BIA/MS of Epitope-Tagged Peptides Directly from E. Coli Lysate: Multiplex Detection and Protein Identification at Low-Femtomole to Subfemtomole Levels

ARTICLE *in* ANALYTICAL CHEMISTRY · AUGUST 1999

Impact Factor: 5.64 · DOI: 10.1021/ac990089v · Source: PubMed

4 AUTHORS, INCLUDING:

Jonathan W Jarvik
Carnegie Mellon University
57 PUBLICATIONS **1,895** CITATIONS
SEE PROFILE

Kemmons Tubbs
University of South Alabama
36 PUBLICATIONS **1,279** CITATIONS
SEE PROFILE

Anal. Chem. **1999,** *71,* 2858−2865

BIA/MS of Epitope-Tagged Peptides Directly from *E. Coli* Lysate: Multiplex Detection and Protein Identification at Low-Femtomole to Subfemtomole Levels

Randall W. Nelson,*,† Jonathan W. Jarvik,‡ Bruce E. Taillon,‡ and Kemmons A. Tubbs†

Intrinsic Bioprobes, Inc., 2009 East 5th Street, Ste. 11, Tempe, Arizona 85281, and Department of Biological Sciences, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213

The use of biomolecular interaction analysis mass spectrometry to selectively isolate, detect, and characterize epitope-tagged peptides present in total cell lysates is demonstrated. Epitope-tagged tryptic peptides were captured via affinity interactions with either chelated Ni^{2+} or monoclonal antibodies and detected using surface plasmon resonance biomolecular interaction analysis (SPR-BIA). After SPR-BIA the tagged peptides were either eluted from the biosensor chips for mass spectrometric analysis or analyzed directly from the biosensor chip using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF). Protein database searches were performed using the masses of the tagged tryptic peptides, resulting in identification of the protein into which the epitope tag was inserted. Detection limits for both SPR-BIA and MALDI-TOF were at the low-femtomole to subfemtomole level. The approach represents a (multiplexed) high-sensitivity chip-based technique capable of identifying epitope-tagged proteins as they are present in complex mixtures.

A fundamental and recurring goal in cell biology and biochemistry is the detection and analysis of individual protein species against a complex background of numerous other molecular species present in the cell, tissue, or biological fluid. The process generally requires the isolation/purification of the protein of interest followed by appropriate characterization. An approach is of particular value if it is highly sensitive, is able to sense the presence of an analyte with high selectivity, and provides defining structural information about the protein. To meet these stringent criteria, high-performance analytical approaches are often multiplexed, using more than one instrument for a single analysis.

Two instrumental techniques that are capable of being interfaced, surface plasmon resonance biomolecular interaction analysis (SPR-BIA)[1−4] and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF),[5−8] have emerged in the past decade to find frequent use in bioanalytical laboratories. The two techniques are viewed as highly complementary in that each is performed for a different analytical purpose. SPR-BIA is used to investigate interactions between a surface-immobilized receptor and solution-borne ligand, in real time, yielding information about the association/dissociation kinetics involved in the interaction, information potentially equating to dissociation constants. Put to different use, SPR-BIA can be used as a biosensor to quantitatively detect the presence of specific molecules. MALDI-TOF has found much application in a variety of analyses where an accurate molecular weight determination is used to assist in the structural characterization of a biomolecule. A number of applications exist ranging from confirmation of protein sequence through an accurate molecular weight determination, to the identification of proteins via sequence database search.[9] Alternate uses of MALDI-TOF extend to detection and quantitation of target analytes (and variants) selectively retrieved from biological carriers via immunoaffinity isolation.[10,11]

When combined to form a concerted analysis, SPR-BIA and MALDI-TOF are highly compatible. Because the optical detection of SPR is nondestructive, a secondary analysis of species retained during SPR-BIA is possible using MALDI-TOF. The analysis, biomolecular interaction analysis mass spectrometry (BIA/MS), is capable of distinguishing the number and nature (specific or nonspecific) of species bound to a sensor chip and the relative contribution of each of the species to the composite binding curve. Furthermore, unknown species retained through specific interaction with the immobilized receptor can be characterized to the point of identification using mass spectrometric data in combination with protein/genome sequence database searches. In all, an analytical approach capable of the real-time optical viewing of binding events, determination of the number of participants in

† Intrinsic Bioprobes, Inc.

‡ Carnegie Mellon University.

(1) Malmqvist, M. *Nature* **1993**, *361*, 186−187.
(2) Szabo, A.; Stolz, L.; Granzow, R. *Curr. Opin. Struct. Biol.* **1995**, *5*, 699−705.
(3) Myszka, D. G. *Curr. Opin. Biotechnol.* **1997**, *8*, 50−57.
(4) Malmqvist, M.; Karlsson, R. *Curr. Opin. Chem. Biol.* **1997**, *3*, 378−383.
(5) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299−2301.
(6) Hillenkamp, F.; Karas, M.; Beavis, R. C.; Chait, B. T. *Anal. Chem.* **1991**, *63*, 1193A−1203A.
(7) Mann, M.; Talbo, G. *Curr. Opin. Biotechnol.* **1996**, *7*, 11−19.
(8) Fenselau, C. *Anal. Chem.* **1997**, *9*, 661A−665A.
(9) O'Connell K. L.; Stults J. T. *Electrophoresis* **1997**, *18*, 349−359.
(10) Nelson, R. W.; Krone, J. R.; Bieber, A. L.; Williams, P. *Anal. Chem.* **1995**, *67*, 1153−1158.
(11) Wang, R.; Sweeney, D.; Gandy, S. E.; Sisodia, S. S. *J. Biol. Chem.* **1996**, *271*, 31894−31902

10.1021/ac990089v CCC: $18.00
Published on Web 05/25/1999

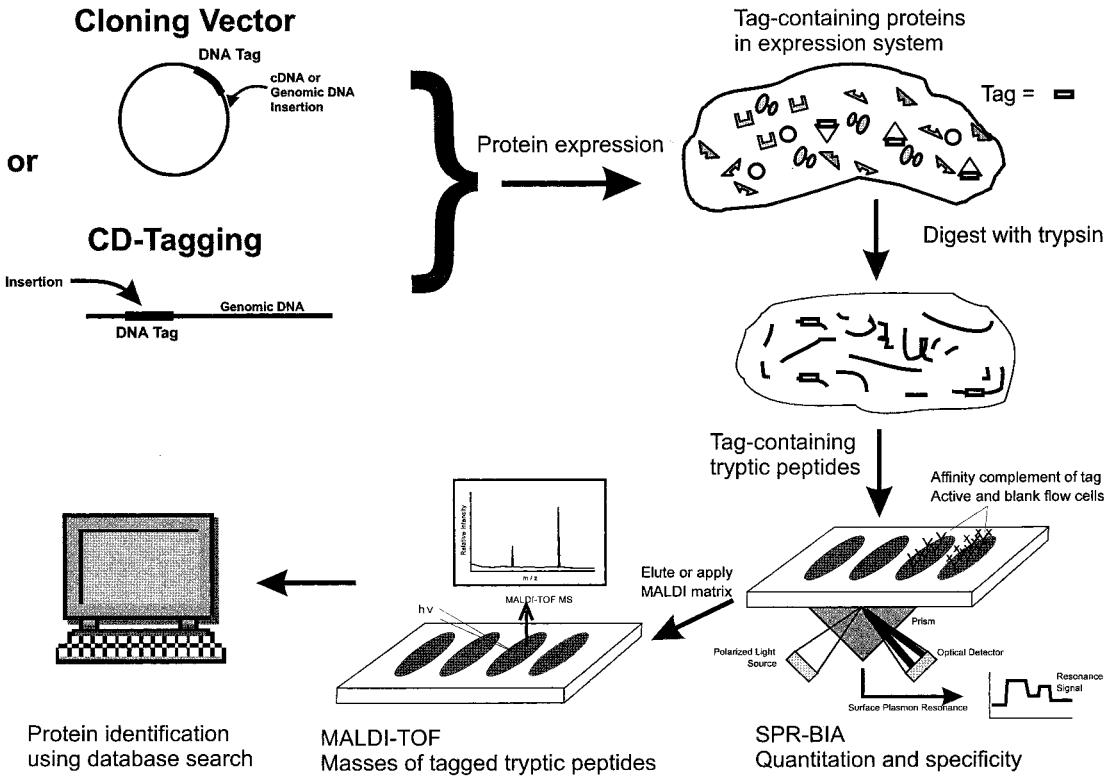

Figure 1. Overview of experimental approach to BIA/MS analysis of epitope-tagged proteins present in complex biological mixtures. Tags are fused or inserted into genes using cloning vectors or insertion vehicles such as CD-tagging.[18] After expression, the entire biological system is digested with trypsin resulting in epitope-tagged peptides. SPR-BIA is used to selectively isolate, detect, and quantitate the peptides from the digested media. MALDI-TOF mass spectrometry is used to characterize the tagged peptides and identify the protein resulting from the gene.

the biomolecular interaction, and mass spectral characterization of interacting species is enabled by linking SPR-BIA with MALDI-TOF. Accordingly, we have devoted much effort to interfacing SPR-BIA with MALDI-TOF for the express purpose of multiplexed biomolecular characterization.[12−17]

The applicability of BIA/MS can be increased even further when used in conjunction with gene-tagging techniques. In this regard, BIA/MS becomes a general technique to isolate, detect, and characterize tag-containing biomolecules. Moreover, BIA/MS would suitably complement cloning vector epitope-tagging techniques and gene discovery vehicles, such as CD tagging,[18] by providing a means of readily retrieving and analyzing gene products postexpression. A general overview of the approach is shown in Figure 1. Using various tagging approaches, tags are fused or inserted into nominally unknown genes for the purpose of tracking biomolecules throughout expression. Regarding the resulting protein, an affinity complement to the epitope tag may be used to selectively isolate the expressed protein or the tag-containing proteolytic fragments from the expression system. SPR-BIA is used during the affinity isolation to differentiate specific from nonspecific binding and to accurately quantify the amount of tagged polypeptide(s) retrieved from the expression system. MALDI-TOF is used after SPR-BIA to accurately determine the masses of the tagged polypeptides, and these mass values are then used to fuel a database search capable of identifying the tagged protein.

Presented here are pilot studies investigating the feasibility of using BIA/MS to detect and characterize small quantities of epitope-tagged polypeptides present in complex biological mixtures. Objectives of the studies were as follows.

(1) To process tagged proteins directly from complex mixtures in a manner favorable for gene/protein identification without severe losses in analytical specificity or sensitivity: The intention here was to develop means of sample preparation and analysis capable of detecting the presence of interacting species that are favorable for obtaining mass spectrometric data sufficient for protein identification via sequence database search. Since database search does not require that information on the protein originates from the intact protein, proteolytic fragments of the protein (containing the epitope tag) were the target of the analysis. These fragments were generated by trypsin digestion of unfractionated *Escherichia coli* whole lysate containing the tagged protein, and BIA/MS was used to analyze the tagged proteolytic fragments present in the digest mixture.

(2) To compare two different forms of BIA/MS in the analysis of epitope-tagged peptides: The first approach utilized a method in which retained peptides were simply eluted from the surface

(12) Krone, J. R.; Nelson, R. W.; Dogruel, D.; Williams, P.; Granzow, R. *Anal. Biochem.* **1997**, *244*, 124−132.
(13) Nelson, R. W.; Krone, J. R. *J. Mol. Recognit.* **1999**, *12*, 77−93.
(14) Nelson, R. W.; Krone, J. R.; Jansson, O. *Anal. Chem.* **1997**, *69*, 4363−4368.
(15) Nelson, R. W.; Krone, J. R.; Jansson, O. *Anal. Chem.* **1997**, *69*, 4369−4374.
(16) Nelson, R. W.; Krone, J. R.; Dogruel, D.; Tubbs, K. A.; Granzow, R.; Jansson, O. In *Techniques in Protein Chemistry VIII*; Marshak, D. R., Ed.; Academic Press: San Diego CA, 1997; pp 493−504.
(17) Nelson, R. W.; Krone, J. R.; Dogruel, D. Tubbs, K. A. In *New Methods for the Study of Molecular Aggregates*; Ens, W., Standing, K. G., Chernushevich, I. V., Eds.; NATO ASI Series C, Vol. 510; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1998; pp 225−238.
(18) Jarvik, J. W.; Adler, S. A.; Telmer, C. A.; Subramaniam, V.; Lopez, A. J. *Biotechniques* **1996**, *20*, 896−904.

of the biosensor chip using microliter volumes of a MALDI matrix solution. Once eluted, the matrix solution/peptide mixture was recovered and transferred to a mass spectrometer target for subsequent MALDI-TOF analysis. Using the second method, retained peptides were analyzed directly from the surface of the biosensor chip by application of nanoliter volumes of MALDI matrix to interaction areas (flow cells; FC) present on the chip followed by insertion of the entire chip into the mass spectrometer and MALDI-TOF analysis of the individual flow cells.

(3) To determine the practical limit of detection of the BIA/MS approaches: Critical to achieving the lowest limits of detection is the ability to distinguish targeted (tagged) species from nontargeted species. A dual-background-subtraction approach was employed to distinguish specifically retained species from non-specifically retained species. The first subtraction routine involved comparing or subtracting the SPR-BIA signals of blank flow cells from those of active flow cells. The process yielded an accurate quantitative estimate of the amount of material retained during SPR-BIA. The second subtraction routine involved subtracting mass spectra obtained when blank flow cells were analyzed from those obtained during analysis of active flow cells. This second subtraction distinguished nonspecific contributions to the SPR-BIA signal from specific contributions.

(4) To investigate the use of different affinity recognition systems in retrieving tagged polypeptides: The affinity tag used during these investigations contained seven histidines within a 16-residue section of the target protein. Additionally, there are no trypsin cleavages sites within the tag, allowing the histidine-rich tag to be conserved throughout treatment of the lysate with trypsin. Tryptic peptides, containing the tag fused to native sequence of the protein under investigation, can therefore be addressed using two different affinity recognition systems: a Ni^{2+} chelate system or a monoclonal antibody system. Both systems were investigated.

EXPERIMENTAL PROTOCOL

Expression of Tagged GST. The 91-base oligonucleotide GGATCCAAGATCTGGTACCCCACACCACACCACACCACA-CCACACCACACCACACCACACCACACCACACCACACCACAA-GATCTGAATTC was synthesized and cloned into the *Bam*HI/*Eco*RI sites in vector pUC118 (Midland Certified Reagent Co., Midland, TX). The sequence was confirmed by manual dideoxy sequencing using forward and reverse pUC118 primers. The insert was excised from the pUC118 construct with *Bam*HI and *Eco*RI and cloned into the GST expression vector pGEX-5X-3 (Pharmacia Biotech, Uppsala, Sweden) that had been digested with *Bam*H1 and *Eco*RI. An *E. coli* lysate containing epitope-tagged GST were produced by growing XL-1 Blue cells containing the pGEX construct in a 37 °C shaker to an optical density $A_{600\ nm}$ of 0.35 in Luria broth/2% glucose, inducing for 120 min with 0.3 mM IPTG, concentrating 25-fold by centrifugation, and lysing by sonication. The presense of induced epitope-tagged GST fusion protein in the lysate was confirmed by SDS gel electrophoresis and by Western blot.

The polynucleotide encoded a 27-amino acid peptide cloned into the *Shistosoma japonicum* glutathione-S-transferase (GST) gene in the vector pGEX-5X-3. The peptide contained the 16-amino acid sequence HTTPHHTTPHHTTPHH that contains three copies (two overlapping) of the epitope HTTPHH recognized by monoclonal antibody 7G8 (mAb7G8, see below). The amino acid sequence of the recombinant GST polypeptide (MW 30 893) is shown below. The inserted amino acids are shown in italics, with the epitope-containing sequence underlined.

MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEG
DKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIIRYIA
DKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSK
DFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVT
HPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEA
IPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLIE
GRGI*QDLVP<u>HTTPHHTTPHHTTPHH</u>TTPQDL*NSRVDSS
GRIVTD

Trypsin Digestion. A 10-μL aliquot of the lysate was mixed with 90 μL of either HBS buffer (0.01 M HEPES, 0.15 M NaCl, 3 mM EDTA, 0.005% v/v surfactant P20, pH 7.4) or NTA buffer (0.01 M HEPES, 0.15 M NaCl, 50 uM EDTA, 0.005% v/v surfactant P20, pH 7.4) containing 0.0005 mg/mL trypsin. Digestion was allowed to proceed for 30 min (40 °C) prior to quenching by the addition of bovine pancreatic trypsin inhibitor (final concentration ~0.001 mg/mL).

Preparation of Monoclonal Antibody. Mouse monoclonal antibody 7G8 (mAb7G8) was generated against the MAP-conjugated peptide PHHTTPHHTTPHHTT using standard methods (American Qualex Inc., San Clemente, CA). Hybridoma supernatants were screened by ELISA against the immunogen and by Western blot against the pGEX-5X-3 construct described above. An analysis of antibody binding to individual synthesized peptides contained within PHHTTPHHTTPHHTT indicated that the smallest continuous sequence sufficient for efficient antibody binding is HTTPHH (Spots analysis, Genysis). mAb 7G8 was isolated from ascites fluid using a protein A/agarose microcolumn. Fluid containing ~20 mg/mL protein was diluted by a factor of 10 with HBS and drawn through the column, after which the column was rinsed three times (500 μL) with HBS and once (500 μL) with doubly distilled water. Monoclonal antibody 7G8 was eluted from the microcolumn using two 25-μL aliquots of 10 mM sodium acetate (pH 2.0). The eluted IgG was then mixed with 100 μL of sodium acetate buffer (pH 4.5) in preparation for coupling to the biosensor chip.

Surface Plasmon Resonance Biomolecular Interaction Analysis. NTA/Ni^{2+}: Elution from Chip and Indirect MALDI-TOF Analysis. Stabilizing agents were washed from a Biacore (Biacore AB, Uppsala, Sweden) NTA (nitrilotriacetic acid) chip by rinsing the entire surface of the chip with five successive 200-μL aliquots of distilled water. The chip was then introduced into a Biacore X biosensor and subjected to NTA buffer flowed serially through the two flow cells at a continuous flow of 20 μL/min. Both flow cells were conditioned for SPR-BIA by injecting 20 μL of regeneration buffer (0.01 M HEPES, 0.15 M NaCl, 35 mM EDTA, 0.005% v/v surfactant P20, pH 7.4) over the flow cells (to remove any metal ions chelated to the surface of the flow cell). For blank analyses, 30 μL of digested lysate (in NTA buffer) flowed over the Ni^{2+}-free flow cells at a rate of 3 μL/min. When the injection was finished, the flow rate of the system was returned to 20 μL/min in order to more rapidly clear the flow cell of residual (unbound) digest components. After 3 min of rinsing the flow rate

was stopped, and the chip was undocked from the biosensor, rinsed with two successive 200-μL aliquots of distilled water, and allowed to air-dry. After drying, ~1.5 μL of ACCA matrix was applied to cover the area of both flow cells and allowed to stand for ~30 s before removal and application to a stainless steel mass spectrometer target. The chip was then rinsed with distilled water and returned to the biosensor. The same experimental procedure was used for Ni^{2+}-active analyses with the exception of exposing the flow cells to 20 μL of Ni^{2+}-charged NTA buffer (0.01 M HEPES, 0.15 M NaCl, 50 uM EDTA, 0.005% v/v surfactant P20, pH 7.4 containing 500 μM $NiCl_2$) resulting in a NTA-chelated nickel surface sufficient for binding histidine-rich polypeptides. Note: Although, in this experiment, MALDI matrix solution was used to elute retained compounds, other eluents capable of breaking Ni^{2+} interactions (e.g., EDTA or imidazole) would serve the same purpose.

CM5/mAb7G8: MALDI-TOF Analysis Directly from the Sensor Chip Surface. SPR-BIA analyses were performed using Biacore X biosensor. Stabilizing agents were washed from a CM5 (carboxylated dextran) sensor chip by rinsing the entire surface of the chip with five successive 200-μL aliquots of distilled water. The chip was then introduced into the biosensor, and both flow cells were subjected to HBS buffer at a continuous flow of 10 μL/min). Both flow cells (FC 1, FC 2) on the sensor chip were prepared using a serial derivatization procedure in which the flow cells were activated through exposure to *N*-hydroxysuccinimide (0.1 M prepared in HBS) containing 0.1 M *N*′-ethyl-*N*′-(dimethylaminopropyl)carbodiimide), derivatized by exposure to mAb 7G8 and then blocked by exposure to 1 M ethanolamine hydrochloride (in HBS adjusted to pH 8.5). After derivatization, FC 1 was deactivated (for use as a blank) by exposing the surface of the flow cell to 0.005 mg/mL trypsin dissolved in HBS (5 min; 10 μL/min). The flow cells were then allowed to equilibrate by continuously flowing HBS buffer (10 μL/min) through the biosensor for a period of ~12 h.

BIA proceeded by routing the trypsinized *E. coli* lysate (in HBS buffer) serially across the flow cells. Both flow cells were monitored in real time using SPR, and the signal from the blank flow cell (FC1) was subtracted from that of the active flow cell (FC2). A 30-μL aliquot of the analytical solution was driven across the flow cells at 10 μL/min before the injection was terminated. The chip was undocked from the biosensor, rinsed with two successive 200-μL aliquots of distilled water, and allowed to dry. The chip was prepared for mass spectrometry by applying ~100 nL of α-cyano-4-hydroxycinnamic acid (~50 mM dissolved in 1:2 acetonitrile/1.5% trifluoroacetic acid (ACCA)) to each of the flow cells using a thin-gauge wire.

Mass Spectrometry. Mass spectrometry was performed using a MALDI-TOF mass spectrometer and experimental methods that were described previously.[14]

RESULT AND DISCUSSION

NTA/Ni^{2+}: Elution from Chip and Indirect MALDI-TOF Analysis. As a point of reference, a MALDI-TOF spectrum of the digested lysate mixture used during these analyses is given in Figure 2. Although a number of signals are observed in the spectrum, no particularly useful information can be gathered from the analysis. Given the proportionally high number of histidines in the epitope tag, studies were undertaken to evaluate the possibility of selectively isolating tagged tryptic peptides using a Ni^{2+} chelate surface. Figure 3 shows sensorgrams resulting from SPR-BIA of trypsinized *E. coli* lysate using a Biacore X biosensor equipped with an NTA biosensor chip. Four sensorgrams are given: two from analyses using Ni^{2+}-active flow cells and two from Ni^{2+}-free flow cells. The Ni^{2+}-active analyses exhibit behavior characteristic of the specific binding of ligands to the affinity surface of biosensor chip. A plateau is observed in both sensorgrams representing the loading of the NTA surface with Ni^{2+}, followed by a gradual increase in response, during exposure to digested lysate, indicating the selective retention of material to the surface of the flow cell. After exposure to the digested lysate, solution flow was returned to NTA buffer in order to rinse the flow cell free of unbound digest components. The SPR response shows a relatively slow rate of dissociation (~3.5×10^{-4} s^{-1}), indicating a reasonable avidity between the epitope tag and the

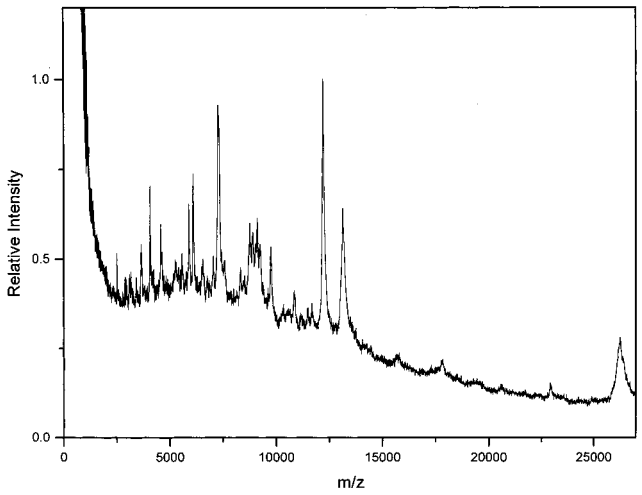

Figure 2. MALDI-TOF mass spectrum of unfractionated, trypsinized *E. coli* lysate. Although a number of signals are observed, no useful data can be gathered from the spectrum.

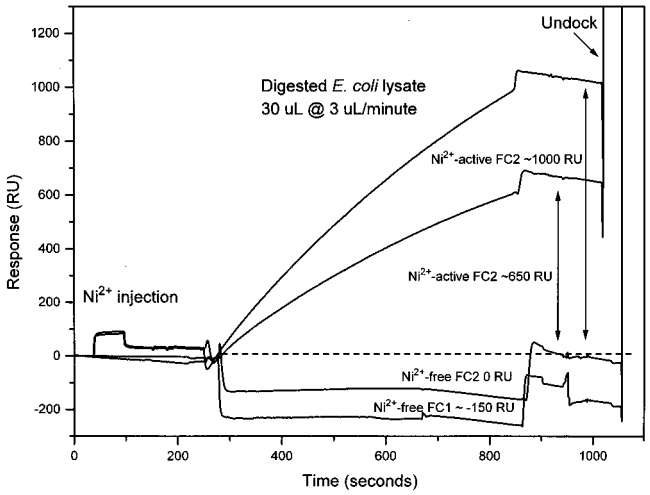

Figure 3. Sensorgrams resulting from SPR-BIA of the unfractionated, trypsinized *E. coli* lysate using an NTA sensor chip in the presence (Ni^{2+}-active) or absence (Ni^{2+}-free) of Ni^{2+}. Specific binding is observed in the presence of Ni^{2+}, presumably due to selective interaction of the histidine-rich epitope with the Ni^{2+} chelate surface. Approximately 1650 pg of material is estimated to have been selectively isolated on the two flow cells of the sensor chip during Ni^{2+}-active SPR-BIA (1000 RU = 1 ng of protein/flow cell).

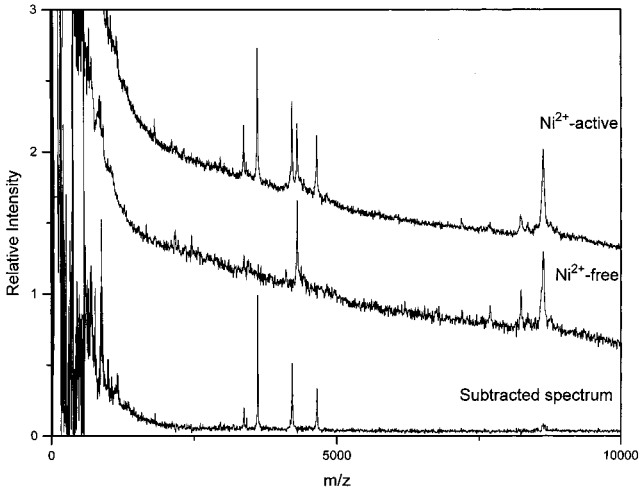

Figure 4. MALDI-TOF mass spectra of material eluted from the NTA sensor chip using a MALDI matrix solution. Several peptide species are observed in the spectrum resulting from Ni^{2+}-active SPR-BIA, whereas fewer are observed to result from Ni^{2+}-free SPR-BIA. The subtracted spectrum shows four signals resulting from specific interaction with the Ni^{2+} chelate surface. A detection level of ~100 fmol/peptide is estimated from these data and the data shown in Figure 3.

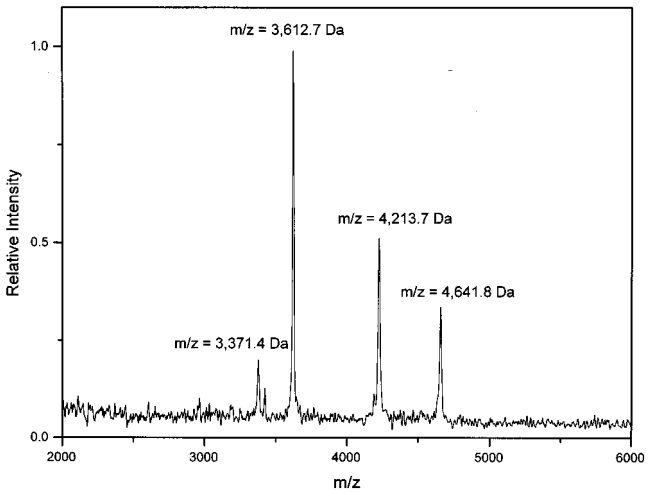

Figure 5. Expanded view of the subtracted spectrum shown in Figure 4. Four peptides at m/z = 3371.4, 3612.7, 4213.7, and 4641.8 Da are retrieved from the unfractionated, trypsinized *E. coli* lysate as a result of specific interaction with the Ni^{2+} chelate. These mass values were used in database searches identifying the native protein as glutathione-S-transferase, GST (see text).

Ni^{2+} surface. Response readings taken just prior to undocking of the sensor chip indicate ~1000 (1000 pg) and 650 RU (650 pg) of material selectively retained in flow cells 1 and 2, respectively. The blank flow cells exhibit no, or a slightly negative, response as a result of incubation with digested lysate, indicating little or no retention of lysate compounds during the analysis (the negative response is likely due to an incubation anomaly occurring at ~690 s into the analysis). The low response changes observed in these blank sensorgrams suggests that the ~1650 pg of material bound in the active flow cells is due to ligands specific to the Ni^{2+} chelate.

Figure 4 shows the MALDI-TOF mass spectra of the material eluted from the sensor chip. The mass spectrum taken of the material eluted from the Ni^{2+}-activated flow cells (both flow cells were addressed in a single elution with the matrix solution) shows strong ion signals at m/z = 3371.4, 3612.7, 4213.7, 4303.1, 4641.8, 8205.7, and 8594.0 Da. The mass spectrum of material retrieved from the Ni^{2+}-free sensor chip show ion signals at m/z = 4303.6, 8206.1, and 8594.9 Da. The third spectrum shown in Figure 4 is the background-subtracted spectrum resulting from subtraction of the Ni^{2+}-free spectrum from the Ni^{2+}-active spectrum. Background signals are observed to cancel out leaving four ion signals, at m/z = 3371.4, 3612.7, 4213.7, and 4641.8 Da, unique to the Ni^{2+}-active SPR-BIA analysis (Figure 5). These signals are presumably due to peptides containing the His-rich eptitope. Signals from neither the digested lysate (see Figure 2) nor the background components are observed in the background-subtracted spectrum. (As near as we can ascertain, the background signals observed in the Ni^{2+}-free spectrum are polypeptides due to bacterial contamination of water supplies that have a general affinity for the Biacore sensor chip surfaces. It is worth noting that these species have been observed repeatedly in many different water supplies and are somewhat independent of the affinity system present on the Biacore chip surface. Our present procedure of eliminating these mass spectral interferences is to use Biacore certified HBS buffer (the buffer is free of contaminants) or to filter water and buffers using a 500 MW cutoff membrane the day of use.)

Given that no response change was observed during the Ni^{2+}-free SPR-BIA (i.e., the background contaminants were present on the sensor chip before the analysis), the ~1650 pg of material retained during Ni^{2+}-active SPR-BIA is represented mass spectrometrically by the four ion signals in the ~4-kDa range. Using both the SPR-BIA and MALDI-TOF data, it is estimated that ~100 fmol of each of the ~4-kDa species was isolated onto the Ni^{2+}-active sensor chip. No estimates of the extraction efficiency from the chip can be made at this time.

CM5/mAb7G8: MALDI-TOF Analysis Directly from the Sensor Chip Surface. Figure 6 shows the activation and derivatization of the two flow cells (FC1, FC2) of a CM5 sensor chip with mAb7G8. An antibody density of ~8 ng/flow cell was determined by SPR-BIA, indicating ~50 fmol of IgG present on the surface of each of the flow cells. After the serial derivatization, FC1 was addressed with trypsin in order to deactivate the antibody. Figure 7 shows the sensorgrams resulting from the serial routing of trypsinized *E. coli* lysate across the surface of the flow cells. Three sensorgrams are shown, one for each of the two flow cells and one of the real-time background subtraction of the signal from the blank flow cell from that of the active flow cell (FC2 − FC1). Changes in response of 30 and 10 RU are observed for the active (FC2) and blank (FC1) flow cells, indicating the retention of 30 and 10 pg of material, respectively. The background-subtracted sensorgram shows a response change of 20 RU, indicating the retention of ~20 pg of material unique to FC2.

Figure 8 shows the mass spectra produced during the MALDI-TOF analysis of material directly from the surface of the flow cells (FC2, FC1) and the spectrum resulting from the background subtraction routine (FC2 − FC1). Minor signals at m/z ~ 2800 and ~6500 Da are observed to cancel out as a result of the subtraction, suggesting that the compounds responsible for these signals were retained through interactions with something other

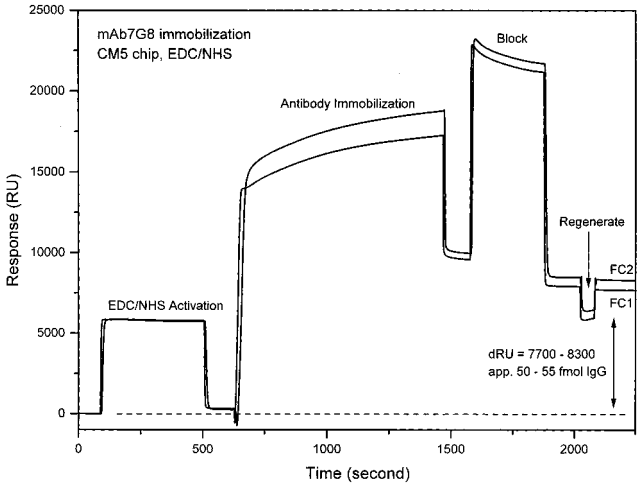

Figure 6. Serial activation and immobilization of mAb7G8 to the two flow cells (FC1, FC2) present on the surface of a CM5 biosensor chip. The flow cells were first activated using EDC-mediated binding of NHS to the carboxylated dextran of the chip. Monoclonal 7G8 was immobilized through coupling of primary amines to the NHS-activated surfaces. Antibody densities of ~50 fmol/flow cell are estimated from the SPR signal. Antibody on FC1 was deactivated after immobilization by exposure of the flow cell to trypsin.

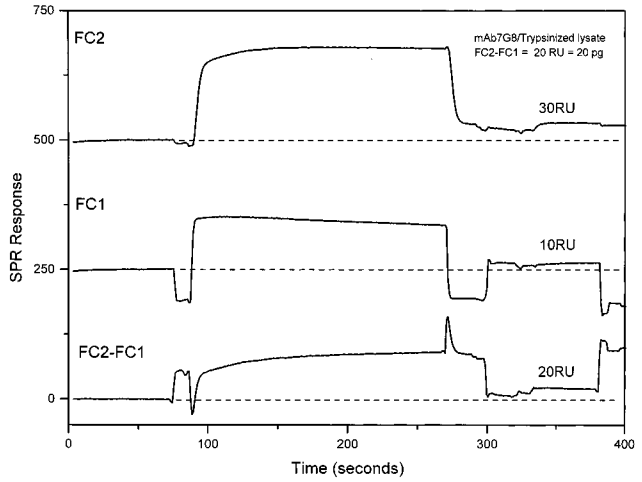

Figure 7. Sensorgrams showing the SPR response observed during the serial routing of unfractionated, trypsinized *E. coli* lysate across the surfaces of the active (FC2) and blank (FC1) flow cell. A 30-μL sample of trypsinized lysate flowed over the flow cells at a rate of 10 μL/min. Responses of 30 and 10 RU indicate the retention of 30 and 10 pg of material on the FC2 and FC1, respectively. The sensorgram resulting from real-time background-subtraction routine (FC2−FC1) indicates the retention of 20 RU (20 pg) of material unique to the active flow cell.

than the immobilized antibody. The multimeric signal in the m/z ~1800-Da range, due to residual detergent (present in the incubation buffer), is also observed to cancel out during the subtraction routine. Figure 9 shows an expanded view of the peptide region of the background-subtracted spectrum. As during the Ni^{2+} studies, ion signals are observed at $m/z = 3370.9$, 3612.3, and 4213.5 Da. Other signals are observed at higher mass, one at m/z ~4642 Da, as observed in the Ni^{2+} studies, and two others at m/z ~5 kDa.

A simple estimate of the mole amount of the peptides present on the biosensor chip can be made by considering data from both SPR-BIA (the mass of all the retained peptides, 20 pg) and MALDI-

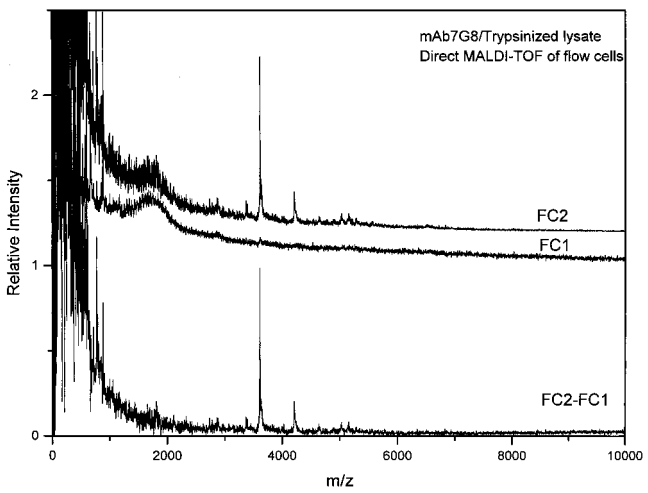

Figure 8. MALDI-TOF mass spectra obtained by targeting the active (FC2) and blank (FC1) flow cells. Ion signals are observed unique to the active flow cell (background subtraction, FC2 − FC1) due to specific interaction with mAb7G8. Considering these data and the data shown in Figure 7, the amount of each peptide present on the sensor chip is estimated at between 800 amol and 5.5 fmol.

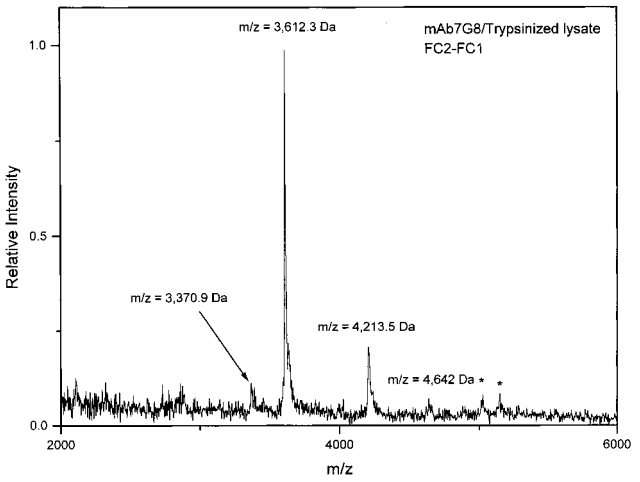

Figure 9. Expanded view of the peptide region of the background-subtracted spectrum shown in Figure 8. Ion signals are observed at m/z = 3370.9, 3612.3, 4213.5, and 4642 Da, consistent with the results shown in Figure 5. The two other ion signals (marked with asterisks) are also observed.

TOF (the number of peptide species and molecular weights, six peptides of average MW ~4300). Dependent on how the mass amount of retained material is divided between the six ion signals, the mole amount of peptides present on the chip is estimated to range between 5.5 fmol (high end of range − if the 20 pg of retained material was represented by only the 3.6-kDa ion species) to 800 amol (low end of range − if the 20 pg of material was distributed evenly between all six ion signals).

Protein Identification. It is known from other affinity-based mass spectrometric investigations performed on the tagged GST *E. coli* lysate system that the m/z ~3370-Da species is retrieved from lysate independent of treating the lysate with trypsin.[19] Essentially, when either a Ni^{2+} or mAb7G8 affinity reagent is used

(19) McLean, M. A.; Nelson, R. W.; Tubbs, K. A.; Krone, J. R.; Jarvik, J. W. *Proceedings of the 46th ASMS Conference on Mass Spectrometry and Allied Topics*, Orlando, FL, 1998; p 1155 (full manuscript in preparation).

in a mass spectrometric immunoassay protocol[10] applied to undigested lysate, intact (tagged) GST and the $m/z = 3370$ species are both retrieved from solution and register in the mass spectrum. Thus, the assumption can be made that the $m/z =$ 3371.4-Da species is not a product of trypsin digestion. Therefore, the three ion species at 3612.7, 4213.7, and 4641.8 Da can be considered tryptic fragments of the fusion protein that contain the epitope.

A number of methods for protein identification could be performed using these signals as starting points. The first of these methods would be to actively (i.e., during mass spectrometric analysis) derive partial sequence information from the peptide fragments using tandem mass spectrometric methods. Regarding the peptides eluted from sensor chips in high amounts, electrospray (ESI) techniques (micro- or nanospray) coupled to a variety of tandem MS instruments (Q-TOF, ion trap, triple-quadrupole) could be used to partially sequence the tagged peptides.[20] Regarding lower amounts of peptide and MALDI-TOF directly from the sensor chip surface, postsource decay (PSD) and in-source decay (ISD) methods have shown great promise in the partial sequencing of peptides. In either approach, the partial sequence data could be used as input into a database search capable of identifying the tagged protein.[21] Tandem mass spectrometric methods, however, were beyond the scope of our current instrumentation and therefore could not be attempted. An alternative for deriving sequence information, conducive to continuous-extraction linear TOF instrumentation, is to digest (ladder) the peptides while they are retained by the affinity receptor by using select exopeptidases. The exopeptidase treatment results in ragged end peptides, still retained on the affinity media, that can be rinsed free of reagents and then analyzed by protein ladder sequencing.[22] We have used this method quite successfully in studies of the current affinity system and a full report is forthcoming.[19]

Two other methods for protein identification that do not involve partial sequence determination can be used with the data available from these studies. The first of these is to subtract the mass of the epitope tag from the masses determined in the mass spectra. The resulting numerical values represent masses of tryptic fragments of the original, nontagged protein. These mass values may be used to fuel database searches capable of identifying the protein. However, a few precautions must be taken to ensure an accurate database search. Specifically, the exact mechanism of tag insertion and, more importantly, variations in gene sequence due to, for example, restriction enzyme specificity must be known. For instance, if the insertion mechanism of the epitope tag results in a perturbation of gene sequence (e.g., dropping of a codon or frame shift), then the protein sequence in the region of the tag is disrupted and the determined mass values will not accurately depict the original gene sequence. In this specific example of epitope tag insertion, two codons (coding for Pro229 and Arg230) are removed from the original sequence of the GST when treated with *Bam*HI and *Eco*RI. To adequately compensate for this deletion, the mass of the two deleted residues must be taken into account. Subtracting the mass of the epitope tag (3066.4 Da) from those determined via MALDI-TOF and then adding in the mass of the deleted dipeptide (253.5 Da) results in search values of 799.6, 1400.6, and 1828.7 Da. Database searching using these values (inserted into the search routine as tryptic fragments) with a 1-Da error window (ProFound, http//www.proteometrics.com) results in the top-rated match being gi595742, glutathione-S-transferase, found in the Genepept database. An advantage to performing the search in this manner is that, because affinity isolation is a component of the analysis, all of the tryptic peptides under analysis must form a nested set within the native protein. Better put, all of the tryptic peptides share a common region of the expressed protein—the region contains the epitope and a constant number of residues of the native protein (i.e., to the first tryptic cleavage site on either side of the epitope tag). Incomplete digestion of the expressed protein results in variable-length peptides that contain this constant region. As a result, an additional restriction is imposed on the database search, that the tryptic fragments must all have an overlapping sequence within the sequence of the native protein. It is for this reason that all but the single match listed above can be eliminated from the other potential candidate proteins retrieved during database search (by checking the coverage maps available through ProFound).

Another method for database search is to simply use the mass differences between the ion signals to fuel the database search. This method requires the addition of 18 Da to each mass difference in order to correspond to the in-silico numerical values. Three search values are calculated from the masses determined during MALDI-TOF: 446.1, 619.0, and 1047.1 Da. Database search using these values (PeptideSearch, http//www.mann.embl-heildelberg.de) with a 0.5-Da error window results in three versions of GST, of which one, accession number U13858, is the correct version. There are two issues of note regarding this approach. First, the precautions regarding knowledge of the epitope insertion mechanism are still in force with this approach. A second consideration is that this approach will work most readily if the epitope-containing tryptic fragments share a common terminus, which they did in this particular example. For more general application, however, it can be expected that peptides with N- and C-terminal heterogeneity may result from tryptic digestion, essentially complicating the data evaluation process. Possible solutions to the terminal heterogeneity problem are to insert tags, or to construct cloning vectors so that the tag is either terminus of the expressed protein, or to include a specific proteolytic cleavage site (e.g., thrombin or factor Xa) into the tag that can be cleaved, pre-trypsin treatment, to create epitope-containing peptides with a common terminus.

CONCLUSION

The experiments presented here were performed to investigate the use of BIA/MS as a general means of isolating, detecting, and identifying epitope-tagged polypeptides present in natural carriers. The data presented throughout this paper clearly demonstrate the use of BIA/MS to analyze compounds directly from complex biological mixtures and the ability to distinguish targeted compounds from nontargeted compounds. Although BIA/MS has already been used for such "ligand fishing", previous experiments were performed on mixtures of limited complexity, e.g., a specific toxin present at ~5% (w/w) in the whole venom of a rattlesnake,[12] on analytes in the presence of a large excess of a

(20) Yates, J. R., III *J. Mass Spectrom.* **1998**, *33*, 1−19.
(21) Reiber, D. C.; Grover, T. A.; Brown, R. S. *Anal. Chem.* **1998**, *70*, 673−683.
(22) Chait, B. T., Wang, R., Beavis, R. C.; Kent, S. B. H. *Science* **1993**, *262*, 89−92.

single exogenous carrier (albumin),[14−17] or on analytes "doped" into cell lysates.[13] By contrast, the system used in these experiments represents a highly complex biological mixture containing the analyte as an expressed protein. As demonstrated, BIA/MS is capable, using the background-subtraction routines, of the accurate quantitative and qualitative evaluation of compounds retrieved from the complex mixture. The background-subtraction routines support the conclusion that the major ion signals seen in Figures 5 and 9 are due to specific interaction with immobilized Ni^{2+} or antibody and therefore represent tagged polypeptides—an important point when considering the use of BIA/MS in the identification unknown species that interact with surface-immobilized receptors.

Although BIA/MS, or other affinity-based mass spectrometric techniques, can be performed using virtually any epitope tag, a feature of the pentanucleotide CCACA coding sequence used for this tag is that after at least three repeats it yields the same repeating amino acid sequence (HTTPH) no matter in which of the three possible reading frames it is translated. This property makes the tag especially useful in gene and protein discovery vehicles that randomly insert the sequence into genes, such as by delivering a CD cassette[18] containing the coding sequence into genomic DNA using transposon or retroviral delivery vectors. Alternatively, the gene/protein discovery process may be accomplished by insertion of cDNA or genomic DNA, from, e.g., genomic libraries, into appropriately designed cloning vectors that contain the repeating pentanucleotide or other epitope coding sequence. In either case, once the epitope coding sequence is incorporated into an unknown DNA sequence and the corresponding protein is produced, BIA/MS may be used to detect and identify the resulting tag-containing polypeptides using the methods describe above.

For the purpose of identifying proteins, there are at least three advantages that stem from operating on tag-containing proteolytic fragments rather than the intact protein. One advantage is that mass spectrometric sensitivity is generally greater for small molecules. A second advantage is MS/MS techniques are more ideally suited for tag-containing peptides than for proteins—as has been shown most elegantly by Gaskell et al. using immunoprecipitation of epitope-containing peptides in preparation for mass spectrometric characterization.[23] A third advantage is that antibody access to a linear epitope is less restricted when the epitope is part of a peptide—resulting in the more efficient capture of the analyte. One drawback of operating on peptides, however, is that the molar sensitivity of SPR-BIA is less for small molecules than for large molecules. It is worth noting that the SPR-BIA signal for the tagged GST is increased by a factor of ∼10 by performing the analysis on undigested whole lysate (data not shown). The increase in SPR-BIA signal was due to binding the heavier, intact analyte (MW ∼31 000) rather than the proteolytic fragments (MW ∼4000). It is questionable, however, whether a gain in SPR-BIA sensitivity is advantageous over the potential of further analyzing the peptides using mass spectrometric methods capable of protein identification.

As demonstrated, the BIA/MS approaches presented above are quite sensitive. The level of detection experienced using the elution method is equal to that experienced by other groups using a BIA/MS method in which retained species were eluted from the sensor chip and analyzed from a conventional mass spectrometer target.[24] However, the elution method described above is significantly simpler in that extensive washing of the biosensor unit is not required. An advantage of the elution method is reuse of the sensor chips. We have performed over 20 analyses using a single NTA chip without significant loss in activity of the sensor surface. A further advantage would be the possibility of analyzing the eluent using ESI mass spectrometry. However, retention of analyte at the 100-fmol level, and elution and sample transfers without loss of analyte, should not be expected as a universal result of SPR-BIA analyses. Under the circumstance where analyte is retained at the 10-fmol level, direct MALDI-TOF analysis from the sensor chip is needed for greater sensitivity. The level of detection experienced during these studies, in which bound peptides were analyzed directly from the sensor chip, is a factor of ∼10 better than that demonstrated in our previous studies of BIA/MS using MALDI-TOF directly from the sensor chip surface.[12−17] Detection of tagged species at the low-femtomole to subfemtomole levels is considered highly significant, especially considering they originated in the *E. coli* system. Given an antigen retrieval efficiency of ∼10%, this amount of sample represents operating on small volumes at analyte concentrations of ∼1 nM (∼30 fmol/30 μL). Although the ability to perform a multiplexed analysis at this concentration is viewed as significant, we do not feel that we have yet tested the limits of sensitivity for BIA/MS analysis of epitope-tagged proteins. When the possibility of improving the analysis by using higher immobilized ligand densities, lower K_d antibodies, slower incubation flow rates, and/or repetitive incubation of the analytical fluid are considered, the prospects for developing BIA/MS as a routine multiplexed technique for the analysis of epitope-tagged polypeptides in small samples are encouraging.

ACKNOWLEDGMENT

The technical support of Dr. Daniel Moothart at American Qualex Inc. (San Cemente, CA) is greatly appreciated. Biacore AB (Uppsala, Sweden) is acknowledged for use of the Biacore X instrument. This publication was supported in part by Grant 1 R43 CA82079-01 from the National Cancer Institute. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute.

Received for review January 28, 1999. Accepted April 13, 1999.

AC990089V

(23) Yu, L.; Gaskell, S. J.; Brookman, J. L. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 208−215.

(24) Sonksen, C. P.; Nordhoff, E.; Jansson, O.; Malmqvist, M.; Roepstorff, P. *Anal. Chem.* **1998**, *70*, 2731−2736.