Oncopeptidomics: A Useful Approach for Cancer Diagnosis?

In this issue of *Clinical Chemistry*, Mary Lopez and colleagues *(1)* describe novel methods for isolation of protein-bound peptides from serum and their characterization by mass spectrometry. Lopez et al. used selected peptide combinations to develop a new profiling method for ovarian cancer diagnosis. To put this advance into perspective, I will briefly summarize relevant previous literature on diagnostic applications of serum proteomic and peptidomic profiling by mass spectrometry.

Approximately 5 years ago, a new approach for diagnosing ovarian cancer, by use of SELDI-TOF mass spectrometry, was proposed by the coauthors of the article under discussion *(2)*. It was then hypothesized that proteins or protein fragments released by tumor cells or their microenvironment may enter the general circulation. By the use of a SELDI chip, proteins or peptides could be extracted from crude serum and used for diagnostic purposes with the aid of mass spectrometry and a mathematical algorithm. Similar methods have subsequently been used to diagnose numerous other malignancies, such as breast, prostate, bladder, pancreatic, head and neck, lung, liver, and nasopharyngeal cancers, as well as gliomas and melanomas, with impressive diagnostic sensitivities and specificities. This method has enjoyed ample coverage in scientific journals, the media, and international conferences *(3)*.

This author and others have criticized this diagnostic approach for methodological shortcomings and bioinformatic artifacts *(4–9)*. During the last 5 years, healthy debates have been conducted in journals (including this one) and at conferences *(10, 11)*. An independent validation study would be the best test for this technology. As yet, however, no published validation is available to confirm that these methodologies are working. The proponents of the serum proteomic profiling methods for cancer diagnostics have now turned their attention to another possibility, proposing to interrogate the serum peptidome (the collective peptide population of serum) as a source of putative novel cancer biomarkers *(12)*. Others have also postulated that the serum peptidome has potential for diagnostics *(13)*. Significant literature is now forming around this principle, which is known as peptidomics, or when applied specifically to cancer diagnostics, as oncopeptidomics. In clinical chemistry practice, we already use a number of serum peptides for diagnostics, including insulin and C-peptide for diabetes; parathyroid hormone, calcitonin, and collagen fragments for osteoporosis; probrain-type naturetic peptide for congestive heart failure; progastrin-releasing peptide for small cell lung carcinoma, β-amyloid 1–42 for Alzheimer disease, and angiotensin 2 for hypertension. The theory that peptides carry important diagnostic information seems reasonable because they can act as surrogate endpoints of protein synthesis, processing, and degradation. Interrogation of peptides for diagnostics requires a related array of technologies to be in place: (a) optimized sample collection and processing; (b) peptide extraction, chromatographic separation, and analytical detection; (c) appropriate bioinformatic tools for multiparametric analysis of potentially thousands of informative peptides and methods for comparing differential peptide patterns from healthy individuals and patients with various diseases; and (d) methods for positive peptide identification through sequence determination to derive clues about their possible biological function.

In early reports on peptidomics for diagnostics, Marshall et al. *(14)* claimed that peptides from the sera of healthy individuals and of patients who suffered myocardial infarction can produce MALDI-TOF patterns useful for diagnosis. Lopez and collaborators have previously demonstrated the use of peptidomic analysis for diagnosis of Alzheimer disease and ovarian carcinoma *(15, 16)*. More recently, Villanueva et al. *(17)* used peptidomic analysis to diagnose breast, prostate, and ovarian cancers. The limitations of the latter approach have been summarized recently *(18)*. The method of Villanueva et al. differs from other diagnostic peptidomic methods in that it apparently uses the enzymatic activities of the coagulation and complement cascades for ex vivo generation of informative peptides. This approach should await independent reproduction before any definitive conclusions on its validity are drawn.

Previously *(19)*, I questioned whether the serum peptidome actually exists. Indeed, Koomen et al. *(20)* reported more than 250 peptides, Villanueva et al. *(17)* more than 650 peptides, and Lowenthal et al. *(16)* more than 1200 peptides in serum. It is clear that a large number of peptides exist in serum and plasma and that these can be extracted with reliable techniques *(1)*. A caveat for this large peptide load in serum, however, as shown by Koomen et al. *(20)*, is that most of these peptides, including the high abundance ones, are generated by a surprisingly small number of proteins, owing to proteolytic digestion of high abundance proteins by common enzymes such as thrombin, plasmin, and complement proteins, followed by aminopeptidase and carboxypeptidase processing. We are thus coming down to the critical question on peptidomics for diagnostics: Can peptides originating as products of exoproteolytic and endoproteolytic digestion of high abundance proteins have value as biomarkers of disease? Given that the parent proteins originate mostly from the liver and/or in the process of acute-phase reactions, would these putative biomarkers have the necessary sensitivity and specificity for early detection of cancer and other diseases? Obviously, these questions can be answered only by well-designed, blinded, bias-free and, preferably, prospective investigations in which sample collection and storage are highly standardized and well documented.

Let us now turn our attention to the recent paper by Lopez et al. *(1)*. Their contribution has several distinct merits. The described method for biomarker discovery is

unbiased and is not restricted by any prior hypothesis. It is also a multiparametric approach, which renders itself to multiplexing, and, it is hoped, to better diagnostic sensitivity and specificity. The methods of Lopez et al. for sample preparation and processing are high throughput, allowing large numbers of samples to be analyzed simultaneously. No immunologic or other specific reagents are necessary. The adoption of high-resolution MALDI-Tandem mass spectrometry for sequence determination of identified peptides is a major asset. Lopez et al. developed and used software that allowed identification of discriminating peptides by comparing spectra from normal and cancer patients. Following a strict sample collection protocol to avoid biases as much as possible, the authors identified 162 peptides that might be useful as cancer biomarkers. Their biomarker panels, consisting of >10 peptides per panel, produced sensitivities and specificities of ~90%. These are good numbers compared with CA125, but not yet good enough for screening the general population.

The reported approach is also associated with a number of shortcomings. The sample preparation method isolates peptides that are bound to albumin and other high-abundance proteins. It is almost certain that non–protein-bound peptides with diagnostic value are lost during this process.

The concentration differences identified for peptides used to differentiate between cancer and noncancer patients are relatively small (up to 3.6-fold increases and up to 2.6-fold decreases) compared with the range of concentrations seen in the best classical biomarkers (10- to 1000-fold increases in patients with cancer). Another concern regarding concentrations is that we do not yet have good hypotheses for explaining why a putative biomarker will decrease, rather than increase, in serum of cancer patients (19, 21).

One question requiring further investigation is whether the discriminatory peptides, which originated mainly from high abundance proteins and the coagulation cascade, are specific to cancer patients and thus useful for differentiating them from patients without cancer and whether these peptides are cancer-type specific and can be used to differentiate patients with various types of cancer. As mentioned earlier, these high abundance proteins, usually synthesized by the liver, are likely altered due to generalized effects such as cancer cachexia, malnutrition, and inflammation. These changes are unlikely to be specific for cancer and may represent epiphenomena. For example, more than 40 years ago, it was reported (22) that the coagulation cascade might be defective in cancer patients. In my opinion, it is unlikely that alterations in nonspecific proteins and their fragmentation patterns will produce robust algorithms for early cancer detection. The relatively small changes in concentration seen with these biomarkers suggest that the derived algorithms may not be sufficiently reproducible if tested in an independent series of samples.

In the past, I made specific recommendations for publishing future serum proteomic profiling data for diagnosis (21). More recently, I made similar proposals for peptidomics (19). Here, I wish to reemphasize some points. Because most of the previous proteomic methods (2), as well as genomic methods, for cancer diagnosis and subclassification did not pass validation (23). Future publications, even if describing highly promising data, should be viewed with caution (24) until independent validations are in place. It is fortunate that organizations such as the Early Detection Research Network (EDRN; edrn.nci.nih.gov) and the Ontario Cancer Biomarker Network (OBCN; www.obcn.ca) have shown interest in validating in a blinded fashion promising technologies such as the one reported by Lopez et al. Studies on promising technologies that attract media attention should be published so that laboratorians and clinicians, as well as the public, receive information that otherwise might fail to reach patients because of validation problems.

I will conclude with a cautionary note. The reservations outlined above are not targeting mass spectrometry and its potential in diagnostics (25, 26). The published data so far have created euphoria, as well as confusion. It is our collective responsibility to find out where we stand now, so that we can better plan for the future.

References

1. Lopez MF, Mikulskis A, Kuzdzal S, Golenko E, Petricoin III EF, Liotta LA, et al. "Top-down" discovery and identification of serum carrier protein-bound peptide biomarker candidates in ovarian cancer samples using a novel, high-throughput workflow. Clin Chem 2007;53:1067–74.
2. Petricoin III EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, et al. 2002. Use of proteomic patterns in serum to identify ovarian cancer. Lancet 2002;359:572–7.
3. Check E. Proteomics and cancer: running before we can walk? Nature 2004;429:496–7.
4. Diamandis EP. Analysis of serum proteomic patterns for early cancer diagnosis: drawing attention to potential problems. J Natl Cancer Inst 2004;96:353–6.
5. Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations. Mol Cell Proteomics 2004;3:367–78.
6. Diamandis EP, van der Merwe D-E. Plasma protein profiling by mass spectrometry for cancer diagnosis: opportunities and limitations. Clin Cancer Res 2005;11:963–5.
7. Baggerly KA, Morris JS, Edmonson SR, Coombes KR. Signal in noise: evaluating reported reproducibility of serum proteomic tests for ovarian cancer. J Natl Cancer Inst 2005;97:307–9.
8. Banks RE, Stanley AJ, Cairns DA, Barrett JH, Clarke P, Thompson D, et al. Influences of blood sample processing on low-molecular-weight proteome identified by surface-enhanced laser desorption/ionization mass spectrometry. Clin Chem 2005;51:1637–49.
9. Karsan A, Eigl, BJ, Flibotte S, Gelmon K, Switzer P, Hassell P, et al. Analytical and preanalytical biases in serum proteomic pattern analysis for breast cancer diagnosis. Clin Chem 2005;51:1525–8.
10. Diamandis EP. Point: Proteomic patterns in biological fluids: do they represent the future of cancer diagnostics? Clin Chem 2003;49:1272–8.
11. Petricoin E 3rd, Liotta LA. Counterpoint: The vision of a new diagnostic paradigm. Clin Chem 2003;49:1276–8.
12. Liotta LA, Ferrari M, Petricoin E. Clinical proteomics: written in blood. Nature 2003;425:905.
13. Schulte I, Tammen H, Selle H, Schulz-Knappe P. Peptides in body fluids and tissues as markers of disease. Expert Rev Mol Diagn 2005;5:145–57.
14. Marshall J, Kupchak P, Zhu W, Yantha J, Vrees T, Furesz S, et al. Processing of serum proteins underlies the mass spectral fingerprinting of myocardial infarction. J Proteome Res 2003;2:361–72.
15. Lopez MF, Mikulskis A, Kuzdzal S, Bennett DA, Kelly J, Golenko E, et al. High-resolution serum proteomic profiling of Alzheimer disease samples

reveals disease-specific, carrier-protein-bound mass signatures. Clin Chem 2005;51:1946–54.

16. Lowenthal MS, Mehta AI, Frogale K, Bandle RW, Araujo RP, Hood BL, et al. Analysis of albumin-associated peptides and proteins from ovarian cancer patients. Clin Chem 2005;51:1933–45.

17. Villanueva J, Shaffer DR, Philip J, Chaparro CA, Erdjument-Bromage H, Olshen AB, et al. Differential exoprotease activities confer tumor-specific serum peptidome patterns. J Clin Invest 2006;116:271–84.

18. Diamandis EP, Kulasingam V, Sardana G. Letter to the Editor about differential exprotease activities confer tumor-specific serum peptidome. J Clin Invest 2006;464 http://www.jci.org/cgi/eletters/116/1/271 (accessed December 2006).

19. Diamandis EP. Peptidomics for cancer diagnostics: present and future. J Proteome Res 2006;5:2079–82.

20. Koomen JM, Li D, Xiao L-C, Liu TC, Coombes KR, Abbruzzese J, et al. Direct tandem mass spectrometry reveals limitations in protein profiling experiments for plasma biomarker discovery. J Proteome Res 2005;4:972–81.

21. Diamandis EP. Serum proteomic profiling by matrix-assisted laser desorption-ionization time-of-flight mass spectrometry for cancer diagnosis: Next steps. Cancer Res 2006;66:5540–1.

22. S. Carlsson, Fibrinogen degradation products in serum from patients with cancer, Acta Chir Scand 1973;139:499–502.

23. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. Lancet 2005;365:488–92.

24. Diamandis EP. Quality of the scientific literature: all that glitters is not gold. Clin Biochem 2006;39:1109–11.

25. Hortin GL, Jortani SA, Ritchie JC Jr, Valdes R Jr, Chan DW. Proteomics: a new diagnostic frontier. Clin Chem 2006;52:1218–22.

26. Hortin GL. The MALDI-TOF mass spectrometric view of the plasma proteome and peptidome. Clin Chem 2006;52:1223–37.

Eleftherios P. Diamandis

*Department of Pathology
and Laboratory Medicine
Mount Sinai Hospital
600 University Avenue
Toronto, ONT M5G 1X5 Canada*

*Department of Laboratory Medicine
and Pathobiology
University of Toronto
Toronto, ONT
Canada
E-mail ediamandis@mtsinai.on.ca*