

mouse. However, some technical and practical issues remain. Tissue distribution of the oligonucleotides will change as distinct organs form in a developing animal. For example, to date no antisense chemistries have been identified that cross the blood/brain barrier. Also, successful use of antisense is still an art. Finding effective antisense target sites within an mRNA requires screening a number of oligonucleotides in cell culture systems using specialized delivery vehicles. Finally, nonspe-

cific effects do occur and vary depending on the stabilizing chemistry used. Such nontargeted effects are often overlooked and can be quite misleading<sup>6,7</sup>. Because the antisense learning curve is still steep, a number of companies specializing in this technology are making their know-how and proprietary reagents available on a service basis. Such antisense reagents are now addressing the needs in the functional genomics field for tools that can interrogate gene function in high-throughput cell culture screens and

validate gene targets in animal models of human disease.

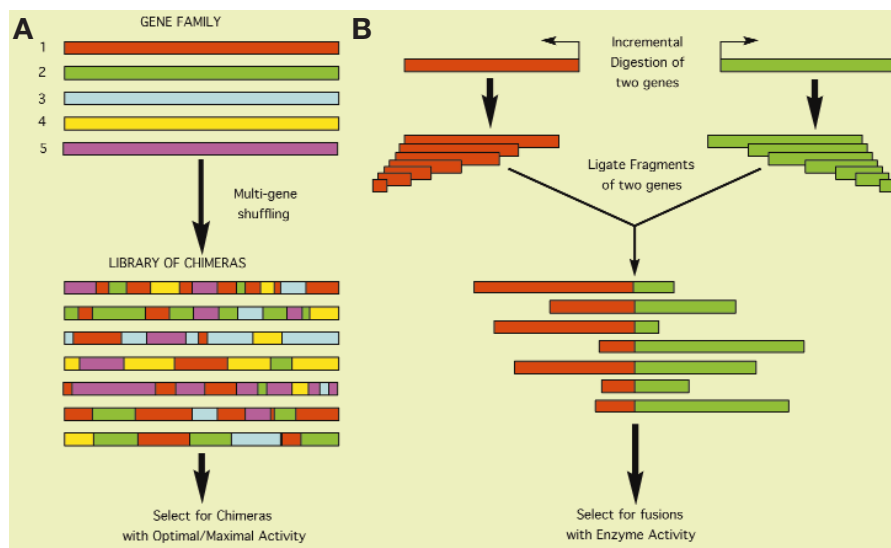
1. Driver, S. et al. *Nat. Biotechnol.* **17**,1184–1187 (1999).
2. Tabara, H. et al. *Cell* **99**, 123–132 (1999).
3. Xie, Y. et al. *Proc. Natl. Acad. Sci. USA* **94**, 13777–13781 (1997).
4. Zhao, J.J & Pick, L. *Nature* **356**, 448–451 (1993).
5. L'Huillier, P.J. et al. *Proc. Natl. Acad. Sci. USA* **93**, 6698–6703 (1996).
6. Stein, C.A. *Nat. Med.* **1**, 1119–1121 (1995).
7. Wagner, R.W. *Antisense Res. Dev.* **5**, 113–114 (1995).

## “Itching” for new strategies in protein engineering

Stephen W. Michnick and Frances H. Arnold

Recent *in vitro* recombination or “DNA shuffling” experiments have demonstrated that exchanging fragments of closely homologous genes provides an efficient way to generate proteins with new traits<sup>1,2</sup>. The resulting molecules are very different, at least in sequence, from those that might be obtained by more local searches of protein space, for example by random mutagenesis. The DNA shuffling method, which relies on homologous recombination during the PCR reassembly of gene fragments from multiple parents, generates crossovers at points of high sequence identity (Fig. 1A). Would crossovers between regions of low or no homology also be useful for molecular breeding of highly similar genes, or to shuffle more distantly related and even unrelated sequences? Would this approach generate a more efficient search of sequence space and, practically, more rapid adoption of novel protein properties? In this issue, Ostermeier et al. report a strategy to perform nonhomologous recombination that should eventually allow these questions to be answered empirically<sup>3</sup>.

The recombination strategy, called iterative truncation for the creation of hybrid enzymes (ITCHY), is based on generation of N- or C-terminal fragment libraries of two genes by progressive truncation of the coding sequences with exonuclease III followed by ligation of the products to make a single-crossover hybrid library (Fig. 1B). ITCHY was used to obtain fusions of glycinamide ribonucleotide transformylases encoded by *Escherichia coli* (PurN) and human (GART) genes sharing only 50% DNA sequence identity. N-terminal fragments of PurN were fused



**Figure 1. Schematic representation of *in vitro* strategies for recombination of gene sequences. (A) “DNA shuffling” strategy for recombining members of a gene family with high sequence homology. (B) The “Itchy” strategy for recombining two genes with lesser or no sequence homology.**

to C-terminal fragments of GART, and active clones were selected using a survival-selection in an *E. coli* auxotroph lacking either of two GAR transformylase activities. For comparison, Ostermeier et al. also performed a DNA shuffling experiment in which the same N-terminal (PurN) and C-terminal (GART) fragments were used as templates for reassembly of a complete shuffled hybrid.

Comparison of active clones produced by the two methods showed that ITCHY crossovers occurred over a wider range of residues than with shuffling, in which all the crossovers were found in the coding sequences of just four residues in a region of identical amino acid and virtually identical DNA sequence. With ITCHY, crossovers occurred throughout the coding region. And, since all sizes of the PurN and GART fragments were generated, fusions of diverse lengths were also possible. However, the functional proteins were fusion products in which the PurN N-terminal sequence was followed exactly by the

residue in GART predicted from the sequence alignment. Thus, it seems that there is value to maintaining the “proper” gene length and finding the small subset of fusions where the crossovers correspond to sequence alignments (which is an integral part of the DNA shuffling strategy, but not of ITCHY). Furthermore, the crossovers occurred in core regions of the protein, in particular near the active sites of the enzymes. One might have expected to see crossovers in unstructured loops, where, in principle, substitutions, deletions, or insertions should be less disruptive. Clearly, this was not the case, and Ostermeier et al. offer the explanation that the structural complementarity of the PurN and GART protein fragments are highest in these highly conserved regions.

None of the clones studied had activities as good as either parent PurN or GART, but this is no surprise. No selection for maximum activity was performed, and the library products only had to provide enough catalytic activity for cells to survive; evidently a small

Stephen W. Michnick is assistant professor with the département de biochimie, Université de Montréal (Stephen.Michnick@umontreal.ca). Frances H. Arnold is professor with the division of chemistry and chemical engineering, California Institute of Technology (frances@cheme.caltech.edu).

## ANALYSIS

fraction of the natural activity is sufficient. None of the fusions generated by DNA shuffling had activities as high as the best ITCHY fusions. Again, however, without a specific selection pressure, one cannot conclude that ITCHY performs any better than DNA shuffling for optimizing enzyme function.

The GAR transformylases (and related enzymes) are modular in both structure and function: residues that bind substrate are found on N-terminal subdomains, whereas those that contribute to catalysis are found in the C terminus. Since the fusions generated in these studies consist of N-terminal PurN and C-terminal GART fragments, one might naively predict that the hybrid enzyme would bind substrate like PurN and catalyze the formyl transfer reaction like GART. Remarkably, this is close to how the hybrids behaved. It might be desirable to dissect and reconstruct enzymes with specific kinetic characteristics or even different substrate specificities as a starting point in a test-tube evolution experiment, and the ITCHY strategy could be a useful starting point.

Protein family histories tell interesting tales of adaptation to radically new environments and sometimes even document the acquisition of new functions. Although sequences diverge more rapidly than function, structural features

can remain long after all evidence of sequence homology has disappeared. Nature achieves this remarkable divergence of function and properties by rearranging the coding sequences of genes using a host of combinatorial strategies, and running the results through the filter of natural selection. These strategies are diverse: mutation, recombination, horizontal gene transfer, and many other mechanisms contribute to the plasticity of a genome and an organism's ability to adapt. However, none of these strategies can access more than an infinitesimal fraction of the enormous space of possible protein sequences ( $20^{300}$  for a typical 300-amino acid protein), leading one to wonder whether some parts of this space are more interesting than others, perhaps richer in function or in the opportunity for innovation.

Nature certainly uses nonhomologous recombination to solve problems fast. Examples include creation of antibody repertoires, acquired antibiotic resistance<sup>4</sup> and the evolution of new strains of bacteria<sup>5</sup>. Thus it is attractive to add such a tool to the test-tube evolution repertoire. Unfortunately, it seems that our ability to generate new molecular diversity is rapidly outpacing our ability to find the interesting molecules therein. To avoid being buried in useless sequences, we sorely need strategies that will land us in the parts of

sequence space that are rich in the desired function(s). DNA shuffling does this by a relatively conservative mechanism: homologous recombination of sequences fine-tuned by evolution. We eagerly await future experiments that will tell us to what extent nonhomologous recombination also generates useful diversity. There is clearly a long way to go to demonstrate the utility of ITCHY or any other nonhomologous recombination strategy for evolving proteins with desirable traits. It would be interesting to see how proteins tolerate the swapping of smaller subdomains and whether the types of crossovers observed correspond to structural subelements such as individual or collective secondary structures. The direct extension of the current method to making multiple crossovers, however, is not obvious, and the utility of single crossovers will be limited. Without doubt, further work on methods will iron out the ITCHY wrinkles and give us a smoother picture of how to wander through sequence space.

1. Cramer, A., Raillard, S.A., Bermudez, E. & Stemmer, W.P. *Nature*. **391**, 288–291 (1998).
2. Ness, J.E. et al. *Nat. Biotechnol.* **17**, 893–896 (1999).
3. Ostermeier, M., Shim, J.H. & Benkovic, S.J. *Nat. Biotechnol.* **17**, 1205–1209 (1999).
4. Shapiro, J.A. *Trends Genet.* **13**, 98–104 (1997).
5. Lawrence, J.G. *Trends Microbiol.* **5**, 355–359 (1997).