Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Biotechnology

In situ synthesis of protein arrays

Mingyue He, Oda Stoevesandt and Michael J Taussig

In situ or on-chip protein array methods use cell free expression systems to produce proteins directly onto an immobilising surface from co-distributed or pre-arrayed DNA or RNA, enabling protein arrays to be created on demand. These methods address three issues in protein array technology: (i) efficient protein expression and availability, (ii) functional protein immobilisation and purification in a single step and (iii) protein on-chip stability over time. By simultaneously expressing and immobilising many proteins in parallel on the chip surface, the laborious and often costly processes of DNA cloning, expression and separate protein purification are avoided. Recently employed methods reviewed are PISA (protein *in situ* array) and NAPPA (nucleic acid programmable protein array) from DNA and puromycin-mediated immobilisation from mRNA.

Addresses
Technology Research Group, The Babraham Institute, Cambridge CB22 3AT, UK

Corresponding author: He, Mingyue (mingyue.he@bbsrc.ac.uk), Stoevesandt, Oda (oda.stoevesandt@bbsrc.ac.uk) and Taussig, Michael J (mike.taussig@bbsrc.ac.uk)

Current Opinion in Biotechnology 2008, **19**:4–9

This review comes from a themed issue on
Analytical Biotechnology
Edited by Thomas Joos and Paul E. Kroeger

Available online 18th January 2008

0958-1669/$ – see front matter

DOI 10.1016/j.copbio.2007.11.009

Introduction

Protein arrays, one of the most active areas in biotechnology today, are miniaturised, solid-phase ligand binding assay systems using immobilised proteins. They are capable of highly multiplexed analyses with the advantage over classical ELISA of an ultra-economical use of reagents and samples. They fall into two general categories, namely (a) capture arrays, where multiplexed affinity reagents, often but not exclusively antibodies, are used to detect and quantitate analytes in complex mixtures such as plasma or tissue extracts, and (b) target protein arrays, in which large numbers of purified proteins are used to monitor biochemical functions or to detect and characterise antibodies. Protein microarrays have potential in diagnostics for high throughput determinations of biomarkers or antibodies in plasma, while in proteomics research they find applications in protein expression profiling and in the expansive field of protein interaction analysis [1–3].

A major challenge in protein arraying is to express and purify the maximum diversity of proteins for array construction, whether they are antibodies or other proteins. The usual starting point is a collection of soluble, purified proteins, which are covalently or noncovalently attached onto suitable surfaces such as derivatised glass slides or beads. Recombinant proteins from an expression library are often used and while many can be expressed in cellular systems, there are many others (including antibodies) for which expression can be a major problem, because of insolubility, disulphide bonds or toxicity for the cellular host [4]. In such cases there are considerable advantages in cell free protein expression from DNA, using commercially available *in vitro* systems.

In addition to protein availability, a second issue is to maintain folding and function in an immobilised state over long periods of storage, bearing in mind that proteins are by nature heterogeneous, each with particular physicochemical properties and stability characteristics. A solution can be found by developing cell free production as *in situ* (on-chip) methods, where the proteins are immobilised simultaneously with their production, enabling arrays to be created as and when required. Problems of availability and stability do not arise with DNA arrays, where the material is readily synthesised, physically homogeneous and reliably stable. *In situ* protein arraying attempts to makes use of this by converting nucleic acid arrays into the less stable protein arrays on demand.

Cell free protein expression

Cell free systems accomplish protein synthesis by means of cell extracts (lysates) containing all the essential elements for transcription and translation. Such expression systems have been made from several different species and cells, of which *Escherichia coli*, rabbit reticulocyte and wheat germ are commonly used. Others include lysates extracted or prepared from hyperthermophiles, hybridomas, *Xenopus* oocytes, insect, mammalian and human cells [5•,6•,7,8]. Thus, proteins can be produced in different environments (prokaryotic, eukaryotic) and temperatures [5•]. Transcription and translation can be coupled, allowing PCR or plasmid DNA to be directly used as the template, providing a rapid tool for high-throughput protein synthesis [9•]. In the PURE system, cell free expression is reconstituted using purified individual proteins, tRNAs, essential substrates and required enzymes [10]. In general, proteins made in cell free systems are soluble and functional and the open, flexible nature of the systems permits addition of

further components, providing an adjustable environment for protein folding, disulphide bond formation, modification or activity. Cell free systems can also directly label proteins during translation in readiness for downstream detection: isotopes, fluorophores, biotin and photoreactive groups can be incorporated at defined positions during protein synthesis [11•,12]. For on-chip production, they can be miniaturised down to a nanoliter level by piezoelectric dispensing [13].

The level of protein expression in the cell free system is clearly an important factor when making *in situ* arrays. While many proteins are expressed readily, others need modifications to the composition of the system itself, for example energy sources or amino acid concentration [14,15], or to the construct, for example fusion of a well-expressed N-terminal tag sequence. Somewhat surprisingly, we have recently observed that C-terminal fusion of the constant domain of immunoglobulin κ light chain consistently improves the expression level of many proteins which are otherwise difficult to express in the *E. coli* S30 system and can have a dramatic effect on yield [16•].

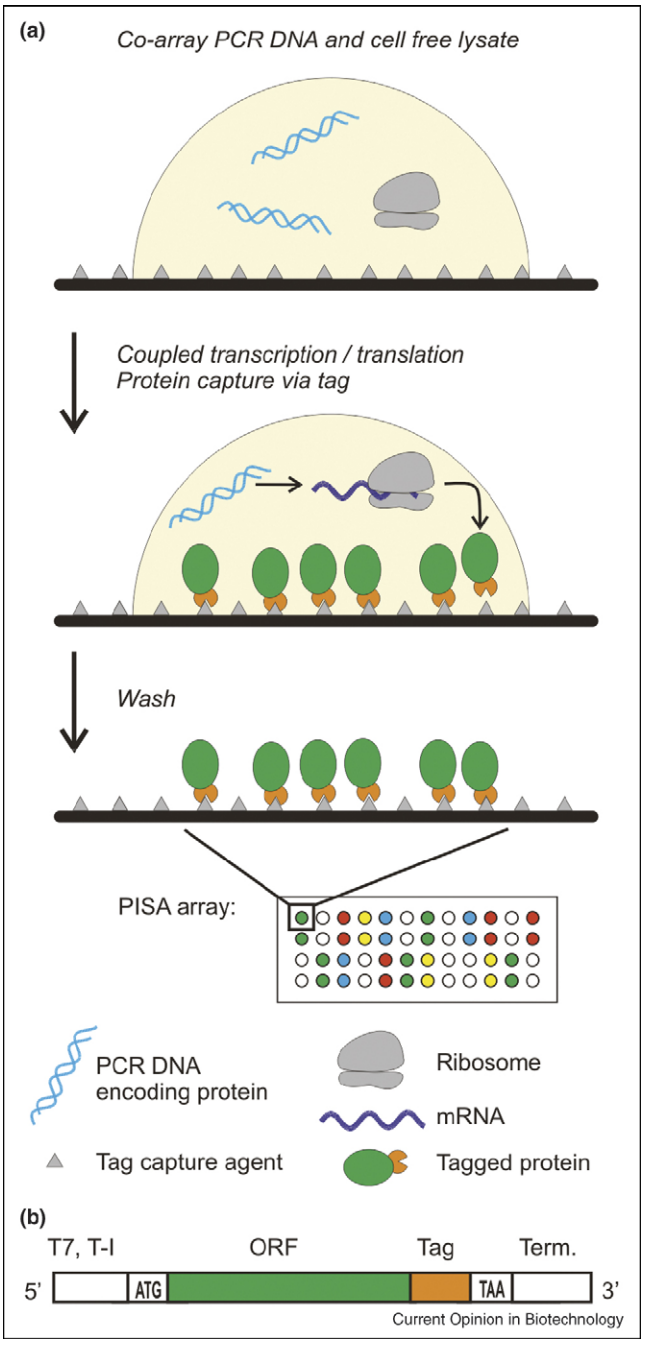
In situ methods for protein arraying

PISA

In the cell free protein array production method termed PISA (protein *in situ* array), proteins are made directly from DNA, either in solution or immobilised, and become attached as they are made to the array surface through recognition of a tag sequence (Figure 1a) [17]. DNA constructs encoding the proteins are generally produced by PCR, using primers designed from sequence (genomic) information, though plasmids can also be used. The constructs include the T7 promoter and sequences for *in vitro* transcription/translation initiation, together with an N- or C-terminal tag sequence for immobilisation (Figure 1b). The proteins are expressed in parallel *in vitro*, commonly utilising the rabbit reticulocyte or *E. coli* S30 systems to perform coupled transcription and translation. The key feature of the method is that the protein expression reaction is carried out on a surface which is precoated with an immobilising agent capable of binding the tag. Thus, after translation, proteins become fixed rapidly and specifically to the surface and unbound lysate material can be washed away. By multiplexing, it is possible to go in a single step from a set of PCR DNA fragments to a protein array. The concept has a number of advantages, in particular the ability to convert DNA sequence information quickly into immobilised functional proteins, without bacterial cloning or expression systems being involved or the need to purify the protein separately. In principle all that is needed in prior information is the DNA sequence in order to design PCR primers. Thereafter, starting from arrayed PCR DNA, the PISA procedure can be completed in three to four hours.

Our preferred method for *in situ* immobilisation is via hexahistidine $(His)_6$ sequences binding onto Ni-NTA (nickel nitrilo-triacetic acid) coated surfaces. As is well known, a single $(His)_6$ sequence tag will bind by chelation

Figure 1



In situ protein arraying by PISA. **(a)** Schematic of the PISA process. **(b)** PCR DNA constructs used, containing T7 promotor (T7), Shine-Dalgarno or Kozak sequences for prokaryotic or eukaryotic translation initiation, respectively (T-I), start codon ATG, gene open reading frame (ORF), tag-coding sequence (Tag), stop codon TAA and transcription termination region (Term.). The tag sequence may also be positioned at the N-terminus.

to the Ni^{2+} ion, with micromolar affinity. We have designed a stronger binding form in which two hexahistidine sequences are separated by an 11-amino acid spacer, double-$(His)_6$ [17,18••]. BIAcore analysis has confirmed the improved binding strength, showing a very slow dissociation and an affinity of at least 10-fold greater than a single-$(His)_6$ [18••]. Moreover, binding to Ni-NTA surfaces is sufficiently strong for the immobilised proteins to be re-used if required, after stripping off the detection molecules [17]. The double-$(His)_6$ sequence has been applied to immobilisation of non-purified antibody fragments on nano- and microarrays leading to enhanced molecular performance on the array [19•]. Ni-NTA coated surfaces are available as microtiter plates, magnetic agarose beads, BIAcore chips and glass slides. A useful feature is that the double-$(His)_6$ tag is detectable by anti-$(His)_6$ antibodies even after binding to Ni-NTA [18••].

The PISA method was originally demonstrated with a small set of proteins (antibody fragments, luciferase) immobilised in microtiter wells or onto magnetic beads, and with the DNA distributed in solution [17,20]. In this form, PISA is a macro method in that 25 μl of mixture (DNA plus cell free extract) were placed in each well. More recently it has been considerably reduced in scale (40 nl) and adapted to produce microarrays directly onto glass slides. In this 'mini-PISA' method, DNA is mixed with the cell free lysate system before spotting onto slides. After 2 h at 30 °C, the slides are washed and probed with a developing reagent (antibody, amplification) for fluorescent visualisation (He *et al.*, unpublished).

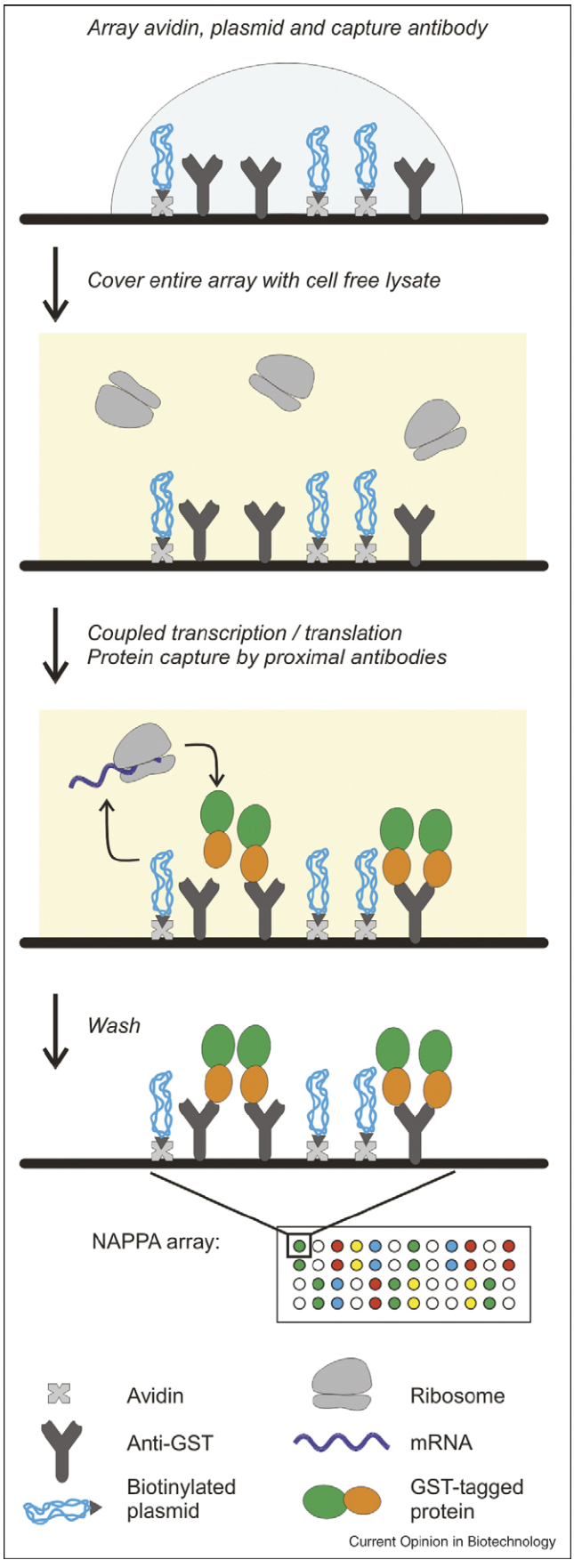
Hoheisel and colleagues [21••] have further developed a highly miniaturised, on-chip system based on a multiple spotting technique (MIST). The DNA template is deposited in a first spotting step (350 pl) and the cell free transcription and translation mixture (*E. coli*) transferred on top of the same spot in a second spotting step; rehydration allows the expression reactions to proceed in the physically separated sites. Using GFP as a model protein well-expressed in cell free systems, with a C-terminal single-$(His)_6$ tag, enough protein was produced from a plasmid to yield signals that were comparable to 300 μg/ml of spotted protein. Aminopropyltrimethoxysilane (APTES)- and Ni chelate-coated slides were used. With unpurified PCR products as templates, as little as 35 fg of PCR DNA (~ 22,500 molecules) was sufficient for the detectable expression of full-length GFP in subnanoliter volumes. They adapted the system to the high-throughput expression of protein libraries by designing a common primer pair for the introduction of the required T7 promoter and terminator and demonstrated *in situ* expression using 384 PCR products amplified from a human fetal brain library. By this means, high-density protein microarrays with (in principle) up to 13,000 spots per slide could be produced from a variety of different sources in an uncomplicated and inexpensive manner.

NAPPA

Transcription and translation from an immobilised (as opposed to in solution) DNA template is a further desirable development of on-chip technologies which would allow conversion of DNA arrays to protein arrays. This has been initiated by LaBaer and colleagues with the nucleic acid programmable protein array (NAPPA) system [22••,23]. Biotinylated plasmids encoding the proteins as GST fusions were printed onto an APTES slide, together with avidin and an anti-GST antibody acting as the protein capture entity. The plasmid DNA array was then covered with rabbit reticulocyte lysate in order to express the proteins, which became trapped by the antibody within each spot, thereby generating a protein array colocalising with the DNA (Figure 2). The procedure was shown to generate discrete protein spots with limited lateral spreading, though there was some evident variation in spot size and quality. In the first demonstration, eight expression plasmids encoding cell cycle proteins were immobilised at a density of 512 spots per slide (0.9 mm spacing) and an estimated 10 fmol (~675 pg) of protein was captured per spot (with a larger range of proteins, expression ranged from 4 to 29 fmol per spot). This was sufficient for functional studies and the report went on to analyse binary protein interactions by including a 'query' plasmid, encoding a potentially interacting protein, with a distinct detection tag, in the lysate. The potential interactor was thus made at the same time as the array and detected by an anti-tag antibody. (This method could pose a problem of standardisation, since different proteins express to different levels.) They were able to map 110 pairwise interactions, 63 previously undetected, among 29 human replication initiation proteins and others, using each in turn as a query probe against an array of the complete set [22••]. Since many known interactions were confirmed, this demonstrated the functional integrity of the *in situ* arrayed proteins.

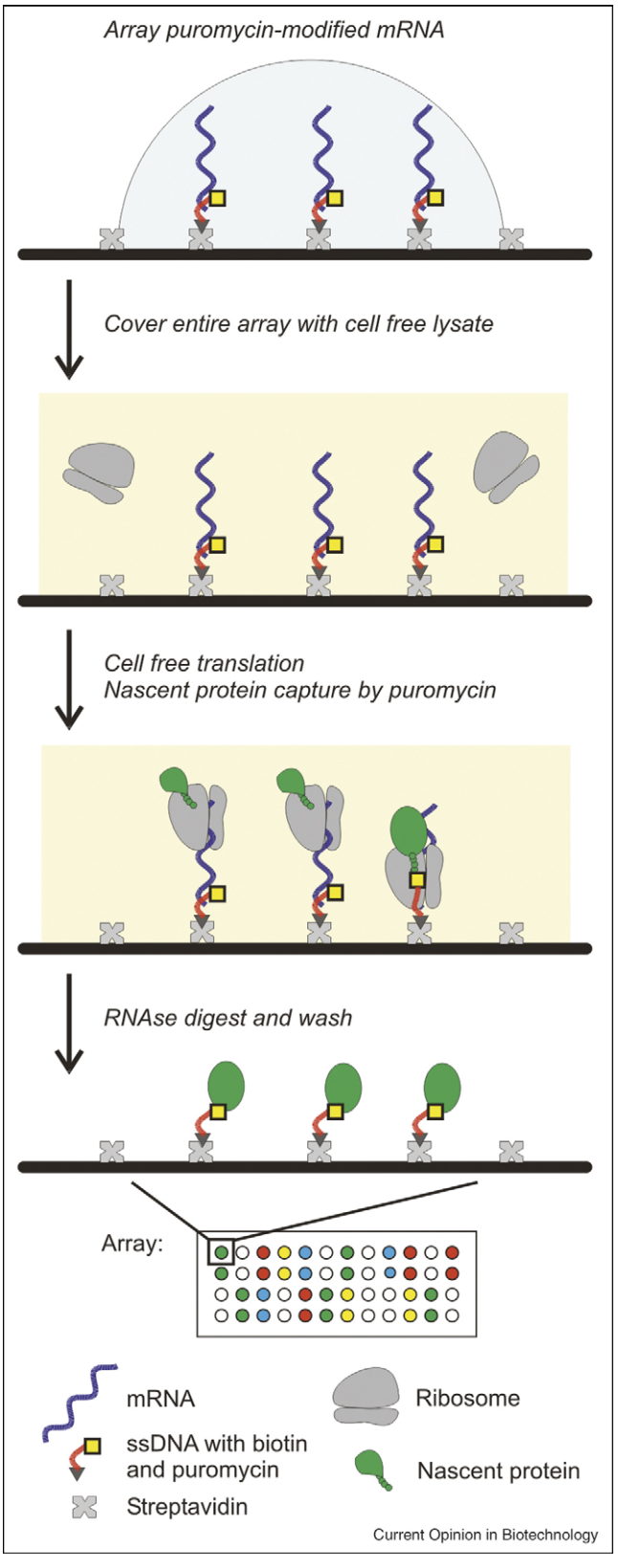
As with PISA, the advantages of NAPPA are that the need to purify the proteins, and concerns about protein stability during storage, are removed. There is also a clear advantage in having an immobilised DNA template, which can be stored, distributed and translated into a protein array as required. Nevertheless, NAPPA as described has some intrinsic drawbacks. For example, expression of the proteins from plasmids carrying the gene of interest as GST fusions requires cloning of cDNAs and immobilisation of the plasmid through a chemical biotinylation reaction designed to enable binding to the avidin surface while avoiding interference with transcription. Moreover, the technology does not generate a 'pure' protein array, but rather one in which proteins are colocalised in the same spots with both the plasmid

Figure 2



Schematic of protein arraying by NAPPA.

Figure 3



Schematic of protein arraying by puromycin-capture from mRNA arrays.

DNA (potentially removable with DNase) and the capture antibody. In addition, as in PISA, each DNA array only produces one protein array.

In situ puromycin-capture from mRNA arrays

In an adaptation of mRNA display [24], Tao and Zhu [25••] describe an ingenious strategy to fabricate protein chips through capture of nascent polypeptides by puromycin. Before arraying, PCR DNA is transcribed into mRNA *in vitro* and the 3′-end of the mRNA is hybridised with a ssDNA oligonucleotide modified with biotin and puromycin. Thus prepared, the mRNAs are arrayed on a streptavidin-coated slide and the array is covered with cell free lysate for *in situ* translation. The translating ribosomes stall upon reaching the RNA/DNA hybrid region and the release of the nascent protein is slowed. The puromycin attached to the DNA is then able to capture the nascent protein and immobilise it onto the slide surface. After the translation reaction, the mRNA is removed by digestion with RNase, leaving a pure protein array (Figure 3). This was exemplified with GST and two kinases, as well as two transcription factors which retained DNA binding function on the chip. An advantage is the sharply defined, nondiffused protein spots, because of the precise localisation of the puromycin with the mRNA and the 1:1 stoichiometry of mRNA and produced protein. On the other hand, extra manipulations are required to transcribe and modify the mRNA separately before printing and protein yields are limited by the amount of mRNA spotted. A similar approach, exemplified by solid phase immobilisation on beads, has been described by Biyani *et al.* [26•].

Conclusions

Existing *in situ* technologies allow protein arrays to be produced in a rapid and economical way, including those of proteins which are hard to express in cell based systems or not cloned. They score over conventional arraying in terms of time and cost and avoid lengthy cloning and protein purification. As PCR technologies, they also allow individual protein fragments or domains to be generated rapidly from sequence information. They are amenable to different templates, including PCR products, plasmids and mRNA. By producing the proteins on the chip just before use, effects of storage on function are minimised. By designing PCR primers from genomic sequence data to make the constructs, they provide a direct route to functional proteome analysis.

In situ protein arraying is showing its promise and performance in the systems described above. A new development in progress is a 'DNA array to protein array' (DAPA) procedure in which an immobilised DNA array is the re-usable template for printing multiple copies of protein arrays on separate slides as and when required. A prototype system achieving this objective has been developed in our laboratory (He et al., *Nat Methods*, in press).

Acknowledgements

Work in the authors' laboratory at The Babraham Institute is supported by the Biotechnology and Biological Sciences Research Council and by the European Commission 6th Framework Integrated Project MolTools.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Bertone P, Snyder M: **Review: advances in functional protein microarray technology**. *FEBS J* 2005, **272**:5400-5411.

2. Barry R, Soloviev M: **Review: quantitative protein profiling using protein arrays**. *Proteomics* 2004, **4**:3717-3726.

3. Michaud GA, Bangham R, Salciu M, Predki PF: **Functional protein microarrays for pathway mapping**. *DDT: Targets* 2004, **3**:238-245.

4. Stevens RC: **Design of high-throughput methods of protein production for structural biology**. *Struct Fold Des* 2000, **8**:R177-R185.

5. • Endoh T, Kanai T, Sato YT, Liu DV, Yoshikawa K, Atomi H, Imanaka T: **Cell-free protein synthesis at high temperatures using the lysate of a hyperthermophile**. *J Biotechnol* 2006, **126**:186-195.

In contrast to systems operating at moderate temperatures, this paper reports preparation of a cell-free lysate operative at high temperatures, from the hyperthermophilic archaeon, *Thermococcus kodakaraensis*. With this *in vitro* translation system, active chitinase was produced at temperatures between 40 and 80 °C (optimum 65 °C). A system for coupled transcription and translation at high temperatures, using a combination of *T. kodakaraensis* lysate and thermostable T7 RNA polymerase, was also developed.

6. • Mikami S, Kobayashi T, Yokoyama S, Imataka H: **A hybridoma-based *in vitro* translation system that efficiently synthesizes glycoproteins**. *J Biotechnol* 2006, **127**:65-78.

The authors developed an efficient system derived from a monoclonal antibody-producing hybridoma for cell-free synthesis of glycoproteins. It was successfully used to produce a biologically active, multichain glycoprotein (human choriogonadotropin).

7. Keller C, Hyrien O, Knippers R, Krude T: **Site-specific and temporally controlled initiation of DNA replication in a human cell-free system**. *Nucleic Acids Res* 2001, **30**:2114-2123.

8. Landsverk HB, Hakelien AM, Kuntziger T, Robl JM, Skalhegg BS, Collas P: **Reprogrammed gene expression in a somatic cell-free extract**. *EMBO Rep* 2002, **3**:384-389.

9. • Langlais C, Guilleaume B, Wermke N, Scheuermann T, Ebert L, Labaer J, Korn B: **A systematic approach for testing expression of human full-length proteins in cell-free expression systems**. *BMC Biotechnol* 2007, **7**:64.

The authors compared cell free expression in *E. coli* and wheat germ systems for a subset of 87 full-length human proteins which, out of 960 evaluated in a pre-screen, showed no product in *E. coli in vivo*. Through appropriate optimisations, a success rate of 93% was achieved *in vitro*. Wheat germ expression with a two compartment system emerged as the method of choice for high yield of soluble proteins.

10. Shimizu Y, Kanamori T, Ueda T: **Protein synthesis by pure translation systems**. *Methods* 2005, **36**:299-304.

11. • Oyama R, Takashima H, Yonezawa M, Doi N, Miyamoto-Sato E, Kinjo M, Yanagawa H: **Protein–protein interaction analysis by C-terminally specific fluorescence labeling and fluorescence cross-correlation spectroscopy**. *Nucleic Acids Res* 2006, **4**:e102.

The authors present a cell free system containing diluted puromycin derivatised with biotin, together with either Cy5 or RhG fluorophores. The expressed labelled proteins are purified via the biotin to provide fluorophore-labeled proteins for interaction studies by FCCS.

12. Ozawa K, Wu PS, Dixon NE, Otting G: **N-Labelled proteins by cell-free protein synthesis**. *FEBS J* 2006, **273**:4154-4159.

13. Angenendt P, Nyarsik L, Szaflarski W, Glökler J, Nierhaus KH, Lehrach H, Cahill DJ, Lueking A: **Cell-free protein expression and functional assay in nanowell chip format**. *Anal Chem* 2004, **76**:1844-1849.

14. Spirin A: **High-throughput cell-free systems for synthesis of functionally active proteins**. *Trends Biotechnol* 2004, **22**:538-545.

15. Calhoun K, Swartz JR: **An economical method for cell-free protein synthesis using glucose and nucleoside monophosphates**. *Biotechnol Prog* 2005, **21**:1146-1153.

16. Palmer E, Liu H, Khan F, Taussig MJ, He M: **Enhanced cell-free**
• **protein expression by fusion with immunoglobulin C_κ domain**. *Protein Sci* 2006, **15**:2842-2846.
A novel strategy is reported for promoting synthesis of 'difficult-to-express' proteins in the *E. coli* cell free system, through C-terminal fusion of the constant domain of human immunoglobulin κ light chain (C_κ). Four proteins (two single-chain antibody fragments, a GTP-binding protein and FK506 binding protein) all showed highly enhanced expression as C_κ fusions.

17. He M, Taussig MJ: **Single step generation of protein arrays from DNA by cell-free expression and *in situ* immobilisation (PISA method)**. *Nucleic Acids Res* 2001, **29**:e73.

18. Khan F, He M, Taussig MJ: **A double-His tag with high affinity**
•• **binding for protein immobilisation, purification, and detection on Ni-NTA surfaces**. *Anal Chem* 2006, **78**:3072-3079.
This paper characterises the double-$(His)_6$ tag, in which two $(His)_6$ sequences are linked by an 11-amino acid spacer. BIAcore analysis showed binding to Ni-NTA with 10-times slower dissociation than single-$(His)_6$. The double-$(His)_6$ tag was also detected with greater sensitivity than single-$(His)_6$ by anti-$(His)_6$ antibodies and a variety of Ni-NTA conjugates in western blotting, ELISA and protein arrays. It was applied in single-step protein immobilisation and purification from crude mixtures.

19. Steinhauer C, Wingren C, Khan F, He M, Taussig MJ,
• Borrebaeck CA: **Improved affinity coupling for antibody microarrays: engineering of double-$(His)_6$-tagged single framework recombinant antibody fragments**. *Proteomics* 2006, **6**:4227-4234.
The authors examined construction of arrays of tagged human recombinant single-chain antibody fragments. Substitution of the standard single-histidine $(His)_6$-tag with the consecutive double-$(His)_6$-tag (ref. 18 above), significantly improved binding to Ni-NTA-coated surfaces. This enabled non-purified probes to be applied directly to the surface, eliminating pre-purification before the immobilisation, and created better long-term functional on-chip stability.

20. He M, Taussig MJ: **DiscernArray technology: a cell-free method for the generation of protein arrays from PCR DNA**. *J Immunol Methods* 2003, **274**:265-270.

21. Angenendt P, Kreutzberger J, Glokler J, Hoheisel JD: **Generation**
•• **of high density protein microarrays by cell-free *in situ* expression of unpurified PCR products**. *Mol Cell Proteomics* 2006, **5**:1658-1666.
This describes *in situ* production of protein microarrays by a multiple spotting technique, first spotting PCR DNA and then the cell free system on top, and allowing cell free transcription and translation to occur in each spot. Protein yields were analysed. The method was adapted to high throughput library expression and demonstrated with 384 randomly chosen clones. Protein microarrays with up to 13,000 spots per slide can be generated *in situ*.

22. Ramachandran N, Hainsworth E, Bhullar B, Eisenstein S, Rosen B,
•• Lau AY, Walter JC, LaBaer J: **Self-assembling protein mircoarrays**. *Science* 2004, **305**:86-90.
This paper describes the NAPPA technique for converting DNA arrays into protein arrays, in which immobilised plasmids encoding fusion proteins are used as templates for on-chip protein synthesis by cell free transcription and translation. The proteins are captured by antibody, cospotted with the plasmid and directed against the fusion tag. As well as demonstrating protein microarray production, pairwise interaction studies with 29 human replication initiation proteins were performed.

23. Ramachandran N, Hainsworth E, Demirkan G, LaBaer J: **On-chip protein synthesis for making microarrays**. *Methods Mol Biol* 2006, **328**:1-14.

24. Roberts RW, Szostak JW: **RNA-peptide fusions for the *in vitro* selection of peptides and proteins**. *Proc Natl Acad Sci U S A* 1997, **94**:12297-12302.

25. Tao S-C, Zhu H: **Protein chip fabrication by capture of nascent**
•• **polypeptides**. *Nat Biotechnol* 2006, **24**:1253-1254.
Thie authors performed on-chip production of protein microarrays from modified *in vitro* transcribed mRNA immobilised on the chip surface. Cell free translation was performed on the slide, during which a puromycin-oligonucleotide, hybridised to the mRNA, displaced the full length nascent protein, both linking it to the mRNA and trapping it on the surface. The mRNA was then removed enzymatically. The method produced arrays with well-defined spots of GST and functional transcription factors.

26. Biyani M, Husimi Y, Nemoto N: **Solid-phase translation and**
• **RNA-protein fusion: a novel approach for folding quality control and direct immobilisation of proteins using anchored mRNA**. *Nucleic Acids Res* 2006, **34**:e140.
A similar, 'RNA array to protein array' approach to that of ref. [25] using modified mRNA-protein fusion and exemplified with GFP and a functional enzyme, aldehyde reductase.