Chem Soc Rev

Cite this: *Chem. Soc. Rev.*, 2011, **40**, 2131–2145

www.rsc.org/csr

**TUTORIAL REVIEW**

# Studying protein–protein interactions using peptide arrays†

**Chen Katz,[a] Liron Levy-Beladev,[a] Shahar Rotem-Bamberger,[a] Tiago Rito,‡[b] Stefan G. D. Rüdiger[b] and Assaf Friedler*[a]**

***Received 30th June 2010***
**DOI: 10.1039/c0cs00029a**

Screening of arrays and libraries of compounds is well-established as a high-throughput method for detecting and analyzing interactions in both biological and chemical systems. Arrays and libraries can be composed from various types of molecules, ranging from small organic compounds to DNA, proteins and peptides. The applications of libraries for detecting and characterizing biological interactions are wide and diverse, including for example epitope mapping, carbohydrate arrays, enzyme binding and protein–protein interactions. Here, we will focus on the use of peptide arrays to study protein–protein interactions. Characterization of protein–protein interactions is crucial for understanding cell functionality. Using peptides, it is possible to map the precise binding sites in such complexes. Peptide array libraries usually contain partly overlapping peptides derived from the sequence of one protein from the complex of interest. The peptides are attached to a solid support using various techniques such as SPOT-synthesis and photolithography. Then, the array is incubated with the partner protein from the complex of interest. Finally, the detection of the protein-bound peptides is carried out by using immunodetection assays. Peptide array screening is semi-quantitative, and quantitative studies with selected peptides in solution are required to validate and complement the screening results. These studies can improve our fundamental understanding of cellular processes by characterizing amino acid patterns of protein–protein interactions, which may even develop into prediction algorithms. The binding peptides can then serve as a basis for the design of drugs that inhibit or activate the target protein–protein interactions. In the current review, we will introduce the recent work on this subject performed in our and in other laboratories. We will discuss the applications, advantages and disadvantages of using peptide arrays as a tool to study protein–protein interactions.

[a] *Institute of Chemistry, The Hebrew University of Jerusalem, Safra Campus, Givat Ram, Jerusalem 91904, Israel. E-mail: assaf@chem.ch.huji.ac.il; Fax: +972 2-6585345; Tel: +972 2-6585746*
[b] *Cellular Protein Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands*
† Electronic supplementary information (ESI) available: From macro to micro - a comparative study. See DOI: 10.1039/c0cs00029a
‡ Present address: Department of Statistics, University of Oxford, OX1 3TG, UK.

## Origins of the peptide array technique

The concept of parallel synthesis of multiple components on a solid support was introduced in the early eighties by both Ronald Frank and Mario Geysen. In 1983, Frank *et al.* demonstrated the first parallel synthesis of oligonucleotide chains on cellulose discs packed in a column.[1] In 1984, Geysen *et al.* demonstrated the first parallel synthesis of

**Chen Katz**

*Chen Katz obtained her BSc degree in Chemistry and Biology in 2004 from the Hebrew University of Jerusalem. She is currently pursuing a PhD in Chemistry under the supervision of Prof. Assaf Friedler at the Hebrew University of Jerusalem. Her research is focused on using peptides to study protein–protein interactions in apoptotic pathways, for example the interaction of ASPP2 with Bcl2. She used peptide arrays to identify the binding regions in both proteins and design inhibitory peptides, which will serve as a basis for anti-cancer drug design.*

**Liron Levy-Beladev**

*Liron Levy-Beladev obtained a BSc degree in Chemistry and Biology in 2004 from the Hebrew University of Jerusalem. She obtained an MSc degree in Chemistry in 2006 from the Hebrew University of Jerusalem. She is currently pursuing a PhD in Chemistry under the supervision of Prof. Assaf Friedler at the Hebrew University of Jerusalem. Her research is focused on using peptides to study protein interactions. For example, she studied the interaction of CFTR with Hsp90. Peptide arrays were used to identify the binding regions of Hsp90 within CFTR.*

hundreds of peptides on plastic pins.[2] In 1992, Frank extended his approach and established the method that later became a milestone in this field, the SPOT™ synthesis. Using this method, peptides are synthesized simultaneously on a solid support (in this case cellulose sheets) by simply dispensing small droplets onto the planar surface of a porous membrane, and generating an open reactor for chemical synthesis attached to the membrane support.[3] Around the same time, another breakthrough method for preparing peptide arrays was reported: the synthesis of peptide arrays using photolithography by Fodor *et al.*[4] The spot technology was so "easy to make" and "easy to detect" method, and it became favourite and widely used. The new technology has encouraged the development of numerous arrays from various types: peptide arrays, protein arrays, combinatorial chemical libraries, *etc*. The arrays that made progress in the past two decades were the DNA/RNA arrays. DNA arrays are used for gene expression analysis, genotyping of individuals, genotyping point mutations, single point mutations (SNPs), short tandem repeats (STRs), and numerous other applications (all reviewed in ref. 5). Overall, DNA arrays became a standard research tool in the genomic field. Yet the information one can obtain from DNA and RNA arrays is limited, since they do not provide information about the proteins mechanism of action.

In contrast to DNA/RNA arrays, the proteomic field did not have the same momentum in using arrays for a long time. Protein arrays did evolve but they are much more difficult to produce. Unlike DNA/RNA, proteins are biochemically diverse and their functionalities extremely depend on their exposed surface, correct folding and conformation. Proteins are also much more sensitive to many conditions and chemicals, so the detection of protein binding is also more complicated. When seeking for therapeutic applications and influencing cellular pathways, it is required to work at the

**Shahar Rotem-Bamberger**

*Shahar Rotem-Bamberger obtained a BSc degree in Biology in 2004 from the Hebrew University of Jerusalem. She obtained a MSc degree in Biochemistry in 2006 from the Hebrew University of Jerusalem. She is currently pursuing a PhD in Biochemistry under the supervision of Prof. Assaf Friedler, at the Hebrew University of Jerusalem. Her research is focused on the intra- and inter-molecular interactions of the ASPP2 protein.*

**Tiago Rito**

*Tiago Rito graduated in Biochemistry from the University of Coimbra, Portugal. Part of this degree included an ERASMUS-funded internship in Stefan Rüdiger's group at the Bijvoet Center, Utrecht University, The Netherlands, where he continued to work for one year as a researcher. Currently, he is reading for a PhD in Systems Biology at the University of Oxford, UK.*

**Stefan G. D. Rüdiger**

*Stefan Rüdiger graduated in Chemistry at Heidelberg University, Germany. Afterwards he joined Bernd Bukau's laboratory and in 2000 he obtained the PhD at Freiburg University, Germany. After a postdoc in Alan Fersht's laboratory at the Centre for Protein Engineering, Cambridge University and MRC Centre, Cambridge, UK, he became in 2004 faculty member at the Bijvoet Center of Utrecht University, The Netherlands. In 2006 he received tenure and became Marie Curie Excellence Team Leader and High Potential of Utrecht University. His research group focuses on the Hsp90 chaperone machine and the role of disordered sequences in protein folding.*

**Assaf Friedler**

*Assaf Friedler did his undergraduate and PhD studies in chemistry at the Hebrew University of Jerusalem. His PhD studies were performed under the supervision of Prof. Chaim Gilon. Between 2000–2004 he did his postdoctoral research at the MRC centre for protein engineering, Cambridge, UK, in the lab of Prof. Sir Alan Fersht. Since 2004 he runs his independent research group at the Institute of Chemistry in the Hebrew University of Jerusalem, Israel (http://chem.ch.huji.ac.il/~assaf/). In 2010 he received tenure and promoted to Professor. Prof. Friedler is currently the head of the School of Chemistry at the Hebrew University. His major research interests are using peptides to study protein–protein interactions in health and disease, and developing peptides as drugs that modulate these interactions. Prof. Friedler recently won a starting grant from the ERC (European Research Council) as well as the outstanding young scientist prize by the Israeli Chemical Society.*

 

protein level, since ultimately proteins are the main mediators in cellular processes. Therefore, during the last decade scientists have tried to overcome the relative complexity of the protein world by adopting the linear thinking of the DNA/RNA world. This led to the evolvement of peptide arrays to explore protein–protein interactions.

Peptide arrays resemble nucleotide arrays in the sense that they are relatively easy and inexpensive to prepare at high purity. In addition, they enable to "zoom in" and narrow down the huge interaction interface to a specific binding site. Peptides have several advantages in protein–protein interactions research: (i) Peptides are easy to make and purify by chemical synthesis, unlike recombinant proteins. (ii) Peptide synthesis avoids the limitations of using only the 20 naturally occurring amino acids. It allows using non-natural and/or modified amino acids as building blocks in chemical synthesis, enabling for example the studies of post-translational modifications. (iii) Probes or linker groups such as biotin, fluorescein or His-Tag can be easily introduced into peptides at any required position, and be used for protein detection or immobilization that is often required in binding assays. (iv) The peptides resulting from the screening can sometimes act as inhibitors or activators of the target proteins, and thus serve for diagnostic or therapeutic purposes. (v) The resulting binding peptides can serve as a basis for modeling of protein–protein interactions[6] (Fig. 1). The peptide arrays described in the current review represent only one type of combinatorial peptide library. There are many other types of combinatorial peptide libraries, for instance: biological peptide libraries, such as phage-display,[7] and synthetic combinatorial libraries such as those generated using the pin technology.[2] Another example is the synthesis of numerous peptides with random sequences based on the split-and-mix[8] strategy using the "one bead one peptide" approach.[9] Other known strategies in the combinatorial peptide libraries field include the "tea bags"[10] approach, and the positional scanning approach.[11] Each strategy has its own advantages and disadvantages. This was reviewed extensively elsewhere and will not be discussed here in detail.[12] The current review will examine the use of peptide arrays to study protein–protein interactions. We will only briefly introduce the synthetic schemes, since this issue was extensively reviewed in ref. 13. We will address the issue of choosing between macro-arrays and micro-arrays and among the commercial arrays available today. Guidelines for peptide array design and for how to technically perform the screening assay will be presented, and data analysis options will be discussed. Finally, we will highlight some recent and innovative developments in the peptide array field.

## Peptide array synthesis

Currently, there are two common ways to prepare peptide arrays: (i) synthesizing functionalized peptides in advance and then covalently attaching them to the support; (ii) synthesizing the peptides sequentially directly on the solid support, usually by using SPOT synthesis or photolithography.

Using the first method, attachment of the peptide to the solid support can be done by several methods for peptide immobilization. For example the use of cysteine-terminated peptide on glyoxylyl glass surfaces,[14] or surfaces presenting functional groups such as bromomethylketone or disulfides[15,16] (for more immobilization methods see ref. 17). This method is expensive and time consuming. It requires peptide purification and high amounts of solvents and reagents relative to the SPOT synthesis. Yet the advantage is clear: the purification of the peptide reduces false-positive results of synthesis by-products, and allows the immobilization of the same exact concentration of pure peptides to each spot the array surface. This method is more suitable for arrays with a relatively small number of peptides or when several copies of the same array are required. The second method of synthesizing peptides sequentially directly on a solid support is much more robust, and also very economical regarding the reagents and solvents. This method is more suited for rapid examination of many sequences, which is currently the most common use for peptide arrays.

The SPOT synthesis technique is the most common and frequently used technique for synthesizing peptides sequentially directly on a solid support. The basic principle of the SPOT synthesis technique is to use the circle that forms when a droplet is dispensed on a planar surface, as a reaction vessel. The circle of the droplet creates limited boundaries that can be addressed individually by manual or automated delivery of the corresponding reagents. Multiple separate spots can be arranged in this way to an organized array. The reduction in solvent volume is enormous because the surfaces of the droplets spots are small by nature, and the synthetic parts that are similar for all the peptides can be done simultaneously by rinsing the whole surface. The size of the spots is determined mostly by the volume of the solvent dispensed, the absorption capacity of the membrane and both membrane/solvent surface tension properties.[18] The most obvious drawback of this process is that peptides do not undergo purification, and it is impossible to determine their identity and purity. If synthesis fails or has low yield this will automatically leads to false negative results. The issue of synthesis and purity is under debate. There are several reports on this topic[19–22] but the results are not conclusive and there is a broad range of reported purities for such peptide arrays, ranging between 50–92% for the tested peptides. The reason is that every peptide is synthesized with a different yield, depending on the peptide sequence, hydrophobicity, length and conformation. Hence the level of purity cannot be predicted. The rule of thumb accepted for the SPOT synthesis method is using relatively short peptides (up to 15 AA). Wenschuh *et al.* found that short peptides synthesized by the SPOT technique had the same purity level as short peptides synthesized using standard solid phase synthesis.[23] This is in good agreement with reports from the commercial company JPT (Jerini Peptide Technologies GmbH), that reports typical purity of $>70\%$ for average 6–15 AA.[24] Yet, the fact that different peptides can be present at different yields is very problematic, especially in terms of results quantification. Sometimes peptides that are difficult to synthesize will have low yield on the array and therefore, even if they are strong binders, due to their low concentration we will see them as very weak dots.

The peptides on the arrays are synthesized using Fmoc solid phase synthesis.[25] Hydroxy groups on the solid support are used for coupling a reactive amino function (commonly Fmoc

β-alanine) in an ester-type bond. The available spot synthesis protocols slightly vary in the procedures for amino acid activation, coupling steps or the use of different solvents. For more detailed protocols see recent reviews.[3,13,26,27] Currently,

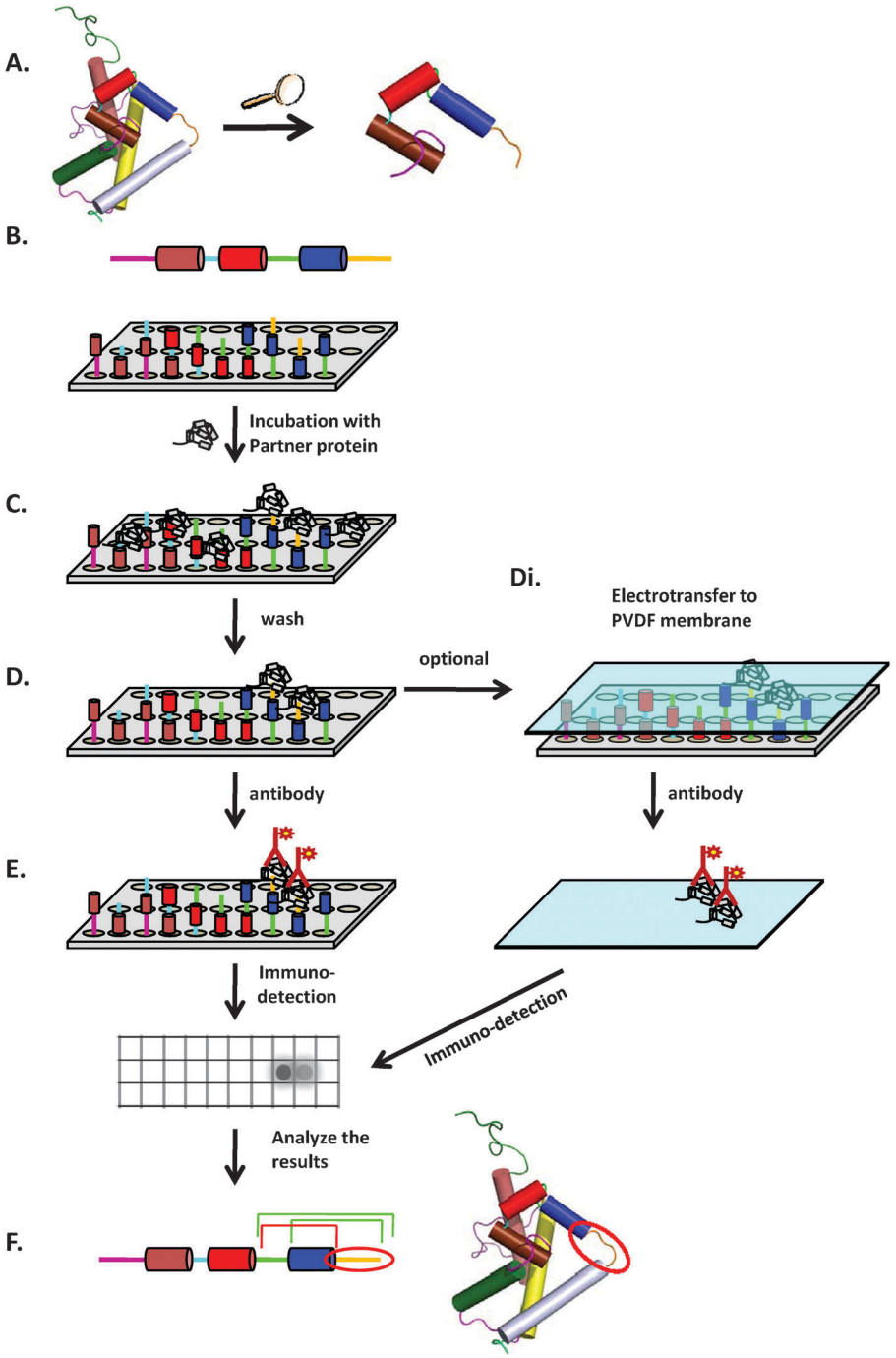


**Fig. 1** A general scheme of peptide array design and screening. (A, B) An example for peptides array design; zooming in on a small structured region of a protein and designing multiple partly overlapping peptides while maintaining the secondary structures (helices and loops) intact. In the limit of length (<25 mer), each secondary element should overlap with the next secondary element, and also with the previous element. For example, one peptide will contain the first purple loop together with the first brown helix, and the second peptide will contain the same brown helix along with the next cyan loop (B). (C–E) The basic scanning assay includes incubation of the binding protein with the peptides array (C), followed by a few washing steps (D), in some cases (mostly for macro-arrays) an additional step takes place before the detection: the binding protein is electrotransferred to a secondary membrane (mostly nitrocellulose or PVDF) (Di), to allow the reuse of the same array several times. Detection of the binding protein is done using specific antibodies, usually conjugated to an enzyme such as peroxidase or alkaline phosphatase, which can catalyse a color reaction. Each dark spot represents binding of the protein to a specific peptide (E). (F) Analyzing the results; marking on the secondary structures all the observed binding peptides (green) and the non-binding peptides (red), in order to minimize the binding site to the shortest binding element, common to all the observed partly overlapping peptides.

 

most of the above process can be carried out automatically using peptide synthesizers and spotting robots. This leads to a smaller spot size and thus a smaller surface for the entire array, along with higher concentration of peptide per spot. For example, JPT, where most of the process is automated, reports that a typical spot has a diameter of approx. 2–3 mm on each cellulose peptide array and that a typical synthesis is expected to yield between 5–10 nmol (6–12 μg for an average 10-mer peptide).

Choosing the appropriate support is also important. The aim of the array screening is to determine binding between peptides and proteins. Thus the support should be chemically compatible with the peptide anchoring and synthesis, stable towards acidic/basic environment, and also suitable for biological screening processes such as electroblotting. The support also should be as flat as possible, homogeneous, and allow easy protein access. Cellulose filter sheets have been proven to be well-matched for use in peptide arrays. Cellulose is inexpensive, hydrophilic, easy to control during synthesis, and stable under a wide range of reaction conditions.[28]

The second most frequently used method for sequential peptide synthesis on a solid support is photolithography. Peptide synthesis by photolithography on a glass surface was first reported by Fodor *et al.*[4] in 1991. Since then many modifications were done to make the method less complicated and more efficient (reviewed in ref. 29). The original synthesis was performed using special photo-labile protected amino acids as building blocks and light irradiation through a photomask. However, the photo-labile amino acids and the photomask were time consuming to use and expensive to make, and their reaction efficiency was insufficient. Techniques were evolved for using photogenerated reagents (PGR) that enable the use of conventional compounds for synthesis such as *t*-Boc amino acid. The most common of the photogenerated reagents are the photogenerated acids (PGA). These compounds form an acid upon light irradiation, creating an acidic environment essential for the deprotection of the *t*-Boc group and subsequent formation of amide bonds.[29] This technology has been commercialized and further developed by LC Sciences company.[30]

## The binding assay

Scanning peptide arrays assay is an extremely useful technique. It is technically quite simple, while generating a large mass of results in a relatively short time. The basic protocol for peptide array screening includes several steps (Fig. 1): wetting of the dry peptide array followed by washing for the disposal of all the preservative materials; incubation of the binding protein with the peptide array, followed by a few washing steps and finally detection of the protein-bound peptides. Small modifications are usually performed in each step in order to optimize the assay for each particular biological system.

The first step, which includes wetting of the dry peptides array and disposal of all preservative materials, consists of washing steps whose particular nature depends on the array type and the materials the array is made of. For example, cellulose macro-arrays (supplied by JPT) are recommended to be soaked for 10 minutes in methanol, rinsed one more minute in ethanol, and then rinsed three more times with TBS, (Tris Buffered Saline: 50 mM Tris Base, 150 mM NaCl, pH 7.5) for ten minutes each. Peptides micro-arrays on glass slides (supplied by INTAVIS) are recommended to be washed 1–4 hours with TBST (TBS-Tween 20: 50 mM Tris Base, 150 mM NaCl, 0.05% Tween 20, pH 7.5).

The second step, including incubation of the binding partner with the peptides array, is the most important one for the success of the experiment. Incubation can be done with different types of binding partner solutions such as an antiserum, cell extract, recombinant protein, peptide, *etc*. The incubation solution usually contains blocking agents, in order to prevent non-specific binding, although in some cases these were not used since they could mask weak binding (see for example ref. 31). There are several commonly used blocking agents such as: (1) 2–5% BSA in PBST or TBST; (2) 2–5% skimmed milk powder in PBST or TBST; (3) 2–5% sucrose in PBST or TBST; (4) MBS (25 mM MES, 150 mM NaCl, pH 6.5), with 50% (v/v) horse serum. The blocking conditions are critical for the success of the experiment and should be optimized for each biological system since different proteins react differently to each solution. Problems may occur when the partner protein binds one of the blocking ingredients or precipitates in it. If this occurs, several blocking solutions should be tested to find the most suited for the protein used. In our lab, blocking solution number (2) works well in most cases. Two main parameters that can be modified when optimizing the assay are the percentage of blocker (BSA, skimmed milk, sucrose, *etc*.) in the solution and the partner protein concentration. These parameters should be modified according to the strength of the interaction. From our experience, when the binding affinity of the interaction is low, the concentration of the partner protein should be increased and the blocker percentage should be reduced in order to avoid the loss of detection of peptides that bind weakly (false negatives). For interactions with high binding affinity, the concentration of the partner protein should be relatively low and the percentage of blocker should be increased. Another variable that can be altered in order to obtain better or more indicative results is the total ionic strength of the incubation solution. In principle, the total ionic strength should be kept close to physiological (150 mM). However, if the binding (between the partner protein and the peptides on the array) has an electrostatic nature, lowering the total ionic strength will result in tighter binding, enabling the detection of weaker interactions. Another important ingredient that should be added to the incubation solution is a reducing agent such as DTT or βME, which should be added in cases where cysteine residues are present in the binding partner protein. The reducing conditions are essential in cases when either the formation of disulfide bonds between the protein and the peptides on the array may lead to false results, or when a reducing environment is essential for stability of the protein. It should be noted that not all peptide arrays are suited for these conditions. For example, JPT company offers two types of arrays: arrays to which peptides are linked *via* β-alanine, and arrays to which peptides are linked *via* PEG. The PEG linker is stable under reducing conditions and is recommended for use when the buffers contain DTT or similar reagents.

After incubation, the excess partner protein should be washed from the peptide array. If the interaction is of low

affinity, it is not recommended to perform many washing steps. This is because each washing step can slightly shift the binding equilibrium towards the unbound state, resulting in abolishing weak binding. This is true especially in the case of micro-array slides, where the detection is performed directly on the peptide array.

In the case of cellulose sheets (macro-arrays) it is very common to add another step of electrostatic transfer of the binding protein (after the incubation) to PVDF membrane. The PVDF membrane is being replaced several times within one experiment to improve the signal to noise ratio. This may also be an effective method to remove unspecific background binding, since affinity to the cellulose support is often lower than to specific peptides, resulting in faster transfer. This process allows the reuse of the same array several times. The analysis of several successive membranes also allows averaging of the results from several membranes, which leads to higher data quality. It also allows judging the amount of protein left on the peptide array at the end of the experiment, which is important for re-using of the array. More detailed information about peptide array transfer can be found at ref. 31 and 32. The transfer has the advantage that after the binding occurs, the protein is immobilized on the PVDF membrane for the detection step. Thus, the washings and detection steps do not shift the binding equilibrium towards complex dissociation. This is especially important for cases of weak binding, where in assays performed directly on the array the washings can lead to complete dissociation of the complex due to shift of the binding equilibrium.

The last step of the assay is detection of the peptide–protein binding. Detection is performed mostly by labeled probe methods, such as fluorescence, chemiluminescence, electro-chemiluminescence and radioactivity detection. It is possible to use either a primary antibody against the partner protein itself combined with a secondary labeled antibody, or a specific antibody against a fused tag such as His-, Biotin-, GST, *etc.* Antibodies against the tags are usually HRP conjugated (preferable for saving the need of a second incubation and a few washing steps). The most common detection method to read and assess the signal intensities is chemiluminescence, since it is inexpensive, highly sensitive and widespread and does not require any special equipment. In our lab, chemi-luminescence works well for most cases, and fluorescence was used for very low affinity interactions since it is a more sensitive method. The clear disadvantage of the antibody-based detection systems is the multiple steps involved, unspecific binding and cross-reactions of the peptide with the antibodies. Another disadvantage one should take into account is when testing oligomerization or intramolecular interactions within proteins: the antibody against the protein may also recognize the peptides on the array, which are derived from the same protein. This disadvantage can be overcome by using a tagged protein and an antibody directed against the tag, in that case one had to ensure that the tag does not influence specificity. Another option to overcome this problem is by the electroblotting technique, which uncouples the peptide array from the antibody-based detection assay. In addition, incubation of the peptide array with the detection agent (antibody) alone should always be performed as a control experiment. Another way of solving this problem is using a directly labeled partner protein. Direct labeling of the protein can be achieved by using a fluorescent dye, radioactive isotopes or by addition of detectable enzyme such as peroxidase or phosphatase. The main drawback of the direct labeling method is that the detection relies on the assumption that the protein is uniformly labeled. This is difficult to ensure, due to the complexity and lack of specificity of labeling process. Moreover, a study on protein arrays that examined six different detection strategies (including fluorescent-labeled, radioactively labeled, and anti-body-based detection systems) found that antibody-based detection provided the best signal to noise ratio.[33]

## From macro- to micro-arrays

Micro-arrays and macro-arrays are the terms that are used to differentiate between array sizes and/or the number of spots on the support. The term "macro-array" is used for supports with larger area and with a smaller number of spots per array. The term "micro-array" is usually used for smaller support area and a higher number of spots. Most macro-arrays are produced on various types of membrane sheets (usually cellulose) while micro-arrays are usually produced on standard microscope glass slides due to their smaller and more convenient format.[34]

Three main reasons were behind the advance from macro- to micro-arrays: miniaturization, higher densities and a large number of array copies. Making smaller arrays is essential to reduce the amount of reagents required for the binding assays. Most importantly, miniaturization reduces the amounts of the proteins and antibodies used, which are expensive and sometimes very difficult to produce. While the typical macro-array has densities of up to 20 spots per $cm^2$, micro-arrays on glass slides have order of magnitude higher densities with about 200 spots per $cm^2$[34] (Fig. 2). In the macro-arrays, the peptides on one spot have a diameter of approx. 2–3 mm,[24] whereas in micro-arrays the diameter of one spot is approximately 0.7 to 1 mm.[35] This significantly reduces the volume of solvent and reagents consumed in the experiment. Furthermore, if the protein tends to aggregate, the difference between the peptide densities on the micro- and macro-arrays can become an important factor. In addition, the peptide micro-array assay is performed directly on the glass slide, allowing the use of coverslips. This leads to experiments in which much less volume of protein/antibody solution is required (about 100 µl with the coverslip *vs.* 2 ml without the coverslip). The concentration of peptide per spot also increases significantly in the transition from macro to micro:

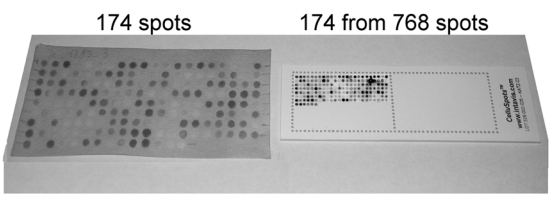


**Fig. 2** Size comparison of macro-array *versus* micro-array. Cellulose-bound macro-array containing 174 peptides (left) and microscope slide micro-array containing 768 peptides (right). 174 peptides on the micro-array were highlighted to ease the comparison. The relative original size differences are presented.

 

A typical synthesis of a macro-array is expected to yield peptide amounts at the pmol range, whereas for a micro-array such yield is expected to be at the nmol range.[36] This high peptide density in each spot is advantageous for identification of interactions with relatively low binding affinities. However, in some cases lower peptide concentrations had to be used since the protein (*e.g.* the chaperone DnaK) did not enter the spots of the high density arrays.[31]

A significant difference between the macro- and micro-arrays is the number of times they are used. While the micro-arrays are single-use glass slides, the macro-array cellulose membranes are made to be reused a number of times, after the appropriate washing and regeneration steps. To reuse the membranes, they must be stripped from the protein already bound to them in the first experiment. This is not a problem for proteins that bind peptides with high dissociation rates or when fractionated electrotransfer would fully remove the protein from the array, as it was the case in studies with molecular chaperones.[31,37–39] This might be different for proteins that are designed to bind peptides with very high affinity, such as antibodies.

General problems arise when arrays are re-used: if the time between two experiments exceeds a few days, the membranes should also be dried and regenerated for storage. Little is known about the effect of the stripping, regeneration and storage conditions on the array performance. Studies on diverse impact of washing, drying and storage conditions on the array activity have been reported, but these studies are limited only to antibody arrays and protein arrays.[40,41] In cases where the stripping is not complete it would not be possible to reuse the membrane.

The above problems led to a need for an efficient production method for multiple replicas of the same array of peptides, for a one-off use of each copy. This was crucial in order to avoid the need to use the same array several times. The use of one array per experiment also has a technical advantage while performing the binding assay: for a disposable micro-array, the detection is performed directly on the glass slide, while for reproducible macro-array in order to strip the array for future reuse the detection involves electrotransfer of the bound protein to a secondary membrane (mostly nitrocellulose or PVDF). The advantage of a disposable array is obvious: there are fewer steps and thus there is less space for technical errors.

An example of a direct side-by-side comparison between macro- and micro-arrays (which examines the binding between the chaperone Hsp90 and peptides derived from the CFTR protein) can be found in the ESI.†

## Preparing peptide micro-arrays

There are several techniques to prepare peptide micro-arrays on solid support. The preparation of peptide micro-arrays is complicated, expensive and requires special equipment. Therefore, their development and production is performed mainly by commercial companies. The differences between the commercial companies are detailed in Table 1.

An example for a high-quality technique used is the SC2 method,[27] which has been commercialized by INTAVIS.[35] This method is a modification of the SPOT synthesis technique. The synthesis begins as a standard Fmoc-based SPOT synthesis on solid support, but with a major difference, which is the use of acid labile cellulose based disks as the support. After synthesis, the disks are dissolved in acidic cleavage solution to obtain only the peptide covalently bound to the cellulose polymer. The polymers are precipitated in ether, re-dissolved in DMSO and then spotted onto coated glass slides. After evaporation of the DMSO, dried peptide–cellulose conjugates are obtained on each spot of the array. These peptide–cellulose–polymers bonds cannot dissociate in aqueous buffers. The chemical nature of the surface is similar to the standard surface of ordinary cellulose sheets, and the cellulose based polymers are suitable for most kinds of biological assays.[35]

Over the years, the photolithography technique also progressed towards reduction of the array size and the creation of high density peptide micro-arrays. Along with the photogenerated acids (PGA), the use of *t*-Boc based synthesis was also introduced.[29] Another improvement emerged in the field: a digital optical unit for light pattern projection onto the surface. Light irradiation with selective patterns enables the selective activation of a specific location, enabling the large scale parallel synthesis of different peptides without the need for expensive, inconvenient micro-fabricated photomasks. This technology has been commercialized by LC Sciences (for a scheme see the web site of the LC Sciences).[30] LC Sciences also introduced a unique micro-fluidic device, which provides picolitre scale reaction chambers (pico-array reactor). These types of peptide micro-arrays have the additional advantages of controlled environment. Parameters such as time, temperature, flow rate and flow direction can be preset and controlled. In addition, the binding assay is performed in a closed system (as opposed to open slides). This removes the risks of dye oxidation and deterioration. The benefit of not using pre-synthesized peptides combined with the spotting process is highly uniform spots, and not having stock-pile with an expiration date. LC Sciences reports that the spot density obtained using this technology can be ten-fold higher than that of spotted chips.[30] This method results in significant saving of reagents and samples due to the new micro-fluidic device that enables the use of minimal volume: subnanolitre to picolitre level per reaction.[30]

Another approach to generate high density peptide micro-arrays is by using particle-based synthesis. This method involves the use of charged solid Fmoc amino acid–pentafluorophenyl ester particles as building blocks, which are positioned onto the slide using an electrical field produced by either a laser printer[42] or a computer chip.[43] The printing technology has been commercialized by PEPperPRINT.[44] Briefly, the charged amino acid is embedded within the toner particles and brought to specific locations on a glass slide. This is followed by simultaneous melting of the whole amino acid-loaded particle layer, to initiate the coupling reaction. The melting enables the amino acid derivatives to diffuse and undergo coupling. This way the desired amino acids are integrated one on top of the other to form the required peptide. As in conventional solid-phase synthesis, each synthetic cycle is completed with the washing of excess amino acid and Fmoc deprotection (for a scheme see ref. 42). This technique further overcomes the limits of speed, quality and amount of peptides that can be synthesized on one array. Printing time is

**Table 1** Comparing peptide arrays offered by commercial companies

| | INTAVIS | JPT | | PEPperPRINT | LC Sciences | PEPSCAN |
|---|---|---|---|---|---|---|
| Product | CelluSpots micro-array | PepSpot macro-array | Pepstar micro-array | PEPperCHIP micro-array | PepArray micro-array chip | PepChip micro-array |
| Technology | SC2 technique: Standard Fmoc SPPS using the SPOT synthesis technique on cellulose discs | Standard Fmoc SPPS using the SPOT synthesis technique | Standard SPOT synthesis on glass slides. Peptides are attached *via* a reactivity tag or linker | Standard SPPS amino acids are placed on the slide using an electrical field produced by a laser printer | Digital photolithography using specialized microfluidics equipment | Standard Fmoc SPPS |
| Support | Glass slides | Cellulose sheets | Glass slides | Chips/Glass slides | Chips | Glass slides |
| Attached to the support | *via* C-terminus | *via* C-terminus | *via* N-terminus | *via* C-terminus | *via* C-terminus | *via* C-terminus |
| Peptide purification and verification | No individual purification Routine sample testing by HPLC/MS | No individual purification 5% of ordered peptides plus internal controls are verified by HPLC/MS | No individual purification 5% of ordered peptides plus internal controls are verified by HPLC/MS | No individual purification Routine sample testing by HPLC/MS | No individual purification Calculated purity for 8 mers is ~75–80%, for 12 mers ~65–70% | Different purities are available. Peptide can undergo individual purification using LC/MS |
| Scanning assay | Standard screening with antibody-based detection. Performed directly on the array | Standard screening with antibody-based detection, *via* electrotransfered blotting to PVDF membrane | Standard screening with antibody-based detection, by creating micro array-chip-sandwich directly on the slide | Standard screening with antibody-based detection, directly on the array. In house services for staining and read-out are optional | Screening by micro-fluidic system. Binding protein solution is re-circulated through the chip, and assay signals are collected by fluorescence scan | Standard screening with antibody-based detection, directly on the array |
| Possible regeneration | No | Yes | No | No | No | No |
| Need of special materials/ equipment | No | No | No | No | Yes, limited to in house services | No |
| Recommended peptide length | Up to 15 mer | Up to 15 mer | Up to 15 mer | up to 20 mer | Between 8–12 mer | >20 mer (depends on the purification choice) |
| Maximum peptides per array | Up to 384 × 2 duplicated peptides | Up to 800 | Up to 6912 × 3 duplicated peptides | Two scales: 156 000 on chip/5440 on glass slide | Two basic format of 4 K and 30 K chips | >500 duplicate on one glass slide |
| Cost (For 10–15 mer array) (information correct for the time the review was written) | € 6.25 per peptide, for 20 replicas € 800 for additional 20 replicas | € 0.6 per residue € 263 setup fee per spotted array | € 13.5 per peptide; € 9 for control peptide spots; € 88 for replica | € 0.135 per peptide + 10% for control peptide spots | 1150$ for synthesis of standard 4 K chip + 500$ for service and data extraction | € 25 per peptide € 150 per spotted slide |
| Website | www.intavis.com | http://www.jpt.com | | www.pepperprint.com | http://www.lcsciences.com | http://www.pepscanpresto.com |

seconds (*versus* hours of spotting) and it can yield up to 5440 individual peptide spots per microscope slide.[44]

## Designing peptide arrays—decision making time

As mentioned above, the SPOT technology is currently well commercialized, making it is much cheaper and quicker to order the peptide array from a commercial company rather than perform in-house synthesis, for which most laboratories are not equipped. The major task before ordering is the design of the required peptide array. This includes deciding how many peptides the array will contain, their length, their sequences and whether to introduce any modification into some of the peptides.

There are two main approaches for designing the peptides in the array: rational design and combinatorial design.

 

Combinatorial design is based on the synthesis of many peptides, usually derived from one specific motif (which is suspected to be the interactor), with systematic small modifications in the parent sequence. This results in a wide range of diverse peptides that are based on the same motif and are screened for binding in order to find the optimal interacting sequence.[12,45] On the other hand, randomized peptides can be also combined with specific defined positions (see ref. 31 for an example). In the rational design, the array includes peptides derived from sequences in partner proteins that are known to bind the protein of interest, or from sequences of proteins that are related or homologous to each other and act *via* the same biological pathway. This approach is applied both for identifying the precise binding sites in the partner proteins, and for identifying peptides that can serve as ligand-mimics for therapeutic applications. After deciding on the best approach for determining the sequences of the peptides in the array, six rules should be followed:

(i) The optimal length of the peptides should not exceed 15 residues. This is since the typical purity of an average 6–15 AA peptide is $>70\%$.[24] For longer peptides purity cannot be guaranteed, since the purity and identity of each individual peptide is not tested. Longer peptides can be made in addition to the shorter peptides, but in any case the binding should be confirmed by studies with the purified peptides, to avoid false positive results caused by binding of the synthesis by-products.

(ii) To increase the reliability of the screening results, the designed peptides should be partly overlapping. When binding of the protein to partly overlapping peptides is observed, it strongly indicates a specific binding region rather than one random binding sequence. Moreover, it can help minimizing the binding site to the shortest binding sequence common to all the observed partly overlapping peptides (Fig. 1).

(iii) When choosing a sequence composition it is important to notice the surface electrical charge and the hydrophobic nature of the peptide. Peptides containing a large percentage of hydrophobic residues, such as Leu, Val, Ile, Met, Phe, and Trp, usually have solubility problems in aqueous solutions (which are the buffers using for the screening of the arrays). In addition, sometimes hydrophobic peptides tend to have non-specific interactions due to aggregation effect, and this is a potential source for false screening results. Cys, Met, and Trp residues can also cause synthetic problems because they are susceptible to oxidation and side reactions.[46] If possible, it is advisable to choose sequences that contain a minimal number of these residues. On the other hand, solid phase peptide arrays do not have problems with solubility and allows mapping of interactions that is not possible in solution, as shown for example for the specificity of molecular chaperones. Many high affinity binders of DnaK turned out to be insoluble in solution while they were accessible on the peptide array.[47]

(iv) When performing rational design of peptides derived from the sequence of a known protein for mimicking protein–protein interactions, it is highly recommended checking if its structure has been solved. If so, the secondary structures of the protein should be considered during the design. For example, helices or beta strands should be taken as whole, even with 2–3 flanking residues from each side, to allow the peptide to populate its native secondary structure in its conformational ensemble (Fig. 1). If the peptide containing the full secondary structure element is longer than 15 residues, it is advisable to synthesize both the short and long peptides.

(v) When doing rational design of peptides derived from the sequence of a protein, their N-terminus should be acetylated, eliminating the positive charge on the N-terminal amine. This acetylated uncharged state represents the physiological condition where the peptide is derived from a protein sequence in which the N-terminus is participating in an amide bond. Acetylated peptides are also more stable to degradation.

(vi) Last but not least, always insert a few control peptides with known binding affinity to the protein of interest. It gives a comparative measure to the binding affinity of the unknown peptides, since the binding assays results are only semi-quantitative. Moreover, the peptide arrays are expensive, and it is better to use such an internal control than to use a whole array for calibration of the system. Calibrating the assays according to a control peptide also speeds up the process.

## The next step—what to do with the screening results

The spot patterns obtained from peptide array screening can be documented and semi-quantitatively evaluated using standard modern image analysis readers such as the ones commonly used in 2D analysis of gel electrophoresis and blotting. Weiser *et al.* presented a comprehensive study on the reliability and accuracy of dissociation constant calculation based on the signal intensities.[48] They showed that the data analysis for array screening is not trivial, since many variables may influence the signal intensities. These include for example the array type, peptides composition, number of spots on the array, antibody concentration, incubation time, inhomogeneity of the array, the possibility of competition among the peptides for the binding to the partner protein and the array functionality quotient. The array functionality quotient depends on the ratios between the solutions (Fmoc-β-alanine-OPfp and *N*-acetyl-β-alanine-OPfp in dimethyl sulfoxide) in the second coupling step of the next anchor position.

Each experiment is an individual collection of the above variables and thus we recommend evaluating the observed intensity of the peptides within each array only in comparison to the other peptides on the same array, but not to compare the intensity with the screening results of other copies of the same array or other arrays. The intensity control is internal for each array. Weiser *et al.* found that the standard deviation of the signal intensities varies in the range of 8–22%. They found a way to improve the standard deviation by taking regional trends in the membrane surface into account. They normalized (using a particular equation) all the binding intensities according to two types of reference peptides: peptides with high-affinity and peptides with low-affinity to the binding partner.[48] Since in many cases there is no prior knowledge about the binding affinities, they also established an equation that allows using the background signal around each spot as an alternative to a low-affinity reference.[48] In a similar way signal intensities can also be normalized by reference peptides with known dissociation constants.[37,39]

The bottom line is that the peptide array screening method is only semi-quantitative. Comparisons of the spots intensity can

be made but the information gained will be of relative affinities and not of absolute $K_d$ values. The reasons for that are mainly technical: for example, since the peptides on the array are usually not purified, different peptides can be present at different yields. The results could be influenced by unsuccessful synthesis with low yields or high amounts of by-products. In addition, observed binding at the peptide level does not prove with certainty that the interaction also occurs at the protein level. To verify this, experiments at protein level are also required.

In a study that identified binding between PLCγ1 SH3 domain and peptides (derived from 12 different proteins) containing proline rich motifs on a macro-array the peptides binding was validated at two levels: (i) At the protein level: the full length proteins from which the observed peptides are derived were tested for binding to PLCγ1 SH3 domain by GST pull-down or co-IP. The results showed 75% success, indicating that peptide array screening is a good method for discovering protein–protein interactions. (ii) At the peptide level: the observed peptides were synthesized and their binding affinities towards PLCγ1 SH3 were measured using fluorescence polarization. The results demonstrated good correlation between the strength of the signal on the array and the dissociation constant. It was clear that peptides with high signal intensity on the array also had stronger binding in solution ($K_d < 50$ μM). These peptides also demonstrated interactions at the protein level. Peptide with low signal intensity on the array had weaker binding in solution ($K_d > 50$ μM) and did not show any detectable interaction at the protein level.[49]

In our laboratory, we also test the binding affinities in solution for the binding peptides observed in the array. We synthesize and purify the binding peptides that had high signal intensity on the array and then measure their binding affinity using quantitative methods as fluorescence anisotropy or SPR.[6,32] In parallel, some of the non-binding peptides are synthesized and tested for binding in solution as control. For example, in studies of the interaction between $ASPP2_{Ank\text{-}SH3}$ and peptides derived from the anti-apoptotic Bcl2 family members, the results showed that some of the peptides gave a weak binding affinity (>50 μM), meaning that the physiological significance of their interactions with the parent protein is doubtful, especially in cases of linear binding motifs. On the other hand, some of the peptides gave a relatively strong binding affinity ($K_d < 1$ μM). In consistence with the previous study, we conclude that the range of binding affinities covered by the array screening method is very wide and therefore validation of the screening results is crucial for data evaluation. The results reliability increases when using several independent methods gives the same affinity. For example, we observed that for all the peptides that exhibited a relatively strong binding affinity ($K_d < 1$ μM) similar $K_d$ values were obtained using several independent methods.[6] Affinities also depend on the protein–protein interaction that is studied. In the case of the DnaK chaperone, for example, all peptides that showed affinity on a peptide array showed affinity in solution, with $K_d$s covering a range of 0.1–3 μM, while non-binding peptides on the array did not show detectable affinity in solution.[47] We conclude that from our experience the peptide array is a good tool for rapid screening of many sequences. But it must be accompanied by another biophysical quantitative method in order to reveal the precise binding affinities.

## Peptide array screening: examples for new applications and developments

The peptide array screening is currently a well established method. In this section, we will present some of the recent unique developments and applications in the field.

All the peptide arrays mentioned so far are suitable only for cases in which the interaction occurs between a protein and linear epitopes or interfaces from the partner protein, which can be represented by linear peptides. Such a system is not necessarily suitable for testing all kinds of interactions between proteins. For example, if the binding site consists of several discontinuous sequences that are spatially close only when the protein is folded, each peptide on the array will represent only part of the binding site. Thus, an individual representative liner peptide on a micro-array would not necessarily demonstrate binding. Several techniques have been developed to solve this problem, for example the duotope scan and the CLIPS™ (Chemical LInkage of Peptides onto Scaffolds) technology.[50,51] The duotope scan is based on the idea that a discontinuous binding site can be mimicked by connecting the potential binding regions (represented by linear peptides) using a linker that resembles their distance in the native 3D protein structure. The linker provides the necessary flexibility for the region to fold correctly.[50] The CLIPS™ technology allows the fixation of peptides into stable secondary structures using a selective chemical reaction between peptides containing homocysteine residues and a synthetic scaffold containing a benzyl bromide group ($PhCH_2Br$). For further details and figures see ref. 51. Using this technology, protein–protein interactions mediated by discontinuous binding sites can be mapped, even if the epitopes are found in oligomeric states such as dimer or tetramer. In addition, it has been shown that using the CLIPS™ method, mapping of interaction sites can be done at the level of single amino acids. This was demonstrated for mapping the mAbs protein binding sites with several proteins—hFSH, hCG, CD20, and CCR5.[51]

Espanel *et al.* have designed another approach for using peptide arrays to study interactions of several regions simultaneously and/or cooperatively. They developed a method to synthesize more than one peptide at each spot. In an early study, the synthesis of two peptides in one spot, using orthogonal protection strategy, was demonstrated. A mixture of Fmoc-β-alanine and Alloc-β-alanine was used in the first cycle, permitting selective deprotection of the Fmoc and Alloc by different reagents (piperidine for Fmoc *vs.* bis(triphenylphosphine) palladium(II)chloride $(PdCl_2(PPh_3)_2$ for Alloc). This selective deprotection allowed a separate synthesis on each β-alanine in the same spot. The synthesis could not be performed in parallel, and the second peptide had to be synthesized after the acetylation of the first peptide.[52] Recently, the same group further developed the technique for the synthesis of up to four peptides in one spot. An additional protecting group, ivDde (1-[4,4-dimethyl-2,6-dioxocyclohexylidene]3-methylbutyl), was added on the side

 

chain of the first C-terminal lysine. This group is stable under piperidine treatment, therefore enabling the separate synthesis of two peptides on each Fmoc and Alloc-β-alanine. The ivDde is removed by 2% hydrazine monohydrate in DMF allowing the second peptide synthesis and making it possible to synthesize a total of four peptides in the same spot.[53]

In standard SPPS the peptide C-terminus is immobilized to the solid support and the synthesis is carried out towards the N terminus, resulting in a peptide with a free N-terminus. A new application, which was evolved in order to address cases where there is a need for a free C-terminus (such as interactions of PDZ domains), is named "inverted peptide arrays". In this case, the N terminus of the peptide is modified so that after the end of a standard synthesis (from the C to the N terminus), it can undergo intramolecular cyclization with the solid support. That support usually contains Cys (Trt) and the cyclization occurs upon treatment with cesium carbonate. The C terminus is then detached from the solid support by hydrolysis (and simultaneous side chain deprotection) leaving a complete peptide attached only by its N terminus (for a detailed scheme see ref. 26). Inverted peptide arrays have been used for mapping several interactions of PDZ domains.[54–56]

A novel type of peptide array is the dry peptide array.[57] This method includes drying peptides in wells of glass plates (using vacuum pump), then adding a solution of the partner protein labeled with fluorescein. This leads to the dissolution of the dried peptides as well. After a short incubation (about 15 minutes), the wells are being dried again. Fluorescent traces will remain only in the wells where binding took place. This fluorescence can then be quantified. The advantage of this system is the use of non-immobilized peptides, which may better represent the native structure of the peptide allowing its correct folding upon ligand binding. Additional benefits are that peptides remain more stable under dry conditions, the assay procedure does not include washing steps, and it is considered to be much simpler than standard screening. The dry peptide arrays approach was demonstrated with several proteins including calmodulin (CaM) and a mini-array consisting of 20 designed basic amphiphilic α-helical peptides.[57] It is a disadvantage, though, that the protein has to be labeled with fluorescein or comparable dyes, which may change the properties or solubility of the protein.

A common application for peptide array screening is enzymatic assays. Peptide arrays can be used to determine enzyme kinetics, to screen rapidly for enzyme inhibitors, and to test biological activity in a dose-dependant manner.[58] A variety of enzymes have been examined using peptide microarrays, the most studied are kinases and proteases. These applications are usually straightforward because the label can be covalently linked to the enzyme (for reviews see refs. 59 and 60). One recent unique example for using the peptide array for enzyme profiling is shown in ref. 61. The authors addressed the problem of quantitative array analysis and demonstrated biological activity (kinase phosphorylation) on a peptide array.[61] They prepared an array of peptides derived from the tyrosine kinase c-Src. The peptides contained N-terminal cysteines and were immobilized onto the amino-modified glass chip *via* glutaraldehyde linkage. Activity assays were carried out by phosphorylation of the peptides on the array by c-Src kinase followed by treatment with fluorescently (Cy5)-labeled anti-phosphotyrosine antibody. Two c-Src peptides on the array were designed to be positive and negative controls for a quantitative assay: a peptide synthesized with a phosphotyrosine served as a positive control, while the same peptide with a Phe residue in the same position served as a negative control. Taking the positive control as a reference for 100% phosphorylation and the negative control as a reference for 0% phosphorylation, they obtained the actual on-chip phosphorylation ratio for the substrate. In this study, on top of the regular calibrations, different amounts of peptides were immobilized on the same array, leading to optimization of the immobilized peptides concentrations. This step is important for the activity assay since it indicates that there is correlation between the fluorescence intensity and the phosphorylated substrate in the measured concentration range. This is an important difference between this method and the cases described so far, for which the levels of the immobilized peptide were identical for all spots.

Another new type of micro-array that emerged in past ten years is the carbohydrate micro-array. Those arrays have become important tools to investigate binding events that involve sugars. Carbohydrate arrays simulated the natural arrangement of carbohydrates on the cell surface, and can be used to study carbohydrate interactions with a wide range of binding partners including proteins, RNA and viruses. This tool has many diverse applications in research and diagnosis and it has already proved itself in the glycobiology field. For more details see ref. 62.

In addition to the new types of arrays described above, new scanning methods are also being developed. For example, Greving *et al.* presented the two-dimensional scanning of peptide arrays.[63] Using a flow chamber and fluorescence imaging, they were able to measure in a single experiment both the kinetic and thermodynamic parameters for peptides binding. They demonstrated the new technique with a labeled TNF-α protein and a micro-array containing 800 peptides. The micro-array was placed in a chamber equipped with a micro-channel that transports the fluids and with a low vacuum around the slide edge, which provides a flat imaging plane. Fluorescently labeled TNF-α was incubated with the peptide array for 50 min, and then dissociated by flowing buffer through the chamber. Fluorescence images revealed the binding peptides and the dissociation phase provided the dissociation curves. The reliability of the dissociation constants was verified using SPR for 180 peptides (out of the 800 peptides in the array). 60 peptides displayed observable TNF-α binding in the SPR experiments. Of these, 54 have also shown significant binding in the micro-array. 26 of the 60 peptides that showed significant kinetic dissociation constants ($k_{off}$) in the SPR experiments have showed also significant $K_{off}$ in the micro-array analysis. The $k_{off}$ determined on the micro-array were on average ~20-fold slower when compared to the $k_{off}$ determined by SPR. Therefore, this two-dimensional scanning method is only offered as an extensive scan tool for rapid search of binding peptides. Binding should then be confirmed and quantified using additional methods. This application provides an insight into the binding mechanism in addition to mapping the potential binding sites.[63]

Integration between computational analysis/modeling and scanning of peptide arrays is becoming widely used. Peptide scans can provide much information, and sometimes even hundreds of peptides are considered as potential interactors based on the screening results. It is necessary to thoroughly analyze this information and focus only on a few best binding

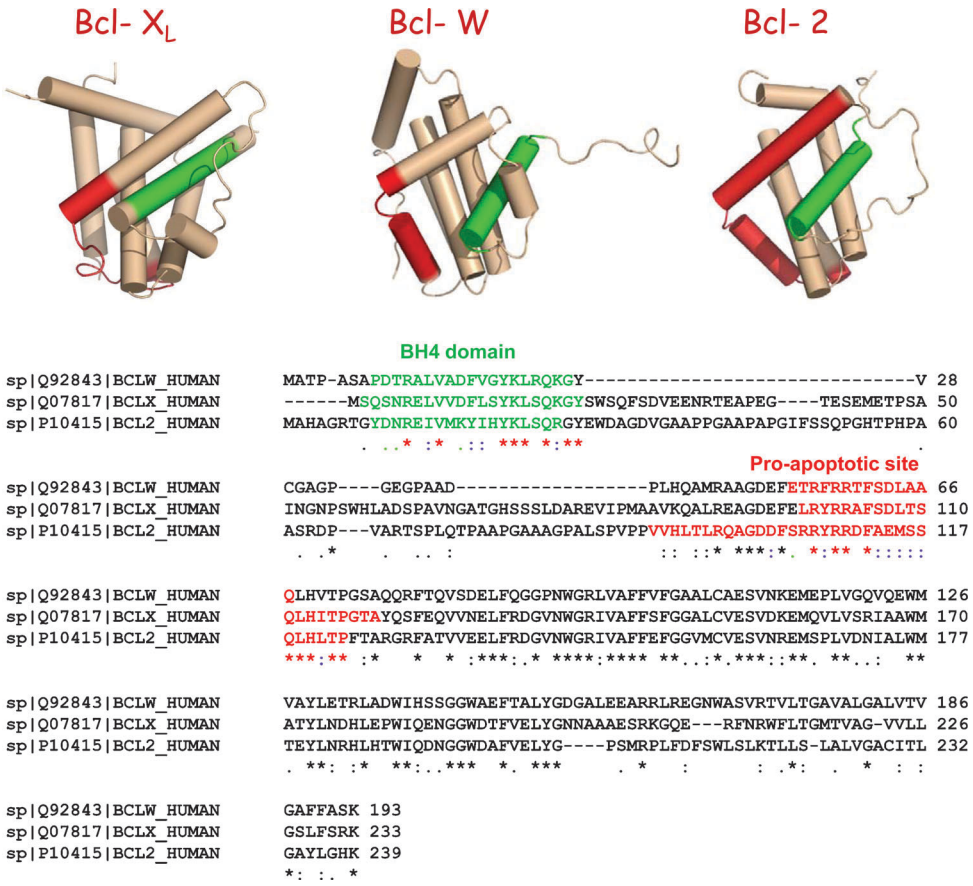


**Fig. 3** Peptide array screening results of the ASPP2$_{Ank\text{-}SH3}$–Bcl2 interaction.[6] Using peptide array screening ASPP2$_{Ank\text{-}SH3}$ found to bind two homologous sites in all the three anti-apoptotic Bcl2 family proteins (Bcl-2, Bcl-W and Bcl-X$_L$): (i) the conserved BH4 motif; colored green; (ii) a binding site for pro-apoptotic regulators, colored red. The homologous sites colored both on the known 3D structures (Bcl-2 PDB: 1ysw (23), Bcl-X$_L$ PDB: 1g5j (20), Bcl-W PDB:1o0l (33)) and on the three aligned sequences.

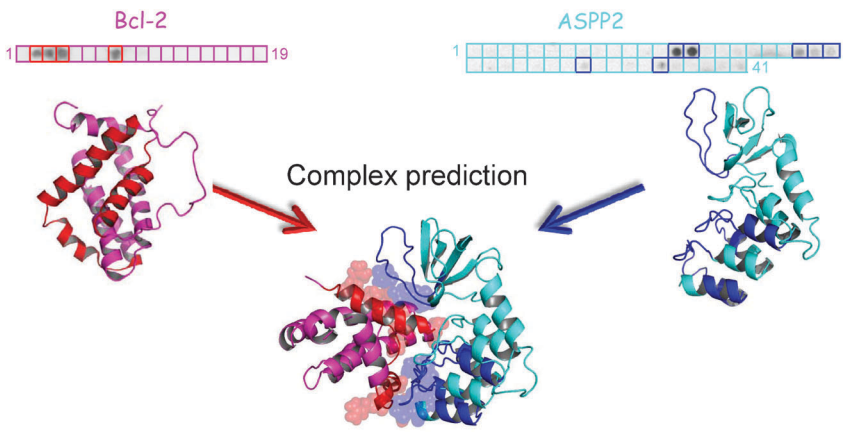


**Fig. 4** Docking model for the complex ASPP2$_{Ank\text{-}SH3}$-Bcl2 based on the peptide array screening results of both Bcl2 and ASPP2$_{Ank\text{-}SH3}$.[6] On the left an array consisting of partly overlapping peptides derived from Bcl2 protein was screened for binding ASPP2$_{Ank\text{-}SH3}$. Each dark spot represents binding of ASPP2$_{Ank\text{-}SH3}$ to a specific peptide. The ASPP2 binding sites, as observed in the peptide array screening, are highlighted in red on the known 3D structure of the Bcl2 colored pink (PDB: 1ysw). On the right an array consisting of partly overlapping peptides derived from ASPP2$_{Ank\text{-}SH3}$ protein was screened for binding Bcl2. Each dark spot represents binding of Bcl2 to a specific peptide. The Bcl2 binding sites, as observed in the peptide array screening, are highlighted in blue on the known 3D structure of the ASPP2$_{Ank\text{-}SH3}$ colored in cyan (PDB: 1ycs). In the middle, binding peptides served as a base (data restriction) for a docking model for the interaction between the ASPP2$_{Ank\text{-}SH3}$ and Bcl2 at the full protein level. Figures were generated using PyMol (34).

 

peptides to avoid unnecessary work. Using computational algorithms and models, it is possible to dramatically improve the array design as well as screening analysis and focus on the best potential interactors, thereby saving lots of materials, money and time. Several such computationally assisted peptide screening and design tools are already known (see ref. 64). Recently, this was demonstrated for the design of cell adhesive peptides for cell culture scaffolds[64] and for the detection of ZnO-binding peptides.[65] In both cases, a fuzzy neural network (FNN) algorithm was used by comparison of high and low affinity peptides. FNN is a type of artificial neural network that automatically constructs complex model structures by learning the hidden relationships between input and output data and thus functions as a predictor.

Improvement of the predictive strength of the model was achieved by repeating and refining the experiments. Computationally assisted screening can be very useful in finding leading peptides with minimum trials.

## Conclusions

Peptide arrays screening is becoming extensively used both in the genomic and proteomic fields. This is one of the best methods of performing a rapid scan through a lot of peptide sequences, using a relatively simple technique, which yields masses of high quality information. Putting an emphasis on proper peptide design and creating partly overlapping peptide sequences are essential in order to obtain a system with all the required internal controls.

Peptides are excellent tools for studying protein–protein interactions. Since they are chemically synthesized and not recombinantly expressed, they are easier to produce and handle relative to proteins. Peptides serve as good models for binding studies of proteins domains, because they often undergo induced fit upon ligand binding and gain their native structure. This was shown for example for peptides derived from the BH4 domain of Bcl2[66] and from Bak.[67] Peptide synthesis is also essential in order to incorporate modified or non-natural amino acid with 100% specificity. This enables for example the design of peptide arrays to systematically test the effect of post-translational modifications.

Synthesis of peptides immobilized to a solid phase has several advantages compared to peptides in solution: (i) they are cheaper than making the individual peptides; (ii) they enable comparable analysis of hundreds or thousands of peptides in one experiment under identical conditions. This can be very useful for comparing between several homologous proteins of the same family. For example in the case of $ASPP2_{Ank\text{-}SH3}$ protein, we revealed that $ASPP2_{Ank\text{-}SH3}$ binds two homologous sites in all three anti-apoptotic Bcl2 family members tested[6] (Fig. 3); (iii) no solubility problems; (iv) easy variation of sequences; (v) combination of randomized and non-randomized peptides is made easy; (vi) peptide array screening also provides a highly efficient way for identifying the binding sites in both interaction partners, and these results can serve as a base for modeling prediction for the whole complex at the full proteins level, as we did for the Bcl2-$ASPP2_{Ank\text{-}SH3}$ complex[6] (Fig. 4). Combination of peptides

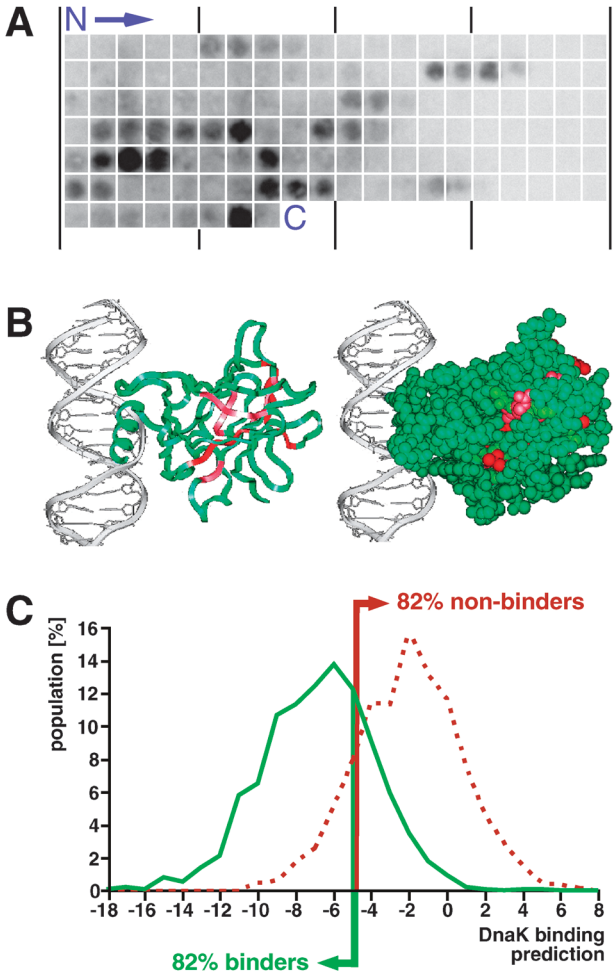


**Fig. 5** Disclosing the substrate specificity of a molecular chaperone by peptide screening (modified from Rüdiger *et al.*).[31] (A) Binding of the Hsp70 chaperone DnaK to peptides derived from the sequence of human p53. Macro-array, 13 mer peptides, each peptide is shifted 3 amino acids towards the C-terminus compared to the peptide before. (B) Plotting of the DnaK binding sites onto the structure of the p53 DNA binding domain demonstrates that the chaperone binding sites are mainly buried inside the hydrophobic core. The backbone (left) or the side chains (right) of the Dnak binding sites are indicated in red. (C) An algorithm to predict DnaK binding and non-binding peptides. The algorithm was developed based on the analysis of binding data of more than 4000 peptides. Score values of −5 and lower indicate DnaK binders. The experimentally verified peptides are plotted against the algorithm score value: green, DnaK binding peptides; red, DnaK non-binding peptides.

arrays with alanine scan can further "zoom in" and provide the minimum binding sites at the single residue level.

A key advantage of peptide arrays compared to the analysis of peptides in solution is the cheap and fast analysis of a large amount of peptides. This does not only allow mapping of protein–protein interaction sites or developing lead compounds, it also enables identification of binding motifs of peptide binding proteins, such as molecular chaperones of the Hsp70 class.[31] For example, in the case of the Hsp70 chaperone DnaK, a prediction algorithm could be developed based on peptide array data that allows prediction of binding sites in protein sequences with more than 80% accuracy (Fig. 5).

Such a study requires 2000–4000 peptides. It is a significant advantage of peptide arrays over methods such as phage display libraries that they allow not only to identify binders but also non-binders.

Technical developments in recent years led to minimization of the system, moving from macro-arrays to micro-arrays. Micro-arrays not only give more results per experiment, but also improve the quality of the results. Preparation of the arrays is not a complicated procedure and it usually requires only standard materials and equipment, yet it is a long and expensive process. Thus, it is better in our opinion, to order the arrays from commercial companies rather to make it in-house, and focus in the lab on the peptide array design, the assay performance, optimization and the results analysis.

The main limitation of the peptide array screening method remains the quantification of the observed results. In our opinion, despite the efforts made for quantifying peptide array data, it is still best to treat the method as a semi-quantitative method that provides preliminary results but needs verification and quantification by other methods. The capability of the system to mimic the physiological structure of a peptide/binding site is limited due to the immobilization of the peptides to a solid support. In addition, it is essential to verify that the observed peptide–protein interaction also occurs in the full length proteins level. In summary, the peptide array scanning method is excellent for mapping protein binding sites of previously known interactions, and not for discovering new interaction between proteins.

## Abbreviations

HOBt, *N*-hydroxybenzotriazole ester; Boc, *tert*-butoxy-carbonyl; Fmoc, 9-fluorenylmethyloxycarbonyl; PVDF, polyvinylidene difluoride; PGR, photogenerated reagents; CFTR, cystic fibrosis transmembrane conductance regulator; Hsp90 protein, heat shock protein 90.

## Acknowledgements

AF is supported by a starting grant from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC Grant agreement no. 203413 and by a grant from the Israeli Ministry of Health. SGDR was supported by a Marie-Curie Excellence grant of the EU, a VIDI career development grant of the Netherlands Organisation for Scientific Research NWO and the High Potential Program of Utrecht University.

## References

1 R. Frank, W. Heikens, G. Heisterberg-Moutsis and H. Blocker, *Nucleic Acids Res.*, 1983, **11**(13), 4365–4377.
2 H. M. Geysen, R. H. Meloen and S. J. Barteling, *Proc. Natl. Acad. Sci. U. S. A.*, 1984, **81**(13), 3998–4002.
3 R. Frank, *Tetrahedron*, 1992, **48**(42), 9217–9232.
4 S. P. Fodor, J. L. Read and M. C. Pirrung, *et al.*, *Science*, 1991, **251**(4995), 767–773.
5 F. F. Bier, M. von Nickisch-Rosenegk and E. Ehrentreich-Forster, *et al.*, *Adv. Biochem. Eng. Biotechnol.*, 2008, **109**, 433–453.
6 C. Katz, H. Benyamini and S. Rotem, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**(34), 12277–12282.
7 J. K. Scott and G. P. Smith, *Science*, 1990, **249**(4967), 386–390.
8 A. Furka, F. Sebestyen, M. Asgedom and G. Dibo, *Int. J. Pept. Protein Res.*, 1991, **37**(6), 487–493.
9 K. S. Lam, S. E. Salmon and E. M. Hersh, *et al.*, *Nature*, 1991, **354**(6348), 82–84.
10 R. A. Houghten, *Proc. Natl. Acad. Sci. U. S. A.*, 1985, **82**(15), 5131–5135.
11 C. T. Dooley and R. A. Houghten, *Life Sci.*, 1993, **52**(18), 1509–1517.
12 R. Liu, A. M. Enstrom and K. S. Lam, *Exp. Hematol. (N. Y., NY, U. S.)*, 2003, **31**(1), 11–30.
13 K. Hilpert, D. F. Winkler and R. E. Hancock, *Nat. Protoc.*, 2007, **2**(6), 1333–1349.
14 J. R. Falsey, M. Renil, S. Park, S. Li and K. S. Lam, *Bioconjugate Chem.*, 2001, **12**(3), 346–353.
15 M. Takahashi, K. Nokihara and H. Mihara, *Chem. Biol.*, 2003, **10**(1), 53–60.
16 G. J. Wegner, H. J. Lee and R. M. Corn, *Anal. Chem. (Washington, DC, U. S.)*, 2002, **74**(20), 5161–5168.
17 M. Kohn, *J. Pept. Sci.*, 2009, **15**(6), 393–397.
18 R. Frank, *J. Immunol. Methods*, 2002, **267**(1), 13–26.
19 A. Kramer, U. Reineke and L. Dong, *et al.*, *J. Pept. Res.*, 1999, **54**(4), 319–327.
20 B. Ay, M. Streitz and P. Boisguerin, *et al.*, *Biopolymers*, 2007, **88**(1), 64–75.
21 F. Molina, D. Laune, C. Gougat, B. Pau and C. Granier, *Pept. Res.*, 1996, **9**(3), 151–155.
22 M. Takahashi, A. Ueno and H. Mihara, *Chemistry*, 2000, **6**(17), 3196–3203.
23 H. Wenschuh, R. Volkmer-Engert and M. Schmidt, *et al.*, *Biopolymers*, 2000, **55**(3), 188–206.
24 J. P. T. GmbH, http://www.jpt.com/index.htm.
25 G. B. Fields and R. L. Noble, *Int. J. Pept. Protein Res.*, 1990, **35**(3), 161–214.
26 R. Volkmer, *ChemBioChem*, 2009, **10**(9), 1431–1442.
27 U. Beutling, K. Stading, T. Stradal and R. Frank, *Adv. Biochem. Eng./Biotechnol.*, 2008, **110**, 115–152.
28 M. D. Bowman, R. C. Jeske and H. E. Blackwell, *Org. Lett.*, 2004, **6**(12), 2019–2022.
29 X. Gao, J. P. Pellois and Y. Na, *et al.*, *Mol. Diversity*, 2004, **8**(3), 177–187.
30 L. Sciences, http://www.lcsciences.com.
31 S. Rüdiger, L. Germeroth, J. Schneider-Mergener and B. Bukau, *EMBO J.*, 1997, **16**(7), 1501–1507.
32 S. Rotem, C. Katz and H. Benyamini, *et al.*, *J. Biol. Chem.*, 2008, **283**(27), 18990–18999.
33 R. Hurst, B. Hook and M. R. Slater, *et al.*, *Anal. Biochem.*, 2009, **392**(1), 45–53.
34 U. Reimer, U. Reineke and J. Schneider-Mergener, *Curr. Opin. Biotechnol.*, 2002, **13**(4), 315–320.
35 B. I. Intavis, http://www.intavis.com/en/.
36 H. E. Blackwell, *Curr. Opin. Chem. Biol.*, 2006, **10**(3), 203–212.
37 S. Rüdiger, J. Schneider-Mergener and B. Bukau, *EMBO J.*, 2001, **20**(5), 1042–1050.
38 H. Patzelt, S. Rüdiger and D. Brehmer, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**(25), 14244–14249.
39 N. T. Knoblauch, S. Rüdiger and H. J. Schönfeld, *et al.*, *J. Biol. Chem.*, 1999, **274**(48), 34219–34225.
40 P. Wu and D. W. Grainger, *J. Proteome Res.*, 2006, **5**(11), 2956–2965.
41 N. Nath, R. Hurst and B. Hook, *et al.*, *J. Proteome Res.*, 2008, **7**(10), 4475–4482.
42 V. Stadler, T. Felgenhauer and M. Beyer, *et al.*, *Angew. Chem., Int. Ed.*, 2008, **47**(37), 7132–7135.
43 M. Beyer, A. Nesterov and I. Block, *et al.*, *Science*, 2007, **318**(5858), 1888.
44 PEPperPRINT, http://www.pepperprint.com.
45 M. Cardo-Vila, R. J. Giordano and R. L. Sidman, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**(11), 5118–5123.
46 B. I. Intavis, http://www.intavis.com/en/Custom_Peptide_Synthesis/peptides/faq_technical.php.
47 J. S. McCarty, S. Rüdiger and H. J. Schönfeld, *et al.*, *J. Mol. Biol.*, 1996, **256**(5), 829–837.
48 A. A. Weiser, M. Or-Guil and V. Tapia, *et al.*, *Anal. Biochem.*, 2005, **342**(2), 300–311.

 

49 C. Wu, M. H. Ma and K. R. Brown, *et al.*, *Proteomics*, 2007, **7**(11), 1775–1785.
50 U. Reineke, R. Sabat and U. Hoffmüller, *et al.*, *Identification of miniproteins using cellulose-bound duotope scans American Peptide Symposia*, Springer, Netherlands, 2006, vol. 6, pp. 167–169.
51 P. Timmerman, W. C. Puijk, R. S. Boshuizen, P. van Dijken, J. W. Slootstra, F. J. Beurskens, P. W. H. I. Parren, A. Huber, M. F. Bachmann and R. H. Meloen, *Open Vaccine J.*, 2009, **2**, 56–67.
52 X. Espanel, S. Walchli and T. Ruckle, *et al.*, *J. Biol. Chem.*, 2003, **278**(17), 15162–15167.
53 X. Espanel and R. Hooft van Huijsduijnen, *Methods*, 2005, **35**(1), 64–72.
54 P. Boisguerin, R. Leben and B. Ay, *et al.*, *Chem. Biol.*, 2004, **11**(4), 449–459.
55 U. Wiedemann, P. Boisguerin and R. Leben, *et al.*, *J. Mol. Biol.*, 2004, **343**(3), 703–718.
56 M. Joshi, C. Vargas and P. Boisguerin, *et al.*, *Angew. Chem., Int. Ed.*, 2006, **45**(23), 3790–3795.
57 K. Usui, K. Y. Tomizaki, T. Ohyama, K. Nokihara and H. Mihara, *Mol. Biosyst.*, 2006, **2**(2), 113–121.
58 D. P. Funeriu, J. Eppinger, L. Denizot, M. Miyake and J. Miyake, *Nat. Biotechnol.*, 2005, **23**(5), 622–627.
59 A. Thiele, J. Zerweck and M. Schutkowski, *Methods Mol. Biol. (Totowa, NJ, U. S.)*, 2009, **570**, 19–65.
60 J. L. Reymond, *Ann. N. Y. Acad. Sci.*, 2008, **1130**, 12–20.
61 H. Xiaoming, S. Syuhei and Y. Takayuki, *et al.*, *Anal. Biochem.*, 2008, **372**(1), 106–115.
62 T. Horlacher and P. H. Seeberger, *Chem. Soc. Rev.*, 2008, **37**(7), 1414–1422.
63 M. P. Greving, P. E. Belcher and C. D. Cox, *et al.*, *Anal. Biochem.*, 2010, **402**, 93–95.
64 C. Kaga, M. Okochi, Y. Tomita, R. Kato and H. Honda, *Biotechniques*, 2008, **44**(3), 393–402.
65 M. Okochi, M. Ogawa and C. Kaga, *et al.*, *Acta Biomater.*, 2009, **6**, 2301–2306.
66 L. C. Lee, J. J. Hunter, A. Mujeeb, C. Turck and T. G. Parslow, *J. Biol. Chem.*, 1996, **271**(38), 23284–23288.
67 M. Sattler, H. Liang and D. Nettesheim, *et al.*, *Science*, 1997, **275**(5302), 983–986.