NIH Public Access
Author Manuscript
*Clin Chim Acta*. Author manuscript; available in PMC 2014 July 21.

Published in final edited form as:
*Clin Chim Acta*. 2007 May ; 381(1): 93–97. doi:10.1016/j.cca.2007.02.017.

# Enzymes and Related Proteins as Cancer Biomarkers: a Proteomic Approach

**Shu-Ling Liang** and **Daniel W. Chan**[*]
Department of Pathology, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, Maryland 21287, USA

## Abstract

**Background—**The discovery of cancer biomarkers has become a major focus of cancer research, which holds promising future for early detection, diagnosis, monitoring disease recurrence and therapeutic treatment efficacy to improve long-term survival of cancer patients. Most of the functional information of the cancer-associated genes resides in the proteome. Since cancer is a complex disease, it might require a panel of multiple biomarkers in order to achieve sufficient clinical efficacy.

**Methods—**Serum/plasma is the most accessible biological specimen collected from patients. Therefore, serum proteomic diagnostics would be the most promising new test for cancer. With the advent of new and improved proteomic technologies, such as protein chips and mass spectrometry coupled with advanced bioinformatic tools, it is possible to develop potential cancer biomarkers. However, specimen collection, handling, study design and data analysis are essential components for successful biomarker discovery and validation. Multi-center case control study should be conducted with extensive clinical validation to minimize the impact of possible confounding variables (non-biological).

**Conclusions—**Enzymes and related proteins, such as inhibitors, are promising candidates for cancer diagnostics.

### Keywords

clinical proteomics; cancer biomarkers; mass spectrometry; prostate specific antigen

## Early Detection in Cancer

Despite advances in diagnostic imaging technology, surgical management, and therapeutic modalities, cancer remains a major cause of mortality. A small number of clinically approved biomarkers are available for early diagnosis and/or monitoring of treatment and

© 2007 Elsevier B.V. All rights reserved.

[*]Corresponding author: Daniel W. Chan, Ph.D., DABCC, FACB, Department of Pathology, Room3M05, CRB2, Building, Johns Hopkins Medical Institutions, 1550 Orleans St., Baltimore, MD 21231, USA Tel: +1-410-955-2674, Fax: +1-443-287-6388, dchan@jhmi.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

relapses in order to reduce mortality rates and increase overall survival for cancer patients (1).

Early detection remains the most promising approach to improve long-term survival of cancer patients. The definitive diagnosis of cancer is still relied on histological evaluation of tissues. An ideal cancer biomarker would be a protein or protein fragment that can be easily detected in the patient's blood or urine, but not detected in a healthy patient. Today, the most common use of cancer biomarkers is for the detection of the recurrent disease and monitoring therapy. In the future, better tests that may predict the response of tumors to particular therapeutic drugs in different individuals may be developed, which can be used in patient-tailored therapy and personalized medicine [2].

## Cancer Biomarkers: a Proteomic Approach

### Why single marker is not sufficient?

Most of the functional information of the cancer-associated genes resides in the proteome. Potentially, there could be tens of thousands of intact, modified, and cleaved proteins in the human serum proteome. Proteome is an exceptionally complex biological system involving several proteins that function through posttranslational modifications and dynamic intermolecular collisions with partners. These protein complexes can be regulated by signals emanating from cancer cells, their surrounding tissue microenvironment, and/or from the host. Some proteins are secreted and/or cleaved into the extracellular milieu and may represent valuable serum biomarkers for diagnostic purpose. It is estimated that the cancer proteome may include over 1.5 million proteins as a result of posttranslational processing and modifications. Therefore, finding the elusive single disease-related protein would require the laborious separation and identification of every protein biomarker. It is possible that the discovery and use of these elusive single biomarkers for early detection of disease will not occur because any final clinical application of these biomarkers will be applied to a human population, constituted by vast heterogeneity in individual pathophysiology [3,4].

## Enzymes as biomarkers

Until recently, the low- and ultra-low molecular weight range (<15 kDa) of the circulatory proteome and metabolome went largely uncharacterized. However, this collection of low molecular weight molecules promises to contain a rich source of previously undiscovered biomarkers [5]. This is because of biological processes that give rise to cascades of biomarker fragments from enzyme-generated proteolytic activity. The rich information archive can arise within the unique disease-tissue microenvironment. These low molecular weight molecules exist below the range of detection achieved by conventional two-dimensional gel electrophoresis, since they cannot be efficiently separated by gel-based techniques As a result, investigators have turned to mass spectroscopy, which exhibits its optimal performance in the low molecular weight range. Such complexity clearly highlights the need for ultra-high resolution proteomic technology for robust quantitative protein measurements and data acquisition (6).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

## Experimental Approach of Proteomics

Initial attempts to use mass spectroscopy for the identification of biomarkers for cancer have been very promising [7]. The rapid development of mass spectrometric technologies applied to protein research has catalyzed entirely new experimental approaches in which that comprehensive protein databases for individual conditions can be used to characterize each individual patients and different disease states by studying systems biology [8]. Unlike past attempts that start with a known single-marker candidate and study the mechanism of pathogenesis, proteomic pattern analysis begins with high dimensional data, usually produced by high-throughput mass spectrometry. This approach does not begin with a specific activity or question and is therefore not hypothesis driven (or the hypotheses are very broad, i.e., “there will be a change in protein expression between two patient states”), which limits the need for analyte specific assays prior to the search for biomarkers. One variation of this approach is the use of proteomic profiles without the identities of the actual proteins in a clinical sample such as serum as a diagnostic fingerprint [9]. However, this approach could be affected by the biological variation and specimen collection biases.

## Limitations

In searching of cancer biomarkers using proteomic technology, three factors are crucial for its success: discovery, validation and translation [10,11]. Discovery of candidate biomarkers is easy, since mass spectrometry is such a powerful tool. However, validation of biomarkers in terms of finding biomarkers with consistent clinical significance in a general population is difficult; so is to evaluate the diagnostic performance of the biomarkers with characteristics of a clinical assay. Currently, most serum proteome investigations are discovery driven, in which a large number of proteins are identified with the hope that one or more proteins are uniquely associated with a specific disease state. The discovery, identification, and validation of these disease-associated serum proteins is conceivably a difficult and laborious task, which often requires hundreds, if not thousands of samples to be analyzed and generates enormous amount of data to be processed.

The biomarker research employing proteomic patterns is capable of discriminating between disease (e.g., ovarian cancer, prostate cancer, etc.) and healthy cohorts with relatively high sensitivity and specificity. However, critical assessment of those results showed significant shortcomings and uncertainties in regard to the reproducibility of the findings, identity of the proteins behind the pattern peaks, and validation of the results. Interlaboratory SELDI experiments performed recently alleviated some of the reproducibility concerns [12]. However, validation of the newly discovered biomarkers remained as the most challenging aspect of clinical proteomics. Critical issues that need to be addressed for the validation studies include the specificity ad reproducibility of the marker as well as experimental design and appropriate controls, such as specimen collection, handling, study design and data analysis. Furthermore, it has yet to be established whether the patterns identified reflect cancer-specific phenomena, or epiphenomena related to general inflammatory responses or metabolic disturbances. Nor is it certain that results from different laboratories are comparable, and the effects of sample handling on the patterns obtained should be investigated [13].

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript



## Our approaches to cancer biomarker discovery and validation

- Select the right technologies to discover biomarkers: protein chips and mass spectrometry.
- Use well characterized clinical specimens for discovery and validation.
- Develop bioinformatics tools for data analysis and multiplexing of biomarkers since most diseases (e.g. cancer) are heterogeneous.
- Design multi-center case control study with extensive clinical validation to minimize the impact of possible confounding variables.
- Discover and identify biomarkers (profile is not sufficient) with biological (clinical) significance.

## PSA: New application for an “old” enzyme?

Prostate cancer is the leading cause of cancer in the U.S. The diagnosis of prostate cancer is most often suspected after finding an elevated serum prostate specific antigen (PSA) or abnormal findings on digital rectal examination (DRE). Prostate biopsy is the gold standard for prostate cancer diagnosis. PSA testing has revolutionized prostate cancer detection. The introduction of widespread PSA screening and earlier detection can potentially result in decreased prostate cancer mortality associated with a decline in metastatic disease.

PSA is a protease. It is a 27-kD glycoprotein that is produced by the secretory cells that line the prostate glands (acini). PSA is released from the prostate epithelial cell as a pro-enzyme (pPSA) with a 7-amino acid peptide chain attached to it [14]. As it is secreted into the lumen, the propeptide is removed by human kallikrein 2 to generate active and mature form of PSA. This molecule then undergoes proteolysis to generate inactive PSA, which enters the bloodstream and circulates in an unbound state (free PSA). A small amount of active PSA diffuses into the circulation and is rapidly bound by protease inhibitors, including a-1-antichymotrypsin (ACT) and a-2-macroglobulin [15]. Serum PSA that is measurable by current clinical immunoassays was found to exist primarily as either the free form (free PSA) or as a covalently bound complex withACT (complexed PSA).

Free PSA in serum is now known to be composed of at least three distinct forms of inactive PSA. One form has been identified as the proenzyme, pPSA, and is associated with cancer [14]. In addition to pPSA, there are also significant levels of truncated pPSA, which refers to pPSA in which any of the normal 7 amino acids in the proleader peptide have been removed. The truncated pPSA forms containing proleader peptides of 4 and 2 amino acids, [−4]pPSA and [[−2]pPSA, respectively, are of particular interest. Truncated pPSA forms are more resistant to activation to mature PSA than the intact pPSA with the 7-aa proleader peptide. The truncated pPSA forms are therefore more stable since they cannot be converted to PSA. The sum or all pPSA forms represents about a third of the free PSA typically present in cancer serum.

A second form of PSA, termed benign PSA (BPSA) is an internally cleaved or degraded form of PSA that is more highly associated with benign prostatic hyperplasia (BPH) [16].

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Studies have also demonstrated that this isoform of PSA has been identified in BPH tissues and seminal plasma, which has been shown to have higher levels of internal peptide bond cleavages and is more enzymatically inactive. BPSA concentrations were relatively lower in cancer tissue from the same prostate. BPSA is highly correlated with the presence of BPH nodules in the prostate, the primary pathological feature of BPH.

Although serum PSA is a prostate-specific marker, elevations can be caused by both cancer and benign conditions such as BPH]. Malignant prostate tissue generates more PSA than normal or hyperplastic tissue, probably because of the increased cellularity associated with cancer. Moreover, cancerous tissue may disrupt the prostate-blood barrier, further increasing the serum concentration of PSA.

The traditional cutoff for an abnormal PSA level in the major screening studies has been 4.0 ng/ml. At this level, the sensitivity of PSA has been estimated to be about 70–80%, while the specificity is estimated to be about 30 to 40%. PSA has poorer discriminating ability in men with symptomatic benign prostatic hyperplasia. The test performance statistic that has been best characterized by screening studies is the positive predictive value: the proportion of men with an elevated PSA who have prostate cancer. Overall, the positive predictive value for a PSA level >4.0 ng/ml is approximately 30%, meaning that slightly <1 in 3 men with an increased PSA will have prostate cancer detected on biopsy. For PSA levels between 4.0 to 10.0 ng/ml, the positive predictive value is about 25%; this increases to 42 to 64% for PSA levels >10 ng/ml.

However, nearly 75% of cancers detected within the "gray zone" of PSA values between 4.0 to 10.0 ng/ml are organ confined and potentially curable. The proportion of organ-confined cancers drops to <50% for PSA values >10.0 ng/ml. Thus, detecting the curable cancers in men with PSA levels <10.0 ng/ml presents a diagnostic challenge because the high false-positive rate leads to many unnecessary biopsies. Free PSA is recommended for total PSA in this gray zone. Lower %fPSA is associated with higher probability of prostate cancer. Preliminary research indicated that ProPSA could improve the diagnostic utility of PSA (17). In addition, the use of multiple marker panels has been proposed to improve clinical utility of tumor markers. With the rapid development of mass spectrometry, it is possible that proteomic technologies may also help to lead us into the new frontier of multiple analyte testing, or multiplexed assays, could gain significant new insights into prostate disease management.

## Enzyme related proteins as biomarkers for ovarian cancer- A case study

Early detection remains the most promising approach to improve long-term survival of patients with ovarian cancer. In our biomarker study in ovarian cancer, serum was collected and proteomic expressions were analyzed from a patient population consist of >500 patients with invasive epithelial ovarian cancer, benign pelvic masses and healthy controls from 5 separate cancer institutes [18]. Data from patients with early stage ovarian cancer and healthy women at 2 centers were analyzed independently and the results cross-validated to discover potential biomarkers. The results were validated using the samples from two of the remaining centers. After protein identification, biomarkers were tested on samples from the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript



fifth center, which included healthy women, ovarian cancer and patients with breast, colon, and prostate cancers. Three biomarkers were identified as follows: (a) apolipoprotein A1 (down-regulated in cancer); (b) a truncated form of transthyretin (down-regulated); and (c) a cleavage fragment of inter- atrypsin inhibitor heavy chain H4 (ITIH4) (up-regulated). In independent validation to detect early stage invasive epithelial ovarian cancer from healthy controls, the sensitivity of a multivariate model combining the three biomarkers and CA125 [74% (95% CI, 52–90%)] was higher than that of CA125 alone [65% (95% CI, 43–84%)] at a matched specificity of 97% (95% CI, 89–100%). When compared at a fixed sensitivity of 83% (95% CI, 61–95%), the specificity of the model [94% (95% CI, 85–98%)] was significantly better than that of CA125 alone [52% (95% CI, 39–65%)]. These biomarkers demonstrated the potential to improve the detection of early stage ovarian cancer.

Among the 3 biomarkers, the cleaved form of ITIH4 is the most interesting. Several proteolytically derived fragments from the proline-rich region (PRR) of human ITIH4 have been identified by surface-enhanced or matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF-MS or MALDI-TOF-MS) as potential disease markers [19]. Therefore, we employed this high-throughput approach to quantify and characterize the extensive fragmentation within the PRR of human serum ITIH4 and determined its association with different disease conditions. The ITIH4-related fragments were first immunocaptured by beads coupled with peptide-specific antibodies. The eluates were analyzed by SELDI-TOF-MS. In addition, freshly collected and immediately processed serum and plasma samples were used to analyze the ex vivo stability of these ITIH4 fragments. Human serum ITIH4 was shown to be extensively proteolytically processed within the PRR, and its fragmentation patterns were closely associated with different disease conditions. Fragmentation patterns were generally consistent with cleavages by endoprotease followed by exoprotease actions. Observed fragments changed little under different assay conditions or blood collection and processing procedures. The fragmentation patterns within the PRR of human serum ITIH4 are associated with different disease conditions and may hold important diagnostic information. These fragmentation patterns could be useful as potential biomarkers for detection and classification of cancer. Enzymes, inactive proteolytic fragments of enzymes or protease inhibitors have become important and promising in the search of cancer biomarkers.

## Conclusion

The future of cancer diagnostics will be based on a panel of proteomic and genomic biomarkers. They could be used to detect cancer at a early stage, predict and direct therapies. Enzymes and related proteins are important biological molecules which could serve as cancer biomarkers. These biomarkers could be intact or fragments of proteins. The challenge is to be able to find and validate these potential biomarkers as clinical diagnostics. With the advances in proteomic technologies, we are closer than ever to find these “new” enzymes or application for the “old” enzymes. The translation of newly discovered biomarkers could revolutionize the new era of personalized medicine.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

## References


1. Chan, DW.; Schwartz, MK. Tumor markers: Introduction and general principles. In: Diamandis; Fritsche; Lilja; Chan; Schwzrtz, editors. Tumor markers: Physiology, patholbiology, technology and clinical applications. AACC Press; 2002. p. 9-17.
2. Chatterjee SK, Zetter BR. Cancer biomarkers: knowing the present and predicting the future. Future Oncol. 2005; 1:37–50. [PubMed: 16555974]
3. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics. 2002; 1:845–867. [PubMed: 12488461] McDonald WH, Yates JR 3rd. Shotgun proteomics and biomarker discovery. Dis Markers. 2002; 18:99–105. [PubMed: 12364816]
4. Colantonio DA, Chan DW. The clinical application of proteomics. Clin Chimica Acta. 2005; 357:151–158.
5. Tirumalai RS, Chan KC, Prieto DA, Issaq HJ, Conrads TP, Veenstra TD. Characterization of the low molecular weight human serum proteome. Mol Cell Proteomics. 2003; 2:1096–1103. [PubMed: 12917320]
6. Rodland KD. Proteomics and cancer diagnosis: the potential of mass spectrometry. Clin Biochem. 2004; 37:579–583. [PubMed: 15234239]
7. Rai AJ, Zhang Z, Rosenzweig J, Shih I-M, Pham T, Fung E, Sokoll LJ, Chan DW. Proteomic approaches to tumor marker discovery: Identification of biomarkers for ovarian cancer. Arch. Pathol. Lab. Med. 2002; 126:1581–1526.
8. Aebersold R. A mass spectrometric journey into protein and proteome research. J Am Soc Mass Spectrom. 2003; 14:685–695. [PubMed: 12837590]
9. Xiao Z, Prieto D, Conrads TP, Veenstra TD, Issaq HJ. Proteomic patterns: their potential for disease diagnosis. Mol Cell Endocrinol. 2005; 230:95–106. [PubMed: 15664456]
10. Zhang Z, Chan DW. Cancer proteomics: in pursuit of "true" biomarker discovery. Cancer Epidemiol Biomarkers Prev. 2005; 14:2283–2286. [PubMed: 16214905]
11. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. Nat Biotechnol. 2006; 24:971–983. [PubMed: 16900146]
12. Semmes OJ, Feng Z, Adam B-L, Banez LL, Bigbee WL, Campos D, Cazares LH, Chan DW, Grizzle WE, Izbicka E, Kagan J, Malik G, McLerran D, Moul JW, Partin A, Prasanna P, Rosenzweig J, Sokoll LJ, Srivastava S, Srivastava Sudhir, Thompson I, Welsh MJ, White N, Winget M, Yasui Y, Zhang Z, Zhu L. Evaluation of serum protein profiling by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry for the detection of prostate cancer: I. Assessment of platform reproducibility. Clin. Chem. 2005; 51:102–112. [PubMed: 15613711]
13. Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations. Mol Cell Proteomics. 2004; 3:367–378. [PubMed: 14990683]
14. Mikolajczyk SD, Millar LS, Wang TJ, et al. A precursor form of prostate-specific antigen is more highly elevated in prostate cancer compared with benign transition zone prostate tissue. Cancer Res. 2000; 60:756–759. [PubMed: 10676664]
15. Chen Z, Chen H, Stamey TA. Prostate specific antigen in benign prostatic hyperplasia: purification and characterization. J. Urol. 1997; 157:2166–2170. [PubMed: 9146608]
16. Mikolajczyk SD, Song Y, Wong JR, Matson RS, Rittenhouse HG. Are multiple markers the future of prostate cancer diagnostics? Clin Biochem. 2004; 37:519–528. [PubMed: 15234233]
17. Sokoll LJ, Chan DW, Mikolajczyk SD, et al. Proenzyme psa for the early detection of prostate cancer in the 2.5–4.0 ng/ml total psa range: preliminary analysis. Urology. 2003; 61:274–276. [PubMed: 12597929]
18. Zhang Z, Bast RC Jr, Yu Y, et al. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. Cancer Res. 2004; 64:5882–5890. [PubMed: 15313933]
19. Song J, Patel M, Rosenzweig CN, et al. Quantification of fragments of human serum inter-alpha-trypsin inhibitor heavy chain 4 by a surface-enhanced laser desorption/ionization-based immunoassay. Clin Chem. 2006; 52:1045–1053. [PubMed: 16574760]