2 DE as part of integrated proteomics: A full system's biology approach

Erich Gombocz, Robert Stanley

I/O Informatics, Inc., 2000 Powell Street, Suite 520, Emeryville, CA 94608, USA

Joint 20th Annual Meeting of the American Electrochemistry Society (AES) and American Institute of Chemical Engineers (AIChE), November 16-21, 2003, San Francisco, CA

Correspondence: egombocz@io-informatics.com

Short title: 2DE integrated proteomics: System's biology approach

Keywords: two-dimensional electrophoresis, proteomics, system integration, Sentient desktop, database query, LIMS, sample tracking, audit

Summary

Despite a variety of new approaches, proteomics still relies heavily on two-dimensional electrophoresis as a central tool. Integration of instruments, applications and data in a flexible, content-relevant and dynamic fashion and treat them as a single, homogenous "virtual database" is demonstrated.

Introduction

The complexity of proteomics' functional data dependencies requires unorthodox approaches as integration of all available information is the key to effective knowledge management. While two-dimensional electrophoresis still remains a focal point in multi-parametric separation, the need for uniform and instant access to secondary analytical steps such as selective, multiplexed visualization, mass spectrometry and sequencing from cut spots, the need to view 2DE as a part of integrated proteomics became apparent. The challenge to apply a full systems biology approach in order to interpret the complexity of protein pattern and their interactions in the context of their metabolism or to detect and define biomarkers for diseases is an unorthodox and new approach for knowledge management, information linking, sample tracking across instruments and laboratories and integration of interdisciplinary collaboration. The poster presents several "real life" examples applying a new commercially available "intelligent" desktop solution, which provides for integration of instruments across vendor barriers, existing applications and disparate data. In this "sentient" environment, the 2DE analysis module proofs not only magnitude improvements in performance, but acts through active linking to relevant data and querying of public, subscription-based and client data bases in the dynamically changing life science environment as a powerful systems biology approach to provide valuable, otherwise inaccessible scientific insights, sample tracking, experiment auditing and data management for any size laboratory. This new technology vectorized n-dimensional data access methods and provides enhanced LIMS sample tracking functions, and its benefits for 2DE proteomics as a "system's approach" are demonstrated and discussed.

Materials & Methods

All shown results were obtained using a Dell Inspiron 8500 laptop, 512 MB RAM (Dell Corp., Ft. Lauderdale, FL) and an internet connection. Operating system was Windows XP Professional (Microsoft, Redmond, WA). Used software was Sentient Desktop™ 2.1 (I/O Informatics Inc., Emeryville, CA) for 2D analysis. All screen captures were taken in wide (16:9) mode at 1680x1050 resolution, 32bit color, 96 dpi. Plug-ins 3 party tools shown in the examples were Weblab Viewer Lite 4.0 (Molecular Simulations, Inc.) and StarTime Viewer 1.1 (Insight Software Inc.). Other tools used were publicly available plugins on the web like Tag editor (available for SwissProt/TrEMBL via FlyBase, Swiss Institute of Bioinformatics). Experiment 2DE image data were obtained from various experimental sources including public databases (SwissProt 2D, Swiss Institute of Bioinformatics). Populated data relationship views are courtesy of Dr. Frank Witzmann, IUPUI Indianapolis, IN. Spot detection, deconvolution and multi-parametric normalization for gel experiments with different gel sizes and visualization techniques were performed using Sentient 2DE module 2.1. Content definitions and environment settings were used from the default life science ontology without modifications. Databases for real-time querying were used as pre-configured in the Sentient Desktop without modifications.

Results

A drag-&-drop query from a pI/molecular weight calibrated 2DE protein spot (sample: human, kidney) was used to generate a "universal" query across a subset of >900 biological databases and to obtain "active" annotation. When clicked on the gel spot, these annotations allow for automatic launch of the linked corresponding information in its native application as data viewer as well as for hotlinks directly to web pages. In distinct difference to other approaches, this allows to also incorporate e.g. search scripts, sub-queries, etc., within the web link Views for proper expansion when additional data becomes available. Results from a simultaneous query to 92 databases were obtained typically within 90 seconds and displayed in the user's browser of choice.

- Fig. 1: Drag-&-drop query (Note: Depending on internet connection and availability of external database, this may range between 48 sec. - 2 min. when automatically re-routing to mirrors)
- Fig. 2: Multiple active annotation links for each spot allow for direct launch of relevant data within their native application. Different annotation types are color-coded.
- Fig. 3: Literature information retrieved via single mouse-click (launched in MS Word)
- Fig. 4: Data from different experimental methods in their relationship using a content-based browsing view for clustering.
- Fig. 5: Experiment linking for 2DE as part of integrated laboratory activities: The tree is searchable and interactively responding to view relationships between sets of experiments (Courtesy: Dr. Frank Witzmann, IUPUI Indianapolis, IN)

Acknowledgments

The authors wish to thank Dr. Frank Witzmann, IUPUI Indianapolis, IN, for applying data relationship views on data sets from their core proteomics facility, and Dr. Kelvin Lee, Cornell University, Cornell, NY, Dr. David Garfin, Kensington, CA, and Dr. Robert Stevenson, Lafayette, CA, for many helpful discussions.

References