**SHORT COMMUNICATION**

**Selective gene amplification**

Bernard T. Kelly1,3 and Andrew D. Griffiths1,2,4

1Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK and 2Institut de Science et d’Ingénierie Supramoléculaires (ISIS), Université Louis Pasteur, CNRS UMR 7006, 8 allée Gaspard Monge, BP 70028, F-67083 Strasbourg Cedex, France

3Present address: Cambridge Institute for Medical Research, Hills Road, Cambridge CB2 2XY, UK

4To whom correspondence should be addressed.

E-mail: griffiths@isis.u-strasbg.fr

We describe a system for directed evolution based on in vitro compartmentalisation in which amplification of a gene is coupled to the formation of product by the enzyme it encodes. This approach mimics the process of natural selection; ‘fitter’ genes — encoding more efficient enzymes — have more ‘offspring’. It allows selection for any activity so long as a product-specific ligand (e.g. an antibody) is available.

Keywords: directed evolution/emulsion/in vitro compartmentalisation/IVC/PCR

**Introduction**

Recently, in vitro compartmentalisation (IVC) (Tawfik and Griffiths, 1998) has emerged as a powerful new technique for the directed evolution of proteins and for massively parallel polymerase chain reactions (PCRs)—for a review see (Griffiths and Tawfik, 2006). In IVC, individual genes are isolated within microcompartments (aqueous microdroplets) formed by dispersing an aqueous phase in an oil phase to form a water-in-oil emulsion (Tawfik and Griffiths, 1998); since the aqueous phase is composed of a coupled transcription–translation system [or a transcription system alone, for production of ribozymes (Agresti et al., 2005; Levy et al., 2005)], the isolated genes may be expressed and the RNAs and proteins they encode are kept within the same microcompartment. In this way, a linkage between genotype and phenotype may be achieved.

Two variants of IVC have been used to select for catalysis. In the first, the product of the catalytic reaction ends up linked to the gene, either directly (Tawfik and Griffiths, 1998; Lee et al., 2002; Cohen et al., 2004; Doi et al., 2004; Agresti et al., 2005), or via a microbead (Griffiths and Tawfik, 2003; Levy et al., 2005) and it is the genes themselves that are selected after breaking of the emulsion. Very large libraries (>10^8) of mutant genes may be screened in this manner. In the second, the unbroken emulsion droplets are sorted, together with the genes they contain using a fluorescence activated cell sorter (FACS), allowing direct sorting of compartments based on the turnover of a fluorogenic substrate (Aharoni et al., 2005; Mastrobattista et al., 2005).

Here, we propose an extension of the IVC technique, which we term as ‘Selective Gene Amplification’ or SGA. Our approach is to link the amplification of a gene to the product turnover mediated by the enzyme or ribozyme it encodes, in a manner which mimics the process of natural selection; ‘fitter’ genes — encoding higher-turnover enzymes — have more ‘offspring’ and so become enriched over successive generations.

Another variant of IVC, termed compartmentalised self-replication (CSR), which also couples the activity of the enzyme to the amplification of the gene that encodes it has also been described (Ghadessy et al., 2001; Ghadessy et al., 2004). CSR is based on a polymerase amplifying its own gene in an emulsion compartment, and this technique can only be used either to select polymerases or to select enzymes whose activity can be coupled to the production of substrates for the polymerase reaction (e.g. nucleoside diphosphate kinase).

SGA, however, may be used to select a wide range of reactions, requiring only an antibody (or other ligand) which specifically binds the product of the reaction being selected.

**Results and discussion**

SGA begins with a conventional IVC system in which a gene library encoding enzyme variants is immobilised on microbeads such that, statistically, few beads carry more than one gene. The microbeads are dispersed into individual microdroplets in a water-in-oil emulsion and the genes are transcribed and translated in the microdroplets to synthesise the encoded enzyme. In SGA, the microbeads also have multiple copies of enzyme substrate attached such that the product of the desired enzymatic reaction also remains attached to the bead as outlined in Fig. 1, steps a to d (Levy et al., 2005). Alternatively, the substrate can initially be in solution in the droplets and the product molecules (and any unreacted substrate) become attached to the microbead only after enzymatic conversion, as described previously using caged-biotinylated substrates, which only bind to streptavidin-coated beads after uncaging by UV-irradiation (Griffiths and Tawfik, 2003). The novel feature of SGA is that after the first emulsion is broken, the recovered beads are incubated with a product-specific ligand, for example an antibody, which has been derivatised with an oligonucleotide primer via a cleavable disulphide linkage (Fig. 1e, Primer A). This primer is designed to anneal to the genes encoding the library of mutants which is to be selected. In this way, microbeads carrying genes encoding active enzymes, which catalysed the formation of the product molecules now co-attached to the same microbead, become loaded with antibody-oligonucleotide conjugates, whereas microbeads carrying inactive or less active genes are not bound by, or are bound by fewer antibody-oligonucleotide conjugates. These microbeads are then mixed with the ingredients of a PCR which can amplify the mutant gene library, but the primer carried by the conjugate is not included in the reaction mixture. The microbeads are then dispersed amongst the aqueous compartments of a thermostable water-in-oil
emulsion (Ghadessy et al., 2001) (Fig. 1f–g), the antibody-oligonucleotide linkage cleaved by addition of DTT through the oil, and the mixture thermocycled (Fig. 1h). Because the primer carried by the conjugate is limiting, only genes encoding active enzymes will be associated within a compartment with a large number of both primers, thus allowing exponential amplification, whereas genes encoding inactive enzymes will only experience linear amplification. Indeed, beyond ~1000 limiting primers per microbead, there is a roughly linear correspondence between the amount of limiting primer present and the degree of amplification of the gene, up to about 100 000 primers per microbead (the binding capacity of the microbeads) (Fig. 2). Below ~1000 primers per microbead, there is little amplification

**Fig. 1.** Selective gene amplification scheme. (a) A repertoire of genes encoding enzyme variants is linked to beads carrying multiple copies of the substrate for the reaction to be selected at, on average, less than one gene per bead. (b) The beads are compartmentalised in a water-in-oil emulsion to give, on average, less than one bead per compartment. (c) The genes are transcribed and translated in vitro in the compartments. In a compartment containing a gene encoding an enzyme with the desired catalytic activity, the substrate is converted into product, which remains attached to the bead. In other compartments, in which the genes do not encode an active enzyme, the intact substrate remains attached to the bead. (d) The emulsion is broken. (e) The beads are incubated with anti-product antibodies (or other product-binding molecules), chemically coupled to a forward oligonucleotide primer (Primer A) via a cleavable linker, which are recruited to the microbead-bound product molecules. (f) The pool of beads is mixed with a PCR reaction mixture lacking Primer A, supplied by the anti-product antibodies, but containing the backward oligonucleotide primer (Primer B), and dispersed amongst the compartments of a thermostable emulsion. (g) In an individual compartment, Primer B (supplied in excess in the PCR reaction mixture), Primer A molecules recruited via anti-product antibodies, the gene and all the other reagents necessary for PCR are present and isolated from other compartments. (h) The linkage between Primer A and the anti-product antibody is cleaved, and, upon thermocycling, amplification of the gene takes place. However, since the emulsion compartments are isolated from each other, amplification is limited by the amount of Primer A recruited to the microbead, which is dictated by the amount of product molecules produced by the action of the encoded protein. (i) The emulsion is broken to retrieve the DNA. (j) DNA obtained is retrieved and may be used in further cycles of selection.
and ΔOPD genes, the biotinylated substrate and product, and the anti-product antibodies are described in (Griffiths and Tawfik, 2003).] A subsequent PCR reaction, using primers annealing within the PCR fragment produced in the first reaction, was performed and analysed by gel electrophoresis (Fig. 3a), demonstrating that genes associated with product molecules are amplified preferentially, leading to an approximately 30-fold enrichment of those genes. See Supplementary data available at PEDS online for detailed experimental protocols.

To test the SGA approach further, we also performed model selections with the enzyme Biotin Protein Ligase (BPL), encoded by the BirA gene. The BPL enzyme catalyses the ATP-dependent biotinylation of a specific protein domain of acetyl-CoA carboxylase, termed BCCP. Microbeads were derivatised with substrate molecules (BCCP) and were further loaded with at most one gene each; the gene encoded either the BirA gene, or a gene encoding a biotinylation-inactive protein (glutathione S-transferase, GST). Microbeads loaded with the BirA gene were further incubated with purified BPL, ATP and biotin, allowing turnover of substrate into product (biotinylated BCCP), whereas microbeads loaded with the GST gene were not. The microbeads were then mixed together in various ratios, once again mimicking steps a–d of Fig. 1; again, the two genes differ in size but are amplified with the same primers, and may be distinguished by gel electrophoresis. After incubation with a streptavidin–oligonucleotide conjugate (composed of streptavidin linked via a disulphide bond to an oligonucleotide primer), which specifically binds to the reaction product (biotinylated BCCP), the microbeads were emulsified as described, subjected to thermocycling and recovered from the emulsions. After emulsion amplification, the expected yield of DNA if all genes encoded highly active enzymes would be in the order of several hundred femtomoles (due to limits on the number of ligand-oligonucleotide conjugates which may be recruited). Thus, to facilitate electrophoretic analysis of the amplification products, we performed a second, ‘nested’ PCR reaction, using primers annealing within the PCR fragment produced in the first reaction, which has the added advantage of mitigating against primer-dimer and single-primer artifacts. Nested PCR reactions were analysed by gel electrophoresis (Fig. 3b), demonstrating that genes associated with biotinylated BCCP molecules are amplified preferentially, leading to an approximately 150-fold enrichment of those genes in the case of a 1:100 starting ratio. The improvement in the enrichment compared to the OPD model selection is most likely due to the better specificity of the streptavidin–biotin system compared to the anti-phosphodiester antibodies used for the OPD selections, which also bind to the substrate phosphotriester with low, but detectable affinity (Griffiths and Tawfik, 2003). Streptavidin, in contrast exhibits a very high affinity for the product (biotinylated-BCCP) and almost complete absence of binding to the substrate (unbiotinylated-BCCP).

We went on to perform an experiment to test the complete SGA selection procedure as described in Fig. 1, in which BirA or GST genes were mixed in a 1:100 ratio (BirA:GST) and coated onto BCCP-coated microbeads at a ratio of 1 gene per 3 beads before emulsification. After an 8 h incubation, compartmentalised in an emulsion, to allow expression of the attached genes, and conversion of substrate to product...
by BPL, the microbeads were recovered, incubated with streptavidin–oligonucleotide conjugate, which specifically binds to the reaction product (biotinylated BCCP), washed, and emulsified in a thermostable emulsion. After thermocycling, the microbeads were recovered from the emulsions; a subsequent PCR reaction, using primers annealing within the PCR fragment produced in the first reaction, was performed and analysed by gel electrophoresis (Fig. 3c), demonstrating that the BirA gene is amplified preferentially and hence enriched over the GST gene by about 20-fold. The lower enrichment factor compared to the experiment above, in which the BirA and GST genes were bound separately to beads before the second emulsification is likely due to two factors. First, even at a ratio of one gene per three beads, 4.5% of beads will still carry more than one gene (assuming a Poisson distribution), secondly, and probably more importantly, even though the number of genes used in the experiment \((2 \times 10^8)\) in the first emulsion; \(10^7\) in the second) was well below the estimated number of compartments in both the first \((\sim 10^{10})\) and second \((\sim 10^8)\) emulsions, the fact that both the primary emulsions (Miller et al., 2006) and the secondary, thermostable emulsions (Ghadessy, Ong and Holliger, 2001), show significant polydispersity, means that there may be co-compartmentalisation of genes in both the first and second emulsions. However, it should be possible to eliminate this problem by using microfluidic systems to produce emulsions with \(\leq 1.5\%\) polydispersity (Anna et al., 2003), as discussed in Griffiths and Tawfik (2006).

**Conclusion**

SGA shares with other variants of IVC a series of advantages for selection and directed evolution of enzymes compared to *in vivo* selections in cells and selections based on binding or single-turnover, intramolecular reactions [discussed in Griffiths and Tawfik (2000)]. Furthermore, the very high throughput of SGA \((\sim 10^8\) genes per 600 µl emulsion) should make it a useful alternative to conventional high throughput screening technologies, where the maximum rate is not much more than 1 s\(^{-1}\) \((\sim 10^5\) day\(^{-1}\)). SGA should also prove useful as an alternative to FACS in the selection of enzymes and ribozymes by IVC, where the screening rate, although high \((20 000\) s\(^{-1}\)), still limits the size of library that is selectable to \(\sim 10^9\) genes, whereas SGA is quite trivially scaled up by the use of larger volumes of emulsion (Zaher and Unrau, 2007). Furthermore, the reagents required for the selective amplification step are relatively inexpensive, and no expensive instrumentation is required.

**References**


Received September 26, 2007; revised September 26, 2007; accepted September 28, 2007

Edited by Peter Hudson