SENSORS AND PROBES

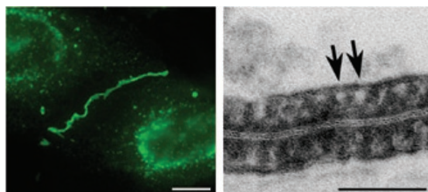
The right chromophore for the job

A normally undesirable property of chromophores is used in the creation of a genetically encoded tag for correlated light and electron microscopy.

Most jobs are best filled by people who are personable, considerate and polite, but some jobs require characteristics that would be undesirable elsewhere. GFP is an excellent example of the former: efficient, nonperturbative and easy to work with—properties that have made it the workhorse of fluorescence microscopy.

But chromophores can also have undesirable properties, such as generating damaging reactive-oxygen species upon exposure to light. Similar to abrasive people, this property can be useful in certain situations. Singlet oxygen generation can be used for light-induced inactivation of nearby proteins, photoablation of cells or—if the chromophore is also fluorescent—correlative light and electron microscopy.

In correlative microscopy, fluorescence imaging of a chromophore-labeled protein allows micrometer-scale imaging. Subsequent light-mediated generation of singlet oxygen by the chromophore is used to photoconvert diaminobenzidine into a osmiophilic polymer. Addition of osmium tetroxide to a fixed sample produces an electron-opaque stain that then allows nanometer-scale imaging by electron microscopy.



Connexin 43 gap junctions labeled with miniSOG. Shown are fluorescence microscopy (left) and electron microscopy (right) images. Scale bars, 10 μm (left) and 100 nm (right). Arrows mark putative negatively stained gap junction channels. Images adapted and reprinted from *PLoS Biology*.

Up until now, the available singlet oxygen–producing chromophore probes for correlative microscopy had drawbacks. GFP is too well behaved and does not produce enough singlet oxygen. Conventional small-molecule chromophores must be delivered by antibodies that require cell permeabilization with detergents, thus perturbing the cell ultrastructure. Recently, the biarsenical probe ReAsH, which binds to small tetracysteine domains inserted in proteins, was used for correlative microscopy. But ReAsH is intrinsically toxic and prone to high background.

Now Shu *et al.* describe a singlet-oxygen generator that exhibits none of these disadvantages. Xiaokun Shu, formerly a post-doc in Roger Tsien's laboratory at the University of California, San Diego and now an assistant professor at the University of California, San Francisco, says, "We wanted to create a fully genetically encoded singlet oxygen generator to produce [electron microscopy] contrast in multicellular organisms." They knew the chromophore generated by amino acids in GFP proteins is highly inefficient at singlet-oxygen production and decided to focus instead on the light-oxygen-voltage 2 (LOV2) flavin-binding domain from phototropin in plants.

Flavin efficiently produces singlet oxygen and exists in all cells. But normally the small 15-kDa LOV2 protein diverts light energy absorbed by flavin to create a covalent bond with a cysteine. To prevent this, Shu *et al.*

© 2011 Nature America, Inc. All rights reserved. npg

GENOMICS

From pseudogenes to proteins

Combining proteome analysis with genome sequencing improves gene annotation and yields evidence that some genes presumed to be noncoding are actually expressed.

Sequencing a genome is one thing; annotating all protein-coding genes is a bigger challenge yet. Efforts of automated and manual annotation by many research groups are directed toward finding locations of genes in the genome and determining which of these are actually translated into proteins.

Tim Hubbard, head of the Vertebrate Genome Analysis Project at the Wellcome Trust Sanger Institute, was looking for ways to improve current annotation efforts. Together with Jyoti Choudhary, head of protein mass spectrometry at the Institute, and Jennifer Harrow, who leads the Institute's manual annotation team, he devised a strategy that would incorporate shotgun proteomics data into the annotation effort of the mouse genome.

"We were looking to take mass spec data from any instrument," says Choudhary, "and have it go through one pipeline in a statistically grounded framework." To lay the groundwork, Markus Bosch, a joint PhD student between the groups of Hubbard and Choudhary, wrote an algorithm that dealt with the key problem in mass spectrum analysis, that of a high false positive rate in the matching of spectra to peptides. Bosch combined the database search engine Mascot with the machine-learning algorithm Percolator to weed out incorrect peptide spectrum matches. This combination provided high sensitivity and allowed a significance measure of every single matched peptide.

Choudhary stresses the importance of being conservative in peptide assignments. "What people have tried to do in the past," she notes, "is to maximize spectrum matching and say they identified a large number of peptides. But this is not representative of the quality of the call." Adding probability scores to a call allows one to set a higher threshold for bona fide matches.

The core of the team's pipeline is the GenoMS database, which contains peptides derived from *in silico* digests of well-annotated proteins in public databases as well as computational *ab initio* predictions of proteins. The team then analyzed entries from existing proteomic datasets, stored in the Peptide Atlas, and new spectra generated from mouse embryonic stem cells and brain cells by running them against the GenoMS database. They validated over 30% of known protein-coding genes with high probability scores.

For Harrow, the most exciting findings did not lie in the robust validation of existing annotations but in the discovery that a small

performed saturation mutagenesis of the cysteine and fused the mutant LOV2 proteins to a singlet oxygen–sensitive infrared fluorescent protein. Finding LOV2 mutants that produced singlet oxygen was as simple as picking bacterial colonies that displayed the greatest decrease in infrared fluorescence after blue-light illumination.

Success came in the form of two colonies with a 70% decrease in infrared fluorescence. Subsequent improvement of the endogenous green fluorescence of the mutated LOV2 protein resulted in a mini singlet-oxygen generator (miniSOG) protein that produced singlet oxygen with more than ten times the efficiency of ReAsH (>100 times better than GFP chromophore) while still producing substantial fluorescence.

Shu *et al.* fused miniSOG to a variety of intracellular proteins expressed in mammalian cell culture, worms and mouse brain. They observed no perturbation of protein localization and normal fluorescence in all cell compartments, suggesting that flavin was never limiting and the fluorescence was not quenched during cell fixation. As expected, miniSOG was very efficient at photoconverting diaminobenzidine for electron microscopy labeling. In fact, it worked almost too well. Shu says, "Because the singlet oxygen generation quantum yield was so high, we had to limit the exposure time to make sure it didn't produce too many [diaminobenzidine] polymers and degrade resolution." Because miniSOG out-competes endogenous peroxidase for diaminobenzidine polymerization, background is very low.

Shu is optimistic the technology can be improved using fixation that avoids dehydration at room temperature. Through collaborations with colleagues he is also actively extending the use of miniSOG to other applications that directly use generated singlet oxygen.

Daniel Evanko

RESEARCH PAPERS

Shu, X. *et al.* A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms. *PLoS Biol.* **9**, e1001041 (2011).

number of pseudogenes are actually translated into proteins. "We have always seen pseudogenes expressed," Harrow says, "but we have never seen translated ones." Of the ~10,000 pseudogenes present in the mouse genome, they saw evidence of translation for 19 of them. Validation of these expressed pseudogenes is ongoing, but Harrow sees it as reassuring that they only found a few, which indicates that the results are specific.

Of course, finding evidence of pseudogene translation does not say anything about the function of the protein; by common definition a pseudogene is a nonfunctional defective DNA segment that only resembles a gene. Choudhary plans to delve into analyzing the functional role of these newly discovered 'pseudoproteins' by looking at their localization and identifying their binding partners.

Having established the reliability of the pipeline, Hubbard plans to expand it to the human genome. "What we focused on up until now [in genome analysis] was entirely based on annotation by transcription," he says; "it is worthwhile trying to integrate the proteins."

Because some pseudogenes may not be 'pseudo' after all, it is possible that the same holds true for noncoding RNAs.

Nicole Rusk

RESEARCH PAPERS

Brosch, M. *et al.* Shotgun proteomics aids discovery of novel protein-coding genes, alternative splicing and 'resurrected' pseudogenes in the mouse genome. *Genome Res.* **5**, 756–767 (2011).

NEWS IN BRIEF

GENOMICS

Genome-wide map of the sixth base

Although 5-hyrdoxymethyl-cytosine (5hmC) is thought to be important for genome function in certain cells, a lack of tools has made profiling difficult. Xu *et al.* used a 5hmC-specific antibody for methylated DNA immunoprecipitation (MeDIP) and found distinct genome-wide 5hmC patterns in the mouse. The researchers saw 5hmC enrichment in gene bodies rather than at promoters and in intergenic regions. They also uncovered a role for the ten-eleven translocation (Tet) family of 5mC hydroxylases.

Xu, Y. *et al. Mol. Cell.* advance online publication (20 April 2011).

STEM CELLS

An eye in the Petri dish

Eiraku *et al.* report the formation of a primitive eye structure starting from a three-dimensional culture of mouse embryonic stem cells. The group grew embryonic stem cell aggregates on Matrigel and added factors that promote their differentiation toward retinal lineages. Under these conditions the cellular aggregates spontaneously formed hemispheral epithelial vesicles and progressively formed a structure reminiscent of the embryonic optic cup seen *in vivo*.

Eiraku, M. *et al. Nature* **472**, 51–56 (2011).

MOLECULAR ENGINEERING

Evolving RNA polymerase ribozymes

Wochner *et al.* describe the evolution of an RNA polymerase ribozyme capable of synthesizing up to 95-nucleotide RNA strands, much longer than the capabilities of natural ribozymes. They used water-in-oil emulsion technology with microbead display, linking ribozyme genes to multiple copies of the corresponding ribozyme; this allowed them to select ribozymes based on primer-extension capability.

Wochner, A. *et al. Science* **332**, 209–212 (2011).

IMAGING

High-throughput subcellular imaging of worms

High-resolution *in vivo* time-lapse assays in whole *Caenorhabditis elegans* worms require immobilizing the worms before imaging. Rohde *et al.* developed a system that allows doing this in a high-throughput manner using multiwell plates. They built an array of individually addressable cooling elements that fit into multiwell plates, and sequentially cooled down and imaged each worm using inverted epi-fluorescence microscopy. They used this technology to perform laser microsurgery of single neurons *in vivo*.

Rohde, C.B. *et al. Nat. Commun.* **2**, 271 (2011).

MOLECULAR BIOLOGY

A cellular system for epitope-tagged PrP

Goold *et al.* introduce an epitope tag into the cellular prion protein, PrP^C, expressed in a neuroblastoma cell line in which endogenous PrP has been knocked down by RNA interference; the tagged PrP^C supports prion replication and allows for the generation of infectious misfolded PrP^{Sc}. This approach allowed Goold *et al.* to distinguish newly formed PrP^{Sc} from inoculum-derived PrP^{Sc}; they observed very rapid infection of cells and found that prion conversion occurs primarily at the plasma membrane.

Goold, R. *et al. Nat. Commun.* **2**, 281 (2011).

© 2011 Nature America, Inc. All rights reserved.