14. Warmuth, M., Damoiseaux, R., Liu, Y., Fabbro, D., and Gray, N. (2003). Curr. Pharm. Des. *9*, 2043–2059.
15. Paez, J.G., Janne, P.A., Lee, J., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F.J., Lindeman, N., Boggon, T.J., et al. (2004). Science. Published online April 29, 2004. 10.1126/science.1099314

Chemistry & Biology, Vol. 11, May, 2004, DOI 10.1016/j.chembiol.2004.05.005

Recombination of Fragmented Proteins

Directed evolution is a powerful method for generating novel molecules with desirable properties. In developing a new sensor to screen for protein-protein interactions, Tafelmeyer et al. [1] report a clever strategy to evolve heterodimeric "split proteins" from a monomer in this issue of *Chemistry & Biology*.

Two individually inactive protein fragments can sometimes recombine to form active heterodimers. Perhaps the most notable example of such a split-protein system is ribonuclease S (RNase S), a tightly associated heterodimer that is generated from two proteolytic fragments produced by the limited digestion of ribonuclease A by the protease subtilisin [2]. RNase S is a longstanding workhorse in both protein chemistry [3] and biotechnology [4]. In addition, a rationally designed split in ubiquitin was used to develop a sensor for screening protein-protein interactions [5] and has also inspired other sensors based on complementation of fragment pairs derived from a few other proteins. Muir et al. recently developed a split-protein application that harnesses a naturally heterodimeric intein to facilitate the in vivo ligation of an endogenously produced protein with an exogenously produced molecule [6]. Despite their diverse uses, the number of available split proteins appears quite limited.

The challenge of finding suitable fragment pairings may have thus far prevented more widespread application of split-protein systems. By performing activity selections on combinatorial libraries of fragment pairs, Tafelmeyer et al. [1] have surmounted this obstacle. In this way, they have identified a series of fragmentation sites within the $(\beta/\alpha)_8$-barrel enzyme N-(5′-phosphoribosyl) anthranilate isomerase, Trp1p, from *Saccharomyces cerevisiae* that can be used to make a sensor for protein-protein interactions (Figure 1). In essence, their strategy for library construction adapts an established method for generating random circular permutations of a given protein sequence. In this method, a circularized gene is randomly linearized to generate new termini [7]. Tafelmeyer et al. modified this procedure by afterwards inserting a DNA sequence containing a successive terminator and promoter between the sequence encoding the original N and C termini of Trp1p. By this scheme, a circular permutation generates a single break in the primary sequence and is followed by the unmasking of the original termini, effectively splitting the expressed protein in two. This process generates a library of plasmids that encode all possible fragment pairs (224 in all for Trp1p).

To isolate sensors of protein-protein interactions, the N- and C-terminal fragments of Trp1p were each produced as C-terminal and N-terminal fusions, respectively, to independently interacting polypeptides. The plasmid library was transformed into *trp1* yeast (which require an added source of Trp1p activity for survival in the absence of tryptophan) to select for functional split-Trp1p pairs. In this selection system, the only cells that grow are those in which the fragment pairs can reconstitute Trp1p activity.

The sequences identified by functional selection must be further screened to identify false positives (Figure 1). In this case, positives were subjected to a second round of selection after removing the fusion partner from the N-terminal fragment of Trp1p. Those split Trp1ps that either associate too tightly or possess one fragment large enough to possess activity by itself will complement the mutation independent of the interaction between the fusion partners and can thus be eliminated. Out of a library of approximately 1600 clones, Tafelmeyer et al. identified four Trp1p fragment pairs that survived the first selection but not the second. These four split Trp1ps thus function as sensors of protein-protein interactions [1].

Trp1p is a smart choice for this proof-of-principle experiment for generating tailored split proteins. In addition to the availability of a selection system for yeast cell growth, there is a wealth of knowledge about the structure and engineering of this $(\beta/\alpha)_8$-barrel scaffold. The existence of enzymatically active variants that have been either circularly permuted (in the homolog of Trp1p from *Escherichia coli*) [8] or rationally fragmented (in *S. cerevisiae* Trp1p) [9] augured well for the success of this approach.

For developing a split Trp1p that can act as a sensor for protein-protein interactions, the trick is to isolate fragments capable of only conditional heterodimerization. Heterodimers with a sufficiently high K_d in vivo require the interaction of fusion partners to raise the effective concentration of one fragment relative to the other. The rationally designed split Trp1p engineered by Eder and Kirschner [9] forms a complex between a large, well-structured N-terminal fragment and a small, poorly structured C-terminal fragment with a K_d of $\sim$0.2 μM and near-native catalytic efficiency. This observation may explain why most of the false positives identified in the search for sensors of protein-protein interactions were heterodimers that possessed relatively large N-terminal fragments [1]. The affinity and activity of Eder and

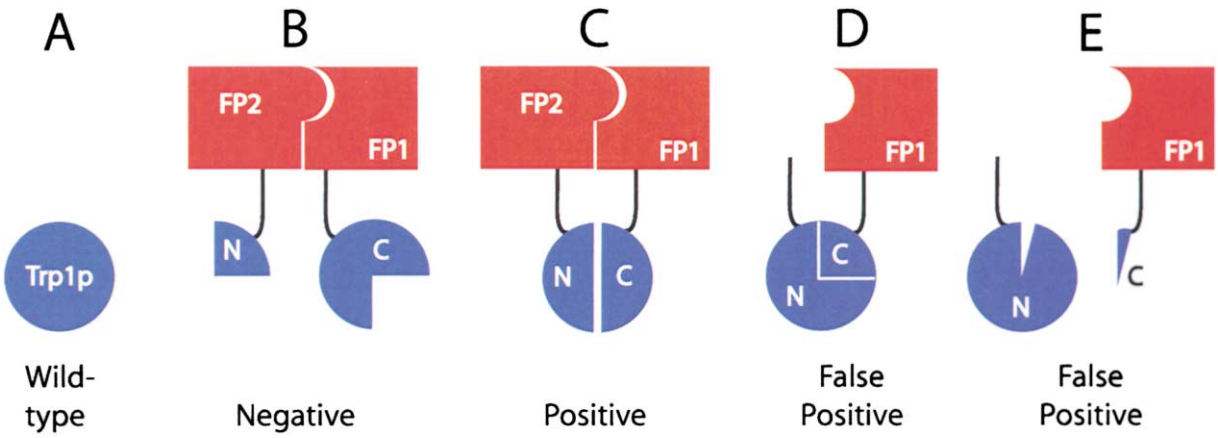

Figure 1. Possible Associations of the Split-Trp1p Fusion Proteins during Selections

(A) A cartoon version represents wild-type Trp1p as a circle.

(B) The N- and C-terminal fragments of Trp1p (N and C, respectively) have no detectable affinity for each other, despite the interaction of the fusion partners (FP1 and FP2). These fragments do not provide complementation in the selection system.

(C) N and C form a native-like complex when FP1 and FP2 are associated and thus provide complementation (i.e., growth of *trp1* yeast in the absence of tryptophan).

(D) Some N and C pairs may have a high enough affinity to reconstitute the Trp1p active site despite deletion of FP2.

(E) Alternatively, some N or C fragments may be so large that they can form a competent active site as monomers. The two complementation scenarios described by associations (D) and (E) will not produce sensors of protein-protein interactions. Split-Trp1p pairs that complement during a second round of selection, in which fragment N lacks attachment to FP2, can be discarded to eliminate false positives for this application.

Kirschner's complex are altogether too high for it to be useful as a sensor of protein-protein interactions in vivo. For Trp1p, the evolutionary strategy of random fragmentation coupled with activity selection yielded multiple, nonobvious solutions to the problem of finding Trp1p heterodimers with just the right mixture of affinity and activity.

As discussed, the split-Trp1p pairs were designed to detect protein-protein interactions in vivo, which is an important application. Protein-protein interactions lie at the heart of myriad biological processes. Detection of associated proteins in the cell is essential for elucidating biological mechanisms, for identifying therapeutic targets, and for evaluating potential drugs. The yeast two-hybrid system [10] and split-ubiquitin system [5] are both powerful tools that are used for pairwise screening of either protein-protein interactions or the targeted disruption of such associations. This split-Trp1p system represents a promising new option for performing such screens.

The split-Trp1p system does not depend on the identity of the associated fusion partners. Conditional reconstitution was detected using two different fusion partner systems: a soluble, dimeric α-helical coiled-coil and the specifically associating integral membrane proteins Sec62p and Sec63p [1]. It remains to be seen how useful this split-Trp1p system is in an actual screen for interacting protein partners. One possible advantage of the split-Trp1p system relative to split ubiquitin is that the reporter activity is intrinsic to the reconstituted protein and does not require further processing by endogenous accessory proteins (such as ubiquitin binding protein or associated proteases) to generate a signal. Sensitivity is always a concern in such screens. While the (relative) sensitivity of the split-Trp1p sensors is not yet known, the identification of multiple split-Trp1p sensors may provide an additional advantage in offering a range of sensitivities to meet the needs of a particular target interaction.

Tafelmeyer et al. [1] point out that the real advantage of this split-Trp1p system may lie in its generation. This method for developing novel split-protein systems is not limited to Trp1p. Any protein for which there is a selectable or screenable phenotype can be split using this strategy to generate novel sensors of protein-protein interactions. While Trp1p activity is limited to the cytosol, these new sensors can be targeted to currently inaccessible areas of the cell, such as the mitochondrial matrix or the secretory pathway.

Furthermore, by omitting the fusion partners, this method could be easily adapted to find tailored split proteins that form high-affinity, native-like complexes from inactive fragments. One possible application could be to test the evolutionary hypothesis that small exons are residual evidence of primordial protein products that possessed compact structural elements and assembled into active complexes [11]. This model is based in part on the observation of fragment complementation in chicken triose phosphate isomerase [the prototypical $(\beta/\alpha)_8$ barrel enzyme] that has been split precisely at exon/exon boundaries [12]. The random fragmentation method of Tafelmeyer et al. [1] might be used to examine in greater detail whether proteins split at exon/exon boundaries are more likely to reconstitute native-like heterodimers than those split elsewhere in the primary structure. Randomly split proteins might also be used to test the suggestion, based in part on the heterodimerization of rationally split proteins, that modern $(\beta/\alpha)_8$ barrels descend from $(\beta/\alpha)_{2N}$ precursors [13–15]. This technique may also find application in protein engineering.

For example, β-lactamase hybrids have been constructed from recombined protein fragments [16]. The identification of promising fragmentation sites could assist in the design of modules for use in DNA shuffling experiments to improve protein optimization by directed evolution.

The ready access to split proteins will likely facilitate their emergence as versatile tools in chemical biology. The strategy for their creation highlights methods development as another useful area in which to apply the methodology of directed evolution [17, 18].

Kenneth J. Woycechowsky
Laboratory of Organic Chemistry
Swiss Federal Institute of Technology
ETH Hönggerberg
CH-8093 Zurich
Switzerland

Selected Reading

1. Tafelmeyer, P., Johnsson, N., and Johnsson, K. (2004). Chem. Biol. *11*, this issue.
2. Richards, F.M. (1955). C R Trav. Lab. Carlsberg [Chim.] *29*, 322.
3. Richards, F.M. (1997). Annu. Rev. Biophys. Biolmol. Struct. *26*, 1–25.
4. Raines, R.T., McCormick, M., Van Oosbree, T.R., and Mierendorf, R.C. (2000). Methods Enzymol. *326*, 362–376.
5. Johnsson, N., and Varshavsky, A. (1994). Proc. Natl. Acad. Sci. USA *91*, 10340–10344.
6. Giriat, I., and Muir, T.W. (2003). J. Am. Chem. Soc. *125*, 7180–7181.
7. Graf, R., and Schachman, H.K. (1996). Proc. Natl. Acad. Sci. USA *93*, 11591–11596.
8. Luger, K., Hommel, U., Herold, M., Hofsteenge, J., and Kirschner, K. (1989). Science *243*, 206–210.
9. Eder, J., and Kirschner, K. (1992). Biochemistry *31*, 3617–3625.
10. Fields, S., and Song, O.-K. (1989). Nature *340*, 245–246.
11. de Souza, S.J., Long, M., Schoenbach, L., Roy, S.W., and Gilbert, W. (1997). Gene *205*, 141–144.
12. Bertolaet, B., and Knowles, J.R. (1995). Biochemistry *34*, 5736–5743.
13. Lang, D., Thoma, R., Henn-Sax, M., Sterner, R., and Wilmanns, M. (2000). Science *289*, 1546–1550.
14. Höcker, B., Beismann-Driemeyer, S., Hettwer, S., Lustig, A., and Sterner, R. (2001). Nat. Struct. Biol. *8*, 32–36.
15. Gerlt, J., and Raushel, F.M. (2003). Curr. Opin. Chem. Biol. *7*, 252–264.
16. Voigt, C.A., Martinez, C., Wang, Z.-G., Mayo, S.L., and Arnold, F.H. (2002). Nat. Struct. Biol. *9*, 553–558.
17. Taylor, S.V., Kast, P., and Hilvert, D. (2001). Angew. Chem. Int. Ed. Engl. *40*, 3310–3335.
18. Lin, H., and Cornish, V.W. (2002). Angew. Chem. Int. Ed. Engl. *41*, 4402–4425.