A Reagent Resource to Identify Proteins and Peptides of Interest for the Cancer Community

A WORKSHOP REPORT*

Brian B. Haab,[a,b] Amanda G. Paulovich,[b,c] N. Leigh Anderson,[d,e] Adam M. Clark,[e,f] Gregory J. Downing,[e,f,g] Henning Hermjakob,[e,h] Joshua LaBaer,[e,i] and Mathias Uhlen[e,j]

On the basis of discussions with representatives from all sectors of the cancer research community, the National Cancer Institute (NCI) recognizes the immense opportunities to apply proteomics technologies to further cancer research. Validated and well characterized affinity capture reagents (*e.g.* antibodies, aptamers, and affibodies) will play a key role in proteomics research platforms for the prevention, early detection, treatment, and monitoring of cancer. To discuss ways to develop new resources and optimize current opportunities in this area, the NCI convened the "Proteomic Technologies Reagents Resource Workshop" in Chicago, IL on December 12–13, 2005. The workshop brought together leading scientists in proteomics research to discuss model systems for evaluating and delivering resources for reagents to support MS and affinity capture platforms. Speakers discussed issues and identified action items related to an overall vision for and proposed models for a shared proteomics reagents resource, applications of affinity capture methods in cancer research, quality control and validation of affinity capture reagents, considerations for target selection, and construction of a reagents database. The meeting also featured presentations and discussion from leading private sector investigators on state-of-the-art technologies and capabilities to meet the user community's needs. This workshop was developed as a component of the NCI's Clinical Proteomics Technologies Initiative for Cancer, a coordinated initiative that includes the establishment of reagent resources for the scientific community. This workshop report explores various approaches to develop a framework that will most effectively fulfill the needs of the NCI and the cancer research community. ***Molecular & Cellular Proteomics 5:1996–2007, 2006.***

From the [a]Van Andel Research Institute, Grand Rapids, Michigan 49503, [c]Fred Hutchinson Cancer Research Center, Seattle, Washington 98109-1024, [d]Plasma Proteome Institute, Washington, D. C. 20009, [f]Office of Technology and Industrial Relations, NCI, National Institutes of Health, Bethesda, Maryland 20892, [h]European Molecular Biology Laboratory-European Bioinformatics Institute, Hinxton, Cambridge CB10 1SD, United Kingdom, [i]Harvard Institute of Proteomics, Harvard University, Cambridge, Massachusetts 02141-2023, and [j]Royal Institute of Technology, 106 91 Stockholm, Sweden

Received, April 12, 2006, and in revised form, July 21, 2006

Published, MCP Papers in Press, July 24, 2006, DOI 10.1074/mcp.T600020-MCP200

BACKGROUND

Cancer is a leading cause of death worldwide, and the evolving nature of tumors challenges investigators who wish to understand the myriad molecular processes that govern tumor formation. Tumors often metastasize before they can be detected, making them difficult to effectively diagnose early, treat, and control. One potential solution to this problem is to develop clinical protein-based systems that can detect and monitor cancer processes. To be clinically useful, however, these high throughput proteomics technologies must identify low abundance proteins linked to cancer processes, be sufficiently specific and sensitive to support diagnostic monitoring applications, and be reproducible and scalable for clinical use.

The National Cancer Institute (NCI)[1] convened the "Proteomic Technologies Reagents Resource Workshop" in Chicago, IL on December 12–13, 2005 to identify the cancer

[1] The abbreviations used are: NCI, National Cancer Institute; ATCC, American Type Culture Collection: a non-profit repository and worldwide distributor of various cell lines and primary cell types used for cell biology research; CPTI, Clinical Proteomics Technologies Initiative for Cancer: an initiative launched by the National Cancer Institute to accelerate advances in the prevention, diagnosis, and treatment of cancer through the use of proteomics technologies; HPA, Human Protein Atlas: an initiative funded by the Knut and Alice Wallenberg Foundation designed to allow the systematic exploration of the human proteome through affinity (antibody) proteomics, combining high throughput generation of affinity-purified (monospecific) antibodies with protein profiling using tissue arrays; HPR, Human Proteome Resource: located in Stockholm and Uppsala, Sweden, the HPR Center oversees the HPA and produces specific antibodies to human target proteins using a high throughput method involving the cloning and protein expression of protein epitope signature tags; IHC, immunohistochemistry; IP, intellectual property; PrEST, protein epitope signature tag: PrESTs are made *in vitro* from predicted coding regions of the human genome and are used to generate antibodies as part of an affinity-based proteomics strategy; SISCAPA, stable isotope standards and capture by anti-peptide antibodies: a methodology for quantitating peptides in complex digests using anti-peptide antibody chromatography and electrospray mass spectrometry; SOP, standard operating procedure: a set of instructions that serve as a guideline for those features of operations that lend themselves to a definite or standardized procedure without loss of effectiveness; XML, extensible markup language: a subset of standard generalized markup language, XML is a way to represent data that facilitates data sharing across different systems.

research community's expressed needs for validated and well characterized affinity capture reagents (*e.g.* antibodies, aptamers, and affibodies) to advance proteomics research platforms for the prevention, early detection, treatment, and monitoring of cancer. The workshop brought together leading scientists in proteomics research to discuss model systems for evaluating and delivering affinity reagents to the research community to support proteomics-based research. This workshop represented the latest effort in an ongoing dialogue between the NCI and the scientific community to enhance the applications of these technologies in discovery and translational research (1–4).

On the basis of discussions with representatives from all sectors of the cancer research community, the NCI recognizes the immense opportunities to apply proteomics technologies to mission-critical problems in cancer research. In particular, the Institute addressed the community's concerns for access to affordable, well characterized, highly validated affinity reagents. A community resource that supports such reagents would accelerate biomarker discovery, cancer diagnostics development, and therapeutics monitoring. Based on these community needs, the Workshop Steering Committee invited speakers and guests to participate in discussions on a variety of topics, including the following.

- A vision for a shared proteomics reagents resource
- Opportunities to advance proteomics technologies for cancer research
- Proposed models for an antibody reagent resource
- Applications of affinity capture methods in cancer research
- Affinity capture reagents and quality control/validation
- Target selection
- Database development
- Next steps toward building a shared resource

The meeting included presentations and discussions from leading industry representatives and academic investigators on state-of-the-art technologies and capabilities to meet the user community's needs. This workshop was developed as a component of the NCI's Clinical Proteomics Technologies Initiative for Cancer (CPTI; proteomics.cancer.gov), a coordinated initiative intended to evaluate, optimize, and advance proteomics technologies, informatics, and reagents to improve reliability and analytical validation (for a perspective on a proteomics initiative that incorporates a reagents resource, see Aebersold *et al.* (5)). Launched in 2006, the CPTI is a 5-year, 104 million-dollar initiative comprised of three core technology development programs: the Clinical Proteomic Technology Assessment for Cancer, the Advanced Proteomic Technologies and Computational Sciences programs, and the Proteomic Technologies Reagents and Resources Core. Together these three programs are designed to increase the reliability and reproducibility of proteomics research results, thereby enabling these technologies to be translated ultimately to the clinical setting. The Clinical Proteomic Technology Assessment for Cancer and Advanced Proteomic Technologies and Computational Sciences programs are intended to fund research teams to rigorously assess and optimize current proteomics platforms and to push the envelope for innovative technologies and data analysis schemes, respectively. By contrast, the Proteomic Technologies Reagents and Resources Core will serve the broader biomedical and life science communities by organizing tools and reagents and enabling technologies to support protein/peptide measurement technology development efforts. These highly purified, standardized, and characterized reagents will be used to support improved approaches to sample preparation, fractionation, separation, detection, and quantitation for proteomics research.

The CPTI is not itself a cancer biomarker discovery initiative but rather a technology-focused initiative designed to address current difficulties with the reproducibility of measurements, or the *analytical validation*, of proteomics platforms. In the absence of reliable measurements, it is difficult (if not impossible) to enable subsequent *clinical validation* of candidate biomarkers discovered using proteomics technologies. Such candidate biomarkers are only as reliable as the tools used to measure them. Reproducible proteomics research results will rely on high quality, well characterized, and easily accessible reagents, particularly affinity capture reagents such as antibodies. Optimized reagents will require detailed performance measurements across multiple platforms. This report will examine various aspects of affinity capture reagents and resources that will help fulfill the needs of the NCI and clinical proteomics researchers and enable the greatest impact on the cancer research community.

AFFINITY CAPTURE REAGENTS IN DISCOVERY AND TRANSLATIONAL RESEARCH

The mapping of the human genome and advances in proteomics technologies have spurred interest in the application of molecular diagnostics (*e.g.* DNA- and protein-based biomarkers) to detect, diagnose, and treat various cancers. Successful utilization of DNA-based diagnostics in the treatment of cancer has been demonstrated by targeted antitumor agents such as trastuzumab (6), imatinib mesylate (7), erlotinib (8), and gefitinib (8). However, measurement of DNA-based molecular markers requires invasive sampling of tumor tissue, thus limiting approaches for targeted diagnostic and therapeutic applications of these markers. Hence despite these success stories, applications of DNA-based biomarkers of cancer to the prevention, early detection, treatment, and monitoring of disease have been limited to date. In contrast, circulating protein-based biomarkers offer a minimally invasive option to aid in the early detection of disease.

Although molecular diagnostics represent critical elements for personalized disease screening and treatment, the successful translation of a diagnostic biomarker from discovery to routine clinical application remains relatively rare. The clin-

ical value of a diagnostic tool is grounded in its robustness and reliability; the ability to measure a disease marker accurately (*e.g.* with high sensitivity and specificity), reproducibly, and rapidly ultimately determines the commercial application of the marker. Large scale assays and bibliometric searches have identified hundreds of candidate biomarkers for various cancers, creating a plethora of preliminary genomics and proteomics data. Sifting through this vast amount of information to test priority candidates requires sophisticated, selective, and high throughput molecular methodologies that include well characterized and validated assays and reagents.

Profiling of differentially expressed proteins in normal and malignant tissues will support the development of a catalog of candidate biomarkers for various disease states (9, 10). Proteomics technologies enable the identification and measurement of these potential biomarkers of disease in serum, plasma (11), urine (12), tissue (13), and tumor interstitial fluid (14). Combinatorial proteomics applications contribute to the understanding of the functional organization of the human proteome through characterizations and measurements of protein abundance, post-translational modifications, protein-protein interactions (15), and correlation of disease phenotype with protein profiles (13). Applied clinical proteomics technologies offer a strong potential for early cancer detection and a strategy to evaluate tumor progression, response to treatment, metastasis (16), and recurrence (12).

However, reliably measuring the concentrations of these candidate cancer proteins at low levels (ng/ml to pg/ml) in plasma and other body fluids presents a bottleneck in the development of protein-based molecular diagnostics. Protein-based tests require precise, high throughput measurements enabled by highly characterized, validated affinity reagents. Yet biologically available proteins and potential biomarkers are being reported at a rate that outpaces the production and characterization of antibodies using conventional methodologies (17). Low throughput screening methods for monoclonal antibodies, arguably the optimal reagent of choice for proteomics analyses, have pushed the development of a number of automated, higher throughput approaches for antibody production (17–20). As these techniques become more refined, reagent production should increase, making the characterization and analytical validation of these reagents increasingly crucial for broad application in proteomics.

Well characterized and validated affinity capture reagents will be valuable and integral components in the development of advanced proteomics technology platforms. Innovative antibody-based methods to improve protein measurements, such as stable isotope standards and capture by anti-peptide antibodies (SISCAPA), are already being applied in mass spectrometry platforms to quantify peptides in complex mixtures (21). Additionally the combination of microarray technology and proteomics reagents has led to the development of "proteome chips," which offer the potential for the multiplexed analysis of hundreds to thousands of proteins in parallel (22, 23). These chips represent a promising application for large scale, cost-effective screening for numerous proteomics applications, including identifying protein-drug and protein-lipid interactions and post-translational modifications (24).

Although antibodies are currently the most mature reagents for large scale applications, variability in design and production parameters and a shortage of economical, renewable capture reagents such as monoclonal antibodies have hampered extensive technological development and assessment of this platform. The development of affinity capture assays capable of surpassing the current validation standard, ELISA, is both costly and labor-intensive. Equally daunting is the reality that commercially available antibodies are often poorly annotated and are not validated for specific applications, leaving researchers guessing as to whether a particular antibody is appropriate to their research needs and resulting in a needless waste of time, money, and effort. The widespread availability of well characterized affinity reagents will greatly accelerate biomarker discovery and validation by facilitating time- and resource-intensive development of immunoassays to measure specific biomarker candidates.

The field of proteomics and protein-based diagnostics has reached a stage for rapid advancement; tremendous opportunity exists to develop useful, meaningful resources to support future technology development and commercialization. A centralized (virtual) repository of affinity reagents with publicly available characterization/validation data will help to ensure reliable results and facilitate interlaboratory data comparison. Such a resource will greatly accelerate the development of proteomics technology platforms to identify protein biomarkers for the early detection of cancer and serve the greater research community as a hub to communicate and disseminate data and information.

KEY CONSIDERATIONS FOR A SHARED REAGENTS RESOURCE FOR CANCER RESEARCH

The workshop participants agreed that, to maximize impact, a shared reagents resource should 1) serve the broader scientific community; 2) coordinate efforts in characterization, validation, annotation, and database development; and 3) incorporate novel, inexpensive, scalable, high throughput technologies for future expansion to target the entire human proteome. These focal points advocate that the NCI and other funding institutions should use available resources to strike a balance between characterizing existing capture reagents, coordinating target selection for new reagents, and supporting advances in technology that will accelerate the production of new reagents.

The current challenges that cancer investigators face with respect to affinity capture reagents may be illustrated by the case of monoclonal antibodies, currently the most mature affinity capture methodology. Numerous antibodies and commercial pipelines for antibody production are currently in

TABLE I
Application-independent validation methods for antibodies (adapted from Uhlen et al. (26))
SPR, surface plasmon resonance; RNAi, RNA interference; GFP, green fluorescent protein.

Method	Description	Examples	Advantages	Disadvantages
Antigen-based	Assays based on the antigen used for immunization	ELISA, protein arrays, Biacore SPR, antigen adsorption	Can be combined with affinity-based validations	Need for pure and well characterized antigen; the binding to peptides or protein fragments might not be relevant for "real" applications
Target-based	Analysis of native or partially denatured protein from natural sources (such as cell lysates)	Western blot, immunohistochemistry, immunocapture	Does not require the antigen used for immunization	In the absence of the purified target, it is difficult to determine whether the antibody is binding to the target; usually relies on denatured targets
RNA-based	Comparison of expression levels at the protein and RNA levels	Transcript profiling, *in situ* hybridizations	A huge set of data already publicly available	Difficult to know whether RNA levels correlate with protein levels
Genetics-based	The use of genetic mutants or recombinant constructions to validate the target	Transgenetics, RNAi, GFP fusions (subcellular localization)	If protein levels are observed to be increasing with an antibody or decreasing with an RNAi, then one can be relatively certain that the antibody is binding to the target.	GFP fusions may be subject to artifacts
DNA-based	Bioinformatics analysis using predictive algorithms (as compared with experimental data)	Signal peptide, transmembrane regions, localization signals	No experimental evidence needed	Can only be used as supportive evidence; must be complemented with experimental data
Affinity-based	Determination of the kinetic parameters for the antibody	Biacore SPR, competition assays	Gives binding parameters	Usually done on antigens from "non-natural" sources
Epitope-based	Comparison of two or more antibodies directed to different parts of the same target	Antibodies to PrESTs or synthetic peptides	The ultimate validation because identical patterns in various assays give strong support for specificity and lack of cross-reactivity	Requires two independent antibodies to each target and also requires knowledge about the respective epitopes

place. However, the majority of antibodies are poorly characterized and not adequately validated for the variety of applications of interest to the research community (Table I). As such, the user must navigate through an increasingly complex marketplace to determine whether data are available on the binding characteristics of an antibody and whether the antibody is suitable for a specific application. For example, a query of Biocompare, an on-line search tool for scientific products and resources, for monoclonal antibodies to "p53" returns more than 1300 choices representing more than 50 vendors. These antibodies vary widely in application, design controls, validation parameters, supporting documentation, and cost. Conversely a search for emerging candidate biomarkers, such as "CA27.29," a recently identified candidate biomarker for breast cancer (25), yields no results. Additionally a query for "CA15.3" identifies several vendors with available monoclonal antibodies to this candidate breast cancer marker, whereas a syntactical change to "CA 15.3" returns no results.

This example identifies several needs of the cancer research community. First, a set of well characterized and validated capture reagents with readily accessible supporting data will conserve resources by enabling investigators to rapidly determine reagent suitability for specific applications. Second, identifying the gaps in reagent needs will support the development of well characterized antibodies to prospective and emerging targets for which there is no established commercial market. Third, appropriate annotations and ontologies will be needed for each of these resources to accommodate the variety of search terms used to query for the same resource and enable researchers to locate the reagents that best meet the needs of their research projects. In addition,

streamlined, high throughput methods of reagent preparation can improve the cost efficiency of production and characterization, thereby resulting in the development of advanced proteomics analysis platforms. A centralized shared reagents resource will therefore serve the research community most effectively by coordinating and maintaining an open access database of primary characterization and validation data for new and existing affinity reagents.

Among the key considerations in developing this central reagent resource is how to balance support for currently accepted technologies and procedures with that for developing platforms. Efforts must be carried out that aim to generate novel, high throughput technologies that will lower the cost of future affinity reagents and analysis platforms without sacrificing the quality of results. The resource should therefore support popular applications (*e.g.* Western blots, ELISA, immunofluorescence/fluorescence-activated cell sorting, immunoprecipitation, and immunohistochemistry (IHC)) as well as emerging technologies (*e.g.* SISCAPA and microarrays). The resource should also provide antibody validation data that address the characterization of reagents (*e.g.* sensitivity, accuracy of binding, and target specificity) for use in specific applications.

Target selection and reagent tracking require the construction and maintenance of a database of existing affinity reagents linked to their associated characterization/validation data. A viable shared reagents resource should therefore guide identification of targets for reagent development, track reagents through a "reagent pipeline," and provide publicly available screening and validation data of these reagents. A number of suitable paths exist to identify protein targets for the reagent pipeline, including literature searches, proteomics discovery efforts, microarray studies, pathway analysis, and systems biology. A database of accessible affinity reagents may be collected from catalogs of existing commercial antibodies or on-line resources. A shared reagents resource should also provide a user interface scheme for prioritizing candidate and new reagents. Possible prioritization schema for candidate reagents include bibliometric scoring (the number and impact factor of literature citations), requests from the user community supported with scientific data, and strategic input from the resource steering committee. In addition, target applications can be prioritized and placed in the queue within several months of start-up.

Creating a reagent pipeline requires the development of a high throughput mechanism for generating each antibody type (*e.g.* expressed whole target protein antigens, protein epitope signature tag (PrEST), protein domain antigens, or synthetic tryptic peptide antigens coupled with carriers). The initial phase of this process may begin by acquiring, screening, and validating existing commercial antibodies against high priority candidates. Additionally vendor pipelines for the streamlined production of antigens and complementary monoclonal antibodies should be established. This approach offers the opportunity for selected hybridoma clones to then be banked with the American Type Culture Collection (ATCC) under an open intellectual property (IP) policy so that investigators can obtain the clones and produce antibodies independently if desired.

The integrity of this pipeline rests on the open availability of accurate and reliable screening and validation results. In this and other NCI-sponsored workshops, the user community has reiterated its need for a minimal data set in the primary screening and validation of antibodies. To evaluate the specificity of antibody binding within a complex mixture, three criteria should preferably be assessed: 1) *the affinity of the antibody to the denatured protein or epitope (and actual epitope amino acid sequence when known)*, 2) *the affinity of the antibody to the target protein in vivo*, and 3) *the selectivity of the antibody to pull down the bound protein from a complex biological mixture.* These criteria could be met through a minimum set of curated data that includes Western blots, IHC, and immunoprecipitation data, respectively. An example of a virtual validation antibody resource is illustrated by the Swedish Human Proteome Resource's (HPR) Human Protein Atlas (HPA; www.proteinatlas.org), which requires validation through IHC applied to tissue arrays (26). Additional supporting validation for the reagent pipeline through antigen-based assay (*e.g.* ELISA), adsorption tests, flow sorting, and immunoprecipitation could also be included as supplementary validation components.

This model offers several options when considering infrastructure design for a shared resource. Based on these three criteria, primary screening and validation centers that focus on minimal requirement applications (*e.g.* Western blot, IHC, and immunoprecipitation) could be established to screen large numbers of antibodies. Secondary screening and validation centers may also be established to further characterize promising antibodies selected in primary screening for detailed affinity characterization (*e.g.* binding affinity kinetics, specificity, and cross-reactivity) or application-specific characterization (*e.g.* ELISA, microarray, and SISCAPA). These centers need not only be government facilities charged with conducting the assessment but could also be independent facilities with the proven capabilities to conduct and document the performance criteria. Although it is likely that different capture reagents will vary in their performance across different assays, it will be essential for protocols used at the testing sites to be transparent and include all associated data from the different assays. Regardless of how the testing reagents and resources are configured, raw experimental data should be stored in a publicly accessible database. As such, curation will become critical to maintain the quality and integrity of all user-submitted data. For example, to balance the data deposited through the user community, the HPA revalidates all submitted antibodies to determine suitability for acceptance and subsequent publication in its database. Whether the NCI should support a similar model by conduct-

ing validation internally or by certifying external validation centers is an issue that remains open for consideration.

This reagents resource will integrate with many existing NCI resources, including the Cancer Biomedical Informatics Grid™ (caBIG™; cabig.nci.nih.gov) to link the data associated with these resources and the Early Detection Research Network (edrn.nci.nih.gov) to provide high quality reagents for clinical validation studies. Additionally this resource will complement ongoing NCI programs that support the goals of the CPTI, such as the development of standardized guidelines for biospecimen collection, analysis, and storage currently underway at the Office of Biorepositories and Biospecimen Research (biospecimens.cancer.gov/).

MAXIMIZING VALUE TO THE RESEARCH COMMUNITY

Factors to be considered when designing a shared reagents resource include the selection of protein targets and reagent platforms, production and distribution parameters, and an informatics infrastructure to support the collection, storage, analysis, and distribution of data generated by resource users. Considerations for these elements are discussed in the sections below.

Optimizing Target Selection—Selection of proteins for further characterization and reagent development must follow a well defined and logical scheme, and these target proteins may be identified using several strategies based on the current state of knowledge in the field. One approach is to incorporate proteins currently under active investigation by the scientific community. For example, target proteins with supporting data that suggest their relevance to the biology or pathogenesis of cancer may be recommended for detailed characterization. A user market exists for antibodies produced to these candidate cancer proteins, thus providing a commercial incentive to companies that develop affinity reagents. Alternately the deliberate choice of novel candidates uncommonly studied by the scientific community may fill gaps in research resources for promising clinically relevant proteins that do not yet have affinity reagents available. Such targets may include low abundance serum proteins, extracellular proteins, or proteins that are predicted to be soluble, secreted, or localized on cellular surfaces. Because limited data would exist on these candidates, duplication of effort may be minimized by strategically aligning the development and characterization of antibodies directed at these targets. However, the absence of a well developed market may discourage companies from investing in the research and development necessary to produce these new antibodies without added incentives or support from the NCI. Regardless it is unlikely that one affinity capture reagent will be able to meet all the needs for protein measurements in the variety of assays available. Therefore, panels of reagents to the same target protein should be developed and selected for binding to different regions or epitopes within this protein. Doing so will also provide flexibility in research platforms and support innovation by providing a variety of options to the end user. The choice for target selection and the number of capture reagents designed should also be balanced by reagents that are already commercially available.

Data sources to support target candidate selection include the following.

- Literature surveys
- Databases such as MedGene
- Swedish HPR data (as it becomes available)
- Mining of microarray, comparative genomic hybridization, or proteomics data sets
- Published and validated data that identify novel candidate biomarkers or molecular disease targets

Anderson *et al.* (27) have used a combination of literature search with chromatographic and mass spectrometric methodologies on human plasma samples to generate a nonredundant list of 1175 distinct gene products experimentally identified in human plasma, suggesting the presence of a number of biologically available candidate marker proteins. Using a similar bibliometric approach, citation frequency analysis of candidate cancer biomarker proteins can be used to score protein targets. These data can be tracked and updated to yield a target selection method based on the publication history of a particular putative cancer biomarker. Such scoring serves as a surrogate measure of interest within the scientific community for characterization of the protein. A second method for selecting targets involves evaluating candidates through the MedGene database (hipseq.med.harvard.edu/MEDGENE), which allows users to enter a disease target and retrieve a rank-ordered list of genes associated with that disease based on Medline co-citation analysis. A third option for candidate selection could be mining publicly available data sets (*e.g.* microarray data) for potentially interesting targets. These approaches would provide the expression and possible localization of proteins in a variety of normal human tissues and cancer cells to suggest novel targets for disease. Candidates consistently discovered from all of these sources could be pooled to provide a list of at least 1000 protein targets.

In addition to literature and database surveys, candidate selection may also be driven by organized groups of investigators who submit research proposals to the NCI that contain a prioritized list of potential candidates. Such a model is currently underway in the NCI's Mouse Proteomic Technologies Initiative, a program that supports two independent consortia to evaluate the robustness and reproducibility of different proteomics technology platforms in multiple mouse models of human cancer (proteomics.cancer.gov/technology/mouse.asp). These consortia generated lists of target proteins identified through different MS-based approaches in multiple mouse models of cancer. These MS data will be tested and compared with data from other measurement platforms such as IHC, Western blot, ELISA, and microarray. Multiple monoclonal and polyclonal antibodies to different epitope regions

on a subset of candidate proteins are being generated for performance testing and validation of data across various platforms. It is anticipated that these data will help to outline parameters necessary to enhance current testing of proteins found in biological specimens.

To ensure impact to the research community, suitable guides, such as those detailed in the NCI's Mouse Proteomic Technology Initiative, could be placed on proposals for affinity capture reagents funded through the CPTI to limit them to organized groups of investigators who are studying a particular pathway implicated in cancer or to require demonstration of relevance to clinical research targets. This would ensure that all new antibodies generated are associated with an investigator who has a vested interest in studying the target protein, who has a need for high quality affinity reagents, and whose research will likely advance measurement technologies in cancer research. Alternately partnerships or agreements could be developed whereby users could submit requests and payments for antibodies on line whenever a particular research interest emerges, provided the antibody production companies ensure reasonable cost structures for production of reliable antibodies.

Selecting the Affinity Reagent Platform—At present, antibodies represent the most mature affinity capture technology, although the quality among developed antibodies varies dramatically. Tens of thousands of monoclonal antibodies are commercially available, and an established company can produce hundreds of monoclonal antibodies per year. However, many of these antibodies are poorly characterized and not readily qualified for multiuse applications or advanced, high throughput proteomics applications such as microarray- or bead-based technologies (28).

Although alternate platforms such as high affinity aptamers (29) and affibodies (30) have potential to substitute for antibodies in contemporary antibody microarray applications, these technologies are not sufficiently mature for large scale application. Few of these reagents have been produced to date, and even fewer are in routine use. The production of high affinity reagents with these platforms often requires multiple rounds of selection and amplification. Likewise alternative affinity capture platforms may integrate slowly into broader research applications due to restrictive intellectual property rights. The NCI could facilitate the incorporation of novel technologies into broader scale applications through open agreements with developers of these technologies. Until these alternative affinity reagents can routinely approach the level of binding specificity currently met with antibodies and have wide ranging availability and application, however, their incorporation into large scale, high throughput pipelines will be limited.

Therefore, it should be evaluated whether resources are best invested in the production of monoclonal antibodies, polyclonal antibodies, or both. Monoclonal antibodies, which have a defined single binding epitope, remain the standard for diagnostic and therapeutic applications. These antibodies offer a consistent and renewable resource as they do not require rescreening once validated. Although monoclonal antibodies are more costly to produce than polyclonal antibodies, recent innovations have decreased the time and labor involved in the hybridoma production process (17, 18). For example, antigens fused to specific proteins involved in the immune response of the immunized animal have reduced the immunization protocol from several months to several weeks (31). Moreover a recent report suggests that transgenic mice that harbor a mutant temperature-sensitive simian virus 40 large tumor antigen produce monoclonal antibodies (32), which may eliminate the need for hybridomas altogether. As these monoclonal antibody production technologies and engineering approaches continue to advance, the cost and time associated with these reagents may continue to decrease, thereby strengthening the case to develop a strategic plan based on monoclonal antibodies as the sole reagent of choice.

Nonetheless it is feasible that multiple monoclonal antibody clones (*e.g.* 20–100) may ultimately be required per protein because of differences in performance of these antibodies in different applications and differences in the epitope to which the antibody binds. High throughput methods will be needed in each application to permit screening of multiple antibodies per target even in the presence of prescreening processes that are capable of assessing isotype and eliminating unsuitable clones.

Polyclonal antibodies, although currently a non-renewable affinity capture resource, can be generated more rapidly and at lower cost than monoclonal antibodies, suggesting the potential application as an initial screening method to identify candidates for further reagent development (33). For example, polyclonal antibodies can be produced and applied initially to screen for differences in proteins between a diseased and a normal state at which point several monoclonal antibodies can be made against different epitopes of the identified target. Concerns over the specificity of the polyclonal antibodies can be addressed by creating monospecific antibodies through affinity purification (19). However, these resources will still be limited by the lack of a clearly defined epitope and by their finite amounts. As such, production of polyclonal antibodies as the sole reagent for the shared reagents resource may not prove to be the appropriate strategy for the NCI. Although a strategy that incorporates mono- and polyclonal antibodies may be used to triage the effort and cost associated with generating affinity reagents, such an approach would reduce the funding available to acquire additional renewable monoclonal antibody resources.

Advancing Assay Platform Development—The inherent limitations in measuring and quantifying large sets of proteins with current technology applications such as ELISA hinder wide scale utility for many clinical proteomics research applications. However, the field of affinity capture research is

rapidly evolving, and a wide variety of novel, multiplexed assay platforms for rapid screening and production of affinity capture reagents continue to be developed and optimized (34–37). Determining the most promising and robust platform(s) for proteomics applications may be facilitated by the comparative development and/or validation of a standard set of antibodies in a predetermined time frame by a number of funded laboratories. This approach supports the notion of a standard suite of renewable, well characterized antibodies to target proteins that could be provided as a resource for technology development applications.

Establishing Database Parameters—Incorporation of these resources and associated data into a publicly available system will require detailed annotation and aligned strategies across multiple databases. To maximize their usefulness to the research community, identified targets should be defined with respect to a well known, comprehensive, annotated, classified, publicly accessible resource such as the Universal Protein Resource consortium (UniProt; www.uniprot.org) (38). UniProt provides several non-redundant sequence databases with extensive query interfaces and an archive updated daily from numerous public databases. The consortium encourages submission of data from users, thus providing a forum for evolving discussion and scientific interchange. As such, it provides a useful model for the development of a shared reagents database resource.

To achieve maximum impact on the scientific community, annotation parameters for an affinity capture reagents database should include core data (*e.g.* associated gene, epitope sequence, accession number, version of the target, and species attribution of the protein) and other critical information, such as the target ranking, associated evidence, and antigen source and sequence. For antibodies, common attributes (*e.g.* name, source, price, quantity, mono-/polyclonal status, validation experiments, and system) should be stored. Clear definitions of all technologies used for validation may be enabled through descriptions with controlled vocabularies and a hierarchical structure. Antibody validations may require the capture and annotation of the following information.

- Relevant antigen (sequence) features (*e.g.* PrESTs)
- Technology
- Laboratory
- Score
- Supporting evidence (*e.g.* Western blot images and IHC)
- Quantitation data
- Technology-dependent parameters
- Repeated validations for one antibody/antigen pair
- Relationship between antibody, antigen, and technology
- Use of multiple antibodies in a validation assay (*e.g.* sandwich assays)

Access control must also be considered as certain users will require read-only access, whereas internal curators and others will need write access to capture anonymous opinion statements and target requests. Automated access to such complex data is currently best facilitated through an extensible markup language (XML) interface. Data elements are already incorporated in a common data exchange format, the XML schema, in programs developed through the Human Proteome Organization, Proteomics Standards Initiative-Molecular Interaction format (psidev.sourceforge.net/mi/rel25/) (39). The NCI should work with these partners to develop strategies for the facile exchange of data for its programs funded through this initiative. A coordinated effort between the NCI and its partners and collaborators could vastly improve data collection, storage, and analysis for the international proteomics community at large.

CONSIDERATIONS FOR IMPLEMENTING AN AFFINITY REAGENTS RESOURCE

The goals of the reagents resource will be to coordinate selection of targets for reagent generation, to provide a user-friendly virtual repository of antigens and affinity reagents, to ensure proper validation of reagents for specific applications, and to facilitate agreements between the reagent providers and the user community to allow cost-effective access to reagents and IP protection for development of multiplexed assays.

Making Antigens Available—Development of antigens represents a bottleneck in the development of new affinity reagents. It is therefore strongly recommended that a shared affinity capture resource develop characterized antigens in parallel to the capture resources. At a minimum, providers of capture reagents should be required to include the precise amino acid sequence that binds to the affinity reagent or is used as the immunogen. Access to high quality antigens, including peptides, proteins, and post-translationally modified proteins, is essential because they are required for antibody production and subsequent validation. A central repository that stores antigens and associated data produced by individual laboratories will benefit private sector contributors and the user community. Although such a facility could also perform protein expression, microarray fabrication, and distribution, the logistics of organizing and managing such a facility require additional input. One option for antigen production could involve a partnership between the user community and the NCI in which the Institute would pay for the commercial development of a monoclonal antibody after a user has produced and characterized a high quality antigen.

Supporting Antibody Production—Although companies may be willing to partner with academic laboratories to generate antibodies to "popular" proteins at low cost, they must be provided with additional incentives to produce and validate antibodies to infrequently studied targets. Validation procedures may be costly, and the level of characterization required will directly influence the user cost. The NCI could possibly mitigate the financial risks assumed by antibody producers by sharing the cost of production or of validation or by providing companies with a list of recommended proteins based on

their predicted potential to serve as useful biomarkers. Should the NCI accept the associated costs, manufacturers would be required to provide the antibodies to the scientific community at reasonable costs.

Enabling Antibody Validation—Validation of new and existing antibodies may proceed through two operational models: 1) a "user validation" model that incorporates initial validation of commercially available antibodies by end users with characterization data deposited into a centralized database and 2) validation at a central facility operated under a regulated set of standard operating procedures (SOPs). The former option benefits researchers who have a vested interest in developing the appropriate conditions for optimal antibody performance. As use of the antibody continues, more application-specific data can be added to the web site by other users. This virtual model provides a web-based rating and review system that offers an incentivized, cost-effective, and timely means of providing information on existing antibodies.

Through this user validation model, NCI-funded investigators could submit the characterization and validation data developed through NCI-approved SOPs as a component of their research projects. This approach would allow the antibodies to be evaluated through NCI-approved protocols while also providing investigators with the freedom to modify or optimize procedures if the SOPs fail to generate positive results. However, it should be noted that this formula will be more appropriate for monoclonal antibodies given the renewable nature of the resource and the amount of reagent that may be necessary to optimize the protocols.

Possible disadvantages of this approach include the difficulty in ensuring consistency across individual user laboratories and the lack of an enforcement mechanism to ensure that data from the user community are deposited in the database. These potential drawbacks could be countered by a centralized facility that performs select functions, such as resolving discrepancies in user validations and reviews, monitoring quality assurance/quality control as a function of lot number or storage time, and providing official independent certification through routine standardized assays. Alternately a partnership between vendors and the academic research community may be created to develop an appropriate standardized validation procedure and SOPs.

The second model uses a centralized facility or a group of certified laboratories to perform standardized application-independent validations for new and existing antibodies. This model would contract characterization work to these laboratories, which would be responsible for characterizing the antibodies and depositing the data in a timely fashion. This model would serve as an unbiased assessment of the antibodies but in doing so may not necessarily retain a strong impetus to optimize protocols. A combination of the two models also represents a viable option, depending on the application. The ultimate choice of an operational validation model or combination of models will be driven by cost. The suggested choice of applications that should be validated routinely include Western blot, ELISA, IHC, and immunoprecipitation as these tests are incorporated into standard validation procedures for all reagents sold by many commercial vendors. However, applications that foster high throughput uses also should be validated.

Supporting Distribution of Antibodies—Distribution of affinity reagents may follow two mechanisms. The first involves a centralized repository and distribution center. With this model, a standard set of quality assurance/quality control validation parameters could be used for all antibodies produced. The logistics of developing high throughput applications also would be simplified by a centralized distribution center containing all of the required reagents. Various IP issues requiring the use of multiple agreements also may be avoided or streamlined under this model. However, the present lack of infrastructure for "privatized" centralized distribution represents a possible disadvantage of this approach.

The second mechanism uses an existing commercial distribution center that could be "virtually centralized" through the NCI. This virtual repository would link the queried antibodies and reagents to the providers. Such a distribution system could align the NCI antibody characterization data to the sources of the antibodies so that researchers could view the data and determine whether the antibody is appropriate for their applications.

Developing Multiplex Affinity Capture Platforms—High throughput analysis and development within the field of genomics was advanced by the adaptation of common protocols, reporting mechanisms, and databases. Currently several barriers impede such a breakthrough in the field of proteomics, including the lack of comprehensive and comparable sets of validated affinity reagents, the prohibitive price of purchasing large numbers of commercial antibodies, and difficulties with incorporating innovative proteomics platforms imposed by IP rights.

The current catalog prices of antibodies makes it economically unfeasible to fabricate high density antibody microarrays for research purposes; it is currently not beneficial for companies to package and distribute the extremely small quantities of antibodies needed for microarray applications. New partnership strategies between industry and academic laboratories will be necessary to provide these small aliquots. A virtual centralized repository could also help to mediate or reduce the cost burden to individual researchers while at the same time providing an opportunity for companies to broaden their distribution of products. NCI-funded antibodies could be stipulated to be sold in microvolume amounts when purchased as large sets of antibodies. The NCI could play a role in mediating these public-private partnerships or provide a standard suite of well characterized antibodies to help develop pilot projects for innovative array formats. Pinpointing mechanisms for lowering the costs of acquiring diverse sets of commercially available affinity reagents (*e.g.* a cooperative,

bulk purchasing agreement) such that large scale multiplexing of assays can begin to help to overcome the present scalability problem that limits strategic profiling of the proteome would also be beneficial.

In parallel, the NCI should work to minimize barriers stemming from IP rights and manage issues associated with individual reagents and platforms efficiently such that they do not limit the scale or scope of multiplexing. This could be accomplished by supporting public-private partnerships in technology development and implementations. Such partnerships could be used to incorporate and characterize innovative technology concepts as critical components of NCI-funded research programs.

Successful design and application of multiplexed protein and affinity capture analysis platforms will require extremely high quality reagents capable of meeting the following minimal standards.

- High and well defined specificity
- High affinity
- The ability to be produced as a renewable resource
- The ability to be engineered readily for diverse platforms and applications
- Minimal to no IP barriers that would inhibit multiplexing
- An economical price, especially when purchased in large sets

It is understood that there are legitimate concerns with the viability of large scale arrays. Differences between low and high abundance proteins, the variable physicochemical properties of different antibodies, and the difficulty in achieving well defined specificity (*i.e.* each antibody binds only one protein) represent major problems to overcome. Likewise it may not be reasonable to expect IP barriers to disappear as a result of an NCI-funded initiative given the tremendous investment that has been made in protecting IP associated with affinity reagents and platforms. To overcome this obstacle, parties with IP rights should work together to ensure that royalties do not scale linearly with the number of proteins being measured on a single microarray. The use of NCI-organized consortia, clearinghouses, and cross-licensing opportunities could help overcome this potential royalty stacking problem.

Advanced reagent technologies may also need to be developed to meet the needs for effective affinity capture multiplex platforms. These reagents should be flexible (*e.g.* able to work with a variety of different chemistries), scalable, inexpensive, and portable. Although proteins and nucleotides can selectively capture targets for proteomics, they are currently of limited application, and a period of time and effort will be required before truly monospecific small molecules with a high affinity for given proteins are broadly applied. However, these reagent technologies can still be applied to address many challenges in clinical proteomics. For example, these reagents could be used to select for more than individual protein binding. Small molecules with undefined, broader specificity could be very useful for sample prefractionation or as secondary affinity reagents. Likewise combinations of individual small molecules with relatively high specificities and affinities could be linked to yield additive binding energies and the product of their affinities. The current challenge is the development of a truly high throughput method for this linked small molecule approach.

PUTTING IT ALL TOGETHER: NEXT STEPS AND FUTURE DIRECTIONS

As investigators begin to compile data on the human proteome, it is becoming clear that proteins will reveal many promising biomarkers for cancer and other diseases. However, without established procedures to characterize and validate antibodies and other affinity capture reagents, data generated will not be globally useful. A virtual, centralized reagents resource will aid in biomarker discovery efforts that may ultimately lead to the early detection, diagnosis, and treatment of cancer. Such a resource will benefit the larger research community by providing an evolving, trusted source for reagents and supporting characterization data.

The NCI workshop highlighted many of the challenges and opportunities for developing such a community-wide resource. The endeavor will require the cooperative efforts of academic laboratories, industry, and federal representatives. Many companies have expressed an interest in receiving help in choosing optimal targets, screening existing hybridomas, validating commercially available antibodies, and participating in a centralized database. Whether the NCI will fund the development of a physical center to house and characterize affinity reagents or whether it is better for the NCI to work with reagent producers and vendors to virtually link affinity capture reagents, details, and associated data from a common use portal seamlessly connecting the user to the source remains a point of discussion. A central facility will require the development of an appropriate infrastructure, which may be cost-prohibitive given that vendors currently have facilities designed to meet these needs. Regardless the NCI could provide great value to the community by developing a web query-capable interface to provide access to the inventory of reagents available and their associated performance data. Programs such as the Mouse Proteomic Technology Initiative could serve as a next step pilot project to examine the variables involved in storage and distribution of a central facility for NCI-funded capture reagents.

The NCI has a unique opportunity to act as a liaison to spur the development of a public resource that serves the research community and commercial vendors and to develop and impose a set of antibody validation standards. Well characterized affinity capture reagents will be of value in individual assays and can enable array-based technology development by providing a renewable source of materials to compare different antibodies targeting a single protein on a single array or to measure panels of different cancer-related proteins. This resource will reduce the financial risks to antibody producers

and lower prices for end users while serving as a model for future platforms based on affinity reagents and emerging technologies.

To this end, the NCI supports a plan to advance specific action items. The Institute is committed to developing guidelines or recommendations for antibody characterization and validation standards that will support a web-based, accessible, centralized database to store data and user comments obtained under a set of standardized parameters. The development of these new resources and *in vitro* technology platforms will have an impact beyond that of the cancer community by fostering an interactive, global hub for biomarker research and discovery.

Acknowledgments—We thank Samir Hanash and Anna Barker for input and advice in planning the Proteomic Technologies Reagents Resource Workshop and for work in developing the National Cancer Institute's Clinical Proteomic Technology Initiative for Cancer.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[b] Co-chair of the National Cancer Institute Proteomic Technologies Reagents Resource Workshop Steering Committee.

[e] Member of the National Cancer Institute Proteomic Technologies Reagents Resource Workshop Steering Committee.

[g] To whom correspondence should be addressed: Office of Technology and Industrial Relations, Office of the Director, National Cancer Institute, Bldg. 31A, Rm. 11A-30, 31 Center Dr., Bethesda, MD 20892. Tel.: 301-496-1550; Fax: 301-496-7807; E-mail: downingg@mail.nih.gov.

REFERENCES

1. *Proteomic Technologies for Early Cancer Detection*, *Chantilly, VA, April 22–23, 2003*, National Cancer Institute, Bethesda, MD
2. *Clinical Proteomics and Biomarker Discovery in Cancer Research, Bethesda, MD, September 24, 2004*, National Cancer Institute, Bethesda, MD
3. *Clinical Proteomics and Biomarker Discovery in Cancer Research (2), Menlo Park, CA, November 5, 2004*, National Cancer Institute, Bethesda, MD
4. *Proteomic Technologies Informatics Workshop, Seattle, WA, February 8–9, 2005*, National Cancer Institute, Bethesda, MD
5. Aebersold, R., Anderson, L., Caprioli, R., Druker, B., Hartwell, L., and Smith, R. (2005) Perspective: a program to improve protein biomarker discovery for cancer. *J. Proteome. Res.* **4,** 1104–1109
6. Bartlett, J. M. (2005) Pharmacodiagnostic testing in breast cancer: focus on HER2 and trastuzumab therapy. *Am. J. Pharmacogenomics* **5,** 303–315
7. Deninger, M. W., and Druker, B. J. (2003) Specific targeted therapy of chronic myelogenous leukemia with imatinib. *Pharmacol. Rev.* **55,** 401–423
8. Comis, R. L. (2005) The current situation: erlotinib (Tarceva) and gefitinib (Iressa) in non-small cell lung cancer. *Oncologist* **10,** 467–470
9. Hudelist, G., Singer, C. F., Kubista, E., and Czerwenka, K. (2005) Use of high-throughput arrays for profiling differentially expressed proteins in normal and malignant tissues. *Anti-Cancer Drugs* **16,** 683–689
10. Patwardhan, A. J., Strittmatter, E. F., Camp, D. G., II, Smith, R. D., and Pallavicini, M. G. (2005) Comparison of normal and breast cancer cell lines using proteome, genome, and interactome data. *J. Proteome Res.* **4,** 1952–1960
11. Berhane, B. T., Zong, C., Liem, D. A., Huang, A., Le, S., Edmondson, R. D., Jones, R. C., Qiao, X., Whitelegge, J. P., Ping, P., and Vondriska, T. M. (2005) Cardiovascular-related proteins identified in human plasma by the HUPO Plasma Proteome Project pilot phase. *Proteomics* **5,** 3520–3530
12. Grossman, H. B., Soloway, M., Messing, E. , Katz, G., Stein, B., Kassabian, V., and Shen, Y. (2006) Surveillance for recurrent bladder cancer using a point-of-care proteomic assay. *J. Am. Med. Assoc.* **295,** 299–305
13. Zhuang, Z., Huang, S., Kowalak, J. A., Shi, Y., Lei, J., Furuta, M., Lee, Y. S., Lubensky, I. A., Rodgers, G. P., Cornelius, A. S., Weil, R. J., Teh, B T., and Vortmeyer, A. O. (2006) From tissue phenotype to proteotype: sensitive protein identification in microdissected tumor tissue. *Int. J. Oncol.* **28,** 103–110
14. Celis, J. E., Gromov, P., Cabezon, T., Moreira, J. M., Ambartsumian, N., Sandelin, K., Rank, F., and Gromova, I. (2004) Characterization of the interstitial fluid perfusing the breast tumor microenvironment: a novel resource for biomarker and therapeutic target discovery. *Mol. Cell. Proteomics* **3,** 327–344
15. Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F. H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., Timm, J., Mintzlaff, S., Abraham, C., Bock, N., Kietzmann, S., Goedde, A., Toksoz, E., Droege, A., Krobitsch, S., Korn, B., Birchmeier, W., Lehrach, H., and Wanker, E. E. (2005) A human protein-protein interaction network: a resource for annotating the proteome. *Cell* **122,** 957–968
16. Everley, P. A., and Zetter, B. R. (2005) Proteomics in tumor progression and metastasis. *Ann. N. Y. Acad. Sci.* **1059,** 1–10
17. Chambers, R. S. (2005) High-throughput antibody production. *Curr. Opin. Chem. Biol.* **9,** 46–50
18. De Masi, F., Chiarella, P., Wilhelm, H., Massimi, M., Bullard, B., Ansorge, W., and Sawyer, A. (2005) High throughput production of mouse monoclonal antibodies using antigen microarrays. *Proteomics* **5,** 4070–4081
19. Nilsson, P., Paavilainen, L., Larsson, K., Odling, J., Sundberg, M., Andersson, A. C., Kampf, C., Persson, A., Al-Khalili Szigyarto, C., Ottosson, J., Bjorling, E., Hober, S., Wernerus, H., Wester, K., Ponten, F., and Uhlen, M. (2005) Towards a human proteome atlas: high-throughput generation of mono-specific antibodies for tissue profiling. *Proteomics* **5,** 4327–4337
20. Angenendt, P., Wilde, J., Kijanka, G., Baars, S., Cahill, D. J., Kreutzberger, J., Lehrach, H., Konthur, Z., and Glokler, J. (2004) Seeing better through MIST: evaluation of monoclonal recombinant antibody fragments on microarrays. *Anal. Chem.* **76,** 2916–2921
21. Anderson, N. L., Anderson, N. G., Haines, L. R., Hardie, D. B., Olafson, R. W., and Pearson, T. W. (2004) Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). *J. Proteome Res.* **3,** 235–244
22. Ng, J. H., and Ilag, L. L. (2002) Biomedical applications of protein chips. *J. Cell. Mol. Med.* **6,** 329–340
23. Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M., and Snyder, M. (2001) Global analysis of protein activities using proteome chips. *Science* **293,** 2101–2105
24. Ptacek, J., Devgan, G., Michaud, G., Zhu, H., Zhu, X., Fasolo, J., Guo, H., Jona, G., Breitkreutz, A., Sopko, R., McCartney, R. R., Schmidt, M. C., Rachidi, N., Lee, S. J., Mah, A. S., Meng, L., Stark, M. J., Stern, D. F., De Virgilio, C., Tyers, M., Andrews, B., Gerstein, M., Schweitzer, B., Predki, P. F., and Snyder, M. (2005) Global analysis of protein phosphorylation in yeast. *Nature* **438,** 679–684
25. [No authors listed] (1998) 1997 update of recommendations for the use of tumor markers in breast and colorectal cancer. Adopted on November 7, 1997, by the American Society of Clinical Oncology. *J. Clin. Oncol.* **16,** 793–795
26. Uhlen, M., Bjorling, E., Agaton, C., Szigyarto, C. A., Amini, B., Andersen, E., Andersson, A. C., Angelidou, P., Asplund, A., Asplund, C., Berglund, L., Bergstrom, K., Brumer, H., Cerjan, D., Ekstrom, M., Elobeid, A., Eriksson, C., Fagerberg, L., Falk, R., Fall, J., Forsberg, M., Bjorklund, M. G., Gumbel, K., Halimi, A., Hallin, I., Hamsten, C., Hansson, M., Hedhammar, M., Hercules, G., Kampf, C., Larsson, K., Lindskog, M., Lodewyckx, W., Lund, J., Lundeberg, J., Magnusson, K., Malm, E., Nilsson, P., Odling, J., Oksvold, P., Olsson, I., Oster, E., Ottosson, J., Paavilainen, L., Persson, A., Rimini, R., Rockberg, J., Runeson, M., Sivertsson, A., Skollermo, A., Steen, J., Stenvall, M., Sterky, F., Stromberg, S., Sundberg, M., Tegel, H., Tourle, S., Wahlund, E., Walden, A., Wan, J., Wernerus, H., Westberg, J., Wester, K., Wrethagen, U., Xu, L. L., Hober, S., and Ponten, F. (2005) A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell. Proteomics* **4,** 1920–1932
27. Anderson, N. L., Polanski, M., Pieper, R., Gatlin, T., Tirumalai, R. S., Conrads, T. P., Veenstra, T. D., Adkins, J. N., Pounds, J. G., Fagan, R.,

and Lobley, A. (2004) The human plasma proteome: a nonredundant list developed by combination of four separate sources. *Mol. Cell. Proteomics* **3,** 311–326
28. Michaud, G. A., Salcius, M., Zhou, F., Bangham, R., Bonin, J., Guo, H., Snyder, M., Predki, P. F., and Schweitzer, B. I. (2003) Analyzing antibody specificity with whole proteome microarrays. *Nat. Biotechnol.* **21,** 1509–1512
29. Blank, M., and Blind, M. (2005) Aptamers as tools for target validation. *Curr. Opin. Chem. Biol.* **9,** 336–342
30. Ronnmark, J., Gronlund, H., Uhlen, M., and Nygen, P.-A. (2002) Human immunoglobulin A (IgA)-specific ligands from combinatorial engineering of protein A. *Eur. J. Biochem.* **269,** 2647–2655
31. Kilpatrick, K. E., Kerner, S., Dixon, E. P., Hutchins, J. T., Parham, J. H., Condreay, J. P., and Pahel, G. (2002) In vivo expression of a GST-fusion protein mediates the rapid generation of affinity matured monoclonal antibodies using DNA-based immunizations. *Hybrid. Hybridomics* **21,** 237–243
32. Pasqualini, R., and Arap, W. (2004) Hybridoma-free generation of monoclonal antibodies. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 257–259
33. Uhlen, M., and Ponten, F. (2005) Antibody-based proteomics for human tissue profiling. *Mol. Cell. Proteomics* **4,** 384–393
34. der Maur, A. A., Zahnd, C., Fischer, F., Spinelli, S., Honegger, A., Cambillau, C., Escher, D., Pluckthun, A., and Barberis, A. (2002) Direct in vivo screening of intrabody libraries constructed on a highly stable single-chain framework. *J. Biol. Chem.* **277,** 45075–45085
35. Alluri, P. G., Reddy, M., Bachhawat-Sikder, K., Olivos, H. J., and Kodadek, T. (2003) Isolation of protein ligands from large peptoid libraries. *J. Am. Chem. Soc.* **125,** 13995–14004
36. Feldhaus, M. J., Siegel, R. W., Opresko, L. K., Coleman, J. R., Feldhaus, J. M., Yeung, Y. A., Cochran, J. R., Heinzelman, P., Colby, D., Swers, J., Graff, C., Wiley, H. S., and Wittrup, K. D. (2003) Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces cerevisiae surface display library. *Nat. Biotechnol.* **21,** 163–170
37. de Wildt, R. M. T., Tomlinson, I. M., Ong, J. L., and Holliger, P. (2002) Isolation of receptor-ligand pairs by capture of long-lived multivalent interaction complexes. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 8530–8535
38. Apweiler, R., Bairoch, A., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M. J., Natale, D. A., O'Donovan, C., Redaschi, N., and Yeh, L. S. (2004) UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* **32,** D115–D119
39. Hermjakob, H., Montecchi-Palazzi, L., Bader, G., Wojcik, J., Salwinski, L., Ceol, A., Moore, S., Orchard, S., Sarkans, U., von Mering, C., Roechert, B., Poux, S., Jung, E., Mersch, H., Kersey, P., Lappe, M., Li, Y., Zeng, R., Rana, D., Nikolski, M., Husi, H., Brun, C., Shanker, K., Grant, S. G., Sander, C., Bork, P., Zhu, W., Pandey, A., Brazma, A., Jacq, B., Vidal, M., Sherman, D., Legrain, P., Cesareni, G., Xenarios, I., Eisenberg, D., Steipe, B., Hogue, C., and Apweiler, R. (2004) The HUPO PSI's molecular interaction format—a community standard for the representation of protein interaction data. *Nat. Biotechnol.* **22,** 177–183