Affinity Proteomics for Systematic Protein Profiling of Chromosome 21 Gene Products in Human Tissues*

Charlotta Agaton‡, Joakim Galli§, Ingmarie Höldén Guthenberg§, Lars Janzon§, Marianne Hansson§, Anna Asplund¶, Eva Brundell§, Susanne Lindberg§, Irene Ruthberg§, Kenneth Wester¶, Dorothee Wurtz§, Christer Höög||, Joakim Lundberg‡, Stefan Ståhl‡, Fredrik Pontén¶, and Mathias Uhlén‡**

Here we show that an affinity proteomics strategy using affinity-purified antibodies raised against recombinant human protein fragments can be used for chromosome-wide protein profiling. The approach is based on affinity reagents raised toward bioinformatics-designed protein epitope signature tags corresponding to unique regions of individual gene loci. The genes of human chromosome 21 identified by the genome efforts were investigated, and the success rates for de novo cloning, protein production, and antibody generation were 85, 76, and 56%, respectively. Using human tissue arrays, a systematic profiling of protein expression and subcellular localization was undertaken for the putative gene products. The results suggest that this affinity proteomics strategy can be used to produce a proteome atlas, describing distribution and expression of proteins in normal tissues as well as in common cancers and other forms of diseased tissues. Molecular & Cellular Proteomics 2:405–414, 2003.

A crucial challenge in the post-genomic era is to utilize the genome information for better understanding of protein expression and function. Most frequently, this has been done using protein separation techniques coupled with mass spectrometry analysis methods (1–4). Recently, several methods have been described for “genome-based” proteomics approaches aimed to enumerate and functionally catalogue all the components of the proteome by a gene-by-gene approach. Initially, these studies have involved the analysis of networks of protein interactions, using genetic tools such as two-hybrid systems (5–7) and protein complex pull-outs (8–10). Furthermore, structural genomics strategies have been developed to allow streamlined unit operations for protein structure determination (11). As a complement to these efforts, a more general tool for genome-based proteomics has been proposed (12–14) in which protein-specific affinity reagents are generated, which can be used in a stepwise manner for a wide range of functional and biochemical studies.

There are several challenges for such a strategy, here called affinity proteomics. First, the success rates for recombinant expression of human proteins in bacteria are normally relatively low (11). Second, both polyclonal and monoclonal antibodies often give cross-reactivity to other proteins or show high background binding in tissue sections (14). Third, high-throughput schemes have been difficult to employ both with regard to automation and cost-efficiency. These problems have seriously hampered whole proteome applications to generate antibodies and have prompted us to try new approaches for protein expression and antibody generation.

Here we describe a strategy for affinity proteomics based on the generation of protein epitope signature tags (PrESTs).¹ The protein fragments (PrESTs) are designed to contain unique epitopes present in the native protein suitable for triggering the generation of antibodies of high selectivity. A strategy based on the generation of polyclonal antibodies has been followed. The antibodies are affinity purified using the target protein as ligand. The use of polyclonal antibodies, instead of monoclonal antibodies, makes the generation relatively cost-effective and increases the probability of specific recognition of the target protein during a variety of denaturing conditions. Although monoclonal antibodies are attractive for routine use, such as diagnostics, therapeutics, or protein arrays, the use of polyclonal antibodies is well suited for the protein profiling, in which the proteins have been denatured in different conditions, such as with formalin for the paraffin-imbedded tissues or with SDS for the protein extracts analyzed on Western blots.

As a pilot project, we describe an analysis of the putative gene products of human chromosome 21. Although the project was designed to provide a general proof-of-concept for

¹ The abbreviation used is: PrEST, protein epitope signature tag.
whole proteome analysis, specific information gained from the chromosome 21-encoded proteome may be valuable for the studies of a range of common complex diseases that map to this chromosome, also including disorders such as cancer and Down syndrome that result from deletion or duplication of sequences on this chromosome. All genes of human chromosome 21 (15) with open reading frames larger than 100 amino acid residues were subjected to PrEST design and production. A set of PrESTs representing proteins with both known and unknown function were selected for further analysis, and antibodies were generated and affinity-purified using the PrESTs as ligands. The results suggest that this affinity proteomics strategy can be used for systematic generation of protein profiles and subcellular localization data on a genome-wide level.

EXPERIMENTAL PROCEDURES

Cloning—RT-PCR was performed using Superscript One Step RT-PCR with Platinum Taq (Life Technologies, Rockville, MD) using a human Total RNA Panel IV (CLONTECH, Palo Alto, CA) as template. Flanking restriction sites NotI and Ascl, respectively, were introduced into the fragments through the specific primers to allow “in frame” cloning into the expression vector pAft8bc (13). The downstream primer was biotinylated to allow solid-phase cloning as previously described (13). The resulting biotinylated PCR products were immobilized onto Dynabeads M280-streptavidin (Dynal Biotech, Oslo, Norway). The fragments were released from the solid support by NotI-Ascl digestion, ligated into pAft8bc, and transformed into Escherichia coli BL21(DE3) cells. The sequences of the clones were verified by dye-terminator cycle sequencing of purified plasmid DNA.

Expression and Purification of Fusion Proteins—BL21(DE3) cells, harboring the different expression constructs, were inoculated into culture medium as described before (13) and induced at A600 nm = 1.0 with isopropyl-β-D-thiogalactopyranoside (Sigma Aldrich, St. Louis, MO) at a final concentration of 1 mM, and the incubation was continued overnight at 25 °C. The cells were harvested by centrifugation, and the pellet was resuspended in 5 ml lysis buffer (7 M guanidinium-HCl, 47 mM NaH₂PO₄, 2.65 mM NaH₂PO₄, 10 mM Tris-HCl, 100 mM NaCl, pH 8.0, 20 mM β-mercaptoethanol) and incubated for 2 h at 37 °C. The solutions were sheared with a bore 0.8-mm needle (Becton Dickinson, Franklin Lakes, NJ) and a 10-ml syringe (Becton Dickinson) to reduce the viscosity. After centrifugation, the supernatants, containing the denatured and solubilized gene products, were filtered (0.45 μm, Sartorius AG, Goettingen, Germany) prior to the immobilized metal ion affinity chromatography procedure. The immobilized metal ion affinity chromatography purification of the His₆-tagged fusion proteins was performed on TALON metal (Co²⁺)⁺ affinity resins and gravity columns (CLONTECH) as recommended by the manufacturer.

Immunizations and Affinity Enrichment of Antibodies—New Zealand and rabbits were immunized with the purified PrEST fragments by AgriSera AB (Vännäs, Sweden) in accordance with the national guidelines (Swedish permit no.A125/00). The rabbits were immunized subcutaneously with 200 μg of antigen in Freund’s complete adjuvant as the primary immunization and boosted three times in 3-week intervals with 100 μg of antigen in Freund’s incomplete adjuvant. Enrichment of antibodies reactive to the fusion proteins was performed as described previously (16). Western blots with proteins extracted from 14 different human tissues were prepared by separation on Novex Nu-page gels (4–12%, 4-morpholineethanesulfonic acid buffer system, Novex, Invitrogen, San Diego, CA) followed by transfer to polyvinylidene difluoride membranes (Immobilon-P, Millipore, Bedford, MA) in Nupage transfer buffer (according to the manufacturer’s protocol). The human tissues were skin (SK), breast (BR), omental fat (OF), tonsil (TO), placenta (PL), lung (LU), kidney, (KI), liver (LI), gallbladder (GB), colon (CO), ileum (IL), duodenum (DU), ventricle (VE), and esophagus (ES). The membranes were incubated with the affinity-enriched affinity reagents (1:75) and washed and probed with horseradish peroxidase-conjugated anti-rabbit IgG (Sigma), SuperSignal West Dura (Pierce, Rockford, IL) was used as detection system, and digital images of the chemiluminescence were monitored using a ChemiImager (Alpha Innotech, San Leandro, CA).

RESULTS

PrEST Selection—As a starting point, we performed bioinformatics analysis of all the putative genes of human chromosome 21 found by Hattori et al. (15). Out of the 225 putative genes mapped to the chromosome, 168 of them were found to contain an open reading frame longer than 100 amino acids (300 base pairs). The results are summarized in Fig. 1. Sequence analysis algorithms were used to select suitable PrEST coding regions for all of these genes. Transmembrane spanning regions were identified using transmembrane hidden Markov model algorithms (17, 18), and these regions were avoided in the selection of PrESTs. To minimize cross-reactivity of the affinity reagents, regions and domains with homology to other human proteins were omitted using homology search programs, such as BLAST (19). The protein fragments, with a size of 100–150 residues, with the lowest homology to other proteins in the human proteome were thus selected. A total of 168 PCR primer pairs were designed to flank the selected (PrEST) regions.
Source of Chromosome 21 Genes—To clone the PrESTs, we used a pool of RNA from various human tissues in an RT-PCR approach. Use of a single RNA pool consisting of material from human brain, spinal cord, placenta, and liver proved to be a valuable approach allowing us to amplify 120 out of the 168 genes (71%) in a first attempt. By designing new primers for those genes that failed, a second round yielded PCR products from 22 additional genes. The overall success rate for the two rounds of RT-PCR was thus 142 out of 168 genes (85%).

Cloning and Expression of the PrESTs—The 142 amplified gene fragments were cloned into an expression vector in frame with a dual affinity tag consisting of a hexahistidyl tag, allowing purification on nickel columns, in frame with an immunopotentiating albumin binding domain (20). All of the amplification products were successfully cloned and sequence verified. The E. coli clones expressed in liquid cultures were lysed in the presence of guanidinium chloride, and the cleared lysates were passed over nickel-containing matrices. In a majority of cases, more than 0.5 mg of proteins of the right size was obtained from 50-ml cultures. A summary of the results of all the putative chromosome 21 genes is shown in Fig. 1. The overall success rate for the protein production using a single expression event was 108 out of 142 (76%).

Generation and Analysis of the Affinity Reagents—Two rabbits were immunized with each of the purified PrESTs and the polyclonal antisera were individually affinity purified using the specific PrESTs as ligands to obtain PrEST-specific antibodies. As a first step in the protein expression analysis, and to validate functionality of the antibodies, screening for expression of the various gene products was performed using protein extracts from a panel of human tissues in a Western blotting format. In Fig. 2, the results before and after affinity purification are shown for a protein with known function (SOD1) and a protein with unknown function (KIAA0539). In both cases, the affinity purification yields a protein pattern with bands of the expected sizes (16 kDa and 252 kDa, respectively). It is noteworthy that several smaller bands displaying the same tissue distribution were obtained using the KIAA0539 affinity reagent. Plausible explanations for this could be the existence of splice variants, proteolysis, protein modifications, such as glycosylation, or simply cross-reactivity of the affinity reagent. The difference in size makes this protein interesting for further studies to also analyze the possibility of mRNA splice variants or protein modifications. In Fig. 3, examples from eight antibodies with known (CCT8, SOD1, CBR1, and ATP50) and unknown (KIAA0539, SAMSN-1, WDR9, and ZNF294) function are shown. For the four known gene products (Fig. 3, A–D), a band of the expected size was obtained in all cases. Furthermore, the expression profiles of the known genes were in accordance with previously published data (21–23). For the four gene products of unknown function, bands with expected sizes were also detected, as deduced from the amino acid sequences (Fig. 3, E–H). A summary of the results of the tissue Western analysis is shown in Fig. 1. The overall success rate of affinity-purified antibodies as determined by a specific band in the tissue Western analysis was 54/96 (56%).

Analysis of Cellular Localization using a Tissue Array—To allow for high-throughput protein profiling, tissue arrays consisting of a variety of human tissues were constructed for analysis of cellular localization of the gene products. The arrays were produced by transferring cylinders of tissues from a multitude of donor blocks consisting of archival paraffin-embedded tissues to a single recipient paraffin block (24, 25). The recipient block was subsequently used to produce a large number of slides containing 52 different tissues as triplicate spots with a diameter of 1.0 mm (Fig. 4). The tissue array permitted simultaneous analysis of both cellular and subcellular localization of expressed proteins in more than 50 different tissues. Fig. 4 shows an example of the tissue array analysis for the unknown human protein KIAA0539. The analysis showed that the protein was expressed in glandular cells of the colon and in smooth muscle cells beneath the colonic mucosa. A more detailed analysis of crypts revealed that the protein was localized in the Goblet cell vacuoles.

In Fig. 5, illustrative examples of the tissue array results of affinity reagents to the eight selected gene products are shown. The protein profiles of the four gene products having known function are well in accordance with the intracellular and tissue distribution suggested by earlier studies. For example, the detoxification enzyme, superoxide dismutase (SOD1), was highly expressed in the liver, whereas ATP50, a subunit of the ATP synthetase in the mitochondrial F1 complex, was expressed in virtually all cells and visualized with an idiosyncratic microgranular staining pattern. The affinity reagents recognizing proteins with unknown function gave distinct staining patterns yielding specific expression and localization profiles. ZNF294, for example, shows a predominantly cytoplasmic staining with variable distribution in different tissues. A strong immunostaining was found in testis, where seminiferous tubules including germinative epithelia and Sertoli as well as Leydig cells reacted with the ZNF294 affinity reagents. A crescent-shaped immunoreactivity was found in spermatids, although mature sperms appeared negative. The WDR9 affinity reagent was negative in a majority of tissues, but a focal and distinct granular immunoreactivity was found in hepatocytes and zona reticulosa cells in the adrenal gland. The distribution and pattern of immunoreactivity resemble that of lipofuscin and suggest that WDR9 has some relation to the lysosomal deposition of lipofuscin. KIAA0539 was expressed by mucus-producing tissues, such as colon and the small intestine and the uterine cervical mucosa. In addition, smooth muscle cells of the prostatic gland showed high KIAA0539 expression.

In Fig. 6, the expression patterns for 12 different gene products are shown for two brain tissues (cerebrum and cerebellum). Distinct protein expression profiles can be ob-
served, such as the dendritic/axonal positivity of MCM3, the expression of ZNF294 and C21orf33 in Bergman astrocytes, and the mosaic-like pattern of GABPA in cerebrum.

**DISCUSSION**

Here we have analyzed the putative genes predicted from the genome sequence of chromosome 21. Chromosome 21 was chosen to establish the applicability of a chromosome-wide study. The original 225 genes described by Hattori et al. (15) were included in this pilot study, although recent predictions have indicated a slightly larger gene content of 238 genes (26). We chose to analyze proteins with an open reading frame larger than 100 amino acids, but obviously smaller proteins could be included in an extended study. Using a combination of bioinformatics, recombinant protein expression, and cost-effective antibody production, we have generated a large set of affinity ligands in the form of monospecific polyclonal antibodies. We show that these affinity-purified antibodies are useful tools to explore protein expression profiles using human tissue arrays. Together, these methodologies allow for a systematic approach to generate and use affinity reagents without dependence on clone repositories or cumbersome expression screening procedures.

The affinity proteomics strategy is based on the generation of PreESTs, which are amenable both to recombinant protein expression and the generation of specific antibody reagents.

**FIG. 1.** A schematic overview of the putative human chromosomal 21 genes and the results from the PreEST protein production and antibody generation. The chromosomal locations (in Mb from the centromere) of the putative genes predicted by Hattori et al. (15) are shown as genes with known function (green), putative genes with open reading frames of more than 100 amino acids (blue), and putative genes with no open reading frame larger than 100 amino acids (red). The boxes in the left lane represent the results of the protein expression approach with successful protein production (green), no protein production (red), or no clones obtained from the RT-PCR approach and thus not used for protein production (blue). The boxes in the right lane represent the results of the immunization protocol as determined by tissue Western analysis with the results indicated by the color of the boxes with functional antibodies (green) and no apparent specific antibodies (red).
The availability of the whole genome sequence allows the most unique (nonhomologous) region within a specific protein to be selected for protein expression. Comparative algorithms are used to ensure that highly homologous protein domains in the human proteome are not being used for protein expression and thus not included as immunogens during immunization. Second, transmembrane regions are omitted, because they are difficult to express and purify (27) and are not likely to be suited for immunolocalization studies due to less accessibility for the affinity reagents. Third, protein fragments of 100–150 amino acid residues are chosen to facilitate cloning and protein expression and to possibly provide conformational epitopes not obtained using shorter peptides. In most cases, the PrESTs will not have a fully native fold. However, the fact that the subsequent protein profiling is performed on denaturated proteins makes it attractive to generate affinity reagents toward partly denaturated proteins. It is in this context an advantage to use polyclonal affinity reagents with a multiple of binding epitopes to increase the probability that some epitopes are present during the conditions for the different protein profiling procedures.

Fig. 3. Tissue distribution of the gene products as determined by Western blot using protein extracts from human tissues. The order of the tissues is the same as in Fig. 2. CCT8, SOD1, CBR1, ATP5O, KIAA0539, SAMSN-1, WDR9, and ZNF294 were detected using the PrEST specific affinity reagents. The molecular mass of each gene product deduced from the amino acid sequence is given.

Fig. 4. Photomicrograph illustrating the cellular and tissue distribution in a tissue array slide. The arrays were probed with the affinity-purified antisera as affinity reagents for the specific gene products (brown), and hematoxylin was used for counter staining (blue). Low-power magnification shows the arrangement of “spots” on the array slide. Increasing magnification shows an example of two sections representing normal colon (left) and two sections representing normal ovary (right). At higher magnifications the colonic mucosa is highlighted, and a strong immunoreactivity against the KIAA0539 gene product is found in epithelial cells aligning the crypts and colonic surface. The high-power view suggests a granular distribution of immunoreactivity in mucin-containing vacuoles of Goblet cells.
An important finding is that a single pool of RNA extracted from four human tissues can be used for efficient RT-PCR cloning. Using this strategy, rapid cloning was performed without dependence on availability of clones through various clone repositories. This is of particular importance for whole proteome efforts, because many of the putative genes are solely predicted from the genome sequence. An alternative would be to use chromosomal DNA as a source of genes. However, the size of an average human exon is only 48 amino acids (28), and the production of PrESTs would require a cumbersome assembly of several exons using “PCR splicing” techniques (29, 30). The genes encoding the PrEST regions can also be synthesized in vitro using solid-phase gene assembly (31), but this requires the synthesis of a multitude of oligonucleotides for each PrEST. Using a single complex RNA pool as a source of genes, the overall success rate for the two rounds of RT-PCR was 148 out of 168 genes (85%). The genes not amplified could be wrongly predicted or may rep-
resent transcriptionally inactive pseudogenes or genes not expressed in the tissues used for RNA sampling. These possibilities could be resolved by further studies using comparative analysis with a genome of a closely related species (e.g., the mouse) to detect functional conservation between orthologs or with extended RNA profiling studies using RNA derived from a larger set of human tissues and cell types.

A single E. coli recombinant protein expression system has been used that was developed to allow a streamlined procedure to collect both soluble and insoluble (inclusion bodies) proteins. This process yielded a higher success rate for the protein production as compared with collecting only soluble proteins, because a high proportion of the PrESTs formed inclusion bodies (data not shown). An additional albumin-binding domain was included in the dual affinity tag to increase the immune response as described earlier (13). The expression vector allowed us to produce PrESTs for 76% of the cloned genes (108/142) using a single-pass procedure. For proteins annotated by SwissProt, the success rate for "transmembrane proteins" was 13 out of 28, for "transcription factors" four out of six, while for the category "DNA-binding" or "nuclear" 19 out of 24 were successfully produced.

The protocol described here involves affinity purification of the generated antisera. As shown in Fig. 3, the affinity purification is necessary to avoid cross-reactivity in the protein profiling. The tissue Western analysis suggests that functional antibodies are obtained from 56% (54/96) of the PrESTs (Fig. 1). Obviously, the quality assurance of the generated antibodies is an important issue, and additional assays may be considered in the future, such as the use of mass spectrometry and/or protein arrays. In this study, we have used affinity-purified polyclonal antibodies, but it might be advantageous to also produce monoclonal binding reagents, preferably using in vitro selection methods. Monoclonal antibodies or in vitro-selected binding moieties would allow unlimited access to validated affinity reagents and would be valuable for development of medical diagnostics or high-throughput protein quantification analysis. Such reagents could be single-chain recombinant antibodies (32) or alternative binding moieties such as the affibody molecules (33) or aptamers (34). How-

![Expression patterns for 12 different gene products shown in brain tissues.](image)

**Fig. 6.** Expression patterns for 12 different gene products shown in brain tissues. a, cerebrum; b, cerebellum. α-KIAA0539 (1) staining was completely negative, and in tissues stained with α-C21orf63 (9) only very weak cytoplasmic immunoreactivity was found. Staining with α-ZNF294 (2) showed a weak positivity in vessel walls and nuclei of glial cells with hinted accentuation of Bergmann astrocytes (border between molecular and granular layer of the cerebellum). α-ATPSO (3) showed a general cytoplasmic positivity with a microgranular staining pattern consistent with mitochondrial localization of the antigen. α-SOD1 (4) showed a dominating nuclear staining pattern with moderate intensity and clear positivity in axons of the cerebellum. α-CBR1 (5) showed a strong immunoreactivity in astrocytic cells with nuclear positivity and dendritic/axonal positivity, while nuclei of neural cells and a subset of glial cells appeared negative. α-SMT3H1 (6) showed weak cytoplasmic staining with some accentuation of cell membranes, and a more distinct staining could be seen in cells of the granular layer of cerebellum. α-MCM3 (7) showed a strong nuclear positivity in all cells; the vessels appeared negative. α-RA4 (8) showed a moderate intense diffuse nuclear and cytoplasmic staining in cells of both cerebrum and cerebellum; the vessels appeared negative. α-PRED4 (10) staining showed a diffuse cytoplasmic positivity in glial cells from cerebrum, and in cerebellum a subset of cells in the granular layer as well as cells in the molecular layer appeared positive. α-C21orf33 (11) showed a microgranular positivity in cytoplasm of glial and neural cells of the cerebrum, and the Bergman astrocytes appeared with a distinct and strong staining pattern. α-GABPA (12) showed an apparent cytoplasmic staining with stronger staining in cell membranes leading to a mosaic-like pattern of immunoreactivity in cerebrum. A more diffuse pattern with axonal accentuation could be seen in cerebellum, and Purkinje cells appeared negative.
ever, it is not likely that the monoclonal reagents would be suitable for the initial protein profiling as described here. The range of denaturation conditions in the various biological assays makes it cumbersome and challenging to obtain monoclonal antibodies that can be used as single and defined reagents across the whole platform of functional assays.

An important long-term goal is to use the affinity reagents for simultaneous detection of nonlabeled target proteins using microfabricated devices, such as protein arrays. The detection of nonlabeled target proteins is a technical challenge, but might be achieved using surface plasmon resonance (35), mass spectrometry (35), or fluorescent resonance energy transfer dyes (36). The PrEST strategy described here is well suited for such efforts, and the PrESTs can thus be used for the generation of both monospecific polyclonal antibodies as well as monoclonal reagents.

The tissue arrays contained triplicates of 52 human tissues of both normal and diseased origin and gave a large number of suggestions for further functional exploration, in particular for the proteins with previous unknown function. As an example, the analysis suggested that the KIAA0539 protein is expressed by mucus-producing cells of the gastrointestinal tract and cervix, but also by small muscle cells in the prostate. A complex expression pattern was found for ZNF294 with a strong and peculiar crescent-shaped localization in spermatides, contrasted by a gradient-like expression in umbrella cells of the urinary bladder and a completely desmosomal distribution in tongue. Another interesting observation was a distinct expression profile of WDR9 in liver and adrenal glands. The strong staining found in adrenal glands is consistent with lipofuscin deposits in patients suffering from severe illness. The analysis also gave new insights about proteins with known function, for example the chaperone CCT8, which was strongly expressed in tumor cells from a mixed germ cell tumor compared with the surrounding stroma. Furthermore, the carbonyl reductase, CBR1, was expressed in the cerebellar cortex, specifically in Purkinje cells. For these proteins, it will be interesting to perform further analysis of specialized tissue arrays, for example derived from human individuals having Down syndrome.

The value of the produced affinity reagents can be further extended by using them for affinity capture of native or denatured proteins from complex biosamples followed by biochemical studies, such as mass spectrometry studies to analyze protein modification or protein interactions using protein “pull-outs.” In this manner, it might be possible to study additional protein characteristics not included in the present approach, such as alternative spliced forms, and post-translational modifications, such as critical phosphorylation patterns of proteins. The polyclonal nature of the affinity reagents might also make them suitable for comparative proteomics to study expression of the functional protein homologue also in related species, such as mice, rats, and chimpanzees. This would allow genomics to be converted into biological knowledge, including information regarding protein size, structures, modifications, interactions, expression patterns, and cellular localizations.

The question arises whether this strategy can be used for high-throughput affinity proteomics on a genome-wide basis. The results presented here are, in this context, very encouraging. The overall success rate was 36% for a single-pass procedure including cloning, protein production, and antibody generation. These yields suggest that at least a 50% success rate can be achieved using an iterative procedure where the fraction of proteins not obtained in the first round can be investigated using a PrEST selected from a different region of the gene. In addition, it is also possible to complement with alternative immunizations schemes, such as alternative animals or in vitro selection procedures. A high-throughput scheme can thus be envisioned to continuously process hundreds of new putative genes and gene products every week. Thus, the whole human proteome can possibly be analyzed with affinity proteomics within a time frame of 5–10 years.

Acknowledgments—We thank Marie-Laure Yaspo, Angela Scott, Per-Åke Nygren, and John Bell for valuable discussions and Karin Bergstrand, Karin Johansson, and Mats Lindskog for technical and computational assistance.

* This work was supported by grants from the Knut and Alice Wallenberg Foundation, the Swedish Foundation for Strategic Research, and the Swedish Research Council. 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.

** To whom correspondence should be addressed: Department of Biotechnology, AlbaNova University Center, Royal Institute of Technology, Roslagsvägen 30 B, 106 91 Stockholm, Sweden. Tel.: +46855378325; Fax: +46855378482; E-mail: mathias.uhlen@biotech.kth.se.
A. Taylor, P., Bennett, K., Boutliker, K., Yang, L., Wolting, C., Donaldson, I., Schandorff, S., Shewnarane, J., et al. (2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature 415, 180–183

10. Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., et al. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415, 147–147

11. Hammarström, M., Heilig, N., van den Berg, S., Berglund, H., and Härd, T. (2002) Rapid screening for improved solubility of small proteins produced as fusion proteins in Escherichia coli. Protein Sci. 11, 313–321

12. Larsson, M., Brundell, E., Nordfors, L., Höög, C., Uhlen, M., and Ståhl, S. (1996) A general bacterial expression system for functional analysis of cDNA-encoded proteins. Protein Expr. Purif. 7, 447–457

13. Larsson, M., Gräslund, S., Yuan, L., Brundell, E., Uhlen, M., Höög, C., and Ståhl, S. (2000) High-throughput protein expression of cDNA-products as a tool in functional genomics. J. Biotechnol. 80, 143–157

14. Gräslund, S., Falk, B., Brundell, E., Höög, C., and Ståhl, S. (2002) A high-stringency proteomics concept aimed for generation of antibodies specific for cDNA-encoded proteins. Biotechnol. Appl. Biochem. 35, 75–82

15. Hattori, M., Fujiyama, A., Taylor, T. D., Watanabe, H., Yada, T., Park, H. S., Toyoda, A., Ishii, K., Totoki, Y., Choi, D. K., Soeda, E., Ohki, M., Takagi, T., Sakaki, Y., Taudien, S., et al. (2000) The DNA sequence of human chromosome 21. Nature 405, 311–319

16. Starborg, M., and Höög, C. (1995) The murine replication protein P1 is differentially expressed during spermatogenesis. Eur. J. Cell Biol. 68, 206–210

17. Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. Proc. Int. Conf. Intell. Syst. Mol. Biol. 6, 175–182

18. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J. Mol. Biol. 305, 567–580

19. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402

20. Ståhl, S., Nilsson J., Hober, S., Uhlen, M., and Nygren PÅ. (1999). Affinity fusions, gene expression, in Bioprocess Technology: Fermentation, Biocatalysis and Bioseparation (Flickinger, M. C., and Drew, S. W., ed.) pp. 8–22. John Wiley & Sons, New York

21. Golden, T. R., and Pedersen, P. L. (1998) The oligomycin sensitivity conferring protein of rat liver mitochondrial ATP synthase: Arginine 94 is important for the binding of OSCP to F1. Biochemistry 37, 13871–13881

22. Forrest, G. L., and Gonzalez, B. (2000) Carbonyl reductase. Chem. Biol. Interact. 129, 21–40

23. Brown, C. R., Doxsey, S. J., Hong-Brown, L. Q., Martin, R. L., and Welch, W. J. (1996) Molecular chaperones and the centrosome. A role for TCP-1 in microtubule nucleation. J. Biol. Chem. 271, 824–832

24. Rimm, D. L., Camp, R. L., Charette, L. A., Costa, J., Olsen, D. A., and Reiss, M. (2001) Tissue microarray: A new technology for amplification of tissue resources. Cancer J. 7, 24–31

25. Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, M. J., Sauter, G., and Kallioniemi, O. P. (1998) Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat. Med. 4, 844–847

26. Gitton, Y., Dahmane, N., Baïk, S., Ruiz i Altaba, A., Neidhardt, L., Scholze, M., Herrmann, B. G., Kahlken, P., Benkahlia, A., Schirrner, S., Yildirim, R., Herwig, R., Lehrah, H., and Yaspo, M. L. (2002) A gene expression map of human chromosome 21 orthologues in the mouse. Nature 420, 586–590

27. Mioux, B., and Walker, J. E. (1996) Over-production of proteins in Escherichia coli: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J. Mol. Biol. 260, 289–298

28. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., et al. (2001). Initial sequencing and analysis of the human genome. Nature 409, 860–921

29. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K., and Pease, L. R. (1989) Engineering hybrid genes without the use of restriction enzymes: Gene splicing by overlap extension. Gene 77, 61–68

30. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77, 51–59

31. Nord, K., Nilsson, J., Nilsson, B., Uhlen, M., and Nygren, P. Å. (1995). A combinatorial library of an alpha-helical bacterial receptor domain. Protein Eng. 8, 601–606

32. Liu, B., and Marks, J. D. (2000) Applying phage antibodies to proteomics: selecting single chain Fv antibodies to antigens blotted on nitrocellulose. Anal. Biochem. 286, 119–128

33. Nord, K., Gunneriussén, E., Ringdahl, J., Stahl, S., Uhlen, M., and Nygren, P. Å. (1997). Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain. Nat. Biotechnol. 15, 772–777

34. Brody, E. N., and Gold, L. (2000) Aptamers as therapeutic and diagnostic agents. J. Biotechnol. 74, 5–13

35. Weinberger, S. R., Morris, T. S., and Pawlak, M. (2000) Recent trends in protein biochip technology. Pharmacogenomics 1, 395–416

36. Karlström, A., and Nygren, P. Å. (2001) Dual labeling of a binding protein allows for specific fluorescence detection of native protein. Anal. Biochem. 295, 22–30

Affinity Proteomics on Human Chromosome 21

414 Molecular & Cellular Proteomics 2.6