for a powerful selection system."

This is exactly what the scientists went on to develop. They took a shotgun approach to cloning carrier proteins that are part of the polyketide and nonribosomal polypeptide synthetic pathways in *B. subtilis*, as well as in the myxobacterium *Myxococcus xanthus*. They displayed shotgun libraries of the bacterial genome on the surface of M13 phage, subjected the expressed proteins to *in vitro* biotinylation and fished out putative carrier proteins by iterative selection (**Fig. 1**). After sequencing the genes encoding these proteins, they mapped them to the full-length clusters.

In *B. subtilis*, 85% of the genes recovered after five rounds of selection encoded carrier proteins of the relevant biosynthetic pathways; in all, 50% of the known carrier protein domains were cloned from a single genomic library. The *M. xanthus* genome has not been fully annotated, but here as well 22 carrier protein–encoding inserts were recovered, including as yet unannotated genes, in addition to six new sequences. As Sfp is known to have a fairly wide substrate specificity, and because it is only one of a family of enzymes that could be used for carrier protein selection, it is likely that this approach will be useful in several other bacterial species.

Undoubtedly, the application of this approach to genomes from unsequenced and ultimately from uncultured microorganisms will pose substantial new challenges. But Yin is optimistic. "I hope we can find a new cluster," he says. "The thing is, there is so little research on this. Even random sequencing of metagenomic samples has turned up some domains that may be involved in synthesizing new natural products, so I think we have a chance."

Natalie de Souza

RESEARCH PAPERS

Yin, J. *et al.* Genome-wide high-throughput mining of natural-product biosynthetic gene clusters by phage display. *Chem. Biol.* **14**, 303–312 (2007).

from a crude biological mixture was possible, but importantly, "Because venom changes depending on the season of the year that it's collected, and geographical reasons [and so forth], we found single nucleotide polymorphism variants in the sample as well," says Bandeira.

Though slow and laborious, the present gold standard for protein sequencing is Edman degradation. "Implicitly, we have nothing against Edman degradation, but we feel that with this technique, Edman degradation becomes unnecessary," says Pevzner. "The number of amino acids we find in a single experiment is in the thousands;…with Edman degradation no one is able to reach anything close."

Bandeira and Pevzner are confident that their concept of spectral networks will become an important new paradigm in MS-based proteomics, as they have welcomed quite a bit of interest from new collaborators. "While we have demonstrated these methods for mixtures of proteins, these are still somewhat small mixtures of proteins," says Bandeira. "It will be exciting to see how these tools scale to whole proteomes."

Allison Doerr

RESEARCH PAPERS

Bandeira, N. *et al.* Protein identification by spectral networks analysis. *Proc. Natl. Acad. Sci. USA* **104**, 6140–6145 (2007a).

Bandeira, N. *et al.* Shotgun protein sequencing: assembly of tandem mass spectra from mixtures of modified proteins. *Mol. Cell. Proteomics*; published online 19 April 2007b.

NEWS IN BRIEF

CHEMICAL BIOLOGY

Genetically encodable aldehyde tag

New bioorthogonal tags are always welcome additions to the chemical biologist's toolbox. Carrico *et al.* describe a new such handle, a genetically encodable aldehyde tag. A 6-amino-acid motif is recognized by formylglycine-generating enzyme, which oxidizes cysteine to the aldehyde-containing formylglycine. The aldehyde serves as a convenient attachment site for aminooxy- and hydrazine-functionalized labels.

Carrico, I.S. *et al. Nat. Chem. Biol.*; published online 22 April 2007.

PROTEOMICS

The Drosophila protein catalog

Brunner *et al.* present a catalog of 63% of the predicted proteome of the model organism *Drosophila melanogaster*, using shotgun mass spectrometry. They obtained high coverage by using diverse samples, an extensive fractionation strategy and a statistical bioinformatic approach called analysis-driven experimentation, which allowed them to optimize experimental conditions to target under-represented portions of the proteome.

Brunner, E. *et al. Nat. Biotechnol.* **25**, 576–583 (2007).

GENE REGULATION

A tool to upregulate gene expression

Whereas RNAi has come into its own as a means to knock down gene expression, no good corresponding tool has been available to do the opposite: upregulate gene expression. Xiao *et al.* describe the development of a cell-permeable synthetic transcription factor mimic, which activates gene expression in living cells by binding to a specially designed promoter. They were able to achieve a fivefold upregulation of expression in HeLa cells.

Xiao, X. *et al. Angew. Chem. Int. Edn.* **46**, 2865–2868 (2007).

SPECTROSCOPY

Aligning with nanotubes

Liquid crystalline media are used to weakly align protein molecules during an NMR experiment, facilitating the measurement of residual dipolar couplings, which aid in structure solution. These media, however, are generally incompatible with the detergents used in membrane protein preparations. Douglas *et al.* report the design and synthesis of detergent-resistant DNA nanotubes as a viable alternative for the weak alignment of membrane proteins in NMR experiments.

Douglas, S.M. *et al. Proc. Natl. Acad. Sci. USA*, **104**, 6644–6648 (2007).

GENOMICS

A gene expression resource for fission yeast

With the goal of understanding the multiple levels of regulation of gene expression, Lackner *et al.* present genome-wide data sets identifying key gene expression intermediates in the fission yeast *Schizosaccharomyces pombe*. Using microarray analysis, they collected data under standardized conditions, revealing new insights about systems-level regulation from transcription to translation.

Lackner, D.H. *et al. Mol. Cell* **26**, 145–155 (2007).