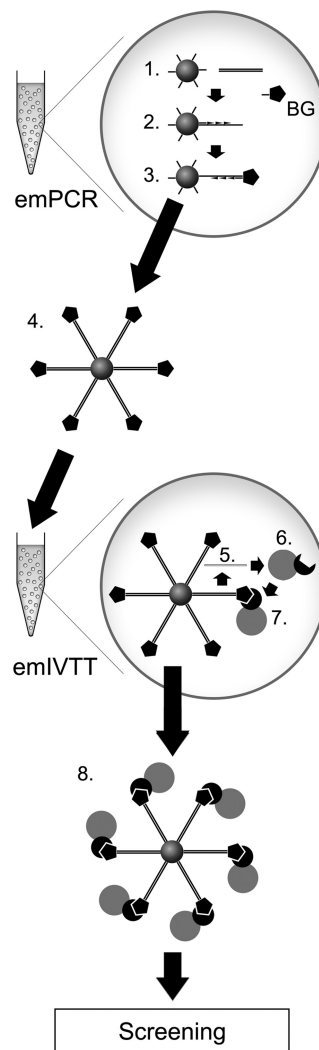
To determine the translational activity of the different T7 promoter sequences, DNA was transcribed to RNA using the Ampliscribe High Yield T7 transcription kit (Epicentre) as recommended by the manufacturer. The purified RNA was further treated with TURBO DNA-free™ (Ambion) to completely remove residual DNA. The RNA transcripts were quantified photometrically. Two nanograms of purified RNA transcripts were then added to the TNT T7 Quick for PCR DNA kit (Promega). Expression of luciferase was quantified as described previously.

## RESULTS

The overall concept of the multi-copy bead display approach is outlined in Figure 1. Template DNA encoding a fusion protein between the protein under selection and SNAP, which reacts specifically with BG (Figure 1), is clonally amplified in picolitre reactors of an emulsion PCR. One of the PCR primers is covalently coupled to magnetic microbeads, whereas the other contains a BG modification. Multiple BG-labelled copies of the same amplicon are, therefore, captured by beads that are present in the same picolitre reactor. If the template DNA consists of a pool of different variants, a library of beads is generated with each bead carrying multiple copies of the same DNA variant and different beads representing different variants. After the amplification, the beads are purified from the emulsion and added to a second emulsion for an IVTT reaction. The SNAP domain of the fusion proteins expressed in each picolitre reactor links the protein under selection to the BG moieties coupled to the encoding DNA by irreversible transfer of the alkyl group from BG to one of its cysteine residues (Supplementary Figure S1) (36). This results in beads carrying multiple copies of the same DNA variant and multiple copies of the protein encoded by the particular DNA variant. The beads can then be screened and selected by different methods, including flow cytometric analysis with antibodies directed against the protein under selection.

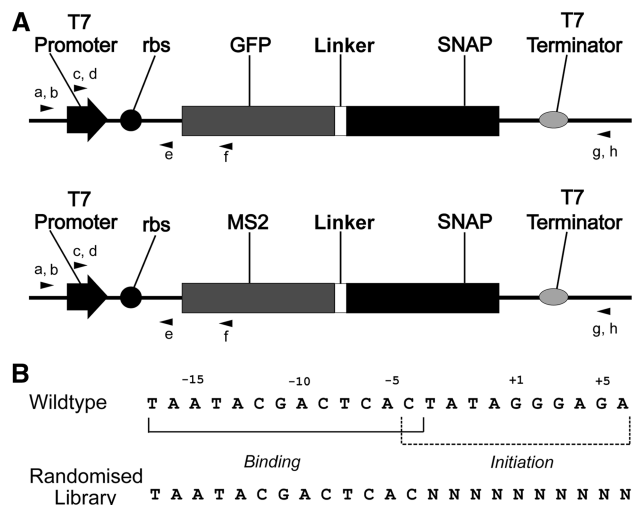
### Coupling of proteins expressed by IVTT to beads

To establish the multi-copy bead display approach, we constructed a template DNA containing the T7 promoter upstream of the open reading frames of the GFP and SNAP (Figure 2A). To confirm the functional activity of the fusion protein, an amplicon spanning the expression cassette from the T7 promoter to the T7 terminator was subjected to an IVTT. BG-coupled agarose beads were

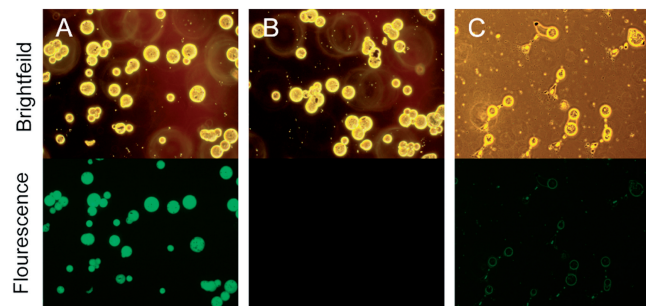


**Figure 1.** Principal steps of the multi-copy bead display approach. The bead display approach involves two steps of compartmentalization using water-in-oil emulsions. In the first step, single template molecules of a DNA library are amplified in each picolitre reactor of an emulsion PCR (emPCR) with a forward primer coupled to microbeads and a reverse primer covalently coupled with BG moieties (1–3). As a result, many copies of the same DNA, each carrying a BG substrate, are immobilized on the bead. These beads are then recovered from the water-in-oil emulsion (4) and are used in a second step in an IVTT reaction also performed in emulsion. The amplicons contain a phage promoter (e.g. T7) upstream of an open reading frame encoding a fusion protein between the protein under selection and the modified DNA repair protein SNAP that reacts specifically with BG moieties incorporated into the amplicons. During the transcription–translation reaction, the fusion proteins synthesized in each picolitre reactor of the emulsion (5,6) are covalently coupled to the BG moieties of the beads present in the same picolitre reactor (7). The beads can then be isolated from the emulsion (8) and can be used in subsequent screening assays.

incubated with an aliquot of the IVTT reaction before fluorescence microscopy. Fluorescently labelled beads were detected in the presence of the IVTT reaction, but not in its absence (Figure 3A and B). If an aliquot of the IVTT reaction is preincubated with an excess of soluble BG, the fluorescence intensity of the beads is greatly diminished (Figure 3C), further confirming the functionality and specificity of the SNAP domain of the fusion protein.



**Figure 2.** Templates for multi-copy bead display. (A) Maps of the different expression cassettes. The open reading frames for GFP or the coat protein of the bacteriophage MS2 are fused in frame through a flexible linker region to the open reading frame of SNAP. The cassettes are flanked by the T7 promoter, a ribosome binding site (rbs) and a T7 terminator sequence. The binding sites of primers used (Table 1) are also indicated (a-h). (B) Sequence of the wild-type and the randomized T7 promoter. The polymerase binding site and the transcription initiation site are indicated.



**Figure 3.** Expression and functionality of the GFP-SNAP fusion protein. PCR products spanning the GFP-SNAP expression cassette shown in Figure 2 were expressed in a cell-free IVTT reaction. BG-coupled agarose beads were incubated with the IVTT reaction for 30 min at room temperature and analysed by fluorescence microscopy. The upper panel shows bright field images of the beads, the lower panel shows the images obtained by fluorescence microscopy. (A) BG-coupled beads after incubation with the GFP-SNAP-primed IVTT reaction. (B) BG-coupled beads without the GFP-SNAP-primed IVTT reaction. (C) GFP-SNAP-primed IVTT reactions were first incubated with an excess of BG before adding the IVTT reaction to the BG-coupled beads.

### Coupling of DNA to beads by emulsion PCR

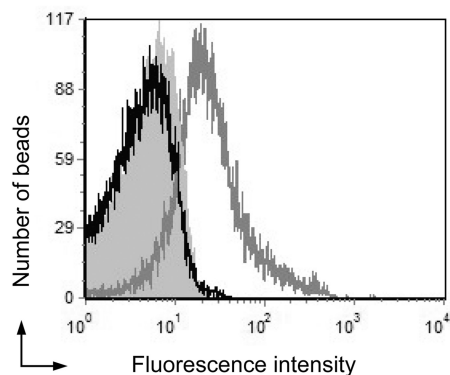
In subsequent experiments, we optimized the PCR amplification of the 1.6 kb expression cassette in emulsion. The emulsion was generated as previously described (34) and contained, as analyzed by light microscopy,  $\sim 10^{10}$ – $10^{11}$  water-in-oil droplets per 100  $\mu$ l of water phase, with the diameter of the droplets mostly ranging from 2 to 5  $\mu$ m. The number of template DNA molecules ( $3 \times 10^9$  copies) added to the emulsion PCR (100  $\mu$ l) was chosen 0.3 times, the minimal number of picolitre reactors. Assuming  $10^{10}$

picolitre reactors and adding  $3 \times 10^9$  template DNA molecules results in a probability of having two template DNA molecules in the same picolitre reactor of 0.09 ( $=0.3^2$ ). This ensures that the majority of picolitre reactors do not contain more than one DNA copy. For efficient amplification in emulsion, a rather high concentration of nucleotides, primers and Taq polymerase had to be used as specified in the ‘Materials and Methods’ section. Using the optimized emulsion PCR conditions,  $\sim 10^8$  magnetic microbeads coupled with the sense primer as capture oligonucleotide were also added per 100  $\mu$ l of water phase. The amount of amplicons captured by the beads could also be increased by including uncoupled sense primer at a concentration of 10 nM. The immobilization of the sense primers by coupling them to the beads might slow down the hybridization kinetics particularly at low numbers of template DNA molecules present during the first cycles of the emulsion PCR. The presence of additional uncoupled sense primer might, therefore, increase the amplification efficiency even at a rather low concentration. To quantify the number of amplicons bound to the beads, the beads were recovered from the emulsion PCR, and 5000 of them were analysed by real time PCR. Under the optimized condition, this revealed  $\sim 900$  copies of amplicons per bead.

### Coupling of DNA and encoded proteins to beads by emulsion PCR and emulsion IVTT

The coupling of the GFP-SNAP fusion protein to the beads from the emulsion PCR is limited by the number of BG containing amplicons captured by the beads. To explore whether sufficient amounts of BG residues were captured, the DNA-coupled beads were incubated with GFP-SNAP fusion protein derived from independent IVTT reactions. In contrast to the experiments with BG-labelled agarose beads (Figure 3), which are larger and contain BG residues directly coupled to the bead, the fluorescence intensity emitted by the GFP bound to DNA-coupled microbeads was not sufficient to reveal binding of the GFP-SNAP fusion protein (data not shown), indicating that the number of BG moieties of the beads is too low.

The number of amplicons detected per bead (900) suggested that a substantial fraction of the capture oligonucleotides of the bead was not fully extended. This allows to increase the number of BG molecules of the beads by simply hybridizing a BG-labelled oligonucleotide to the capture oligonucleotide of the bead after the emulsion PCR. Emulsions for the IVTT reaction were, therefore, formed with the DNA-coupled beads hybridized to the BG-labelled oligonucleotide. To avoid the possibility of covalent linkage of an excess of expressed SNAP fusion proteins of one picolitre reactor to free BG moieties on beads from other reactors during breaking up of the emulsion, excess amounts of free BG were added at the recovery step of the beads from the IVTT emulsion reaction. To further enhance the signal and to make the screening system independent from the fluorescence of the GFP protein, beads were stained with anti-GFP antibodies and a fluorescently labelled secondary



**Figure 4.** Detection of the GFP-SNAP fusion protein after multi-copy bead display by flow cytometry. Dark grey line: beads were loaded with DNA encoding GFP-SNAP and added to a compartmentalized IVTT reaction. After recovery of the beads from the IVTT reaction, the beads were stained with an anti-GFP antibody and a fluorescently labelled secondary antibody before flow cytometric analysis. After gating on the single beads, the fluorescence intensity of single beads is plotted as a histogram. Filled light grey: beads displaying GFP-SNAP proteins were incubated with fluorescently labelled secondary antibody in the absence of the anti-GFP antibody. Black line: beads were processed through emulsion PCR and IVTT without a GFP-SNAP DNA template and then stained for GFP.

antibody. Using the optimized emulsion PCR, emulsion IVTT and staining conditions, a population of GFP-positive beads can clearly be detected by flow cytometry with an anti-GFP antibody and a fluorescently labelled secondary antibody (Figure 4). The specificity of the signal was confirmed by omitting the GFP-SNAP DNA for the emulsion PCR or the anti-GFP antibody for the staining (Figure 4). Thus, DNA and the encoded protein can be coupled to beads at sufficient levels by emulsion PCR and emulsion IVTT.

#### Clonality and sensitivity of the multi-copy bead display approach

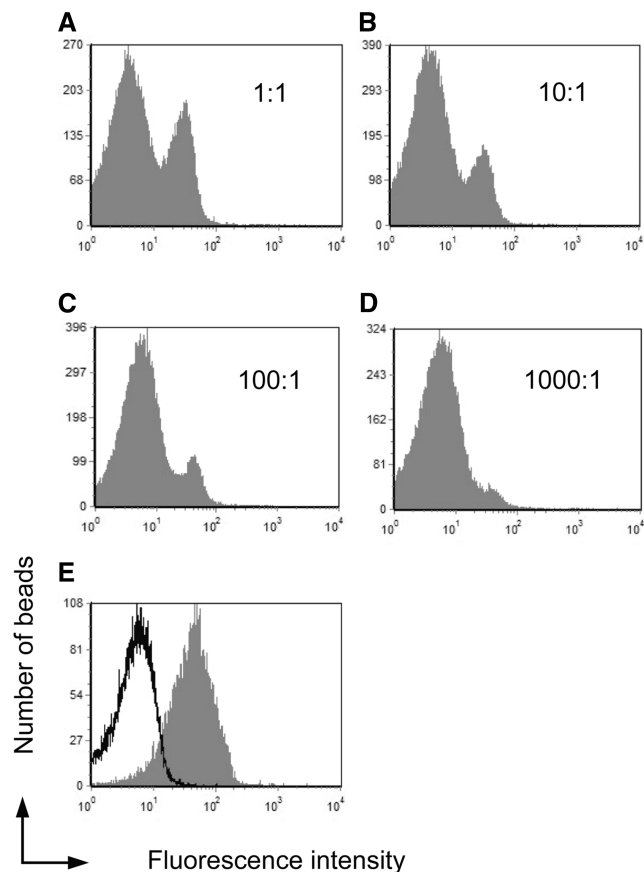
After having established the principal methods for the multi-copy bead display, mixing experiments were performed to confirm clonal amplification of the template and to determine the sensitivity for the detection of GFP-SNAP in the presence of an excess of a control SNAP fusion protein. An expression cassette encoding the coat protein of the bacteriophage MS2 fused to SNAP was used as a control construct (Figure 2A). Adding same amounts of MS2-SNAP or GFP-SNAP separately to emulsion PCR and emulsion IVTT resulted in similar expression levels as determined by western blot analysis with a SNAP antiserum (Supplementary Figure S2). The MS2-SNAP template was then mixed at different ratios with the template encoding GFP-SNAP. After multi-copy bead display, the beads were stained for GFP. At a 1:1 ratio, there are clearly two populations of beads, one being GFP positive, the other GFP negative (Figure 5A). Reducing the amount of GFP-SNAP template gradually decreased the GFP-positive bead population, but not the mean fluorescence intensity of the positive beads. However, even at a 1000-fold excess of the MS2-SNAP template, a GFP-positive bead population remains detectable (Figure 5D). These results

indicate that at the copy numbers of template added to the emulsion PCR, each bead is loaded with DNA encoding only one of the SNAP fusion proteins. It also confirms that coupling of SNAP fusion protein produced in one picolitre reactor to the bead of another picolitre reactor does not occur at detectable levels.

To evaluate whether a population of beads carrying MS2-SNAP and MS2-GFP could be discriminated from populations carrying either MS2-SNAP or MS2-GFP, we generated beads carrying both proteins by adding increasing amounts of soluble DNA encoding for MS2-SNAP to an IVTT in emulsion with beads carrying the GFP-SNAP expression cassette. After IVTT in the same picolitre reactor, both proteins compete for binding to the BG residues on the beads. In the absence of MS2-SNAP competitor DNA (Figure 6A) and in the presence of  $10^{11}$  copies of MS2-SNAP DNA (Figure 6B), a single GFP-positive bead population is detectable. Increasing the amount of MS2-SNAP DNA to  $4 \times 10^{11}$  and  $1 \times 10^{12}$  copies reduces the mean fluorescence intensity of the GFP-positive bead population by a factor of  $\sim 2$  and to background levels (Figure 6C and D). This experiment clearly shows that a population of beads carrying both proteins can be detected if present and confirms that clonal amplification occurs in our multi-copy bead display.

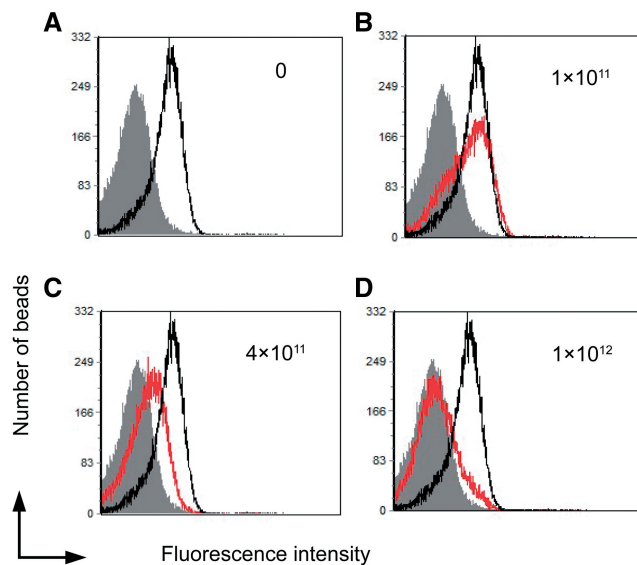
#### Selection of T7 promoter variants with enhanced activity

As a first application of the multi-copy bead display approach, we wanted to select a T7 promoter with enhanced activity in IVTT reactions. As the wild-type T7 promoter evolved for efficient expression of a phage gene in prokaryotes, we reasoned that there may be promoters with higher activity in IVTT reactions. We, therefore, randomized 10 nucleotides spanning the T7 promoter transcription initiation site (37) of the GFP-SNAP template (Figure 2B) by a PCR-based approach. The maximum complexity of the template library is  $4^{10}$  ( $= 1,048,576$ ). The multi-copy bead display was then performed by adding  $3 \times 10^9$  copies of the template library and  $10^8$  beads. The low number of DNA copies ( $\sim 0.3$  copies per picolitre reactor) was used to avoid that more than one template DNA molecule is present in each picolitre reactor. Only  $10^8$  beads were added, as this number was sufficient to cover the complexity of the library ( $10^6$ ). To avoid saturation of GFP-SNAP expression in the emulsion IVTT reaction, the reaction time for the IVTT was reduced to 30 min, which resulted in half-maximal mean fluorescence intensities after staining for GFP and flow cytometric analysis (Supplementary Figure S3). To select beads with the highest GFP fluorescence and to avoid co-purification of negative beads, a two-step flow cytometric sorting strategy was used. During a first yield sort,  $\sim 1\%$  of the beads with the highest fluorescence intensities was selected and then submitted to a more accurate purity sort selecting again  $\sim 1\%$  of the remaining beads with the highest fluorescence intensity. Fluorescence intensity distribution of the beads during yield sort and purity sort clearly reveals an increase in mean fluorescence intensity from yield to purity sort (Figure 7). The efficacy of the purity sort was not directly analysed, as the number of recovered beads was low ( $< 10^4$ )



**Figure 5.** Detection of GFP-SNAP containing beads in the presence of an excess of a competing SNAP fusion protein. (A–D) Emulsion PCRs containing beads were performed with the indicated ratios of MS2-SNAP to GFP-SNAP template DNA. Beads were then transferred to compartmentalized IVTT reactions. After recovery of the beads from the IVTT reaction, the beads were stained with an anti-GFP antibody and a fluorescently labelled secondary antibody before flow cytometric analysis. After gating on the single beads, the fluorescence intensity of single beads is plotted as a histogram. (E) Overlay of histograms of beads loaded with either GFP-SNAP (filled grey area) or MS2-SNAP (black line).

and would have been further reduced. Instead, the beads recovered from purity sort were used to amplify the randomized T7 promoter region by PCR. The amplified T7 promoter region was fused by overlap extension PCR to the remaining GFP-SNAP expression cassette regenerating the entire GFP-expression cassette shown in Figure 2A. This PCR product was excised from the gel, and a second round of multi-copy bead display was performed. During the subsequent yield sort, a small population of GFP-positive beads was detectable (Figure 7). Analysing the 1% of beads selected in the yield sort during the purity sort demonstrates a further enrichment of GFP-positive beads. The T7 promoter region of the beads recovered from the purity sort during the second round was amplified as described previously and was used to reconstitute the entire expression cassette. Single GFP-positive beads obtained after a third round of the multi-copy bead display were directly sorted into microtitre PCR plates. The T7 promoter region could be successfully amplified from 31 of the 40 single beads selected.



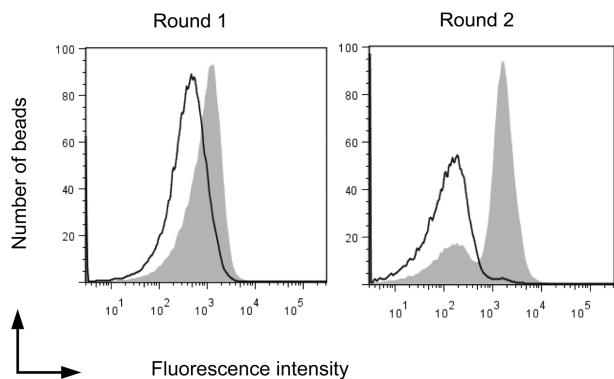
**Figure 6.** Fluorescence intensities of beads containing GFP-SNAP and MS2-SNAP. Beads loaded with the GFP-SNAP expression cassette by emulsion PCR were added to emulsion IVTT in the absence (black lines) or presence of  $1 \times 10^{11}$  (B),  $4 \times 10^{11}$  (C) or  $4 \times 10^{12}$  (D) copies of the MS2-SNAP expression cassette (red lines). After recovery of the beads from the IVTT reaction, the beads were stained with an anti-GFP antibody and a fluorescently labelled secondary antibody before flow cytometric analysis. Overlay of histograms of the fluorescence intensities of the different bead populations are shown in comparison with the background staining of beads processed through emulsion PCR and emulsion IVTT in the absence of T7 promoter expression cassettes (filled grey area).

### Characterization of T7 promoter variants

The activity of the selected T7 promoter regions was determined by fusing the promoter to the luciferase gene by overlap extension PCR. The gel-purified PCR fragments were then used in IVTT reactions, which were stopped after 15 min to avoid saturation effects. About half of the amplifiable T7 promoter regions had an activity that was as high or higher than the wild-type T7 promoter included as a positive control (Supplementary Figure S4). We also determined the activity of 40 randomly picked clones of the T7 promoters of the original library. The four random clones with the highest activity had  $\sim 20\%$  of the wild-type promoter's activity (data not shown). Thus, the high activity of the T7 promoters selected by the multi-copy bead display is not a random effect but is because of the selection procedure.

The six selected T7 promoters with the highest activity were then sequenced directly from amplicons. As two sequences were represented by two independently selected beads, four different T7 promoter sequences were obtained with a higher activity than the wild-type promoter sequence (Table 2). The only nucleotide of the randomized region conserved between the wild-type T7 promoter and all four selected sequences is the guanine at position +1 (Table 1). The fact that the sequence of the randomized nucleotides of selected T7 promoters could be obtained unambiguously from amplicons also confirms that the clonality is maintained at sufficient levels throughout the entire multi-copy bead display procedure.





**Figure 7.** Selection of variants from the T7 promoter library. Beads were loaded with amplicons from the randomized T7 promoter library upstream of the GFP-SNAP coding region by emulsion PCR, transferred to compartmentalized IVTT reactions and stained for GFP. Overlays of histograms of flow cytometric analyses during yield (black line) and purity sort (grey area) of the first and second round of selection are shown.

**Table 2.** Sequence of T7 promoter variants

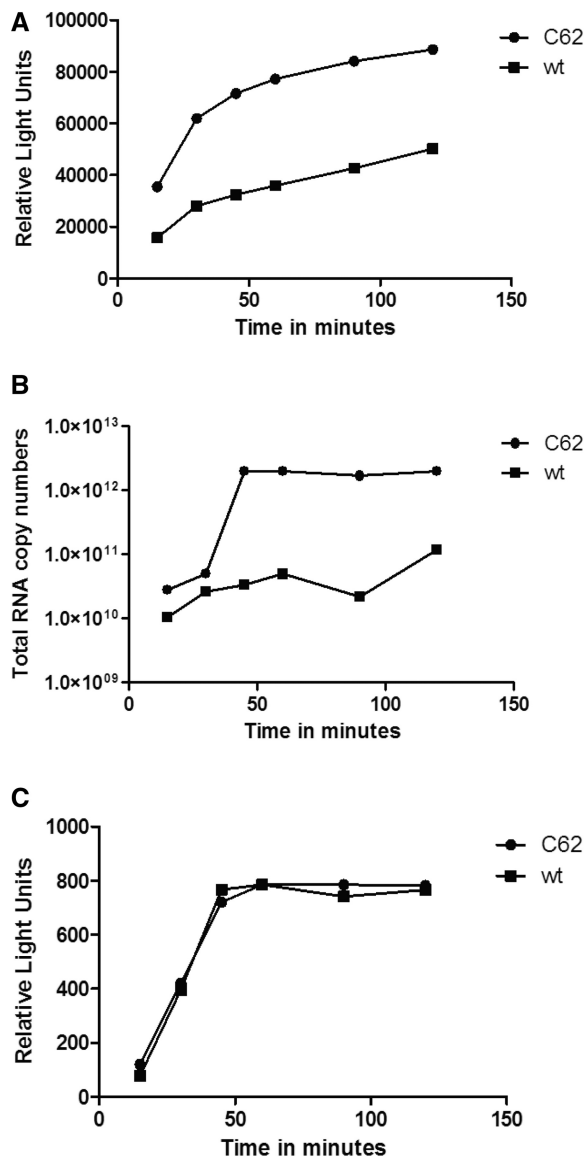
Variants	Sequence	Activity Relative to wild-type T7 promoter <sup>a</sup>
Wild-type	TAATACGACTC <b>ACTTATAGGG</b> GAGA	100
A6, B1	TAATACGACTC <b>ACTCCGGCAATC</b>	151
A7	TAATACGACTC <b>ACTTCGGCAACC</b>	118
A10	TAATACGACTC <b>ACTACGAGCGGCA</b>	138
C15, C62	TAATACGACTC <b>ACTAATCGCGGAG</b>	193

Nucleotides from randomized library are in bold; conserved residue (predicted transcription start site) is underlined.

<sup>a</sup>Mean values from two independent experiments.

The T7 promoter variant with the highest activity (C62) was further characterized in direct comparison with the wild-type T7 promoter. The luciferase activity was determined at different time points after starting the IVTT. The C62-T7 promoter led to at least 2-fold higher protein expression levels at each of the time points tested (Figure 8A). Luciferase expression levels obtained by wild-type T7 promoter after 1 h were already obtained with the C62-T7 promoter after 15 min.

Although we had randomized the transcription initiation site of the T7 promoter, it could not be excluded that the modifications affected not only the transcriptional activity but also the translational efficiency. Therefore, we analysed both activities separately. To assess the relative transcriptional activity, RNA was isolated from IVTT reactions with the wild-type and the C62-T7 promoter at different time points, and the amount of RNA was determined by real time PCR. After a lag phase of 30 min, RNA levels were at least 10-fold higher for the C62-T7 promoter (Figure 8B). To analyse whether the promoter affects translational efficiency, we added defined copy numbers of *in vitro* transcribed luciferase encoding RNA to the IVTT without any DNA. At different time points, translation was stopped, and luciferase expression levels were determined. No differences



**Figure 8.** Characterization of the C62 T7 promoter variant. (A) Time course of the luciferase activity in IVTT reactions of the luciferase gene driven by the C62 or wild-type (wt) T7 promoter. (B) Time course of luciferase RNA levels in *in vitro* transcription reactions of the luciferase gene driven by the C62 or wild-type T7 promoter. (C) Time course of luciferase activity of *in vitro* translation reactions of equal amounts of luciferase mRNA containing the first transcribed nucleotides either of the wild-type or of the C62 T7 promoter. Results of one representative experiment of two independent experiments are shown.

in luciferase activities between wild-type and C62-T7 promoter were observed (Figure 8C), indicating that the C62-T7 promoter has a substantially higher transcriptional activity than the wild-type promoter without affecting the translational efficiency.

## DISCUSSION

One of the T7 promoters selected by the bead display approach led to 10-fold higher luciferase RNA levels without any changes in the concentration of the reagents

required for the IVTT reaction. As the novel T7 promoter was initially selected by enhanced expression of the GFP–SNAP protein, it seems unlikely that the enhanced transcriptional activity of the novel promoter is restricted to the luciferase gene. We expect that the novel T7 promoter sequence may, therefore, be used to optimize *in vitro* RNA synthesis and reduce costs in a variety of biotechnological and therapeutic applications.

As exemplified by the successful T7 promoter screen, the multi-copy bead display approach combines a number of advantages that should provide broad applicability and robustness. Microbeads were used as the unit of selection. As they are rather inert, they can be transferred from one compartmentalized reaction to the next. This allowed to first generate beads in emulsion PCR reactions carrying multiple copies of the DNA variant and then to transfer the beads with its DNA load to compartmentalized IVTT reactions. The beads could then be recovered from the compartmentalized IVTT reaction, which now contain multiple copies of the DNA variant and multiple molecules of the proteins encoded by each bead's DNA. The average number of DNA molecules coupled to each bead was determined to be in the range of 900. The precise number of protein molecules also coupled to the beads is not known. However, as the sensitivity of flow cytometric detection of cell surface protein expression is in the range of 1000 molecules per cell, a similar number of protein molecules should be displayed on the beads. In the present study, we used fluorescently labelled ligands to the encoded protein to select beads containing the highest amount of protein by flow cytometry. Mixing templates encoding either a binding or a non-binding partner of the ligand clearly revealed two different populations of the beads (Figure 5), indicating that a flow cytometric screen could also be used to identify protein variants with increased binding affinity to a fluorescently labelled ligand. It should also be feasible to transfer the protein and DNA loaded beads recovered from the IVTT emulsion to cell-like microcapsules for independent enzymatic screening reactions and subsequent microfluidic selection methods (28,29) extending the potential applicability of the multi-copy bead display to all kind of enzymatic optimizations in the field of synthetic biology.

In comparison with previously described DNA display approaches (14,17,22–26,31), our multi-copy bead display should have a number of advantages. The coupling of the DNA and the encoded protein on beads allows easy removal of IVTT reagents by a simple magnetic separation step. Thus, the selection conditions can be chosen independent from the IVTT reaction providing greater flexibility and avoidance of high background signals. In contrast to the DNA display approach, encoded proteins not coupled to DNA can also be removed avoiding competition and interference during the selection step. As multiple molecules of the same protein are concentrated on a single bead, the signal used for the selection step is also strong enough for flow cytometric screening. This is advantageous, as it is possible to precisely define what percentage of the library is selected

during each round. In addition, flow cytometric screening has the potential to be multi-parametric. For example, beads could be selected that are positive for one ligand but are negative for a second ligand. As the number of beads recovered at each step can be easily determined, it is also possible to monitor maintenance of the complexity of the library. A practical advantage of having multiple copies of DNA on each bead is that the genotype of single beads as units of selection can be easily determined by a one-step PCR and subsequent sequence analysis of the amplicons without the need for any additional cloning steps. However, side-by-side comparisons are needed to determine the reliability and success rate of the different display approaches.

Another interesting strategy to link multiple copies of DNA with multiple copies of the encoded protein has recently been described (30). In this approach without beads, the droplets from emulsion PCR were fused with droplets containing IVTT reagents and reagents for the subsequent fluorogenic microfluidic screening assay. One of the difficulties in this approach might be to prevent interference of PCR or IVTT reagents with the screening assay, as removal of reagents is not possible. In contrast and as discussed previously, the use of beads as solid units of selection allows to perform the screening assay independently from the reaction conditions of the emulsion PCR and the IVTT reaction.

The multi-copy bead display approach does not involve any bacterial cloning step. This is particularly important for the generation of the DNA library, as inefficient transformation can constitute a bottleneck limiting the complexity of the library [reviewed in (38)]. Instead of bacterial cloning steps, PCR amplifications were used to insert a chemically synthesized oligonucleotide with randomized residues into a 1.6 kb expression cassette. In this first proof-of-principle study, the library screened had a complexity of  $\sim 10^6$ . Under the experimental conditions described, the  $10^8$  beads used in the multi-copy bead display approach limit the maximum number of variants that can be screened. However, as our emulsions contain at least  $10^{10}$  picolitre reactors per 100  $\mu$ l of water phase, and as the number of template molecules added is  $3 \times 10^9$ , a 10–100-fold increase of the number of beads and simple upscaling of the total volumes should allow to increase the number of beads that can be recovered from the IVTT reaction for subsequent screening steps by several orders of magnitude. The number of beads that can be processed in the selection step might then become the limiting factor.

Because of the advantages discussed previously, the multi-copy bead display approach should be a useful addition to existing *in vitro* evolution methodologies, with broad applicability in protein engineering and synthetic biology.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figures 1–4.

## ACKNOWLEDGEMENTS

The authors would like to thank Bastian Grewe and Klaus Sure for providing the MS2-SNAP plasmid and technical assistance, respectively.

## FUNDING

FoRUM program of the Medical Faculty of the Ruhr-University Bochum and institutional resources. Funding for open access charge: Institutional resources.

*Conflict of interest statement.* The authors S.P., A.S. and K.Ü. are inventors on a patent application that is currently under preparation.

## REFERENCES

- Kopylov, A.M. and Spiridonova, V.A. (2000) Combinatorial chemistry of nucleic acids: SELEX [Article in Russian]. *Mol. Biol.*, **34**, 1097–1113.
- Wilson, D.S. and Szostak, J.W. (1999) In vitro selection of functional nucleic acids. *Ann. Rev. Biochem.*, **68**, 611–647.
- Griffiths, A.D. and Tawfik, D.S. (2006) Miniaturising the laboratory in emulsion droplets. *Trends Biotechnol.*, **24**, 395–402.
- Griffiths, A.D. and Tawfik, D.S. (2003) Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization. *EMBO J.*, **22**, 24–35.
- Lee, Y.F. (2002) Investigating the target recognition of DNA cytosine-5 methyltransferase HhaI by library selection using in vitro compartmentalisation. *Nucleic Acids Res.*, **30**, 4937–4944.
- Cohen, H.M., Tawfik, D.S. and Griffiths, A.D. (2004) Altering the sequence specificity of HaeIII methyltransferase by directed evolution using in vitro compartmentalization. *Protein Eng. Des. Sel.*, **17**, 3–11.
- Ghadessy, F.J., Ong, J.L. and Holliger, P. (2001) Directed evolution of polymerase function by compartmentalized self-replication. *Proc. Natl Acad. Sci. USA*, **98**, 4552–4557.
- McCafferty, J., Griffiths, A.D., Winter, G. and Chiswell, D.J. (1990) Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*, **348**, 552–554.
- Bratkovic, T., Lunder, M., Popovic, T., Kreft, S., Turk, B., Strukelj, B. and Urleb, U. (2005) Affinity selection to papain yields potent peptide inhibitors of cathepsins L, B, H, and K. *Biochem. Biophys. Res. Commun.*, **332**, 897–903.
- Georgiou, G., Stathopoulos, C., Daugherty, P.S., Nayak, A.R., Iverson, B.L. and Curtiss, R. (1997) Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. *Nat. Biotechnol.*, **15**, 29–34.
- Lunder, M., Bratkovic, T., Doljak, B., Kreft, S., Urleb, U., Strukelj, B. and Plazar, N. (2005) Comparison of bacterial and phage display peptide libraries in search of target-binding motif. *Appl. Biochem. Biotechnol.*, **127**, 125–131.
- Boder, E.T. and Wittrop, K.D. (1997) Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.*, **15**, 553–557.
- Roberts, R.W. (1997) RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc. Natl Acad. Sci. USA*, **94**, 12297–12302.
- Tawfik, D.S. and Griffiths, A.D. (1998) Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol.*, **16**, 652–656.
- Roberts, R.W. (1999) Totally in vitro protein selection using mRNA-protein fusions and ribosome display. *Curr. Opin. Chem. Biol.*, **3**, 268–273.
- Lipovsek, D. and Plückthun, A. (2004) In-vitro protein evolution by ribosome display and mRNA display. *J. Immunol. Methods*, **290**, 51–67.
- Doi, N. and Yanagawa, H. (1999) STABLE: protein-DNA fusion system for screening of combinatorial protein libraries in vitro. *FEBS Lett.*, **457**, 227–230.
- Brody, E.N. and Gold, L. (2000) Aptamers as therapeutic and diagnostic agents. *Rev. Mol. Biotechnol.*, **74**, 5–13.
- Geyer, C.R. and Brent, R. (2000) Selection of genetic agents from random peptide aptamer expression libraries. *Methods Enzymol.*, **328**, 171–208.
- Ellington, A.D. and Szostak, J.W. (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature*, **346**, 818–822.
- Keefe, A.D. and Szostak, J.W. (2001) Functional proteins from a random-sequence library. *Nature*, **410**, 715–718.
- Bertschinger, J. and Neri, D. (2004) Covalent DNA display as a novel tool for directed evolution of proteins in vitro. *Protein Eng. Des. Sel.*, **17**, 699–707.
- Stein, V., Sielaff, I., Johnsson, K. and Hollfelder, F. (2007) A covalent chemical genotype-phenotype linkage for in vitro protein evolution. *Chembiochem*, **8**, 2191–2194.
- Yonezawa, M. (2003) DNA display for in vitro selection of diverse peptide libraries. *Nucleic Acids Res.*, **31**, e118.
- Yonezawa, M. (2004) DNA Display of biologically active proteins for in vitro protein selection. *J. Biochem.*, **135**, 285–288.
- Kojima, T., Takei, Y., Ohtsuka, M., Kawarasaki, Y., Yamane, T. and Nakano, H. (2005) PCR amplification from single DNA molecules on magnetic beads in emulsion: application for high-throughput screening of transcription factor targets. *Nucleic Acids Res.*, **33**, e150.
- Jongsma, M.A. and Litjens, R.H. (2006) Self-assembling protein arrays on DNA chips by auto-labeling fusion proteins with a single DNA address. *Proteomics*, **6**, 2650–2655.
- Dittrich, P.S., Jahnz, M. and Schwill, P. (2005) A new embedded process for compartmentalized cell-free protein expression and on-line detection in microfluidic devices. *Chembiochem*, **6**, 811–814.
- Song, H. and Ismagilov, R.F. (2003) Millisecond kinetics on a microfluidic chip using nanoliters of reagents. *J. Am. Chem. Soc.*, **125**, 14613–14619.
- Fallah-Araghi, A., Baret, J.C., Ryckelynck, M. and Griffiths, A.D. (2012) A completely in vitro ultrahigh-throughput droplet-based microfluidic screening system for protein engineering and directed evolution. *Lab Chip*, **12**, 882–891.
- Sumida, T., Doi, N. and Yanagawa, H. (2009) Bicistronic DNA display for in vitro selection of Fab fragments. *Nucleic Acids Res.*, **37**, e147.
- Chen, Y., Mandic, J. and Varani, G. (2008) Cell-free selection of RNA-binding proteins using in vitro compartmentalization. *Nucleic Acids Res.*, **36**, e128.
- Connolly, B.A. and Rider, P. (1985) Chemical synthesis of oligonucleotides containing a free sulphhydryl group and subsequent attachment of thiol specific probes. *Nucleic Acids Res.*, **13**, 4485–4502.
- Williams, R., Peisajovich, S.G., Miller, O.J., Magdassi, S., Tawfik, D.S. and Griffiths, A.D. (2006) Amplification of complex gene libraries by emulsion PCR. *Nat. Methods*, **3**, 545–550.
- Miller, O.J., Bernath, K., Agresti, J.J., Amitai, G., Kelly, B.T., Taly, V., Magdassi, S., Tawfik, D.S. and Griffiths, A.D. (2006) Directed evolution by in vitro compartmentalization. *Nat. Methods*, **3**, 561–570.
- Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H. and Johnsson, K. (2003) A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.*, **21**, 86–89.
- Rong, M., He, B., McAllister, W.T. and Durbin, R.K. (1998) Promoter specificity determinants of T7 RNA polymerase. *Proc. Natl Acad. Sci. USA*, **95**, 515–519.
- Hoogenboom, H.R. (2005) Selecting and screening recombinant antibody libraries. *Nat. Biotechnol.*, **23**, 1105–1116.