

Up Front

Yeast Is Yeast

Cataloging the first atlas of an organism's proteins, scientists created a powerful tool for understanding these "workhorse" molecules.

Using high-tech robots and old-fashioned hard labor, HHMI researchers have measured the abundance and pinpointed the cellular locations of more than 4,000 proteins in yeast.

The proteins catalyze reactions, transport molecules within the cell, and switch genes on and off. Measuring the abundance of yeast proteins and identifying their cellular locations will be invaluable in helping to understand the complex biology of a relatively simple organism. Moreover, the effort exemplifies a shift in biological research toward understanding how changes in the "proteome"—the interacting global network of proteins in a cell—can influence cellular behavior.

"We have now made the yeast proteome accessible in a way it simply wasn't before," says HHMI investigator Jonathan S. Weissman at the University of California, San Francisco (UCSF), who collaborated on the effort with fellow HHMI investigator Erin K. O'Shea, also at UCSF. Researchers can now "measure the abundance of proteins and follow their location with a degree of sensitivity that was never possible for any proteome in any organism," Weissman observes.

Because the functional level of proteins in a cell can vary enormously, measuring the level is critical to understanding their properties, the researchers point out. Likewise, mapping the location of individual proteins in the cell is also essential to understanding protein function, says O'Shea.

Weissman adds that the new atlas helps to put yeast itself in the scientific spotlight. "We believe that this capability really strengthens the status of yeast as the premier organism for the systems-biology approach to a coherent, comprehensive understanding of how the cell works," he says.

O'Shea, Weissman, and colleagues published their results in two articles in the October 16, 2003, issue of *Nature*.

The researchers developed a method for

tagging each protein by first introducing the DNA for a specific tag into each of the genes that specify each yeast protein. They then synthesized about 13,000 gene sequences that would target one of the tags to the end of each gene in the yeast genome. Taking advantage of the available *Saccharomyces cerevisiae* genome sequence, the researchers developed a method for tagging the known 6,234 yeast gene segments thought to specify each yeast protein, called open reading

Jonathan Weissman (right) and Erin O'Shea believe that mapping proteins is critical to understanding their functional properties.





frames, or ORFs. They synthesized short DNA fragments that carried, at one end, a “tag” that allows the resulting protein to be detected and, at the other end, a sequence specific for each ORF, which allows the tag to be inserted into each yeast gene.

The researchers created two different libraries of yeast cells, each cell producing a different tagged protein. One library consisted of roughly 4,200 yeast strains, each producing a protein tagged with a sequence that makes it easy to purify using

an antibody to the tag. This method permitted the researchers to use antibodies to quantify the level of protein in the cell. The other library, used for protein-localization studies, consisted of strains in which each protein was tagged with a sequence that produced a green fluorescent protein visible under a microscope.

RARE BUT CRITICAL

“The experimental highlights for me are, first, that we can detect more than 80 percent of the yeast proteins in the cell—and that’s a high fraction of the genome to be expressing at one time,” says O’Shea. “And equally interesting is that we can see a vast range of protein abundance, from fewer than 50 molecules per cell to more than 1 million.” In contrast, previous methods of detecting yeast proteins identified only the more abundant proteins, missing rare but critical proteins, such as those that switch on genes.

“We just didn’t know how much of the proteome the cell needed during growth,” says Weissman. “We might have guessed that the cell had many genes in reserve for other purposes. So the 80 percent expression level we detected was a bit of a surprise. Also, although we had hints that proteins existed at a wide range of abundances, until these

measurements were made, there was no way to quantify that range.”

In addition to identifying many thousands of functional genes, the researchers also quantified the number of spurious ORFs, which are DNA segments that appear to be genes but are not. Their findings, they said, agreed with other studies of potentially spurious ORFs that do not yield detectable protein products.

O’Shea and Weissman next plan to use their protein libraries to explore how protein levels change over time. Obtaining this kind of dynamic information will be critical to efforts to model the action of proteins as the cell grows and adapts to changing conditions. Such models, they say, will give biologists the scientific equivalent of a movie of the cell’s machinery, as opposed to the snapshots available today.

“Although it’s interesting to know under standard lab conditions how much of each protein is present,” says O’Shea, “what you really want to know in order to understand more about biology and biological processes is how the amounts of the proteins change in response to different perturbations, like changes in the environment. Or if you make a mutation in the cell, how the abundance of the proteins changes.”

Little Green Protein *A single protein (green) glows in the mitochondrion—the site of energy production—of a dividing cell. Using unique DNA tags that were also phosphorescent, UCSF scientists were able to identify and pinpoint the locations of most of the proteins in yeast. They found that more than 500 proteins—about 1 in 8 in the organism—are active in the mitochondrion.*



WHERE PROTEINS LIVE

The localization studies were performed at two levels of specificity, notes Weissman. “First, we examined the cells using microscopy techniques, and for many proteins, that was enough. We could tell that many proteins—more than a thousand—existed in the nucleus or the cytoplasm. But for other proteins, we saw a punctate pattern that only told us that the protein was concentrated in a specific place.”

To pinpoint the locations of these proteins more specifically, the researchers introduced reference proteins tagged with red fluorescent molecules that were known to localize to one or another specific cell structure. Such structures might include, for example, the mitochondria—the cell’s power plants—or the Golgi apparatus, which is a network of internal cell membranes. When the researchers saw both red and green fluorescence at a given point, they knew that the protein concentrated in that structure.

According to O’Shea, the results of the localization studies were gratifying, and are already beginning to have an impact. “We were surprised that we could see as many proteins as we did and that the quality of the data was so good. Also, we were surprised

that more than 1,800 proteins have at least a part of their localization in places other than the cytoplasm or the nucleus. So, from this study, we’ve gained a lot of information about potential new functions of these proteins.”

O’Shea notes that her future studies of localization, like those quantifying protein levels, will concentrate on dynamic changes in the cell. “In this study, we’ve only provided a static view of localization under one condition,” she says. “But protein localization is dynamic, in many cases, and I think that the big challenge now is to use this library of strains to study how protein localization changes in response to environmental conditions.”

—DENNIS MEREDITH