Use of ProteinChip™ Array Surface Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (SELDI-TOF MS) to Identify Thymosin β -4, A Differentially Secreted Protein from Lymphoblastoid Cell Lines

Deborah L. Diamond

Ciphergen Biosystems Inc., Fremont, California, USA

Yanni Zhang, Alexander Gaiger, Molly Smithgall, Thomas S. Vedvick, and Darrick Carter

Department of Antigen Discovery, Corixa Corporation, Seattle, Washington, USA

The identification of proteins differentially expressed between cancer and normal cells is vital for the development of cancer diagnostics, therapeutics and vaccines. Using a ProteinChip Biomarker System (Ciphergen Biosystems, Fremont, CA) which combines ProteinChip™ technology with time-of-flight mass spectrometry, we have developed a simple method to screen and identify differentially secreted proteins from tumor cell lines. Mass spectra of the range of proteins secreted from normal B-cells were generated along with those secreted from Epstein-Barr virus transformed B-cells. A mass peak at $m/z = 4972.1$ that was highly over-represented in the transformed B-cell line was chosen for identification and purified by reversed phase chromatography with concomitant monitoring of fractions by SELDI-TOF MS. The resulting purified protein was digested with trypsin and the peptide masses derived from the SELDI-TOF spectrum were used to search the public databases for protein identification. Fragment matching of the resulting peptides identified the protein as thymosin β -4. Using LC-electrospray ionization MS/MS, the identity of this protein was confirmed. Thymosin β -4 is a known marker in LCLs establishing the utility of this method to discover and identify proteins differentially expressed between cancers and their matched normal counterparts. (J Am Soc Mass Spectrom 2003, 14, 760–765) © 2003 American Society for Mass Spectrometry

Modern cancer therapy and diagnosis relies on the availability of targets unique to tumor cells. Development of novel therapeutics and diagnostic agents is greatly facilitated by the discovery of proteins that are differentially expressed between cancerous and normal cells. These proteins can then be used as targets for the development of small molecule inhibitors, ELISA-based diagnostics, and functional monoclonal antibodies that may block vital pathways, deliver a toxic compound to the tumor, or cause apoptosis by stimulating a tumor-specific surface molecule. Such rationally designed therapies are far less likely to have the broad toxicities associated with current che-

motherapy and may replace standard chemotherapy in the future. Indeed, some of these targeted cancer treatments are already on the market or in advanced clinical trials and are showing great promise in efficacy and reduced toxicity [1].

A major hurdle for developing new drugs of this type is finding a target that is specific for the tumor. Advances in the use of DNA biochips and real time PCR have accelerated the discovery of new candidate markers, but are not optimal for a number of reasons: increased transcription of a gene does not necessarily correlate with higher protein levels, differentially modified targets will be missed by this approach, and it is difficult to select a sub-cellular localization to focus discovery efforts on. Using combinations of mass spectrometry and 2-D gels different groups have been successful in directly identifying differentially expressed proteins between various samples [2, 3].

Published online May 29, 2003

Address reprint requests to Dr. D. Carter, Department of Antigen Discovery, Corixa Corp., 1124 Columbia St., Suite 200, Seattle, WA 98104. E-mail: carter@corixa.com

A simplification of this type of method may be the use of Ciphergen's ProteinChip™ Array technology which utilizes the surface enhanced laser desorption/ionization (SELDI) process to integrate protein capture, purification, characterization and quantitative protein interaction assays on a single biochip based platform [4]. Briefly, ProteinChip™ Arrays contain multiple spots derivatized with chemical or biological "docking sites" which are used for the selective capture of proteins and peptides from complex mixtures. After the sample is applied to the surface, unbound proteins and interfering substances (i.e., salts, detergents, urea) are washed away, an energy-absorbing molecule (EAM), or matrix, is added and the array is analyzed by laser desorption ionization time-of-flight mass spectrometry. Prior to spotting, a biochemical fractionation can be used to select proteins of a specific type such as membrane or secreted proteins. The use of this ProteinChip™ Biomarker System to profile differential protein expression has been described for other cancers such as prostate, breast, and ovarian cancer [5–8]. In this report we combine high pressure liquid chromatography (HPLC), SELDI, and electrospray ionization-tandem MS to screen, isolate, and identify differentially secreted targets from a model tumor lymphoblastoid cell line (LCL).

Materials and Methods

Chemicals and Reagents

Tris[hydroxymethyl]aminomethane (Tris), sodium chloride, ammonium acetate, phosphate buffered saline (PBS), glycine, methanol, silver nitrate, and sodium thiosulfate were obtained from Sigma (St. Louis, MO). Acetonitrile was from Mallinckrodt Laboratory Chemicals (Phillipsburg, NJ). All other reagents were of the highest quality commercially available.

Enrichment of CD19⁺ Cells from PBMC

Peripheral blood mononuclear cells (PBMC) were enriched for CD19⁺ cells by depletion of T lymphocytes, monocytes, and NK cells. PBMC from two donors were pooled and depleted of adherent cells by incubation for 2 h at 37 °C in tissue culture flasks. The non-adherent population was collected in a 15 mL conical tube and incubated with monoclonal antibodies (mAb) against CD4, CD8, CD14, CD16, and CD56 (Pharmingen, San Diego, CA) at 1 µg mAb/10⁶ cells in PBS for 1 h on ice. The cells were then collected by centrifugation, washed in PBS to remove unbound Ab, and resuspended in PBS at 10 × 10⁶ cells/mL. Goat anti-mouse magnetic IgG Dynabeads (DynaL, Norway), washed according to the manufacturer's recommendation, were then added at 10⁷ beads/mL of cells, and the mixture was incubated for 30 min at room temperature while rotating. The negatively selected cells were then removed from the cell suspension by placing the tube in a Dynal Magnetic

Particle Concentrator (MPC), and the supernatant containing the CD19⁺ enriched cell population was collected.

Immortalization of B lymphocytes by Epstein-Barr Virus (EBV)

10 × 10⁶ PBMC, from the same donors that were used to enrich for CD19⁺ cells, were pelleted and resuspended in 9 mL of culture supernatant from the marmoset cell line B95-8, which is EBV transformed and releases high titers of transforming EBV. An additional 9 mL of Roswell Park Memorial Institute (RPMI) medium (Gibco Life Technologies, Gaithersburg, MD) containing 10% FBS (HyClone, Logan, UT), plus 1 µg/mL cyclosporin A (Sandoz, Basel, Switzerland) was added to the culture. The cells were cultured in a T-25 tissue culture flask, standing upright. After approximately three weeks in culture, the B-lymphocytes will be immortalized to an LCL, which was expanded conservatively in RPMI/10% FBS.

SELDI Analysis of Secreted Proteins

Ten µL of LCL and B-cell supernatants were spotted onto a reversed phase H4 (hydrophobic surface) ProteinChip™ Array in 5% acetonitrile, 0.05% trifluoroacetic acid, and incubated in a humid environment for 20 min at room temperature. A 10 µL wash with 5% acetonitrile, 0.05% trifluoroacetic acid was performed. After drying, 0.5 µL of a saturated solution of sinapinic acid in 50% acetonitrile(vol/vol), 0.5% trifluoroacetic acid (vol/vol) was applied to the spots and allowed to dry prior to mass analysis in a ProteinChip™ Reader (Ciphergen Biosystems). Calibration was performed in parallel using protein standards spanning molecular masses from 8564.8 Da (Ubiquitin) to 147,300 Da (Bovine IgG).

SELDI Analysis of RP-HPLC Fractions

Peak fractions from reversed phase separation of LCL and B cell supernatants were spotted directly on an H4 ProteinChip™ Array surface in order to monitor purification of the $m/z = 4972$ species. 1 µL of diluted sample was put onto the array and allowed to dry. Under other conditions in which buffers incompatible with subsequent mass analysis (i.e., detergents, urea, salts) are present, the ProteinChip™ Array is washed to remove interfering substances. 0.5 µL of saturated α -cyano-4-hydroxy cinnamic acid in 0.5 % trifluoroacetic acid, 50 % acetonitrile and 49.5% water was then added to each spot and allowed to dry prior to mass analysis as described above.

Protein Purification

One hundred µL of secreted protein supernatant (see above) was loaded onto a Supelco (Bellefonte, PA) Discovery cyano reversed phase column (10 cm × 3

Figure 1. SELDI mass spectrum for differentially expressed molecules. B-cells (a) and EBV-transformed B-cells (b) were spun down, washed three times in PBS and incubated in PBS for 1 h at room temperature. 2 μ L of each sample was spotted onto the array and treated as described in Materials and Methods. The m/z range from 0 to 22,000 is shown here. Strong signal for a molecule at m/z = 4972 was present in EBV-transformed B-cells. This peak (labeled with an asterisk) was selected for further analysis. The peak is seen at a slightly higher molecular mass than in the following experiments due to the fact that standardization was performed here for a wide range of masses from 8 to 147 kDa, but is still an acceptable deviation from the expected mass (0.14%).

mm, 5 μ m). Unbound compounds were washed out in buffer A (water, 0.05% TFA) and proteins were eluted with a linear gradient to 100% buffer B (acetonitrile, 0.05% TFA) over 17 column volumes. Peak fractions were dried onto an H4 ProteinChipTM Array and analyzed by SELDI-MS as described above. Purified fractions containing the mass of interest were pooled and used for further study.

Trypsin Digestion and Analysis by SELDI

Twenty μ L of purified protein were heated for 1 min at 100 °C. Protein was rapidly cooled on ice and condensate spun down. The pH was brought to 7.5 with 1M Tris pH = 10.0 and sequencing grade modified porcine trypsin (Promega, Madison, WI) was added to 0.05 μ g/mL. The protein was digested for 1h and the reaction stopped by spotting and drying onto an H4 ProteinChipTM Array. 0.5 μ L of alpha-cyano-4-hydroxy cinnamic acid in 50% acetonitrile (vol/vol), 0.5% trifluoroacetic acid (vol/vol) was added and mass analysis performed. External Calibration was performed using a mixture of peptide standards in the same matrix on a different spot of the array. Three molecular mass standards were used (parentheses indicate predicted molecular masses): Dynorphin A (2147.5 Da), Glu1-Fibrinopeptide B (1570.6 Da), and human ACTH (2933.5 Da).

Capillary LC-Electrospray Ionization-Tandem Mass Spectrometry (LC-ESI-MS/MS)

A capillary column (100 μ m i.d., 12 cm long) was filled with C18 resin. Peptides were bound to this column and eluted by a gradient from 5 to 65% B over 20 min (A,

0.2% acetic acid in water; B, 80% acetonitrile in A). Eluted peptides were introduced into the LC-Q ion-trap mass spectrometer (Finnigan, CA) by electrospray via an electrospray ionization interface (Cytopeia, Seattle, WA) and analyzed by data dependent MS and MS/MS scans. The collision induced dissociation spectra generated during the experiment were searched against a human protein database using the SEQUEST program of the Xcalibur software package version 1.2 [9] to identify possible sequence matches. A second database was created using possible variants of thymosin β -4 following initial identification to account for removal of the N-terminal methionine and acetylation.

Results and Discussion

After incubation of the cells with PBS, cell supernatants were analyzed on a variety of ProteinChipTM Array chemical surfaces to characterize protein expression patterns. The use of on-spot chromatography, binding the proteins to the reversed phase matrix on the ProteinChipTM Array, and washing off salts and other impurities allowed for the quick analysis of crude secreted material. The SELDI mass spectrum revealed different protein profiles of the two cell samples chosen (Figure 1). The major differences between the two samples are tabulated in Table 1. Due to the fact that the isolated B cell population from patients did not completely consist of B-cells, proteins that are seen only in the patient B-cell sample and not the LCL sample cannot be attributed to down-regulation of these proteins in LCL. Therefore, while it is possible that the major peaks seen in the B-cell isolate (Table 1) are really down regulated in LCL it is equally possible that these

Table 1. Major peaks with different intensities between LCL and the B-cell population

Higher in LCL ^a	Higher in B-cell
4972.1	—
9964.6	—
—	6896.2
—	10427.9
—	10841.9
—	13740.4
—	14008.3

^aExamination of the spectra from Figure 1 reveals major differences between LCL and the normal B-cell population. Peaks that are much higher in the B-cell population could be derived from impurities in the isolated native B-cells. Thus, the ones that showed higher levels in the LCL were examined more closely. Of these, the peak at m/z of 4972.1 was the most striking difference.

are artifacts derived from isolating native cells. However, proteins that are seen in the LCL and not in the B-cell population are LCL unique as this is a B-cell line free of other cell types. Thus, we set out to validate our method by choosing one of the LCL specific molecules and looking to see if it has a known implication in EBV-transformed B-cells.

The most prominent peak identified was the species at $m/z = 4972.1$ indicated in Figure 1. We wanted to use multiple methods to determine the identity of this peak: LC-ESI-MS/MS and SELDI analysis of tryptic fragments followed by database searching of the tryptic fingerprint. However, the unpurified LCL supernatant was deemed too crude for further analysis and was subjected to reversed phase HPLC to enrich and purify

the target peak. Fractions were monitored by SELDI-TOF MS on material dried directly onto a ProteinChip™ Array after elution. The mass of interest was identified in a peak eluting early off of the reversed phase column (Figure 2). This sample was relatively free of other detected masses and used for subsequent analysis.

A more exact mass of the protein was determined by standardizing the ProteinChip™ Reader using smaller peptides rather than the large range of peptide and protein standards used in the initial screen. This led to the determination that there are two dominant peaks in the purified fraction with $m/z = 4962.3$ and $m/z = 4981.8$ (Figure 3a). Subsequent analysis (below) revealed that the protein is thymosin β -4 and exists in a major oxidized state. The predicted masses for the full protein and its oxidized counterpart are 4964.5 and 4980.5 respectively, in good agreement with the determined m/z values.

Following characterization of the parent protein we wanted to use trypsin digestion to identify the protein by its mass fingerprint (Figure 3b). The protein in the HPLC fraction was digested with trypsin and subjected to SELDI-TOF MS and m/z values tabulated. The determined m/z after digestion are indicated in Table 2, and were used to search a database for matches using the MS-FIT algorithm [10]. As the protein was not digested under denaturing conditions and was trypsinized for a relatively short time we allowed for a number of missed cleavages in the MS-FIT menu. A match for thymosin β was found in the results of the initial search. This was

Figure 2. Reversed Phase Chromatography of cell supernatant. 100 μ L of LCL supernatant in PBS were injected onto a SUPELCO Cyano reversed phase column and eluted with a water 0.5% TFA to acetonitrile 0.5% TFA gradient. Fractions were collected and analyzed on an H4 ProteinChip™ Array for signal at $m/z = 4972$. The peak fraction containing highly purified peptide (indicated by an asterisk) was used for all further analyses.

Figure 3. (a) SELDI Mass Spectrum of Purified Protein. Purified protein from the reversed phase chromatography was diluted 1:10 and spotted as described. Determined masses (*m/z*): ThyB: 4962.3, ThyB-Ox: 4981.8 (resolution, *t*/2*dt* = 532.7); predicted masses (Da): ThyB: 4964.5; ThyB-Ox: 4980.5. In secreted forms of thymosin β -4 it has been observed that there is a significant portion of oxidized protein present. (b) SELDI Mass Spectrum of tryptic digest of the purified protein. 40 ng of trypsin were added to 100 μ L of purified protein in ammonium bicarbonate and digested for 1 h at 37 $^{\circ}$ C. A control of trypsin alone was run in parallel. Peaks unique to the ThyB digest are marked with capital letters and are tabulated in Table 2.

Table 2. Predicted and determined *m/z* values for thymosin β -4 tryptic digest

	<i>m/z</i> Determined ^a	<i>m/z</i> Predicted ^b	Error (%)	Amino acids in thymosin β -4 ^c	Comments
A	1695.9	1695.9	0.00	1–14 (acetyl)	N-terminus acetylated
B	1711.4	1711.0	0.023	1–14 (acetyl + Ox)	Oxidized Met
C	1985.9	1986.2	0.015	26–43	
*	—	—	—	—	Trypsin autolysis
D	2230.8	2230.4	0.018	20–38	
*	—	—	—	—	Trypsin autolysis
E	2358.9	2358.6	0.013	19–38	
*	—	—	—	—	Trypsin autolysis
F	2600.5	2600.0	0.019	17–38	
G	2703.2	2702.9	0.011	20–43	
H	2831.6	2831.1	0.018	19–43	
I	3072.7	3072.4	0.010	17–43	

^aTabulated are *m/z* values from the spectrum in Figure 3b.
^bPredicted masses are those derived from a MS-Digest [10] run on the published sequence of thymosin β -4 while accounting for the N-terminal acetylation.
^cPeptide fingerprinting searches with the non-tryptic fragments using the MS-FIT program [10] pull out human thymosin β -4 as atop hit matching all input fragments.

Table 3. Peptides determined by LC-ESI-MS/MS

Amino acids in thymosin β -4 ^a	Matching sequence for MS/MS	Corresponding peptide from SELDI Analysis
1-11 (acetyl)	*SDKPDMAEIEK	–
1-14 (acetyl)	*SDKPDMAEIEKFDK	A
20-38	TETQEKNPLPSKETIEQEK	D
26-38	NPLPSKETIEQEK	–
19-43	KTETQEKNPLPSKETIEQEKQAGES-	H
20-43	TETQEKNPLPSKETIEQEKQAGES-	G
Known sequence 1-43	*SDKPDMAEIEKFDKSKLKKTTETQEKNPLPSKETIEQEKQAGES^b	

^aHighly significant hits from SEQUEST search of collision induced dissociation fragmentation of LC-ESI-MS tryptic digest of peak fraction. Shown in bold is the published sequence of thymosin β -4. Identified peptides: 1-14 (as acetyl derivative), 19-38, 19-43 (C-terminus), 20-38, 26-38, 20-43 (C-terminus), 1-11 (as acetyl derivative), 20-31. Total coverage: 88%

^b*indicates N-terminal acetylation, -indicates the C-terminus.

the only protein which had an appropriate parent mass and good matches for the generated peptides. After allowing for N-terminal modifications [11] and methionine oxidation [12] thymosin β -4 was the top protein pulled out by peptide fingerprinting.

To further verify the identity of the protein as thymosin β -4, LC-ESI-MS/MS was performed on the purified and digested protein fraction. Numerous peptides derived from thymosin β -4 were identified and gave coverage of almost 90% after allowing for N-terminal acetylation (Table 3). The biochemical profile of the material isolated from LCL secretions is one of a protein without its initiating methionine that is modified by acetylation of the N-terminal serine. The methionyl residue at position 6 is in a mixed state partially oxidized to its sulfoxide equivalent. These findings are in good agreement with the known properties of thymosin β -4: it is acetylated at the N-terminal serine and exists in a mixed form between thymosin β -4 and thymosin β -4 sulfoxide when secreted (see also Figure 3a).

Intracellular thymosin β -4 is thought to be a G-actin sequestering molecule [13], while when secreted [14], especially as the thymosin β -4 sulfoxide, the molecule appears to act as a signaling molecule functioning in a number of ways including wound healing [11], inflammation [12], and hematopoiesis [15]. Previous studies on cell lines have shown that thymosin β -4 is overexpressed on the mRNA and protein level in various cell lines [16]. LCL had the highest amount of thymosin β -4 in these studies. Although the function of thymosin β -4 in cancer has not been completely elucidated, it is known that numerous tumor types overexpress this protein possibly for its involvement in apoptosis [17] or angiogenesis [11]. The results of this study show that the use of SELDI technology on relatively crude samples of tumor derived material can lead to the rapid identification of tumor specific markers when a matched normal control sample is present. The Protein-Chip™ technology can be combined with other biochemical and mass spectrometric methods to purify and accurately characterize a target within a short period of time. Further studies combining these techniques

should lead to the identification and characterization of novel markers to be used as targets for rational design of tumor-specific therapies.

References

- Kaminski, M. S.; Zelenetz, A. D.; Press, O. W.; Saleh, M.; Leonard, J.; Fehrenbacher, L.; Lister, T. A.; Stagg, R. J.; Tidmarsh, G. F.; Kroll, S.; Wahl, R. L.; Knox, S. J.; Vose, J. M. *J. Clin. Oncol.* **2001**, 19(19), 3918–3928.
- Smith, M. A.; Bains, S. K.; Betts, J. C.; Choy, E. H.; Zanders, E. D. *Clin. Diagnost. Lab. Immunol.* **2001**, 8(1), 105–111.
- Hanash, S. M.; Strahler, J. R.; Neel, J. V.; Hailat, N.; Melhem, R.; Keim, D.; Zhu, X. X.; Wagner, D.; Gage, D. A.; Watson, J. T. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, 88(13), 5709–5713.
- von Eggeling, F.; Junker, K.; Fiedle, W.; Wollscheid, V.; Durst, M.; Claussen, U.; Ernst, G. *Electrophoresis* **2001**, 22(14), 2898–2902.
- Adam, B. L.; Qu, Y.; Davis, J. W.; Ward, M. D.; Clements, M. A.; Cazares, L. H.; Semmes, O. J.; Schellhammer, P. F.; Yasui, Y.; Feng, Z.; Wright, G. L., Jr. *Cancer Res* **2002**, 62(13), 3609–3614.
- Li, J.; Zhang, Z.; Rosenzweig, J.; Wang, Y. Y.; Chan, D. W. *Clin. Chem.* **2002**, 48(8), 1296–1304.
- Rai, A. J.; Zhang, Z.; Rosenzweig, J.; Shih, I.; Pham, T.; Fung, E. T.; Sokoll, L. J.; Chan, D. W. *Arch. Pathol. Lab. Med.* **2002**, 126(12), 1518–1526.
- Paweletz, C. P.; Trock, B.; Pennanen, M.; Tsangaris, T.; Magrant, C.; Liotta, L. A.; Petricoin, E. F., III. *Dis. Markers* **2001**, 17(4), 301–307.
- Gygi, S. P.; Corthals, G. L.; Zhang, Y.; Rochon, Y.; Aebersold, R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, 97(17), 9390–9395.
- Baker P. R.; Clauser K. R., MS-FIT 2001 Computer Program.
- Huff, T.; Muller, C. S.; Otto, A. M.; Netzker, R.; Hannappel, E. *Int. J. Biochem. Cell Biol.* **2001**, 33(3), 205–220.
- Young, J. D.; Lawrence, A. J.; MacLean, A. G.; Leung, B. P.; McInnes, I. B.; Canas, B.; Pappin, D. J.; Stevenson, R. D. *Nat. Med.* **1999**, 5(12), 1424–1427.
- Carlier, M. F.; Jean, C.; Rieger, K. J.; Lenfant, M.; Pantaloni, D. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, 90(11), 5034–5038.
- Huang, W. Q.; Wang, Q. R. *Exp. Hematol.* **2001**, 29(1), 12–18.
- Li, W. M.; Huang, W. Q.; Huang, Y. H.; Jiang, D. Z.; Wang, Q. R. *Cytokine* **2000**, 12(7), 1017–1023.
- Hannappel, E.; Leibold, W. *Arch. Biochem. Biophys.* **1985**, 240(1), 236–241.
- Iguchi, K.; Usami, Y.; Hirano, K.; Hamatake, M.; Shibata, M.; Ishida, R. *Biochem. Pharmacol.* **1999**, 57(10), 1105–1111.