FEBS 23898 FEBS Letters 480 (2000) 49–54

Minireview

Domain repertoires as a tool to derive protein recognition rules

Adriana Zucconi, Simona Panni, Serena Paoluzi, Luisa Castagnoli, Luciana Dente, Gianni Cesareni*

Department of Biology Enrico Calef, University of Rome Tor Vergata, Via Della Ricerca Scientifica, 00133 Rome, Italy

Received 19 May 2000

Edited by Gianni Cesareni

Abstract Several approaches, some of which are described in this issue, have been proposed to assemble a complete protein interaction map. These are often based on high throughput methods that explore the ability of each gene product to bind any other element of the proteome of the organism. Here we propose that a large number of interactions can be inferred by revealing the rules underlying recognition specificity of a small number (a few hundreds) of families of protein recognition modules. This can be achieved through the construction and characterization of domain repertoires. A domain repertoire is assembled in a combinatorial fashion by allowing each amino acid position in the binding site of a given protein recognition domain to vary to include all the residues allowed at that position in the domain family. The repertoire is then searched by phage display techniques with any target of interest and from the primary structure of the binding site of the selected domains one derives rules that are used to infer the formation of complexes between natural proteins in the cell.

Key words: Binding specificity; Molecular repertoire; Phage display; Protein interaction module; SH3 domain

1. Strategies to assemble a protein interaction map

Genomic databases, once translated into protein sequences, can be viewed as puzzle boxes waiting to be assembled into a coherent picture that resembles a cell. Given any protein segment, there are two types of strategies that are helpful in placing the corresponding protein in the appropriate slot in the cell puzzle (Fig. 1).

The first approach is straightforward and comparable to the strategy utilized by 3-year-old children when they tackle their first jigsaw puzzle. It consists of a systematic screening of the limited repertoire of natural peptides in a search for partners that display a complementary surface. This approach does not involve any a priori understanding of the rules governing protein interaction. In principle, this would be the approach of choice but it is limited by the difficulty of obtaining a complete, equally represented, collection of the proteins in a cell. It has been suggested that a high throughput implementation of the yeast two-hybrid method and of MALDI mass spectrometry could be exploited to develop a complete protein interaction map of an organism [1–4].

On the other hand, one could take the alternative approach of deriving a set of rules that eventually would allow one to infer the binding partners from the primary structure of a related query protein. This second, perhaps more general, strategy could be implemented by an experimental approach that permits exploration of all the sequence space in order to extract the subset of sequences that have the potential to bind to the selected bait with an affinity above a certain threshold. In the hypothesis of a single solution, or a small subset of related solutions, to the problem of binding a given protein target, this information can in turn be used to derive the 'ligand signature' and use it to probe the databases of natural peptide sequences in search of potential in vivo partners.

It is hoped that by repeating this approach with a manageable number of baits one learns general rules that can be applied to the whole proteome without the need to perform the selection experiment for each single protein. Although this approach is less direct, it offers the advantage of searching synthetic repertoires where all the elements are equally represented. Furthermore, ligands that bind to the target with affinity higher than the natural ligand may be selected in the process thus providing valuable tools to probe the consequences of disrupting the protein complex in vivo. The limitation of this approach is that a complete search of the sequence space of peptides of reasonable size (say a small domain of 50 amino acids) is far beyond the technical limits imposed by currently available techniques. However, a judicious selection of the sequence space to be explored (see below) might render this approach feasible. This review and the accompanying one [5] will focus on the exploitation of phage display to implement this second strategy.

2. Searching artificial repertoires

Screening of peptide repertoires has been used by several groups as a tool to shed light on the mechanisms underlying specificity of ligand recognition in the assembly of macromolecular complexes inside the cell [6–20]. In particular, phage display libraries of relatively short peptides (6–15 residues) of random sequence displayed on either pIII or pVIII have been panned with different protein recognition domains in order to deduce the sequence of a consensus ligand. An accompanying review [5] reports these results in detail and discusses the implications of the finding that, often, these experiments yield biologically relevant ligands. That is, the derived consensus sequences, when used as templates to search a protein data-

*Corresponding author. Fax: (39)-6-2023500.
E-mail: cesareni@uniroma2.it

0014-5793/00/$20.00
PII: S0014-5793(00)01777-4

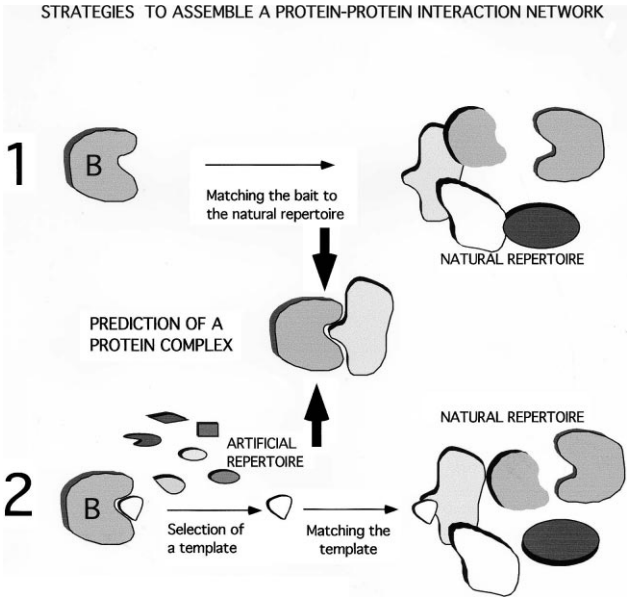

Fig. 1. The cartoon illustrates two strategies that may be used to assemble the protein–protein interaction network in the cell. The structure indicated with a capital B is the bait.

base, often retrieve partners that can be demonstrated to be physiologically relevant interactors by independent techniques.

3. Protein–protein interaction is often asymmetric

The success of the published experiments that make use of random peptide repertoires raises the question of whether this can be considered a general approach to describe protein recognition specificity.

In our group, using a synthetic nonapeptide library displayed on the major capsid protein of filamentous phage [21], we have scored a success rate of approximately 50%. Approximately half of the protein domains that we have tested find, in a random library of 10^8 linear peptides, a ligand that binds in the high micromolar affinity range. Since our library contains peptides with all the possible combinations of residues in any six out of the nine randomized positions, fewer than six key residues are sufficient to confer to a nonapeptide the capacity to bind to this class of domains with at least 100 μM affinity.

Our conclusion is that the molecular surfaces involved in protein–protein recognition behave differently when tested in our panning assay. Some succeed in finding short ligands while others do not. Antibodies, for instance, in the vast majority of cases find a partner in these repertoires [22–26]. Furthermore, the deduced *consensus* sequence is often related to the sequence in the antigen that was used to raise the antibody. This means that, although from the crystallographic point of view, most epitopes are conformational as they involve interaction with regions of the protein that are far apart in the primary sequence [27], a large fraction of the binding energy is contributed by the binding to a relatively short linear peptide. Whenever the epitope is purely conformational or it is non-peptidic in nature, the antigen binding site selects a peptide that mimics the conformational epitope [28,29].

When considering protein–protein interaction in general, the two models that we have in mind are represented in the cartoon in Fig. 2. In the first interaction type the two partners contribute equally, with a relatively large, more or less flat, surface. The second type, on the contrary, is asymmetric with one of the two partners (R, receptor) forming a molecular pocket that can host a relatively simple structure of the second partner (L, ligand). Partner R has a more complex involvement since the design of this pocket is likely to require the structural contribution of relatively distant residues and structures. In contrast, the participation of partner L is less complex, from a structural point of view, and consists in a relatively short peptide that makes specific contacts in the pocket.

Our panning test allows us to make a clear operational distinction between the two interaction types, although in nature no clear threshold is likely to exist. In the case of a symmetric interaction, neither of the two partners will find ligands in our peptide collection since both binding surfaces are extended and cannot be reduced to the simple linear peptides that form our repertoire. By contrast, in an asymmetric interaction, one of the two partners, the receptor, will be able to extract a family of ligands from the peptide library while the other, the ligand, will not.

This discrimination of protein surfaces, operationally defined by the ability of the protein to select a clone from a repertoire of short peptides, depends on the complexity of the repertoire itself. If we were able to assemble and screen a repertoire containing all the possible sequences of a peptide of 100 amino acids, all the proteins would be classified as 'receptors' according to this criterion.

Summarizing, one can exploit short linear peptide repertoires displayed by filamentous phage only to investigate protein recognition mediated by 'receptor' proteins, that is by proteins or domains that have evolved a molecular pocket that can bind to relatively short linear peptides. This strategy can help inferring the chemical characteristics of the ligands of any receptor molecule.

4. Artificial domain repertoires

Since the available techniques do not permit the assembly of a complete repertoire of peptides of the size of an average receptor domain, this approach cannot be easily extended to the identification of the preferred receptor(s) of any given natural peptide. Fortunately the problem of complexity is somewhat reduced by the recognition, supported by numerous reports in the last decade, that the structure of natural proteins is often modular. Many of these module families are found repeated several times in the proteome and are often

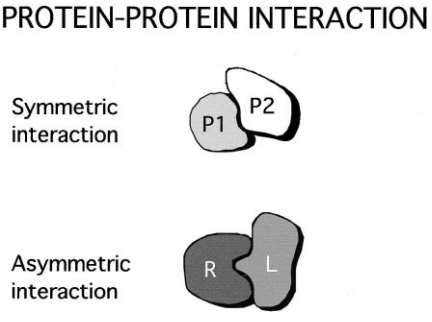

Fig. 2. Schematic representation of symmetric and asymmetric binding in protein–protein interaction.

involved in protein recognition [30,31]. This is simplified by the structure of the Src kinase where one can identify three domains (SH1, SH2, SH3) that are found repeatedly in a number of other proteins [32,33]. Although new modules, proposed to be involved in protein interaction, are continuously reported in the literature, it is clear that their total number does not exceed a few hundred. In any case they are far from including all the sequence and/or structure space. At the moment the SMART database includes approximately 400 entries and a sizeable fraction of these motifs are involved in protein recognition [34]. Thus, the fraction of peptides in the sequence space that have had any chance of experiencing natural selection is small and those that have been selected tend to cluster in a small number of defined regions of the sequence/structure space.

These considerations lead us to a second approach, complementary to the one that aims at building artificial peptide repertoires as large and as complete as possible. Instead of using a very general, but inevitably incomplete repertoire, we aim at mirroring the natural proteome which, for the scope of modulating protein interaction, has selected a limited number of different scaffolds, each specialized in recognizing specific molecular characteristics of the protein surfaces. For instance the SH3 (or WW) scaffold to bind poly-proline peptides, the SH2 (or PTB) scaffold for phospho-tyrosine peptides and the PDZ scaffold for carboxy-terminal peptides.

One could think of assembling a limited number of different domain libraries by randomizing the ligand binding surface of one representative domain for each of the natural families. Since, in this case, the ligand sequence space that we are aiming for is much more restricted, the problem becomes accessible. Furthermore we can learn from the molecular scaffolds that have been selected during evolution and we are likely to have a better chance of success in assembling a collection of repertoires that is complete, and contains receptors for any given physiologically significant peptide ligand.

The question 'which peptides in the sequence space bind to the query peptide' can then be rephrased into two simpler questions: (i) 'which subset of peptide sequence space represented in the proteome (domain repertoire) contains a ligand for a query peptide?' and (ii) 'which element of this restricted repertoire has the potential to bind to the query peptide?' The answer, obtained by panning the different domain repertoires, could provide hints about the possible natural ligands and help in the development of specific protein interaction inhibitors, 'perturbagens', to be used to interfere with cell pathways [35–40].

5. A general repertoire to bind poly-proline peptides

To illustrate the approach in more detail, we will describe the strategy that we have utilized to design a collection of molecules that bind to poly-proline containing peptides (Fig. 3, Panni et al., in preparation). The selected scaffold consists of the four β-strand sandwich that characterizes the SH3 domain family [41]. An alternative scaffold evolved to bind proline-rich peptides is the WW domain scaffold [7,42].

SH3 domains bind to their targets by accommodating the ligand in a binding surface formed by three molecular pockets [43–45]. Two of them are hydrophobic and host the PXXP motif which is considered the mark of SH3 ligands. The third pocket is normally negatively charged and determines the specificity and the orientation of the ligand by hosting a positively charged residue that either precedes or follows the PXXP core motif. Peptides characterized by the RXXPXXP motif bind in an orientation called 'class I' orientation, while peptides displaying the PXXPXR motif accommodate the carboxy-terminal R in the hydrophilic pocket and bind in the opposite orientation (class II) [45].

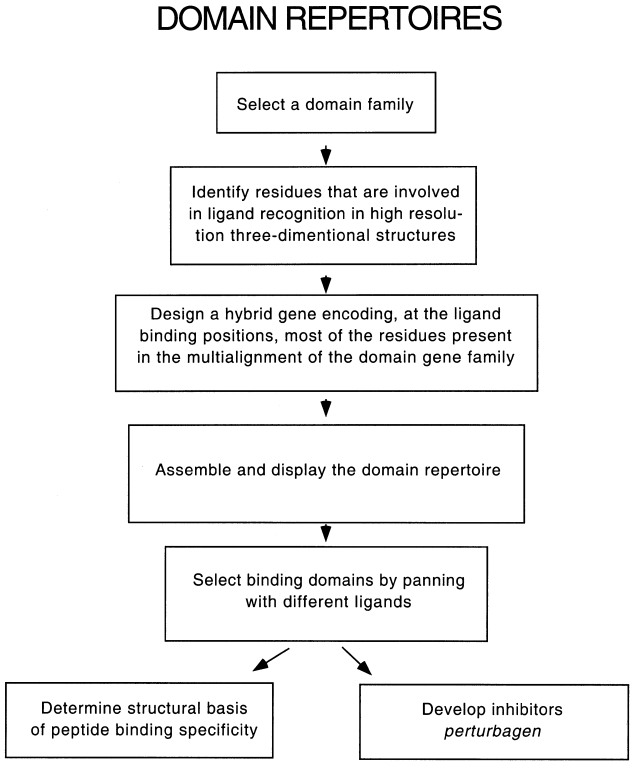

Fig. 3. Strategy to design, assemble and exploit a domain repertoire.

More than 1000 SH3 domains can be identified in protein sequence databases [34] and those that have already been characterized are known to bind to poly-proline stretches contained in a variety of proteins involved in signal transduction, cytoskeleton organization and membrane traffic [46]. Poly-proline sequences are relatively common in the proteome. For instance, yeast proteins containing a PXXP core motif, which identifies potential targets for SH3 domains, are 3300 out of a total 6148 predicted open reading frames [47]. We decided to design a molecular repertoire that would contain elements that could bind to any of these potential targets with sufficient affinity and specificity.

In order to assemble an SH3 domain repertoire, we selected a structurally well characterized SH3 domain scaffold, the one belonging to the protein Abl [43], and identified all the residues that form the binding surface and are in contact with a peptide ligand. By varying the residue composition of the binding surface in a combinatorial fashion, one should obtain a collection of SH3 molecules with a wide specificity spectrum. However, because of the large number of possible combinations, such a repertoire would be largely incomplete and would probably include many elements that do not fold properly. Thus, we restricted the variation of the chemical characteristic of the binding surface to the residues that are found at the corresponding positions in natural SH3 domains.

Two hundred and ninety SH3 domains in the SMART

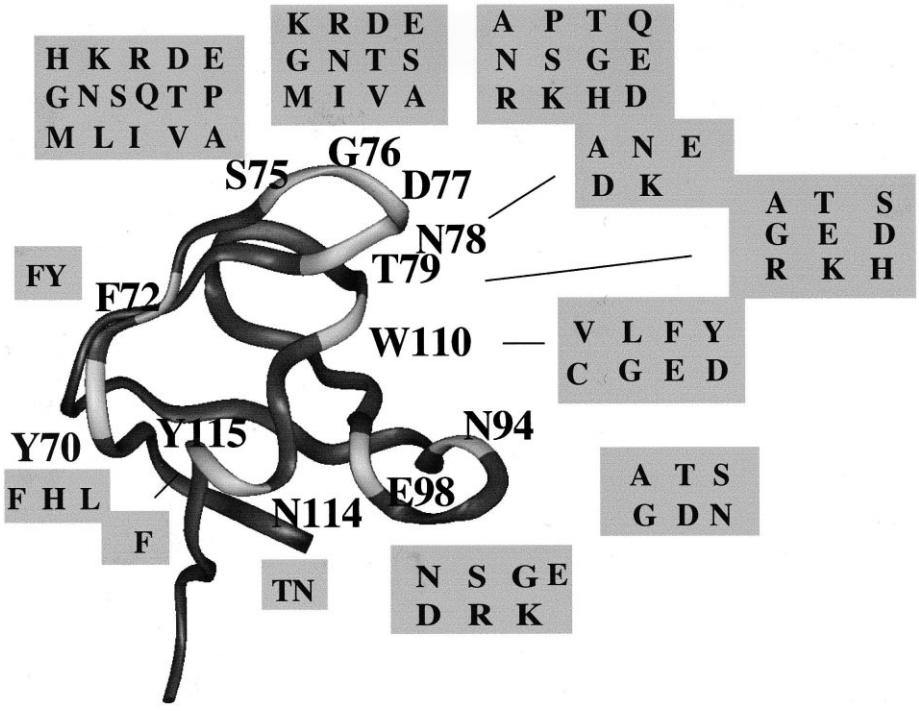

Fig. 4. SH3 repertoire. The three-dimensional structure of the SH3 domain of the Abl kinase is represented as a ribbon. The scaffold is shaded in dark while the residues that have been shown to make contact with the target peptides are in a lighter shade. These include 13 residues: Y70, F72, S75, G76, D77, N78, T79, N94, E98, W99, W110, N114, Y115. The letters in the gray boxes represent all the amino acids that are encoded by a gene of degenerate sequence in the corresponding positions.

database were aligned and the frequencies of all the residues at each of the 13 sites that make contact with the target peptide were determined. Some of these positions are very conserved. At position 110 the vast majority of SH3s have Trp while in the RT loop (S75–T79) almost all residues are represented in natural SH3s. Then we designed a degenerate gene sequence that encodes hybrid SH3s displaying, on the Abl scaffold, a collection of binding surfaces that include most of those that are found in natural SH3 domains.

The selection of residues that are allowed at each position is somewhat limited and biased by the characteristics of the genetic code. However, a judicious choice of the degeneracy of the nucleotide sequence makes it possible to assemble a repertoire that contains most of the residues present in natural SH3 in the peptide binding surface.

The residues that are allowed, at each position, by the sequence of the degenerate gene that we have designed are reported in Fig. 4. The total number of possible SH3 domains in this combinatorial repertoire is 1.5×10^9, while the actual size of the library that we have constructed is 10^8.

The first question that we wanted to address is whether this collection of artificial SH3 domains contain elements that can bind to most of the peptides folding in poly-proline helix II conformation. To answer this question the SH3 repertoire, displayed by fusion to the carboxy-terminus of the λ D protein [48,49], was panned with several peptides that were used as baits.

These experiments, whose results are summarized in Fig. 5, and similar results obtained with different peptide baits, indicate that most, possibly all, poly-proline peptides that fold in a helix II conformation have a ligand partner in our SH3 domain library. Furthermore, SH3 domains that have a high degree of specificity can be selected by this approach. We have also shown that, by using as baits folded proteins exposing poly-proline peptides (for instance the HIV protein Nef), one can select partners that have high affinity and specificity (Panni et al., in preparation). By this approach one can identify the residues that are involved in binding specificity and extract recognition rules that help in associating any poly-proline peptide to a set of specific residues in the SH3 binding surface. The assumption is that the specific scaffold that is selected, in this case the Abl-SH3 domain, plays a minor role in ligand recognition and that residues that are buried in the domain structure, or far away from the ligand, can be, in a first approximation, neglected.

The above mentioned experiments were done by panning in vitro an SH3 repertoire displayed by fusion to the COOH-terminus of the D protein of bacteriophage λ. The same repertoire, however, can be expressed in vivo, in practically any organism, by inserting the degenerate gene sequence into a suitable expression vector. Thus, by exploiting appropriate selection systems, the repertoire can be used to isolate elements that inhibit biological pathways by interacting with proteins that display a poly-proline helix on their surface [35–40].

Similar experiments were carried out by Schneider and collaborators [50] who utilized the two-hybrid selection system to extract artificial PDZ domains recognizing different carboxy-terminal peptides from a PDZ degenerate library. Furthermore the same authors demonstrated that such artificial PDZ domains can be used to target intracellular proteins to different subcellular compartments. Another domain, whose binding specificity can be changed by minor alterations of the ligand binding residues, is the WW domain [51].

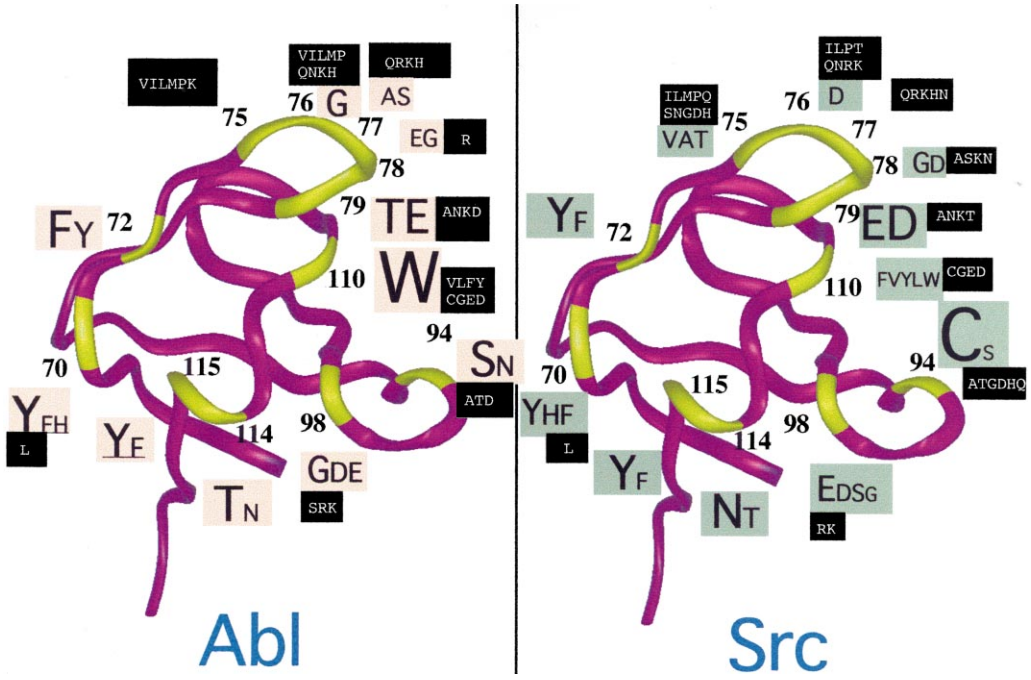

Fig. 5. Sequence characteristics of SH3 domains that bind to Src-type or Abl-type peptides. We have characterized approximately 20 domains that bind to a peptide that displays a consensus similar to peptides recognized by the Abl SH3 domain (left) or by the Src SH3 domain (right). The font size in the gray boxes are proportional to the frequency of occurrence of the specific residue at that variable position of the repertoire in the selected domains. The residues that, although present in the repertoire, were never found in the selected domains are in black boxes. The first peptide (Abl) is homologous to many peptides recognized by the Abl wild type SH3 while the second (Src) is a preferred target for SH3 domains of Src-type kinase. Both peptides, after a couple of panning cycles, could select ligands from the SH3 domain repertoire. The two sets of SH3 domains, however, were distinct and SH3 domains that recognize the Abl peptide would not bind to the Src peptide and vice versa. By comparing the sequence in the degenerate regions of the SH3 binding to the Abl or to the Src peptide, it was possible to identify regularities. For instance, all the SH3 domains that bind to the Abl peptide had a Trp at position 110 while most hydrophobic residues were tolerated at this position by Src peptide ligands. By contrast a Cys was strongly preferred at position 94 in Src peptide ligands.

Thus we conclude that general repertoires of protein binding domains can be assembled by randomizing specific positions of the domain ligand binding surface. These repertoires can be searched, for instance by phage display techniques, with any bait of interest and from the chemical characteristics of the selected domains one can learn rules underlying protein recognition specificity and utilize these rules to predict ligands of newly discovered natural domains.

6. Perspectives

The last decade has witnessed a flourishing of new tools to study protein–protein interaction [52,53]. The speed with which binding information is acquired by these methods is such that, whatever the scientific interest, most groups have adopted and are using them to obtain hints about the protein interaction web in the specific physiological niche of their interest. The impact of these tools in molecular biology research is such that one can hardly come across grant or fellowship applications where these methods are not mentioned as important tools in the proposal. Some recent reports or unpublished work, which we have described here, suggest that the construction and panning of artificial domain repertoires can help decipher the rules underlying protein recognition specificity mediated by small structurally conserved protein modules. These rules in turn can be used to identify putative targets of newly discovered elements of these domain families. Sometimes these rules are simple and can be easily incorporated in a pattern search algorithm. For instance in the case of the WW domain a single residue in the target recognition surface was shown to mediate ligand specificity [51] (Kasanov and Kay, personal communication). By contrast, extensive work on the SH3 domain has not revealed a small number of residues that make it possible to infer ligand preference in a straightforward manner. Also in this case, however, the data collected from panning both peptide or domain repertoires can be organized in a domain-specific residue to residue contact matrix to be used with a score function to evaluate the likelihood that a given domain would bind to any specific poly-proline peptide [54].

The general consensus is that eventually only 'traditional', and time consuming, genetic experiments will provide the definite proof that the interactions, discovered by these artificial tools, have any functional relevance in vivo. However, the large number of reports that describe the successful exploitation of display methods in the discovery of new interaction partners has established them as unavoidable approaches to restrict the number of putative ligands. One must realize, however, that the intrinsic limit of these methods is that they only provide information about the affinity of the two putative partners under investigation. Affinity is only one of the factors that determine whether a specific complex is formed in vivo. Local protein concentration mediated by covalent linkage to localization signals or other protein interaction domains often plays a more important role than absolute affinity. Nevertheless, a 'complete', biochemically naive, interaction map, where each link established between two proteins by any of the methods described in this issue is associated

with a 'reliability coefficient', would be very much welcomed by the community of biologists and by commercial companies looking for new drug targets. Some recently developed computational tools for genome-wide prediction of protein interaction maps could also contribute to the assessment of such coefficients [55–58].

Acknowledgements: The work described in this manuscript is supported by the Italian Association for Cancer Research (AIRC), Telethon and the Istituto Superiore di Sanità (Programma Nazionale per la Ricerca sull'AIDS). A.Z. and S.P. are supported by a F.I.R.C. fellowship.

References

[1] Bartel, P.L., Roecklein, J.A., SenGupta, D. and Fields, S. (1996) Nature Genet. 12, 72–77.
[2] Fromont-Racine, M., Rain, J.C. and Legrain, P. (1997) Nature Genet. 16, 277–282.
[3] Andersen, J.S. and Mann, M. (2000) FEBS Lett. 480, 25–31.
[4] Legrain, P. and Selig, L. (2000) FEBS Lett. 480, 32–36.
[5] Kay, B.K., Kasanov, J., Knight, S. and Kurakin, A. (2000) FEBS Lett. 480, 55–62.
[6] Sparks, A.B., Quilliam, L.A., Thorn, J.M., Der, C.J. and Kay, B.K. (1994) J. Biol. Chem. 269, 23853–23856.
[7] Linn, H., Ermekova, K.S., Rentschler, S., Sparks, A.B., Kay, B.K. and Sudol, M. (1997) Biol. Chem. 378, 531–537.
[8] Bunnell, S.C., Henry, P.A., Kolluri, R., Kirchhausen, T., Rickles, R.J. and Berg, L.J. (1996) J. Biol. Chem. 271, 25646–25656.
[9] Rickles, R.J., Botfield, M.C., Weng, Z., Taylor, J.A., Green, O.M., Brugge, J.S. and Zoller, M.J. (1994) EMBO J. 13, 5598–5604.
[10] Rickles, R.J., Botfield, M.C., Zhou, X.M., Henry, P.A., Brugge, J.S. and Zoller, M.J. (1995) Proc. Natl. Acad. Sci. USA 92, 10909–10913.
[11] Mosser, E.A., Kasanov, J.D., Forsberg, E.C., Kay, B.K., Ney, P.A. and Bresnick, E.H. (1998) Biochemistry 37, 13686–13695.
[12] Yamabhai, Y., Hoffman, N.G., Hardison, N.L., McPherson, P.S., Castagnoli, L., Cesareni, G. and Kay, B.K. (1998) J. Biol. Chem. 273, 31401–31407.
[13] Paoluzi, S. et al. (1998) EMBO J. 17, 6541–6550.
[14] Salcini, A.E. et al. (1997) Genes Dev. 11, 2239–2249.
[15] Dente, L., Vetriani, C., Zucconi, A., Pelicci, G., Lanfrancone, L., Pelicci, P.G. and Cesareni, G. (1997) J. Mol. Biol. 269, 694–703.
[16] Cestra, G. et al. (1999) J. Biol. Chem. 274, 32001–32007.
[17] Mongiovì, A.M., Romano, P.R., Panni, S., Mendoza, M., Wong, W.T., Musacchio, A., Cesareni, G. and Di Fiore, P. (1999) EMBO J. 18, 5300–5309.
[18] Songyang, Z. et al. (1993) Cell 72, 767–778.
[19] Songyang, Z. et al. (1994) Mol. Cell. Biol. 14, 2777–2785.
[20] Songyang, Z. et al. (1997) Science 275, 73–77.
[21] Felici, F., Castagnoli, L., Musacchio, A., Jappelli, R. and Cesareni, G. (1991) J. Mol. Biol. 222, 301–310.
[22] Scott, J.K. and Smith, G.P. (1990) Science 249, 386–390.
[23] Cwirla, S.E., Peters, E.A., Barrett, R.W. and Dower, W.J. (1990) Proc. Natl. Acad. Sci. USA 87, 6378–6382.
[24] Vaccaro, P. et al. (1997) Brain Res. 52, 1–16.
[25] Iannolo, G., Minenkova, O., Gonfloni, S., Castagnoli, L. and Cesareni, G. (1997) Biol. Chem. 378, 517–521.
[26] Cesareni, G., Castagnoli, L. and Cestra, G. (1999) Comb. Chem. High Throughput Screen. 2, 1–17.
[27] Braden, B.C., Goldman, E.R., Mariuzza, R.A. and Poljak, R.J. (1998) Immunol. Rev. 163, 45–57.
[28] Luzzago, A., Felici, F., Tramontano, A., Pessi, A. and Cortese, R. (1993) Gene 128, 51–57.
[29] Phalipon, A., Folgori, A., Arondel, J., Sgaramella, G., Fortugno, P., Cortese, R., Sansonetti, P.J. and Felici, F. (1997) Eur. J. Immunol. 27, 2620–2625.
[30] Pawson, T. and Scott, J.D. (1997) Science 278, 2075–2080.
[31] Sudol, M. (1998) Oncogene 17, 1469–1474.
[32] Biscardi, J.S., Tice, D.A. and Parsons, S.J. (1999) Adv. Cancer Res. 76, 61–119.
[33] Mayer, B.J. (1997) Curr. Biol. 7, R295–R298.
[34] Schultz, J., Milpetz, F., Bork, P. and Ponting, C.P. (1998) Proc. Natl. Acad. Sci. USA 95, 5857–5864.
[35] Cohen, B.A., Colas, P. and Brent, R. (1998) Proc. Natl. Acad. Sci. USA 95, 14272–14277.
[36] Colas, P., Cohen, B., Jessen, T., Grishina, I., McCoy, J. and Brent, R. (1996) Nature 380, 548–550.
[37] Caponigro, G., Abedi, M.R., Hurlburt, A.P., Maxfield, A., Judd, W. and Kamb, A. (1998) Proc. Natl. Acad. Sci. USA 95, 7508–7513.
[38] Hannon, G.J., Sun, P., Carnero, A., Xie, L.Y., Maestro, R., Conklin, D.S. and Beach, D. (1999) Science 283, 1129–1130.
[39] Souroujon, M.C. and Mochly-Rosen, D. (1998) Nature Biotechnol. 16, 919–924.
[40] Norman, T.C. et al. (1999) Science 285, 591–595.
[41] Musacchio, A., Gibson, T., Lehto, V.P. and Saraste, M. (1992) FEBS Lett. 307, 55–61.
[42] Sudol, M., Chen, H.I., Bougeret, C., Einbond, A. and Bork, P. (1995) FEBS Lett. 369, 67–71.
[43] Musacchio, A., Saraste, M. and Wilmanns, M. (1994) Nature Struct. Biol. 1, 546–551.
[44] Yu, H., Chen, J.K., Feng, S., Dalgarno, D.C., Brauer, A.W. and Schreiber, S.L. (1994) Cell 76, 933–945.
[45] Lim, W.A., Richards, F.M. and Fox, R.O. (1994) Nature 372, 375–379.
[46] Kay, B.K., Williamson, M.P. and Sudol, M.J. (2000) FASEB J. 14, 231–241.
[47] Cherry, J.M. et al. (1998) Nucleic Acids Res. 26, 73–79.
[48] Santini, C., Brennan, D., Mennuni, C., Hoess, R.H., Nicosia, A., Cortese, R. and Luzzago, A. (1998) J. Mol. Biol. 282, 125–135.
[49] Mikawa, Y.G., Maruyama, I.N. and Brenner, S. (1996) J. Mol. Biol. 262, 21–30.
[50] Schneider, S. et al. (1999) Nature Biotechnol. 17, 170–175.
[51] Espanel, X. and Sudol, M. (1999) J. Biol. Chem. 274, 17284–17289.
[52] Phizicky, E.M. and Fields, S. (1995) Microbiol. Rev. 59, 94–123.
[53] Mendelsohn, A.R. and Brent, R. (1999) Science 284, 1948–1950.
[54] Brannetti, B., Via, A., Cestra, G., Cesareni, G. and Citterich, M.H. (2000) J. Mol. Biol. 298, 313–328.
[55] Marcotte, E.M., Pellegrini, M., Ng, H.L., Rice, D.W., Yeates, T.O. and Eisenberg, D. (1999) Science 285, 751–753.
[56] Pellegrini, M., Marcotte, E.M., Thompson, M.J., Eisenberg, D. and Yeates, T.O. (1999) Proc. Natl. Acad. Sci. USA 96, 4285–4288.
[57] Marcotte, E.M., Pellegrini, M., Thompson, M.J., Yeates, T.O. and Eisemberg, D. (1999) Nature 402, 83–86.
[58] Enright, A.J., Iliopoulos, I., Kyrpides, N.C. and Ouzounis, C.A. (1999) Nature 402, 86–90.