

Snapshot imprinting: rapid identification of cancer cell surface proteins and epitopes using molecularly imprinted polymers



Stanislav S. Piletsky^{a,*}, Elena Piletska^b, Marta Poblocka^c, Salvador Macip^{c,d}, Donald J.L. Jones^{e,f,g}, Marta Braga^h, Thong H. Cao^{f,g}, Rajinder Singh^e, Alan C. Spivey^a, Eric O. Aboagye^h, Sergey A. Piletsky^b

^a Department of Chemistry, Molecular Sciences Research Hub, White City Campus, Imperial College London, London W12 0BZ, UK

^b School of Chemistry, College of Science and Engineering, University of Leicester, Leicester, UK

^c Mechanisms of Cancer and Ageing Laboratory, Department of Molecular and Cell Biology, University of Leicester, Leicester, UK

^d FoodLab, Faculty of Health Sciences, Universitat Oberta de Catalunya, Barcelona, Spain

^e Leicester Cancer Research Centre, Leicester Royal Infirmary, University of Leicester, Leicester, UK

^f Department of Cardiovascular Sciences, University of Leicester, Leicester, UK

^g National Institute for Health Research, Leicester Biomedical Research Centre, Glenfield Hospital, Leicester, UK

^h Department of Surgery and Cancer, Imperial College London, Hammersmith Campus, Du Cane Road, London, UK

ARTICLE INFO

Article history:

Received 10 May 2021

Received in revised form 29 July 2021

Accepted 18 September 2021

Available online xxxx

Keywords:

Cancer
Proteomics
Molecularly imprinted polymers
Epitopes
Cells

ABSTRACT

Proteomic mapping of cell surfaces is an invaluable tool for drug development and clinical diagnostics. This work describes a new 'snapshot imprinting' method designed to obtain proteomic maps of cell surfaces, with the aim of identifying cell surface markers and epitopes for diagnostic and therapeutic applications. The analysis of two cancer cell lines, HN5 and MDA-MB-468, is described herein as a proof of concept, along with the selective targeting of three identified epitopes of epidermal growth factor receptor using molecularly imprinted polymer nanoparticles. 438 proteins were identified using this technique, with 283 considered to be transmembrane or extracellular proteins. The major advantage of the molecular imprinting approach developed here is the ability to analyse cell surface proteins without tedious fractionation, affinity separation or labelling. We believe that this system of protein analysis may provide a basic molecular diagnostics toolbox for precise, personalised treatment of cancer and other diseases.

© 2021 The Authors. Published by Elsevier Ltd.
CC_BY_4.0

Introduction

The exposure of proteins and their spatial organisation on the cell surface is a highly complex phenomenon, influenced by protein expression, protein stability, and the micro and extracellular environment. Protein complexes on cell membranes are constantly being formed and resolved, and proteins are constantly shuttling between subcellular and extracellular locations to execute biological processes. These molecules are involved in signal transduction, transmembrane transport, cell-cell communication, cell adhesion to the extracellular matrix, and many other processes [1]. Their importance is underscored by the fact that membrane proteins are the targets of at least half of all currently approved drugs [2]. Identification of membrane proteins, however, is a challenging task due to their hydrophobicity and often low expression levels [3].

There are a number of established methods available to monitor the expression of membrane proteins, which can be used to interrogate the cell surface-associated proteins. Commonly used examples include biotinylation of surface amine residues followed by proteolysis and affinity purification of labelled extracellular peptides [4,5], 'shaving' approaches based on controlled proteolysis of live cells to collect surface proteins [6], and filter-aided sample preparation (FASP) methods, compatible with both of the above [7,8]. These protocols are highly useful for drug development, particularly for the development of immunotherapies. They can be supplemented with phenotypic screening of cell interactions with antibodies, for example using hybridoma technology or phage display [9]. However, these approaches do not identify peptides based on their immunogenic properties, but rather, on their susceptibility to trypsinisation. As a result, these peptides are not necessarily appropriate targets for antibodies and other binding agents.

Recently, we described a new approach using molecular imprinting to identify exposed peptide sequences on protein surfaces [10]. Molecularly imprinted polymers (MIPs) are synthetic receptors

* Corresponding author.

E-mail address: stanislav.piletsky14@imperial.ac.uk (S.S. Piletsky).

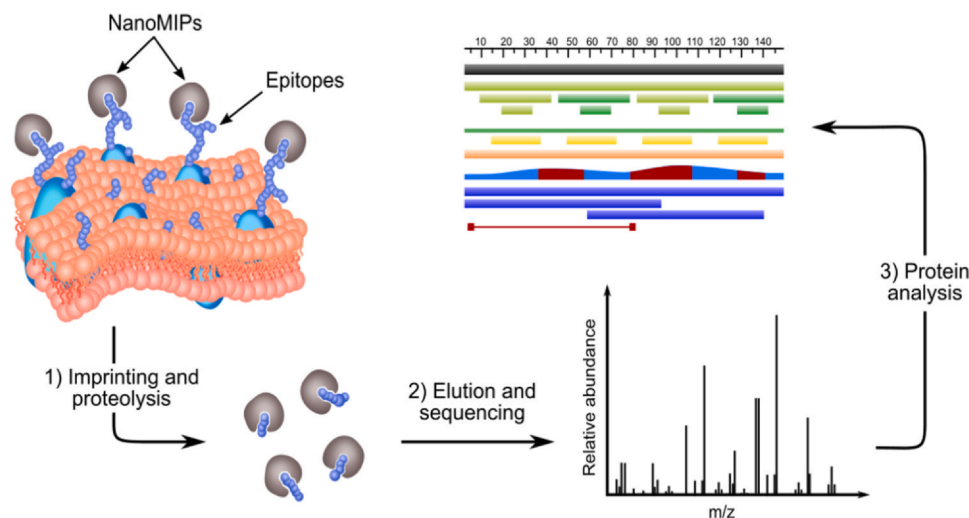


Fig. 1. Diagram depicting snapshot imprinting, wherein cell surface proteins are imprinted, the resultant nanoMIPs are collected, and the template peptides eluted and sequenced in order to identify novel biomarkers.

generated by forming polymers in the presence of templates such as small molecules, proteins or cells. Removal of the templates results in cavities which are structurally and electrostatically complementary for their template [11–13].

The protocol described herein, dubbed ‘snapshot imprinting’, expands upon the concept of MIP-based protein mapping by synthesising imprinted polymer nanoparticles, or nanoMIPs, in the presence of whole cells. This is followed by partial proteolysis of the protein bound to the polymer, and subsequent sequencing of peptides that were bound to the MIPs (Fig. 1, Fig. S1). This requires the following key steps:

1. Growth of the cell culture.
2. Addition of monomer mixture to the cell culture, initiation of polymerisation and synthesis of polymer nanoparticles in the presence of adherent cells.
3. Removal of non-imprinted polymers, monomers and extra-cellular proteins.
4. Trypsinisation and collection of nanoMIPs.
5. Filtering of nanoMIPs from cell fragments.
6. Elution of peptide templates.
7. Sequencing of eluted peptides and analysis of MS data.

The central concept behind this protocol relies on the assumption that MIPs synthesised in the presence of cells would only bind to exposed epitopes of surface proteins, and would be able to protect the imprinted peptide sequences from proteolysis. This approach provides the possibility of locating regions of the protein surface which have not yet been identified as epitopes, but which may offer improved affinity for natural and synthetic receptors such as antibodies, MIPs and aptamers.

Here we use this approach for the characterisation of two cancer cell lines: HN5 (squamous cell carcinoma of the tongue) and MDA-MB-468 (adenocarcinoma of the breast). These cancer subtypes are aggressive and often have few treatment options, demonstrating the need for the identification of new biomarkers and the development of new therapeutic agents [14,15].

Materials and methods

Materials

All reagents used in this project were obtained from Sigma-Aldrich Company Ltd (Poole, UK), Fisher Scientific Ltd

(Loughborough, UK), or Waters Corporation (Milford, USA) unless otherwise stated. Peptides were synthesised by Zhejiang Ontores Biotechnologies Co. Ltd. (Hangzhou, China).

MDA-MB-468 cells were provided by Dr Larissa Lezina at the Department of Genetics and Genome Biology, University of Leicester. HN5 cells were provided by the Division of Radiotherapy and Imaging & Cancer Therapeutics at the Institute of Cancer Research (ICR). MDA-MB-468 cells were grown in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 10% Foetal Bovine Serum (FBS) (GIBCO) and 50 units mL⁻¹ of penicillin-streptomycin (GIBCO) and cultured in a 37 °C humidified incubator with 5% CO₂. HN5 cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% Foetal Bovine Serum (FBS) (GIBCO) and 50 units mL⁻¹ of penicillin-streptomycin (GIBCO) and cultured in a 37 °C humidified incubator with 5% CO₂.

Snapshot imprinting

The growth media was carefully decanted from the flasks containing MDA-MB-468 and HN5 cells. The cells were then washed with phosphate buffered saline (PBS, pH 7.4, 10 mM) (4 × 20 mL) followed by the addition of 20 mL of monomeric mixture in PBS. The monomeric mixture consisted of *N*-isopropylacrylamide (NIPAm) (19.5 mg, 0.18 mmol), *N,N'*-methylene-bis-acrylamide (MBAA) (3 mg, 19.5 μmol), *N*-tert-butylacrylamide (TBA) (15 mg, 0.12 mmol) dissolved in 200 μL of ethanol, acrylic acid (1.1 μL, 16.0 μmol), *N*-(3-aminopropyl)methacrylamide hydrochloride (3 mg, 16.8 μmol) dissolved in 50 mL of PBS. Before addition to the cells the monomeric mixture was briefly sonicated in an ultrasonic bath and purged with nitrogen for 20 min. Polymerisation was initiated via the addition of potassium persulfate (KPS) (12 mg, 44.4 μmol) and *N,N,N',N'*-tetramethylethylenediamine (TEMED) (6 μL, 40.3 μmol) dissolved in PBS (400 μL). The polymerisation was carried out for 1 h at 20 °C.

After polymerisation the liquid was carefully removed from the flasks. The flask surface with adherent cells was washed with PBS (4 × 20 mL) in order to remove the unreacted monomers and unbound polymers. Digestion of the unprotected proteins was carried out via addition of trypsin (0.2 mg, trypsin from porcine pancreas, 1000–2000 BAEE units mg⁻¹ solid, Sigma-Aldrich) in PBS (20 mL) and left to digest for 72 h at 20 °C. The mixtures within the flasks were then collected with brief shaking and transferred to 50 kDa centrifugal filters (Amicon®Ultra-15, Merck, UK).

In order to remove trypsin and unbound peptides, the samples were filtered through these centrifugal filters for 20 min at 2355 g

(3500 rpm) using a Sigma 3–16 P centrifuge (SciQuip, UK). The cartridge was washed with HPLC-grade water (4 × 15 mL). Finally, the peptides bound to MIPs were eluted using hot water (95 °C, 3 × 1 mL) followed by centrifugation. The solution of the eluted peptide was frozen and sent for sequencing.

Nano ultra performance liquid chromatography (NanoUPLC)

Sample analysis was performed using a Waters NanoAcquity UPLC system (Waters Corporation, Milford, US). The peptides were initially loaded onto a Waters 2 G-V/M Symmetry C18 trap column (180 µm × 20 mm, 5 µm) to desalt and chromatographically focus the peptides prior to elution onto a Waters Acquity HSS T3 analytical UPLC column (75 µm C 250 mm, 1.8 µm). A 2 µL injection volume was used. Single pump trapping was used with 99.9% solvent A and 0.1% solvent B at flow rate of 5 µL min⁻¹ for 3 min. Solvent A was LC-MS grade water containing 0.1% formic acid and solvent B was acetonitrile containing 0.1% formic acid. The following 50 min run time gradient was used: 0 min:3% B, 30 min:40% B, 32 min:85% B, 40 min:85% B, 41 min:3% B and 50 min:3% B. For the analytical column the flow rate was set at 0.3 µL min⁻¹ and the temperature maintained at 40 °C.

Nano electrospray ionisation mass spectrometry

The NanoAcquity UPLC was coupled to a Waters Synapt G2 HDMS mass spectrometer (Waters Corporation, Milford, US). The instrument was operated in positive electrospray ionisation (ESI) mode. The capillary voltage was set at 2.40 kV and cone voltage at 30 V. PicoTip emitters (10 µm internal diameter, New Objective, US) were used for the nanostage probe. A helium gas flow of 180 mL min⁻¹ and ion mobility separator nitrogen gas flow of 90 mL min⁻¹ with a pressure of 2.5 mbar were used. The IMS wave velocity was set at 650 m s⁻¹ and the IMS wave height at 40 V. During the high-definition mass spectrometry (HDMSE) acquisition, a low collision induced dissociation (CID) energy of 2 V was applied across the transfer ion guide. For the high CID energy acquisition, a ramp of 27–50 V was applied. Argon was used as the CID gas. Lockspray provided mass accuracy throughout the chromatographic run using [Glu1]-Fibrinopeptide (GFP) with *m/z* 785.8427. The data were acquired using MassLynx 4.1.

Peptide and protein analysis

Peptide and protein analysis was performed using Progenesis QI for Proteomics version 4.2 (Nonlinear Dynamics, Manchester, UK), Microsoft Excel, and GraphPad Prism. Progenesis QI allowed analysis of LC-MS data, enabling identification of peptides and proteins within samples and quantitative comparison between samples. Progenesis QI also provides quality control metrics to give confidence in the experimental conditions, instrument set up and data analysis. The human UniProtKB database (November 2013) was employed in FASTA format. Strict trypsin cleavage rules were used and two missed cleavages were allowed. A minimum of two fragments per peptide, a minimum of five fragments per protein and a minimum of two unique peptides per protein were applied. A maximum rate of 1% was set for the false discovery rate (FDR) at the peptide and protein level. The Hi-3 relative quantitation method was used, in which the top three most abundant peptides for each protein were employed for protein quantitation. Finally, the results generated from using the Progenesis QI were exported to Microsoft Excel for further data analysis.

In order to analyse subcellular localisation of detected proteins, the proteins were analysed by accession number against UniProt's database. A script was used to search for each accession number and record the locations as noted by both the UniProt Annotation panel

and the UniProt Gene Ontology Project (GO). When considering whether a protein was extracellular, the UniProt GO listings of observed and predicted sublocations were used. The sublocations considered to be extracellular or otherwise exposed were 'cell surface', 'plasma membrane', 'integral component of plasma membrane', 'extracellular region', 'extracellular space', 'extracellular vesicle', and 'extracellular exosome'. When calculating frequency of appearance of each sublocation within the resultant datasets, the majority of proteins were found in multiple sublocations and so the total frequency sums to over 100%.

During analysis of MDA-MB-468 proteins previously identified via label-free deep proteome analysis [16], the identified proteins were searched by accession number against the UniProt GO database as described above. Proteins were only considered for analysis if they were found within one of the two replicates, and only if the protein had at least one known sublocation within the UniProt GO database.

Synthesis of MIP nanoparticles – preparation of solid phase

MIPs were prepared using a modified version of an existing solid-phase synthesis approach [12]. The immobilisation of peptides was performed as follows. Glass beads (60 g) were boiled in NaOH (1 M, 200 mL) for 15 min, washed with water (5 × 200 mL), washed with PBS (2 × 100 mL), washed with water (2 × 200 mL), washed with acetone (1 × 100 mL) and allowed to dry. They were then incubated in 4% (v/v) (3-iodopropyl)trimethoxysilane in anhydrous toluene (25 mL) overnight at room temperature, protected from light. The beads were then washed with acetone (3 × 100 mL) and allowed to dry. The peptide (epitope-1, epitope-2 or epitope-3) (10 mg) was dissolved in borate buffer (pH 9.2, 30 mM sodium tetraborate, 25 mL), added to the glass beads and incubated overnight, protected from light. The beads were then washed with water (3 × 200 mL) and acetone (1 × 100 mL) and allowed to dry.

The immobilisation of biotin was performed as follows. Glass beads (60 g) were boiled in sodium hydroxide in water (1 M, 200 mL) for 15 min, washed with water (5 × 200 mL), washed with PBS (2 × 100 mL), washed with water (2 × 200 mL), washed with acetone (100 mL) and allowed to dry. They were then incubated in 4% (v/v) *N*-(6-aminohexyl)-aminomethyltriethoxysilane in anhydrous toluene (25 mL) overnight at room temperature. The beads were then washed with acetone (3 × 100 mL) and allowed to dry. Biotin (10 mg, 41 µmol, 1 eq.), *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC) (79 mg, 0.4 mmol, 10 eq.) and *N*-hydroxysuccinimide (NHS) (71 mg, 0.6 mmol, 15 eq.) were dissolved in water (10 mL), incubated for 15 min, added to PBS (15 mL) and transferred to the amine-functionalised glass beads. This mixture was incubated overnight, then the beads were washed with water (3 × 200 mL) and acetone (1 × 100 mL), and allowed to dry.

Synthesis of MIP nanoparticles – polymerisation

The following monomers were dissolved in PBS (50 mL): *N*-isopropyl acrylamide (200 mg, 1.8 mmol), *N*-tert-butylacrylamide (165 mg, 1.3 mmol), *N,N'*-methylenebis(acrylamide) (30 mg, 0.2 mmol), *N*-(3-aminopropyl)methacrylamide hydrochloride (30 mg, 0.17 mmol) and acrylic acid (11 µL, 0.16 mmol). 60 g of template-functionalised glass beads were added to the monomer solution, which was then bubbled with nitrogen for 20 min. Polymerisation was initiated through the addition of ammonium persulfate (30 mg, 0.13 mmol) and *N,N,N',N'*-tetramethyl ethylenediamine (30 µL, 0.2 mmol) in water (500 µL). The mixture was shaken and incubated for 1 h before being transferred to a solid phase extraction cartridge fitted with a 20 µm polyethylene frit. Unreacted monomers and low affinity polymers were removed from the glass beads by washing with water (5 × 100 mL). High affinity polymers were collected with hot ethanol (65 °C, 2 × 25 mL), reduced

to 5 mL under vacuum and dialysed against water for 1 week with regular change of water.

Surface plasmon resonance (SPR)

SPR measurements were performed using a Biacore 3000 (Cytiva, UK) at 25 °C using phosphate-buffered saline (PBS) (10 mM phosphate buffer, 2.7 mM KCl, 137 mM NaCl, pH 7.4) as the running buffer at flow rate 35 $\mu\text{L min}^{-1}$. The self-assembled gold sensor chip was cleaned using plasma and placed in a solution of mercaptododecanoic acid in ethanol (1.1 mg mL^{-1}) where they were stored until use. Before assembly the sensor chip was rinsed with ethanol and water and dried in a stream of air. The EGFR protein (10 $\mu\text{g mL}^{-1}$ solution in PBS) was immobilised on the surface of the carboxylated chip using EDC/NHS coupling (0.4 mg mL^{-1} and 0.6 mg mL^{-1} , respectively). The nanoMIPs prepared for different peptide template and for biotin using solid phase synthesis approach were diluted with PBS in the concentration range averaging between 0.2 nM and 0.01 nM. Sensorgrams were collected sequentially for all analyte concentrations running in KINJECT mode (injection volume 100 μL and dissociation time 120 s). Dissociation constants (K_d) were calculated from plots of the equilibrium biosensor response using the BiaEvaluation v4.1 software using a 1:1 binding model with drifting baseline fitting.

Results

Analysis of MDA-MB-468 and HN5 proteins

Snapshot imprinting revealed a significant number of cell membrane proteins and corresponding epitopes on the surface of both cell lines that may serve as promising therapeutic targets, such as EGFR, 14-3-3 proteins, CD44 and basigin [17–20]. A total of 438 proteins were identified across these two cell lines. 241 proteins were common to both cell lines, 91 were unique to HN5 cells and 106 to MDA-MB-468 cells (Fig. 2A). These proteins may therefore serve as possible antigens for cell-line-specific targeting. The sublocations of these proteins do not vary widely between the cell lines (Fig. 3), demonstrating the consistency of this technique and the similarity in the proteomes of these two cell lines.

As a control experiment, MDA-MB-468 cells underwent this mapping procedure without the addition of the polymerisation initiator. In this case, proteins exposed on the cell surface were digested without the protection afforded by MIP nanoparticles. This experiment resembles the classical 'shaving' protocol used in surface protein analysis.

Fewer proteins were found within the MIP-free mapping control experiment as compared to snapshot imprinting. 347 unique proteins were found during snapshot imprinting of MDA-MB-468 cells, compared to 115 proteins found within the non-imprinted control.

70 proteins were common to both samples, and 45 were found in the control that were absent from the snapshot imprinting results (Fig. 2B). The number of median peptides per protein in the MIP-free control experiment was also smaller (6 versus 10). The reason for the absence of 45 proteins from the snapshot imprinting results is not fully understood. These proteins are amongst those with the lowest abundance found within the control experiment; as such it is possible that they were not detected in the snapshot imprinting results in the presence of signals from more abundant, MIP-enriched proteins. Furthermore, though our selection criteria required that at least two unique peptides be found for a protein to be included, 9 of these 45 proteins were found during snapshot imprinting with only 1 unique peptide. The sublocations of the 115 proteins found within the non-imprinted control do not deviate significantly from those of the 347 found during snapshot imprinting, indicating no particular sublocation was excluded or enriched (Fig. S2). As our methodology involved a washing step, the peptides remaining on the centrifuge cartridge filter in the MIP-free control likely adsorbed onto the filter membrane itself. It is possible that the presence of MIPs on the filter membrane during snapshot imprinting interferes with this adsorption, resulting in the loss of some proteins from the snapshot imprinting results that were present in the MIP-free control.

Comparison of MDA-MB-468 proteins found via snapshot imprinting to those found via deep proteome analysis

The proteins identified via snapshot imprinting were sorted by sublocation using the UniProt database, as described within Material and Methods. This process was repeated for 8792 proteins of MDA-MB-468 found via label-free deep proteome analysis [16].

As shown in Fig. 4, compared to proteins found via label-free deep proteome analysis, a number of cellular sublocations are greatly over-represented within the data obtained from snapshot imprinting. Notably among these are the membrane, extracellular exosome, extracellular region, focal adhesion and extracellular space sublocations. Furthermore, of the 8792 MDA-MB-468 proteins described within literature, only 2380 (27%) were found within sublocations considered to be extracellular/exposed by the criteria described within Material and Methods. This is in contrast to the results obtained via snapshot imprinting, in which 238 out of 347 proteins (69%) were found to be extracellular/exposed. This implies a degree of enrichment of exposed proteins during snapshot imprinting, potentially caused by exposed proteins acting as better templates for MIP synthesis due to greater accessibility to monomers and polymer nanoparticles. This enrichment is beneficial when selecting proteins that can serve as targets for antibodies, MIPs or aptamers. Interestingly, extracellular exosomes were also highly represented within the snapshot imprinting results, despite the expectation that exosomes will be washed away prior to collection of cell lysate. This may be attributed to a high degree of colocalisation

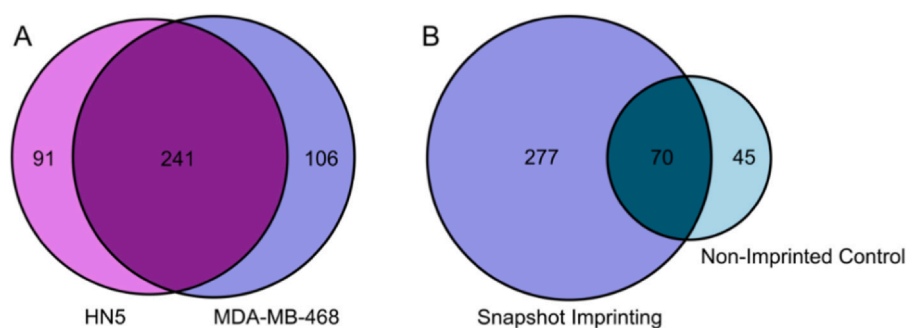


Fig. 2. Venn diagrams showing: (A) the number of proteins discovered in HN5 (left) and MDA-MB-468 (right) cell lines using snapshot imprinting; (B) the number of proteins discovered in MDA-MB-468 cell line with (left) and without (right) imprinted polymers.

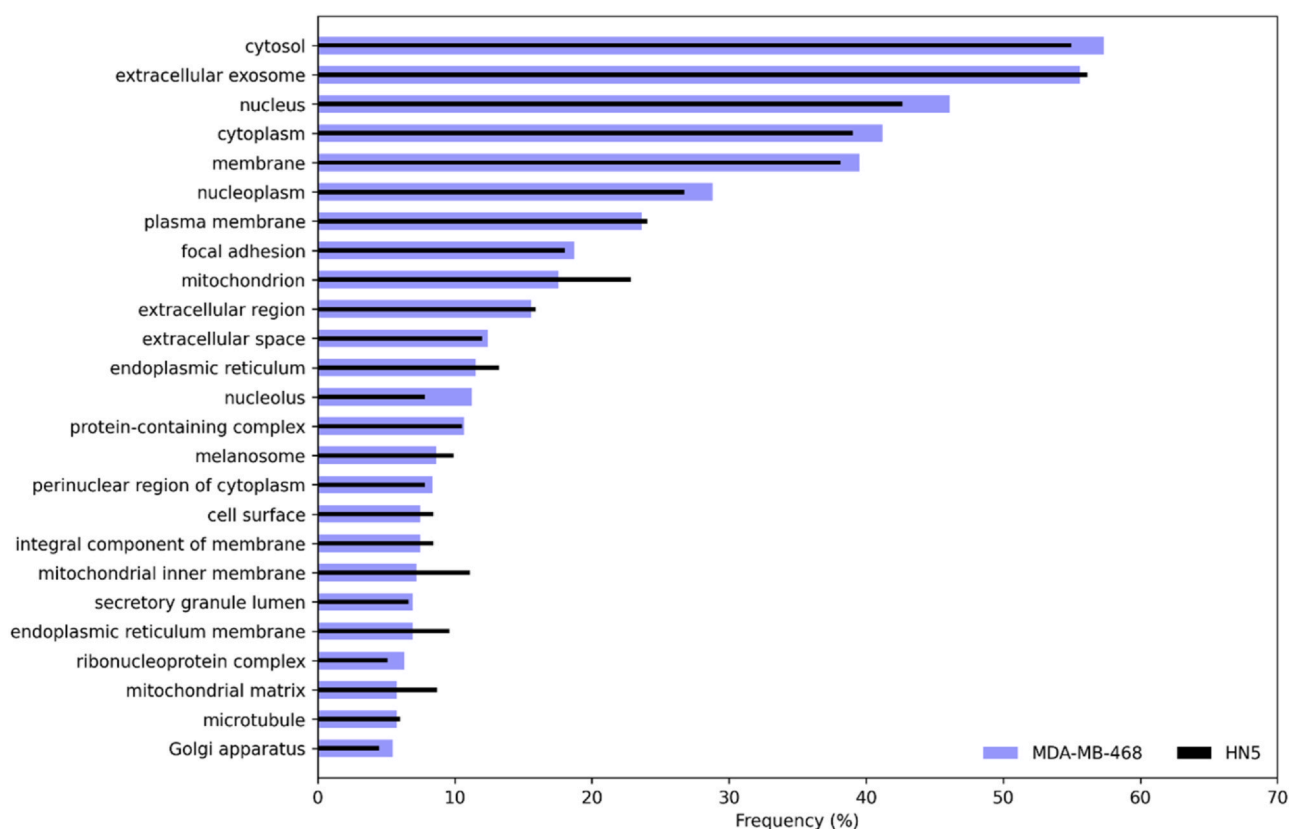


Fig. 3. Bar chart showing the frequencies of the 25 most common cellular sublocations for MDA-MB-468 proteins (blue) as compared to HN5 proteins (black) identified via snapshot imprinting.

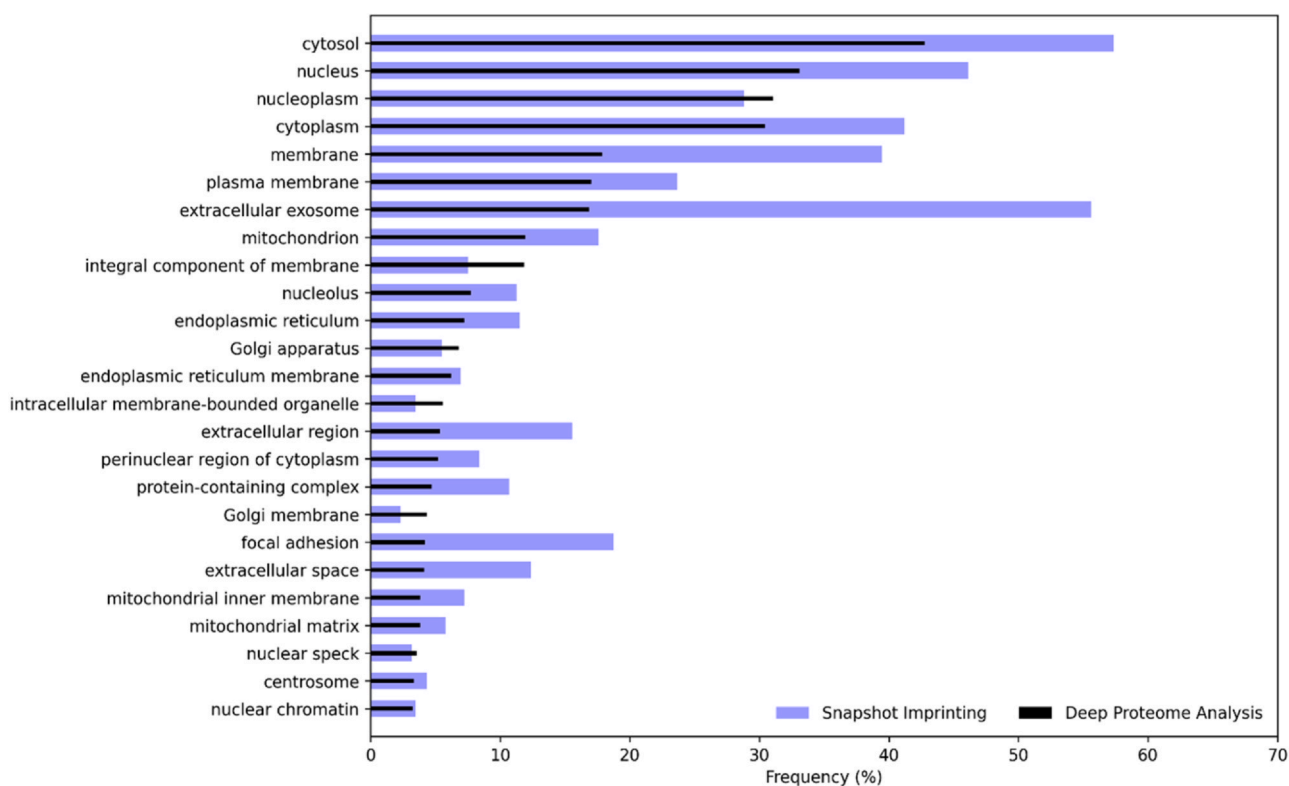


Fig. 4. Bar chart showing the frequencies of the 25 most common cellular sublocations for MDA-MB-468 proteins found via snapshot imprinting (blue) as compared to those found using label-free deep proteome analysis (black) [16].

of exosomal proteins with membrane, focal adhesion proteins and other exposed sublocations.

The discovery of non-integral membrane proteins

A large number of proteins identified in this work, as well as in previous proteomics experiments, are not integral membrane proteins (IMPs). These are typically associated with the endoplasmic reticulum/Golgi apparatus, mitochondria, peroxisomes, lysosomes, cytoskeleton and nuclei [3]. In our study, we observed the presence of major cytosolic (alpha-enolase, GAPDH, glucose-6-phosphate isomerase), cytoskeletal (actin, tubulin) and also high-abundance nuclear (histones, small nuclear ribonucleoproteins) proteins. These same proteins were previously identified via fractionation of transmembrane proteins [3]. The discovery of non-IMPs both in our experiments and by others using the classical proteolytic shaving approach is not fully understood. Contamination of the membrane fractions with intracellular proteins during trypsinisation is certainly possible, however this should not occur during surface protein analysis that relies on biotinylation as this takes place prior to cell lysis. The fact that intracellular proteins are detected in every protocol used in cell surface protein analysis suggests that the presence of non-IMPs in the cell surface is genuine. The authors of the shaving approach explained the presence of intracellular proteins on the cell membrane by invoking unspecified exporting/secretory cellular machinery. There are a number of viable mechanisms that could be responsible for this phenomenon, such as canonical and non-canonical secretion, the integration of exosomal and ectosomal proteins, and adsorption of proteins released during cell death or mitosis (Fig. S3) [21–31].

Epitope mapping – a case study on EGFR

In order to demonstrate the value of this technique in identifying not only surface proteins but also epitopes suitable for targeting, epitopes of epidermal growth factor receptor (EGFR) were investigated as a case study. EGFR was selected both because of its clinical significance as a common cancer biomarker, and due to its overexpression on both investigated cell lines [32].

The Immune Epitope Database and Analysis Resource (https://www.iedb.org/home_v3.php) has, at the time of writing, 88 epitopes of EGFR on record, collated from 32 sources. This includes overlapping epitopes and both linear and conformational epitopes. A total of 426 amino acids are considered to contribute towards at least one epitope.

During snapshot imprinting, 36 EGFR peptides were identified between the two cell lines. Taking into account that some peptides were subsections of larger peptides and that two peptides may overlap, 18 EGFR sequences were identified as possible epitopes,

totalling 394 amino acids (Fig. 5). Of the 18 sequences found via snapshot imprinting, 13 overlapped with known epitopes of EGFR. Of the 394 constituent amino acids, 137 are found within known epitopes. As is to be expected, neither known epitopes within literature nor the results of snapshot imprinting show epitopes sequentially close to the transmembrane region of EGFR (residues 646–668).

To demonstrate the suitability of these epitopes as potential therapeutic targets, three peptides from the extracellular portion of EGFR were selected (Fig. 6, Table 1) and used as templates for making MIPs using a modified variant of the solid phase synthesis approach developed by Canfarotta et al. [12]. The binding affinity of these three peptides for whole EGFR was then assessed using surface plasmon resonance (SPR), along with MIPs made for a non-peptide template (biotin) as a control.

The dissociation constant (K_d) of each MIP with whole EGFR was measured as listed in Table 1. All three EGFR epitope-imprinted MIPs showed significantly higher affinity for whole EGFR (16–48 nM) than a control MIP imprinted with biotin, which showed no discernible binding at nanomolar concentration ranges (Fig. S4). This demonstrates the utility of snapshot imprinting in the identification of protein epitopes suitable for generation of selective binders such as antibodies, MIPs and aptamers.

Discussion and conclusion

The process of snapshot imprinting can be presented as the preservation of exposed fragments of cell proteins through their complexation with monomers in polymeric networks. The term ‘snapshot’ does not describe the speed of the process (which is fast but not instantaneous), but rather that the proteins which are imprinted are representative of all those present during imprinting. The monomeric mixture, comprising of various acrylamide-based monomers, was previously optimised for protein imprinting [33,34]. The synthesised nanoparticles are non-toxic to cultured cells [35–37], and similarly in our experiments the addition of monomer mixture did not appear to trigger cell lysis (Fig. S5).

It is useful to discuss our results in relation to those obtained using the previously described protocols. The total number of distinct proteins identified for cancer cells using a shotgun proteomics approach was 12,775, though of course these offer no selectivity for surface proteins [16]. The total number and type of cell membrane proteins identified in proteomic studies is highly dependent on the isolation technique and varies from 69 to 629 proteins depending on the cell line and the discovery process used (Cell Surface Protein Atlas- <http://wlab.ethz.ch/cspa/#downloads>) [38]. Accordingly, 438 proteins were identified via snapshot imprinting for both cell lines, with 283 considered to be exposed or extracellular. While appreciating the differences in lineage, these numbers are significantly higher than the 188 proteins previously identified in surface protein

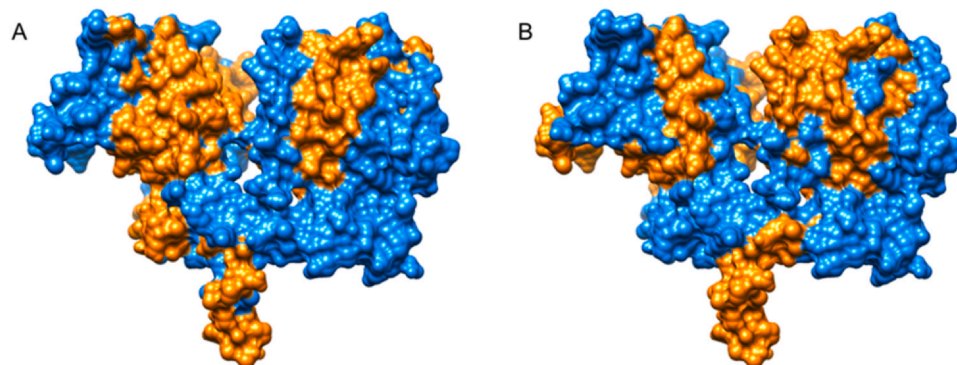


Fig. 5. Extracellular region of EGFR; regions considered to be epitopes highlighted in orange, remainder in blue. Epitopes selected from: (A) label-free deep proteome analysis; (B) snapshot imprinting.

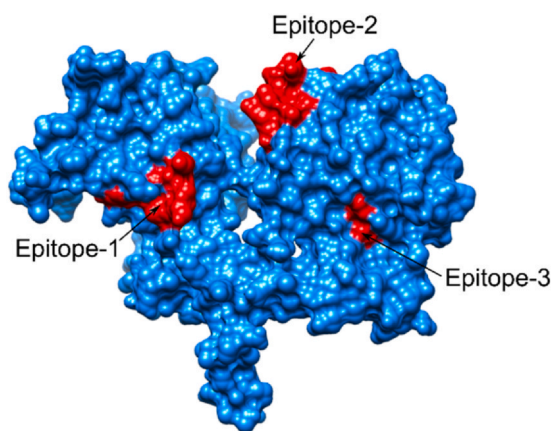


Fig. 6. Extracellular region of EGFR; epitopes selected for imprinting highlighted in red, remainder in blue.

Table 1

MIP binding affinities. Dissociation constants of four MIPs against whole EGFR as determined by surface plasmon resonance measurement.

MIP	Template	Dissociation Constant, nM (χ^2)
MIP-1	Epitope-1 (CGKLFGTSGQK)	47.1 (3.36×10^{-5})
MIP-2	Epitope-2 (CGLTQLGTGFEDHFLSLQR)	16.8 (0.4)
MIP-3	Epitope-3 (CGNLQEILHGAVR)	31.9 (4.15×10^{-5})

analysis of MDA-MB-231 cells [9]. Using the traditional shaving approach, the number of discovered membrane proteins varied from 178 to 237; a similar number to the 115 proteins discovered within our similar MIP-free control group [39,40]. This snapshot imprinting approach can therefore compete with previously reported studies of cell surface proteins, which employed elaborate fractionation or affinity purification of membrane proteins [40]. The total number of membrane proteins identified with either snapshot imprinting or the shaving approach was very low as compared with the total estimated number of membrane proteins present [41]. The most likely explanation for this is the low abundance of most membrane proteins, and the possibility that some of these proteins are not ideal targets for imprinting, for example, due to poor accessibility.

The two cell lines mapped using this technique, HN5 and MDA-MB-468, were originally selected for the purpose of identifying epitopes of epidermal growth factor receptor (EGFR) for a separate study on MIP-based EGFR targeting, due to their overexpression of EGFR. While information about the cell proteome is available for MDA-MB-486 cells, only genomic information can be found for HN5 cells [42]. We have demonstrated that EGFR epitopes identified via snapshot imprinting can act as templates for the synthesis of molecularly imprinted polymers. Further investigations can be carried out in the future as to whether all detected peptides are suitable targets, and whether peptides that were not highlighted are indeed worse targets.

The strength of this approach is that it is not constrained by pre-conceived assumptions of target protein location, allowing the freedom to discover novel targets for drug development and diagnostics. It is relatively simple to perform as no extraction or separation of membrane fractions is required. In contrast to antibody staining, which assesses the presence or absence of proteins, our approach provides sensitive and precise quantitation over a broad range of membrane-associated proteins for which commercial antibodies do not exist. The additional benefits of snapshot imprinting lie in its ability to link sequences of identified peptides with their ability to generate MIP nanoparticles as binders. In contrast, established protocols such as the shaving approach can provide information about exposed peptides and protein structures, but due to a lack of correlation

between the abundance of proteins and their immunogenicity the identified peptides cannot necessarily be used to make corresponding antibodies. As a result, snapshot imprinting represents a new all-in-one system for mapping surface proteins whilst screening their peptides for the ability to serve as epitopes for targeting.

In summary, a new method was developed for identifying cancer cell surface proteins and epitopes via synthesis of molecularly imprinted polymers in the presence of cells, followed by trypsinisation and sequencing of imprinted peptides. This technique screens peptides for their ability to act as targets for molecular imprinting, allowing the selection of epitopes for diagnostic and therapeutic applications. As a proof of concept, two cancer cell lines (MDA-MB-468 and HN5) were mapped, identifying 438 proteins and over 5000 peptides. Molecularly imprinted polymers (MIPs) were successfully generated for three epitopes of epidermal growth factor receptor (EGFR) found via mapping, demonstrating the ability of this technique to identify epitopes suitable for targeting.

CRedit authorship contribution statement

Stanislav S. Piletsky: Investigation, Analysis, Writing – original draft, Writing – review & editing. **Elena Piletska:** Methodology, Investigation. **Marta Poblocka:** Investigation. **Salvador Macip:** Methodology. **Donald J. L. Jones:** Methodology. **Marta Braga:** Investigation. **Thong H. Cao:** Formal Analysis. **Rajinder Singh:** Formal Analysis. **Alan C. Spivey:** Supervision. **Eric O. Aboagye:** Supervision. **Sergey A. Piletsky:** Conceptualization, Formal Analysis, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC, UK) [2133561], the Medical Research Council (MRC, UK) [2136145, MC-A652-5PY80], the Imperial President's PhD Scholarship (UK), the Imperial College NIHR Biomedical Research Centre award (UK) [WSCC_P62585], the Experimental Cancer Medicine Centres (UK) [C37/A7283], and the Cancer Research UK (CRUK) National Cancer Imaging Accelerator (UK) [C2536/A28680].

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.nantod.2021.101304.

References

- [1] L.M. Orre, M. Vesterlund, Y. Pan, T. Arslan, Y. Zhu, A. Fernandez Woodbridge, O. Frings, E. Fredlund, J. Lehtiö, Subcellbarcode: proteome-wide mapping of protein localization and relocalization, *Mol. Cell.* 73 (2019) 166–182.e7, <https://doi.org/10.1016/j.molcel.2018.11.035>
- [2] M.A. Yildirim, K. il Goh, M.E. Cusick, A.L. Barabási, M. Vidal, Drug-target network, *Nat. Biotechnol.* 25 (2007) 1119–1126, <https://doi.org/10.1038/nbt1338>
- [3] O. Vit, P. Man, A. Kadek, J. Hausner, J. Sklenar, K. Harant, P. Novak, M. Scigelova, G. Woffendin, J. Petrak, Large-scale identification of membrane proteins based on analysis of trypsin-protected transmembrane segments, *J. Proteom.* 149 (2016) 15–22, <https://doi.org/10.1016/j.jprot.2016.03.016>
- [4] Y. Zhao, W. Zhang, Y. Kho, Y. Zhao, Proteomic analysis of integral plasma membrane proteins, *Anal. Chem.* 76 (2004) 1817–1823, <https://doi.org/10.1021/ac0354037>
- [5] V. Hung, N.D. Udeshi, S.S. Lam, K.H. Loh, K.J. Cox, K. Pedram, S.A. Carr, A.Y. Ting, Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2, *Nat. Protoc.* 11 (2016) 456–475, <https://doi.org/10.1038/nprot.2016.018>

- [6] R.L. Juliano, V. Ling, A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants, *BBA - Biomembr.* 455 (1976) 152–162, [https://doi.org/10.1016/0005-2736\(76\)90160-7](https://doi.org/10.1016/0005-2736(76)90160-7)
- [7] J.R. Wiśniewski, A. Zougman, N. Nagaraj, M. Mann, Universal sample preparation method for proteome analysis, *Nat. Methods* 6 (2009) 359–362, <https://doi.org/10.1038/nmeth.1322>
- [8] M. Hernandez-Valladares, E. Aasebø, O. Mjaavatten, M. Vaudel, Ø. Bruserud, F. Berven, F. Selheim, Reliable FASP-based procedures for optimal quantitative proteomic and phosphoproteomic analysis on samples from acute myeloid leukemia patients, *Biol. Proced. Online* 18 (2016) 13, <https://doi.org/10.1186/s12575-016-0043-0>
- [9] S. Rust, S. Guillard, K. Sachsenmeier, C. Hay, M. Davidson, A. Karlsson, R. Karlsson, E. Brand, D. Lowne, J. Elvin, M. Flynn, G. Kurosawa, R. Hollingsworth, L. Jermutus, R. Minter, Combining phenotypic and proteomic approaches to identify membrane targets in a 'triple negative' breast cancer cell type, *Mol. Cancer* 12 (2013) 11, <https://doi.org/10.1186/1476-4598-12-11>
- [10] S. Piletsky, E. Piletska, F. Canfarotta, D. Jones, Methods and kits for determining binding sites, GB1704823.2., 2018. (<https://patents.google.com/patent/WO2018178629A1/en>) (accessed February 5, 2020).
- [11] K. Haupt, K. Mosbach, Plastic antibodies: developments and applications, *Trends Biotechnol.* 16 (1998) 468–475, [https://doi.org/10.1016/S0167-7799\(98\)01222-0](https://doi.org/10.1016/S0167-7799(98)01222-0)
- [12] F. Canfarotta, A. Poma, A. Guerreiro, S. Piletsky, Solid-phase synthesis of molecularly imprinted nanoparticles, *Nat. Protoc.* 11 (2016) 443–455, <https://doi.org/10.1038/nprot.2016.030>
- [13] S. Piletsky, F. Canfarotta, A. Poma, A.M. Bossi, S. Piletsky, Molecularly imprinted polymers for cell recognition, *Trends Biotechnol.* 38 (2020) 368–387, <https://doi.org/10.1016/j.tibtech.2019.10.002>
- [14] J. Loeffler-Ragg, I. Schwentner, G.M. Sprinzel, H. Zwiernitz, EGFR inhibition as a therapy for head and neck squamous cell carcinoma, *Expert Opin. Investig. Drugs* 17 (2008) 1517–1531, <https://doi.org/10.1517/13543784.17.10.1517>
- [15] S.K. Pal, B.H. Childs, M. Pegram, Triple negative breast cancer: unmet medical needs, *Breast Cancer Res. Treat.* 125 (2011) 627–636, <https://doi.org/10.1007/s10549-010-1293-1>
- [16] R.T. Lawrence, E.M. Perez, D. Hernández, C.P. Miller, K.M. Haas, H.Y. Irie, S.I. Lee, A.C. Blau, J. Villén, The proteomic landscape of triple-negative breast cancer, *Cell Rep.* 11 (2015) 630–644, <https://doi.org/10.1016/j.celrep.2015.03.050>
- [17] P. Seshacharyulu, M.P. Ponnusamy, D. Haridas, M. Jain, A.K. Ganti, S.K. Batra, Targeting the EGFR signaling pathway in cancer therapy, *Expert Opin. Ther. Targets* 16 (2012) 15–31, <https://doi.org/10.1517/14728222.2011.648617>
- [18] H. Hermeking, The 14-3-3 cancer connection, *Nat. Rev. Cancer* 3 (2003) 931–943, <https://doi.org/10.1038/nrc1230>
- [19] C. Chen, S. Zhao, A. Karnad, J.W. Freeman, The biology and role of CD44 in cancer progression: therapeutic implications, *J. Hematol. Oncol.* 11 (2018) 64, <https://doi.org/10.1186/s13045-018-0605-5>
- [20] X. Hu, J. Su, Y. Zhou, X. Xie, C. Peng, Z. Yuan, X. Chen, Repressing CD147 is a novel therapeutic strategy for malignant melanoma, *Oncotarget* 8 (2017) 25806–25813, <https://doi.org/10.18632/oncotarget.15709>
- [21] A. Olaya-Abriel, I. Jiménez-Munguía, L. Gómez-Gascón, M.J. Rodríguez-Ortega, Surfomics: shaving live organisms for a fast proteomic identification of surface proteins, *J. Proteom.* 97 (2014) 164–176, <https://doi.org/10.1016/j.jprot.2013.03.035>
- [22] J.D. Bendtsen, L. Kiemer, A. Fausbøll, S. Brunak, Non-classical protein secretion in bacteria, *BMC Microbiol.* 5 (2005) 58, <https://doi.org/10.1186/1471-2180-5-58>
- [23] C. Nombela, C. Gil, W.L.J. Chaffin, Non-conventional protein secretion in yeast, *Trends Microbiol.* 14 (2006) 15–21, <https://doi.org/10.1016/j.tim.2005.11.009>
- [24] L. Aguilera, E. Ferreira, R. Giménez, F.J. Fernández, M. Taulés, J. Aguilar, M.C. Vega, J. Badia, L. Balmori, Secretion of the housekeeping protein glyceraldehyde-3-phosphate dehydrogenase by the LEE-encoded type III secretion system in enteropathogenic *Escherichia coli*, *Int. J. Biochem. Cell Biol.* 44 (2012) 955–962, <https://doi.org/10.1016/j.biocel.2012.03.002>
- [25] K. Watson, N.J. Gooderham, D.S. Davies, R.J. Edwards, Nucleosomes bind to cell surface proteoglycans, *J. Biol. Chem.* 274 (1999) 21707–21713, <https://doi.org/10.1074/jbc.274.31.21707>
- [26] E. Hariton-Gazal, J. Rosenbluh, A. Graessmann, C. Gilon, A. Loyter, Direct translocation of histone molecules across cell membranes, *J. Cell Sci.* 116 (2003) 4577–4586, <https://doi.org/10.1242/jcs.00757>
- [27] T.P. Lozito, R.S. Tuan, Endothelial and cancer cells interact with mesenchymal stem cells via both microparticles and secreted factors, *J. Cell. Mol. Med.* 18 (2014) 2372–2384, <https://doi.org/10.1111/jcmm.12391>
- [28] K. Al-Nedawi, B. Meehan, J. Micallef, V. Lhotak, L. May, A. Guha, J. Rak, Intercellular transfer of the oncogenic receptor EGFRvIII by microvesicles derived from tumour cells, *Nat. Cell Biol.* 10 (2008) 619–624, <https://doi.org/10.1038/ncb1725>
- [29] Z. Ma, X. Cui, L. Lu, G. Chen, Y. Yang, Y. Hu, Y. Lu, Z. Cao, Y. Wang, X. Wang, Retracted article: exosomes from glioma cells induce tumor-like phenotype in mesenchymal stem cells by activating glycolysis, *Stem Cell Res. Ther.* 10 (2019) 60, <https://doi.org/10.1186/s13287-019-1149-5>
- [30] S.A. Whelan, J. He, M. Lu, P. Souda, R.E. Saxton, K.F. Faull, J.P. Whitelegge, H.R. Chang, Mass spectrometry (LC-MS/MS) identified proteomic biosignatures of breast cancer in proximal fluid, *J. Proteome Res.* 11 (2012) 5034–5045, <https://doi.org/10.1021/pr300606e>
- [31] M.T. Müller, R. Schempp, A. Lutz, T. Felder, E. Felder, P. Miklavc, Interaction of microtubules and actin during the post-fusion phase of exocytosis, *Sci. Rep.* 9 (2019) 11973, <https://doi.org/10.1038/s41598-019-47741-0>
- [32] D.W. Rusnak, K.J. Alligood, R.J. Mullin, G.M. Spehar, C. Arenas-Elliott, A.M. Martin, Y. Degenhardt, S.K. Rudolph, T.F. Haws, B.L. Hudson-Curtis, T.M. Gilmer, Assessment of epidermal growth factor receptor (EGFR, ErbB1) and HER2 (ErbB2) protein expression levels and response to lapatinib (Tykerb®, GW572016) in an expanded panel of human normal and tumour cell lines, *Cell Prolif.* 40 (2007) 580–594, <https://doi.org/10.1111/j.1365-2184.2007.00455.x>
- [33] A. Poma, H. Brahmabhatt, J.K. Watts, N.W. Turner, Nucleoside-tailored molecularly imprinted polymeric nanoparticles (MIP NPs), *Macromolecules* 47 (2014) 6322–6330, <https://doi.org/10.1021/ma501530c>
- [34] K. Yoshimatsu, H. Koide, Y. Hoshino, K.J. Shea, Preparation of abiotic polymer nanoparticles for sequestration and neutralization of a target peptide toxin, *Nat. Protoc.* 10 (2015) 595–604, <https://doi.org/10.1038/nprot.2015.032>
- [35] Y. Hoshino, H. Koide, T. Urakami, H. Kanazawa, T. Kodama, N. Oku, K.J. Shea, Recognition, neutralization, and clearance of target peptides in the bloodstream of living mice by molecularly imprinted polymer nanoparticles: a plastic antibody, *J. Am. Chem. Soc.* 132 (2010) 6644–6645, <https://doi.org/10.1021/ja102148f>
- [36] M. Gallorini, A. Cataldi, V. di Giacomo, HEMA-induced cytotoxicity: oxidative stress, genotoxicity and apoptosis, *Int. Endod. J.* 47 (2014) 813–818, <https://doi.org/10.1111/iej.12232>
- [37] L. Wang, L. Xiao, R.Z. Zhang, L.Z. Qiu, R. Zhang, H.X. Shi, Effects of acrylate/acrylamide polymers on the adhesion, growth and differentiation of Mouse cells, *Biomed. Mater.* 14 (2019), <https://doi.org/10.1088/1748-605X/aae5cb>
- [38] F. Raimondo, P. Ceppi, K. Guidi, M. Masserini, C. Foletti, M. Pitto, Proteomics of plasma membrane microdomains, *Expert Rev. Proteom.* 2 (2005) 793–807, <https://doi.org/10.1586/14789450.2.5.793>
- [39] X. Lu, H. Zhu, Tube-gel digestion: a novel proteomic approach for high throughput analysis of membrane proteins, *Mol. Cell. Proteom.* 4 (2005) 1948–1958, <https://doi.org/10.1074/mcp.M500138-MCP200>
- [40] W. Dormeyer, D. van Hoof, S.R. Braam, A.J.R. Heck, C.L. Mummery, J. Krijgsvel, Plasma membrane proteomics of human embryonic stem cells and human embryonic carcinoma cells, *J. Proteome Res.* 7 (2008) 2936–2951, <https://doi.org/10.1021/pr800056j>
- [41] M. Ahrm, Z.I. Litou, R. Fang, G. Al-Tawallbeh, Estimation of membrane proteins in the human proteome, *Silico Biol.* 6 (2006) 379–386.
- [42] D.M. Dickinson, G.B. Marshall, G.J. Beran, M. Varela-Garcia, E.A. Mills, M.C. South, A.M. Cassidy, K.L. Acheson, G. McWalter, R.M. McCormack, P.A. Bunn, T. French, A. Graham, B.R. Holloway, F.R. Hirsch, G. Speake, Identification of biomarkers in human head and neck tumor cell lines that predict for in vitro sensitivity to gefitinib, *Clin. Transl. Sci.* 2 (2009) 183–192, <https://doi.org/10.1111/j.1752-8062.2009.00099.x>