JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v2i1.81

Optimisation of Downscaled Tandem Affinity Purifications to Identify Core Protein Complexes

Eric B. Haura*[1], Roberto Sacco [2], Jiannong Li[1], André C. Müller[2], Florian Grebien[2], Giulio Superti-Furga[2], Keiryn L. Bennett*[2].

[1] Department of Thoracic Oncology Program, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida, USA; [2] CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria.

Received: 01 December 2011 **Accepted:** 11 April 2012 **Available Online**: 18 April 2012

ABSTRACT

In this study we show that via stable, retroviral-expression of tagged EGFR del (L747-S752 deletion mutant) in the PC9 lung cancer cell line and stable doxycycline-inducible expression of tagged Grb2 using a Flp-mediated recombination HEK293 cell system, the SH-TAP can be downscaled to 5 to 12.5 mg total protein input (equivalent to 0.5 - 1 × 15 cm culture plate or 4 - 8 × 10^6 cells). The major constituents of the EGFR del complex (USB3B, GRB2, ERRFI, HSP7C, GRP78, HSP71) and the Grb2 complex (ARHG5, SOS1, ARG35, CBL, CBLB, PTPRA, SOS2, DYN2, WIPF2, IRS4) were identified. Adjustment of the quantity of digested protein injected into the mass spectrometer reveals that optimisation is required as high quantities of material led to a decrease in protein sequence coverage and the loss of some interacting proteins. This investigation should aid other researchers in performing tandem affinity purifications in general, and in particular, from low quantities of input material.

Keywords: EGFR; Grb2; Orbitrap; TAP; downscale.

Abbreviations

1D-SG-CID, one-dimensional shotgun collision-induced dissociation; **AGC,** automatic gain control; **BPI**, base peak intensity; **DMEM**, Dulbecco's modified eagle medium; **cDNA**, complementary deoxyribonucleic acid; **EGFR**, epidermal growth factor receptor; **ESI-MS,** electrospray ionisation mass spectrometry; **FBS**, foetal bovine serum; **GFP**, green fluorescent protein; **LC-MS**, liquid chromatography mass spectrometry; **PBS**, phosphate-buffered saline; **PCR**, polymerase chain reaction. **RIC**, reconstructed ion chromatogram; **PCT**, unique peptide counts; **SCT**, spectral counts. **SCV**, percent sequence coverage; **SH-TAP,** streptavidin-binding peptide haemagglutinin tandem affinity purificationl.; **TAP-MS**, tandem affinity purification mass spectrometry; **VSV-G**, vesicular stomatitis virus G.

1. Introduction

Tandem affinity purification (TAP) is a generic two-step purification protocol that enables the enrichment and isolation of non-covalent protein complexes under near-physiological conditions. Coupling of TAP to mass spectrometry (TAP-MS) has proven to be a powerful approach to generate large-scale protein-protein interaction networks in yeast [1-4] and mammalian [5, 6] systems. Although TAP-MS in yeast was relatively straightforward, similar studies were hampered in human cells by: (i) the labour-intensive generation of large collections of cell lines expressing epitope -tagged bait proteins; (ii) the low yield of the protein complexes isolated from the cell lines; and (iii) the limited sensitivity of MS-based protein identification. The recent introduction of the Flp-mediated recombination system in HEK-

***Corresponding authors:** *Keiryn L. Bennett, Ph.D.* CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences. Lazarettgasse 14, AKH, BT 25.3. Vienna, 1090, Austria. Phone: +43-1-40160-70010. Fax: +43-1-40160-970000. Email Address: mail: kbennett@cemm.oeaw.ac.at; *Eric B. Haura, M.D.* Department of Thoracic Oncology Program. H. Lee Moffitt Cancer Center and Research Institute. 12902 Magnolia Drive. Tampa, FL, 33612, USA. Phone: +1-(0)813-903-6287. Fax: +1-(0)813-903-6817. Email Address: eric.haura@moffitt.org.

293 cells combined with the streptavidin-binding peptide haemagglutinin (SH)-tag and gel-free one-dimensional liquid chromatography mass spectrometry (LC-MS) analysis of the purified protein complexes [7] overcame most of the shortcomings proffered by the traditional [8] and improved TAP-tag protocols [9].

To yield sufficient material for multiple LC-MS experiments, the current recommendation is to perform an SH-TAP from approximately 3×10^7 HEK293 Flp-In cells (around 50 mg total protein from 5 × 15 cm culture plates) [7]. Successful identification of abundant protein complexes (*e.g.*, the PPP2R2B complex) has been achieved from 4×10^6 HEK293 cells (25% eluate volume per LC-MS injection) [7]. Although considerably lower quantities of cells are required compared to the previous recommendations of $5 \times 10^8 - 1 \times 10^9$ cells [10, 11] and 5×10^7 cells [9], the design of certain experiments and choice of bait protein (*e.g.*, innate immune sensors requiring cell lines that express the necessary modulating proteins) may prohibit the use of HEK293 Flp-In cell lines. Additionally, the Flp-In system is neither amenable to characterisation of protein complexes from cells that are difficult to cultivate (neuronal, immune) nor from primary cells.

Diseases such as cancer are increasingly recognised as the result of aberrantly-functioning protein networks that drive cells toward a diseased state. Affinity purification and mass spectrometry has the ability to map such disease networks and enable the realisation of so-called 'network medicine' [12-14]. Using this methodology, we are currently undertaking a systematic approach to map the protein-protein interaction network driven by somatic and activating mutations of the epidermal growth factor receptor (EGFR) in lung cancer. Activated EGFR can drive numerous downstream signalling pathways responsible for cancer growth and survival [15]. PC9 cells are widely-used as a model system to study EGFR mutations in lung cancer. These cells are robust and easily-cultured, however, the generation of large amounts of cells is associated with a number of costs and increased time requirements. The use of high quantities of cellular material may be restrictive when specialised cells are required. Such cells may take longer periods to grow in culture and/or require specific media and/or growth factor/cytokine supplementation.

As EGFR biology has been extensively investigated, there is a wealth of information available to assist in supporting experimentally-derived data. Thus, the protein is ideal as a case study to evaluate a downscaling of the SH-TAP methodology and examine the lower limit of cells from which the core complex can still be successfully purified and identified by mass spectrometry. C-terminally-tagged SH-EGFR del (L747-S752 deletion mutant) [16] was expressed in PC9 lung cancer cells via retroviral transduction, and the EGFR del core complex enriched via a modified two-step affinity purification procedure [17] (Figure 1). The proteins from the complex were digested with trypsin and the resultant peptides identified by LC-MS. Different percentages of the eluate volume were analysed by LC-MS to ascertain the outcome on the number and nature of the identified EGFR-interacting proteins. Ultimately, this information can assist researchers performing generic TAP experiments from low quantities of protein input material and in particular, TAP from more difficult cell types.

A common misconception with mass spectrometric analyses of biological samples is that the higher the quantity of digested material injected onto an LC-MS system, the more peptides (and thus proteins) will be identified. This is not necessarily the case and there are numerous publications and reviews that support this [18-20] and references therein. Increasing the quantity of injected material does not always provide more information. There are a number of important factors that can influence peptide identification including: (i) ion suppression and reduced ionisation efficiency [21]; (ii) resolving capacity of the chromatographic system [22]; (iii) composition of HPLC mobile phase system [23]; (iv) loading capacity of the pre-column; (v) mass spectrometer automatic gain control (AGC) settings; and (vi) TAP purification background (protein and other molecular groups). These points are discussed in the context of downscaled tandem affinity

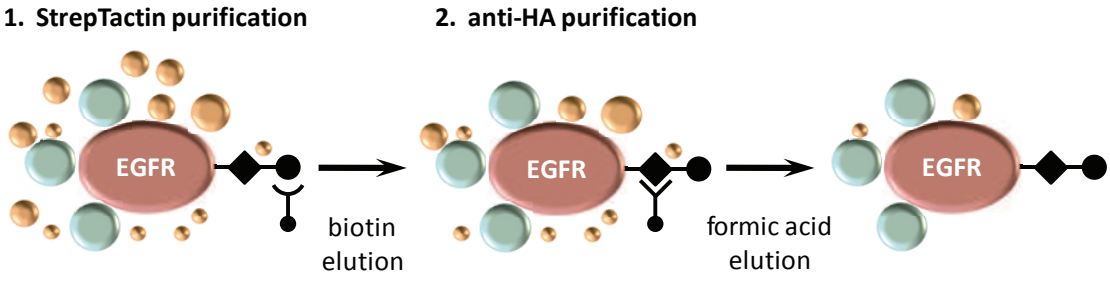

Figure 1. Schematic representation of the tandem affinity purification (TAP) procedure (adapted from Glatter *et al.* [7]). PC9 or HEK293 cells expressing SH-tagged EGFR del or Grb2, respectively, were lysed and purified from total protein extracts using streptavidin sepharose (streptactin beads). After several wash steps, tagged proteins were eluted with 2.5 mM D-biotin for subsequent immunoaffinity purification using anti-HA agarose. After further wash steps, protein complexes were eluted with 100 mM formic acid (pH 2.5) and processed for mass spectrometric analysis.

purification coupled to LC-MS.

The aims of this study were to determine the: (i) optimal quantity of eluate to inject onto the LC-MS system to maximise the number of peptides and proteins identified from the SH-EGFR del purified complex; and (ii) lowest amount of starting input protein required to perform a successful tandem affinity purification. Success is defined as identifying the relevant interacting proteins, as observed from a large-scale purification conducted in our laboratory combined with information from the literature. To confirm that our approach was not restricted to TAP of SH-EGFR complexes in PC9 cells, the study was extended to a second core complex (Grb2) using an alternative system (doxycycline-inducible Flp-In HEK293 cells). The ultimate future prospective of the optimised downscale tandem affinity purification methodology is to discover novel binding partners in specialised cell types. By assessing two well-known and characterised protein complexes at low levels of starting input material, we demonstrate that the discovery of new interactors is attainable in systems where the quantity of available material is limited.

2. Material and Methods

2.1 *Materials*

Iodoacetamide, dithiothreitol, 1 M triethylammonium bicarbonate (TEAB), protease inhibitor cocktail, anti-HA agarose, polybrene, doxycycline, HA7 antibody, 2-propanol, LC-MS grade (SIGMA-Aldrich, St. Louis, MO); trypsin (Promega Corp., Madison, WI); formic acid (HCOOH) (MERCK, Darmstadt, Germany); Strep-Tactin sepharose (IBA TAGnologies, Göttingen, Germany); D-biotin (Alfa Aesar, Karlsruhe, Germany); micro Bio-Spin chromatography columns (Bio-Rad, Hercules, CA); Gateway LR Clonase™ II Enzyme Mix Kit, foetal bovine serum, lipofectamine 2000, pDONR201, pcDNA6/TR, pOG44 (Invitrogen, Carlsbad, CA); methanol, LC-MS grade (Fisher Scientific, Schwerte, Germany).

2.2 *Generation of Stable Cell Lines: Retroviral Infection*

cDNA for EGFR del (L747-S752 deletion mutant) [16] was provided by Dr. William Pao (Vanderbilt University, Nashville, TN). Design of the PCR primers, amplification and pENTR™ TOPO® cloning of EGFR del was performed as previously described [24]. EGFR del was inserted into a gateway-compatible version of pMSCV-C-SH IRES GFP from pENTR™ TOPO® vector by Gateway LR Clonase™ II Enzyme Mix Kit. The retroviral expression clone was verified by DNA sequencing using an Applied Biosystems 3130X1 Genetic analyser (HITACHI) with data analysis performed using Lasergene software V7.2. Phoenix HEK293 cells were obtained from ATCC (catalogue № SD3514, Manassas, VA) and grown in DMEM (4.5 g/L glucose, 2 mM L-glutamine, 50 mg/mL penicillin, 50 mg/mL streptomycin) containing 10% FBS. On day one, 8×10^5 Phoenix cells per well were seeded in a 6-well plate. On day two, cells were transfected with 3 μg VSV-G and 5 μg retroviral plasmids using lipofectamine 2000. Six hours after transfection, the supernatant was replaced with 2 mL DMEM containing 20% FBS and the cells incubated in a 5% CO_2 incubator at 32°C for 48 h. The supernatant (viruses) was collected by centrifugation at 4°C; either stored at -80°C or used immediately to infect the target cells. PC9 cells were maintained in RPMI-1640 medium supplemented with 10% FBS. For retroviral transduction, 2×10^5 cells per well were seeded in a 6-well plate. After overnight incubation, cells were infected with 800 μL virus supernatant plus 6 μg/mL polybrene for 24 h and then supplemented with 4 mL of media per well. Cells were grown continuously until cell sorting. One week after infection, GFP-positive PC9 cells were sorted using a FACSVantage (BD Biosciences, San Diego, CA). GFP positivity and HA expression were assessed by flow cytometry and immunoblot, respectively, before expanding the cells to 10 × 15 cm dishes. When approximately 90% confluent, the EGFR del-tagged PC9 cells were washed with ice-cold PBS containing 1 mM sodium orthovanadate and scraped with a cell lifter on ice. Two cell pellets (each consisting of 5 × 15 cm dishes) were collected in 15 mL conical tubes by centrifugation at 129 × *g* at 4°C and stored at -80°C until required.

2.3 *Generation of Stable Cell Lines: Doxycycline-inducible Expression in HEK293 Flp-In Cells*

To obtain an entry vector, the human full length Grb2 cDNA was cloned by BP Clonase™ recombination into the Gateway®-compatible vector pDONR201. An LR Clonase™ recombination was performed between the entry and destination vector pcDNA5/FRT/TO/SH/GW [7] to generate an expression vector for tetracycline-controlled expression of an N-terminally SH-tagged version of human Grb2. To generate a stable cell line expressing the SH-tagged version of Grb2 in an inducible manner, a Flp-In T-Rex HEK293 cell line stably-transfected with the pcDNA6/TR regulatory vector was cultured in DMEM (containing 10% FBS) supplemented with 100 μg/mL Zeocin™ and 15 μg/mL blasticidin and co-transfected with the expression vector and the pOG44 vector using the FuGENE transfection reagent. Two days after transfection, cells were selected in DMEM supplemented with hygromycin (100 μg/mL) for 2-3 weeks and positive clones were pooled and amplified. After induction with 1 μg/mL doxycycline for 24 h, the inducible expression of the construct was verified by immunoblot.

2.4 *SH-tandem-affinity Purification (adapted from Glatter et al.) [7]*

To minimise experimental and daily variation, *e.g.*, different lot numbers of the streptactin and anti-HA beads, different stock solutions of buffers *etc.*; the 50 mg full pulldowns (n = 2 biochemical replicates) were performed simultaneous-

ly. Note that a biochemical replicate refers to a separate affinity-purification experiment that was performed from the same cell lysate. The series of half pulldowns (20, 12.5, 5 and 2.5 mg) were also performed together (n = 2 biochemical replicates, total 8 purifications) under identical experimental conditions. EGFR del-expressing PC9 cells or Grb2-expressing HEK293 cells were lysed in TNN-HS buffer (50 mM HEPES pH 8.0, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 50 mM NaF, 1.5 mM Na_3VO_4, 1.0 mM PMSF and protease inhibitor cocktail). Insoluble material was removed by centrifugation at 39,443 × *g* for 15 min at 4°C. 200 µL streptactin sepharose (400 µL slurry/pulldown) was transferred to a 14 mL dust-free Falcon tube and washed with 2 × 1 mL TNN-HS buffer. The lysates (50, 20, 12.5, 5 or 2.5 mg) were added to the washed streptactin sepharose and rotated for 20 min at 4°C. The sepharose beads and supernatant were transferred to a micro Bio-Spin chromatography column (Bio-Rad, Hercules, CA) and gravity drained. The sepharose was washed with 4 × 1 mL TNN-HS buffer, and bound proteins eluted from the sepharose with 3 × 300 µL freshly-prepared 2.5 mM D-biotin in TNN-HS buffer into a fresh dust-free 1.5 mL eppendorf tube. 100 µL anti-HA agarose beads (200 µL slurry/pulldown) were transferred into a 1.5 mL eppendorf tube, washed with 1 × 1 mL TNN-HS buffer, centrifuged at 200 × *g* for 1 min at 4°C. The supernatant was removed and the agarose resuspended in 100 µL TNN-HS buffer. The anti-HA agarose beads were added to the biotin eluate and rotated for 1 h at 4°C. The samples were centrifuged at 200 × *g* for 1 min at 4°C and the supernatant removed. The agarose beads were resuspended in 1 mL TNN-HS buffer, the washed beads and buffer loaded into a fresh dust-free micro Bio-Spin column and gravity drained. The anti-HA agarose was washed with 3 × 1 mL TNN-HS buffer and then with 2 × 1 mL TNN-HS buffer consisting of only HEPES, NaCl and EDTA. Retained proteins were eluted from the column with 500 µL 100 mM HCOOH directly into a glass HPLC vial and immediately neutralised with 125 µL 1 M TEABC [17, 25]. Except for the Grb2 half pulldowns, 200 µL (*i.e.*, 33% of the final eluate volume) were routinely removed for silver-stain one-dimensional gel electrophoresis and/or immunoblot analysis as required. The remaining sample was frozen at -20°C until further processing. All volumes in the protocol were halved for the half pulldown series of experiments.

2.5 *Solution Tryptic Digestion*

Proteins from 66% (all EGFR experiments and Grb2 full pulldown) or 100% (Grb2 half pulldowns) of the TEAB neutralised acid eluate were reduced with dithiothreitol, alkylated with iodoacetamide and digested with 2 µg (full pulldown) or 1 µg (half pulldown) modified porcine trypsin. Multiples of 0.25 to 50% of the total eluate volume were desalted and concentrated with customised reversed-phase stage tips [26]. The volume of the eluted sample was reduced to approximately 2 µL in a vacuum centrifuge and reconstituted to 8 µL with 5% formic acid in water and multiples thereof.

2.6 *Anti-haemagglutinin Immunoblot*

The samples were denatured in 4× reducing SDS and boiled for 5 min, separated by 1D-SDS-PAGE and transferred onto a nitrocellulose membrane. The membrane was incubated in blocking solution (5% non-fat milk in PBS/0.05% Tween-20) for 1 h at room temperature, then with murine anti-HA-peroxidase antibody (diluted 1:5000 in blocking solution) for 1 h at room temperature. After thorough washing with PBS/0.05% Tween-20, the antibody signal was visualised by ECL™ Western Blotting Detection Reagents.

2.7 *Liquid Chromatography Mass Spectrometry*

Mass spectrometry was performed on a hybrid LTQ-Orbitrap XL mass spectrometer (ThermoFisher Scientific, Waltham, MA) using the Xcalibur version 2.0.7 coupled to an Agilent 1200 HPLC nanoflow system (dual pump system with one precolumn and one analytical column) (Agilent Biotechnologies, Palo Alto, CA) via a nanoelectrospray ion source using liquid junction (Proxeon, Odense, Denmark). Solvents for LC-MS separation of the digested samples were as follows: solvent A consisted of 0.4% formic acid in water and solvent B consisted of 0.4% formic acid in 70% methanol and 20% 2-propanol. From a thermostatted microautosampler, 8 µL of the tryptic peptide mixture were automatically loaded onto a trap column (Zorbax 300SB-C18 5µm, 5×0.3 mm, Agilent Biotechnologies, Palo Alto, CA) with a binary pump at a flow rate of 45 µL/min. 0.1% TFA was used for loading and washing the pre-column. After washing, the peptides were eluted by back-flushing onto a 16 cm fused silica analytical column with an inner diameter of 50 µm packed with C18 reversed phase material (ReproSil-Pur 120 C18-AQ, 3 µm, Dr. Maisch GmbH, Ammerbuch-Entringen, Germany). The peptides were eluted from the analytical column with a 27 minute gradient ranging from 3 to 30 percent solvent B, followed by a 25 minute gradient from 30 to 70 percent solvent B and, finally, a 7 minute gradient from 70 to 100 percent solvent B at a constant flow rate of 100 nL/min. The analyses were performed in a data-dependent acquisition mode using a top 6 collision-induced dissociation (CID) method. Dynamic exclusion for selected ions was 60 seconds. No lock masses were employed. Maximal ion accumulation time allowed on the LTQ Orbitrap in CID mode was 150 ms for MS^n in the LTQ and 1,000 ms in the C-trap. Automatic gain control was used to prevent overfilling of the ion traps and were set to 5,000 (CID) in MS^n mode for the LTQ, and 10^6 ions for a full FTMS scan. Preview mode for FTMS master scans was enabled to determine the charge state of the intact peptides prior to a scan at 100,000 resolution. Monoisotopic precursor ion selection was enabled and unassigned charge states were not selected for fragmenta-

tion.

2.8 *Data Analysis*

The acquired data were processed with Bioworks V3.3.1 SP1 (ThermoFisher, Manchester, UK). The .dta files were extracted from the .raw files with the extract_msn.exe program. All .dta files were merged into a single peak list file (.mgf) with an internally-developed Perl script. The merged peak list was searched against the human SwissProt database version v57.4 (34,579 sequences, including isoforms as obtained from varsplic.pl) with the search engines MASCOT (v2.2.03, MatrixScience, London, UK) and Phenyx (v2.5.14, GeneBio, Geneva, Switzerland) [27]. Submission to the search engines was via a Perl script that performs an initial search with relatively broad mass tolerances (MASCOT only) on both the precursor and fragment ions (±10 ppm and ±0.6 Da, respectively). High-confidence peptide identifications are used to recalibrate all precursor and fragment ion masses prior to a second search with narrower mass tolerances (±4 ppm and ±0.3 Da). One missed tryptic cleavage site was allowed. Carbamidomethyl cysteine was set as a fixed modification, and oxidised methionine was set as a variable modification. To validate the proteins, MASCOT and Phenyx output files were processed by internally-developed parsers. For MASCOT, two unique peptides with an ion score >18 (plus additional peptides from proteins fulfilling the criteria with an ion score >10) are required. For Phenyx, two unique peptides with a z-score >4.5 and a P-value <0.001 are required (plus additional peptides from proteins fulfilling the criteria with a z-score >3.5 and a P-value <0.001). The validated proteins retrieved by the two algorithms are merged, any spectral conflicts discarded and grouped according to shared peptides. A false positive detection rate (FDR) of <0.25 % and <0.1 % (including the peptides with lower scores) was determined for proteins and peptides, respectively, by applying the same procedure against a reversed database. Comparisons between analytical methods involved comparisons between the corresponding sets of identified proteins. This was achieved by an internally -developed program that simultaneously computes the protein groups in all samples and extracts statistical data such as the number of distinct peptides (PCT), number of spectra (SCT), and sequence coverage (SCV).

3. Results and Discussion

The results reported in this study are divided into three areas. Firstly, using typical starting quantities of cellular lysate, the quantity of eluate required for injection onto the LC -MS system was adjusted to maximise the number of peptides and proteins identified for SH-tagged EGFR del. Secondly, the initial input material was reduced and also the injection quantities varied to determine the minimal protein input required to identify all core components of the EGFR del complex. Finally, the entire strategy was assessed by evaluating another bait in a different cell line to both confirm and generalise our findings.

3.1 *Optimisation of Injection Quantity*

The standard protocol established in our laboratory for the mass spectrometric analysis of streptavidin-binding peptide haemagglutinin (SH)-tagged proteins by one-dimensional gel-free or 'shotgun' collision-induced dissociation (1D-SG-CID), is to inject 3% [17] of the total eluate volume from tryptically-digested, acid-eluted and neutralised protein complexes (50 mg protein input). N-terminally-tagged green fluorescent protein (GFP) and C-terminally-tagged EGFR del were retrovirally-expressed in PC9 cells and analysed accordingly. The protein groups identified from the GFP negative control (TBA1B, TBB5, TBB2C, TBA4B, ADT2, ADT3, ACTB, ACTBL, RS27) were considered as non-specific proteins that were interacting with the SH-tag and/or GFP and were thus removed from the results obtained for EGFR del. The filtered data consisted of 8 proteins (n = 4). EGFR (bait) was the most abundant protein with a maximum sequence coverage (SCV) of 50% (52 unique peptide counts, PCT, from 936 spectral counts, SCT). The proteins identified as part of the core complex of EGFR were: UBS3B, GRB2, ERRFI, HS90A, HS90B, SHC1, and CDC37. A summary of the PCT, SCT and SCV are given in Table 1. UBS3B, GRB2, ERRF1, HSP90, and SHC1 are all *bona fide* interacting proteins with EGFR (www.hprd.org) while CDC37 is a co-chaperone of kinases along with HSP90 [28-30]. GRP78, HSP7C, HSP71 and HS71L were identified in both the EGFR del and GFP experiments, however, the spectral counts were markedly increased in the former thereby indicating that the interaction of EGFR del with some heat shock proteins is enriched and not entirely due to non-specific binding.

Table 1. The core complex proteins identified from the 50 mg protein input SH-EGFR del-TAP (n = 4) following subtraction of proteins identified in the green fluorescent protein (GFP) negative control. PCT (peptide counts); SCT (spectral counts); SCV (percent sequence coverage).

ACCESSION	SWISSPROT	PCT	SCT	SCV
P00533-1	EGFR	52	936	50
Q8TF42	UBS3B	14	77	29
P62993-1	GRB2	8	66	46
Q9UJM3	ERRFI	6	20	20
P07900-1	HS90A	5	19	8
P08238	HS90B	3	8	5
P29353-1	SHC1	3	5	7
Q16543	CDC37	2	3	7

Shown in Figure 2 is the effect on maximum unique PCT when the amount of injected material was increased from 3% of the eluate volume (n = 4) to 5% (n = 4), 10% (n = 4) and 20% (n = 2). The general tendency for the major proteins was a gradual increase in maximum PCT from 3 to 5% eluate injected. UBIQ was identified when a higher quantity of material was analysed, however, CDC37 disappeared with increasing quantity of injected sample. From 5% to 10% to 20%, the tendency was for the maximum PCT to gradually decrease. When 20% of the eluate was injected, however, an additional protein (H90B4) with low PCT was identified. Only two unique peptides were observed for this protein: VILHLKEDQTEYLEER, (shared with HS90A); and IMEESNVK (specific to H90B4). It would appear that although the PCT for all the identified proteins decreased at 20% of the eluate injected, sufficient intensity for the specific peptide from H90B4 was reached to be selected and fragmented in the mass spectrometer.

Unique peptide counts (PCT) and sequence coverage (SCV) (see below, section 3.2) were chosen in this study as the metric for monitoring the effect of increasing the quantity of digested protein that is injected onto the LC-MS system. Contrary to SCT [31], PCT and SCV are by no means quantitative in nature. Nonetheless, both PCT and SCV demonstrated a phenomena that was not observed with SCT. Namely, PCT and SCV have a tendency to decrease with increasing quantities of injected sample, whereas SCT continues to increase. A example of the effect on SCT with increasing quantities of injected material is included in the supplementary section (Supplementary Figure S2).

It has long been established that ESI-MS is a highly-sensitive detection method for polar molecules, however, at elevated concentrations (>10^{-5} M), the approximate linearity of the ESI response is often lost and ion suppression becomes an increasing problem [21]. Shown in Figure 3 is the reconstructed ion chromatogram (RIC) for one of the two unique peptides (LQAEAQQLR, theoretical $[M+2H]^{2+}$ = 528.7935) that were identified from CDC37 across four LC-MS analyses (biochemical replicate 1: 3, 5, 10 and 20% eluate injected). Inset is the *m/z* and isotopic pattern of the ion. The peptide is present in the MS spectrum across all analyses, however, the MSMS spectrum was only matched to the peptide (Figure 3C) from the technical replicate with 3% of the eluate injected (Figure 3B). In all the other analyses, no MSMS spectrum was recorded. The general trend was that the maximal ion intensity of the ^{12}C isotope for LQAEAQQLR actually decreased from 3% to 20% of the eluate injected. More notable is that the isotopic pattern of the ion was also compromised with increasing quantities of injected material. For the spectra shown in Figure 3A and 3D-H, the ion representative of the peptide containing two ^{13}C atoms is not evident. As a consequence, the mass spectrometer did not switch to MSMS. Conversely, the population of ions representative of the peptide with two ^{13}C atoms is only evident in Figure 3B and thus for this analysis only, an MSMS spectrum was recorded (Figure 3C). Thus, it fol-

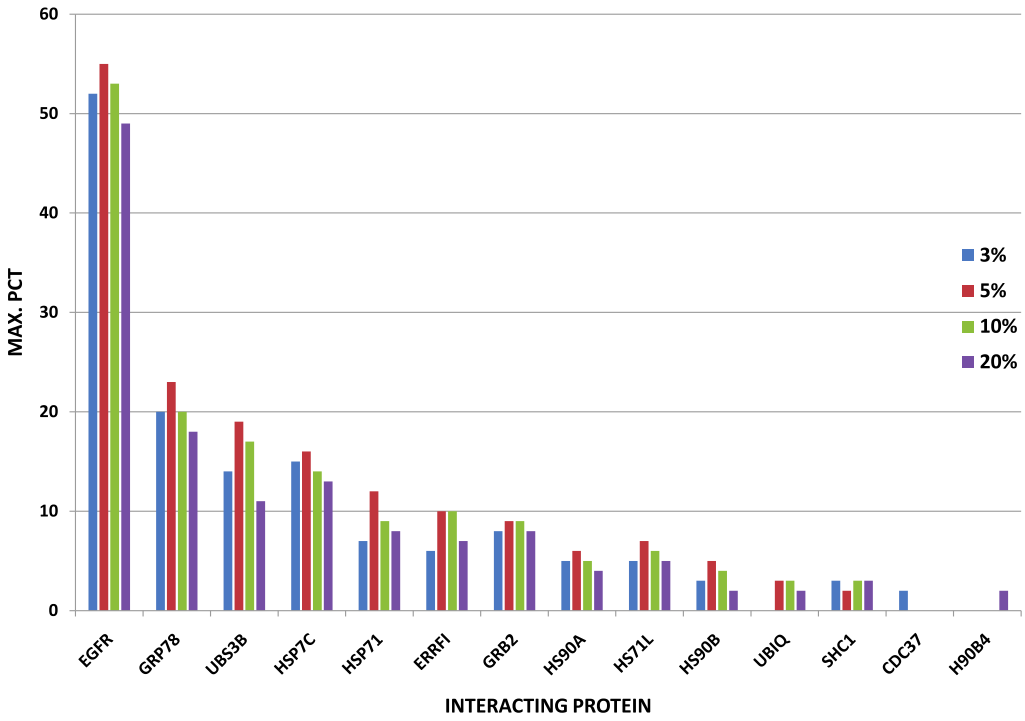

Figure 2. The effect on maximum unique PCT for 50 mg protein input with increasing quantities of injected digested SH-EGFR del-TAP eluate onto the LC-MS system. Blue (3%), red (5%), green (10%) and purple (20%). For the injected quantities of 3, 5 and 10%, n = 4 and for 20%, n = 2.

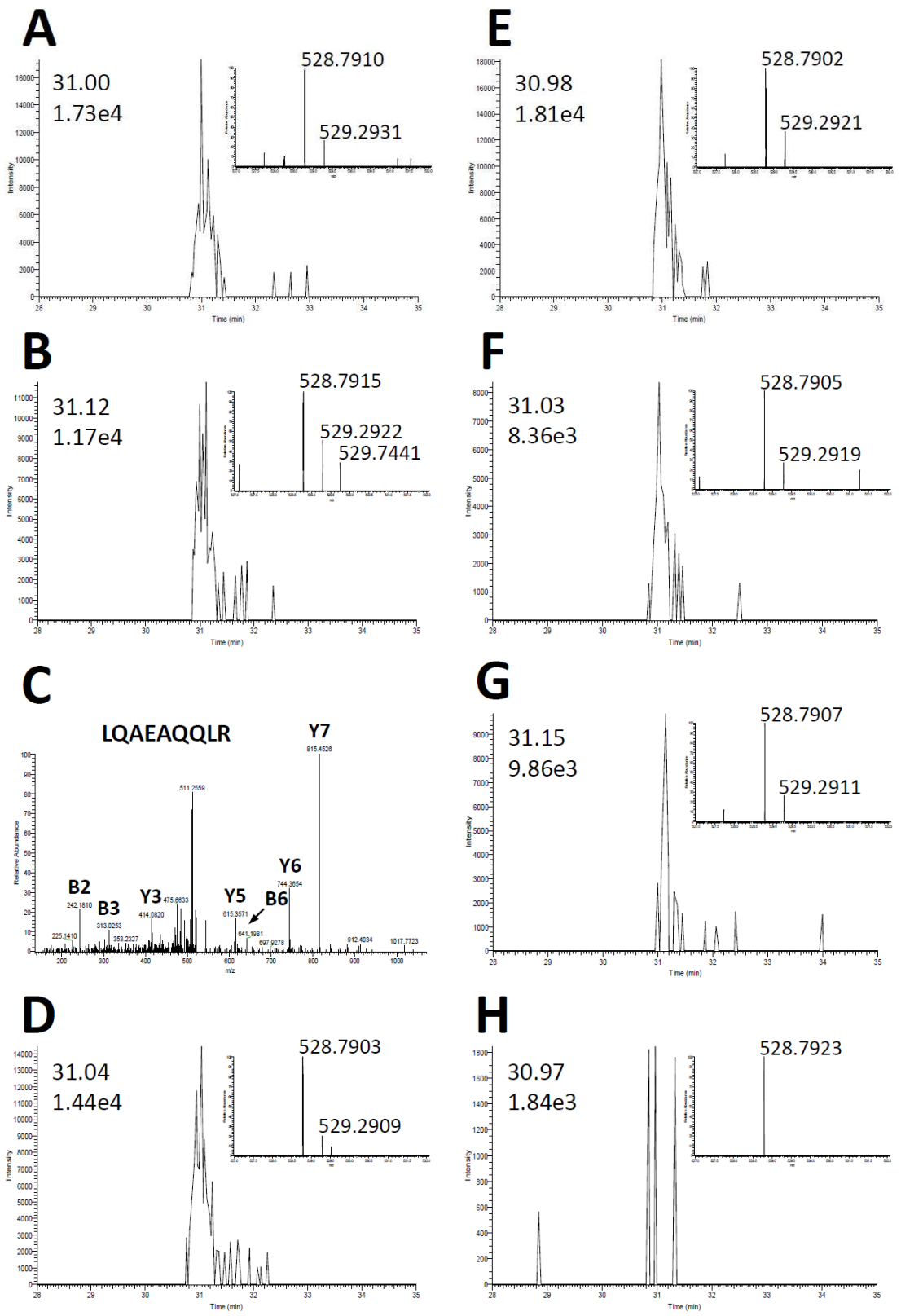

Figure 3. Reconstructed ion chromatogram (RIC) for the peptide LQAEAQQLR (theoretical $[M+2H]^{2+}$ = 528.7935) from CDC37. (A) and (B) 3%; (C) MSMS spectrum from (B); (D) and (E) 5%; (F) and (G) 10% and (H) 20% eluate volume injected. The observed retention time (T_r); and maximum ion intensity are given for each eluate injection percentage. Shown inset is the isotopic distribution and observed *m/z* for the peptide ion.

lows that at a lower concentration of analyte, the apparent ion suppression effect on CDC37 is reduced and the ions are available for selection in the mass spectrometer for fragmentation.

The resolving capacity of the chromatographic system is another factor that can influence the identification of peptides by mass spectrometry. The overall separation efficiency of an analytical column depends on such factors as column length, chosen solvent system, flow rate, gradient profile and type of packing material. Loss of resolution between consecutive chromatographic analyses can be caused by overloading effects or degradation of the reversed-phase material. This can be monitored by changes in peak shape, peak width and altered retention times. The effect of increasing quantities of protein injected on the chromatography is given in Figure 4. Overall the base peak intensity (BPI) chromatograms are relatively weak and the predominant peaks (T) are due to autodigestion products of trypsin. Shown inset is the reconstructed ion chromatogram (RIC) for one of the most intense doubly-charged peptide (NLQEILHGAVR, theoretical $[M+2H]^{2+}$ = 625.3542) from EGFR del (normalised to the maximal intensity of the ion across the four analyses). The peak full-width-at-half-maximum-height (FWHM) was 0.49, 0.42, 0.48 and 0.42 min for 3, 5, 10 and 20% eluate volume injected, respectively. Throughout all analyses in this study, a Zorbax 300SB-C18 5 µm, 5×0.3 mm precolumn was used (see Materials and Methods). As stated by the manufacturer, the loading capacity of this column is in the range of 1 to 10 µg. In chromatography, column overloading effects are observed if there is an increase in peak width of greater than 10%. Since this is not the case, the precolumn capacity is sufficient. The peak profile is unaffected by the increasing amounts of injected material and chromatographic resolution and separation are not compromised. Thus, the quantity of sample injected onto the precolumn does not appear to exceed the specifications.

The composition of HPLC mobile phase system can also play an important role in peptide identification. It is stand-

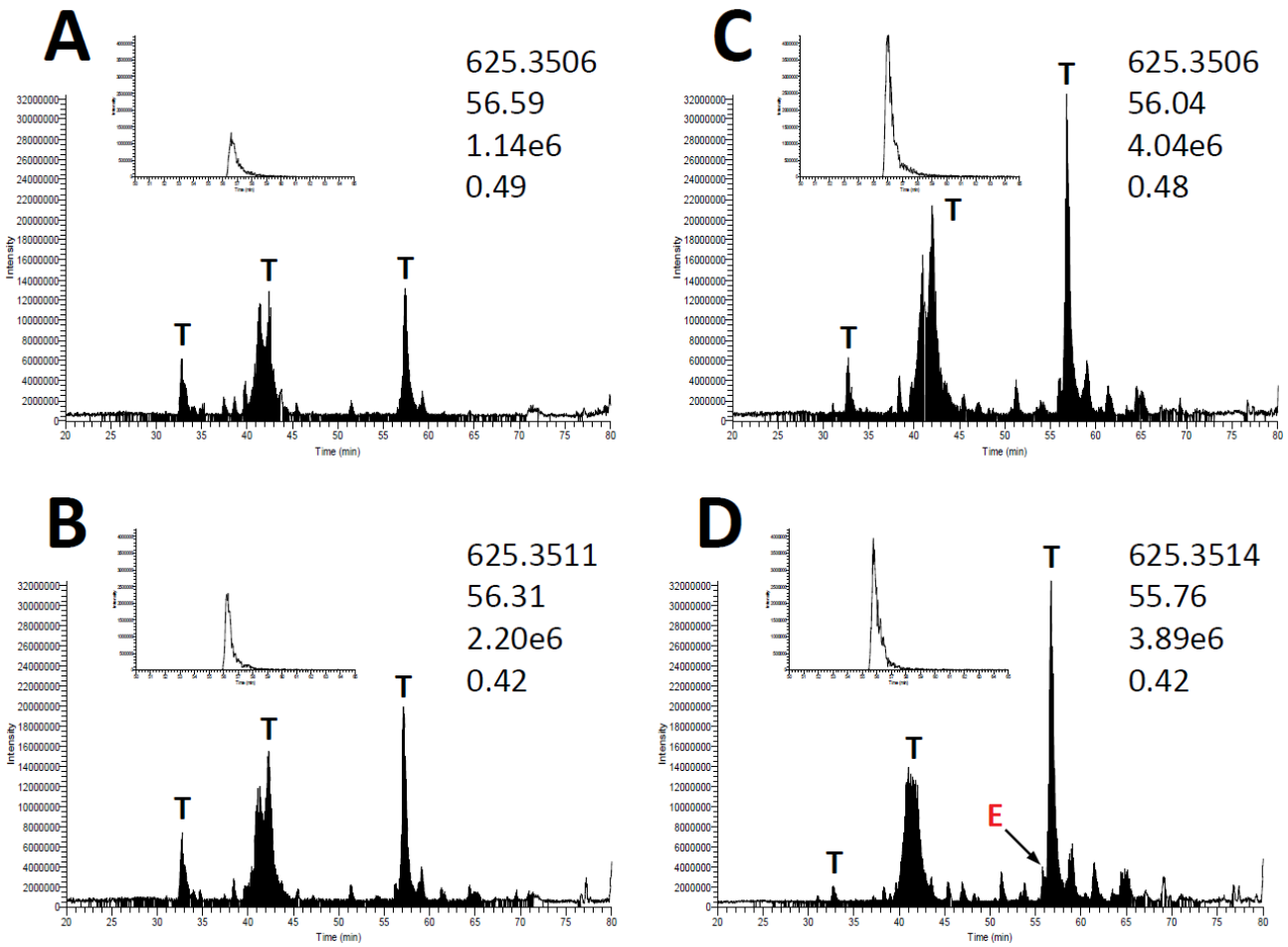

Figure 4. Chromatographic separation of the tryptic digest of the proteins from an SH-EGFR del TAP (50 mg protein input). Shown inset is the reconstructed ion chromatogram (RIC) for NLQEILHGAVR (theoretical $[M+2H]^{2+}$ = 625.3542) from EGFR (normalised to the maximal intensity of the ion across the four analyses). (A) 3; (B) 5; (C) 10 and (D) 20% eluate volume injected. The observed $[M+2H]^{2+}$ retention time (T_r); maximum ion intensity; and peak full-width-at-half-maximum-height (FWHM) are given for each eluate injection percentage. (T), autodigestion products of trypsin; (E) extracted EGFR peptide ion.

ard practice in our laboratory to use a non-traditional mobile phase B containing 70% methanol, 20% 2-propanol (see Materials and Methods). A systematic evaluation of alternative solvent systems compared to the more commonly-used acetonitrile has been performed by our group (unpublished data); however, the data is not included here as this is outside the scope of this manuscript. Nonetheless, in our hands we have found that a 70% methanol, 20% 2-propanol mixture provides superior data compared to other mobile phase combinations (including acetonitrile). Interestingly enough, it is known from the literature that the use of methanol in the mobile phase B actually augments and improves the limits of peptide detection compared to acetonitrile [23].

The AGC settings on an Orbitrap mass spectrometer are used to prevent overfilling of the ion traps. As indicated in the Material and Methods, the AGC was set to 5,000 (CID) in MS^n mode for the LTQ, and 10^6 ions for a full FTMS scan. Naturally, it would have been possible to increase the AGC settings to compensate for the lower quantities of injected material; however, the negative implication is that it takes longer to accumulate more ions. Any gain in adjusting the AGC are counteracted by longer filling times. To maintain consistency during the analyses of the different input quantities of starting material and allow direct comparison between the data sets, the choice was made to limit the number of variables and as such the same AGC settings were maintained throughout the analyses.

The TAP purification background (protein and other molecular groups) can also play a role in the observed decrease in identified peptides and proteins. For the non-specific proteins (TBA1B, TBB5, TBB2C, TBA4B, ADT2, ADT3, ACTB, ACTBL, RS27), the following observations were made (see Table 2). From 3% to 20% of the eluate injected (n = 4) RS27 was identified at the lower injection quantities but was not apparent when 20% of the eluate was analysed. TBB5 and ADT2 were identified across all injection quantities, however, the PCT decreased at 20% injected. ADT3 was only identified when 3% of the eluate was injected. At the higher injection quantities, no unique peptides were identified and the protein was grouped with ADT2. TBB2C, TBA4B, ACTBL and ACTB were only observed at 10% or higher injection percentages. TBA1B was the only non-specific protein that showed a gradual increase in PCT as the amount of eluate injected increased. Overall, the non-specific protein contaminants identified in the SH-EGFR del-TAP exhibited similar traits to the core complex proteins. See Supplementary Table S1 for an overview of the PCT, SCT and SCV for all proteins identified at the different eluate injection volumes. In addition to background non-specifically-interacting proteins, residual non-volatile substances (*e.g.*, salts, buffers, chaotropic agents, stabilisers and detergents) may play a role in the observed phenomena. For example, NP-40 is a pre-requisite component of the cell lysis buffer. The detergent is necessary to not only solubilise membrane and membrane-associated proteins, but also to maintain the structural integrity of the protein-protein complexes prior to

Table 2. The non-specific interacting proteins identified from the SH-EGFR del-TAP (50 mg protein input) with increasing quantities of injected eluate onto the LC-MS system. 3, 5 and 10%, n = 4 and 20%, n = 2. PCT (peptide counts); SCT (spectral counts); SCV (percent sequence coverage).

ACCESSION CODE	SWISSPROT IDENTIFIER	3%		5%		10%		20%									
		PCT		SCT		SCV		PCT		SCT		SCV		PCT		SCT		SCV
P68363	TBA1B	6		18		23		6		24		20		8		37		26		8		17		27	
P07437	TBB5	3		4		9		7		23		19		7		32		21		6		21		19	
P68371	TBB2C	0		0		0		0		0		0		6		28		18		5		15		16	
Q9H853	TBA4B	0		0		0		0		0		0		0		0		0		2		3		4	
P05141	ADT2	3		16		12		4		22		15		4		37		15		5		23		19	
P12236	ADT3	3		9		12		0		0		0		0		0		0		0		0		0	
P60709	ACTB	0		0		0		2		3		9		2		6		8		2		2		3	
Q562R1	ACTBL	0		0		0		0		0		0		0		0		0		2		2		9	
P42677	RS27	2		7		23		2		10		23		2		5		23		0		0		0	

enrichment via TAP. Although the methodology is well-established in our laboratory and particular care is taken with the washing steps to remove detergent (and other non-volatile material) prior to elution with acid, it is feasible that residual detergent (and other components) remain and are concentrated as more of the eluate is injected onto the LC-MS. The negative repercussions are that a very small proportion of the detergent can also lead to reduced peptide intensity as detergents are a scavenger of charge [32]. Additionally, detergents also bind strongly to reversed-phase material which ultimately compromises peak shape and column performance.

When undertaking a TAP-MS experiment, it is important to bear in mind that there are a number of critical factors that can markedly influence the data obtained. These elements play a central role in the quality of the end result. Thus, for novice researchers embarking into protein-protein interactomics, this manuscript serves as a guide to performing TAP and offers plausible explanations for observed anomalies.

3.2 *Optimisation of Protein Input Quantity*

A series of half pulldowns was performed to determine the lowest quantity of input protein from which the major constituents of the EGFR del core complex could be identified. The total input protein quantities were: 20, 12.5, 5 and 2.5 mg and the inject quantities ranged between 0.25 to 20% (Figure 5).

At 20 mg protein input (Figure 5A) and 0.25% of the eluate injected (n = 4), only EGFR del was identified plus 2 peptides from UBS3B. Increasing the percentage of injected eluate to 0.5, 2, 3 and 6% led to the identification of GRP78, GRB2, HSP7C, HS71L (0.5%); HSP71, HS90A (2%); ERRFI (3%), and HS90B (6%), respectively. A further increase in injection quantity to 10% (n = 2) did not lead to the identification of any additional interacting proteins. As observed for the full pulldown experiments, increasing the quantity of injected material decreased the maximum sequence coverage (SCV) of all the identified proteins. When the protein input was 12.5 mg (Figure 5B) and 3% of the eluate volume was injected (n = 4), EGFR del was identified with a maximum SCV of 31% (31 PCT from 321 SCT). GRP78, GRB2, HSP7C, UBS3B and ERRFI were also observed. No specific peptides from HS71L and HSP71 were identified but the proteins are grouped with HSP7C. Increasing the percentage of injected eluate volume to 6% resulted in the identification of HS90A (plus HS90B grouped with HS90A). A further increase, however, in the injected eluate volume to 10 and 15% did not lead to the identification of additional interacting proteins. Conversely, HS90A was no longer apparent. Concurrent with this observation was the plateauing of the maximum SCV for EGFR del and the core complex proteins.

At 5 mg input material (equivalent to approximately 1 × 10 cm plate or 4 × 10^6 cells, Figure 5C) and 15% of the eluate injected (n = 4), EGFR, GRP78, GRB2, HSP7C, UBS3B and

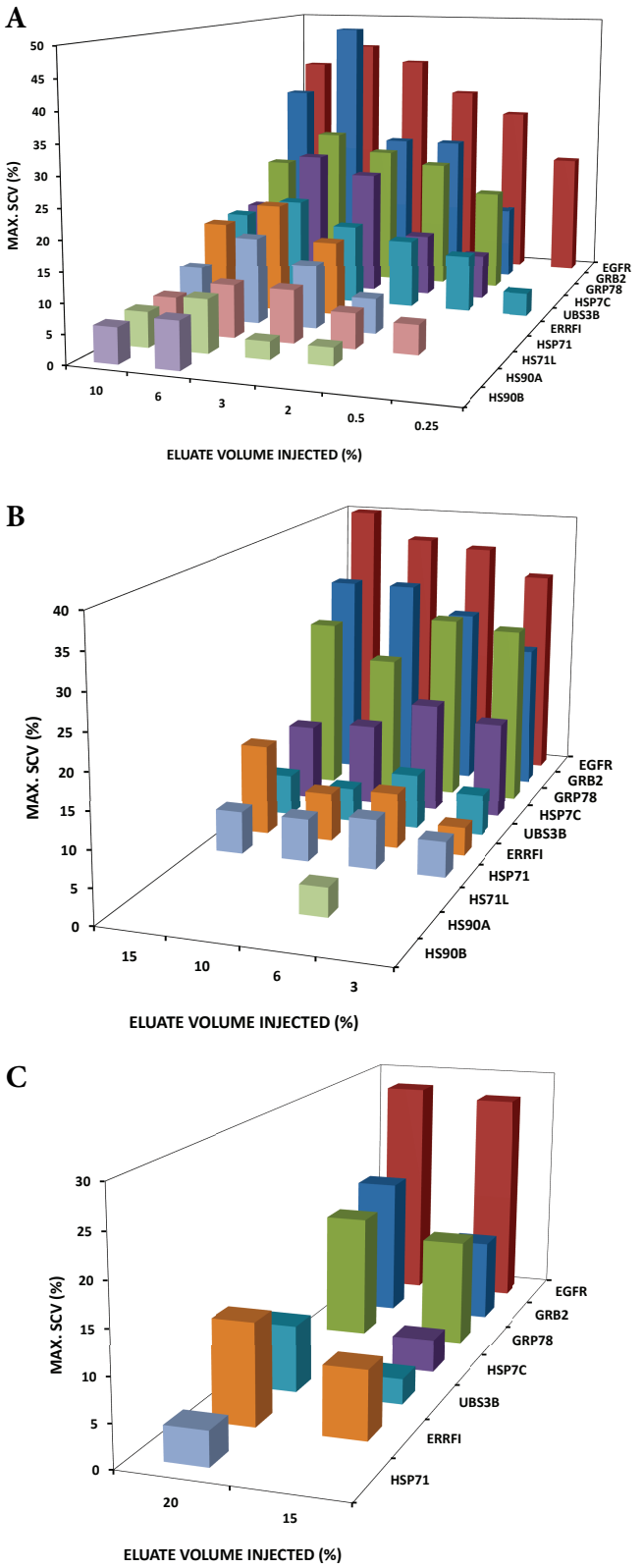

Figure 5. Maximum sequence coverage (SCV) obtained for proteins from the EGFR del core complex at variable total protein input quantities (A) 20 mg; (B) 12.5 mg; and (C) 5 mg. Quantity of injected material (% eluate volume) ranged from 0.25 to 20% per LC-MS analysis.

ERRFI were identified. Increasing the quantity of injected eluate volume to 20% resulted in the appearance of HSP71 but with the concomitant disappearance of HSP7C. Neither HS90A nor HS90B were identified. Overall, from the 5 mg protein input experiments, 6 interacting proteins (plus EGFR) were observed. When the total protein input was decreased to 2.5 mg and 33% of the eluate volume injected per technical replicate (n = 4), only EGFR (20% max. SCV, 17 PCT from 99 SCT) and two non-specifically-interacting proteins (GFAP, S10A9) were identified. Therefore this low quantity of input material is insufficient to purify/observe the EGFR del complex and thus provides a lower cut-off value for a successful protein complex purification from retrovirally-transduced cells using the instrumentation and workflow described in this study. See Supplementary Table S2 for an overview of the PCT, SCT and SCV for all proteins identified at the different protein input quantities and eluate injection volumes.

Across the entire half pulldown series of experiments, 9 EGFR del-interacting proteins were identified in common with the data obtained from the full pulldown series. Only the proteins identified with low PCT were absent (SHC1, CDC37, UBIQ, H90B4). Interestingly, when 6% of the eluate volume was injected from the 12.5 mg protein input sample, basically all the major EGFR-interacting proteins were identified. No specific peptides from HS71L and HS90B were selected for fragmentation by mass spectrometry but the proteins are grouped with HSP71 and HS90A, respectively. The data is comparable with the 3% of the eluate volume from the 50 mg protein input sample. Thus, from four times less input material and only double the injection volume, the same results were obtained.

As for the full pulldown series of experiments, there was also evidence of a decrease in PCT and SCV with a concomitant loss of certain proteins with increasing quantities of injected material. When 10% of the digested eluate from 20 mg protein input sample was analysed by LC-MS, there was a decrease in maximum SCV for all identified proteins compared to the data obtained when 6% of the same eluate was analysed (Figure 5A). The effects were not as pronounced for the 12.5 mg protein input (Figure 5B) and not apparent with the 5 mg protein input (Figure 5C).

For all of the purifications performed with EGFR del, the data was filtered to remove single peptide identifications (Material and Methods, section 2.8). Considering that the level of protein input in these experiments was very low, the discarded data was assessed for know interactors of EGFR del. Overall, the majority of the removed proteins were abundant, non-specifically interacting proteins (*e.g.,* tubulin, ribosomal *etc*), however CDC37 (50, 25 and 12.5 mg input material), HS90A (12.5 mg) and UBS3B (2.5 mg) were also present as single peptide identifications. The inclusion of such peptides could be beneficial, particularly in the discovery of novel interactors.

Comparison of the mass spectrometric data with the anti-HA immunoblot for SH-tagged EGFR del revealed some surprising findings. As observed in Supplementary Figure S1, the levels of the SH-tagged EGFR del following the second step of the purification are extremely low, and at 5 and 2.5 mg protein input (Figure S1G and I) the bait protein is below the detection limit of the anti-HA antibody. Although the immunoblot was very weak for the 12.5 mg protein input, the same amount of eluate volume (6%) analysed by LC-MS unequivocally identified all the core proteins of the EGFR del complex. The whole cell lysate (100 µg) and the biotin eluate from first step of the TAP are also given for comparison. As reported by Glatter *et al.* [7], elution from the streptavidin column with biotin is highly-efficient (>90% recovery) with negligible bait protein remaining on the streptavidin beads following elution with Laemmli buffer. From the signal intensity of the anti-HA immunoblot of the final eluate, these workers also estimated that the overall yield of the TAP was 30–40% of bait protein present in the cell lysate. Based on Supplementary Figure S1, the data presented here reflects some of the findings of Glatter *et al.* [7]. This is particularly evident for the one-step biotin elution for the 20 and 12.5 mg protein inputs. The yield of the second-step formic acid elution, however, appears to be slightly lower than that reported. Nonetheless, regardless of the yield determined by immunoblot techniques, the fact remains that it is feasible to purify and identify non-covalently-interacting protein complexes from low input quantities of protein. This alone is a substantial achievement and opens the possibility of exploring and charting molecular pathways from primary cells and/or cells that are particularly difficult to cultivate.

3.3 *Grb2 Core Complex*

To confirm the observations made with downscaling the EGFR del SH-TAP, a second protein (Grb2) was cloned with an N-terminal-HA tag via the alternative doxycycline-inducible Flp-In system in HEK293 cells [7]. Tandem affinity purifications were performed with 50, 12.5 and 5 mg protein input material. Following subtraction of the protein groups observed in the GFP negative control (HSP71, HSP7C), 12 proteins remained in the purification from 50 mg total protein input and 3% of the eluate injected (n = 2). Grb2 (bait) was the most abundant protein with a maximum sequence coverage (SCV) of 92% (24 unique peptide counts, PCT, from 295 spectral counts, SCT). The proteins identified as part of the core complex of Grb2 were: ARHG5, SOS1, DYN2, ARG35, SOS2, CBLB, CBL, DYN1, WIPF2, IRS4, and PTPRA. A summary of the PCT, SCT and SCV are given in Table 3A. Increasing the quantity of injected eluate to 5% (n = 2) resulted in the identification of the same proteins, plus one additional non-specific interactor (HNRPK). At 10% (n = 2) and 20% of the eluate injected (n = 1), a total of 6 further proteins were identified, albeit with low PCT, SCT and SCV. These were: WASL, ERRFI, DYN3 (interactors), ACACA (SH-TAP purification contaminant); SNX18 and FA59B (likely non-specific binders).

Figure 6 shows the effect on maximum unique PCT when

Table 3. The core complex proteins identified from the SH-Grb2-TAP. (A) 50 mg (n = 2); and (B) 12.5 mg (n = 2) protein input following subtraction of proteins identified in the green fluorescent protein (GFP) negative control. PCT (peptide counts); SCT (spectral counts); SCV (percent sequence coverage). n.i.; not identified.

ACCESSION CODE	SWISSPROT IDENTIFIER	A			B		
		PCT	SCT	SCV	PCT	SCT	SCV
P62993-1	GRB2	24	295	92	17	149	79
Q12774-1	ARHG5	45	95	32	35	66	27
Q07889	SOS1	39	77	29	24	42	19
P50570-2	DYN2	15	23	15	2	2	2
A5YM69	ARG35	9	20	20	8	15	20
Q07890	SOS2	11	18	9	5	7	5
Q13191-1	CBLB	9	17	9	6	12	7
P22681	CBL	11	18	15	8	13	13
Q05193-2	DYN1	8	11	7	n.i.	n.i.	n.i.
Q8TF74-1/A6NGB9	WIPF2/3	3	7	11	2	2	6
O014654	IRS4	4	6	4	2	2	2
P18433-1	PTPRA	5	8	7	6	11	10

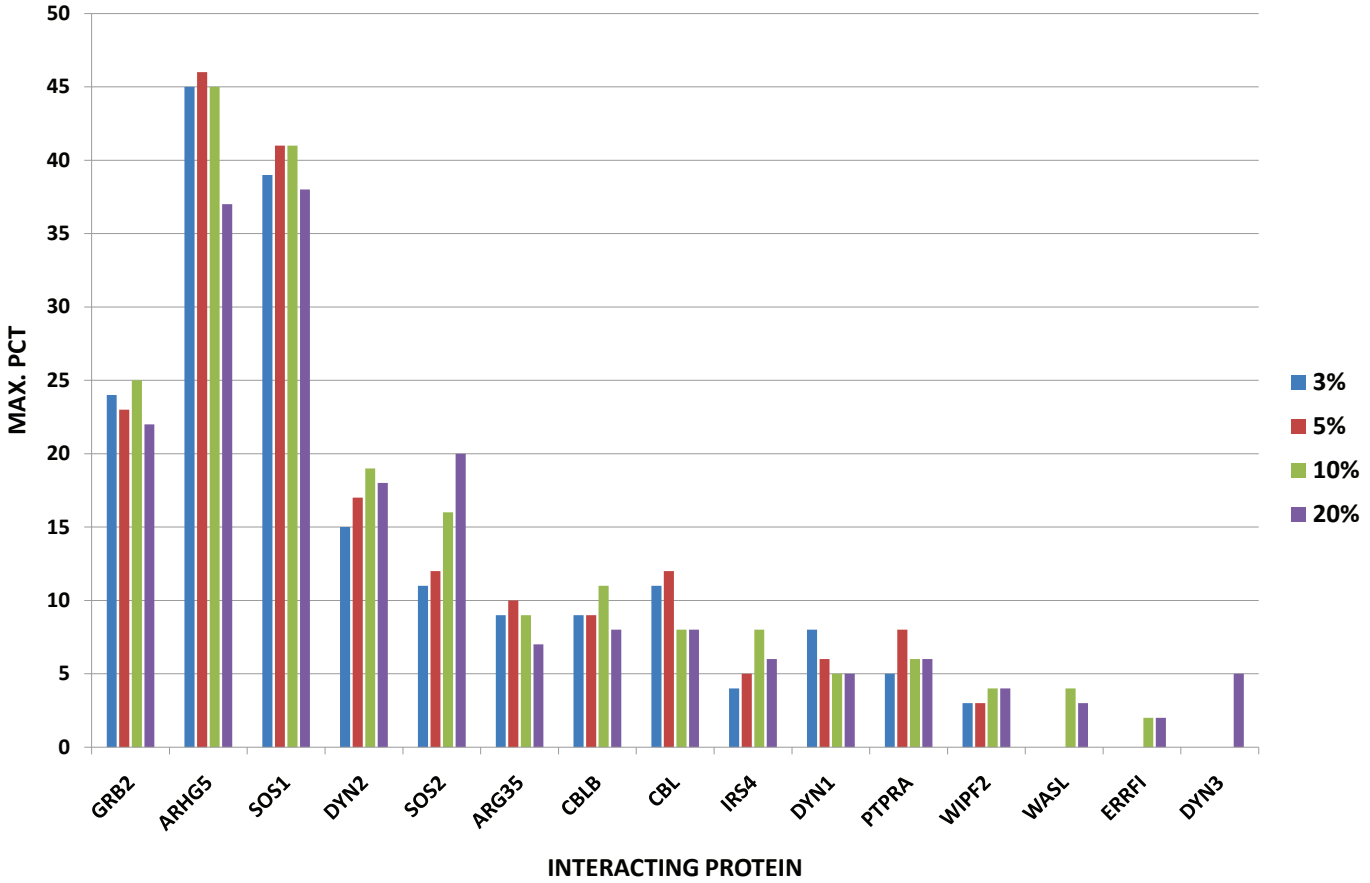

Figure 6. The effect on maximum unique PCT for 50 mg protein input with increasing quantities of injected digested SH-Grb2-TAP eluate onto the LC-MS system. Blue (3%), red (5%), green (10%) and purple (20%). For the injected quantities of 3, 5 and 10%, n = 2 and for 20%, n = 1.

the amount of injected material was increased from 3% of the eluate volume (n = 2) to 5% (n = 2), 10% (n = 2) and 20% (n = 1). A similar pattern to the EGFR experiments was obtained (Figure 2). That is, the general tendency for the major proteins was a gradual increase in maximum PCT from 3 to 5% eluate injected (ARHG5, SOS1, ARG35, CBL, PTPRA) or from 3 to 5 to 10% (DYN2, CBLB, IRS4, WIPF2) and then a decrease in PCT as more eluate was injected. DYN 1 had a maximum number of unique PCT at 3% injected, whilst SOS2 was the only protein that steadily increased in the number of unique PCT with increasing percentage of material injected. The low-abundance proteins WASL and ERRFI were only identified from 10% of the eluate injected and above. DYN3 was only observed at the maximum injected percentage.

Decreasing the protein input quantity to 12.5 mg and injection of 50% of the eluate, resulted in the identification of 17 proteins (n = 2). After removal of the proteins observed in the negative control (HSP71, HSP7C, GRP78) and apparent non-specific binders (H14, NFH, ACTA), 11 of the 15 proteins (ARHG5, SOS1, ARG35, CBL, CBLB, PTPRA, SOS2, WIPF3, DYN2, and IRS4) observed in the 50 mg full pull-down experiments remained. Grb2 had a maximum SCV of 79% (17 PCT from 149 SCT). The PCT, SCT and SCV for the identified proteins are summarised in Table 3B. When the protein input quantity was decreased further to 5 mg, only Grb2 (59% max. SCV, 15 PCT from 79 SCT), HSP71 and NFH were identified, thus providing the lower cut-off value for this protein in the HEK293 Flp-In cells. A summary of the proteins identified is given in Supplementary Table S3. As for EGFR del, the single peptide identifications were also assessed. Two known interactors of Grb2, WIPF3 [33] and M4K5 [34] also became evident. Overall, the data generated with the Grb2 experiments were reflective of the results obtained from the EGFR experiments.

4. Concluding Remarks

The ability to scale down the SH-TAP to low quantities of input material opens numerous possibilities and potential applications and shows that it is feasible to perform an SH-TAP experiment from a single 10 cm or 15 cm plate of cells (equivalent to 4 - 8 × 10^6 cells). From this study, we believe that it is now possible to map protein-protein interaction networks from cells that are difficult to cultivate in large quantities (*i.e.*, primary cells) or cells from which it is not feasible to generate a stable cell line without using viral expression systems or the Flp-FRT recombination system. This may allow elucidation of biological networks in a larger panel of cell types and conditions, *e.g.*, detailed experiments with dynamic network perturbations with drugs or RNA interference molecules. Identification of these networks would be of considerable benefit in delineating disease-related signalling pathways in aberrant tissues obtained from mouse models and/or patient material.

Refinement of the liquid chromatography mass spectrometry system (*e.g.*, lower flow rates, smaller diameter analytical columns, submicron ESI needles) and sample preparation techniques (*e.g.*, fractionation at the peptide or protein level, removal of abundant bait protein) could further increase both sensitivity and dynamic range. This then has the potential to allow the identification of protein complexes from even lower quantities of protein material. To improve TAP-MS analyses based on the results from this study, our recommendation to other researchers that are embarking into the field of protein-protein interactomics is to find the balance between the (i) quantity of starting material for the tandem affinity purification; and (ii) the quantity of eluate injected onto the LC-MS system. As each protein is most certainly different, the results from this manuscript act as a guide for other researchers endeavouring to establish TAP in their laboratory or to perform TAP on proteins from unusual cell lines with low quantities of input material. Additionally, we put forward several plausible explanations to account for the phenomena observed in our study. Namely, that increasing quantities of injected material eventually leads to a decrease in the number of proteins identified. Researchers endeavouring to enter this field of research are thus aware of the anomalies that can occur when working at this low level of sensitivity and are therefore fully-prepared for interpretation of the data.

5. Supplementary material

Supplementary material regarding this manuscript is online available in the web page of JIOMICS.
http://www.jiomics.com/index.php/jio/rt/suppFiles/81

Acknowledgements

The authors would like to thank William Pao for providing EGFR del cDNA, Matthias Gstaiger for supplying the original SH constructs and providing valuable advice in establishing SH-TAP protocols at CeMM, Fumi Kinose for assistance with cell culture, Christian Knoll and Elena Rudashevskaya for helpful discussion and input. Work in our laboratory is supported by the Austrian Academy of Sciences, the Austrian Federal Ministry for Science and Research (Gen-Au projects, APP and BIN) and by the Austrian Science Fund FWF (P22282-B11, FG). This study was also funded in part by NIH (5P50CA119997, EBH).

Competing interests

The authors declare that there are no competing interests.

References

1. Y. Ho, A. Gruhler, A. Heilbut, G. D. Bader, L. Moore, S. L. Adams, A. Millar, P. Taylor, K. Bennett, K. Boutilier, L. Yang,

C. Wolting, I. Donaldson, S. Schandorff, J. Shewnarane, M. Vo, J. Taggart, M. Goudreault, B. Muskat, C. Alfarano, D. Dewar, Z. Lin, K. Michalickova, A. R. Willems, H. Sassi, P. A. Nielsen, K. J. Rasmussen, J. R. Andersen, L. E. Johansen, L. H. Hansen, H. Jespersen, A. Podtelejnikov, E. Nielsen, J. Crawford, V. Poulsen, B. D. Sorensen, J. Matthiesen, R. C. Hendrickson, F. Gleeson, T. Pawson, M. F. Moran, D. Durocher, M. Mann, C. W. Hogue, D. Figeys, M. Tyers, Nature 415 (2002) 180-183.

2. A. C. Gavin, M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J. M. Rick, A. M. Michon, C. M. Cruciat, M. Remor, C. Hofert, M. Schelder, M. Brajenovic, H. Ruffner, A. Merino, K. Klein, M. Hudak, D. Dickson, T. Rudi, V. Gnau, A. Bauch, S. Bastuck, B. Huhse, C. Leutwein, M. A. Heurtier, R. R. Copley, A. Edelmann, E. Querfurth, V. Rybin, G. Drewes, M. Raida, T. Bouwmeester, P. Bork, B. Seraphin, B. Kuster, G. Neubauer, G. Superti-Furga, Nature 415 (2002) 141-147.
3. A. C. Gavin, P. Aloy, P. Grandi, R. Krause, M. Boesche, M. Marzioch, C. Rau, L. J. Jensen, S. Bastuck, B. Dumpelfeld, A. Edelmann, M. A. Heurtier, V. Hoffman, C. Hoefert, K. Klein, M. Hudak, A. M. Michon, M. Schelder, M. Schirle, M. Remor, T. Rudi, S. Hooper, A. Bauer, T. Bouwmeester, G. Casari, G. Drewes, G. Neubauer, J. M. Rick, B. Kuster, P. Bork, R. B. Russell, G. Superti-Furga, Nature 440 (2006) 631-636.
4. N. J. Krogan, G. Cagney, H. Yu, G. Zhong, X. Guo, A. Ignatchenko, J. Li, S. Pu, N. Datta, A. P. Tikuisis, T. Punna, J. M. Peregrin-Alvarez, M. Shales, X. Zhang, M. Davey, M. D. Robinson, A. Paccanaro, J. E. Bray, A. Sheung, B. Beattie, D. P. Richards, V. Canadien, A. Lalev, F. Mena, P. Wong, A. Starostine, M. M. Canete, J. Vlasblom, S. Wu, C. Orsi, S. R. Collins, S. Chandran, R. Haw, J. J. Rilstone, K. Gandi, N. J. Thompson, G. Musso, P. St Onge, S. Ghanny, M. H. Lam, G. Butland, A. M. Altaf-Ul, S. Kanaya, A. Shilatifard, E. O'Shea, J. S. Weissman, C. J. Ingles, T. R. Hughes, J. Parkinson, M. Gerstein, S. J. Wodak, A. Emili, J. F. Greenblatt, Nature 440 (2006) 637-643.
5. T. Bouwmeester, A. Bauch, H. Ruffner, P. O. Angrand, G. Bergamini, K. Croughton, C. Cruciat, D. Eberhard, J. Gagneur, S. Ghidelli, C. Hopf, B. Huhse, R. Mangano, A. M. Michon, M. Schirle, J. Schlegl, M. Schwab, M. A. Stein, A. Bauer, G. Casari, G. Drewes, A. C. Gavin, D. B. Jackson, G. Joberty, G. Neubauer, J. Rick, B. Kuster, G. Superti-Furga, Nat Cell Biol 6 (2004) 97-105.
6. C. Behrends, M. E. Sowa, S. P. Gygi, J. W. Harper, Nature 466 (2010) 68-76.
7. T. Glatter, A. Wepf, R. Aebersold, M. Gstaiger, Mol Syst Biol 5 (2009) 237.
8. G. Rigaut, A. Shevchenko, B. Rutz, M. Wilm, M. Mann, B. Seraphin, Nat Biotechnol 17 (1999) 1030-1032.
9. T. Burckstummer, K. L. Bennett, A. Preradovic, G. Schutze, O. Hantschel, G. Superti-Furga, A. Bauch, Nat Methods 3 (2006) 1013-1019.
10. A. K. Al-Hakim, O. Goransson, M. Deak, R. Toth, D. G. Campbell, N. A. Morrice, A. R. Prescott, D. R. Alessi, J Cell Sci 118 (2005) 5661-5673.
11. J. Gregan, C. G. Riedel, M. Petronczki, L. Cipak, C. Rumpf, I. Poser, F. Buchholz, K. Mechtler, K. Nasmyth, Nat Protoc 2 (2007) 1145-1151.
12. A. L. Barabasi, N Engl J Med 357 (2007) 404-407.
13. T. Pawson, R. Linding, FEBS Lett 582 (2008) 1266-1270.
14. A. L. Barabasi, N. Gulbahce, J. Loscalzo, Nat Rev Genet 12 (2011) 56-68.
15. T. Yoshida, G. Zhang, E. B. Haura, Biochem Pharmacol 80 (2010) 613-623.
16. W. Pao, V. Miller, M. Zakowski, J. Doherty, K. Politi, I. Sarkaria, B. Singh, R. Heelan, V. Rusch, L. Fulton, E. Mardis, D. Kupfer, R. Wilson, M. Kris, H. Varmus, Proc Natl Acad Sci U S A 101 (2004) 13306-13311.
17. E. B. Haura, A. Muller, F. P. Breitwieser, J. Li, F. Grebien, J. Colinge, K. L. Bennett, J Proteome Res 10 (2011) 182-190.
18. K. L. Bennett, L. A. Hick, S. V. Smith, R. J. W. Truscott, M. M. Sheil, J. Mass Spectrom. 30 (1995) 769-771.
19. T. M. Annesley, Clin Chem 49 (2003) 1041-1044.
20. J.-P. Antignac, K. d. Wasch, F. Monteau, H. D. Brabander, F. Andre, B. L. Bizec, Analytica Chimica Acta 529 (2005) 129-136.
21. P. Kebarle, L. Tang, Anal. Chem. 65 (1993) 972A.
22. E. Grushka, Analytical Chemistry 42 (1970) 1142-&.
23. F. Giorgianni, A. Cappiello, S. Beranova-Giorgianni, P. Palma, H. Trufelli, D. M. Desiderio, Anal Chem 76 (2004) 7028-7038.
24. J. Li, U. Rix, B. Fang, Y. Bai, A. Edwards, J. Colinge, K. L. Bennett, J. Gao, L. Song, S. Eschrich, G. Superti-Furga, J. Koomen, E. B. Haura, Nat Chem Biol 6 (2010) 291-299.
25. N. V. Fernbach, M. Planyavsky, A. Muller, F. P. Breitwieser, J. Colinge, U. Rix, K. L. Bennett, J Proteome Res 8 (2009) 4753-4765.
26. J. Rappsilber, Y. Ishihama, M. Mann, Anal Chem 75 (2003) 663-670.
27. J. Colinge, A. Masselot, M. Giron, T. Dessingy, J. Magnin, Proteomics 3 (2003) 1454-1463.
28. K. Kowanetz, N. Crosetto, K. Haglund, M. H. Schmidt, C. H. Heldin, I. Dikic, J Biol Chem 279 (2004) 32786-32795.
29. T. Shimamura, A. M. Lowell, J. A. Engelman, G. I. Shapiro, Cancer Res 65 (2005) 6401-6408.
30. P. J. Gray, Jr., T. Prince, J. Cheng, M. A. Stevenson, S. K. Calderwood, Nat Rev Cancer 8 (2008) 491-495.
31. D. H. Lundgren, S. I. Hwang, L. Wu, D. K. Han, Expert Rev Proteomics 7 (2010) 39-53.
32. O. N. Jensen, A. Shevchenko, M. Mann: Protein analysis by mass spectrometry. In *Book Protein analysis by mass spectrometry* (Editor Hames, B.D.). pp. 29-57. City: Oxford University Press Inc.; 1997:29-57.
33. N. Bisson, D. A. James, G. Ivosev, S. A. Tate, R. Bonner, L. Taylor, T. Pawson, Nat Biotechnol 29 (2011) 653-658.
34. C. Wu, M. H. Ma, K. R. Brown, M. Geisler, L. Li, E. Tzeng, C. Y. Jia, I. Jurisica, S. S. Li, Proteomics 7 (2007) 1775-1785.