Published in final edited form as:

Anal Biochem. 2015 April 1; 474: 8–15. doi:10.1016/j.ab.2015.01.004.

T³: Targeted Proteomics of DNA-Binding Proteins

Linda I. Nagore and Harry W. Jarrett

Department of Chemistry, University of Texas at San Antonio, 1 UTSA Circle, San Antonio, TX 78249. Phone: (210) 458-7053 Fax: (210) 458-7428 Harry.Jarrett@utsa.edu

Abstract

A technique that allows the inclusion of a specific DNA to enrich and direct proteomic identification of transcription factors (TF) while providing a route for high throughput screening on a single platform would be valuable in investigations of gene expression and regulation. Polyvinylpyrrolidone binds DNA avidly while binding negligible amounts of protein. This observation is used in a proof-of-concept method to enrich for TF by combining nuclear extract with a specific DNA sequence and immobilizing the DNA-protein complex on a PVP-coated MALDI plate. Any unbound proteins are washed away and further processed for analysis in a MALDI-TOF/TOF mass spectrometer. Enrichment on a PVP-coated plate gives the unique advantage of purification, enzymatic digestion and analysis on a single platform. The method is termed T³ as it combines Targeted purification on a Target plate with Targeted proteomics. Validation was achieved in model experiments with a chimeric fusion protein, green fluorescent protein-CAAT enhancer binding protein (GFP-C/EBP) with an oligonucleotide containing the CAAT sequence. Both domains were identified with an expectation value of less than 10⁻¹⁵ and over 15% sequence coverage. The same oligonucleotide mixed with HEK293 cell nuclear extract allowed the unambiguous identification of native human C/EBP alpha with 24.3% sequence coverage.

Keywords

Transcription factors; DNA binding proteins; mass spectrometry; MALDI-MS; proteomics

INTRODUCTION

Transcription factors (TF) are present in all living organisms and are vital for numerous cellular processes and comprise approximately 6% of the proteome [1]; less than 5% of those characterized at the protein level. With so much still unknown about how gene expression is regulated, a high-throughput system is essential to identifying every TF for each individual promoter DNA. A model is depicted in Fig. 1 of a proof of concept, where we envisioned using a polymer coated MALDI (matrix assisted laser desorption ionization)

© 2015 Elsevier Inc. All rights reserved.

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

plate to purify low abundant DNA-binding proteins from nuclear extract. Mixing a DNA-binding protein with DNA that contains the corresponding binding sequence forms a complex that is then spotted onto a coated MALDI plate that binds only DNA. In complex samples, proteins that are not bound to the DNA are simply washed away, enriching for DNA-specific binding proteins. This scheme permits complexation, enrichment, and on-plate digestion to be accomplished on a single platform. The prepared sample can be analyzed directly from the plate with a MALDI-TOF/TOF mass spectrometer (MS). The method is called T³ since it combines Targeted purification on a Target plate with Targeted proteomics.

To convert this concept into practice, a suitable coating with specific properties was needed. One being it binds DNA but only retains negligible amounts of protein. It must also be relatively inert to the chemicals needed for sample processing and MS analysis while remaining bound to the plate surface. Several potential candidates were examined before identifying polyvinylpyrrolidone (PVP).

PVP has many unique characteristics that make it optimal for proteomic purification. PVP is biologically inert, nearly chemically neutral, pH stable, and non-ionic making it a suitable environment for reactions to take place without any secondary effects occurring along with producing a suitable surface for MALDI-MS [2; 3; 4]. PVP forms a hydrogel layer capable of absorbing forty times more water than its dry weight and does not dissolve when immersed in most solvents [5]. This is particularly important so that samples can be dried and concentrated on a single spot. The DNA-PVP binding phenomenon is caused either by hydrogen bonding between PVP and the major groove of DNA (pH 4 to 6) or through π stacking interactions between aromatic rings [6]. From previous research where crosslinked PVP has been used to extract caffeine, tannins, phenols, indoleacetic acid, DNA and other aromatic compounds, π -stacking is the likely mode of interaction [7]. For protein, amino acids containing aromatic rings are generally located in the protein's globular interior and inaccessible, allowing PVP to bind to DNA with high affinity but not bind proteins [8; 9].

PVP has been shown useful for desalting DNA samples and MALDI-MS being an instrument that produces primarily singly charged ions it has the ability to easily analyze DNA of less than 50 nucleotides [10] [11]. By changing a few parameters, such as the matrix and the mass range, the same sample used for protein analysis can be used for DNA analysis confirming that it has not been degraded or modified during the experiment.

There are many assays available to investigate protein-DNA interactions. One widely used is Chromatin Immunoprecipitation (ChIP). This assay determines if specific proteins (e.g, TFs) are associated with a specific gene. Unfortunately, like with all techniques, there are limitations. The method utilizes antibodies specific for a given TF. Assuming an antibody is available the method only has the ability to detect one protein at a time, making investigations of unknown proteins or those lacking an antibody slow. Southwestern-blotting uses radiolabeled DNAs to detect multiple DNA-binding elements at a time, however characterization is time consuming and false positives occur. With many of these methods there are multiple steps, sometimes requiring the sample to be transferred from one apparatus to another, creating opportunities for sample loss and making extremely low

abundant proteins undetectable. Non-specific binding is also a concern but can be reduced by using low concentrations of probe DNA (nM) and an excess of competitor synthetic DNA such as poly(dI:dC) [12].

While the method described here can be used for DNA analysis, the focus here is to use the properties of PVP to enrich proteins that bind a specific DNA sequence and also provide adequate sequence coverage for positive identification of that protein. Here, we investigate these concepts and show that not only do irrelevant proteins not bind to the coated plates as contaminants but loss of DNA-binding proteins is minimal.

EXPERIMENTAL DETAILS

Materials

Polyvinylpyrrolidone, PVP-40 (average molecular mass 43 kDa.), dI:dC, diammonium hydrogen citrate (Cat# 61116-5000) were from Sigma Aldrich (St. Louis, MO). Indium Tin Oxide (ITO) coated glass plates (Cat# 237001) were from Bruker Daltonics (Billerica, MA). α -Cyano-4-Hydroxycinnamic Acid (HCCA), Sinapinic Acid (SA), 3-Hydroxypicolinic Acid (3-HPA), and peptide calibration standard mixtures were from Bruker Daltonics (Billerica, MA). Porcine trypsin (Cat# V5280) was from Promega (Madison, WI). Other chemicals were of the highest purity obtainable commercially.

ITO Slide Glass Coating

ITO glass slides were thoroughly washed to ensure that an even layer of PVP is coated onto the surface. The plate is first washed with ethanol, then acetone, and 40% nitric acid, each for one hour on a shaker and the wash solution discarded. To ensure all the nitric acid is removed the plate is washed with H₂O for one minute, repeating twice. Adsorption of PVP to the MALDI plate was accomplished by immersing the plate in 5 mL of a 25 mg PVP/100 mL methanol solution and allowed to stand undisturbed for 72 hours under N₂. After the incubation period the PVP solution is removed and the plate dried under N₂.

DNA Probes

EP18 (5'-GCAGATTGCGCAATCTGA-3') and other oligonucleotides were synthesized by Integrated DNA Technology (Coralville, IW, USA). EP18 is self-complementary and annealing yields the duplex. As irrelevant DNA sequences we used two derived from the c-jun promoter (CJ1, CAGCGGAGCATTACCTCATC, and CJ19, CTCGGGCTGGATAAGGGCTC, both annealed with their complement strand) and two derived from the hTERT promoter (HT1 CCCACGTGGCGGAGGGACT and HT8 CCGCTCCTCCGCGCGGACC). None of these four oligonucleotides contain a CAAT element and should not bind C/EBP.

Preparation of Green Fluorescent Protein - CCAAT/Enhancer Binding Protein

GFP-C/EBP was produced in *Escherichia coli* strain BL21 containing plasmid pJ22-GFP-C/EBP and purified as previously described [13]. The chimera contains 44 amino acids derived from plasmid sequences (encoding a His-tag, T7 antigen, and thrombin cleavage

site), 239 amino acids derived from GFP, and 101 amino acids derived from rat C/EBP for a total of 381 amino acids.

Application of Sample to Coated Plates

Proteins (typically, 25 µg) were mixed with 10 nM of oligonucleotide in binding buffer (50 mM NaCl, 10 mM HEPES, pH 7.5, 10 mM MgCl₂, 1 mM EDTA, 50 µM ZnSO₄, 1 mM DTT, and 30 µg/mL poly dI:dC) to a final volume of 25 µL. Human embryonic kidney cell (HEK293) nuclear extract was prepared as previously described [14] and stored at -85°C. The mixture was either incubated at 20°C for 30 min. or simply spotted directly on the plate and incubated there for 30 min. Either way works equally well. Five microliters was spotted, and after 30 min. the spot absorbed. The spot was then washed twice by applying 5 µL binding buffer and immediately removing it. Finally the spot was washed once with 5 µL H₂O.

On-Plate Digestion

A humidity chamber was assembled by creating a raised platform for the MALDI plates in the center of a sealable container. On the bottom of the container (on each side of the platform) damp tissues were placed and then a cover was placed over the container.

The sample was reduced by adding 5 µL 10 mM DTT to each spot and incubated at 37° for one hour in the humidity chamber. The protein was alkylated by adding 5 µL of 40 mM iodoacetamide to each spot and incubating at 37° for one hour in the dark in the humidity chamber. Finally, 5 µL of 100 ng/µL trypsin in 50 mM NH₄HCO₃ was added and incubated at 37° overnight in the humidity chamber.

The next day, 3 µL of the appropriate matrix was spotted onto each spot and dried.

Matrix Preparation

α -Cyano-4-Hydroxycinnamic Acid (HCCA): A solution of HCCA was made by dissolving 10 mg HCCA in 1 mL of 30:70 (v/v) acetonitrile: 0.1% TFA in H₂O.

Sinapinic Acid (SA): A solution of SA was made by dissolving 10 mg SA in 1 mL of 30:70 (v/v) acetonitrile: 0.1% TFA in H₂O.

3-Hydroxypicolinic Acid (3-HPA): 3-HPA was prepared at a concentration of 50 mg/mL in H₂O / ACN (1:1 v/v). 100 mg/mL diammonium hydrogen citrate (AHC) was mixed with 3HPA at a 10:1 v/v (3-HPA:AHC).

Mass Spectrometry

Matrix assisted laser desorption ionization mass spectrometry (MALDI-MS) was accomplished on an UltrafleXtreme MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Bellerica, MA). Spectra were collected using FlexControl (Compass for FlexSeries 1.4, Version 3.4, Build 119) in reflectron mode for the parent mass peaks and LIFT experiments were utilized for MS/MS data.

An external calibration of the MALDI-TOF-MS was performed using a peptide mixture (Peptide Calibration Standard II, Cat.# 222570, Bruker Daltonics, Billerica, MA). Positive ion MS data collection was performed between m/z 900 and 4500 with 1000 laser shots per spot at 20% laser power and averaged. Fifteen of the most intense ion signals per sample were selected for MS/MS analysis with 2000 laser shots (20% laser power) per selected peak. A separate calibration was performed for higher mass ranges using a protein mixture and oligonucleotide mixture (Protein Calibration Standard II, Cat.# 207234, Oligonucleotide Calibration Standard, Cat.# 206200, Bruker Daltonics, Billerica, MA).

Protein Database Searching

Searches were performed using two software tools, FlexAnalysis (Version 3.4, Build 50, Bruker Daltonics) and BioTools (Version 3.2, Build 4.48, Bruker Daltonics). Tandem mass spectra were searched against the SwissProt database (release 2012_11; 538,577 sequences) with an in-house a 10-node Mascot cluster (version 2.3.02, Matrix Science, London, UK) using TopHat peak picking algorithm in FlexAnalysis. The database search was accomplished with the following parameters: three missed cleavages, enzymatic digestion by trypsin, monoisotopic peptide mass tolerance of 100 ppm, fragment ion mass tolerance of 0.5 Da, confidence interval of 95%, carbamidomethylation set for fixed modifications, oxidation of methionine for variable modifications, and false discovery rate (FDR) less than 5%. Proteins were identified on the basis of two or more peptides whose ion scores exceeded 20. Ion score is calculated based on $-10 \cdot \log_{10}(P)$, where P is the absolute probability.

RESULTS AND DISCUSSION

Model Experiments: Does the PVP coating bind DNA and DNA-protein complexes?

A previously characterized chimeric fusion protein GFP-C/EBP, which contains green fluorescent protein fused at its C-terminus to a 101-amino acid fragment of C/EBP, binds the EP18 oligonucleotide [13; 14]. Here, GFP-C/EBP and EP18 was used to prove the concepts presented in Fig. 1. PVP was used to coat MALDI plates for improved throughput and sensitivity of TF identification. We found that retention of GFP-C/EBP requires both the coating and the specific DNA as shown in Fig. 2 using the MALDI-TOF.

When the protein-DNA mixture was applied to an uncoated surface using the same procedures, little protein is recovered for mass spectrometry (Fig. 2A). While there are a few peaks present, observable at high sensitivity, these contaminants are very low intensity and are not expected to interfere. When a coated plate is used, the results are quite different. When the same mixture as Fig. 2A is applied to a coated plate in Fig. 2B, over 15 peptide peaks were observed and identified as arising from GFP-C/EBP. When the same mixture was prepared except leaving out the DNA (Fig. 2C), again few peaks are found, even at high sensitivity. The only major peaks were identified as trypsin autolysis fragments. The sheer abundance of peptide peaks present in Fig. 2B as well as the absence of peaks in Fig. 2A and 2C proves that the method requires both the PVP coating and DNA.

Is the DNA sequence responsible for binding specificity?

Additional experiments took four duplex oligonucleotides, each comprised of different sequences, none of which contained the CAAT motif. Each was mixed with GFP-C/EBP, applied to the coated plate and subjected to analysis. Again the spectra are similar to Fig. 2A and 2C (data not shown) indicating the sequence of DNA determines the protein binding.

Can sufficient sequence data be obtained to confirm protein identity?

To confirm the identity of GFP-C/EBP, the fifteen most abundant peaks from Fig. 2B were selected to be fragmented by collision induced dissociation (CID) and subsequent TOF/TOF peptide sequence determination. Figures 3 and 4 show the sequence coverage as well as the Mascot score for GFP and C/EBP portions of the fusion protein, respectively; 31.0% of the GFP region and 23.8% of the C/EBP region were sequenced. For the combined sequences derived from either GFP or C/EBP, 28.8% of the sequence was determined.

These results (Figs. 2–4) were compared to the results from Jiang et al. where GFP-C/EBP was purified using separation by EMSA, followed by 2D gel electrophoresis, and then LC-MS/MS (EMSA-3D) [15]. The results were on the same order of magnitude, with 9% sequence coverage using EMSA-3D. The T³ method gives higher sequence coverage with less effort.

Do non-DNA-binding proteins also bind?

An additional experiment investigated the effectiveness of washing to remove non-DNA-binding proteins. 5.5 µg of bovine serum albumin (BSA) was mixed with the EP18 DNA, processed according to the T³ method and analyzed. At our normal laser power setting of 20%, the results were essentially the same as Figures 2A and 2C, where BSA peptides were not detected (data not shown). To further probe the sample and the surface, the laser power was increased to 80%. The results of the increased laser power are shown in Fig. 5 where we start to see PVP fragments. Even with the higher laser power BSA peptides were not found. It should also be noted that there is little to no ion suppression caused by PVP and very little is detected at higher laser power suggesting that it is either firmly bonded to the surface or it does not readily ionize under analysis conditions.

Can the method detect the DNA used?

DNA can also be detected using MALDI-MS as shown in Fig. 6 by applying 1 µL HPA onto any sample that did not previously receive matrix. This allows confirmation of the intactness and identity of the applied oligonucleotide.

Other similar systems have been found useful in detecting low abundant proteins. One such system includes the use of polydimethylsiloxane as the “glue” and titanium dioxide as the “anchor” for phosphopeptide enrichment [16]. Here we have shown that the GFP-C/EBP is only bound in the presence of a specific DNA sequence hence directing proteomic discovery to a specific DNA-binding protein. Enrichment and sequencing a purified protein is hardly notable, but as a proof of concept that DNA-binding proteins can be specifically bound to a surface and successfully analyzed has been adequately modeled.

Can the method be extended to investigate protein identification from spiked nuclear extract?

To extend the model further, a complex mixture was investigated. A nuclear extract from HEK293 cells was spiked with purified GFP-C/EBP and tested using the same stringent parameters as described above. For the entire GFP and C/EBP derived sequences (340 amino acids), 16.1% was sequenced. For the GFP region, 15.1% was sequenced with a Mascot protein score of 166 and the C/EBP region was confidently identified with 17.8% coverage and a Mascot score of 63 (Fig. 7). Compared to the simplified model where purified GFP-C/EBP was identified with 23.8% of the C/EBP region and a Mascot protein score of 159.86 (Fig. 4), the results are still within the realm of acceptance. Thus, even within the milieu of other nuclear proteins, a specific protein can be identified.

An interesting possibility arose from the analysis of the spiked nuclear extract sample is whether only C/EBP sequences from GFP-C/EBP was detected or whether native C/EBP was also detected. Peptides derived from GFP and CEBP are shown in Fig. 7, demonstrating that the fusion protein was pulled out of nuclear extract. Unfortunately, the C/EBP peptide C-terminus sequences analyzed are the same in human (HEK293) and rat (where the C/EBP sequences contained in GFP-C/EBP were derived). Therefore, it is not known whether or not native human C/EBP is also contributing to the results.

Can native C/EBP in nuclear extract also be identified?

This was investigated using a directed proteomic approach on un-spiked nuclear extract. The search was based on peptides identified in the spiked sample as well as peptides not present in the fusion protein but present in the native protein (Fig 8). The identification of native human C/EBP alpha with a Mascot score of 38 and sequence coverage of 24.3% demonstrates that at even the lower concentrations of C/EBP in nuclear extract, proteins of interest can be enriched using their specific DNA-binding partner and identified using the T³ approach. Although the sequence coverage increased for the un-spiked sample, the lower Mascot score can be attributed to the decreased in signal-to-noise ratio. A side-by-side comparison of each of MS/MS spectra obtained from spiked nuclear extract (panel A) and un-spiked nuclear extract (panel B) is shown in Figure 9.

It is expected that T³: Directed Proteomics will produce higher purity of low abundant proteins with minimal sample handling with any DNA sequence. Based on the nanomolar affinity of transcription factors for DNA and the increased speed and accuracy of MALDI-MS this method will provide a new high throughput method with the ability to prepare, screen, and identify transcription factors from 396 samples on a single plate in less than three days. Further automation could significantly reduce analysis time. With this proof of concept, we will next endeavor to increase the signal to noise ratio by improving the coating method via direct polymerization of vinylpyrrolidone onto the plate. We will also explore the use of this technology with longer DNA sequences, such as promoter sequences, to determine if the relevant DNA-binding proteins directing transcription can be identified.

Acknowledgments

FUNDING SOURCES

This work was supported by NIH grant RO1 GM043609, a grant from the National Institute on Minority Health and Health Disparities (G12MD007591) from the National Institutes of Health, NIH/NIGMS MBRS-RISE GM060655, and the National Science Foundation under CHE -1126708.

ABBREVIATIONS

MALDI-TOF-MS	matrix assisted laser desorption - time of flight - mass spectrometry
TF	transcription factors
GFP-C/EBP	green fluorescent protein-CAAT enhancer binding protein
PVP	polyvinylpyrrolidone
CID	collision induced dissociation

REFERENCES

1. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballego RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, et al. The sequence of the human genome. *Science*. 2001; 291:1304–1351. [PubMed: 11181995]
2. Foltmann H, Quadir A. Polyvinylpyrrolidone (PVP) - One of the Most Widely Used Excipients in Pharmaceuticals: An Overview. *Drug Delivery Technology*. 2008; 8:22–27.
3. Liu XL, Sun K, Wu ZQ, Lu JH, Song B, Tong WF, Shi XJ, Chen H. Facile Synthesis of Thermally Stable Poly(N-vinylpyrrolidone)-Modified Gold Surfaces by Surface-Initiated Atom Transfer Radical Polymerization. *Langmuir*. 2012; 28:9451–9459. [PubMed: 22621226]
4. Zhang L, Liang Y, Meng L, Wang C. Characterization of complexation of PVP copolymer with DNA. *Polymers for Advanced Technologies*. 2009; 20:410–415.
5. Peppas NA, Hilt JZ, Khademhosseini A, Langer R. Hydrogels in biology and medicine: From molecular principles to bionanotechnology. *Advanced Materials*. 2006; 18:1345–1360.
6. Nicolaou, M.; Newman, M.; Chang, P. *Polymeric Gene Delivery: Principles and Applications, Use of Poly(N-vinyl pyrrolidone) with Noncondensed Plasmid DNA Formulations for Gene Therapy and Vaccines*. CRC Press; 2004. p. 380-401.
7. Churilina EV, Sukhanov PT, Korenman YI, Il'in AN, Shatalov GV, Bolotov VM. The distribution coefficients of phenol and substituted phenols in the ammonium sulfate-poly-Nvinylpyrrolidone-water system. *Russian Journal of Physical Chemistry A*. 2011; 85:568–572.
8. Matsuda M, Yamamoto K-i, Yakushiji T, Fukuda M, Miyasaka T, Sakai K. Nanotechnological evaluation of protein adsorption on dialysis membrane surface hydrophilized with polyvinylpyrrolidone. *Journal of Membrane Science*. 2008; 310:219–228.
9. Scoble, HA.; Vath, JE.; Yu, W.; Martin, SA. *A Practical Guide to Protein and Peptide Purification for Microsequencing*. Academic Press; 1993. *Mass Spectrometric Strategies for the Structural Characterization of Proteins*; p. 125-153.
10. Smirnov IP, Hall LR, Ross PL, Haff LA. Application of DNA-binding polymers for preparation of DNA for analysis by matrix-assisted laser desorption/ionization mass spectrometry. *Rapid communications in mass spectrometry : RCM*. 2001; 15:1427–1432. [PubMed: 11507755]

11. Kwon Y, Tang K, Cantor C, Koster H, Kang C. DNA sequencing and genotyping by transcriptional synthesis of chain-terminated RNA ladders and MALDI-TOF mass spectrometry. *Nucleic Acids Res.* 2001; 29:E11. [PubMed: 11160913]
12. Larouche K BM, Leclerc S, Guérin SL. Optimization of competitor poly(dI-dC).poly(dI-dC) levels is advised in DNA-protein interaction studies involving enriched nuclear proteins. *BioTechniques.* 1996; 20:439–444. [PubMed: 8679204]
13. Jarrett HW, Taylor WL. Transcription factor-green fluorescent protein chimeric fusion proteins and their use in studies of DNA affinity chromatography. *J Chromatogr A.* 1998; 803:131–139. [PubMed: 9604329]
14. Jiang D, Jia Y, Jarrett HW. Transcription factor proteomics: identification by a novel gel mobility shift-three-dimensional electrophoresis method coupled with southwestern blot and highperformance liquid chromatography-electrospray-mass spectrometry analysis. *J Chromatogr A.* 2011; 1218:7003–7015. [PubMed: 21880322]
15. Jiang D, Jarrett HW, Haskins WE. Methods for proteomic analysis of transcription factors. *J Chromatogr A.* 2009; 1216:6881–6889. [PubMed: 19726046]
16. Chen CJ, Lai CC, Tseng MC, Liu YC, Liu YH, Chiou LW, Tsai FJ. A novel titanium dioxide-polydimethylsiloxane plate for phosphopeptide enrichment and mass spectrometry analysis. *Anal Chim Acta.* 2014; 812:105–113. [PubMed: 24491770]

Figure 1. T³ Schematic

Adsorption of PVP to the MALDI plate was accomplished by immersing the plate in 25 mg PVP/100 mL methanol and incubated undisturbed for 48 hours under N₂. An oligonucleotide containing a sequence of interest was mixed with protein under conditions that encourage DNA-protein complex formation. Then the solution was spotted onto the PVP-coated-target-plate. After incubation, each spot was washed twice with binding buffer to remove any non-specific proteins and then washed once with diH₂O to remove salts present in the binding buffer. Reduction, alkylation and trypsin digestion was carried out on the target plate, then matrix was applied to the digested sample and analysis was accomplished by MALDI-TOF-MS/MS.

Figure 2. Uncoated Plate vs. Coated Plate, with and without DNA

The T^3 technique was evaluated on an uncoated plate (panel A), or a coated plate (panels B and C). In panel B, 10 nM DNA was mixed with the protein, while in panel C DNA was omitted. All panels have a peak at $m/z = 842.5$ which corresponds to trypsin sequence VATVSLPR. In panel B, 13 peptides were identified by MALDI-TOF-MS/MS to arise from GFP-C/EBP. Note: each panel in Fig. 2 is not shown at the same intensity; with the lowest being panel A.

Figure 3. Annotated Spectrum of Green Fluorescent Protein

The GFP region (239 amino acid, N-terminal) received 31% sequence coverage and an ion score of 497.72 by Mascot search engine and SwissProt database. In the upper panel, the peaks with masses labeled corresponds fragments from tryptic peptide DDGNYKTR, matching GFP after CID of $m/z = 1604.81$. In the lower panel, the numbers to the right of the sequences represents the residue number. Grey lettering denotes plasmid-derived sequences, green represents sequences derived from GFP, and black lettering are from C/EBP sequences. Bold and italicized lettering specifies the different sequences determined by MALDI-MS/MS.

Figure 4. Annotated Spectrum of CAAT Enhancer Binding Protein

The C/EBP region (101 amino acid, C-terminal) was identified with 23.8% sequence coverage and ion score of 159.86 by protein database searching with Mascot software and SwissProt database. The upper panel represents a spectra obtained from CID of $m/z = 1622.83$, the peaks with labeled masses denotes fragments from tryptic peptide matching RRVRYENS NK. In the lower panel, the coloring scheme is the same as Fig. 3. Identifications were all found on the C-terminus C/EBP sequences of this chimera.

Figure 5. EP18 and BSA used in Oligo-MALDI-MS

A PVP coated plate with 10 nM EP18 oligonucleotide, 2.8 mg/ml bovine serum albumin, and CHCA matrix with an increased laser intensity of 80%.

Figure 6. Mass Spectrum of EP18 oligonucleotide on PVP coated MALDI plate

Detection of the oligonucleotide allows confirmation of the adsorbed DNA. DNA sequence (5'-GCAGATTGCGCAATCTGA-3') was detected with a +1 charge and $m/z = 5483.8$ using 3-hydroxypicolinic acid as the matrix. The calculated mass for this DNA sequence is 5482.8.

Figure 7. Spectrum of Green Fluorescent Protein and CCAAT Enhancer Binding Protein purified from spiked nuclear extract

The 5 μ l sample spotted on the plate contained 10 nM EP18 oligonucleotide and 5 μ g HEK293 nuclear extract and 0.35 μ g of purified GFP-C/EBP. The GFP region (239 amino acid, N-terminal) was identified with a Mascot protein score of 166 and 15.5% sequence coverage. The C/EBP region (101 amino acid, C-terminal) was identified with 17.8% sequence coverage and a Mascot protein score of 63. Protein database were searched with Mascot software and SwissProt database. In the upper panel, the peaks with green-labeled masses denote peptides derived from GFP and black-labeled masses correspond to C/EBP by MALDI-MS. In the lower panel, the coloring scheme is the same as Fig. 3.

Figure 8. Sequence Coverage of Native CAAT Enhancer Binding Protein Enriched from Nuclear Extract

Human C/EBP alpha was identified with 24.3% sequence coverage and a Mascot protein score of 38 by protein database searching with Mascot software and SwissProt database. Peaks with masses labeled correspond to tryptic peptides matching C/EBP. High sequence coverage and low protein score can be explained by the targeted approach and the spectra collected having low signal to noise.

Figure 9. Comparison of MS/MS Spectra of CAAT Enhancing Binding Protein Enriched from Spiked and Un-Spiked Nuclear Extract

5 μ l samples spotted contained 10 nM EP18 oligonucleotide and either (A) 5 μ g NE spiked with 0.35 μ g GFP-C/EBP or (B) un-spiked NE. Spectra were collected in positive LIFT mode using CHCA matrix. MS/MS of m/z 1127.6 corresponds to sequence RERNNIIVR. The position of fragments in the overall sequence is shown in the scheme at the top of the spectrum.