Developing a natural partnership for rapid, sensitive identification of binding partners
Biacore 3000 and mass spectrometry

Summary
The coupling of label-free protein interaction analysis to the precise mass determination of biomolecules captured on the sensor chip is a powerful and valuable tool for fishing of binding partners. Here, we present a method for automating the process, from capture to recovery, tryptic digestion, and mass spectrometry (MS) sample preparation. Furthermore, the use of reversed-phase columns to concentrate recovered samples prior to MS analysis is discussed and exemplified by the extraction and identification of calmodulin from brain extract, using LC-ESI MS and tandem MS analysis.

Introduction
Complex biological mixtures such as serum, tissue homogenates and cell lysates present a very challenging environment for the identification of specific molecules as they are comprised of thousands of different biomolecules, and the target molecule represents only a minute fraction of the total mixture.

Label-free interaction analysis based on surface plasmon resonance (SPR) is a powerful tool for directly measuring the binding of biomolecules in solution (the analyte) to its binding partner, which is immobilized on the chip surface either by covalent binding or high affinity capture. Biacore™ 3000 allows for the direct measurement of binding, even against the background of complex mixtures. This analysis provides not only confirmation of the presence of the analyte, but also kinetic data on the interaction as well as concentration determination.

Mass spectrometry can accurately measure the mass of a protein or its peptide fragments obtained by enzymatic cleavage. By comparing the masses of fragmented proteins to a database of known proteins, an identification can be made based on the unique mass fingerprint. By applying tandem mass spectrometry (MS/MS), sequence information is obtained that makes the protein identification even more unambiguous and, at the same time, localizes post-translational modifications.

Coupling the two technologies is an obvious and powerful approach, and a number of successful attempts have been reported. Such integration was not always a simple matter, as the operational conditions of the two systems (running buffer, elution buffer, cleavage conditions, matrix addition) differ, making such experiments relatively slow and occasionally troublesome. To further streamline the technology interface we present a method for automatically coupling the two techniques, using a fishing approach in which the ubiquitous signal transduction protein calmodulin is isolated from brain extract prior to MS analysis. The system sequentially and automatically performs all the required steps of capture, recovery, sample deposition onto a MALDI target, matrix deposition, and provides the option of on-target tryptic digestion. Alternatively, the sample can be recovered in a vial prior to the addition of trypsin and analysis by ESI-MS and MS/MS. The combination of label-free interaction analysis with the identification and sequencing ability of MS and MS/MS creates a powerful tool for proteomic studies.

Acknowledgement
This work was carried out by Östen Jansson, Andrei Zhukov & Jos Buis (Biacore AB, Uppsala, Sweden) and Detlev Suckau, Martin Schürenberg & Markus Lübeck (Bruker Daltonics, Bremen, Germany).
Materials and Methods

Reagents
Calmodulin and the calmodulin-binding domain (CBD) of myosin light chain kinase were purchased from Calbiochem (San Diego, California, USA). Bovine brain acetone powder was from Sigma-Aldrich (Stockholm, Sweden). Peptide calibration standards and α-cyano hydroxycinnamic acid (HCCA) were from Bruker Daltonics (Bremen, Germany). Acetonitrile was from Lab-Scan (Dublin, Ireland). Acetone, formic acid and trifluoroacetic acid (TFA) were from Merck (Darmstadt, Germany). Sequencing grade modified trypsin was from Promega (Madison, Wisconsin, USA).

CBD immobilization and analyte capture
All experiments were carried out using Biacore 3000 and Sensor Chip CM5 at 25°C. The CBD was immobilized in all four flow cells using EDC/NHS primary amine linkage with HBS-EP as the running buffer. For immobilization, the CBD was injected at a concentration of 0.1 mg/ml in 10 mM K-phosphate buffer pH 7.4 for 10 minutes, followed by a seven minute injection of 1 M ethanolamine to deactivate any residual active groups. Before use the chip surface was preconditioned with ten 30 second injections of 50 mM NaOH to remove any remaining non-covalently bound CBD from the sensor chip.

The brain extract was prepared by vortexing 10 mg of Bovine brain acetone powder in 1 ml of HBS-N buffer with 2 mM CaCl$_2$ for 5 minutes. The insoluble residue was pelleted by centrifugation and discarded. The resulting supernatant was diluted tenfold in the same buffer to create the final brain extract solution.

The capture and recovery of the target protein as well as sample preparation for MALDI MS were steered entirely by a method written in Biacore Method Definition Language. 200 µl of the brain extract solution was injected at a flow rate of 20 µl/min across all four flow cells using the MS_INJECT command. The flow system was first washed twice with 50 mM NaOH, then rinsed with 50 mM NH$_4$HCO$_3$. The flow cells were rinsed using the MS_WASH command with 50 mM NH$_4$HCO$_3$, 2 mM CaCl$_2$ to remove any non-specifically bound material from the brain extract from the chip surface.

Protein and post-translational modification identification by MALDI MS and tandem MS after on-target digestion of sample
Bound analyte was eluted from the chip with 50 mM NH$_4$HCO$_3$, 2 mM EGTA and deposited onto a 0.8 mm spot of the AnchorChip™ target (Bruker Daltonics) using the MS_RECOVER command. One µl of 15 µg/ml trypsin in acetonitrile:water (60:40) was added using the MICROTRANSFER command, and the sample was allowed to dry by pausing program execution for 20 min by means of the WAIT command. One µl each of 0.5% TFA and 0.5 mg/ml HCCA in ethanol:acetone (2:1) was then place on the spot using the MICROTRANSFER command and allowed to dry.

The mass spectrum was recorded on an Autoflex MALDI-ToF mass spectrometer (Bruker Daltonics) operated in the reflector mode. Bruker Daltonics peptide calibration standard was used for external calibration. The peptide masses were submitted to the MSDB database via the Mascot database search engine (Matrix Science) for protein identification.

The MS/MS spectra were acquired and analyzed on an Ultraflex MALDI-TOF/TOF mass spectrometer and all peaks in the mass fingerprint that were assigned to Calmodulin were subjected to MS/MS analysis using the LIFT procedure (Bruker Daltonics).

In-vial digestion and LC-ESI MS and MS/MS analysis
In this experiment the material bound to the CBD-modified sensor chip was eluted in a vial and trypsin was added to a final concentration of 5 µg/ml and incubated overnight at room temperature. The sample was acidified by adding formic acid to a final concentration of 1% and loaded onto a reversed-phase column (monolithic column PS-DVB, LC-Packings) that was heated to 60°C using an LC system (Ultimate, LC-Packings). The LC system was coupled to an Esquire HCT ion trap (Bruker Daltonics) with an Agilent ESI interface. The material bound to the reversed-phase column was eluted with a 0-80% acetonitrile gradient in 11 minutes and a flow rate of 2.5 µl/min. The eluted material was analyzed by generating single scan spectra while automatically alternating between the MS and MS/MS mode.

For experimental systems where the amount of material captured on the sensor chip is sub-optimal for MS analysis (typically less than 100 fmol), the in-vial digestion protocol offers the advantage of collecting material from multiple capture and recovery cycles. The reversed-phase column, inserted in the LC-loop, provides the means to concentrate and purify the sample prior to ESI analysis. Note that additional sample concentration and purification can also be obtained for MALDI MS by using reversed-phase columns embedded in pipette tips.
Results and Discussion

Analyte capture and recovery

While label-free protein interaction analysis has been integrated with MS in the past to identify binding partners, the goal in this experiment was to produce a system that would more easily integrate the two technologies while producing more useful information about the structure of the target protein. Calmodulin binding to the calmodulin-binding domain (CBD) of myosin light chain kinase was selected as a model system. Calmodulin is a ubiquitous 17 kDa regulatory protein involved in a number of signal transduction pathways. Myosin light chain kinase is an interaction partner to which calmodulin binds in the presence of Ca\(^{2+}\) ions. CBD, a 20-mer fragment of the kinase responsible for the interaction with calmodulin, was immobilized on Sensor Chip CM5 via amine coupling. The goal of the experiment was to capture calmodulin on the chip from a crude biological mixture (bovine brain extract), followed by the recovery, identification and characterization of the captured material by MS.

The 1300 RU of calmodulin recovered from each flow cell corresponds to a total mass of 5.2 ng on the chip surface, or about 300 fmol for a 17 kDa protein. The total amount delivered to the MALDI target may be less due to possible losses within the flow system between the chip and target.

On-target digestion and protein identification by peptide mass fingerprint

For this experiment the calmodulin was eluted from the chip surface by the addition of 50 mM NH\(_4\)HCO\(_3\) with 2 mM EGTA. The EGTA triggered calmodulin release by chelating the Ca\(^{2+}\) ions necessary for CBD binding. The MS_RECOVER command was used to deliver 2 \(\mu\)l of the eluate onto a 0.8 mm spot of a Bruker AnchorChip target mounted on a Biacore MALDI Holder. This was immediately followed by the transfer of 1 \(\mu\)l of trypsin solution in 60% acetonitrile onto the same spot to give a final volume of 3 \(\mu\)l and an acetonitrile concentration of 20%. Previous experiments had shown that it took 15 to 17 minutes for the spot to dry under ambient conditions, so a 20 minute WAIT command after trypsin addition was included in the method. After this time, the sample was acidified and mixed with the matrix by transferring 1 \(\mu\)l each of the TFA and HCCA solutions to the spot.

Protein identification by peptide mass fingerprint usually involves the incubation of the target protein with proteases for several hours in a sealed vial, as performed for the ESI-MS analysis. In addition, this approach requires a series of sample transfer steps inevitably leading to sample losses by adsorption, especially when protein concentrations are low. The automated on-target digestion method described here has the advantages of speed, automation, and reduced sample loss. Previous on-target digestion protocols were difficult to adapt to this new method due to their requirement for an incubation chamber that had high ambient humidity at 37°C for 30–60 minutes. The new experimental design utilizes the presence of organic solvents, which dramatically increase the rate and extent of tryptic digestion, producing extensive cleavage within a few minutes at room temperature. Acetonitrile concentrations as high as 80% can be used with trypsin, which may be necessary for proteins that are more resistant to trypic digest than calmodulin. On a conventional stainless-steel target, liquids with a high fraction of organic solvent tend to spread over a large area, reducing sensitivity. The AnchorChip target avoids this problem through a hydrophobic coating that constrains the sample material to the small hydrophilic anchor spot upon drying, even if the surface tension in the drop is very low.

[Figure 1. Sensorgram of calmodulin capture on and recovery from a chip derivatized with calmodulin-binding domain. Shadowed areas correspond to the steps where the running buffer is automatically exchanged for a user-defined MS-compatible buffer.]

Figure 1 presents a typical sensorgram (plot of binding response versus time) obtained after the injection of brain extract over a CBD-derivatized chip, washing of the fluidic system, and recovery of the bound material. These three operations were accomplished by the consecutive execution of the commands MS_INJECT, MS_WASH, and MS_RECOVER. Note that these three command are automatically executed when the Analyte Recovery wizard is used. The abrupt changes in response immediately before and after the injection are due to bulk refractive index changes when the MS_INJECT routine switches between the running buffer and a user-specified MS-compatible buffer in order to minimize the contamination of the recovered material with the components of the running buffer. In this experiment 50 mM NH\(_4\)HCO\(_3\) with 2 mM CaCl\(_2\) was used as the MS-compatible buffer.
After drying, the sample was analyzed by MALDI-ToF MS. Figure 2 shows the resulting MS spectrum and the sequence coverage map obtained after the eluted protein was identified as bovine calmodulin. The database query included an option of acetylated N-terminus (as is the case with calmodulin), which helped identify two of the peaks (1563.7 and 3389.7 Da). Bovine calmodulin was identified as the top match with a Mowse score of 137. An additional peak (2401.1 Da) was identified by matching the mass spectrum with the in silico digest containing another known modification of calmodulin, trimethylation of Lys 115. The overall sequence coverage was 95%.

Analyte capture on a Biacore sensor chip coupled to MS has previously been shown to effectively extract a protein from a crude extract and identify it based on its mass, but this automated method increases the speed of the process and reduces potential errors and inconsistencies from inter-operator variability. All the steps of the fishing experiment from capture to sample preparation for MALDI-ToF MS were performed in a fully automated regime, making it ideally suited for scaling-up towards large scale proteomics.

**Peptide sequencing and identification of post-translational modifications by MS/MS**

MS/MS is a powerful tool for protein analysis, which uses multiple mass selections in series to gain further sequence information about every peptide in the digest. The first MS separates out ions within a pre-selected mass range. These selected ions are fragmented and then accelerated away for subsequent mass analysis. Peptide sequencing by MS/MS is the most trustworthy method for protein identification for three reasons. Tandem MS allows de novo sequencing, which is important since not all protein species can be found in databases, ii) identification and localization of post-translational modifications, iii) increased certainty in protein identification.

**ESI-MS/MS**

Using LC-ESI for identification of the material eluted from the sensor chip, only a few low-intensity peaks corresponding to the tryptic digest of calmodulin could be found in the MS spectra. These data alone were not enough for a positive identification of calmodulin. The complementary information obtained from the ESI-MS/MS analysis of these peaks, however, made it easy to unambiguously identify calmodulin (with a Mowse score of 253, data not shown). Mowse scores are a measure of the degree of certainty of identification from database queries, and the score of 253 compared with the 137 score obtained from the MALDI-MS mass fingerprint (Figure 2), demonstrates the superior data quality provided by the ESI-MS/MS approach.

**MALDI-MS/MS**

The sequence information obtained from MS-MS analysis also allows for the identification of amino acids that have undergone post-translational modification. As such modifications can significantly affect protein activity, the ability to distinguish modified sub-populations is a critical aspect of deciphering protein function.

Figure 3 shows the MS/MS spectra of the peptides with molecular masses of 2401.1 and 3389.7 Da. In addition to the complete sequence, two modifications have been detected. A lysine in the former peptide corresponding to Lys 114 in bovine calmodulin is either trimethylated or acetylated (these two modifications give the same mass increment), and the N-terminal alanine in the latter peptide is acetylated. This agrees with previous studies showing that Ala 1 is acetylated and Lys 114 trimethylated in calmodulin.
Figure 3. MALDI MS/MS trace showing the sequencing of the peptide fragments generated by the LIFT procedure. By determining the mass of individual amino acids the presence and type of post-translational modifications could be determined. Panels A and B show spectra from the 2401.1 and 3389.7 Da calmodulin peptides. The green vertical bars on the spectra indicate the regions that identified post-translational modifications.

Conclusions

Previous work has shown the power and utility of combining label-free protein interaction analysis with mass spectrometry, but historically differences in running conditions between the two techniques made integration relatively complicated and cumbersome. Using the analyte recovery function in Biacore 3000 makes it possible to simplify and speed the integration process, allowing samples to be affinity separated using the Biacore instrument, then deposited directly on a MALDI target for MS analysis or deposited in a vial, followed by the addition of trypsin for LC-ESI analysis. Additional recovery-dedicated MDL commands enable the writing of methods that allow for on-target tryptic digestion, which when coupled to MS gives a more certain identification of the sample based on the unique set of peptide mass fingerprints. These peptides can also be analyzed by MS/MS to obtain sequence information and to identify post-translationally modified residues. In combination with the sensitivity and identification power of mass spectrometry, this approach yields a powerful and versatile tool for fishing of binding partners and proteomics.