Chapter 21

Protein Function Microarrays for Customised Systems-Oriented Proteome Analysis

Jonathan M. Blackburn and Aubrey Shoko

Abstract

Protein microarrays have many potential applications in the systematic, quantitative analysis of protein function. However, simple, reproducible, and robust methods for array fabrication that are compatible with the study of large, custom collections of potentially unrelated proteins are required. Here, we discuss different routes to array fabrication and describe in detail one approach in which the purification and immobilisation procedures are combined into a single step, significantly simplifying the array fabrication process. We illustrate this approach by reference to the creation of an array of human protein kinases and discuss methods for assay and data analysis on such arrays.

Key words: Protein microarray, Biotinylation, Proteomics, Functional analysis, Surface capture, Protein kinase, Biotin carboxyl carrier protein, Inhibitor specificity

1. Introduction

In the post-genomic era, attention has turned towards the systematic assignment of function to proteins encoded by genomes. Bioinformatics methods are typically now used ubiquitously as an essential first step in assigning predicted function to open reading frames (1). However, while such methods give helpful insights into possible function, there remain many examples of proteins that have closely related sequences and/or structures but which prove to have quite different functions when studied experimentally (2–4). As the number of sequenced genomes expands ever further, there is thus an ever increasing need for experimental methods that enable the determination and/or verification of protein function in high throughput. At the forefront of this monumental task, the field of proteomics can be segregated into discovery- and systems-oriented proteomics (5). Discovery-oriented proteomics

Ulrike Korf (ed.), *Protein Microarrays: Methods and Protocols*, Methods in Molecular Biology, vol. 785,
DOI 10.1007/978-1-61779-286-1_21,

is mainly concerned with documenting the abundance and localisation of individual proteins as well as building a picture of protein–protein interaction networks. This is the realm of 2-hybrid screens, 2D-gel electrophoresis and increasingly powerful more direct, isotope-labelling-based mass spectrometry methods; these latter two methods in particular are commonly used to understand the way in which expression profiles change in response to different stimuli by comparing, for example, diseased and healthy cell extracts. However, these discovery-oriented proteomics methods tell us little directly about the precise function of individual proteins or protein complexes, even when augmented by ever more sophisticated bioinformatic methods. Systems-oriented proteomics takes a different approach; rather than rediscovering each protein in each new experiment, the focus is on a pre-defined set of proteins – in principle up to an entire proteome, but in practice more typically a limited subset thereof – enabling the functionality of each member of that set to be dissected in great detail (6). However, obtaining quantitative and genuinely comparative functional data across large sets of proteins with any degree of accuracy is technically difficult, requiring isolation of each individual protein in an assayable format. We and others have chosen to focus on protein function microarray-based methods because the parallel, high-throughput nature of microarray experiments is attractive for analysing large numbers of protein interactions, while the uniform intra-array conditions both simplify and increase the accuracy of assays (6–13). Additionally, the small volumes of ligand or reaction solution required to perform assays, typically tens to hundreds of microlitres, can provide economic advantages, for example when using expensive recombinant proteins or labelled compounds. The key element to such microarray experiments is that the arrayed, immobilised proteins retain their folded structure such that meaningful functional interrogation can then be carried out. There are a number of approaches to this problem which differ fundamentally according to whether the proteins are immobilised through non-specific, poorly defined interactions or through a specific, set of known interactions. The former approach is attractive in its simplicity and is compatible with purified proteins derived from native or recombinant sources (14, 15) but suffers from a number of risks. Most notable among these is that the uncontrolled nature of the interactions between each protein and the surface might at best give rise to a heterogeneous population of proteins or at worst destroy activity altogether due to partial or complete surface-mediated unfolding of the immobilised protein. In practice, an intermediate situation probably most often occurs, where a fraction of the immobilised proteins either have undergone conformational change as a result of the non-specific interactions or have their binding/active sites occluded by surface attachment; these effects effectively reduce the specific activity of the immobilised protein,

and therefore decrease the signal-to-noise ratio in any subsequent functional assay that is sensitive to conformation. It is, therefore, important to consider the possible effects of unfolding on the intended downstream assay prior to choosing an array surface: for example, an assay in which solution-phase antibodies bind to linear epitopes on the array will be unlikely to be affected by unfolding of the arrayed proteins (indeed, it may even be desirable to deliberately unfold such proteins in order to expose a greater range of potential epitopes); by comparison, an assay in which a solution phase kinase phosphorylates arrayed proteins may well be sensitive to disruption of the relative 3-dimensional arrangement of targeting and substrate domains in the arrayed proteins.

The advantages of controlling the precise mode of surface attachment are that, providing the chosen point of attachment does not directly interfere with activity, the immobilised proteins will have a homogeneous orientation resulting in a higher specific activity and higher signal-to-noise ratio in assays, with less interference from non-specific interactions (16). This may be of particular advantage when studying protein–small-molecule interactions or conformationally sensitive protein–protein interactions in an array format. The disadvantages of this approach though are that it is really only compatible with recombinant proteins or with families of proteins, such as antibodies, which have a common structural element through which they can be immobilised. However, in a systems-oriented approach the disadvantage of working with recombinant proteins is largely outweighed by the problems encountered in individually purifying large numbers of active proteins from native sources. In addition, experimental approaches that facilitate high-throughput expression and purification of many different proteins in parallel have become more generally accessible over recent years, simplifying access to larger, defined collections of recombinant proteins. An important caveat here though is that it is increasingly clear that despite its ease of use, *Escherichia coli* is not an optimal host for recombinant expression of folded, functional mammalian proteins. Furthermore, while cell-free transcription/translation-based protein microarray systems have been described (17, 18), it remains unclear how reproducible such arrays are, or what proportion of mammalian proteins produced by such approaches are properly folded and therefore functional prior to immobilisation.

In this chapter, we therefore describe the high-throughput cloning, expression, purification, array fabrication, and assay of a set of recombinant proteins in which the human proteins are expressed in insect cells and in which the mode of surface attachment is tightly controlled through the use of an appropriate affinity tag. Furthermore, we show how laborious pre-purification of the recombinant proteins prior to array fabrication can be avoided through the use of a suitable array surface combined with a suitable

affinity tag, thus greatly simplifying array fabrication (7). We illustrate this approach to array fabrication with respect to a set of human protein kinases expressed in insect cells as fusions to a polypeptide tag which becomes biotinylated in vivo (7) (see Note 1). In addition, we show representative data from a number of different assays carried out on protein function arrays made in this way.

2. Materials

1. *E. coli Acc* B gene.
2. Human MAPK1 cDNA.
3. *Autographa californica* baculovirus "bacmid" vector pBAC10:KO_{1629}.
4. *E. coli* strain HS996.
5. *Spodoptera frugiperda* SF21 cells.
6. Lipofectin (Invitrogen).
7. InsectXpress media (Lonza).
8. 6-well cell culture plates (Nunc).
9. Foetal bovine serum (FBS; Sigma).
10. 24-well deep well blocks (Nunc).
11. Biotin (Sigma).
12. Phosphate-buffered saline (PBS): 1.5 mM KH_2PO_4, 4.3 mM Na_2HPO_4, 137 mM NaCl, 3 mM KCl, adjust to pH 7.4 with HCl, prepare as 10× stock, and autoclave before storage at room temperature.
13. PBST: Prepare from PBS by adding 0.1% (v/v) Tween20.
14. Freezing buffer: 25 mM HEPES, 50 mM KCl, pH 7.5.
15. Lysis buffer: 25 mM HEPES pH 7.5, 20% glycerol, 50 mM KCl, 0.1% Triton X-100, 0.1% bovine serum albumin (BSA), Protease inhibitor cocktail, 1 mM DTT.
16. Streptavidin–HRP conjugate (Sigma).
17. Anti-cMyc antibody (Sigma).
18. Streptavidin (Sigma).
19. 384-well V-bottomed plate (X7022; GENETIX).
20. Nexterion Slide P (Schott).
21. Lifterslips (Nunc, USA).
22. Wash buffer: 20 mM KH_2PO_4, 0.2 mM EDTA, 5% glycerol, pH 7.4.
23. QArray II microarray robot equipped with 16 × 300 µm solid stainless steel pins (Genetix).

24. Kinase buffer: 25 mM Tris–HCl, 5 mM beta-glycerophosphate, 2 mM dithiothreitol, 0.1 mM Na_3VO_4, 10 mM $MgCl_2$, pH 7.5.
25. Adenosine triphosphate (ATP; Sigma).
26. DNA microarray scanner (Tecan LS Reloaded).
27. Ethanolamine (Sigma).
28. 150 mM Na_2HPO_4 buffer, pH 8.5: Prepared by titrating 0.2 M NaH_2PO_4 into 0.2 M Na_2HPO_4 to pH 8.5 and diluting to a final concentration of 150 mM.
29. Milk/PBST: 5% (w/v) non-fat dry milk in PBST.

3. Methods

The methods described below outline (1) the construction of a representative transfer vector in *E. coli*, (2) co-transfection of insect cells with this transfer vector and bacmid, (3) propagation of recombinant baculovirus, (4) induction of protein expression, (5) the extraction of the protein from insect cells, (6) the printing of a protein microarray, and (7) the assay of the protein microarray for protein kinase activity.

3.1. Construction of the Transfer Vector for Full Length Human MAPK1

The general baculoviral system used here is adapted from the work of Prof Ian Jones (Reading University, UK; (19)). The specific *E. coli* transfer vector system used is derived from pTriEx1.1 (Novagen).

3.1.1. Transfer Vector

3.1.2. Amplification and Cloning of the MAPK1 Gene as an N-terminal Fusion to BCCP

All DNA manipulations were carried out using standard recombinant DNA methods (20) to construct the transfer vector and are accordingly not described here in detail.

The gene encoding the *E. coli* biotin carboxyl carrier protein (BCCP) domain (amino acids 74–156 of the *E. coli accB* gene; Fig. 1) (21, 22) was amplified by PCR from an *E. coli* genomic DNA preparation and was cloned downstream of a viral polyhedrin promoter in an *E. coli* vector to create the transfer vector pJB1 (Fig. 2). Flanking this *polh*-BCCP expression cassette were the baculoviral 603 gene and the 1629 genes (19) to enable subsequent homologous recombination of the construct into a replication-deficient baculoviral genome (Fig. 2).

The full length, *MAPK1* gene was amplified by PCR from a cDNA clone (obtained from the Mammalian Gene Collection) and cloned into the pJB1 transfer vector upstream of and in-frame with the BCCP tag using ligation-independent cloning methods (24), replacing the ORF region between the *Spe* I and *Nco* I sites of pJB1 in the process, to generate pJB1-MAPK1. In the course of the PCR amplification step, the stop codon of the MAPK1 gene was

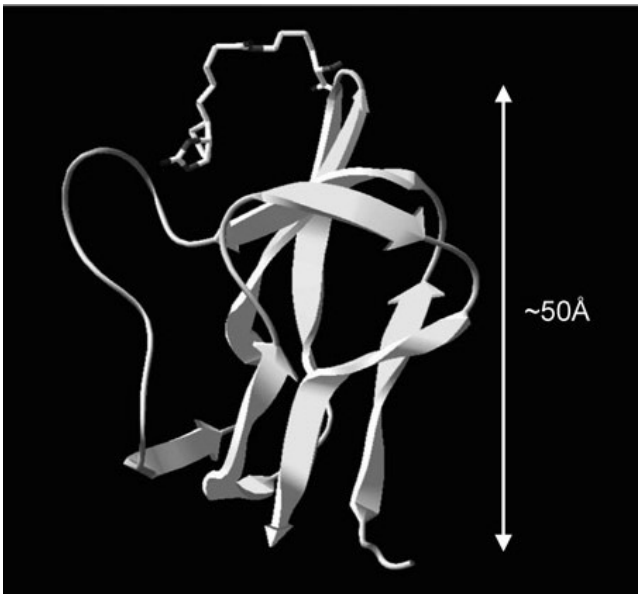

Fig. 1. Structure of the *E. coli* biotin carboxyl carrier protein (BCCP) domain. Residues 77–156 are drawn (coordinate file 1bdo), showing the N- and C-terminii and the single biotin moiety that is attached to lysine 122 in vivo by biotin ligase. Representation produced using SwissPDBViewer (23).

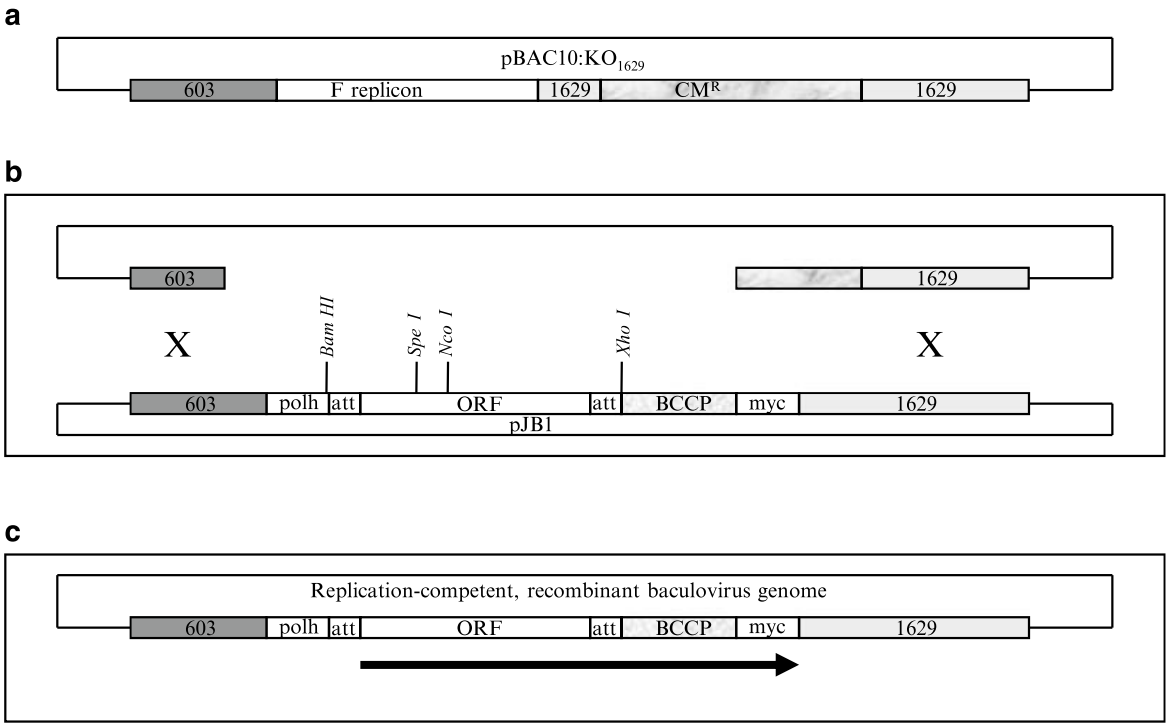

Fig. 2. Schematic of the baculoviral recombination system used in this work. (**a**) Disruption of the essential 1629 gene renders the baculoviral genome replication-deficient. (**b**) Linearised baculoviral genomic DNA and transfer vector are co-transfected into insect cells. (**c**) Homologous recombination regenerates an intact 1629 gene, enabling viral replication. Expression of the BCCP fusion protein is driven from the polh promoter.

removed such that the resulting construct encoded an in-frame MAPK1-BCCP fusion protein; the *MAPK1* gene was sequence-verified against the RefSeq database.

3.2. Insertion of the MAPK1-BCCP Expression Cassette into a Baculoviral Genome

1. Bacmid pBAC10:KO_{1629} (19) was propagated in *E. coli* HS996 cells and bacmid DNA prepared according to standard procedures. pBAC10:KO_{1629} was then linearised by restriction with *Bsu*361 for 5 h at 37°C, after which *Bsu*361 was heat killed at 80°C for 15 min.
2. Linearised pBAC10:KO_{1629} was combined with undigested pJB1-MAPK1 and used to transfect SF21 cells according to standard protocols.
3. 500 ng of linearised pBAC10:KO_{1629} was combined with 500 ng undigested pJB30-MAPK1 and the total volume made up to 12 µl with water.
4. 12 µl lipofectin (diluted 2:1 in H_2O) was then added to this DNA mix and the tube incubated at room temperature for 30 min.
5. 1 ml serum free media (InsectXpress) was added to the lipofectin/DNA mixture.
6. A 6-well plate containing 1×10^6 SF21 cells/well was prepared and incubated at 27°C for 1 h to allow the cells to adhere.
7. Excess media was aspirated from the SF21 cells and replaced with the lipofectin/DNA/serum-free mix.
8. The transfected cells were incubated at 27°C overnight.
   The media was replaced with 2 ml InsectXpress media supplemented with 2% FBS and incubated at 27°C without agitation for a further 72 h.
9. Cells were resuspended by physical agitation and then pelleted by centrifugation at $1{,}000 \times g$ for 10 min.
10. The supernatant containing recombinant baculovirus was transferred to a fresh tube and stored at 4°C; this is the P_0 stock.

3.3. Amplification of Recombinant Baculovirus

Recombinant baculoviral particles were amplified according to standard procedures. Briefly:

1. A 6-well plate was set up with 1×10^6 SF21 cells/well and incubated at 27°C for 1 h.
2. Excess media was removed and replaced with 500 µl of P_0 virus plus 500 µl InsectXpress media supplemented with 2% FBS and incubated at 27°C without agitation for a further 72 h.
3. P_1 virus was harvested as described above.
4. A 150 ml tissue culture flask was seeded with 20 ml of 1×10^6 SF21 cells/ml and incubated at 27°C for 1 h.

5. Excess media was removed and replaced with 500 µl of P_0 virus plus 3 ml InsectXpress media supplemented with 2% FBS and incubated at 27°C for 1 h, after which a further 25 ml InsectXpress media supplemented with 2% FBS were added and cells incubated without agitation for 72 h.
6. P_2 virus was harvested as described above.
7. The titre of the P_2 viral stock was determined by a SybrGreen-based quantitative PCR assay vs. a stock of known titre determined by plaque assay. The titre of the P_2 stock should be *ca.* 10^7 pfu/ml.

3.4. Protein Expression and Extraction

1. Set up a 24-well deep well plate containing 6×10^6 SF21 cells/well suspended in 3 ml InsectXpress media supplemented with 2% FBS and 50 µM biotin (see Note 3).
2. Add 200 µl of P_2 virus and incubate at 27°C for 72 h with agitation.
3. Harvest cells by centrifugation of the 24-well deep well plate prior to lysis.
4. Gently resuspend the cells in 3 ml of PBS buffer, recentrifuge the plate and discard the supernatant; repeat three times in total.
5. Gently resuspend the pellets in 350 µl of freezing buffer ensuring thorough mixing of the cells.
6. Aliquot the cells in 50 µl volumes and store at –80°C until required for cell lysis.
7. Thaw a 50 µl aliquot and add 50 µl lysis buffer plus 10 U Benzonase (Pierce) and shake on ice for 30 min.
8. Remove cell debris by centrifugation, collect the supernatant, and store on ice for up to 24 h before printing.
9. Determine the protein concentration of the soluble, crude protein extract by Bradford assay (25) to confirm that effective cell lysis has occurred (see Note 4).
10. Determine the approximate expression level of soluble BCCP fusion by SDS-PAGE together with Western blot analysis (20) using a streptavidin–HRP conjugate (Fig. 3; see Note 5).
11. To determine the extent of biotinylation of the BCCP fusion protein, carry out a supershift Western blot assay (with an anti-c-Myc antibody) in which equivalent crude lysate samples are pre-incubated with or without streptavidin (0.1 g/ml) (Fig. 4; see Note 6).

3.5. Multiplexed Cloning and Expression of Mammalian Proteins

The procedures described above (Subheadings 3.1–3.4) can clearly be applied to any mammalian cDNA. Moreover, we and others have found that the ligation-independent cloning methods, as well as insect cell transfection, baculovirus amplification, and protein

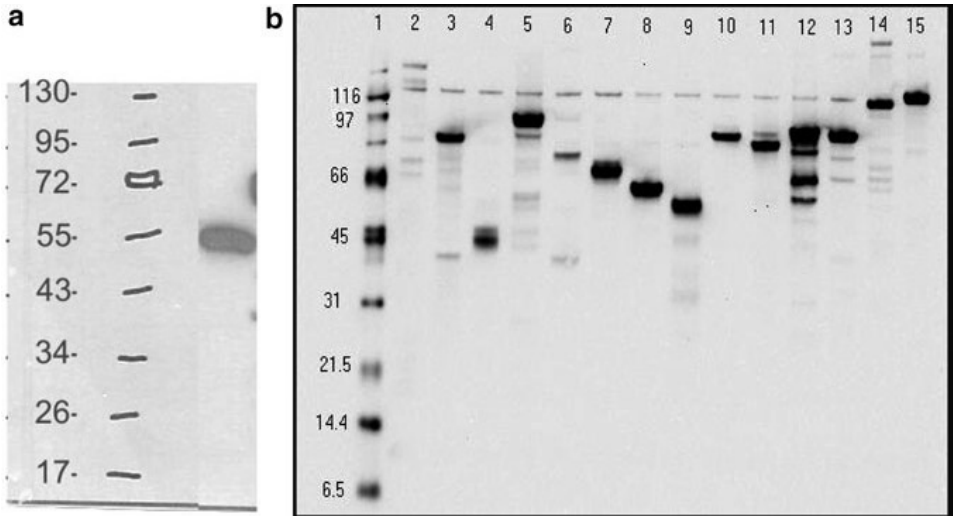

Fig. 3. Western blots of BCCP fusion proteins expressed in SF21 cells. (**a**) MAPK1-BCCP, encoded by pJB1-MAPK1. (**b**) (1) Biotinylated marker, (2) ZNF198-BCCP, (3) FUS-BCCP, (4) SDHB-BCCP, (5) STAT4-BCCP, (6) FH-BCCP, (7) MUTYH-BCCP, (8) PNUTL1-BCCP, (9) GATA1-BCCP, (10) NF2-BCCP, (11) FACL6-BCCP, (12) MSF-BCCP, (13) MSN-BCCP, (14) RAB5EP-BCCP, (15) Control. All westerns were of crude insect cell lysates and were developed using a streptavidin–hrp conjugate. Good expression levels can be observed across a range of unrelated proteins, some in excess of 100 kDa.

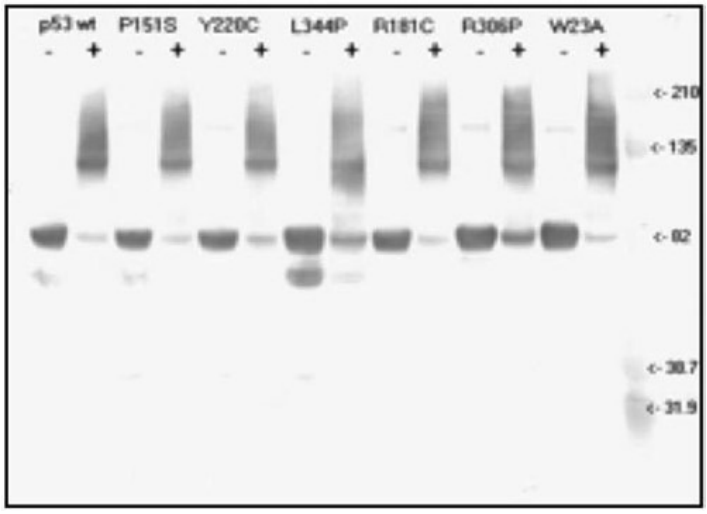

Fig. 4. Supershift assay to determine extent of biotinylation. In a supershift assay, each sample is pre-incubated either with ("+") or without ("−") streptavidin prior to separation by SDS-PAGE. The extent of biotinylation can be estimated by comparison of "+" and "−" samples. Here, we show examples using wild-type and variant p53-BCCP fusion proteins.

expression and extraction are all amenable to multiplexing through the use of appropriate multi-well plate formats. We have observed that by the use of the approach described here, we can easily achieve a ~70% success rate from starting cDNA through to expressed, folded, biotinylated protein suitable for array fabrication. Interestingly, the ~30% overall drop-out rate we observe lies almost entirely in the cloning steps and we have observed a >95% success

rate in progressing from sequence-verified transfer vector through to expression of a folded, biotinylated, arrayable human protein in insect cells; this compares favourably with the much lower success rates observed when attempting to express mammalian proteins in *E. coli* (26).

Using these methods, we have thus been able to assemble collections of hundreds of expressed human proteins in a form ready for array fabrication in just a few months and at low cost.

3.6. Fabrication of Protein Microarrays

In the procedures described below, we do not employ a pre-purification step prior to array fabrication but instead rely on a rapid, single-step immobilisation and purification procedure to create arrays of biotinylated BCCP fusion proteins (Fig. 5; see Notes 1 and 2).

3.6.1. Preparation of Source Plates for Printing

1. Transfer 40 µl of crude protein extract for each BCCP-tagged protein to be arrayed into individual wells of a 384-well V-bottom plate and keep at 4°C. This is the source plate for the print runs.
2. Centrifuge the 384-well plate at 4,000 × *g* for 2 min at 4°C to pellet any cell debris that has carried over. Store plate on ice prior to print run.

3.6.2. Preparation of Streptavidin-Coated Slides for Printing

1. Equilibrate a Nexterion Slide P microarray slide to room temperature and remove from the foil package (see Note 2).
2. Make up a 1 mg/ml streptavidin solution in 150 mM Na_2HPO_4 buffer (pH 8.5).
3. Place a glass microarray "lifterslip" over the microarray surface and pipette 60 µl of the streptavidin solution along the edge of the lifterslip such that the solution is drawn under the coverslip uniformly by capillary action (see Note 2).
4. Leave for 1 h at room temperature in a humidified chamber.

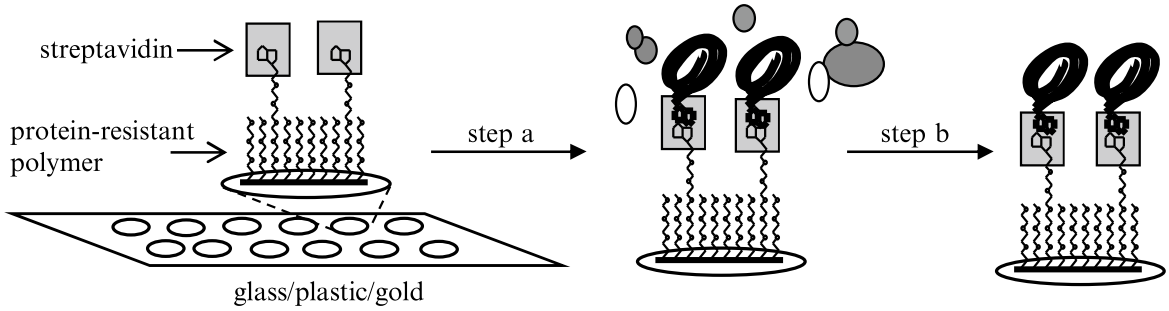

Fig. 5. Schematic of single step immobilisation/purification route to array fabrication. The array surface is intrinsically "non-stick" with respect to proteinaceous material but has a high affinity and specificity for biotinylated proteins. Crude cellular lysates containing the recombinant biotinylated proteins can then be printed onto the surface in a defined array pattern (*step a*) and all non-biotinylated proteins removed by washing (*step b*), leaving the recombinant proteins purified and specifically immobilised via the affinity tag in a single step.

5. Remove the lifterslip and wash the slide for 1 h at RT in 10 ml 150 mM Na_2HPO_4 buffer (pH 8.5) containing 50 mM ethanolamine to deactivate any remaining amine-reactive groups.
6. Wash the slide for 3×5 min in 10 ml wash buffer and then for 5 min in 10 ml water.
7. Place the slide in a 50-ml Falcon tube and centrifuge at 1,000×*g* for 5 min at 20°C to spin dry. Streptavidin-coated slides were placed into slide boxes, sealed in Ziploc bags and stored at –20°C.

3.6.3. Fabrication of Arrays

In general, any microarray printer could be used to print the arrays. The printing procedures can be carried out at room temperature providing the source plate is kept at 4°C and the atmosphere in the print chamber is humidified to ca. 50%. However, preferably the printing device itself should also be cooled. Here, we describe one specific set of parameters which work well on streptavidin-coated glass microarray slides at room temperature using a Genetix QArray II robot equipped with 16×300 µm tipped solid pins (Figs. 6 and 7; see Notes 7–12).

1. Load the 384-well source plate into the QArray II.
2. Load the streptavidin-coated slides onto the print bed of the QArray II.

   Print the arrays using the following key QArray II settings:

   Inking Time (ms)=500

   Microarraying pattern=7×7, 500 µm spacing

   Max stamps per ink=1

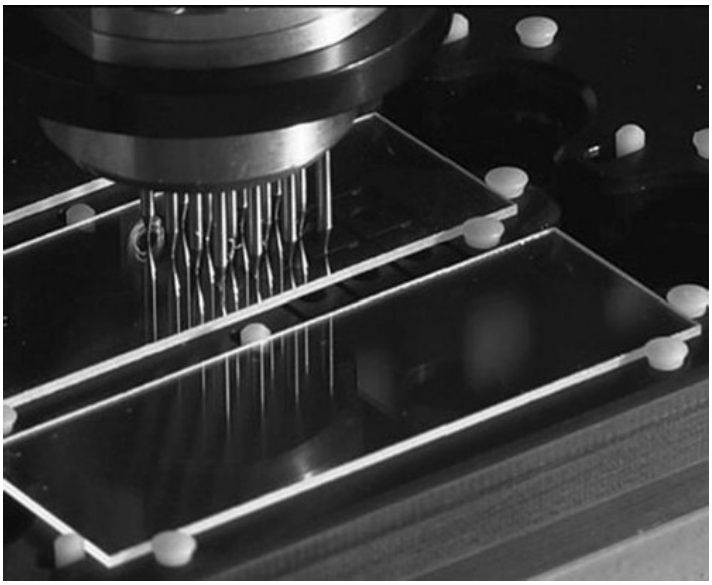

Fig. 6. Printing onto a streptavidin-coated glass microscope slide using 16 solid pins.

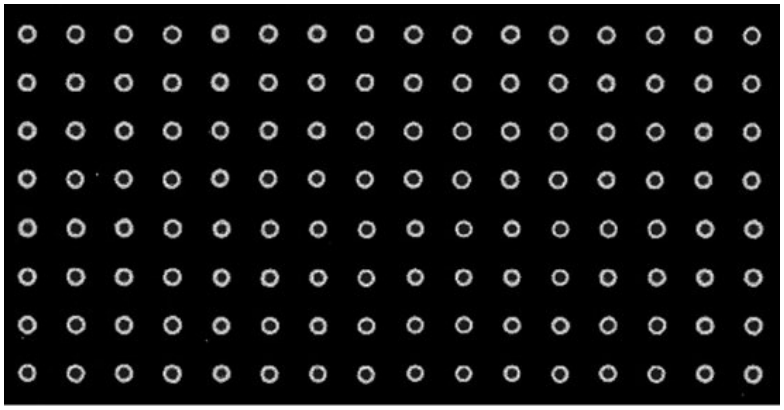

Fig. 7. Highly reproducible arrays with low intra- and inter-array variabilities are key to accurate downstream assays. Here, a single biotinylated protein was printed multiple times using all pins in a 16-pin printing head. The amount of bound protein and spot morphology was determined by imaging the array after binding of a Cy3-labelled anti-His antibody to the immobilised proteins on the array.

No. stamps per spot = 2
Printing depth = 150 μm
Water washes = 60 s wash, 0 s dry
Ethanol wash = 10 s wash, 1 s dry.

3.6.4. Post-printing Processing of Arrays

1. Remove slides from microarraying robot.
2. Wash the slides for 30 min in Falcon tubes containing 50 ml wash buffer supplemented with 100 μM biotin, 0.1 mg/ml BSA (see Note 13).
3. Store in wash buffer at 4°C until use.

3.7. Array-Based Assays for Protein Activity

Once microarrays containing sets of folded, immobilised proteins have been physically fabricated, a wide range of different, systematic, and quantitative assays can in principle be carried out on replica arrays, inter alia: protein–protein interaction assays; protein–nucleic acid interaction assays; protein–small-molecule binding assays; protein–lipid binding assays; and even enzymatic turnover reactions. Through careful experimental design, it is therefore possible now to dissect the properties of diverse collections of unrelated proteins or of specific families of proteins in order to gain a greater understanding of, among others, substrate and inhibitor selectivity. The limiting factor here thus remains the ability to devise an assay that is readable in microarray format for each of the individual arrayed proteins in parallel. Here, we exemplify such assays briefly by reference to arrays of protein kinases.

3.7.1. Verifying that the Individual BCCP-Tagged Proteins Have Indeed Become Immobilised

Following standard protocols for Western blots, it is possible, for example, to probe a replica array with an anti-cMyc antibody followed by a secondary antibody-HRP conjugate, as follows:

1. Dilute a mouse anti-cMyc antibody 1:1,000 in 1 ml milk/PBST.
2. Remove the protein array from wash buffer and equilibrate in PBST at room temperature for 5 min.
3. Drain away the PBST, add 5 ml antibody solution to the array and incubate with gentle agitation at room temperature for 30 min.
4. Wash the array for 3×5 min with 1 ml of PBST.
5. Dilute a goat anti-mouse antibody–hrp conjugate 1:1,000 in 1 ml milk/PBST.
6. Add the antibody solution to the array and incubate with gentle agitation at room temperature for 30 min.
7. Wash the array for 3×5 min with 1 ml of PBST.
8. Add 1 ml chemiluminescent detection reagents (Pierce) to the array.
9. After 1 min, remove slide to a 50-ml Falcon tube and centrifuge for 30 s to dry.
10. In a dark room, place the array against autoradiography film for varying lengths of time before developing the film (Fig. 8; see Note 14).

3.7.2. Measuring the Extent of Autophosphorylation of Each Arrayed Protein Kinase

A subset of protein kinases undergo autophosphorylation at elevated ATP concentrations; for those that do a test to demonstrate whether the arrayed kinases are active or not, based on a simple autophosphorylation assay, can be carried out as follows.

1. Dilute a Cy3-labelled broad-specificity antiphosphotyrosine antibody (Cell Signalling) 1:200 in 2 ml milk/PBST.

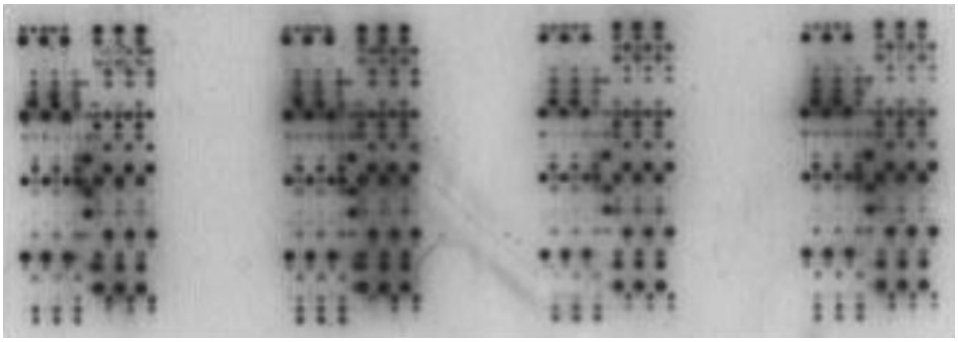

Fig. 8. Array-based qualitative analysis of immobilisation efficiencies. Here, an array of 48 diverse BCCP-tagged human proteins was printed in triplicate, with four replica arrays per slide. Western blot-style analysis, probing with an anti-c-myc antibody, reveals that all proteins are immobilised as expected; apparent gaps on the array are due to the printing pattern used here.

2. Remove two replica protein kinase arrays from wash buffer and equilibrate in PBST at room temperature for 5 min.
3. Drain away the PBST, add 1 ml of kinase buffer to the first array and 1 ml kinase buffer supplemented with 100 μM ATP to the second array. Incubate both arrays at room temperature for 1 h.
4. Wash each array for 3×5 min with 1 ml of PBST containing 0.1% SDS.
5. Drain away the PBST/SDS solution, add 1 ml antibody solution to each array and incubate with gentle agitation at room temperature for 30 min.
6. Wash the arrays for 3×5 min with 1 ml of PBST each.
7. Remove the arrays to a 50-ml Falcon tube and centrifuge for 30 s to dry.
8. Scan the arrays at 550 nm using a DNA microarray scanner and process the data using a DNA microarray data analysis software package (Fig. 9; see Note 15).

3.7.3. Phosphorylation of Arrayed Protein Kinases by Solution Phase FES Kinase

One use of protein kinase arrays is to determine and identify the subset of the kineome that are substrates for a given solution-phase kinase, as follows:

1. Remove two replica protein kinase arrays from wash buffer and equilibrate in PBST at room temperature for 5 min.

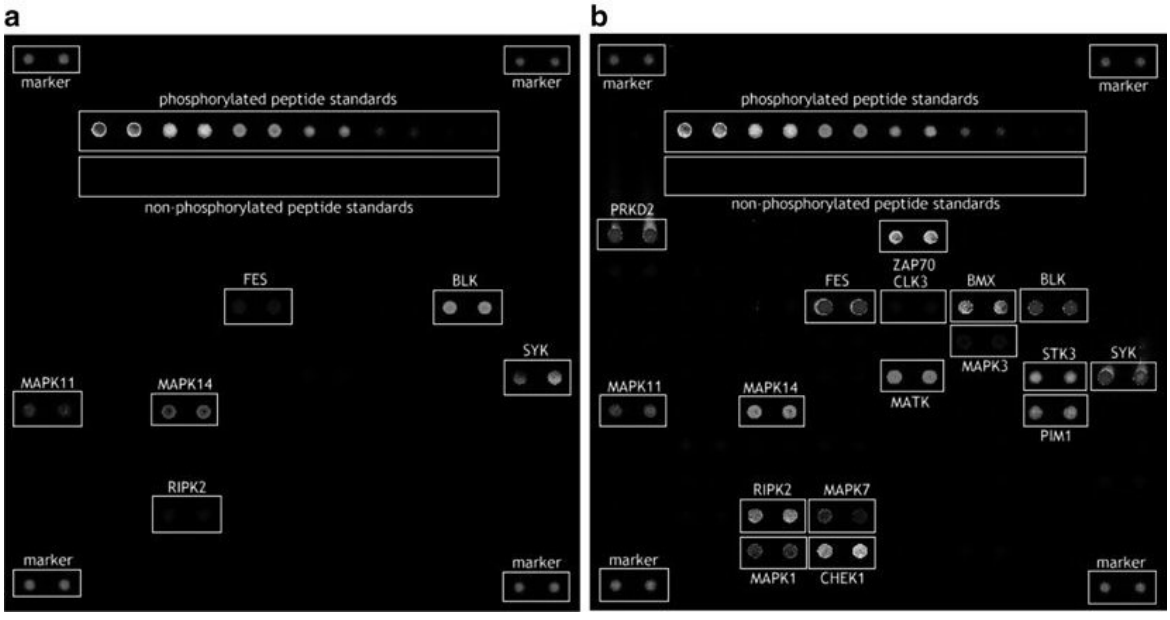

Fig. 9. Array-based autophosphorylation assays. Here, an array of 80 human protein kinases was printed in duplicate and assayed for autophosphorylation activity. (**a**) Kinase buffer only. (**b**) Kinase buffer containing 100 μM ATP. The assays were developed using a fluorescently labelled antiphosphotyrosine antibody and revealed ATP-dependent, on-array autophosphorylation for 17 human kinases, as marked.

2. Drain away the PBST, add 1 ml of kinase buffer supplemented with 100 μM ATP to each array.
3. To one of the arrays, also add 50 nM FES kinase (Upstate) and incubate both arrays at room temperature for 1 h.
4. Wash each array for 3×5 min with 1 ml of PBST containing 0.1% SDS.
5. Drain away the PBST, add 1 ml Cy3-labelled anti-phosphotyrosine antibody solution (from Subheading 3.7.2, step 1) to each array and incubate with gentle agitation at room temperature for 30 min.
6. Wash the arrays for 3×5 min with 1 ml of PBST each.
7. Remove the arrays to a 50-ml Falcon tube and centrifuge for 30 s to dry.
8. Scan the arrays at 550 nm using a DNA microarray scanner and process the data using a DNA microarray data analysis software package (Fig. 10; see Note 16).

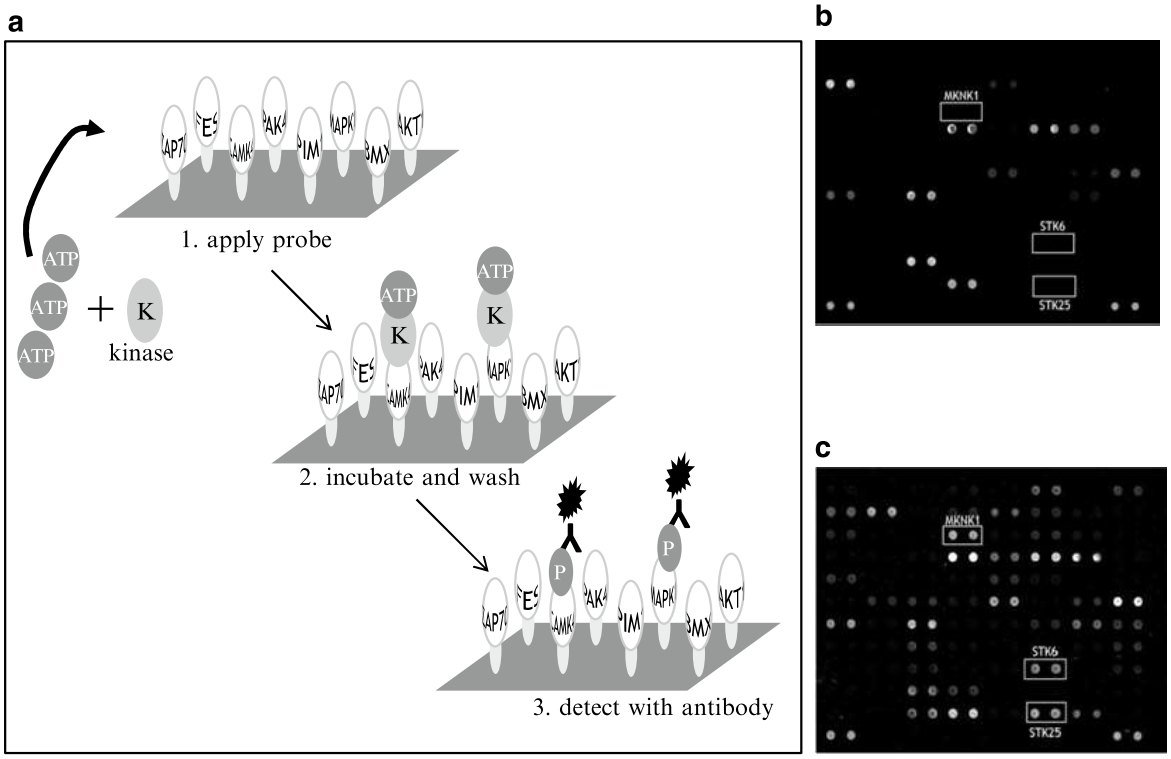

Fig. 10. Array-based phosphorylation assay using an exogenous kinase. An array of 96 human protein kinases was printed in duplicate and assayed in the presence and absence of exogenous Fes kinase. (**a**) Schematic of assay. (**b**) 100 μM ATP plus kinase buffer only. (**c**) 100 μM ATP, kinase buffer plus exogenous Fes kinase. The assays were developed using a fluorescently labelled antiphosphotyrosine antibody and revealed a number of substrates for Fes kinase, including MKNK1, STK6, and STK25, as marked.

3.7.4. Measuring the Binding of an Unlabelled Kinase Inhibitor to Each Arrayed Protein Kinase

Protein kinases are currently regarded as tractable drug targets in many different disease states. Most drug-like current kinase inhibitors target the ATP binding pocket to a greater or lesser degree, yet data from structural biology suggests that the ATP-binding pocket of protein kinases is strongly conserved. Cross-reactivity of protein kinase inhibitors is, therefore, likely to be an issue so low cost, simple, quantitative, yet high-throughput approaches to assess the selectivity of such inhibitors across large panels of human kinases have considerable potential. Protein kinase arrays can be used to provide such data as follows:

1. Remove eight replica protein kinase arrays from wash buffer and equilibrate in PBST at room temperature for 5 min.
2. Drain away the PBST, add 1 ml of kinase buffer to each array.
3. To each array, add an increasing concentration of a fluorescently labelled broad-spectrum kinase inhibitor (e.g. Cy3-labelled staurosporine; a fluorescent-ATP analogue; or fluorescently labelled polycyclic heteroaromatic compounds (27); concentration range from 0.5–50 nM) and incubate all arrays at room temperature for 1 h.
4. Wash each array for 3×5 min with 1 ml of PBST containing 0.1% SDS.
5. Remove the arrays to a 50-ml Falcon tube and centrifuge for 30 s to dry.
6. Scan the arrays at 550 nm using a DNA microarray scanner and process the data using a DNA microarray data analysis software package to determine the K_d for the binding of the fluorescent ligand to each arrayed kinase.
7. Remove a further eight replica protein kinase arrays from wash buffer and equilibrate in PBST at room temperature for 5 min.
8. Drain away the PBST, add 1 ml of kinase buffer to each array.
9. To each array, add fluorescently labelled broad-spectrum kinase inhibitor (10 nM final concentration) plus an increasing concentration of an unlabelled kinase inhibitor (e.g. Geftinib; concentration range from 0.5–50 nM) and incubate all arrays at room temperature for 1 h.
10. Wash each array for 3×5 min with 1 ml of PBST containing 0.1% SDS.
11. Remove the arrays to a 50-ml Falcon tube and centrifuge for 30 s to dry.
12. Scan the arrays at 550 nm using a DNA microarray scanner and process the data using a DNA microarray data analysis software package to determine the IC_{50} for the binding of the unlabelled inhibitor to each arrayed kinase (Figs. 11 and 12).

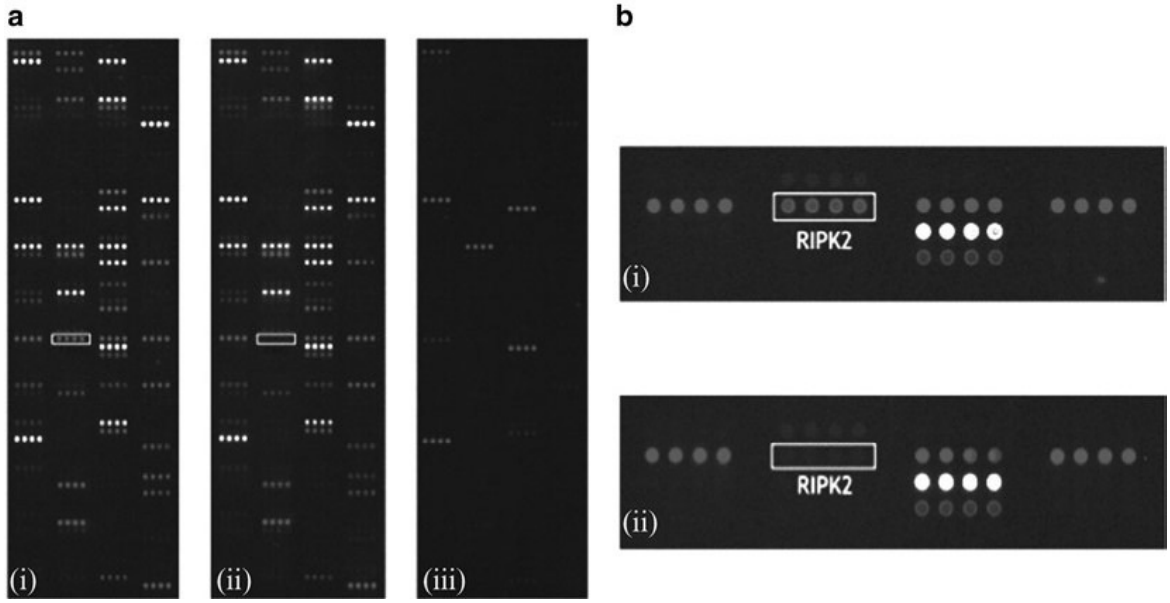

Fig. 11. Array-based inhibitor-binding assays. An array of 150 human protein kinases was printed in quadruplicate, one array per slide, and assayed for binding to a universal fluorescent kinase ligand in the presence and absence of small-molecule kinase inhibitors. (**a**) (i) Universal fluorescent ligand only; (ii) Universal fluorescent ligand plus Iressa, a specific kinase inhibitor; (iii) Universal fluorescent ligand plus staurosporine, a broad-spectrum kinase inhibitor. (**b**) RIPK2 was identified as a target for Iressa because the drug successfully competed with the universal fluorescent ligand for binding to this kinase. Note that the specific "universal" fluorescent ligand used in this assay only bound to 48 out of 150 kinases on the array; by using combinations of such ligands, higher coverage can be achieved.

NB. From the IC_{50} and K_d values, it is possible to derive the K_i value for the unlabelled inhibitor binding to each kinase by standard manipulations.

3.8. Data Analysis: General Principles

3.8.1. Raw Data Extraction

There are in principle numerous different ways in which the raw data can be pre-processed prior to analysis, depending on the precise microarray scanner software package used. We find that the most reliable method is as follows:

1. Use the highest gain setting on the microarray scanner that does not cause any signals to saturate.
2. Set a "pixel inclusion" threshold to exclude any dark pixels (i.e., pixels that strictly form part of the background) from within the analysis area.
3. Extract the raw data and determine the median foreground pixel intensity for each spot, as well as the median local background pixel intensity for each spot.
4. Subtract the median local background pixel intensity from the median foreground pixel intensity for each spot to give the net pixel intensity for each spot.
5. Using the technical replicate data, determine the mean of the median net pixel intensities for each arrayed protein; this is then the data to use in subsequent analyses.

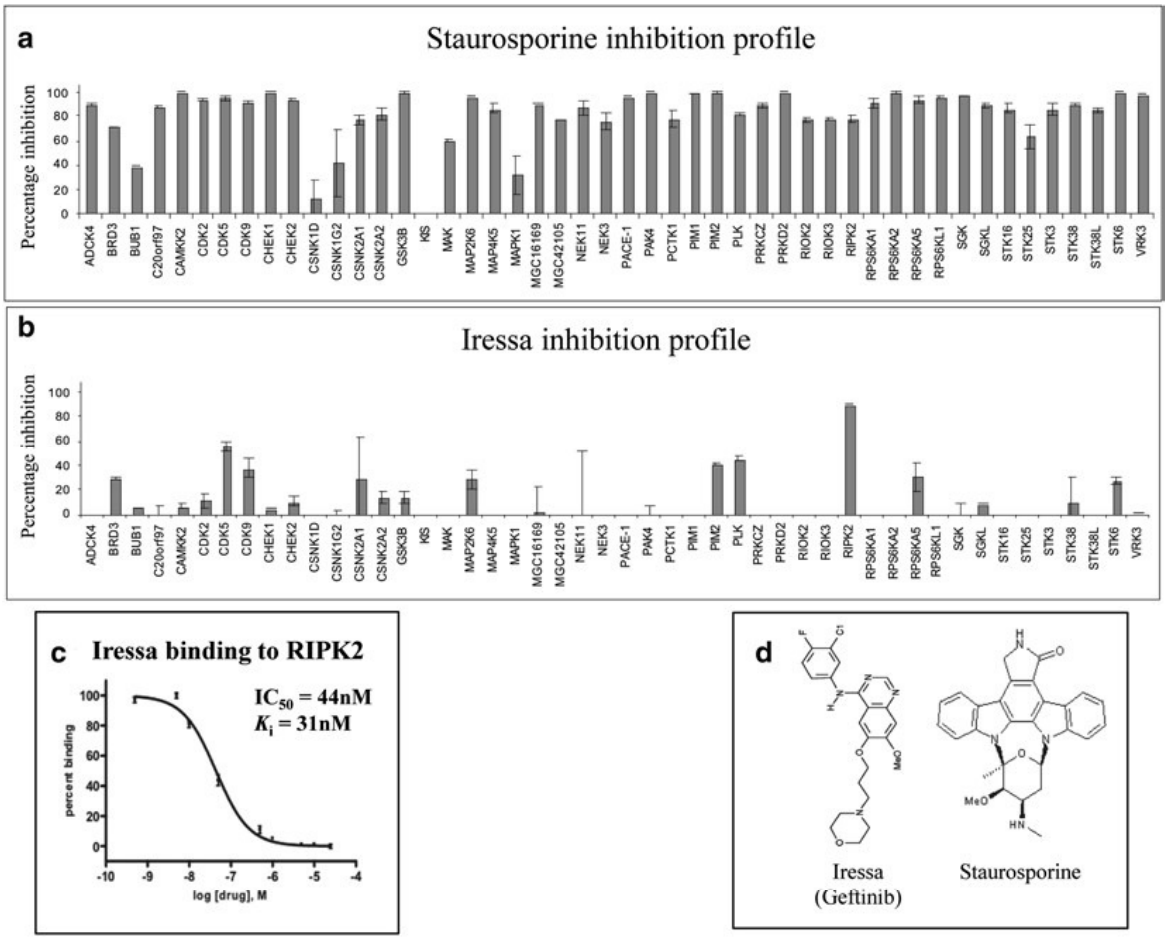

Fig. 12. Data analysis. From the primary data in Fig. 11, the ability of staurosporine and Iressa to compete with the universal fluorescent ligand for binding to each arrayed kinase can be quantified. (**a**) Percentage inhibition of arrayed kinases by 1 μM unlabelled staurosporine. (**b**) Percentage inhibition of arrayed kinases by 1 μM unlabelled Iressa. (**c**) Repeating the assay at different concentrations of Iressa enabled an IC_{50} value of 44 nM to be determined for the Iressa/RIPK2 interaction, in good agreement with literature values. From the IC_{50} value, we calculated that K_i (Iressa/RIPK2) = 31 nM. (**d**) Structures of Iressa and staurosporine, which both bind in the ATP-binding pocket of their target kinases. As expected, at 1 μM concentration, staurosporine competed effectively with the universal fluorescent ligand across the majority of arrayed kinases while Iressa was much more selective in its binding profile.

3.8.2. Data Analysis

Fluorescent ligand-binding data can be analysed in a number of different ways, including the use of global fitting algorithms. Manual analysis can be carried out by the use of Scatchard-type plots as follows:

1. Plot the relative amount of bound fluorescent ligand (FL) as a function of "ligand concentration in solution" for each protein in the array (7) and fit to a simple hyperbolic concentration–response curve according to:

$$R = \frac{B_{\max}[\mathrm{FL}]}{\{K_d + [\mathrm{FL}]\}},$$

where R is the response in relative fluorescence units and [FL] is the ligand concentration. From this, the ligand-binding constants (K_d) and maximum ligand-binding capacities (B_{max}) for each arrayed protein can be determined (7).

2. In competition-binding assays of the type described above (Subheading 3.7.4), the IC_{50} values for a specific protein–ligand interaction are dependent on the intrinsic K_i of the unlabelled ligand for the protein and on the K_d of the fluorescent ligand competitor, as well as on the concentration of the fluorescent ligand in the assay. The relationship is:

$$K_i = \frac{IC_{50}}{\left\{1 + \left(\frac{[FL]}{K_d}\right)\right\}}$$

4. Notes

1. In order to avoid laborious pre-purification of each recombinant protein prior to array fabrication, we have developed a procedure in which we combine the immobilisation and purification stages into a single-step (7). For this approach to work, the array surface itself must have a low capacity for non-specific binding of proteins, yet must have a high specificity and high-binding affinity for the proteins to be arrayed. In principle, a range of different affinity tags could be used here but in practice few actually offer sufficiently high-specificity interactions with the surface affinity matrix. For example, His-tags are not well-suited to such an approach since the specificity of the interaction with Ni^{2+} ligands is too low (particularly, in the context of expression in eukaryotic cells where there are numerous Ni^{2+}-binding proteins) and the intrinsic affinity of the interaction is also relatively low, resulting in leaching of protein from the array even in the absence of imidazole. To circumvent these problems, we have chosen to use an affinity tag that becomes biotinylated in vivo at a single specific residue, allowing us to make use of the very high affinity ($K_d \sim 10^{-15}$M) and specificity of the streptavidin–biotin interaction. Array fabrication thus becomes the simple process of printing crude lysates containing the recombinant biotinylated proteins onto streptavidin-coated surfaces which have a low non-specific protein-binding capacity (see Note 2), followed by washing to remove all non-biotinylated proteins from the array surface (Fig. 5).

   In the absence of any pre-purification step, it might be expected intuitively that host cell proteins which are endogenously

biotinylated would compete with the biotinylated recombinant proteins for the available streptavidin-binding sites on the array surface. Insect cells have five such endogenous proteins but under native conditions we have observed that these proteins do not compete efficiently with biotinylated recombinant proteins for binding to streptavidin. This perhaps reflects low natural expression levels and the fact that in endogenous biotinylated protein, the biotin is typically less solvent accessible in the native protein (28) meaning that the biotin may not be physically accessible to bind to streptavidin.

We have also observed that using the streptavidin–biotin interaction as the basis for array fabrication confers an additional major advantage: the very high affinity of the streptavidin–biotin interaction means that we quickly start to saturate the available biotin-binding sites on the surface so a crude normalisation of protein loading can be achieved without pre-adjusting the concentrations of the crude lysates to compensate for differences in the individual expression levels of the different recombinant proteins (7).

As with many affinity tags, biotinylated affinity tags can be positioned at either the N- or C-terminus of the protein to be arrayed, dependent on the structural and functional characteristics of the protein. In the context of oriented protein microarray fabrication, "biochemical" biotinylation seems preferable to "chemical" biotinylation since the latter offers little control over the site of biotinylation and still requires pre-purification of each protein to remove excess biotinylation reagent. There are currently three main alternatives for biochemical introduction of a biotin moiety into a recombinant protein: two involve affinity tags that can be biotinylated in vivo or in vitro while the third involves an intein-mediated introduction of biotinylated cysteine.

The AviTag (Avidity, Colorado USA) is an in vitro evolved 15 residue peptide that is specifically biotinylated exclusively by the *E. coli* biotin ligase (29). Alternatively, a single biotinylated cysteine moiety can be added to the C-terminus of a recombinant protein during intein-mediated protein splicing of fusion proteins (30). However, it is not entirely clear what advantages this latter route offers over.

We have chosen however to use a compact, folded, biotinylated, ~80 residue domain derived from the *E. coli* BCCP (Fig. 1) (21, 22) since this affords two significant advantages over the AviTag and intein-based tag. Firstly, the BCCP domain is cross-recognised by eukaryotic biotin ligases enabling it to be biotinylated efficiently in yeast, insect, and mammalian cells without the need to co-express the *E. coli* biotin ligase (31–33). Secondly, the N- and C-termini of BCCP are physically separated from the site of biotinylation by ~50 Å (Fig. 1) (21),

so the BCCP domain can be thought of as a stalk which presents the recombinant proteins away from the surface, thus minimising any deleterious effects due to immobilisation.

Vectors for expressing proteins as fusions to a BCCP domain derived from *Klebsiella pneumoniae* are now available from Invitrogen (pET104 DEST Bioease). This *K. pneumoniae* domain is highly homologous to the *E. coli* BCCP protein and confers the same properties.

2. We and others have found that organic polymer coatings, such as those based on dextran or polyethylene glycol (PEG) (either in the form of long chain PEGs or of short chain PEGs supported on a self-assembled monolayer), are considerably superior to proteinaceous blocking agents, such as BSA or powdered milk, in reducing the non-specific-binding background in surface-based assays (34); a number of commercially available surfaces conform to this specification.

   One such commercially available PEG-based streptavidin-coated surface that works well in resisting non-specific protein binding is the Nexterion Slide S from Schott, although we have also used other streptavidin-coated glass or plastic surfaces with similar "non-stick" properties. However, most if not all pre-coated streptavidin slides must be shipped on dry-ice in order to preserve the integrity of the streptavidin layer and the shipping costs can therefore be substantial, depending on location. In order to circumvent this cost issue, we have found that we can custom coat amine-reactive slides (for example, the PEG-based, NHS-activated Nexterion Slide P from Schott, which ships at room temperature and therefore at significantly lower cost) with streptavidin at point of use. Importantly, we have found that our home-made slides show as good uniformity of the streptavidin coating as commercial, pre-made streptavidin surfaces (data not shown); furthermore, we found that utilising an automated hybridisation station in an effort to gain greater control over the streptavidin-coating process gave no obvious advantage over the manual process described here (data not shown).

3. We observed that the addition of free biotin to the growth medium increases the extent of biotinylation of the recombinant BCCP fusion protein. Importantly, we have not found it necessary to also overexpress the *E. coli* biotin ligase in any expression host when using the BCCP tag. The wash step prior to cell lysis is needed to remove free biotin in the media before array fabrication; if the protein is purified before array fabrication, this is then not necessary.

4. When expressing a large number of clones in parallel for array fabrication, the Bradford assay can conveniently be done on all clones in parallel using a microtitre plate format. However, it

would be laborious to carry out SDS PAGE analysis on all clones for each and every repeat expression run, so typically we only assess a selection of clones in this way since the absolute expression level is not critical for array fabrication. For the majority of clones, we have found that a simpler dot-blot assay can form a reliable qualitative indicator that biotinylated recombinant protein has been expressed in the specific set of insect cell cultures in question (data not shown).

5. The growth temperatures and times described here have proven to be directly applicable to all BCCP fusions we have sought to express in insect cells, which number in the hundreds. Since our approach to downstream array fabrication does not require pre-purification of recombinant proteins prior to printing, we have not found it necessary to try to maximise expression levels of any one protein. To the contrary, we have found heuristically that as long as we can observe an expressed, biotinylated recombinant protein by Western blot, there will then be enough recombinant protein to array and assay.
6. The biotinylated component of the sample is supershifted by streptavidin even under denaturing conditions, enabling a simple side-by-side comparison to be carried out. To save time, this assay need not be done on all clones since we have observed that if the recombinant BCCP proteins are expressed and folded, the BCCP domain is efficiently biotinylated.
7. We have found solid pins the easiest to clean rapidly and have observed no carry over between samples using such pins. In addition, we have found that solid pins also work well printing directly onto many different surface types.
8. Each spotting event delivered ~10 nl liquid. We typically use multiple stamps per spot to increase the protein loading at each position in the array, and we have found that these general printing parameters work well with other surfaces, although care must be taken in calibrating the z-height and touch-down velocities on the robot when using fragile surfaces. In addition, by using the same printing parameters under the same conditions of printing (i.e. in a glycerol-containing buffer to reduce evaporation rates), pre-purified proteins can be spotted onto non-selective surfaces which bind proteins by chemical cross-linking (e.g. epoxide or aldehyde-coated glass (14, 15)) or by non-covalent adsorption (e.g. supported nitrocellulose, agarose, or polyacrylamide (15)).
9. Each 50 µl aliquot of harvested expression cells provides enough recombinant protein to print 25 replica slides in 4-plex format, with each protein printed in triplicate in each sub-array. Thus, one 3 ml baculovirus culture yields enough expressed

protein to fabricate 700 replica sub-arrays, or >2,000 replica spots of each protein, which is an important consideration when seeking to minimise array-to-array variability. It is important to note, though, that not all the sample volume loaded into a 96-well V-bottomed source plate can be used for printing because care needs to be taken to ensure that the uniformity of the spot size is not affected by any change in the volume remaining in the source plate.

10. We routinely print each protein in triplicate within each array so that we have three "technical replicates" during downstream data analysis. Furthermore, depending on the desired downstream assay format, it is simple with robotic arrayers to print 1, 2, 4, 8, or 16 replica sub-arrays on each 7.5 × 2.5-cm glass slide. Obviously, the higher the number of sub-arrays per slide, the lower the number of discrete protein spots that can be accommodated in each sub-array. By way of example, using a "4-plex" format, with 8 × 300 μm solid pins it is possible to print eight 8 × 8 panels (=512 discrete protein spots) in each of the four replica sub-arrays with a spot-to-spot spacing of 500 μm; this would enable, for example, 170 individual protein types to be printed in triplicate in each of the four sub-arrays.
11. Importantly, with the ready availability of reusable gasket systems today (e.g., Gentel Biosciences SIM*plex*™ system; Whatman FAST Frame system; etc.), it is possible now to assay each sub-array on any one slide under independent conditions, thus minimising the impact of the cost of the base slides on the "cost per assay" and "cost per data" point to much more reasonable levels. This capability also allows a number of parallel assays to be carried out on a single slide, thus reducing the possible impact of slide-to-slide variability on experimental error.
12. By printing, processing, and assaying arrays using these protocols, we have been able to achieve spot-to-spot CVs of 10% (Fig. 7).
13. The blocking and wash steps should remove all non-biotinylated proteins from the array surface while the biotin in the milk powder blocks any remaining biotin-binding sites on the streptavidin surface such that any excess biotinylated proteins which have not bound within the specific spot cannot then rebind elsewhere on the array. We have found that under the simple storage conditions described here, our arrays remain viable for around 3 weeks, after which the loss of activity of the arrayed proteins starts to be observed. We have made no effort to address this issue though since our interest does not lie in trying to make protein array products here.

14. We have found better performance in this assay when using a chemiluminescent end point than when using a fluorescent end point. The reasons for this are not entirely clear, but may relate to factors surrounding the physical accessibility of the c-Myc epitope in a non-denatured protein array (the c-Myc epitope is on the C-terminus of the BCCP tag, which itself forms the link to the surface, so may be occluded by the fusion protein unless deliberately denatured prior to assay); the signal amplification obtained via chemiluminescence appears to be helpful, but this also suggests that the resultant data from this assay may not be quantifiable in a meaningful sense.
15. We routinely observe that some human protein kinases expressed in insect cells are already phosphorylated prior to assay. However, many other kinases become newly phosphorylated during this assay. There are two obvious explanations for this: firstly, the autophosphorylation of these latter kinases might simply require the higher ATP concentration used in the assay (compared to in vivo concentrations); or secondly, that the autophosphorylation event requires dimerisation of the kinases in question and that such dimerisation is promoted by the surface attachment in the arrays.
16. It is also possible to carry out the same assay using radioactive ATP if desired (we have found 100 μM ATP at 60 Ci/mmol to work well, in which case the detection is by direct autoradiography. This assay format has the immediate advantage of not being restricted by the specificity of the anti-phosphopeptide antibodies. Furthermore, in such cases, it is also then possible to mask competing autophosphorylation events by carrying out a pre-incubation with 100 μM cold ATP.

Acknowledgments

The authors thank Nashied Peton, Natasha Beeton-Kempen, Sarah Joyce, Colin Wheeler, Jens Koopman, Nick Workman, Steve Parham & Mike McAndrew for their help in generating the data detailed herein. We thank Procognia Ltd. (UK) for provision of arrays and also thank the Centre for Proteomic & Genomic Research, Cape Town, for access to equipment. JMB thanks the National Research Foundation (South Africa) for a Research Chair. The research was supported by grants from the National Research Foundation, Procognia Ltd. and Genetix PLC (UK).

References

1. Hunter S., Apweiler R. *et al.* (2009) InterPro: the integrative protein signature database (2009). Nucleic Acids Res. **37**, D224–228.
2. Wise, E.Y., Yew, W.S., Babbitt, P.C., Gerlt, J.A., & Rayment, I. (2002) Homologous $(b/a)_8$-Barrel Enzymes That Catalyze Unrelated Reactions: Orotidine 5'-Monophosphate Decarboxylase and 3-Keto-L-Gulonate 6-Phosphate Decarboxylase. *Biochemistry* **41**, 3861–3869.
3. Schmidt, D.M.Z, Mundorff, E.C., Dojka, M., Bermudez, E. *et al.* (2003) Evolutionary potential of $(b/a)_8$-Barrels: Functional promiscuity produced by single substitutions in the enolase superfamily. *Biochemistry* **42**, 8387–8393.
4. The genome international sequencing consortium (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.
5. MacBeath, G. (2002) Protein microarrays and proteomics. *Nature Genetics* **32**, 526–532.
6. Wolf-Yadlin, A., Sevecka, M. & MacBeath, G. (2009) Dissecting protein function and signaling using protein microarrays. *Current Op. Chem. Biol.* **13**, 398–405.
7. Boutell, J.M., Hart, D.J., Godber, B.L.J., Kozlowski, R.Z. & Blackburn, J.M. (2004) Analysis of the effect of clinically-relevant mutations on p53 function using protein microarray technology. *Proteomics* **4**, 1950–1958.
8. Kodadek, T. (2001) Protein microarrays: prospects and problems. *Chem. Biol.* **8**, 105–115.
9. Predki, P. (2004) Functional protein microarrays: Ripe for discovery. *Curr. Opin. Chem. Biol.* **8**, 8–13.
10. Zhu, H., Klemic, J.F., Chang, S., Bertone, P. *et al.* (2000) Analysis of yeast protein kinases using protein chips. *Nat. Genet.* **26**, 283–289.
11. Zhu, H., Bilgin, M., Bangham, R., Hall, D. *et al.* (2001) Global analysis of protein activities using proteome chips. *Science* **293**, 2101–2105.
12. Michaud, G.A., Salcius, M. Zhou, F. Bangham, R. et al. (2003) Analyzing antibody specificity with whole proteome microarrays. *Nature Biotechnology* **21**, 1509–12.
13. Fang, Y., Lahiri, J. & Picard, L. (2003) G-Protein-coupled receptor microarrays for drug discovery. *Drug Discovery Today* **8**, 755–761.
14. MacBeath, G. & Schreiber, S.L. (2000) Printing proteins as microarrays for high-throughput function determination. *Science* **289**, 1760–1763.
15. Angenendt, P., Glokler, J., Sobek, J., Lehrach, H. & Cahill, D.J. (2003) Next generation of protein microarray support materials: Evaluation for protein and antibody microarray applications. *J. Chromatogr. A* **1009**, 97–104.
16. Koopmann, J.-O. & Blackburn, J.M. (2003) High Affinity Capture Surface for MALDI compatible Protein Microarrays. *Rapid Communication in Mass Spectrometry* **17**, 1–8.
17. He, M. & Taussig, M. (2001) Single step generation of protein arrays from DNA by cell-free expression and *in situ* immobilization (PISA method). *Nucleic Acids Res.* **29**, E73.
18. Ramachandran, N., Hainsworth, E. et al. (2004) Self-Assembling Protein Microarrays. *Science* **305**, 86–90.
19. Zhao, Y. Chapman, D.A.G. & Jones, I.M. (2003) Improving baculovirus recombination. *Nucleic Acids Res.* **31**, e6.
20. Sambrook, J., *MacCallum, P. &* Russell, D. (2001) *Molecular Cloning, A Laboratory Manual*, Third ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
21. Athappilly, F.K. & Hendrickson, W.A. (1995) Structure of the biotinyl domain of acetylcoenzymeA carboxylase determined by MAD phasing. *Structure* **3**, 1407–19.
22. Chapman-Smith, A. & Cronan, J.E. (1999) The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. *Trends Biochem. Sci.* **24**, 359–363.
23. Guex, N. & Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723.
24. Yang,Y-S., Watson, W.J., Tucker, P.W. & and Capra, J.D. (1993) Construction of recombinant DNA by exonuclease recession. Nucleic Acids Res. **21**, 1889–1893.
25. Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.
26. Terwilliger, T.C., Stuart, D. & Yokoyama, S. (2009) Lessons from Structural Genomics. *Ann. Rev. Biophys.* **38**, 371–383.
27. Brown, M. (2007) Novel fluorescent kinase ligands and assays employing the same. Patent application no. WO2008071937.
28. Choi-Rhee, E. & Cronan, J.E. (2003) The biotin carboxylase-biotin carboxyl carrier