RESEARCH ARTICLE

Infrared-based protein detection arrays for quantitative proteomics

Christian Loebke, Holger Suelmann, Christian Schmidt, Frauke Henjes, Stefan Wiemann, Annemarie Poustka and Ulrike Korf

Division of Molecular Genome Analysis, German Cancer Research Center, Heidelberg, Germany

The advancement of efficient technologies to comply with the needs of systems biology and drug discovery has so far not received adequate attention. A substantial bottleneck for the time-resolved quantitative description of signaling networks is the limited throughput and the inadequate sensitivity of currently established methods. Here, we present an improved protein microarray-based approach towards the sensitive detection of proteins in the fg-range which is based on signal detection in the near-infrared range. The high sensitivity of the assay permits the specific quantification of proteins derived from as little as only 20,000 cells with an error rate of only 5%. The capacity is limited to the analysis of up to 500 different samples per microarray. Protein abundance is determined qualitatively, and quantitatively, if recombinant protein is available. This novel approach was called IPAQ (infrared-based protein arrays with quantitative readout). IPAQ offers a highly sensitive experimental approach superior to the established standard protein quantification technologies, and is suitable for quantitative proteomics. Employing the IPAQ approach, a detailed analysis of activated signaling networks in biopsy samples and of crosstalk between signaling modules as required in drug discovery strategies can easily be performed.

Keywords:
Protein arrays / Protein quantification / Reverse phase protein microarray / Signaling networks

1 Introduction

According to the present paradigm of systems biology, biological signaling networks consist of modular and extensively interconnected elements. The elucidation of these complex networks is a key to a full understanding of cellular processes in health and disease. To this end, approaches to map complex signaling networks on the proteome level require the analysis of large sample numbers in parallel and with high sensitivity. The qualitative and quantitative determination of protein abundance, the time-resolved determination of protein turnover rates, and the measurement of protein modifications are of fundamental importance for the understanding of biological pathways and signaling networks [1]. For many years, the measurement of proteins and their PTMs has been performed by standard Western blotting [2]. Electrophoretic transfer of proteins to a hydrophobic membrane was originally developed for the analysis of single proteins [3, 4]. Improvements of this method include the chemiluminescence- or fluorescence-based detection [5], allowing for quantitative protein analysis [6]. However, there are considerable drawbacks of Western blot detection: first, the
amount of total protein consumed per blot is high. Second, Western blotting can not efficiently be adapted to a multiplexed format. Third, nonlinear gel runs and uneven blotting conditions influence the amount of protein transferred to the membrane. Finally, the transfer velocity strongly depends on the physicochemical properties of the proteins, and is therefore strongly dependent on a protein’s size, hydrophobicity, pI and folding properties [7]. Therefore, a total SD of 20–35% is commonly reported for quantitative Western blotting [8]. Despite the shortcomings, innovative technologies for reliable protein quantification are rare.

To overcome these limitations, we have set out to implement the RP protein array technology (RPPA or “lysate array” [9]) for robust quantification of proteins in high throughput format, and named this approach IPAQ (infrared-based protein arrays with quantitative readout). IPAQs carry serial dilutions of complex biological samples, e.g., cell lysates, biopsy samples, which are printed by noncontact piezo-spotting onto NC coated glass slides [10]. The detection of a protein of interest is performed by incubation with a specific antibody. The fraction of captured detection antibody is visualized with a secondary near-infrared (NIR) dye-labeled antibody, directed against its Fc-terminus (Fig. 1a). All of the microarrays are subsequently scanned in the NIR range and signal intensities of the digitized images are quantified. Here, we report on the adaptation of this method to quantitative protein measurements. With a minimal sample consumption of only 10 μL, the lowest amount of detectable protein is in femtogram (10^{-15} g) range. Intra- and interchip signal SDs are below 1 and 3.3%, respectively. In summary, the IPAQ technology is a useful tool for monitoring protein abundance and turnover of PTMs on a large scale, and is ideal for challenge-based analyses in systems biology.

2 Materials and methods

2.1 Cells and cell culture conditions

Human breast cancer cell lines MCF-7 (ATCC; HTB-22™), MCF-12F (ATCC CRL-10783), MCF-10F (ATCC CRL-10318), MCF-12A (ATCC CRL-10781), SK-BR-3 (ATCC HTB-30), T-47D (ATCC HTB-133), and BT-474 (ATCC HTB-20) were cultured under conditions recommended by the ATCC (ATCC serum & medium annotation, http://www.lgcpro mochem-atcc.com). The cells were split two to three times per week in a 1:3 ratio. EGF (Sigma, Munich, Germany) was applied at 25 ng/mL for the stimulation of the cell line MCF-7 after a 48 h period of starvation under serum-free conditions.

2.2 Printing of protein microarrays

Cells were trypsinized and 1 × 10^6 cells were lysed in 1 mL M-PER® reagent (Pierce, Rockford, USA) for 20 min at 4°C. The protein concentration was determined by BCA Protein Assay (Pierce) and adjusted to 2 μg/mL by diluting with M-PER reagent. For the total protein quantification, we used BSA standard (also from Pierce) and prepared ten dilutions in concentrations ranging between 0.05 and 2 μg/mL. Standard samples with recombinant human GST-ERK2 (Invitrogen, Karlsruhe, Germany) were prepared in the range from 16 pg/mL to 2 ng/mL.

Prior to printing, the samples were mixed 1:1 with 2 × Protein Arraying Buffer (Whatman, Brendford, USA). All the samples were printed onto NC FAST-Slides (Whatman) with a noncontact piezo spotter (4-pin head, Biochip-Arrayer, Perkin-Elmer, Wellesley, USA) delivering 333 pL per drop. All samples were spotted in quadruplicate and three drops were delivered per spot. At least two subarrays were printed resulting in eight spots per sample and slide.
2.3 Incubation and signal detection

Slides were blocked overnight at 4°C with a mixture of Odyssey blocking buffer (LI-COR, Lincoln, USA)/PBS, 1:3. Incubations with primary antibody were performed for 2 h in blocking buffer. After washing, a 1:10000 dilution of the IRDye™700DX-labeled secondary antibody (LI-COR) in PBS + 0.02% w/w NP40 was applied for 30 min. The slides were rinsed and dried at room temperature, scanned with the Odyssey NIR scanner, and signal intensities were detected with the Odyssey 2.0 software (LI-COR). The following primary antibodies were employed: anti-ERK-antibody (sc-94, rabbit) (Santa Cruz, Santa Cruz, USA), anti-ERK (06-182, rabbit) and anti-ER-antibody (05-820, rabbit) (Upstate, Charlottesville, USA), anti-STAT3-antibody (9139), and anti-STAT3 (Y705)-antibody (9131) (Cell Signaling Technologies, Danvers, USA), the antibody against EGF-receptor was also from CST. All of the antibodies were diluted 1:500 in 1:3 diluted Odyssey Odyssey blocking buffer.

2.4 Signal analysis and generation of a quantitative readout

The average intensity of each single spot was determined (pixels/cm²). Next, the mean intensity of the control spots was calculated and subtracted from the mean intensity of the sample spots to correct for background and/or noise due to unspecific antibody binding. Mean and SD were calculated for a set of four spots of a serial dilution. For the absolute readout, a calibration curve was generated employing a standard program.

2.5 Total protein quantification assay using Fast Green FCF

Fast Green FCF is commonly used for staining connective tissues in histological samples. This dye (absorption maximum at 625 nm) turned out to be detectable in the NIR range at 700 nm. The slides were rinsed for 1 min with PBS and stained for 2 h with Fast Green FCF staining buffer (0.005% w/w Fast Green FCF (Sigma) in 10% acetic acid, 30% EtOH, and 100% H₂O, v/v/v). Destaining was performed for 2 h with 10% acetic acid, 30% EtOH ad 100% H₂O, v/v/v, and the slides were scanned as described above.

3 Results

3.1 Microarray-based quantification of total protein

The total protein content of single spots on the microarray is determined using Fast Green FCF [11]. Fast Green FCF signals serve as the basis for the subsequent calculation of the total protein content of individual samples. Technically, the exact quantification of total protein on IP AQ requires a serial dilution of BSA covering the range of 0.1–1 μg/μL (Fig. 1). The calibration curve was linear over a concentration range between 1 μg/μL (1 ng BSA per spot) and 100 ng/μL (100 pg BSA per spot) with the latter value corresponding to the protein content of less than one cell (data not shown). The SD between identical spots on the same array was determined to be 1.96%. The Fast Green based quantification of the total protein content also serves as a basis for the normalization of the signals derived from incubations with specific antibodies.

3.2 Quantification of individual proteins

In order to validate our approach, we selected the protein kinases ERK1 and ERK2 and their dual phosphorylation on Threonine and Tyrosine as a test case. ERK1/2 is a key element of the mitogen-activated protein kinase (MAPK) pathway. This pathway plays a prominent role in the transduction of extracellular stimuli through specific cell surface receptors into the cell, followed by a cascade of phosphorylation events, ultimately leading to the activation of gene transcription. In particular, the MAPK pathway is downstream of the ras oncogene whose activation leads to the formation and progression of malignant tumors.

A serial dilution of recombinant GST-tagged ERK2, covering a concentration range of 0.25 to 2 ng/μL, was delivered as a calibrator protein onto NC-coated glass slides and visualized by incubation with an anti-ERK antibody and detection in the NIR range. The combination of Fast Green FCF staining data with data derived from the antibody-based detection permits the normalization of signal intensities, and thus the precise determination of ERK1/2 in the different samples. Six replicate slides were prepared to estimate the signal deviation within (intraslide) and across (interslide) slides from one production batch. In order to determine the intraslide signal variation, signals from four replicate spots on each of the six slides were averaged. The SD was calculated on the basis of the mean of corresponding spots from six slides (24 spots), which was set to the arbitrary value of “1”. Thus, the intraslide SD was calculated to be less than 1%, and the interslide SD to be lower than 3%. Therefore, we conclude that the low interslide SD allows the direct comparison of protein abundance data derived from different slides processed in parallel.

3.3 SD of IPAQ

In order to compare data from IPAQ with those yielded by standard techniques, we isolated total protein from seven human breast cancer cell lines (BT474, MCF-7, MCF-10, MCF-12A, MCF-12F, SK-BR-3, and T47D) and prepared Western blots with 20 μg of total protein for each sample. Two independent blots were incubated with anti-ERK1/2 antibody (Fig. 2), and signal intensities were determined with the LI-COR imaging system. An SD of 18.6% was observed which is in accordance with published data on Western blot techniques for protein quantification [8, 12].
Figure 2. ERK quantification by Western blotting; (a) lysates from different breast cancer cell lines were probed with anti-ERK antibody and analyzed after infrared scanning. (b) comparison of signal intensities from both Western blots. The SD calculates to 18.61%.

For the corresponding analysis by IPAQ, cell lysate samples from the same batch were printed onto NC slides, according to the experimental design described in Fig. 1. Each sample was spotted in quadruplicate and each subarray was printed three times, thus resulting in 12 data points per sample (Fig. 3a). The use of a general ERK1/2 antibody did not discriminate between ERK1 and ERK2, but this would certainly be possible with antibodies exclusively directed against a single form of both related proteins.

In addition, to determine not only the total ERK1/2 expression level (Fig. 3b), but also the activation state of ERK1/2 in the seven breast cancer cell lines, two slides were detected with an antibody recognizing the phosphorylation motif of ERK1/2 (Fig. 3c). With IPAQ, the SD for the determination of total ERK was below 0.1%. In contrast, the SD for the phosphorylated ERK1/2 was higher (3.1%), but still far lower than the signal deviation observed by Western blotting. The excellent slide-to-slide reproducibility of IPAQ makes this approach superior for the quantification of proteins and protein modification as demonstrated by the comparative analysis of the ERK1/2 activation status in seven breast cancer cell lines (Fig. 3d). In detail, ERK1/2 was approximately four-fold higher expressed in MCF-7 than in MCF-12F (Fig. 3b), but in MCF-12F the kinases ERK1/2 were much more strongly phosphorylated (Figs. 3c and d). In order to compare the phosphorylation rate between both cell lines, the pERK1/2 read-out for MCF-12F has to be multiplied to compensate for the different ERK1/2 expression.
level. In this way the activation state of differently expressed proteins can be compared directly (Fig. 3d). With respect to the comparison of breast cancer cell lines, the results suggest that in MCF-7 cells only a small subpopulation of ERK1/2 is constitutively phosphorylated, contrasting the phosphorylation state of ERK1/2 in the cell lines MCF-12A and MCF-12F. Thus, the IPAQ approach can be employed to determine the protein phosphorylation rates between different samples, as shown here, for breast cancer cell lines.

3.4 Determination of detection limit and data quality

The absolute quantification of a single target protein in a complex mixture can only be performed if a suitable recombinant protein is included for the calibration of signal intensities. Commonly, the recombinant protein is mixed with lysis- or SDS-buffer and added to the gel or array and processed along with the samples. For the IPAQ approach, we examined the influence of the sample buffer on the antibody-derived signal intensities. A serial dilution of recombinant inactive ERK2 was prepared in different buffers including cell lysis buffer with or without BSA, and in a diluted MCF-7 lysate as sample preparation buffer. Worth mentioning is the fact, that signal intensities decreased with increasing amounts of BSA added to the sample preparation buffer (Fig. 4b). For example, when the ERK1/2 calibration samples were prepared in a buffer containing 1 μg/μL BSA, the signal intensities were approximately 3.5-fold lower compared with those derived from ERK1/2 standard prepared in a buffer containing 0.5 μg/μL BSA (Fig. 4e).

Even more interestingly, the signals derived from cell lysates spiked with recombinant ERK resulted in the most consistent signal intensities (Fig. 4b): A three-step ERK2 dilution, corresponding to 250, 125, and 62.5 pg ERK2/μL, was mixed with MCF-7 cell lysate corresponding to a total protein concentration of 125 and 62.5 ng lysate/μL. Furthermore, in order to correct for endogenous ERK1/2 in the MCF-7 lysate, a serial dilution of MCF-7 lysate without the recombinant ERK1/2 was spotted as well (Fig. 4c). The signal intensities resulting from the MCF-7 samples were subsequently subtracted from those resulting from the calibration curve. Thus, a strong linear correlation between the amount of calibrator ERK1/2 and the signal intensity was obtained, indicating that spiking a diluted cell lysate with recombinant standard protein is the buffer of choice for precise protein quantification (Figs. 4b and d). Based on data from both standard curves, total ERK1/2 was calculated to be approximately 700 pg per μg of MCF-7 cell lysate. Moreover, this amount was in good agreement with results from quantitative Western blotting which was performed as control in two independent experiments. Data from quantitative Western blotting indicate an ERK1/2 concentration of 740 pg/μg

![Figure 4. Absolute ERK1/2 quantification with IPAQ; (a) breast cancer cell line lysate and different sample buffers were spiked with recombinant ERK2 and analyzed as described. BSA and MCF-7 lysate without ERK2 serve as controls; (b) protein concentration plotted against antibody-derived signal intensities; (c) signal intensities derived from MCF-7 cell lysate spiked with recombinant protein; (d) ERK1/2 quantification based on a serial dilution of ERK2 in diluted MCF-7 lysate. A total ERK abundance of 698 pg/μg lysate was calculated; (e) effect of total protein content on signal intensities shown for a fixed amount of ERK2 added to buffers with 0.25–1 μg/μL BSA.](圖像)
MCF-7 cells, with an SD of 17% (data not shown). Thus, the ERK1/2 concentration determined by Western blot and IPAQ differed by 5% only. However, the microarray approach resulted in numbers with a much lower SD, in this case only 0.1%.

3.5 Time resolved measurements

In order to demonstrate the usefulness of IPAQ for the quantification of signal transduction events in a time-resolved manner, MCF-7 cells were stimulated with epidermal growth factor (EGF). Samples were collected after stimulation with EGF after 0 (negative control), 0.5, 1, 2, 3, 5, 8, 12, 15, 20, 25 and 30 min resulting in 12 time points. The cells were lysed and used for the preparation of protein microarrays. The detection of the total protein content was performed with the Fast Green FCF staining approach, described above. Specific antibodies against EGF receptor (EGFR), estrogen receptor (ER), total ERK1/2, activated ERK1/2, as well as total and activated STAT3 were used to monitor the effect of EGF on intracellular signaling. The signals for EGFR, ER, total ERK1/2, and STAT3 expression were constant throughout the experiment. After 0.5 min, a sharp increase of ERK activation was observed. Maximal activation was reached after 3 min and decreased to reach the level of nonstimulated cells after 20 min (Fig. 5). The ERK1/2 phosphorylation in stimulated cells was found to be 4.5-fold higher compared to unstimulated cells.

4 Discussion

Protein arrays can be used to determine protein abundance reliably. Lysate arrays, also named RP arrays, are the most promising technological approach. In contrast to protein detection arrays [13, 14], requiring two different antibodies directed against the same protein, lysate arrays require only one antibody per analyte or PTM. Therefore lysate arrays are a promising tool for large-scale applications. In contrast to commercially available protein microarray platforms, a labeling reaction of the analyte mixture was omitted. The introduction of a covalently bound fluorescent dye may target a given analyte in different ways, and thus introduce artifact epitopes, or cover an existing epitope in a subset of the analyte population. Therefore, label-free approaches are preferred in protein quantification.

Therefore, we optimized the experimental approach of RP protein arrays by introducing new steps including a modified sample preparation step, and an optional calibration step based on a serial dilution of specific recombinant proteins. Thus, the calculation of protein copy numbers is feasible. The simultaneous printing of multiple slides facilitates a multiplexed quantification of different target proteins and/or their PTMs.

Our approach, named IPAQ, has the potential to overcome the limitations of Western blotting for the quantitative analysis of protein abundance and protein modification in biological processes. It will also be useful for monitoring time dependent events in signal transduction.

We implemented the Fast Green FCF-based total protein quantification assay on IPAQ, which allows for the reliable determination of proteins down to a total protein concentration of 100 ng/μL. The high sensitivity of the Fast Green FCF staining permits the quantification of total protein in a complex protein mixture derived from as little as only 20000 cells, resulting in 2 ng to 100 pg protein deposited per spot. This strategy allows for normalization of the signal intensities and correction for technical variances due to inhomogeneous spotting of samples.

By comparing the relative expression level of active (phosphorylated) and inactive (nonphosphorylated) ERK1/2 in different breast cancer cell lines, we demonstrated that a comparative analysis of protein abundance can be performed with high reproducibility. The SD of signal intensities from microarray experiments is several-fold lower than that achievable with Western blotting protocols: Western blot experiments typically result in an SD of 20–30%, whereas the SD of IPAQ was found to be less than 5%.

We demonstrated that a read-out in the NIR-range allows a specific detection down to 16 fg of specific protein per spot. Furthermore, the quantification of ERK1/2 in the cell line MCF-7 resulted in a concentration 700 pg ERK1/2 per μg MCF-7 lysate, thus leading to an ERK1/2 concentration of 140 fg/μL, or approximately $2 \times 10^6$ molecules ERK1/2 per MCF-7 cell. Thus, the detection of 16 fg corresponds to 11% of the endogenous ERK1/2 in a single cell. The detection of such small amounts of protein per spot is favored by the
excellent S/N of fluorescent dyes in the NIR range, where the autofluorescence of NC membranes is lower than in the visible range [15].

In summary, we demonstrated that IPAQ is a preferable tool for the large-scale analysis of protein abundance in biological samples, and ideal for robust measurements as required in systems biology. In addition, sample volumes and detection threshold are 50-fold and 105-fold, respectively, lower than those achievable with conventional methods. This is particularly true when samples are spiked with a recombinant protein. The sample capacity of microarrays exceeds that of Western Blotting and MS by magnitudes. Due to the high sample capacity, IPAQ has the potential to burst the time-resolved analysis of signal transduction networks as well as the analysis of valuable biopsy samples. In contrast to protein capture arrays [13, 14, 16] IPAQ needs only a single antibody per analyte, thus reducing the effort required for the selection of assay-suitable antibodies. Extending the IPAQ approach to the proteome-wide quantification of target proteins is currently restricted only by the limited availability of sufficiently specific antibodies and of recombinant proteins. However, focusing on the analysis of signaling pathways, where the required tools and standards are available, will definitely lead to a better understanding of biological systems.

We thank Sabrina Schumacher and Julia Büttner for excellent technical assistance. This work was supported by a grant from the National Genome Research Net (NGFN) through the German Federal Ministry for Education and Research (grant no. 01GR0418 to H.S. and A.P., grant no. 01GR0420 to S.W.).

5 References