

Overcoming protein diversity on arrays

Researchers show that to reliably establish protein interaction networks from microarray data it is necessary to measure saturation binding curves for each arrayed protein and its potential binding partners.

Scientists want to know with whom their proteins hang out. Without information on binding and reaction partners they cannot place a protein within a larger context and fully understand its role in the cell. Microarray technology is particularly well suited to query the interactions of large numbers of proteins at clearly defined conditions. But at the same time it prompts the question of whether all proteins behave similarly on an array or whether their different properties skew the results.

Gavin MacBeath and his team at Harvard University faced this question when they established a quantitative protein interaction network for the epidermal growth factor receptor family (Jones *et al.*, 2006). In the course of this study they found that data from arrayed proteins and binding probes at a single concentration were not reliable to correctly predict an interaction. Instead they obtained saturation binding curves for each protein-probe pair and fitted them to an equation that took the dissociation constant and maximum fluorescence at saturation into account. This allowed the researchers to reliably distinguish between specific and nonspecific interactions at various affinity thresholds, and to combine all data to construct a quantitative protein interaction network.

To underscore this finding that saturation binding curves are the only way to a reliable network, the MacBeath team performed a follow-up study, recently published in the *Journal of the American Chemical Society* (Gordus and MacBeath, 2006). MacBeath explains the rationale: “We wanted to investigate how serious the problem of protein diversity is; ‘Do these proteins really behave that differently?’ and if so, [we wanted to] show that the saturation binding curve method circumvents that problem.”

Their results highlight that even closely related proteins do indeed behave very differently on an array. The scientists used pro-

teins that showed similar binding activity in solution and arrayed them using the same volumes and concentrations. They found up to a 50-fold difference in binding activity between proteins. MacBeath cautions that these differences may be the source of mistakes in binary networks that are drawn based on whether or not an arrayed protein binds a probe at a single concentration. He says, “People have been drawing broad conclusions about how evolution works based on the topology of binary networks. In some cases, however, the topology may be an artifact of the binary assay. Differences in protein connectivity may simply arise from the fact that certain proteins behave well in the assay, which is why you see many interactions, and others don’t behave well, which is why you see fewer interactions.”

The MacBeath team then went on to confirm that measurements at multiple probe concentrations and the resulting binding curves yield quantitative information about protein interactions independent of differences on the array.

MacBeath acknowledges that these measurements are time consuming, but he is quick to point out that the ensuing quantitative networks not only contain more information than binary networks—they yield binding information at various affinity levels—but they are also more reliable and not plagued by high false positive and negative rates. Because of the amount of data they require, there is a limit to the number of proteins these quantitative interaction networks can be drawn for. In MacBeath’s experience, they are feasible for up to several hundred thousand interactions. He therefore suggests focusing on subsets of proteins for the time being, rather than genome-wide studies; he advises: “Decrease the numbers and increase the quality; that is more valuable to the community.”

Nicole Rusk

RESEARCH PAPERS

Gordus, A. & MacBeath, G. Circumventing the problems caused by protein diversity in microarrays: implications for protein interaction networks. *J. Am. Chem. Soc.* **128**, 13668–13669 (2006).

Jones, R.B. *et al.* A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature* **439**, 168–174 (2006).