Printing protein arrays from DNA arrays

Mingyue He[1,4], Oda Stoevesandt[1,3,4], Elizabeth A Palmer[1,3], Farid Khan[1,3], Olle Ericsson[2] & Michael J Taussig[1,3]

We describe a method, DNA array to protein array (DAPA), which allows the 'printing' of replicate protein arrays directly from a DNA array template using cell-free protein synthesis. At least 20 copies of a protein array can be obtained from a single DNA array. DAPA eliminates the need for separate protein expression, purification and spotting, and also overcomes the problem of long-term functional storage of surface-bound proteins.

Protein arrays are powerful tools for large-scale, parallel protein analysis, applicable to high-throughput screening of protein activities and interactions. The growing number of applications of protein arrays includes proteome expression profiling, biomarker discovery, detection of protein modifications, and characterization and quality control of binding molecules[1–5]. Despite considerable recent progress, protein array technology still has to overcome several important technical challenges to achieve its maximum capability. One hurdle is that availability of purified, functional proteins for immobilization on arrays often creates a considerable limitation, especially for human proteins. Second, unlike DNA, immobilized proteins may be difficult to store in a functional state over long periods of time. To solve both of these problems, we and others have previously developed methods for the cell-free *in situ* transcription and translation of DNA to form a protein array[6–11]. Here we present a new method, DAPA, for repeatable printing of protein arrays from a single DNA template array, on demand.

In the DAPA concept, a slide with an array of covalently immobilized, PCR-amplified fragments encoding a set of tagged proteins is assembled face-to-face with a second slide, functionalized with the tag-capturing reagent (**Supplementary Methods** online). A permeable membrane carrying a cell-free lysate, capable of performing coupled transcription and translation, is positioned between the two slide surfaces. With protein synthesis originating from the spots of immobilized DNA, the newly synthesized proteins diffuse through the membrane and become rapidly immobilized on the capture slide surface, creating the protein array corresponding to the DNA array template (**Fig. 1a**). We designed and used a simple apparatus for incubating slide pairs (**Supplementary Fig. 1** online).

To demonstrate the DAPA principle, we produced an array of Cy-5–labeled PCR-amplified DNA encoding GFP fused with a C-terminal double-hexahistidine (6His) tag[6,12] together with the

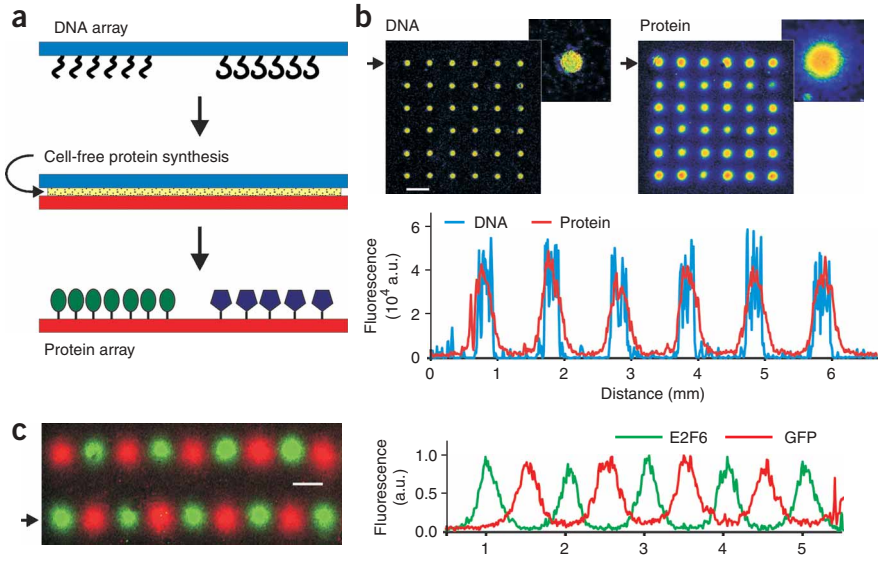

Figure 1 | Principle and examples of DAPA. (**a**) Schematic diagram of the DAPA procedure. Yellow, membrane carrying the cell-free transcription and translation system. (**b**) Array of Cy-5–labeled PCR-amplified DNA encoding double-6His–tagged GFP and a corresponding protein array of GFP on an Ni-NTA slide, after DAPA for 4 h at 30 °C using *E. coli* cell-free protein synthesis system. Immunostaining by biotinylated anti-GFP, HRP-linked streptavidin and tyramide-Cy3. The protein array has been mirror imaged to match individual spot positions of the DNA array. Scale bar, 1 mm. Below the array images, fluorescence intensity profile of DNA and protein spots in rows marked by arrows. (**c**) Specific two-color immunostaining of DAPA spots of GFP (red) alternating with spots of myc-tagged E2F6 (green; left), and normalized fluorescence intensity profile along row of spots marked by the arrow (right). Immunostaining by biotinylated anti-GFP followed by streptavidin-Alexa647, and HRP-linked anti-myc followed by tyramide-Cy3. Scale bar, 500 μm.

[1]Technology Research Group, The Babraham Institute, Cambridge CB22 3AT, UK. [2]Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University, SE-75185 Uppsala, Sweden. [3]Present addresses: Protein Technology Group, Babraham Bioscience Technologies, Babraham Research Campus, Cambridge CB22 3AT, UK (O.S. and M.J.T.), MRC Virology Unit, Glasgow G11 5JR, Scotland, UK (E.A.P), and University of Manchester, Manchester Interdisciplinary Biocentre, Manchester, M1 7DN, UK (F.K.). [4]These authors contributed equally to this work. Correspondence should be addressed to M.J.T. (mike.taussig@bbsrc.ac.uk).

RECEIVED 22 NOVEMBER 2007; ACCEPTED 20 DECEMBER 2007; PUBLISHED ONLINE 20 JANUARY 2008; DOI:10.1038/NMETH.1178

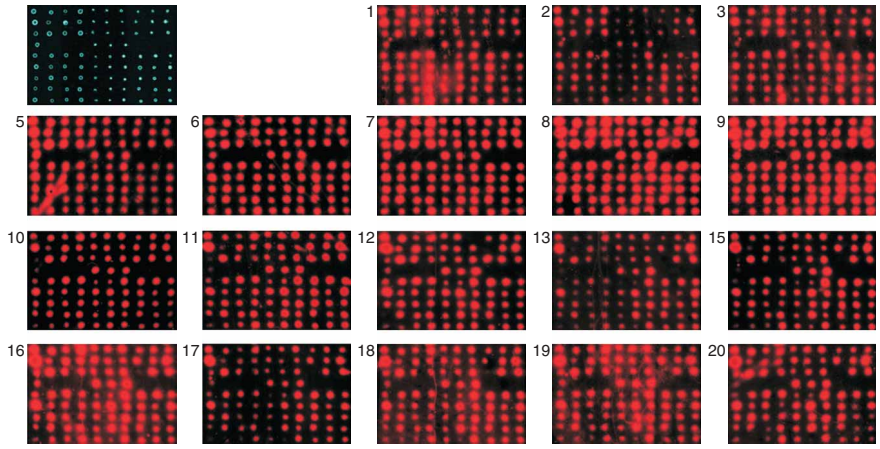

Figure 2 | Repeated protein array printing from a single DNA array template. A series of consecutive replicate DAPA protein arrays was printed from the same DNA array template. Incubation and detection conditions were identical for each print. Eighteen out of 20 attempted repeats were successful (two failed for handling reasons). Blue, PCR-amplified DNA encoding GFP directly labeled with Cy-5. Red, GFP detected by immunostaining as in **Figure 1b**. Protein arrays are mirror imaged to match DNA array.

required sequence elements for *in vitro* protein expression (**Supplementary Methods**). The DNA array was transcribed and translated into a protein array of His-tagged GFP captured on a Ni-NTA–coated slide (**Fig. 1b**). The protein spots, detected by anti-GFP, mirrored the layout of the DNA array and showed a Gaussian profile resulting from the diffusion gradient of the newly synthesized protein originating from the DNA spots (**Fig. 1b** and **Supplementary Fig. 2** online). For DNA spots with a diameter of 250 μm, the average diameter of the protein spots, measured as the width of the Gaussian profile at half height, was 275 μm.

To estimate the amount of protein generated by DAPA, we compared signal intensities of DAPA-printed GFP spots with intensities of directly spotted GFP of known concentration and volume. From 0.1 ng of template DNA, up to 30 fmol of protein per spot could be deposited (data not shown), comparing well with existing technologies for protein array generation[7,13]. In general, the amount of protein made increases with the amount of DNA spotted. We performed specific detection of DAPA-printed GFP and the DNA-binding domain of transcription factor E2F6 on the same array (**Fig. 1c**). Mean deviations from the mean signal intensities of the protein spots were 11% for both GFP and E2F6, compared to 7% and 5% for their respective template DNA spots. We tested various DAPA-printed proteins using template DNA arrays encoding the signal transduction proteins Grb2 and Rab22B, the metalloproteinase inhibitor TIMP1, DNA-binding domains of 16 human transcription factors and two single-chain antibody fragments (**Supplementary Fig. 3** online). The single chain antibody fragments retained their binding specificity, showing that the printed proteins can be functional (**Supplementary Fig. 4** online), though we did not extensively test for this.

To test whether a single DNA array could be reused to print multiple copies of a protein array, we repeated the DAPA procedure 20 times on a DNA array encoding GFP (**Fig. 2**). Although there was some variation in the intensity of protein detection on the individual DAPA arrays, we found no general deterioration or drop in efficiency of the procedure over the 20 repeats. It is likely that even more reprintings could be made. Thus, even in laboratories without access to, or expertise in, routine microarray spotting, protein arrays can be printed rapidly, economically and on demand from a single reusable DNA template microarray. The effort involved is comparable to standard methods such as assembly of a western blot.

The utility of DAPA will be dependent on the expressibility of the proteins in the cell-free system of choice. Although the number of proteins known to be expressed in cell-free systems is large[7,14], the yield tends to decrease for proteins of high molecular weights. The *Escherichia coli* S30 system we used is optimally suited for synthesis of proteins up to 120 kDa. However, low expression can be compensated for by increasing the amount of DNA template or by fusion to a well-expressed tag[11]. To date, only single-chain proteins have been tested with DAPA. It remains to be shown that multimeric proteins assemble effectively.

As well as acting as a new milieu for cell-free protein synthesis, the membrane provides a matrix, stabilizing the protein gradient against thermal mixing or agitation. Compared with conventionally spotted protein arrays, the protein spots made in DAPA show effects of protein diffusion, occupying a larger area than the corresponding DNA spot and with a Gaussian-profile of concentration from the center to the outside (**Supplementary Fig. 2**). Therefore, where identical proteins are expressed in adjacent spots, the spacing should be adjusted appropriately. However, for array layouts with neighboring nonidentical spots, overlap of the diffusion zones of the protein spots does not interfere with individual protein detection or determination of the local background for specific staining (**Fig. 1c**). This can be exploited to reduce further the spot-to-spot spacing of template arrays and increase protein spot number. Hence, diffusional broadening of DAPA spots should not interfere with downstream applications.

In the concept of nucleic acid programmable protein arrays (NAPPA), protein arrays are generated from plasmid DNA arrays by cell-free synthesis with capture of tagged proteins onto the same array surface by an immobilized antibody to the fusion tag[9]. However, in NAPPA, translated proteins become colocalized with the plasmid DNA and the capturing antibody, in what is effectively a mixed DNA, antibody and target protein array. Another *in situ* strategy using cell-free translation is an adaptation of mRNA display, in which protein arrays are generated through puromycin capture of nascent polypeptides and immobilization via biotin and streptavidin[10]. This method yields sharply defined protein spots, but requires additional manipulations to transcribe and modify the mRNA before arraying. The DAPA method has the considerable advantage of generating a pure protein array on a separate surface, avoiding any potential interference of colocalized molecules in downstream applications. Moreover, it has the clear benefit of allowing repeated use of the same DNA array, which is not possible in the NAPPA or puromycin capture designs.

DAPA also offers some substantial advantages over conventional protein array methods, in which purified proteins are spotted. First, printing the proteins from a DNA array eliminates the need for

separate protein expression and purification. Second, the use of DNA arrays to create protein arrays on demand avoids storage of immobilized proteins and associated problems of functional deterioration. Third, the use of PCR fragments as templates for cell-free protein synthesis without the need for DNA cloning provides a rapid means for translating genomic information into functional protein analysis. This could be particularly useful for generating arrays of individual domains, mutant proteins and others. Fourth, cell-free systems allow the conditions of protein synthesis to be modified by introducing non-natural or chemically modified amino acids, or enzymes for post-translational modification[15,16]. We envision that the DAPA method will find important applications in the functional analysis of the proteome.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

We thank H. Liu for expert technical assistance, B. Korn (DKFZ, Heidelberg) and J. Taipele (University of Helsinki) for providing clones. Research at the Babraham Institute is supported by the Biotechnology and Biological Sciences Research Council. Work on this project has been supported through the European Commission 6th Framework Programme Integrated Project MolTools.

AUTHOR CONTRIBUTIONS

M.H. and M.J.T. devised the principle of DAPA; M.H. developed and exemplified the method; O.S. optimized the method, devised the DAPA apparatus and designed and performed the experiments shown in **Figures 1** and **2**; E.A.P. performed initial DAPA experiments, contributing examples to the data shown in **Supplementary Figures 3** and **4**; F.K. characterized the double-6His tag for protein immobilization and protein detection methods; O.E. spotted the DNA template slides for experiments shown in **Figure 1**; M.H., O.S. and M.J.T. prepared the manuscript.

Published online at http://www.nature.com/naturemethods/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Bertone, P. & Snyder, M. *FEBS J.* **272**, 5400–5411 (2005).
2. Wingren, C. & Borrebaeck, C.A. *Expert Rev. Proteomics* **1**, 355–364 (2004).
3. Horn, S. *et al. Proteomics* **6**, 605–613 (2006).
4. Wulfkuhle, J.D., Edmiston, K.H., Liotta, L.A. & Petricoin, E.F. III. *Nat. Clin. Pract. Oncol.* **3**, 256–268 (2006).
5. Stoll, D., Templin, M.F., Bachmann, J. & Joos, T.O. *Curr. Opin. Drug Discov. Dev.* **8**, 239–252 (2005).
6. He, M. & Taussig, M.J. *Nucleic Acids Res.* **29**, e73 (2001).
7. Angenendt, P., Kreutzberger, J., Glokler, J. & Hoheisel, J.D. *Mol. Cell. Proteomics* **5**, 1658–1666 (2006).
8. Nord, O., Uhlen, M. & Nygren, P.A. *J. Biotechnol.* **106**, 1–13 (2003).
9. Ramachandran, N. *et al. Science* **305**, 86–90 (2004).
10. Tao, S.C. & Zhu, H. *Nat. Biotechnol.* **24**, 1253–1254 (2006).
11. He, M., Stoevesandt, O. & Taussig, M.J. *Curr. Opin. Biotechnol.* (in the press).
12. Khan, F., He, M. & Taussig, M.J. *Anal. Chem.* **78**, 3072–3079 (2006).
13. MacBeath, G. & Schreiber, S.L. *Science* **289**, 1760–1763 (2000).
14. Langlais, C. *et al. BMC Biotechnol.* **7**, 64 (2007).
15. Murakami, H., Ohta, A., Ashigai, H. & Suga, H. *Nat. Methods* **3**, 357–359 (2006).
16. Devirgiliis, C., Gaetani, S., Apreda, M. & Bellovino, D. *Biochem. Biophys. Res. Commun.* **332**, 504–511 (2005).