FEBS Letters 586 (2012) 2780–2786

journal homepage: www.FEBSLetters.org

Review

Synthetic peptide arrays for investigating protein interaction domains

Rudolf Volkmer *, Victor Tapia, Christiane Landgraf

Institut für Medizinische Immunologie Berlin, Molecular Libraries and Recognition Group, Charité-Universitätsmedizin Berlin, Hessische Str. 3-4, 10115 Berlin, Germany

ARTICLE INFO

Article history:
Received 22 March 2012
Revised 16 April 2012
Accepted 17 April 2012
Available online 25 April 2012

Edited by Marius Sudol, Gianni Cesareni, Giulio Superti-Furga and Wilhelm Just

Dedicated to the memory of Barbara Ann Townsend Winsor. She passed away on September 29, 2011 after returning to her beloved family from a conference.

Keywords:
SPOT synthesis
Protein interaction domains
SH3 domains
WW domains
PDZ domains
Peptide arrays

ABSTRACT

Synthetic peptide array technology was first developed in the early 1990s by Ronald Frank. Since then the technique has become a powerful tool for high throughput approaches in biology and biochemistry. Here, we focus on peptide arrays applied to investigate the binding specificity of protein interaction domains such as WW, SH3, and PDZ domains. We describe array-based methods used to reveal domain networks in yeast, and briefly review rules as well as ideas about the synthesis and application of peptide arrays. We also provide initial results of a study designed to investigate the nature and evolution of SH3 domain interaction networks in eukaryotes.

1. Protein interaction networks

Living systems are organized by complex, dynamic networks of molecular interactions where proteins are the central network components. Since they can bind not only to other proteins, but also to phospholipids, nucleic acids and small molecules, they link diverse physiological functions of the cell. It is tempting to suggest that a molecular recognition code rules such dynamic networks [1]. In analogy to human language, one could formulate a hierarchical organization, starting from linear sequences of polypeptide chains, then simple three dimensional fold elements and complex structural motifs, up to protein complexes [2–4]. In fact, structural modules and motifs may have isolated functional "meaning" like words in human language [5,6].

To pursue this analogy, we can think of cellular wiring as a masterpiece of evolutionary tinkering, with structural elements used many times in different protein contexts, and trial and error creating some rules of interconnectivity to achieve a favorable feature or message [7]. Therefore, it is not surprising that the idea of independent protein "linguistics" arose in the protein–protein interaction community [8]. Many believe that a complete understanding of the protein–protein interaction network will enable researchers to break the protein recognition code or predict cellular responses, or even positively interfere with the molecular basis of diseases.

The yeast two-hybrid technique [9] was the first method applied in a high-throughput manner to reveal the protein interaction network of a model organism [10,11]. Shortly afterwards, the classical pull-down in combination with mass spectrometry was the next high-throughput strategy to reveal the interactome in the same species [12,13]. The low intra- and inter-technique overlaps between the resulting networks (10–20%) indicate that these experimental approaches suffer from false positives and false negatives. In addition, graphical representation of such interaction networks is hard to understand. Recently Gianni Cesareni's interesting analogy described such a network as a complete road map of a large city, but without any information on traffic flow or which routes represent large traffic arteries and which represent narrow one-way alleys [14].

A further aspect limits interpretation of experimentally obtained networks. The real networks have a so-called scale-free topology, meaning more proteins than expected interact with many partner proteins at a time [15]. It is hard to say how many partners these "hub" proteins can contact at the same time;

* Corresponding author. Fax: +49 (0)30 450 524942.
E-mail address: rve@charite.de (R. Volkmer).
URL: http://immunologie.charite.de (R. Volkmer).

0014-5793 © 2012 Federation of European Biochemical Societies. Published by Elsevier B.V. Open access under CC BY-NC-ND license.
http://dx.doi.org/10.1016/j.febslet.2012.04.028

consequently it is nearly impossible to predict the composition of multiprotein complexes. Fortunately, the modular architecture of proteins can help solve the problem of assignment. A protein is composed of single domains (modules) separated on discrete sequence patterns, which in turn comprise folding motifs. In general, isolated protein modules have the same globular folding as in the whole protein, and therefore a reductionist approach can be applied in practice [16,17].

Protein alignments indicate that even unrelated proteins frequently share sections of sequence similarity [18–20] and these regions often function as independently folded modules or domains of autonomous functionality. It would be more informative if the protein interactome could be split at each protein (node) into covalently linked domains, whereby each domain can interact with several proteins, but only one at a time [14]. In other words, if an existing protein interaction network is supported by a domain network, this might provide enough information about the topology of protein–protein interactions.

2. Protein interaction domains

Structural analysis of functional protein complexes suggests at least two classes of protein–protein interactions [21–23]. In the first class, which reflects the majority of protein–protein interactions, the complementary surfaces of both interacting partners are extensive. This means that the residues involved in each interacting surface only come together upon protein folding (discontinuous binding sites). The second class comprises asymmetric interactions, where a modular protein domain may dock with a short linear sequence motive on the partner protein. Such a modular protein domain is called a protein interaction domain (PID).

Mapping discontinuous binding sites is a challenge; however, the concept of hot spots [24,25] shows the feasibility of interfering with interactions mediated by extensive surfaces. In contrast, PID binding determinants may be mapped to short linear motifs matching the sequence of the ligand peptide. The importance of small protein recognition domains in forming protein complexes that involve binding to short linear peptides was demonstrated in the late 1980s and early 1990s (reviewed in the excellent book [20]).

Such domains preferentially bind to peptides possessing specific sequence or structure characteristics. For example, Sudol and Bork demonstrated that WW domains bind proline-rich peptides sharing a PPxY motif, and folded into a proline type II helix [26]. However, two factors should be noted: firstly, PID interactions are dynamic with affinities mostly in the middle to high micromolar range, and secondly, one has to carefully define the domain borderlines. The first factor often hampers complete characterization of a domain-interacting network when using pull-down based experiments. The second factor is important since domain borderlines are generally defined by bioinformatic methods such as multiple protein alignments without any information about structural data. Therefore, preparing a domain according to the given borderlines might yield an unstructured shape with no function. This is known for the TCERG1 WW-3 domain where extending the bioinformatically defined borderline from a 43-mer up to a 53-mer sequence proved essential for generating a properly folded domain (PDB ID code 2dk7) [27].

3. Phage display

Over the past decade two predominant experimental approaches have been used to investigate PID recognition specificity, namely phage display and SPOT synthesis; both have high throughput potential and can reveal domain interaction networks with information on the stoichiometry of domain-peptide interactions. Phage display is a powerful biological library comprising 10^9–10^{10} peptides of random sequence displayed on bacteriophage capsids. After the pioneering work of Sparks et al. [28] and Rickles et al. [29] many groups have applied this approach to determine the recognition specificity of several domains such as SH3, WW, PDZ, and GYF domains [20,30,31]. Due to the impressive record of successful studies mapping PID recognition specificities, phage display has become the first choice for determining the recognition specificity of new domains in the absence of any a priori information.

Among the first to expand the study of peptide recognition modules to a genome-wide scale were Tong et al. [32]. They combined phage display and yeast two-hybrid to elucidate the peptide ligand consensus of 20 SH3 domains of *Saccharomyces cerevisiae* and draw up an SH3 domain network. Besides the impressive experimental results, some general conclusions could be drawn from this study. An overlap of about 25% interactions in common suggests that both methods lead to over-prediction, whereby a significant part of the interactome remains undiscovered. Interestingly, we notice that exploring the same protein–protein interaction space with orthogonal approaches removes false positives. In contrast to Tong et al., the interactome scanning study of Landgraf et al. [33] combined phage display with a semi-quantitative analysis achieved using SPOT synthesis technology. This study allowed the first comparative interpretation of the yeast SH3 domain interactome: in the devised diagram lines of different thickness correlate with binding strength.

4. SPOT synthesis and synthetic peptide arrays

Array technologies, especially protein arrays, arrived late in the field of protein–protein interactions due to critical factors such as native folding stability or functionality [34,35]. Peptides, in contrast, are easier to handle and retain partial features of protein function. The fact that PIDs recognize short linear peptides perfectly corresponds to the scope of synthetic peptide arrays. Thus, peptide arrays are predestined to support PID recognition studies such as revealing binding specificity, screening for cellular interaction partners, or developing selective PID inhibitors.

Two techniques for chemically synthesizing peptide arrays were published almost simultaneously: Frank presented the SPOT synthesis technique [36], while Fodor and co-workers [37] described the concept of light-directed, spatially addressable chemical synthesis. The latter was only used by the group that originally developed the method. In contrast, the majority of peptide arrays reported to date have been produced using the SPOT synthesis concept. This is due to the fact that SPOT synthesis is a very simple but extremely robust method for highly parallel synthesis of peptides on planar surfaces. The method itself has been reviewed several times, e.g. [38,39] and over a period of 20 years the method has become a widespread and essential tool in biology and biochemistry, with a literature base of more than 400 original, peer-reviewed papers.

Nowadays, we are in the comfortable situation that a diverse collection of SPOT technology methods are available, permitting the application of peptide arrays to a broad spectrum of targets [39]. SPOT technology simplified the chemical synthesis of peptide arrays to the addressable deposition of reagents on a planar cellulose membrane (filter paper). Moreover, chemical synthesis allows one to incorporate phosphorylated [40], methylated or acetylated amino acids [41,42], use non-natural building blocks [43,44], prepare branched and cyclic structures [45] and label with chromophores or short biological tags such as biotin [46,47].

SPOT synthesis can be performed fully automatically with a MultiPep synthesizer (Intavis AG, Köln, Germany) in an analytical or preparative mode. The latter enables parallel synthesis of up to 1000 peptides at amounts of approximately 50–100 nmol of cleaved material, achieved by producing spots with a diameter of 8 mm dispensed in droplets of up to 1 μl. The preparative SPOT synthesis mode has been used mostly for cell-based screening assays [48,49] but also for quality control of spot synthesized peptides. The analytical mode enables the synthesis of about 6000 cellulose membrane-bound peptides by producing spots with a diameter of ~1 mm. Down-sizing the spot diameter below 1 mm failed due to the capillary effect of cellulose membranes. In both cases we use membranes with a dimension of 18.3 × 28.5 cm for a special platform with the fully automated MultiPep synthesizer that allows four coupling steps per day. Fig. 1 summarizes the principal steps of both the analytical and preparative SPOT synthesis mode.

5. Qualitative explorative investigations and quantitative analytical studies

In general, synthetic peptide arrays (prepared by the analytical synthesis mode) can be used for either qualitative exploratory investigations or quantitative analytical studies; the initial choices of array content and assay design often depend on which kind of analysis is required. Qualitative exploratory investigations may involve either simple or comparative functional assays.

In a simple functional assay the non-covalent binding interactions between a single sample and a set of immobilized peptides (probes) are analyzed in parallel to obtain relative peptide-binding

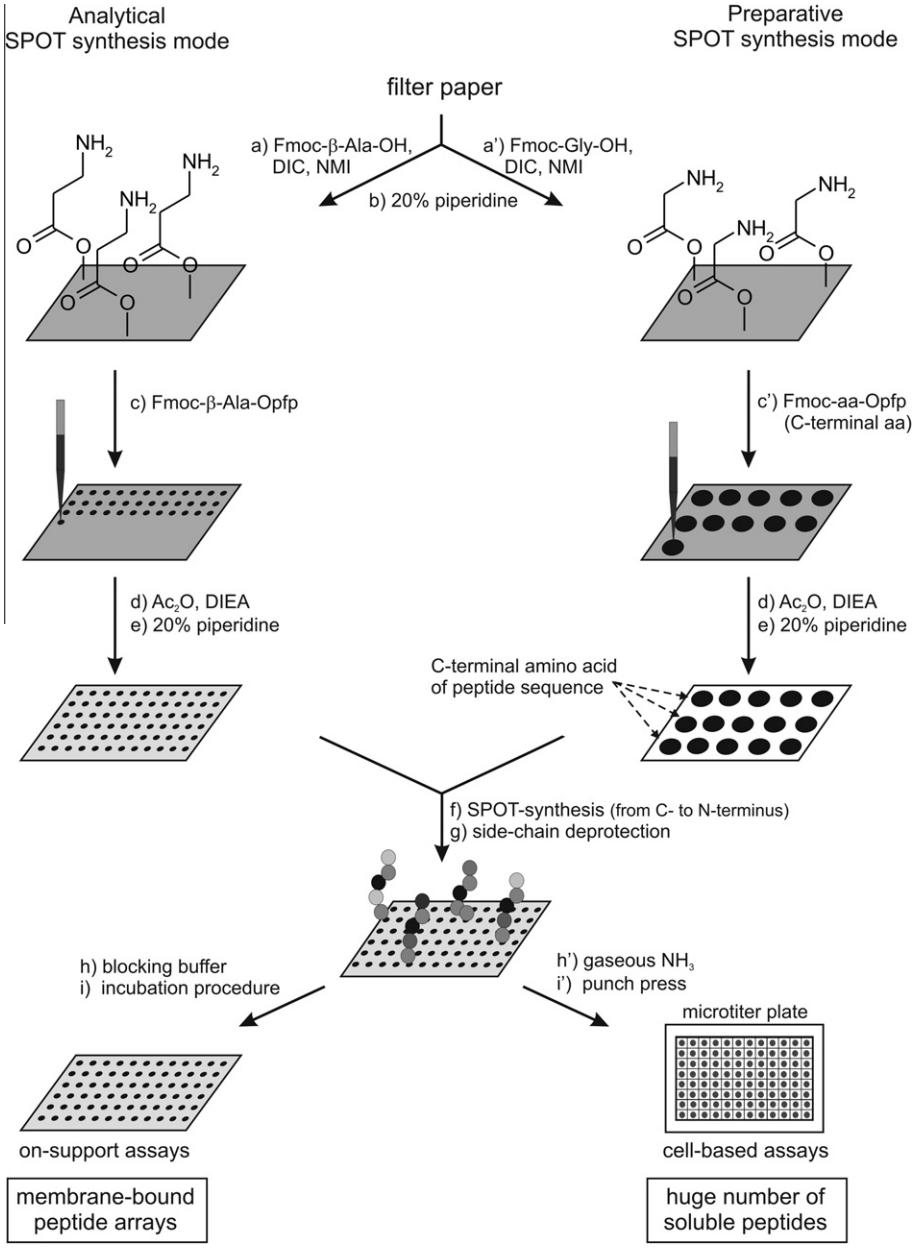

Fig. 1. The principals of SPOT synthesis. The analytical SPOT synthesis mode enables the synthesis of thousands of immobilized cellulose membrane-bound peptides especially for on-support binding studies (left). On the other hand, high numbers of soluble peptides can be generated in sufficient quality and yield by the preparative SPOT synthesis mode suitable for several kinds of solution- and cell-based assays as well as for quality control (right). Fmoc = fluorenylmethyloxycarbonyl, DIC = diisopropylcarbodiimide, NMI = N-methylimidazole, Ac_2O = acetic anhydride, DIEA = N,N-diisopropylethylamine, Opfp = pentafluorphenyl ester.

preferences for the sample. For instance, the sample could be a purified PID, which is challenged with relevant peptide probes collected in a peptide array. These peptides are defined by some shared characteristics, e.g. the targets of a consensus in a sequence data bank such as an array of 12-mer human PPxY sequences [50,51] or an array of human C-terminal protein sequences (''Humlib'') [52]. Further systematic sets of peptides for simple functional assays applied for mapping protein binding are scans of overlapping peptides [53,54], amino acid substitution scans such as the alanine scanning approach [55,56], or (complete) substitution analyses. Here, each amino acid of the original sequence is replaced by all other 19 genetically encoded amino acids. This approach has been used in peptide array technology since the beginning (for more references see [39,57]).

For a simple functional assay, previous knowledge about absolute binding affinity is not required. After performing protein binding experiments by challenging a peptide array with a single protein domain the spot intensities can be measured to yield non-dimensional numerical magnitudes: for instance, in "Boehringer light units" using chemiluminescence as a readout [33] or as non-dimensional values if using densitometry [58]. Comparing intra-array spot signal intensities results in ranking peptide-binding preferences (good binders, medium binders, non-binders) towards a single sample. Unfortunately, no quantitative affinity information is available from a simple functional assay; peptide binding could be in the range of nanomolar, micromolar or even millimolar affinity. However, an impression of affinity can be gained if the binding affinities of selected peptides are quantitatively determined in follow-up studies such as surface plasmon resonance or fluorescence polarization studies. Spot signal intensities can then be correlated to the measured binding affinities, yielding a semi-quantitative binding assay. This approach has been demonstrated with the yeast SH3 interactome [33] and for the CAL PDZ domain interaction [59]. Such follow-up studies are important if peptide arrays are used to develop effective inhibitors or assess the biological relevance of the observed binding events (specific versus unspecific). However, these methods are time consuming and costly since selected peptides must be synthesized by standard solid-phase peptide synthesis before performing exact affinity measurements.

Comparative functional assays are based on inter-array comparison. For example, two peptide arrays with identical content are challenged with two related proteins, e.g. a wild type and related mutant version. Spot signal intensities can then be compared when proteins are applied at identical concentrations and arrays are of equal synthetic quality. The quality of SPOT-synthesized peptides has been investigated by several groups. Takahashi and co-workers [60] reported peptide purity higher than 92%, while Kramer and co-workers [61] reported lower purities. An extended HPLC analysis showed that purities of SPOT-synthesized short peptides of up to 15 amino acids are similar to those synthesized by solid-phase methods in reactors [62]. Ay and co-workers analyzed a huge number of SPOT-synthesized cytomegalovirus deduced non-americ peptides by HPLC/MS and found peptide purity in the range of 50–85% [48]. This is in good agreement with Molina and co-workers reporting peptide purity in the range of 74.4–91.3% [63]. Even longer peptides such as the 34-meric FBP28 WW domain could be SPOT-synthesized with a high quality of 65% purity [64]. Besides the high synthetic peptide quality, equivalent peptide array quality is achieved by applying identical chemical conditions during array synthesis. For this reason we use a special platform for the spot synthesizer, and we strongly recommend taking peptide arrays from the same cellulose membrane (intra-membrane arrays). As far as possible, this ensures generating spots with similar peptide density (peptide concentration in a spot).

Recently, we used such a comparative functional assay to challenge the PQBP WW domain and its Y65C missense mutant with peptide arrays of potential human WW domain binding sequences in order to understand the molecular basis of the Goalbi–Ito–Hall syndrome (GIH) [50]. The two identical peptide arrays were synthesized in parallel on one cellulose membrane. A standard β-alanine membrane [36,62] was used and residues were coupled as amino acid OPfp ester derivatives (triple coupling) since this simplifies the synthesis process, particularly when using SPOT robots. After cleaving the side chain protection groups the peptide arrays were separated by cutting the cellulose membrane. One array was probed with the wild type PQBP1 WW domain, and the other with the Y65C missense mutant (both at identical concentrations).

The inter-array comparison revealed that both PQBP1 WW domains recognize the same peptides, but with different binding strengths: peptide binding of the Y65C mutant was lower than that of wild type. As a conclusion, the GIH syndrome is apparently not ruled by a loss or gain of function but by lower binding affinity of the mutant.

Another comparative functional assay approach is being used to investigate the nature and evolution of SH3 domain interaction networks in the eukaryotes *Schizosaccharomyces pombe*, *Candida albicans*, *Ashbya gossypii* and *S. cerevisiae*. A total of 109 SH3 domains have been investigated so far. Uniform arrays from the same cellulose membrane were used to probe a family of SH3 domains across the four yeast species, and the concentrations of the domains were adjusted accordingly. However, some questions still remain: the low member peptide array used in this study was optimized while investigating the *S. cerevisiae* SH3 domain interaction network [33,65] and it is not clear whether it will also work with the other yeast species; the displayed peptides (proteins) were deduced from the *S. cerevisiae* proteome without knowing whether orthologs, or paralogs exist in the three other yeast species; and finally little is known about the evolution of the deduced peptide (proteins). A preliminary result of this study provided evolutionary domain fingerprints of each SH3 domain, for example as shown for the Abp1-1 SH3 domains of the four yeast species (Fig. 2). Finally, it

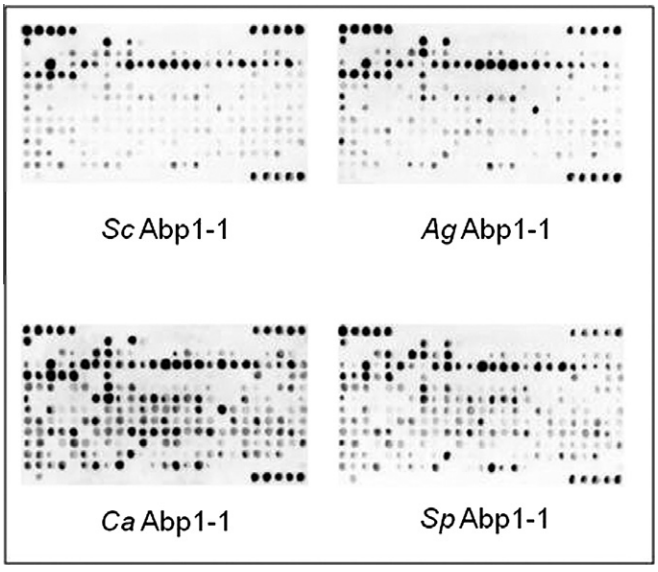

Fig. 2. Evolutionary SH3 domain fingerprints. Four uniform peptide arrays are synthesized on one cellulose membrane. The membrane is cut to yield individual arrays which were then individually probed with Abp1-1 SH3 domains of four yeast species: *Saccharomyces cerevisiae* (*Sc*, upper left), *Ashbya gossypii* (*Ag*, upper right), *Candida albicans* (*Ca*, lower left) and *Schizosaccharomyces pombe* (*Sp*, lower left). Domain concentrations are adjusted and equal assay conditions are applied. In a first glance an individual peptide-binding interaction finger print of each SH3 domain is yielded reflected by the differences in reactive spots. At the upper left, upper right and lower right corner five control spots are placed for better orientation. These spots could not be used for array quality control.

should be mentioned that while intra-membrane arrays are a good choice for repeating simple functional assays, such comparative assays are inadequate for directly calculating quantitative affinity values.

As mentioned above, semi-quantitative binding can be assessed if spot signal intensities can be correlated to measured binding affinities obtained from follow-up studies. However, since such follow-up studies are costly and time consuming an alternative approach would be advantageous. Analytical peptide arrays could perhaps pave the way as real quantitative binding assays. In analogy to a classical ELISA approach, intra-membrane peptide arrays with identical content are probed in parallel with different concentrations of a protein sample. The principle of such an approach is demonstrated using uniform intra-membrane peptide arrays challenged with different concentrations of WW domains and then measuring spot signal intensities (Fig. 3). Plotting spot signal intensities against the applied protein concentrations resulted in sigmoid graphs from which EC_{50} (half effective concentration) values could be calculated for each array peptide (Fig. 3B).

As good laboratory practice the peptide arrays should contain repeated sequences in order to obtain reliable error bars. This has been demonstrated in preliminary work by Weiser et al. [66] where the standard deviation of spot signal intensities measured for several peptide replicas on one membrane varied from 8% to 22%. However, the obtained EC_{50} values were calculated based on the following assumptions: Immobilized peptides were all of the same quality, peptide concentrations were identical at each spot, peptide accessibility was the same for each spot, protein tags did not influence protein-peptide binding, and the measured protein concentration was the effective assay concentration. Obviously, this will never be the case in practice, and strictly speaking, the quantitative EC_{50} values obtained reflect the true dissociation constants (K_D), depending on knowing the quality and concentration of peptides in each spot, as well as the effective concentration of the protein. Nevertheless, as demonstrated above, high peptide synthesis quality, equal peptide density of intra-membrane peptide arrays and the possibility to adjust protein concentration enables calculating EC_{50} values reflecting a binding affinity at a realistic order of magnitude.

Clearly the analytic approach is still under investigation, but we hope our discussion here will encourage others to apply the principle of analytical peptide arrays. At the protein level (protein microarrays) such an approach has been applied to construct a quantitative protein interaction network linking a human proteomic set of SH2 and PTB domains and phosphorylated peptides from ErbB receptor tyrosine kinases [67]. Especially for proteomic approaches, EC_{50} values are extremely helpful for creating quantitative protein interaction networks, and for filling in "the traffic flow information" in Gianni Cesareni's road map of protein interactions.

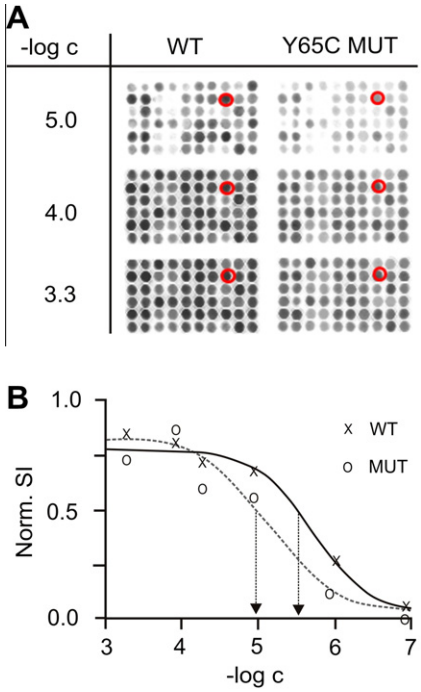

Fig. 3. Principle of an analytical assay. (A) Uniform peptide arrays synthesized in multiple copies on one cellulose membrane. The membrane is cut to yield individual arrays, which are then individually assayed. Here six peptide arrays are shown each displaying the same set of WW domain binding peptides. Left column: arrays were assayed with different concentrations of the wild type (wt) PQBP1 WW domain. Right column: arrays were probed with different concentrations of the PQBP1 Y65C WW domain mutant. Spot signal intensities are measured and as shown spot signal intensities are proportional to the concentration of both WW domains. (B) Exemplarily, calculation of binding affinity is shown for a peptide depict by the red circle. The obtained spot signal intensities (SI) are plotted against domain concentrations and sigmoid graphs for the wild type (solid line) as well as for the Y65C mutant domain (dashed line) are resulted. At the inflection point the binding affinity could be calculated as EC_{50} (half effective concentration) values. As shown by the curve shift of the Y65C mutant the assigned binding affinity is significantly reduced (higher EC_{50} value) compared to the wild type domain.

6. Cellulose membrane modifications and inverted peptide arrays with free C-termini

Cellulose membranes of the ester type [36,62], especially β-alanine and glycine membranes, are standard membranes for SPOT synthesis. Historically, five years after the first report of classical ester-type cellulose membranes, the first amino functionalized ether-type membrane described in 1997 was a cellulose-aminopropyl ether membrane (called CAPE membrane) [68]. CAPE membranes have been predominantly and successfully used for studying SH3 domain interactions [33,54,65,69,70]. They are distinguished by an excellent signal-to-noise ratio during on-support assays due to the extremely low background signal of the membrane itself.

In addition to the membrane type, peptide density is crucial for probing peptide arrays with a protein of interest, and several groups have worked on adjusting amino functionality in order to optimize synthesis or screening [61,71–74]. As a rule of thumb, WW, GYF or BROMO domain interaction studies should be performed on low loaded β-alanine membranes with amino-capacities of 30–120 nmol/cm^2 whereas SH3 domain interaction studies work well on low loaded CAPE membranes.

One kind of protein interaction domain, namely the family of PDZ domains, requires a very special kind of synthetic peptide array. PDZ domains generally recognize the C-terminal four to seven residues of their protein binding partner and require a free C-terminus for ligand recognition. In other words, PDZ domains recognize short linear peptides containing a free C-terminus. Unfortunately, standard SPOT-synthesized peptides lack free C-termini due to their coupling to the cellulose support. The first reliable and robust SPOT synthesis concept for synthesizing inverted peptide arrays with free C-termini was published in 2004 by the Volkmer lab, and the approach was recently improved [52,75]. Such an inverted peptide array requires a special, ether-type cellulose membrane called N-CAPE membrane [76] with amino-loading capacities of 200–800 nmol/cm^2.

More recently, cystic fibrosis research focused on finding a selective CAL PDZ inhibitory peptide by applying an integrated synthetic peptide array approach designed for PDZ domain screening [59]. The screening discovered a peptide that selectively inhib-

its the CAL PDZ domain, which in turn extends the half-life of the ΔF508-CFTR protein responsible for cystic fibrosis. Such an inhibiting peptide might act as a drug or lead structure for drug development, and hence therapy to reduce the adverse effects of cystic fibrosis [77].

7. Combining phage display and SPOT synthesis

A new approach combines the relative strength of selective phage display with the quantitative analysis achieved by SPOT synthesis. Initially, Landgraf et al. [33] applied a strategy to reveal all the peptides in the yeast proteome that have the potential to bind to any domain of interest. Based on the strict consensus sequences identified by phage display, they designed a relaxed consensus; for example the strict consensus sequence of the yeast SH3 domain Rvs167 defined as RxFPRxP was relaxed to R/KxxPxxP. Subsequently, all sequences within the yeast proteome matching a relaxed consensus of a given SH3 domain were identified by computational methods, synthesized on a cellulose membrane, and probed with the SH3 domain of interest. This approach was repeated for eight yeast SH3 domains, readily identifying peptide binding to each domain, and leading to predicting protein partners.

More recently, this approach was extended to the complete SH3 domain interactome of yeast [65]. A consortium comprising the Cesareni, Volkmer, Drubin, Kim, Sidhu, and Boone labs applied a combined approach of orthogonal experimental proteomic tools, such as phage display, yeast two-hybrid and SPOT technology, linked to sophisticated computational and mathematical tools. The results from the three complementary experimental techniques were integrated using a Bayesian algorithm to generate a high confidence yeast SH3 domain interaction map.

8. Concluding remarks

Why have synthetic peptide arrays prepared by SPOT synthesis become so attractive for biologists and the protein domain community? We believe it is due to the robustness, flexibility and simplicity of the economical synthesis method, along with a broad spectrum of well-established and highly sensitive assays (binding affinities detected down to the millimolar range) combined with a semi-quantitative readout of binding affinities. Further good arguments are the opportunities to develop a novel analytical peptide array approach, the high quality of SPOT-synthesized peptides (50–91% purity), a variety of cellulose membranes with diverse physical properties suitable for nearly any kind of binding assay, as well as commercially available membranes (AIMS Scientific Products, Braunschweig, Germany) and equipment (Intavis, Köln, Germany).

However, SPOT technology is limited by the number of peptides that can be synthesized on a membrane of reasonable size, and by the fact that regenerating peptide arrays generally fails. The latter is a severe limitation of SPOT technology, especially for experimental proteomics. Ideally, one would wish to screen a given peptide array several times without any loss of quality, e.g. with different representatives from a protein domain family. One solution could be to use peptide microarrays that could be prepared for a multitude of replicas. Pre-synthesized soluble peptides have been immobilized on glass slides using several methods [78–81]. However, the peptide microarray technique requires expensive equipment and special laboratory conditions. Producing peptide microarrays involves highly parallel and high throughput peptide synthesis, as well as robotic-supported immobilization of pre-synthesized peptide derivatives on glass slides. Hence, SPOT synthesis (preparative mode) is essential for preparing peptide microarrays. Recently peptide microarrays were used for the first time to investigate the protein interaction network mediated by human SH3 domains [82].

The in situ synthesis of high-density peptide microarrays is still a great challenge, and interestingly, the concept of light-directed, spatially addressable synthesis of peptides [37] is once again attractive. The group of Klaus-Peter Stengele at Roche NimbleGen has developed a novel strategy for photolithographic in situ synthesis of thousands of peptides per cm^2 on a glass surface. Finally, to complete the picture, Stadler and co-workers have used a modified color laser printer to "print" the 20 amino acids in the form of solid amino acid toner particles at defined positions on a glass support [83].

Acknowledgements

R.V. thank the Deutsche Forschungsgemeinschaft for their financial support. V.T. appreciates financial support from the FEBS federation. All master and diploma students, PhD students, postdocs and technical assistants of the molecular library and recognition group who cannot be named here personally are thanked for their contribution.

References

[1] Sudol, M. (1998) From Src Homology domains to other signaling modules: proposal of the protein recognition code. Oncogene 17, 1469–1474.
[2] Searls, D.B. (1997) Linguistic approaches to biological sequences. Comput. Appl. Biosci. 13, 333–344.
[3] Searls, D.B. (2002) The language of genes. Nature 420, 211–217.
[4] Phizicky, E., Bastiaens, P.I.H., Zhu, H., Snyder, M. and Fields, S. (2003) Protein analysis on a proteomic scale. Nature 422, 208–215.
[5] Smith, J.M. (1970) Natural selection and the concept of a protein space. Nature 225, 563–564.
[6] Jacob, F. (1994) The Possible and the Actual (Jessie and John Danz Lectures), University of Washington Press, Washington.
[7] Lichtarge, O., Bourne, H.R. and Cohen, F.E. (1996) An evolutionary trace method defines binding surfaces common to protein families. J. Mol. Biol. 257, 342–358.
[8] Gimona, M. (2006) Protein linguistics – a grammar for modular protein assembly? Nat. Rev. Mol. Cell Biol. 7, 68–73.
[9] Fields, S. and Song, O. (1989) A novel genetic system to detect protein–protein interactions. Nature 340, 245–246.
[10] Uetz, P. et al. (2000) A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. Nature 403, 623–627.
[11] Ito, T. et al. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc. Natl. Acad. Sci. USA 98, 4569–4574.
[12] Gavin, A.C. et al. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415, 141–147.
[13] Ho, Y. et al. (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. Nature 415, 180–183.
[14] Santonico, E., Castagnoli, L. and Cesareni, G. (2005) Methods to reveal domain networks. Drug Discov. Today 10, 1111–1117.
[15] Yook, S.H. et al. (2004) Functional and topological characterization of protein interaction networks. Proteomics 4, 928–942.
[16] Das, S. and Smith, T.F. (2000) Identifying nature's protein Lego set. Adv. Protein Chem. 54, 159–183.
[17] Holm, L. and Sander, C. (1998) Dictionary of recurrent domains in protein structures. Proteins 33, 88–96.
[18] Pawson, T. and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains. Science 300, 445–452.
[19] Bork, P., Schultz, J. and Ponting, C.P. (1997) Cytoplasmic signalling domains: the next generation. Trends Biochem. Sci. 22, 296–298.
[20] Cesareni, G. et al., Eds., (2004). Modular Protein Domains, Wiley-VCH Verlags GmbH & KGaA, Weinheim.
[21] Jones, S. and Thorton, J.M. (1996) Principles of protein–protein interactions. Proc. Natl. Acad. Sci. USA 93, 13–20.
[22] Ma, B., Elkayam, T., Wolfson, H. and Nussinov, R. (2003) Protein–protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces. Proc. Natl. Acad. Sci. USA 100, 5772–5777.
[23] Reineke, U., Kramer, A. and Schneider-Mergener, J. (1999) Antigen sequence- and library-based mapping of linear and discontinuous protein–protein-interaction sites by spot synthesis. Curr. Top. Microbiol. Immunol. 243, 23–36.
[24] Bogan, A.A. and Thorn, K.S. (1998) Anatomy of hot spots in protein interfaces. J. Mol. Biol. 280, 1–9.
[25] Clackson, T. and Wells, J.A. (1995) A hot spot of binding energy in a hormone–receptor interface. Science 267, 383–386.
[26] Sudol, M., Chen, H.I., Bougeret, C., Einbond, A. and Bork, P. (1995) Characterization of a novel protein-binding module – the WW domain. FEBS Lett. 369, 67–71.

[27] Fidan, Z., Younis, A., Schmieder, P. and Volkmer, R. (2011) Chemical synthesis of the third WW domain of TCERG 1 by native chemical ligation. J. Pept. Sci. 17, 644–649.
[28] Sparks, A.B., Quilliam, L.A., Thorn, J.M., Der, C.J. and Kay, B.K. (1994) Identification and characterization of Src SH3 ligands from phage-displayed random peptide libraries. J. Biol. Chem. 269, 23853–23856.
[29] Rickles, R.J. et al. (1994) Identification of Src, Fyn, Lyn, PI3K and Abl SH3 domain ligands using phage display libraries. EMBO J. 13, 5598–5604.
[30] Huang, H. and Sidhu, S.S. (2011) Studying binding specificities of peptide recognition modules by high-throughput phage display selections. Methods Mol. Biol. 781, 87–97.
[31] Tonikian, R. et al. (2008) A specificity map for the PDZ domain family. PLoS Biol. 6, E239.
[32] Tong, A.H. et al. (2002) A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. Science 295, 321–324.
[33] Landgraf, C. et al. (2004) Protein interaction networks by proteome peptide scanning. PLoS Biol. 2, E14.
[34] MacBeath, G. and Schreiber, S.L. (2000) Printing proteins as microarrays for high-throughput function determination. Science 289, 1760–1763.
[35] Zhu, H. et al. (2001) Global analysis of protein activities using proteome chips. Science 293, 2101–2105.
[36] Frank, R. (1992) SPOT-synthesis: an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. Tetrahedron 48, 9217–9232.
[37] Fodor, S.P., Read, J.L., Pirrung, M.C., Stryer, L., Lu, A.T. and Solas, D. (1991) Light-directed, spatially addressable parallel chemical synthesis. Science 251, 767–773.
[38] Frank, R. and Schneider-Mergener, J. (2002) in: Peptide arrays on membrane supports – synthesis and applications. (Koch, J. and Mahler, M., Eds.), pp. 1–22. Springer Verlag, Heidelberg.
[39] Volkmer, R. (2009) Synthesis and application of peptide arrays: Quo vadis SPOT technology. ChemBioChem 10, 1431–1442.
[40] Tapia, V., Ay, B., Triebus, J., Wolter, E., Boisguerin, P. and Volkmer, R. (2008) Evaluating the coupling efficiency of phosphorylated amino acids for SPOT synthesis. J. Pept. Sci. 14, 1309–1314.
[41] Rathert, P. et al. (2008) Protein lysine methyltransferase G9a acts on non-histone targets. Nat. Chem. Biol. 6, 344–346.
[42] Filippakopoulos, P. et al. (2012) Histone recognition and large-scale structural analysis of the human bromodomain family. Cell 149, 214–231.
[43] Hoffmann, B., Ast, T., Polakowski, T., Reineke, U. and Volkmer, R. (2006) Transformation of a biologically active peptide into peptoid analogs while retaining biological activity. Protein Pept. Lett. 13, 829–833.
[44] Heine, N. et al. (2003) Synthesis and screening of peptoid arrays on cellulose membranes. Tetrahedron 59, 9919–9930.
[45] Scharn, D., Wenschuh, H., Reineke, U., Schneider-Mergener, J. and Germeroth, L. (2000) Spatially addressed synthesis of amino- and amino-oxy-substituted 1,3,5-triazine arrays on polymeric membranes. J. Comb. Chem. 2, 361–369.
[46] Mahrenholz, C.C., Tapia, V., Stigler, R.D. and Volkmer, R. (2010) A study to assess the cross-reactivity of cellulose membrane-bound peptides with detection systems: an analysis at the amino acid level. J. Pept. Sci. 16, 297–302.
[47] Winkler, D.F. and McGeer, P.L. (2008) Protein labeling and biotinylation of peptides during spot synthesis using biotin p-nitrophenyl ester (biotin-ONp). Proteomics 8, 961–967.
[48] Ay, B. et al. (2007) Sorting and pooling strategy: a novel tool to map a virus proteome for CD8 T-cell epitopes. Biopolymers 88, 64–75.
[49] Kamradt, T. and Volkmer-Engert, R. (2004) Cross-reactivity of T lymphocytes in infection and autoimmunity. Mol. Divers. 8, 271–280.
[50] Tapia, V.E. et al. (2010) Y65C missense mutation in the WW domain of the Golabi–Ito–Hall syndrome protein PQBP1 affects its binding activity and deregulates pre-mRNA splicing. J. Biol. Chem. 285, 19391–19401.
[51] Hu, H. et al. (2004) A map of WW domain family interactions. Proteomics 4, 643–655.
[52] Boisguerin, P. et al. (2004) An improved method for the synthesis of cellulose membrane-bound peptides with free C termini is useful for PDZ domain binding studies. Chem. Biol. 11, 449–459.
[53] Geysen, H.M., Meloen, R.H. and Barteling, S.J. (1984) Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. Proc. Natl. Acad. Sci. USA 81, 3998–4002.
[54] Rottensteiner, H. et al. (2004) Peroxisomal membrane proteins contain common Pex19p-binding sites that are an integral part of their targeting signals. Mol. Biol. Cell 15, 3406–3417.
[55] Cunningham, B.C. and Wells, J.A. (1989) High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. Science 244, 1081–1085.
[56] Bolger, G.B. et al. (2006) Scanning peptide array analyses identify overlapping binding sites for the signalling scaffold proteins, beta-arrestin and RACK1, in cAMP-specific phosphodiesterase PDE4D5. Biochemical J. 398, 23–36.
[57] Reineke, U., Schneider-Mergener, J. and Schutkowski, M. (2005) Peptide arrays in proteomics and drug discovery in: BioMEMS and Biomedical Nanotechnology, Micro and Nano-technologies for Genomics and Proteomics, vol. II (Ozkan, M. and Heller, M.J., Eds.), pp. 161–282, Springer, New York.
[58] Portwich, M. et al. (2007) A network of coiled-coil associations derived from synthetic GCN4 leucine-zipper arrays. Angew. Chem. Int. Ed. 46, 1654–1657.
[59] Vouilleme, L., Cushing, P.R., Volkmer, R., Madden, D.R. and Boisguerin, P. (2010) Engineering peptide inhibitors to overcome PDZ binding promiscuity. Angew. Chem. Int. Ed. 49, 9912–9916.
[60] Takahashi, M., Ueno, A. and Mihara, H. (2000) Peptide design based on an antibody complementarity determining region (CDR): construction of porphyrin-binding peptides and their affinity maturation by a combinatorial method. Chem. Eur. J. 6, 3196–3203.
[61] Kramer, A. et al. (1999) Spot synthesis: observations and optimizations. J. Pept. Res. 54, 319–327.
[62] Wenschuh, H. et al. (2000) Coherent membrane supports for parallel microsynthesis and screening of bioactive peptides. Biopolymers 55, 188–206.
[63] Molina, F., Laune, D., Gougat, C., Pau, B. and Granier, C. (1996) Improved performances of spot multiple peptide synthesis. Peptide Res. 9, 151–155.
[64] Przezdziak, J. et al. (2006) Probing the ligand binding specificity and analyzing the folding state of SPOT synthesized FBP28 WW domain variants. ChemBioChem 7, 780–788.
[65] Tonikian, R. et al. (2009) Bayesian modeling of the yeast SH3 domain interactome predicts spatiotemporal dynamics of endocytosis proteins. PLoS Biol. 7, E1000218.
[66] Weiser, A.A. et al. (2005) SPOT synthesis: reliability of array-based measurement of peptide binding affinity. Anal. Biochem. 342, 300–311.
[67] Jones, R.B., Gordus, A., Krall, J.A. and MacBeath, G. (2006) A quantitative protein interaction network for the ErbB receptors using protein microarrays. Nature 439, 168–174.
[68] Volkmer-Engert, R., Hoffman, B. and Schneider-Mergener, J. (1997) Stable attachment of the HMB-linker to continuous cellulose membranes for parallel solid phase spot synthesis. Tetrahedron Lett. 38, 1029–1032.
[69] Schell-Steven, A. et al. (2005) Identification of a novel, intraperoxisomal Pex14-binding bite in Pex13: Association of Pex13 with the docking complex is essential for peroxisomal matrix protein import. Mol. Cell. Biol. 25, 3007–3018.
[70] Pires, J.R. et al. (2003) The ScPex13p SH3 domain exposes two distinct binding sites for Pex5p and Pex14p. J. Mol. Biol. 326, 1427–1435.
[71] Dürauer, A. et al. (2006) Evaluation of a sensitive detection method for peptide arrays prepared by SPOT synthesis. J. Biochem. Biophys. Methods 66, 45–57.
[72] Gail, R., Frank, R. and Wittinghofer, A. (2005) Systematic peptide array-based delineation of the differential beta-catenin interaction with Tcf4, E-cadherin, and adenomatous polyposis coli. J. Biol. Chem. 280, 7107–7117.
[73] Otvos Jr., L. et al. (2000) In situ stimulation of a T helper cell hybridoma with a cellulose-bound peptide antigen. J. Immunol. Methods 233, 95–105.
[74] Lin, Q., O'Neill, J.C. and Blackwell, H.E. (2005) Small molecule macroarray construction via Ugi four-component reactions. Org. Lett. 7, 4455–4458.
[75] Boisguerin, P. et al. (2007) Characterization of a putative phosphorylation switch: adaptation of SPOT synthesis to analyze PDZ domain regulation mechanisms. ChemBioChem 8, 2302–2307.
[76] Licha, K. et al. (2000) Highly parallel nano-synthesis of cleavable peptide–dye conjugates on cellulose membranes. Tetrahedon Lett. 41, 1711–1715.
[77] Cushing, P.R., Vouilleme, L., Pellegrini, M., Boisguerin, P. and Madden, D.R. (2010) A stabilizing influence: CAL PDZ inhibition extends the half-life of ΔF508-CFTR. Angew. Chem. Int. Ed. 49, 9907–9911.
[78] Beutling, U., Städing, K., Stradal, T. and Frank, R. (2008) Large-scale analysis of protein–protein interactions using cellulose-bound peptide arrays. Adv. Biochem. Eng./Biotechnol. 110, 115–152.
[79] Henderson, G. and Bradley, M. (2007) Functional peptide arrays for high-throughput chemical biology based applications. Curr. Opin. Biotechnol. 18, 326–330.
[80] Tapia, V.E., Ay, B. and Volkmer, R. (2009) Exploring and profiling protein function with peptide arrays. Methods Mol. Biol. 570, 3–17.
[81] Reimer, U., Reineke, U. and Schneider-Mergener, J. (2002) Peptide arrays: from macro to micro. Curr. Opin. Biotechnol. 13, 315–320.
[82] Carducci, M. et al. (2012) The protein interaction network mediated by human SH3 domains. Biotechnol. Adv. 30, 4–15.
[83] Stadler, V. et al. (2008) Combinatorial synthesis of peptide arrays with a laser printer. Angew. Chem. Int. Ed. 47, 7132–7135.