ProtTech™

Identification of Phosphorylation Sites

Determining protein phosphorylation sites is often the first step in the elucidation of a regulation mechanism. Knowledge about a protein phosphorylation sites is the prerequisite for the functional analysis of phosphorylation by mutational analysis. Information about phosphorylation sites is also crucial for the purification of the upstream kinase that phosphorylates the defined sites. Currently, three complimentary approaches are used to determine phosphorylation sites: the bioinformatics approach, the genetics approach, and the biochemical approach.

The bioinformatics approach uses computer algorithms to predict potential phosphorylation sites based on the consensus recognition sequence of a known protein kinase, or based on the known phosphorylation sites in a homologous protein. The predicted sites are then verified by mutational analysis or by immunological methods with antibodies that specifically recognize the predicted phosphorylated site. Unfortunately, the consensus recognition sites of most protein kinases are not stringent enough to be very useful in predicting phosphorylation sties.It becomes impossible to make such prediction when a protein kinase is not known. Since protein kinases most likely recognize three dimensional structures instead of linear sequences in their substrates, a bioinformatics approach that is based on structural information will be more accurate and useful in predicting phosphorylation sties. We can anticipate that the bioinformatics approach will become a more powerful tool in the near future with the progress of structure genomics.

The genetics approach utilizes molecular biology to mutate potential phosphorylation sites and looks for in vivo phenotypes of these mutations. This approach is often used in combination with other approaches, such as bioinformatics or biochemical approach to reveal a biological function of such phosphorylation in a physiological condition. However, a mutant phenotype may be attributed to protein conformation change caused by such a mutation instead of change in its phosphorylation. It is often needed to demonstrate that the site is indeed phosphorylated in vivo by a biochemical method. The genetic approach is one of the most widely used methods to study functions of phosphorylation.

The biochemical approach directly identifies phosphorylation sites. This approach can provide the most conclusive evidence that a protein is phosphorylated at a certain site. A number of different procedures of the biochemical approach have been developed in determining phosphorylation sites. In principle, they all have at least two common steps: proteolytic cleavage of a phosphoprotein in order to narrow down each phosphorylation site to a phosphopeptide, and determination of the exact phosphorylation site in each phosphopeptide. Historically 2-D phosphopeptide mapping with 32P labeling in combination with site directed mutagenesis or Edman sequencing has been the method of choice in the identification of phosphorylation sites.

In the last decade, advancements in mass spectrometry have largely redefined the biochemical approach for the determination of phosphorylation sites. The mass spectrometry based approach offers at least four major advantages over the conventional biochemical approaches. First, it is a very accurate method. There is usually no ambiguity once a phosphorylation site is identified through mass spectrometric sequencing of the phosphopeptide.Second, it is a relatively fast procedure. The cycle of identifying phosphorylation sites is typically one or two days. Third, it does not require prior knowledge about the phosphorylation in the protein of interest. Phosphorylation sites can be identified even for proteins that their upstream kinases are not known. Fourth, it eliminates the use of radioactive materials such as 32P, which often posts a safety concern in most laboratories.

Although mass spectrometry is a relatively new technique in the study of protein phosphorylation, it has

already made significant contributions in the elucidation of regulatory mechanisms of a number of important proteins involved in cell cycle regulation, signal transduction, transcriptional regulation and other cellular processes. These proteins include Sic1, DOS, NPAT, NFAT, BRCA1 and NBS1.In many of these experiments, the ability to identify phosphorylation sites provided the important clues so that a hypothesis could be generated and experiments could be designed to test the functions of these proteins. In the post-genome era, complete dissection of signaling networks will require more effort in the identification of phosphorylation sites as phosphorylation is undoubtedly the most widely used mechanism for regulation. We envision that mass spectrometry based methods for the identification phosphorylation sites will become as important and widely used as those for protein identification.

The method used at ProtTech is based on two highly sensitive mass spectrometric techniques: MALDI-TOF and capillary LC-ESI-MS with an ion trap mass analyzer, coupled with an efficient enzymatic procedure. Such combinations attribute to several important features in this method. First, since each step in this method has compatible sample requirement, the overall procedure only needs subpicomole of sample. This amount can be purified by a micropurification procedure. Second, the protein does not need to be purified to homogeneity since it will be separated by SDS-PAGE before the MS analysis. It is able to map phosphorylation sites of the protein in a mixture in which the phosphoprotein of interest may be present as a minor fraction. Third, this method is applicable to Ser, Thr and Tyr phosphorylation. Fourth, it offers high-speed compared to other alternative methods. It allows the identification of an in vivo phosphorylation site using as little as 20 ng of SDS-PAGE isolated protein in a single day. We have used this method in mapping phosphorylation sites of various biologically important proteins.