Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about *Accepted Manuscripts* in the **Information for Authors**.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard <u>Terms & Conditions</u> and the <u>Ethical guidelines</u> still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



www.rsc.org/loc

Multiplexing of Miniaturized Planar Antibody Arrays for Serum Protein Profiling – Biomarker Discovery in SLE Nephritis

Linn Petersson¹, Linda Dexlin-Mellby¹, Anders A. Bengtsson², Gunnar Sturfelt², Carl A.K. Borrebaeck¹, Christer Wingren¹

¹Dept. of Immunotechnology and CREATE Health, Medicon Village, Lund University, SE-22381 Lund, Sweden.

²Dept. of Clinical Science, Section of Rheumatology, Skåne University Hospital, Lund University, SE-22185 Lund, Sweden.

*Corresponding author:	Christer Wingren
	Dept. of Immunotechnology
	Lund University
	Medicon Village, Building no. 406
	SE-223 81 Lund
	Sweden
	Phone: +46-46-2224323
	E-mail: Christer.wingren@immun.lth.se

Key words: Miniaturization, antibody arrays, SLE, biomarker,

Abbreviations:	AUC	- area under the curve
	CV	- coefficient of variation
	DPN	- dip-pen nanolithography
	LOD	- limit of detection
	Ni ²⁺ -NTA	- Nickel ²⁺ -nitrilotriacetic acid
	ROC	- receiver operating characteristics
	scFv	- single-chain fragment variable
	SLE	- systemic lupus erythematosus
	SLEDAI-2K	- SLE disease activity index 2000

ABSTRACT

In the quest of deciphering disease-associated biomarkers, miniaturized and multiplexed antibody arrays may play a central role for generating protein expression profiles, or protein maps, of crude serum samples. In this conceptual study, we have explored a novel, 4-times larger pen design, enabling us to in a unique manner simultaneously print 48 different reagents (antibodies) as individual 78.5 um^2 (10 μ m in diameter) sized spots at a density of 38,000 spots/cm² using dip-pen nanolithography technology. The antibody array set-up was interfaced with a high-resolution fluorescent-based scanner for sensitive sensing. The performance and applicability of this novel 48-plex recombinant antibody array platform design was demonstrated in a first clinical application targeting SLE nephritis, a severe chronic autoimmune connective tissue disorder as model disease. To this end, crude, directly biotinylated serum samples were targeted. The results showed that the miniaturized and multiplexed array platform displayed adequate performance, and that SLE-associated serum biomarker panels reflecting disease process could be deciphered, outlining the use of miniaturized antibody arrays for disease proteomics and biomarker discovery.

INTRODUCTION

Miniaturized and multiplexed immunoassays [1], represented by antibody-based microarrays, has become a key technology for protein expression profiling and biomarker discovery [2-5]. Targeting crude serum samples, representing the most common clinical sample format, candidate biomarker signatures reflecting e.g. bladder cancer [6], pancreatic cancer [7, 8], and systemic lupus erythematosus (SLE) [9], has been deciphered. However, aiming for global serum proteome screening (>>10,000 targets) [2, 10], antibody microarray characteristics, such as spot (array) size, spot density, and multiplexity, will be critical.

To date, antibody microarrays with an overall foot print of $< 1 \text{ cm}^2$, based on $18 \times 10^3 \text{ µm}^2$ (diameter (\emptyset) of $\sim 150 \text{ µm}$) sized spots at a density of $\leq 2,000 \text{ spots/cm}^2$, and a multiplexity of < 825 different antibodies/array, have mainly been produced and applied [2-5]. Using ink-jet printers to produce the arrays, the antibodies have been spotted in parallel (1 to 4 antibodies), and the multiplexity (the number of spotted reagents (e.g. antibodies)) has been achieved by washing the nozzles and loading them with new probes. In an attempt to simultaneously scaledown (spot size; $< 0.8 \text{ µm}^2$, $\emptyset < 1 \text{ µm}$) and scale-up (spot density; $>100,000 \text{ spots/cm}^2$) the array set-up, conceptual antibody (protein) nanoarrays have been designed [11-21], for review see [22]. Albeit successful, these nanoarray layouts have been found to be associated with three key technological restrictions. Firstly, reducing the spot size below 1 µm will lead to impaired rather than improved assay performance (e.g. sensitivity) [23]. Secondly, the production technologies are, so far, compatible with generating only 1-plex arrays (in rare cases < 5-plex layouts) [11-22, 24-26]. Thirdly, methods (hardware) for sensitive (fluorescent-based) read-out of highdensity nanoarrays remain to be established [22].

These key limitations might, however, be resolved by aiming for miniaturized arrays based on submicron-sized (\emptyset 10 µm) rather than nano-sized (\emptyset < 1 µm) spot features. To this end, a desk-top printer, denoted NanoArrayer 3000, based on dip-pen nanolithography (DPN), has been developed capable of simultaneously inking 12 different reagents, resulting in 0.78 to 78.5 µm² (\emptyset 1 to 10 µm) sized spot features, depending on the tip-surface contact time [27]. As demonstrated for a 1-plex monoclonal antibody array set-up, 10 µm sized spots could readily be visualized using a conventional high-resolution fluorescent-based scanner [28].

Adopting the NanoArrayer 3000 printer, we have recently designed the first generation of a 12-plex planar recombinant antibody array technology platform, based on 78.5 um^2 (\emptyset 10µm) sized spots at a density of 38,000 spots/cm², interfaced with a fluorescent-based read-out [29]. The feasibility of the array platform for reproducible and sensitive profiling of crude, directly labelled serum samples was indicated. Albeit successful, the multiplexity of the arrays

could only be increased in incremental steps of 12 by loading the printer with a new 12-DPN pen and a 12-inkwell prepared with the next set of reagents. Replacing the pen is doable, but the pen then has to be re-positioned, re-orientated, and re-aligned prior to continuing the printing process or the array layout will be compromised. Hence, the size (multiplexity) of the pen will be important when aiming for high-density arrays.

In this conceptual study, we have explored a novel, 4-times larger DPN pen design, enabling us to in a unique manner simultaneously print 48 different reagents (antibodies) as individual 78.5 um^2 (\emptyset 10µm) sized spots at a density of 38,000 spots/cm². The performance and applicability of this novel 48-plex array platform design was then demonstrated in a first clinical application targeting SLE nephritis. SLE nephritis is a most severe form of SLE [30], a chronic autoimmune connective tissue disease, involving the kidneys for which the lack of high performing serum biomarkers represents a key unmet clinical need [31]. Hence, novel biomarker panels, rather than single markers [32], for diagnosis and prognosis as well as monitoring of disease activity would be of great clinical importance [31]. The results showed that our miniaturized 48-plex planar recombinant antibody array set-up could be used to decipher disease-associated serum biomarkers, paving the way for novel miniaturized biomarker discovery initiatives.

EXPERIMENTALS

Clinical Samples

In total, 75 serum samples (denoted S1 to S75) were used in this study. The samples were collected at the Department of Rheumatology, Skåne University Hospital, Lund, Sweden, including 45 SLE nephritis patients and 30 healthy serum samples (Table 1). The SLE patients all suffered from glomerulonephritis, and were collected during both flares and remissions. The clinical disease activity was defined using SLE disease activity index 2000 (SLEDAI-2K) score [33]. The SLE and healthy sample cohorts were matched with respect to gender and age. The gender distribution within a cohort reflected the anticipated distribution. All samples were aliquoted and stored at -80°C until further use. The levels of serum C1q were determined using electroimmunoassay (rocket immunoelectrophoresis). The study was approved by the regional ethics board in Lund, Sweden, and performed in compliance with relevant laws and institutional guidelines.

Labeling of clinical samples

Page 5 of 28

Lab on a Chip

The serum samples were biotinylated using EZ-Link Sulfo-NHS-LC-Biotin (Pierce, Rockford, IL, USA) according to a previously optimized labeling protocol for serum proteomes [34-36]. Briefly, 5 µl serum aliquots were centrifuged at 16 000g for 20 min at 4°C and diluted 1:45 in PBS, resulting in a concentration of about 2 mg/ml. Next, cholera toxin subunit B (CT) (Sigma, St. Louis, MO) was spiked into the serum samples to a final concentration of 1.2 µM. The samples were then biotinylated by adding sulfo-NHS-biotin to a final concentration of 0.6 mM and incubated on ice for 2 h, with careful mixing every 20th minute. The samples were labeled at a molar ratio of biotin: protein (15:1) assuming an average molecular weight of 50 kDa for the serum proteins [37]. Unreacted biotin was removed by dialysis against PBS (pH 7.4) for 72h using a Slide-A-Lyser (MWCO 3.5 kDa) (Pierce). The samples were aliquoted and stored at -20°C until further use. Noteworthy, major efforts have been made by us and others to optimize the labelling (biotinylation) of crude samples (e.g. serum), to ensure reproducible and consistent labelling, and as broad coverage as possible of both low- and high-abundant proteins, for review see. Of note, major efforts have been made by us [34-36] and others [38] to optimize the labelling protocol of crude serum samples to ensure reproducible and consistent labelling, providing as broad coverage as possible of both high and low-abundant proteins, for review see [39]. While the reactivity of some antibodies might be lost since the label could block the affinity binding to the antibodies (epitope masking), we have bypassed this problem, as in this study, by frequently including more than one antibody against the same protein, but directed against different epitopes [3].

Antibodies

Our approach is based on using the immune system as a sensor for disease [3, 9]. To this end, thirty-nine human recombinant single-chain fragment variable (scFv) antibodies, directed against 28 different, analytes mainly involved in immunoregulation, anticipated to reflect events taking place in SLE, and one non-human serum protein, selected from an in-house designed phage-display library, were used as probe source (Table 2) [40]. One intact mouse monoclonal antibody directed against interferon alpha (INF- α) was purchased from Abcam, (Cambridge, UK).

Production and purification of scFv antibodies

All scFv:s were produced in 100 ml of *E. coli* cultures and purified from expression supernatants using affinity chromatography on Ni^{2+} -nitrilotriacetic acid (Ni^{2+} -NTA) agarose (Qiagen, Hilden, Germany). Bound molecules were eluted with 250 mM imidazole (pH 8),

extensively dialyzed against PBS (pH 7.4), and stored at 4°C until further use. The degree of purity and integrity of the scFv antibodies were evaluated using 10% SDS-PAGE (Invitrogen, Carlsbad, CA, USA). The protein concentration was determined by measuring the absorbance at 280 nm (average concentration 0.37 mg/ml, range 0.10 to 0.92 mg/ml).

Production and handling of scFv arrays

The scFv arrays were produced using a desktop nanofabrication system, denoted NanoArrayer 3000, based on DPN technology (NanoInk Inc, Skokie, IL, USA). Prior to dispensing, the tips, DPN Probes Type M-ED Side M2 with 48 modified "A frame" cantilevers (NanoInk Inc.), were plasma cleaned for 1 min at low RF-value, using a gas mixture of Oxygen/Argon (20%/80%) at 1.4 bar using a Plasma Cleaner (PDV-002) (Harrick Plasma, Ithaca, NY, USA). The antibodies (≤0.5 mg/ml) were diluted in printing buffer (NanoInk Inc.), and Alexa-Fluor-555 cadaverine disodium salt (Invitrogen) was added to a final concentration of 56 µg/ml (position marker). Then, 400 nL of the antibody printing buffer solutions were loaded into inkwell cartridge (NanoInk Inc.) and the DPN tips were subsequently inked (dipped) into the antibody solutions by the NanoArrayer 3000. The antibodies were printed on Nexterion Slide E (Schott AG, Mainz, Germany), with a contact time of 200 ms, resulting in 10 µm-sized spots. The dispensing was performed at ambient temperature and humidity. Eight replicates of each scFv were arrayed to ensure adequate statistics. Each slide could hold 18 subarrays, arranged in a 3x6 format (to fit the slide module assembly used), where each subarray was functionalized with 39 different scFv fragments, one positive control (alexa-647 labelled BSA), one positive control scFv against CT, (printed in duplicates at 3 concentrations (0.42 mg/ml, 0.32 mg/ml and 0.22 mg/ml)) and one negative control (PBS). The produced scFv array slides were then placed in a slide module assembly (NanoInk Inc) and dried over night at room temperature. Next, the arrays were blocked with 5% (w/v) fat-free milk powder (Semper AB, Sundbyberg, Sweden) in PBS for 1 h. All incubations were performed at room temperature. The arrays were washed three times with 0.05% (v/v) Tween-20 in PBS (PBS-T) and incubated with 70 µl biotinylated serum sample, diluted 1:10 (resulting in a total serum dilution of 1:450, corresponding to about 0.2 mg/ml) in 1% (w/v) fat free milk powder and 1% (v/v) Tween-20 in PBS (PBS-MT) for 1h. Next, the arrays were washed three times with PBS-T, where after the arrays were incubated with 100 uL 1 µg/mL Alexa-647 conjugated Streptavidin (Invitrogen) in PBS-MT for 1 hour. Finally, the arrays were washed three times with PBS-T and one time with PBS, and directly dried under a stream of nitrogen gas and immediately scanned.

Analysis of miniaturized scFv arrays

The arrays were scanned using a high-resolution fluorescent scanner, denoted InnoScan®900 (Innopsys, Carbonne, France) at five different scanning setting (detection gain at 635nm: 5% 15%, 50%, 80%, and 100%). All arrays were scanned with 1 μ m/pixel resolution and the Mapix software (V.5.5.0) (Innopsys) was used to quantify the median value of each spot using the fixed circle method. The local background was subtracted (median value minus local background signal) and to compensate for any possible local defects, the two highest and the two lowest signal intensities were automatically excluded, and each data point represents the median value of the remaining four replicates. For protein analytes displaying saturated signal intensities, values from lower scanning settings were chosen. The samples (n=75) were analyzed on 81subarrays (including duplicates), on 7 different slides (denoted A to E) during in total 4 days. One sample was later excluded from the study due to high background signals. The antibody array reports back the relative level of each targeted analyte.

Data handling and normalization

Only non-saturated spots were used for data analysis, and any signals below limit of detection were set to 1.1, since only values above zero can be handled in the subsequent normalization step. Log10 values of the signal intensities were used for further analysis. Four different chip-to-chip normalization strategies were performed and evaluated, including i) anti-CT normalization [34], ii) semi-global normalization [9, 36], iii) subtract by geometric group mean normalization, and iv) divide by geometric group mean normalization.

In the case of the anti-CT normalization, an internal standard protein (CT) was spiked into the serum samples prior to the labeling procedure[34]. The normalization factor N_i was calculated for each sample (i) by the formula N_i= $S_{(CT17)}/\mu_{(CT17)}$, where $S_{(CT17)}$ was the average signal intensity from the anti-CT scFv for each chip and $\mu_{(CT17)}$ was the average signal from anti-CT from all the samples. The anti-CT was printed in duplicates at three concentrations (0.22, 0.33, and 0.42 mg/ml), but the normalization factor N_i was based on the average value from the two highest concentrations of anti-CT (in total four values) from each chip, since the lowest concentration of anti-CT was below the detection limit. Chip-to-chip normalization was then performed by dividing the signal from each sample with the normalization factor.

In the semi-global normalization method, the coefficient of variation (CV) was calculated for each analyte and ranked [36, 41]. Thirty percent of the analytes that displayed the lowest CV-values over all samples were identified, corresponding to 12 analytes, and used to calculate a chip-to-chip normalization factor. The normalization factor N_i , was calculated by the formula $N_i = S_i / \mu$ where S_i is the sum of the signal intensities for the 12 analytes for each sample and μ is the sum of the signal intensities for all 12 analytes over all samples. Each data-set generated from one sample was divided with the normalization factor N_i .

In the subtract by geometric group mean approach, a mean value of each analyte per chip was calculated, resulting in a normalization factor N_i per antibody and chip. Chip-to-chip normalization was then performed by subtracting the normalization factor N_i from the signal intensity for each sample and antibody.

The divide by geometric group mean approach was similar to that of subtract by geometric group mean, but instead of using subtraction, the chip-to-chip normalization was performed by dividing the signal intensity for each sample and antibody with the normalization factor N_i .

After each normalization method, the positive- (BSA, CT) and negative- (PBS) controls were removed from the dataset. The limit of detection (LOD) was defined as the analyte concentration corresponding to negative control plus 2x standard deviations. An antibody had to display detectable signal intensities in \geq 30% of all analyzed samples in order to be included in the dataset. This resulted in that nine antibodies (II-1 α (1), IL-2, IL-4 (1), IL-4 (2), MCP-4, INF- γ , C3, C5 and INF- α (monoclonal)) were removed from the dataset. One SLE nephritis sample was also removed from the dataset due to high non-specific background-binding.

Data analysis

The normalizations strategies were evaluated using principle component analysis (PCA) in Qlucore (Qlucore AB, Lund, Sweden). The clustering of the samples were visualized based on day-to-day variations and slide-to-slide variations. The ANNOVA test was used to determine the number of significant antibodies (p<0.05) between days and slides for each normalization strategy.

Data classification was carried out with a support vector machine tool (SVM) in R [42-44], a supervised learning method generating prediction values as output. The classification was performed using a linear kernel, and the cost of constraints was set to 1, which is the default value in the R function SVM. The SVM was trained using a leave-one-out cross-validation procedure. Briefly, the training sets were generated in an iterative process in which the samples were excluded one by one. The SVM was then asked to blindly classify the left out sample as either disease or healthy, or different disease activities, and to assign a SVM decision value, which is the signed distance to the hyperplane. These decision values were then used to construct a receiver operating characteristics (ROC) curve, and the area under the curve (AUC) was calculated. Significantly up- or down- regulated analytes (p<0.05) were identified using Wilcoxon's signed-rank test, and visualized by box plots, using R.

RESULTS AND DISCUSSION

In this conceptual study, we have designed and produced the first 48-plex miniaturized planar recombinant scFv antibody microarray layout using DPN-based printing technology, by exploring and exploiting a novel 48-plex DPN pen design (Fig. 1). Hence, the degree of multiplexity was successfully increased by a factor of 4, compared to our previous miniaturized array set-up [29]. The performance and applicability of the set-up for serum protein profiling was then demonstrated by targeting crude, directly biotinylated serum samples, including 45 SLE nephritis patients and 30 healthy controls (Table 1).

Design, production and characterization of the 48-plex planar array set-up

First, the 48-plex inkwells were loaded with the printing reagents, including 39 recombinant scFv antibodes, 1 intact, full-length monoclonal antibody, 6 control scFv antibodies (duplicates of 1 antibody clone at 3 different concentrations), 1 positive control (fluorescently labelled BSA), and 1 negative control (PBS) (Fig. 2). The inkwells were individually filled with 400 nl of each antibody (0.07-0.69 mg/ml) in printing buffer, and the 3µm high tips were loaded with probe at the small micro channel interface. Compared to our conventional recombinant scFv antibody microarray set-up (30ul, \leq 0.50 mg/ml) [8, 9], the array production step consumed about 75 times less antibody. In other words, 75 times more arrays could be produced using the same amount of antibodies, thereby increasing the overall throughput.

Next, the 48 printing reagents were simultaneously deposited in 8 replicates on Nexterion E slides, resulting in a printing pattern of 8x48 features per subarray, with 18 subarrayes per slide. The overall footprint of a single subarray was ~0.9 mm². To the best of our knowledge, this represents the highest degree of multiplexity that can be simultaneously printed for miniaturized antibody arrays. The multiplexity of the arrays can then be extended in incremental steps of 48 by loading the printer with a new 48-DPN pen and a 48-inkwell prepared with the next set of reagents.

We used a tip-surface contact time of 200 ms, resulting in 10 μ m-sized spots (78.5 μ m²). The spots were orientated at a pitch-to-pitch distance of 40x66 μ m, corresponding to an array density of 38,000 spots/cm². The spot density could be increased to ~510,000 spots/cm² by

minimizing the pitch-to-pitch distance ($14x14 \mu m$). Hence, the size (area) of the individual spot features was reduced 225 times and the array density was increased at 19 times or more, compared to conventional antibody microarrays [3, 8, 9]. By reducing the spot size, the increased spot density was thus well within the range of what might be required even for large-scale serum protein expression profiling efforts [3].

The printing efficiency was evaluated by examining the spot morphology (shape and homogeneity) of the positive control, i.e. directly labelled BSA, based on in total 4,536 spots in 81 subarrays on 7 slides. The results showed that the spot morphology was adequate (visual inspection) (mean spot diameter of 10.8 μ m, coefficient of variation (CV) of 16%) with consistent spot signal intensities, illustrated by a CV of 12%. The performance of the printing process was thus found to be adequate.

Next, the printing efficiency of functional scFv was evaluated by examining the spot morphology (shape and homogeneity) and spot size of printed scFv antibodies observed after capture of biotinylated targets in crude serum samples. The results showed that low-nonspecific background binding, dynamic signal intensities, and adequate spot morphologies (visual inspection) were obtained (Fig. 2). The spot diameter was evaluated for 31 scFv antibodies printed in 8 replicates per subarray, including 81 subarrays on 7 slides, The results showed that the spot diameter ranged between 5.2 and 12.8 μ m (average 8.0), reflecting both antibody concentration and signal intensity (i.e. amount of analyte) (data not shown), giving a coefficient of variation (CV) of 19%. Hence, the printing efficiency of functional scFv was found to be adequate.

The current miniaturized recombinant antibody array layout, based on 40 antibodies, was designed to target 29 high- (μ M to nM range) and low-abundant (pg/ml) serum analytes (Table 2). Antibodies, targeting mainly immunoregulatory analytes, anticipated to reflect the events taking place in SLE nephritis were included on the array, a strategy well supported by previous profiling results [41] (Wingren *et al*, unpublished observations). The limit of detection (LOD) of the array set-up was defined as the analyte concentration corresponding to a signal 2x standard deviations above the negative control. An antibody had to display detectible signal intensities in \geq 30% of all analyzed serum samples (n=74 plus 7 duplicates, 1 sample excluded due to high background signals), in order to be included in the data set. This resulted in that nine antibodies, directed against II-1 α (1), IL-2, IL-4 (1), IL-4 (2), MCP-4, INF- γ , C3, C5 and IFN- α (monoclonal)), were excluded from the data set. Hence, 31 antibodies were used in all further analysis, thus generating protein expression profiles, or protein maps, for 22 unique serum analytes in each sample targeted (see Fig 2).

In total, 70 µL biotinylated serum sample, diluted 450 times, thus corresponding to 0.16 µL starting sample, was consumed per array. Compared to conventional antibody microarrays [3, 8, 9], similar, or smaller, sample volumes were consumed, while targeting at least 19 times more analytes (due to the increased multiplexity, i.e. spot density). It should, however be noted that gaskets, or reaction chambers, of similar sizes were used, while the reduced array foot print was not exploited, outlining distinct opportunities for reducing also the amount of sample required. This is a key feature since the amount of clinical sample often is a limiting factor. The small sample sizes needed opens up novel opportunities for analyzing minute amount of clinical samples, such as fine needle aspirates etc. Taken together, the results outlined the wealth of data that could be generated using miniaturized antibody arrays, while consuming minute amounts of a clinical sample. In addition, the array assay took less < 4 hours to run, further

Normalization of miniaturized antibody arrays

demonstrating the feasibility of the set-up.

The serum samples (n=74 plus 7 duplicates, 1 sample excluded due to high background signals) were randomized and analyzed on 7 slides over 4 days. In order to be able to compare array data generated on different slides and days, the data must be normalized, a procedure relatively well established for conventional antibody microarrays [3, 8, 9, 45], but not yet evaluated for miniaturized antibody arrays. Here, we evaluated 4 strategies for normalizing the data set, including i) anti-CT normalization (based on a spike-in protein), ii) semi- global normalization, iii) subtract by geometric group mean, and iv) divide by geometric group mean.

To this end, we first mapped the day of analysis and number of slide on to the data set, and determined any bias introduced by slide-to-slide and/or day-to-day variations in the raw, non-normalized data set (Fig. 3A). Principle component analysis (PCA) showed that the raw data grouped according to day and slide, and that 31 of 31 antibodies displayed significantly different signal intensities (ANNOVA test, p < 0.05). Hence, this clearly warranted the introduction of a data normalization step.

In the case of both anti-CT normalization (cfs. Figs. 3A and 3B) and semi-global normalization (cfs. Figs. 3A and 3C), the results showed that the data still grouped according to day-to-day and slide-to-slide variations, and that the number of antibodies displaying significantly differential signal intensities was maintained. In addition, by comparing the intensity profile over all samples per antibody, as illustrated for the IL-1 α antibody (Fig. 3A), showed that the normalization strategies had a significant impact on the intensity profiles, which risk transforming the data for some samples in a detrimental manner (cfs. Figs. 3A to 3C). Hence, these two normalization strategies were rejected.

In contrast, the samples did not cluster according to day-to-day and slide-to slide variations when either subtract by geometric group mean (Fig. 3D) or divide by geometric group mean (Fig. 3E) were used for normalization. In addition, no antibodies were longer found to display differential binding patterns (cfs Figs. 3A, 3D, and 3C). While both normalization strategies resulted in condensed intensity profiles, the overall pattern was maintained, indicating the feasibility of both normalization strategies (cfs Figs. 3A, 3D, and 3C). We selected the divide by geometric group mean approach as the main normalization strategy throughout the remaining part of the paper.

Reproducibility of miniaturized antibody arrays

The technical reproducibility of the entire miniaturized antibody array set-up was evaluated by determining the spot-to-spot and chip-to-chip variability after capture of biotinylated targets in crude serum samples. The data was reported in terms of CV values.

First, the spot-to-spot reproducibility was assessed for all antibodies defined to give a detectable signal (n=31), deposited in 8 replicates per subarray targeting 74 serum samples (1 sample was excluded due to high background signals) (18,352 individual data points). The results showed that the overall median CV-value, based on all antibodies, was 4.0% (range 0 to 11.4%) (Fig. 4A). If the antibodies were instead evaluated one-by-one, the median CV-values were still consistent, 3 to 6%, although the range (minimum to maximum) differed (Fig. 4B), indicating on antibody clone and/or sample (target) dependent variability. As for example, the antibody against CD40 often displayed larger variability. Noteworthy, this novel 48-plex set-up was found to display a spot-to-spot variability better than, or similar to, that displayed by 12-plex miniaturized antibody arrays (mean overall CV-value of 12%, range 5 to 27%) [29]. In comparison, the CV-value has been found to be <30% for conceptual antibody nanoarrays [11-21, 28], while conventional antibody microarrays (more established platforms) regularly display CVs of < 10% (recombinant antibody microarrays) [46] and < 20% (polyclonal and monoclonal antibody microarrays [47, 48].

Next, the chip-to-chip reproducibility was assessed by analyzing the same serum sample on 5 different slides, and determining the overall median CV-value for all antibodies after the data had been normalized. Only antibodies giving a detectable on at least 4 of the slides were included (n=25). The results showed that a median CV-value of 5.5% (range 1.0 to 15.7%) (subtract by geometric group normalization) (Fig. 4C) and 5.3% (range 1.2 to 10.2%) (divide

by geometric group mean) (Fig. 4D). Some antibodies (against e.g. CD-40 and IL-13) were found to display larger variability, again indicating on antibody clone and/or sample (target) dependent variability (only 1 sample analysed in this evaluation). In comparison, the CV-values of conventional antibody microarray set-ups have regularly been found to be in the same range [2-5], or better (1.6%) (Wingren *et al*, unpublished data). Hence, the miniaturized array platform was found to display adequate reproducibility.

Serum biomarker discovery in SLE nephritis

Next, the applicability of this novel 48-plex miniaturized antibody array set-up was demonstrated by performing serum biomarker discovery in SLE nephritis. SLE nephritis is one of the most severe manifestations of SLE involving the kidneys [30], for which the lack of validated serum biomarkers represents an unmet clinical need [31, 49, 50]. To this end, we profiled 75 crude serum samples (1 SLE sample excluded due to high background signals), including 45 nephritis patients and 30 healthy controls (Fig. 5). The data was normalized using the divide by geometric group mean strategy. The observed protein expression profiles, covering 22 unique serum proteins present at detectable levels (targeted by 31 antibodies) (see above) were then compared, and any differentially expressed proteins, i.e. potential biomarkers, were deciphered.

First, the serum protein profiling data showed that SLE nephritis vs. healthy controls could be classified running an SVM leave-one-out cross-validation, as illustrated by a ROC AUC value of 0.69 (Figs. 5A and 5B). Using Wilcoxon signed-rank test, the data showed that 3 of 31 (10%) of the antibodies showed significantly differential (p < 0.05) binding patterns (Figs. 5A and 5C). In accordance, we have previously demonstrated differentially expressed serum levels of C1q (down-regulated), IL-6 (up-regulated), and LDL (up-regulated) in lupus nephritis vs. healthy controls, using conventional recombinant antibody microarrays [9]. In fact, de-regulated levels of e.g. C1q and IL-6 has been frequently reported [51-56]. It should be noted that similar results were obtained if the data was normalized using the subtract by geometric group mean strategy (Supplementary Fig. S1). Hence, relevant disease-associated biomarker panels were defined using the multiplex, miniaturized planar recombinant antibody arrays. In this context, it might be of interest to note that it is challenging to pin-point serum biomarkers, single or panels, for SLE (as well as for other diseases) using conventional (proteomic) technologies due to issues with sensitivity, specificity, multiplexity, and resolution, giving the present approach a technological advantage.

Lab on a Chip Accepted Manuscript

The disease activity of SLE can be assessed using various validated activity indexes, of which the SLE disease activity index-2000 (SLEDAI-2K), based on recording 24 descriptors in 9 organs, weighted according to severity [33], represents a commonly used index [50, 57, 58]. Here, we examined whether serum biomarkers reflecting disease activity could be deciphered, that in the long-term run might provide novel, complementary means of monitoring, evaluating, and even predicting disease activity. To this end, the SLE nephritis patients were first divided into two subgroups, reflecting lower disease activity (SLEDAI-2K of 3 to 14) and higher disease activity (SLEDAI-2K of 16-32), and their serum protein expression profiles were compared to those of the healthy controls. While the lower-disease activity cohort vs. healthy controls were difficult to distinguish (ROC AUC of 0.60), the higherdisease activity cohort vs. controls were readily classified (ROC AUC of 0.80) (Figs. 5A and 5B). In accordance, we and others have previously identified the levels of e.g. C1q and IL-6 to be associated with disease activity [9, 51-53, 55, 56]. Again, similar results were obtained whether the data were normalized using divide by geometric group mean or subtract by geometric group men strategy (cfs. Figs. 5 and Supplementary Fig. S1). The low number of differentially expressed analytes that were identified [9] might be explained by heterogeneous samples, too low sample number, and/or the fact this set-up represents the 1th generation of a 48-plex miniaturized recombinant antibody array set-up for serum protein profiling. We will continue the work to develop and establish the next generation of high-performing miniaturized planar recombinant array platform for crude serum protein profiling, similar to what we have already achieved for conventional recombinant antibody microarrays [34, 35].

In an attempt to refine the classification based on disease activity, the SLE nephritis samples were divided into three sub-groups based on SLEDAI-2K denoted: low (3-10) < mid (12-17) < high (18-32), and the serum protein expression profiles were again compared to those of the controls (Fig. 5). Again, the results showed that the classification, as might be expected, improved with increasing disease activity (ROC AUC of 0.47 < 0.72 < 0.84) (Figs. 5A and 5C). Again, C1q, IL-6, and LDL were found to be associated with disease activity.

Finally, we validated the observed expression profile for one of the de-regulated analytes, C1q, using an orthogonal method (Fig. 6). The results showed that down-regulated levels of C1q were validate, targeting the SLE nephritis vs. control comparison, again supporting the observed serum protein expression profiles.

CONCLUSIONS

Taken together, we have designed the first 48-plex miniaturized planar recombinant antibody microarrays for protein expression profiling of crude serum samples. By exploring a novel 48-plex DPN pen design, we were able to in a unique manner simultaneously print 48 different reagents as individual 78.5 um² (\emptyset 10µm) sized spots at a density of 38,000 spots/cm². The miniaturized array set-up was interfaced with a high-resolution fluorescent-based scanner. Compared to conventional antibody microarrays, the current set-up provided several key advantages, including 225 times reduced spot features, ≥19 times increased spot density (multiplexity), 75 times less consumption of antibodies, and the ability to target ≥19 times more analytes, while consuming similar, or smaller, amount of sample. The performance and applicability of this novel technology platform was then demonstrated in a first clinical application targeting SLE nephritis. The results showed that disease-associated serum biomarkers could be deciphered, paving the way for large-scale multiplex profiling of crude serum in a miniaturized fashion applicable to many indications.

ACKNOWLEDGEMENTS

This study was supported by grants from the Alfred Österlund Foundation, Greta and Johan Kock Foundation, Foundation of Strategic Research (Strategic Center for Translational Cancer Research - CREATE Health (www.createhealth.lth.se), and the Swedish Research council (VR-M and VR-NT).

REFERENCES

- 1. R. Ekins, F. Chu and E. Biggart, Ann. Biol. Clin., 1990, 48, 655-666.
- 2. J.D. Hoheisel, M.S.S. Alhamdani and C. Schröder, Proteomics Clin. Appl., 2013, 7, 8-15.
- 3. C.A.K. Borrebaeck and C. Wingren, J. Proteomics, 2009, 72, 928-935.
- 4. S.F. Kingsmore, Nat. Rev. Drug Discov., 2006, 5, 310-321.
- 5. M. Sanchez-Carbayo, *Tumor Biol.*, 2010, **31**, 103-112.
- 6. M. Sanchez-Carbayo, N.D. Socci, J.J. Lozano, B.B. Haab and C. Cordon-Cardo, *Am. J. Pathol.*, 2006, **168**, 93-103.
- C. Schröder, A. Jacob, S. Tonack, T.P. Radon, M. Sill, M. Zucknick, S. Rüffer, E. Costello, J.P. Neoptolemos, T. Crnogorac-Jurcevic, A. Bauer, K. Fellenberg and J.D. Hoheisel, *Mol. Cell. Proteomics*, 2010, 9, 1271-1280.

- C. Wingren, A. Sandström, R. Segersvärd, A. Carlsson, R. Andersson, M. Löhr and C.A.K. Borrebaeck, *Cancer Res.*, 2012, 72, 2481-2490.
- 9. A. Carlsson, D.M. Wuttge, J. Ingvarsson, A.A. Bengtsson, G. Sturfelt, C.A.K. Borrebaeck and C. Wingren, *Mol. Cell. Proteomics*, 2011.
- P.M. Harrison, A. Kumar, N. Lang, M. Snyder and M. Gerstein, *Nucleic Acids Res.*, 2002, 30, 1083-1090.
- J.H. Lim, D.S. Ginger, K.B. Lee, J. Heo, J.M. Nam and C.A. Mirkin, *Angew. Chem. Int. Ed. Engl.*, 2003, **42**, 2309-2312.
- S. Nettikadan, K. Radke, J. Johnson, J. Xu, M. Lynch, C. Mosher and E. Henderson, *Mol. Cell. Proteomics*, 2006, 5, 895-901.
- 13. K. Lee, S. Lee, H. Yu and S.H. Kang, J. Nanosci. Nanotechnol., 2010, 10, 3228-3231.
- J.D. Hoff, L.J. Cheng, E. Meyhofer, L.J. Guo and A.J. Hunt, *Nano letters*, 2004, 4, 853-857.
- N. Backmann, C. Zahnd, F. Huber, A. Bietsch, A. Pluckthun, H.P. Lang, H.J. Guntherodt, M. Hegner and C. Gerber, *Proc. Natl. Acad. Sci. U.S.A.*, 2005, **102**, 14587-14592.
- Y. Arntz, J.D. Seelig, H.P. Lang, J. Zhang, P. Hunziker, J.P. Ramseyer, E. Meyer, M. Hegner and C. Gerber, *Nanotechnol.*, 2003, 14, 86-90.
- G. Zheng, F. Patolsky, Y. Cui, W.U. Wang and C.M. Lieber, *Nat. Biotech.*, 2005, 23, 1294-1301.
- P. Ellmark, S. Ghatnekar-Nilsson, A. Meister, H. Heinzelmann, L. Montelius, C. Wingren and C.A.K. Borrebaeck, *Proteomics*, 2009, 9, 5406-5413.
- 19. S. Ghatnekar-Nilsson, L. Dexlin, C. Wingren, L. Montelius and C.A.K. Borrebaeck, *Proteomics*, 2007, **7**, 540-547.
- 20. A. Bruckbauer, D. Zhou, D.-J. Kang, Y.E. Korchev, C. Abell and D. Klenerman, *J. Am. Chem. Soc.*, 2004, **126**, 6508-6509.
- P.L. Tran, J.R. Gamboa, D.J. You and J.Y. Yoon, *Anal. Bioanal. Chem.*, 2010, **398**, 759-768.
- 22. C. Wingren and C.A.K. Borrebaeck, Drug Discov. Today, 2007, 12, 813-819.
- 23. R.P. Ekins, Clin. Chem., 1998, 44, 2015-2030.
- 24. K.-B. Lee, S.-J. Park, C.A. Mirkin, J.C. Smith and M. Mrksich, *Science*, 2002, **295**, 1702-1705.
- N. Berthet-Duroure, T. Leïchlé, J.-B. Pourciel, C. Martin, J. Bausells, E. Lora-Tamayo,
 F. Perez-Murano, J. François, E. Trévisiol and L. Nicu, *Biomed. Microdev.*, 2008, 10, 479-487.

- A. Meister, M. Liley, J. Brugger, R. Pugin and H. Heinzelmann, *Appl. Phys. Letters*, 2004, 85, 6260–6262.
- 27. J.W. Jang, A. Smetana and P. Stiles, *Scanning*, 2010, **32**, 24-29.
- E.J. Irvine, A. Hernandez-Santana, K. Faulds and D. Graham, *Analyst*, 2011, **136**, 2925-2930.
- 29. L. Petersson, M. Coen, N. Amro, L. Truedsson, C.A.K. Borrebaeck and C. Wingren, *Bioanalysis*, 2014, In press.
- 30. C.C. Mok, Int. J. Women's Health, 2012, 4, 213 222.
- 31. B.H. Rovin, D.J. Birmingham, H.N. Nagaraja, C.Y. Yu and L.A. Hebert, *Bulletin of the NYU hospital for joint diseases*, 2007, **65**, 187-193.
- 32. C.A.K. Borrebaeck and C. Wingren, Exp. Rev. Proteomics, 2009, 6, 11-13.
- 33. D.D. Gladman, D. Ibañez and M.B. Urowitz, J. Rheumatol., 2002, 29, 288-291.
- 34. J. Ingvarsson, A. Larsson, A.G. Sjöholm, L. Truedsson, B. Jansson, C.A.K. Borrebaeck and C. Wingren, *J. Proteome Res.*, 2007, **6**, 3527-3536.
- C. Wingren, J. Ingvarsson, L. Dexlin, D. Szul and C.A.K. Borrebaeck, *Proteomics*, 2007, 7, 3055-3065.
- J. Ingvarsson, C. Wingren, A. Carlsson, P. Ellmark, B. Wahren, G. Engström, U. Harmenberg, M. Krogh, C. Peterson and C.A.K. Borrebaeck, *Proteomics*, 2008, 8, 2211-2219.
- 37. N.L. Anderson and N.G. Anderson, Mol. Cell. Proteomics, 2002, 1, 845-867.
- W. Kusnezow, V. Banzon, C. Schroder, R. Schaal, J.D. Hoheisel, S. Ruffer, P. Luft, A. Duschl and Y.V. Syagailo, *Proteomics*, 2007, 7, 1786-1799.
- 39. C. Wingren and C.A.K. Borrebaeck, Curr. Opin. Biotechnol., 2008, 19, 55-61.
- E. Söderlind, L. Strandberg, P. Jirholt, N. Kobayashi, V. Alexeiva, A.-M. Aberg, A. Nilsson, B. Jansson, M. Ohlin, C. Wingren, L. Danielsson, R. Carlsson and C.A.K. Borrebaeck, *Nat. Biotech.*, 2000, 18, 852-856.
- A. Carlsson, C. Wingren, M. Kristensson, C. Rose, M. Ferno, H. Olsson, H. Jernstrom, S. Ek, E. Gustavsson, C. Ingvar, M. Ohlsson, C. Peterson and C.A. Borrebaeck, *Proc. Natl. Acad. Sci. U.S.A.*, 2011, **108**, 14252-14257.
- 42. C.-C. Chang and C.-J. Lin, ACM Trans. Intell. Syst. Technol., 2011, 2, 1-27.
- 43. N. Cristianini and J. Shawe-Taylor, Cambridge University Press, 2000.
- 44. R. Ihaka and R. Gentleman, J. Comp. Graph. Stat., 1996, 5, 299-314.
- D. Hamelinck, H. Zhou, L. Li, C. Verweij, D. Dillon, Z. Feng, J. Costa and B.B. Haab, Mol. Cell. Proteomics, 2005, 4, 773-784.

- 46. A. Sandstrom, R. Andersson, R. Segersvard, M. Lohr, C.A.K. Borrebaeck and C. Wingren, *Proteomics Clin. Appl.*, 2012, **6**, 486-496.
- 47. M.S. Alhamdani, C. Schroder and J.D. Hoheisel, Proteomics, 2010, 10, 3203-3207.
- 48. R. Orchekowski, D. Hamelinck, L. Li, E. Gliwa, M. vanBrocklin, J.A. Marrero, G.F. Vande Woude, Z. Feng, R. Brand and B.B. Haab, *Cancer Res.*, 2005, **65**, 11193-11202.
- 49. C.C. Mok, J. Biomed. Biotechnol., 2010, 2010.
- 50. B.H. Rovin and X. Zhang, Clin. J. Am. Soc. Nephrol. 2009, 4, 1858-1865.
- J.W. Bauer, E.C. Baechler, M. Petri, F.M. Batliwalla, D. Crawford, W.A. Ortmann, K.J. Espe, W. Li, D.D. Patel, P.K. Gregersen and T.W. Behrens, *PLoS medicine*, 2006, 3, e491.
- J.W. Bauer, M. Petri, F.M. Batliwalla, T. Koeuth, J. Wilson, C. Slattery, A. Panoskaltsis-Mortari, P.K. Gregersen, T.W. Behrens and E.C. Baechler, *Arthritis Rheum.*, 2009, 60, 3098-3107.
- 53. C.C. Liu and J.M. Ahearn, *Best practice & research*, 2009, **23**, 507-523.
- 54. C.H. Suh and H.A. Kim, *Exp. Rew. Mol. Diagnos.*, 2008, **8**, 189-198.
- 55. G. Sturfelt, U. Johnson and A.G. Sjoholm, Scand. J. Rheumatol., 1985, 14, 184-196.
- 56. G. Sturfelt and A.G. Sjoholm, Int. Arch. Allergy Appl. Immunol., 1984, 75, 75-83.
- 57. D.P. D'Cruz, M.A. Khamashta and G.R. Hughes, *Lancet*, 2007, **369**, 587-596.
- 58. A. Rahman and D.A. Isenberg, N. Eng. J. Medicine, 2008, 358, 929-939.

LEGENDS

Figure 1

Schematic overview of the miniaturized 48-plex recombinant antibody array platform.

Figure 2

Design and printing of 48-plex miniaturized recombinant antibody arrays. Close-ups of the 48plex pen, inkwells, printing image (drops) are shown. A representative fluorescently scanned image after the array was exposed to a sample and its corresponding signal intensities are shown. The limit of detection (LOD) was defined as the analyte concentration corresponding to a signal 2x standard deviations above the negative control. An antibody had to display detectible signal intensities in \geq 30% of all analyzed samples (n=74), in order to be included in the data set. The cut-off is shown as a solid line, and the 31 antibodies in bold was used in all further analysis.

Figure 3

Evaluation of array data normalization method. (A) Raw data, (B) The anti-CT normalization method, (C) The semi-global normalization method, (D) Subtract by geometric group mean normalization, (E) Divide by geometric group mean normalization. The serum samples (n=74 plus 7 duplicates, 1 sample excluded due to high background signals) were randomized and analyzed on 7 slides over 4 days. The day of analysis and number of slide was mapped onto the data set, and clustered using Principle component analysis (PCA) and the number of any differentially expressed antibodies were determined using the ANNOVA test (p<0.05). The intensity profile over all samples per antibody was plotted and, a representative plot are shown for the IL-1 α antibody.

Figure 4

Technical reproducibility of the miniaturized array set-up, expressed as CV-values. (A) Spotto-spot variability for all antibodies (n=31), deposited in 8 replicates per array for all 74 samples. (B) Spot-to-spot variability per antibody. (C) Chip-to-chip reproducibility (subtract by geometric group mean normalization). CV-values were calculated for all antibodies (n=25) that gave a detectable signal on 4 of 5 slides (subarrays) for a single serum samples repeatedly analysed. (D) Chip-to-chip reproducibility (divide by geometric group mean normalization). CV-values were calculated for all antibodies (n=25) that gave a detectable signal on 4 of 5 slides (subarrays) for a single serum samples repeatedly analysed.

Figure 5

Protein expression profiling of biotinylated SLE nephritis serum samples using the miniaturized 48-plex antibody array platform. The array data was normalized using the divide by geometric group mean strategy. The SLE nephritis patients was grouped into 2 or 3 groups based on disease activity (SLEDAI-2K); low (3-14) vs. high (16-32) disease activity, or low (3-10) vs. mid (12-17) vs. high (18-32) disease activity. The classification of SLE nephritis patients vs. healthy controls was performed using a leave one out cross validation test, and presented as ROC curves and AUC values. Protein expression profiles, covering 22 unique serum proteins

present at detectable levels (targeted by 31 antibodies) were compared, and any differentially expressed proteins (p<0.05), i.e. potential biomarkers, were identified using Wilcoxon signed ranked test. (A) Summary of the protein expression profiling. (B) ROC curves and AUC values for SLE vs. controls, SLE (16-32 vs. controls), and SLE (18-32) vs. controls. (C) Array signal intensities for the differentially expressed serum analytes in SLE vs. controls, SLE (16-32 vs. controls), and SLE (18-32) vs. controls, SLE (16-32 vs. controls), and SLE (18-32) vs. controls.

Figure 6

Validation of the differentially expressed protein profile for the serum complement protein C1q. The protein expression profile observed using the miniaturized arrays was compared with those determined using an orthogonal method (electroimmunoassay).

Graphical Abstract text

Miniaturization of multiplexed recombinant antibody microarrays for protein expression profiling of crude proteomes.

Table 1.

Patient demographics and clinical parameters of SLE nephritis patients and healthy controls included in the study.

Parameter	SLE	Normals
No. of patients	45	30
Gender (female/male)	40/5	28/2
Age mean (range)	38 (18-70)	43 (19-68)
SLEDAI-2K mean (range)	15 (3-32)	-

Table 2

The different scFv specificities used for the miniaturized recombinant antibody arrays.

Antigens (no. of clones)		
IL-1α (2)		
IL-2		
IL-3		
IL-4 (2)		
IL-6 (2)		
IL-8 (2)		
IL-10 (2)		
IL-11 (2)		
IL-12 (2)		
IL-13 (2)		
VEGF		
TGF-β1		
TNF-α (2)		
IL-18 (2)		
MCP-4		
INF-γ		
RANTES		
LDL (2)		
Lewis x		
Procathepsin W		
GLP-1		
C1q		
C3		
C4		
C5		
IL-16		
CD40		
MCP-1		
INF-a		
Cholera toxin subunit B		



Figure 2.

Chip Accepted Manuscript

σ

ЧО

Lab







Figure 5 Page 27 of 28

Lab on a Chip

Α.

	Significant clones (p<0.05)				
Comparison	AUC	C1q	LDL (1)	IL-6 (1)	IL-6- (2)
SLE vs. Normal	0.69	J.	$\mathbf{\Lambda}$	$\mathbf{\Lambda}$	
SEE VS. Norman	0.05	v	I	I	
SLE (3-14*) vs. Normal	0.60	\downarrow	\uparrow		
SLE (16-32)* vs. Normal	0.80	\downarrow	\uparrow	\uparrow	\uparrow
SLE (3-10)* vs. Normal	0.47	\downarrow	\uparrow		
SLE (12-17)* vs. Normal	0.72	\downarrow	\uparrow	\uparrow	\uparrow
SLE (18-32)* vs. Normal	0.84	\downarrow	\uparrow	\uparrow	
		•			

*= SLEDAI-2K



