doi:10.1006/jmbi.1999.3487 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2000) 296, 821–832

JMB

Sequential Determination of Ligands Binding to Discrete Components in Heterogeneous Mixtures by Iterative Panning and Blocking (IPAB)

Bradley T. Messmer[1]*, Craig J. Benham[2] and David S. Thaler[1]*

[1]*Sackler Laboratory of Molecular Genetics and Informatics, Rockefeller University, 1230 York Ave New York, NY 10021-6399, USA*

[2]*Biomathematical Sciences Department, Box 1023, Mount Sinai School of Medicine, 1 Gustave L. Levy Place, New York, NY 10029, USA*

**Corresponding authors*

Biopanning has been used extensively in conjunction with purified components, but there are also examples in which mixtures of targets have been investigated. This study introduces a methodological innovation, termed iterative panning and blocking (IPAB), to extend the range of specific interactions that can be probed in mixtures. Here this procedure is used to probe a mixture of high molecular mass components of human cord blood with phage-peptide display libraries. The initial panning recovered phage that bore the consensus motif Gly-Pro-Arg-Pro, a known fibrinogen-binding motif. These phage bound specifically to purified fibrinogen. A series of peptides containing the Gly-Pro-Arg-Pro motif efficiently blocked the binding of phage having the same motif, presumably by binding to their common target. A second round of panning was performed against the same target mixture in the presence of this blocking peptide. Phage recovered from this second panning exhibited a motif (Ser-His-Tyr) that was subsequently shown to bind specifically to complement component C1q. A second peptide containing this motif specifically blocked the interaction of the phage with C1q. A third round of panning performed in the presence of both the fibrinogen- and the C1q-blocking peptides yielded phage with a new peptide motif (Asn-Pro-Phe) that also bound specifically to C1q, apparently at a new site. The three motifs isolated through this iterative process were distinct in that each was blocked only by its corresponding peptide. This IPAB strategy can be applied to many high diversity selection procedures that target complex mixtures.

Keywords: phage display; selection strategy; panning; C1q; fibrinogen

Introduction

The fundamental processes of life take place in complex milieus of interacting molecular shapes (Goodsell, 1998; Pauling, 1945). Molecular shape interactions are typically studied in a pair-wise fashion with purified components. Some of the richness of the interactions occurring in living systems may be lost in this pairwise approach even though, in other respects, experimental clarity is gained. Additional information might be obtained by using experimentally precise methods that allow the study of shape interactions in complex mixtures. In conjunction with specific methodologies, a theory must be developed to understand their abilities and limitations when used to probe for specific pairwise interactions in mixtures.

Phage peptide display libraries are a powerful tool for investigating biomolecular interactions, allowing targets to select specific binding motifs from among a high diversity of different peptides (Scott & Smith, 1990; Smith *et al.*, 1998; Smith & Scott, 1993). Peptide display libraries have been used to identify peptide motifs that bind to a particular target. These targets can include antibodies (Bonnycastle *et al.*, 1996; Cook *et al.*, 1998; McConnell *et al.*, 1994; Messmer *et al.*, 1999; Pieczenik, 1999; Stephen *et al.*, 1995), proteins (Lauvrak *et al.*, 1997; Sahu *et al.*, 1996), carbohydrates (Peletskaya *et al.*, 1996), plastic polymer (Adey *et al.*, 1995; Gebhardt *et al.*, 1996), or nucleic

Abbreviations used: IPAB, iterative panning and blocking; mAb, monoclonal antibody; PBS, phosphate buffered saline.

E-mail addresses of the corresponding authors: messmeb@rockvax.rockefeller.edu; thalerd@rockvax.rockefeller.edu

0022-2836/00/030821–12 $35.00/0

acids (Cheng *et al*., 1996; Rebar *et al*., 1996). When seeking ligands that bind to a specific molecule, it is common practice to use the isolated and purified molecule as the target for ''bio-panning'' of the library. On the other hand, peptide motifs that bind to members of multi component mixtures have been successfully recovered. These include polyclonal antibodies (Cortese *et al*., 1994; Germaschewski & Murray, 1996; Tchernychev *et al*., 1997), cell surfaces (Mazzucchelli *et al*., 1999; Szardenings *et al*., 1997), and tissue (Rajotte *et al*., 1998; Van Ewijk *et al*., 1997). These studies focused on finding peptides that mimic a particular antigen, bind a particular surface molecule, or stain a particular tissue type, respectively.

We are interested in developing methods for panning mixtures containing many potential target molecules in order to identify motifs that bind to distinct components of the mixture. A theoretical analysis of procedures for panning libraries against mixtures of targets has been presented in the context of the nucleic acid library SELEX (systematic evolution of ligands by exponential enrichment) protocol (Tuerk & Gold, 1990). This analysis also applies to other library panning procedures in which one target-ligand interaction neither inhibits nor facilitates another. Its results suggest that panning against mixtures should not be significantly different from panning each component or target site separately (Vant-Hull *et al*., 1998). However, this expected parallel, independent enrichment of separate target-ligand pairs only happens occasionally (Morris *et al*., 1998). The more common experience is that one particular target-ligand pair dominates the final pool. Sequencing a reasonable number of members of the final pool often reveals only one family of ligands that bind to a single component of the original mixture.

Different approaches must be used to examine the range of ligand-target interactions in each case. If a complex final mix is found, it must be ''deconvoluted'' using molecular biological techniques such as crosslinking to identify targets for particular ligands from the products of selection (Morris *et al*., 1998). If the final mix is dominated by a single family of ligands that binds to one target, then different strategies are needed. Subtraction involves depleting the library of unwanted ligands prior to panning, while counterselection conducts the panning in the presence of decoys that engage these unwanted ligands (Van Ewijk *et al*., 1997).

Subtractive approaches are most useful when looking for the difference between two roughly isogenic or otherwise very similar targets or mixtures of targets. For example, phage displayed antibody fragments specific for Rh(D)-factor antigens were obtained by incubating an antibody display library with magnetic bead-labeled Rh(D) positive red blood cells (RBCs) in the presence of an excess of unlabeled Rh(D) negative RBCs. The phage attached to the Rh(D) positive cells recovered after magnetic sorting were significantly enriched for specificity to the Rh(D)-specific antigen (Siegel *et al*., 1997). However, to perform this type of counterselection an appropriate subtractive substrate must be identified and available. Because this is not always possible, it is important to develop a different strategy that does not rely on subtraction.

One such strategy is to use blocking peptides to prevent attachment of phage bearing particular motifs to their respective target. Such an approach was used to identify peptide ligands that block interleukin 1 binding to an IL-1 receptor. In that case, initial pannings of the receptor produced a peptide ligand family that specifically bound the receptor but did not block IL-1 binding. Subsequent panning of the receptor in the presence of a peptide from the original family selected novel peptide sequences that were capable of competing with IL-1 (Yanofsky *et al*., 1996).

We have applied the same principles to develop a general method to sequentially identify ligand families that bind to different components of a heterogeneous mixture. First, the dominant ligand family is found by the conventional panning methodology. Then a blocking peptide is constructed based on that family, which competitively binds its target and thereby permits selection of the next most dominant family. As each new ligand family is discovered it can be blocked in further selections. This method differs from standard subtractive strategies in that the targets, not the ligands, are blocked. Because one uses blocking peptides constructed to contain the sequence motif common to the phage ligands, one does not require prior knowledge of its target. The initial application of this method to a mixture of high molecular mass components of human cord and adult sera identified peptides that bound to different components of the mixture as well as distinct peptide motifs that bound to the same target molecule. The iterative panning and blocking (IPAB) procedure enabled sequential identification of a fibrinogen binding motif and two distinct C1q binding motifs. The experimental part of this study uses phage displaying short peptides, but the principles and analysis apply equally to nucleic acids and other systems in which elements are selected from diverse libraries.

Results

Sequential identification of peptide motifs that bind to a complex mixture

High molecular mass cord blood serum components were prepared by size exclusion chromatography in the course of other work. One or more peptide libraries were independently panned against the serum fraction and, following significant enrichment after the third round of panning, candidate phage clones were sequenced. Amino acid sequences of the peptides displayed by those clones are shown in Figure 1(a). The majority of the sequences contained a Gly-Pro-Arg-Pro (GPRP)

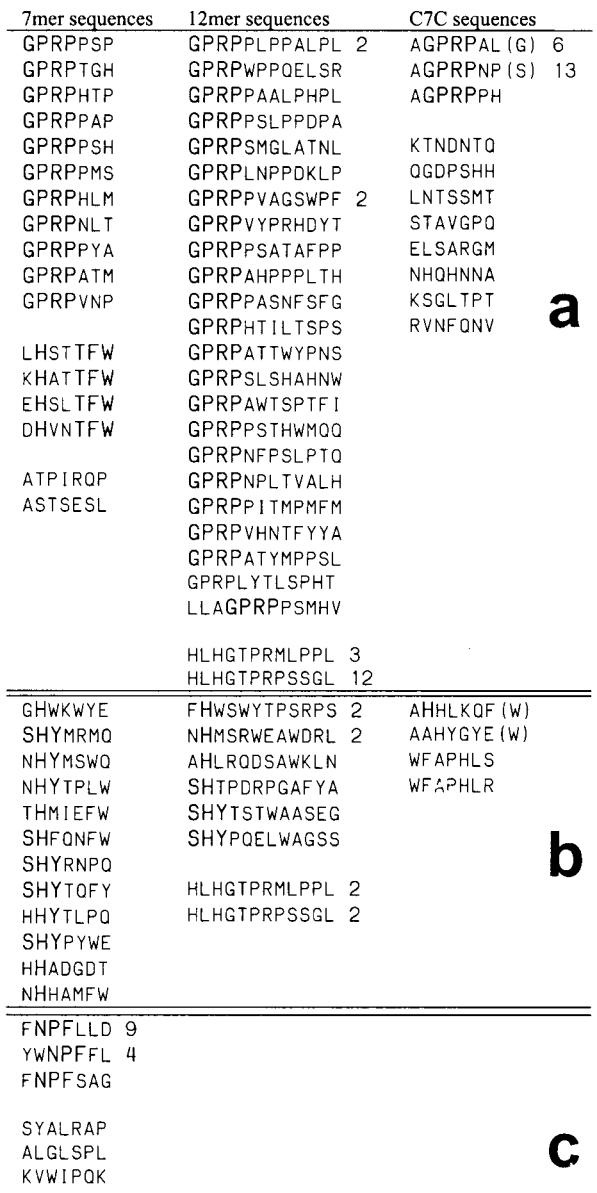

Figure 1. Amino acid sequences of phage displayed peptides selected against high molecular mass fraction of cord blood serum. Numbers after a sequence indicate the total number of independent sequenced phage clones displaying that peptide sequence. The amino acid indicated in parentheses for some C7C clones represents mutation of the C-terminal cysteine residue normally present at that position in this library. (a) Phage displayed amino acid sequences from the first cycle of panning with no peptide. (b) Sequences from the second cycle of panning performed in the presence of 50 μg/ml peptide D058-1 (sequence: GPRPPLNHALS). (c) Sequences from the third cycle of panning performed in the presence of 50 μg/ml peptide D058-1 and 50 μg/ml D093-1 (sequence: LHSTTFWGGGS).

motif at the N terminus of the displayed peptide. Similar sequences were obtained when adult serum fractions were used (data not shown). A second motif, xHxxTFW, was present in a minority of the 7-mer sequences not containing the GPRP motif.

Peptide D058-1 was synthesized based on the sequence of one of the GPRP motif phage. (The sequences of the synthesized blocking peptides used in this study are shown in Table 1.) This peptide specifically blocked the binding of phage displaying GPRP-containing peptides, presumably by competitive binding to the target molecules (Figures 2 and 5(a)). Interactions between another, unrelated phage-displayed peptide and its target, a monoclonal antibody, were not affected, indicating that peptide D058-1 is not a non-specific disrupter of phage displayed peptide/target interactions (Figure 5(a)). Furthermore, although the tetra-peptide NH_2-Gly-Pro-Arg-Pro-COOH (D168) also blocked GPRP-displaying phage binding, blocking did not occur when a peptide (D118) having this motif in reverse order was used (data not shown).

A second cycle of three panning rounds was performed in the presence of peptide D058-1. In each round, the peptide was added to the coated and blocked wells at a concentration of 50 μg/ml one hour before the library phage were added to the solution. The rest of the panning steps were identical with those in the original panning procedure. The sequences of the displayed peptides from a sample of the phages present after the final panning round are shown in Figure 1(b). Almost all of these sequences contain histidine at the second position from the N terminus, reminiscent of the minority motif identified in the original panning. A peptide (D093-1 in Table 1) was synthesized whose sequence is based on one of these histidine-containing motifs. This peptide blocked phage displaying the histidine motif but did not block the GPRP displaying phage (Figure 2). The reverse peptide did not inhibit the phage binding (data not shown).

Both peptides D058-1 and D093-1 were added in a third cycle of panning. After successful enrichment of the 7-mer library when panned against the high molecular mass serum fractions in the presence of both peptides at 50 μg/ml, phage clones were sequenced (Figure 1(c)). The majority of those sequences contained an Asn-Pro-Phe motif.

Phage clones representing each motif were individually tested for binding to the serum fraction in the presence of each of the synthesized blocking peptides. Wells coated with 1 μg of the serum fraction were incubated with peptide at 100 μg/ml for one hour prior to the introduction of approximately 10^{10} phage. Each peptide only blocked the

Table 1. Amino acid sequence of peptides

Peptide	Amino acid sequence
D058-1	GPRPPLNHALHS
D093-1	LHSTTFWGGGS
D118	SHLAHNLPPRPG
D168	GPRP
D246	FNPFLLDGGGS

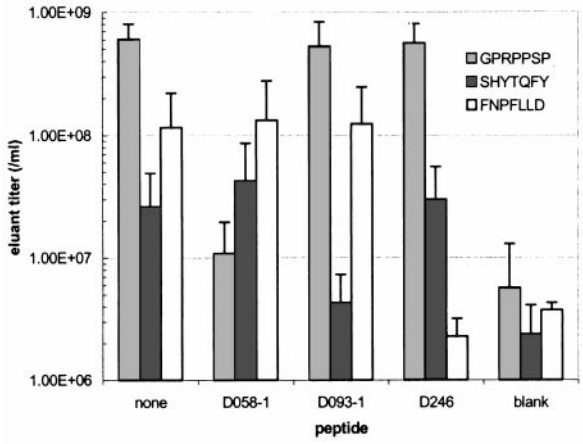

Figure 2. Phage displayed peptides are only blocked from binding to high molecular mass serum fraction by a peptide containing the same motif. Peptides were incubated on coated and blocked wells at 100 μg/ml for one hour prior to the addition of 10^{11} phage. The peptides displayed by the phage clones are shown in the key. The blank refers to the titer recovered from wells with no coating protein and no peptide.

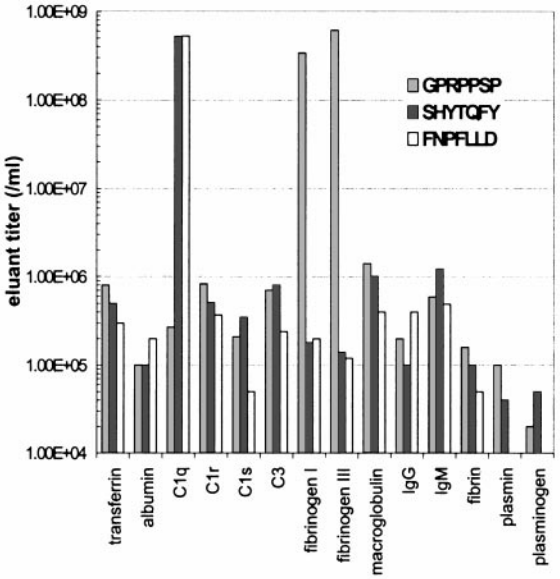

Figure 3. Representative phage clones were tested against wells coated with various serum proteins. All proteins were coated at concentrations of 0.1-1 mg per ml. After blocking with 1% BSA, 10^{11} phage were placed on the wells in PBS. The peptides displayed by the phage clones are shown in the key.

phage bearing the same motif, as shown in Figure 2. None of the phage significantly bound uncoated wells, ruling out potential plastic binding motifs.

Identifying the targets of the peptide motifs

Three phage clones representing each of the three motifs were tested for binding to several isolated serum proteins using the phage-binding assay. Serum proteins were coated on the wells at solution concentrations in the range 0.1-1 mg/ml. The results of these assays are shown in Figure 3. Phage bearing the initially identified GPRP motif bound to fibrinogen coated wells exclusively while the remaining two motifs bound to C1q coated wells exclusively. Binding to other proteins in all cases occurred only at the background level (10^6/ml) corresponding to non-specific attachment. This level of background is typical in our experience and is equivalent for phage with or without a displayed peptide.

SDS-PAGE of two high molecular mass cord blood fractions reveals that there are many proteins present in these samples (Figure 4(a)). Several common serum proteins, including C1q and fibrinogen, were run in parallel on the same gel. Interestingly, there are no clear bands in the serum fraction lanes that correspond to either C1q or fibrinogen. Western blotting reveals the presence of C1q and gives an estimate of approximately 5 μg/ml for sample Z3 and 10 μg/ml for Z4 (Figure 4(b)). Western blot analysis of the fibrinogen is shown in Figure 4(c). There are no visible bands at the height of the intact fibrinogen in the serum sample lane. However, several lower molecular mass bands are present. These likely are degradation products of fibrinogen subsequent to thrombin cleavage that occurred at the time the blood sample was clotted. A quantitative estimate of the fibrinogen concentration is not possible in this case.

Analysis of factors relevant to panning of mixtures

Once the targets of the various motifs were established, we used the purified target proteins to investigate several aspects of the experimental method. A dilution series of the peptide D058-1 was added to wells coated with 100 ng of the high molecular mass serum components, then our standard panning method was performed using phage displaying the GPRP motif. The results are shown in Figure 5(a). A solution concentration of approximately 10 μg/ml of peptide was required before significant blocking of phage binding occurred. At peptide concentrations around 100 μg/ml, the number of bound phage was reduced to the background level (typically 10^5-10^6 per ml in our experience). This is a several thousand-fold molar excess of free peptides relative to the phage-displayed peptides present on the roughly 10^{10} phage used in this assay. Other phage-target pairs also have shown blocking peptide ratios of this magnitude (those results are included with those described here in Figure 5(b)). It should be noted that the peptide solution was not washed from the well prior to the addition of

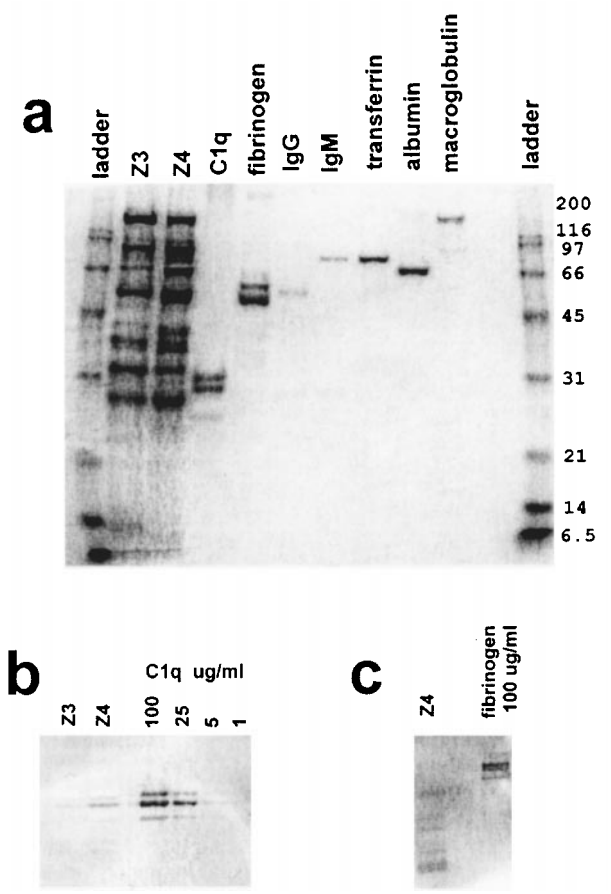

Figure 4. (a) SDS-PAGE of cord serum fractions Z3 and Z4 and various serum proteins. Ladder masses are shown on the right. All purified proteins were at 100 μg/ml. (b) Western blot of serum fractions Z3 and Z4 with anti-C1q antisera. Purified C1q was run in parallel at the concentrations indicated. (c) Western blot of serum fraction Z4 with anti-human fibrinogen antibody. Purified fibrinogen at 100 μg/ml was run on the right.

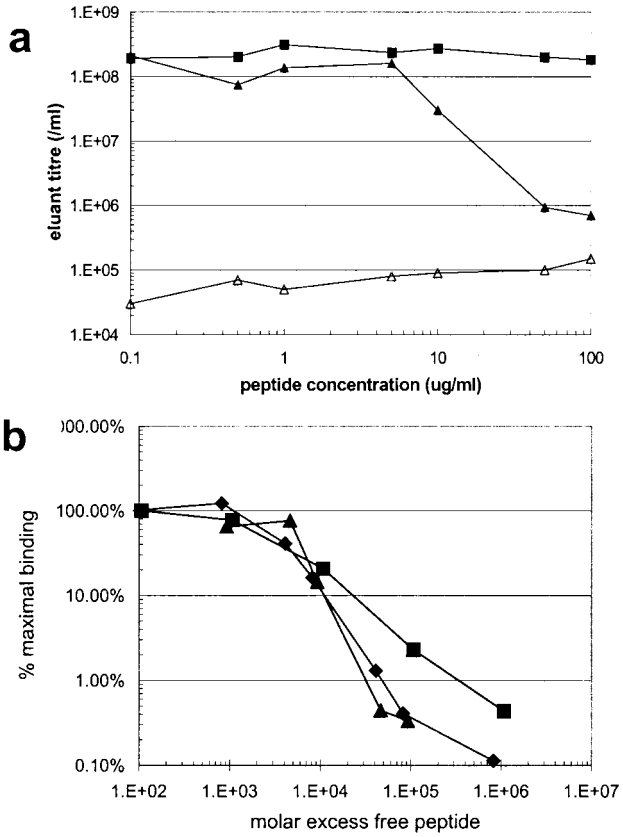

Figure 5. (a) Peptide D058-1 (sequence: GPRPPLNHALS) specifically blocks a phage displaying the same peptide. Peptides were incubated on coated and blocked wells at 100 μg/ml for one hour prior to the addition of 10^{11} phage. (▲) Phage displaying GPRPPLNHALS and (△) cclac+ f1 phage were tested for binding against wells coated with 100 ng of high molecular mass serum fraction. An unrelated phage (■) known to bind mAb RWL1 (previous study) (Messmer *et al.*, 1999) was tested for binding against that antibody in the presence of D058-1. (b) Molar excess of free peptide required to block phage displaying peptides that bind to (▲) high molecular mass serum fraction, (◆) a monoclonal IgM mAb RWL1, and C1q (■).

the phage. We have found that a washing step prior to the addition of the phage eliminates the free peptide's ability to block phage binding (data not shown). This suggests that the free peptide does not have a high enough affinity to remain bound over the course of the washing step whereas the phage displayed peptides, with greater potential avidity, are retained.

The relative and absolute concentration of any given component of a mixture could be expected to play a role in determining the stage at which ligands binding that target will be recovered. To address this question we investigated the relationship between coating concentration of the purified target and the recovered phage titer. Phage specific for fibrinogen, C1q, and the IgM class mAb RWL1 (previously described by Messmer *et al.*, 1999) were panned by our standard method against wells coated with a dilution series of each respective target protein. These three proteins of different valency (i.e. the number of potential peptide binding sites per molecule) were assayed over a wide range of coating concentrations. IgM mAb molecules each have ten antigen binding sites, while C1q has 6-fold symmetry and fibrinogen has 2-fold symmetry. In each case we found a critical coating solution concentration below which the number of phage recovered decreased rapidly (Figure 6(a)). In all cases the maximum rate of decrease was approximately quadratic, having a slope of two on the displayed log-log plot. However, this critical concentration is different for each target protein. Differences in the coating efficiency and/or differences in valency between proteins could contribute

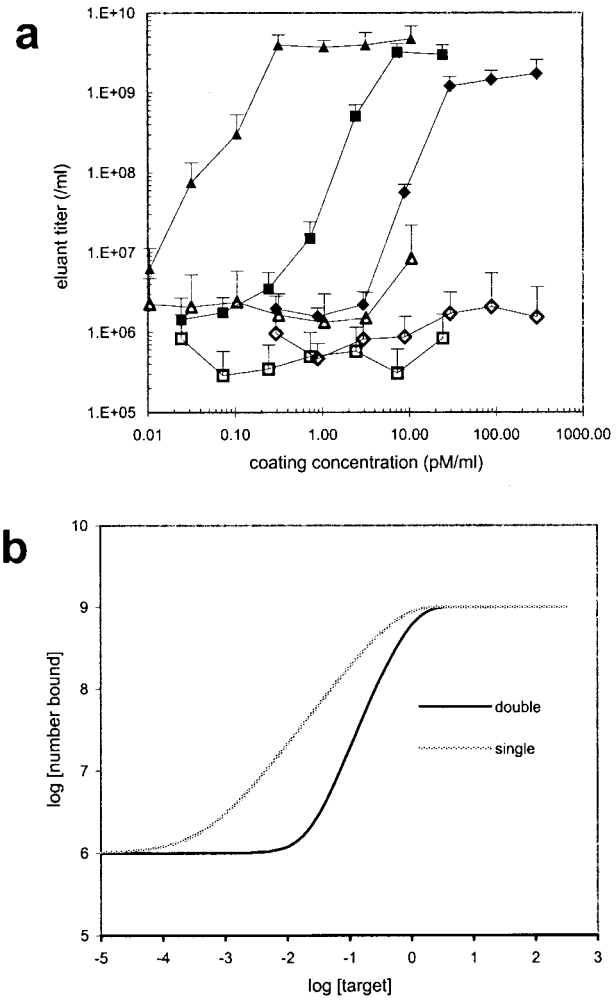

Figure 6. (a) Several peptide displaying phage were panned against wells coated with different concentrations of their target proteins. In all cases 10^{11} peptide displaying phage were mixed with 10^{11} cclac+ f1 phage. The eluant titers of the cclac+ f1 phage included in each experiment are shown with open symbols. Phage displaying GPRPPSP (◆) were panned against fibrinogen, phage displaying FNPFLLD (■) were panned against C1q, and phage displaying WGSNVYNSPFHS (▲) were panned against mAb RWL1, a monoclonal IgM antibody previously described (Messmer *et al.*, 1999). (b) Results of sample calculations analyzing the number of ligands that are bound to a target as the target concentration increases. The dotted line assumes that phage retention requires only a single binding event, while the continuous line assumes that a phage must be bound to two or more distinct target molecules to be retained. This difference is seen to change the maximum slope of the curve involved. In these log-log coordinates the maximum slope is 1 in the former case, and 2 in the latter.

to the different critical concentrations observed. The consistent quadratic slopes could derive from a non-linearity in the efficiency with which these proteins adhere to the polystyrene plates. However, we doubt this explanation because it implies that each protein binds to the plate with the precise cooperativity that would cause the addition of ten-fold more protein (in the coating solution) to result in 100-fold more protein attached on the plate (as indicated by the increased phage retention). We believe the consistency of the slopes observed reflects an important aspect of the phage-target interactions. The impact of these results on the application of the IPAB method led us to undertake a theoretical analysis.

Theoretical analysis

This analysis makes the following assumptions: (1) the target molecules attach to the polystyrene walls of the well at random according to Poisson statistics; (2) the phage titer recovered depends linearly on the number of phage that attach to the well; (3) non-specific attachment occurs at a constant background density (number per unit area), while the level of specific binding of phage to target molecules depends on the details of the attachment. We note that both the phage and all target molecules have the potential for multivalent interactions. Each phage has five identical displayed peptides, while IgM has ten potential binding sites, C1q has six and fibrinogen has two. This means phage could, in principle, be multiply bound to a single target. Alternatively, binding at one site on a molecule might sterically preclude binding at other sites, in which case multiple binding events would have to involve multiple target molecules.

The experiments described above suggest that the binding of a single peptide to a single target is relatively weak, regardless of whether the peptide is free or displayed by a phage. For this reason we explicitly examine two models in which the eluant titer depends on binding either to a single target molecule or to multiple molecules.

Suppose that the target proteins have average coated concentration λ so there are an average of λA target particles in area A. Here we interpret A as the area which one phage can probe. For present purposes we need not specify its size. However, as the diameter of the phage is 60 Å, this area might be a circle having approximately that radius. Then according to Poisson statistics the probability that there will be n target molecules in area A is:

$$P(n) = (\lambda A)^n \exp(-\lambda A)/n!$$

Then the probability that there is at least one target particle in area A is:

$$P(n > 0) = 1 - \exp(-\lambda A)$$

and the probability that there are two or more target particles in that area is:

$$P(n > 1) = 1 - \exp(-\lambda A) - \lambda A \exp(-\lambda A)$$

We assume the number of non-specifically bound phage after panning and washing is f_n, and the number of phage that could be specifically bound

after these procedures is f_s. We examine two cases, where the phage need only bind to a single target, and where it must bind to two separate targets. In either case the total number F of bound phages will be:

$$F = f_n + f_s P$$

where P is either of the probabilities in the previous two equations. Because the results in Figure 6(a) are shown in log-log coordinates, we change variables to $y = \log_{10}(F)$ and $x = \log_{10}(\lambda)$. Figure 6(b) shows plots of x *versus* y in each of our two cases, as calculated from the above equations with $f_n = 10^6$ and $f_s = 10^9$. One sees that these plots have the same general attributes as the experimental results. They both are asymptotically flat at low and high values of x, and are monotonically increasing. However, their maximum slopes differ. The maximum slope is 1 when binding to a single molecule suffices for retention, but it is 2 when binding to two or more distinct target molecules is required. The latter slope is consistent with the experimental results for each of the three target molecules tested. This strongly indicates that phage are retained to contribute to the eventual eluant titer only if they are multiply bound to their target, and that this multiple binding must involve more than one target molecule. It is not surprising that more than one peptide epitope on the same phage can interact with target. It is somewhat more surprising that only one site on each multivalent target is able to bind a particular phage.

Discussion

The IPAB method

An IPAB procedure that sequentially recovers specific ligands from a heterogeneous mixture of target molecules has been developed. The IPAB approach identifies dominant ligand families individually, then uses peptide analogs of those ligands to block their recovery in future panning cycles. In this way, dominant ligand families are blocked from subsequent selections, allowing minority ligand families to become over-represented.

The IPAB approach described here contrasts with previously developed subtractive methods in several important ways (Siegel *et al.*, 1997; Vant-Hull *et al.*, 1998). First, this method does not require any *a priori* knowledge regarding either the composition of the target mixture or the identity of the target to which a motif binds. This contrasts with subtraction methods, where one must either have some information regarding the identity and/or location of target molecules or be looking for the difference between two known sets in order to devise suitable subtractive substrates. Iterative panning only requires that displayed peptides be blocked from binding by homologous free peptides. Second, IPAB does not deplete the library of members that bind to a particular target. The use of small peptides to occupy specific interaction sites and thereby block phage binding does not preclude the subsequent enrichment for ligands that bind to other sites on the same molecule, as was found in the case of C1q.

One limitation of this panning strategy is that the protein targeted by the isolated motif must be identified by a subsequent procedure. Here we identified the targets recognized by our three peptide motifs simply by screening known serum proteins. More elegant approaches, such as Western blotting, crosslinking of ligands to targets, or, in some lucky cases, database homologies, could also be used either to determine the target uniquely or to identify a set of candidate proteins. In any event, this deconvolution of the ligand families is an important step that we believe can provide essential insights into the properties of the mixture being panned. Other implications of this limitation are discussed below.

Generality of IPAB

All methods of panning share genetic reasoning applied to biochemical scenarios. IPAB is equally applicable to phage display of peptides, antibodies, or proteins, as well as nucleic acid and other library selections and screens. All panning methodologies use libraries that consist of individual members. The three components of each library member are: (1) a carrier or vector that is identical for all members of the library; (2) a unique segment that is potentially selectable for its binding or activity; (3) an identifier or label that identifies and/or encodes the unique segment in (2). The distinction between (2) and (3) is not the same for different library types.

In phage display libraries, the vector is the phage or plasmid that allows for *in vivo* growth of selected members of the library. Segment (2) is the expressed polypeptide displayed on the phage coat. Segment (3) is the nucleic acid sequence that encodes the selected polypeptide. In nucleic acid selection methods (Green *et al.*, 1990; Kinzler & Vogelstein, 1989; Oliphant & Struhl, 1987; Robertson & Joyce, 1990; Tuerk & Gold, 1990) the vector is the constant sequence on either side of the library sequence in the center of each member of the library. The constant regions that flank the library inserts allow amplification either *in vivo* (following transformation, Oliphant & Struhl, 1987) or *in vitro* (by PCR, Szostak, 1992). In other cases, the library is constructed on beads with each bead containing a unique sequence for selection and an identifier that is easier to read on a microscale than is the actinic, or selected material (Brenner & Lerner, 1992). Conceptually the latter method is analogous to replica plating (Lederberg, 1989).

The IPAB method is applied after first panning a mixture of targets with a library in the standard way. This first layer panning identifies the domi-

nant epitope. In order to reach the second and subsequent layers, the first layer is selectively blocked and the panning is repeated. Selective blocking is accomplished by separating the epitope component from the vector of the library member selected in the previous panning. An excess of this epitope is added to saturate the target mixture and a subsequent round of panning is then conducted in the continuous presence of the blocking epitope. In the application of phage panning, blocking may be accomplished with a synthetic peptide or with ''special'' phage that display the relevant peptide. Special phage are killed or counter selectable in a way that leaves the displayed peptide intact, e.g. grown on a non-modifying host (Kunkel, 1985), or grown on a vector containing other counter selectable features such as ambers, or killed, e.g. with UV. In the case of nucleic acid libraries, the specific blocking epitope could be polynucleotides of the sequence of only the central portion of the previously selected library member. Alternatively, the sequence of primer binding sites could be changed for the next round of panning. For other libraries, such as those on beads, the principle is the same: the actinic epitope is separated from its label and carrier. An excess of the first layer's epitope is added to the target mixture and the panning repeated.

IPAB can delve beyond the second layer by blocking simultaneously with the first two layers and then looking for a third. There is no intrinsic limit to how many layers can be peeled by this method. Neither is there a limitation on the mixing of libraries that might be used in panning complex mixtures; e.g. a peptide library could be used to probe a mixture blocked with nucleic acids and so on.

Specific aspects of the motifs recovered

The Gly-Pro-Arg-Pro motif is a well-known homolog of the N-terminal portions of the fibrinogen alpha and beta chains that are exposed after thrombin cleavage (Bale *et al*., 1985; Furlan *et al*., 1982; Kuyas & Doolittle, 1986). These N-terminal ''knobs'' bind non-covalently to clefts within the D-domains of the corresponding chains of a fibrinogen or fibrin molecule, and thereby initiate formation of fibrin soft clots. The GPRP tetrapeptide has been shown to interfere with fibrin polymerization (Harfenist *et al*., 1982; Kuyas & Doolittle, 1986). The crystal structure of a GPRP peptide bound to fibrinogen also has been solved (Everse *et al*., 1998). Because so much information already was available, we did not further investigate this peptide.

A previous study identified C1q-binding phage by directly panning against C1q in solution. Those investigators used several libraries, and recovered some phage with our NPF motif (Lauvrak *et al*., 1997). The present study expands on that work by showing that the presence of a free peptide blocks binding to the homologous phage-displayed peptide. We currently are investigating both of the C1q binding motifs isolated in this study to determine what if any biological effects they might have. The fact that the GPRP fibrinogen-binding motif identified here mimics an essential aspect of fibrinogen's function suggests that other protein-binding motifs (in particular those of the C1q-binding peptides) may also interact at biologically active sites.

Several interesting observations were made from the cysteine-containing libraries. The two cysteine residues in this library are thought to form a disulfide bond, thus constraining the displayed peptide. However, it seems several of the peptides escaped this constraint by mutating the C-terminal cysteine-encoding codon (Figure 1). Sequencing of the phage DNA revealed that in all cases where the C-terminal cysteine was replaced, it was due to a single nucleotide change. We have also encountered this phenomenon in the course of other work (unpublished observation). Cysteine-constrained libraries have been suggested to benefit from the conformational limitations imposed by the formation of disulfide bonds, presumably due to the fewer degrees of freedom and consequently lower entropy of the peptide (Bonnycastle *et al*., 1996). However, our results suggest that some motifs, such as the GPRP motif, may not retain their target binding ability when the cysteine constraint is imposed. It should be noted that we do not know whether the mutations were present in the initial library pool or whether they arose during the selection process.

While the same observation was made for some of the histidine motifs recovered in our second round of panning (in those cases the C-terminal cysteine was replaced with tryptophan), a second effect was also seen. Two of the sequences recovered from the cysteine-containing library appeared to have reversed the histidine motif, i.e. the aromatic residues were near the N terminus and the histidine near the C terminus of the displayed peptide. This was not seen in either of the unconstrained libraries and as such may reflect a reaction to the imposed structure of the disulfide bond. Given these anomalies, we feel it is preferable to use both constrained and unconstrained libraries whenever possible.

Models and relationship to other work

Biopanning can be considered as a type of chromatography. Panning a library against a mixture of targets is conceptually equivalent to passing the library through a chromatography column that contains the mixture bound to a matrix. If the interaction of a member of the library with its target is independent of other library-target interactions, panning of a mixture of targets is equivalent to a set of parallel pannings, each conducted against a single component of the mixture. Chromatography theory is well developed, and the predicted output of such a system is a mixture of

library elements, each of which has been selected against one of the targets in the mixture.

Experimental results (this study and Morris *et al*., 1998) seem, at first, to contradict this prediction because a single dominant epitope, which binds to only one component of the mixture, was found. Others have examined this apparent paradox both experimentally and mathematically, and have suggested an insightful interpretation (Vant-Hull *et al*., 1998). They find that after panning a mixture of targets the library is indeed enriched for epitopes that bind to each of the separate targets in the mixture. However, the mixture contains different numbers of epitopes against each of the targets because the panning procedure selectively favors those having the highest target binding affinity. This affinity-induced bias is enhanced through each round of panning, yielding epitopes that bind to their targets with highest affinity. The analysis in that work concentrates on differences in affinity of a large set of target and ligand interactions. While target-ligand affinity is important, it is not the only determinant of the emergence of a dominant epitope-ligand pair, and may indeed not be the dominant factor.

Trivial factors influencing the order in which ligand-target pairs are obtained by panning a mixture of targets only have relevance because of the nature of a specific protocol. Other factors are general in that they are a consequence of chromatography theory applied to panning. Finally, there may be biological information implicit in the order of ligands obtained by panning against mixtures. To the extent that trivial and chromatographic factors can be understood and ruled out, biological insights may emerge from the order in which ligand-target pairs are found.

Trivial factors include selection of components of the mixture that bind to the non-biological matrix more tightly and in an orientation amenable to interaction with the library. This study used an ELISA plate of which there are several types of surfaces. Panning may also use targets bound to polystyrene beads. There are special features of the physics of boundary layers between solid and liquid phases where the key interactions between library and targets occur. The special physics of the boundary layer could obscure the interpretation of panning of protein mixtures that are normally in a liquid phase. In the panning of cell surfaces this is inescapably relevant.

Chromatography theory and the composition of libraries for biopanning are somewhat more general factors in the emergence of ligand-target pairs. The concentration of a particular target in the mixture will have clear effects. Both the absolute and the relative concentration of a candidate target in the mixture determine the potential for cooperative interactions of a target with members of the library (see Figure 6). Cooperative interactions will be a decisive factor in multivalent libraries such as gene III or gene VIII peptide display but are not expected to be important with monovalent antibody display libraries. Polynucleotide libraries are not in principle multivalent, but if avidity interactions are important and the binding motif is small enough, then library members that contain repeats of the interaction domain might be selected.

The mathematical analysis presented by others (Vant-Hull *et al*., 1998) assumes a continuous distribution of binding affinities within the library for any particular target. This assumption need not hold. Its importance in practice depends on whether the selection method can distinguish higher affinity ligands. Our panning methodology essentially selects for those phage with sufficiently slow disassociation rates such that they remain bound through the washing step. As such, our system may select only for those ligands above a minimum affinity, with no distinctions made beyond that level. Multiple epitope binding may also degrade the importance of affinity. That is, there will be a threshold affinity below which the phage will wash off and above which will not (presuming singly bound phage are washed off). In many cases, the consensus sequence for binding to a particular target is smaller than the size of the random peptides present in individual library members. This introduces a bias towards targets that require a smaller consensus sequence since more library members will contain it.

Finally one may consider the relevance of the order in which library-target pairs emerge from the panning of a biologically important mixture. In this study of umbilical cord serum fractions, the three targets revealed by biopanning were, first, fibrinogen, and subsequently two distinct sites on C1q. The fibrinogen binding motif yielded by panning in a ''naive'' unbiased manner is in fact the key site for fibrinogen polymerization in blood clotting. The second and third targets identified were on C1q. During infection of a naive host, C1q is the first molecule to bind to invading bacteria and to initiate a host response. Target discrimination is important for C1q function as binding to host tissue is associated with autoimmunity. The two motifs identified as binding to C1q are clearly solvent accessible, and may play biological roles. The second motif identified is similar to a sequence on C4, a complement component that has been shown to interact with C1q but has never been shown to bind to it. A peptide with the second motif is able to modify the activity of C1q (data to be presented elsewhere). Neither C1q nor fibrinogen comprised a large fraction of the sample examined (see Figure 4) yet they are the dominant targets that emerged from biopanning.

These considerations suggest that biopanning of mixtures can do more than yield pairs of specific target-library interactions. Applied to biologically relevant mixtures it has the potential to reveal which shape-specific interactions at a particular size scale are primary.

Materials and Methods

Reagents

All purified serum proteins were purchased from Sigma (St. Louis, MO). All peptides were purchased from Biosynthesis (Lewisville, TX) at the ''discovery'' scale. Peptides were analyzed by mass spectrometry by the manufacturer and the spectra provided. Table 1 shows the amino acid sequence of the various peptides used in this study.

Libraries and phage

Three phage displayed peptide libraries were purchased from New England Biolabs. The random peptide is displayed at the N terminus of the gene III protein (pIII). There is a small leader sequence separating the peptide from the protein as shown below.

Ph.D.7 NH_2-XXXXXXX-GGGS-pIII
Ph.D.12 NH_2-XXXXXXXXXXX-GGGS-pIII
Ph.D.C7C NH_2-ACXXXXXXXC-GGGS-pIII

These libraries were built in M13mp18 and as such contain the *lacZ*α fragment. They will not form blue plaques on *E. coli* strain S26*lacZ*$^-$ (Bachman, 1987) when plated on media containing Xgal. All phage stocks were grown on strain K91, an S26 derivative that has been cured of lambda.

Phage strain cclac+ f1 is an f1 phage containing a *lacZ* insertion. This phage forms blue plaques when plated on a *lacZ*$^-$ strain on media containing Xgal. The phage was a gift from Dr C. Cupples and is described elsewhere (Cupples & Miller, 1988).

Preparation of high molecular mass serum fractions

Serum was obtained from clotted blood, diluted 1:1 in PBS, and applied to a centricon C100 column. A Sephacryl S100 fast flow column was prepared as per manufacturer's instructions (Pharmacia, Bridgewater, NJ) and washed with 10× volume of PBS (pH 7.2) without $CaCl_2$ or $MgCl_2$: 5 ml of serum sample was applied to the column bed and eluted at 1 ml per hour for ca 62 hours. Fractions were read spectrophotometrically at 280 nm. The first 12 fractions containing significant amounts of protein were then pooled and reconcentrated to a volume of 2 ml. SDS-PAGE with 2-ME was performed on the final sample using 15% Tris-glycine/polyacrylamide gels and the buffer system of Laemmli (1970). The two umbilical cord serum fractions thus prepared were labeled Z3 and Z4.

Phage displayed peptide library panning

Our basic panning procedure has been described in detail elsewhere (Messmer *et al*., 1999). For each library panning, a single well of a high capacity ELISA plate (Corning cat. no. 25805-96) was coated with 1-5 μg of high molecular mass serum fraction in 100 μl of PBS (pH 7.4). Plates were left at 4 °C overnight. Coating solution was removed and 200 μl of blocking buffer (1% BSA) added. Plates were left at room temperature for at least one hour. After washing, 10^{11} library phage in 100 μl of PBS were added. After one hour at room temperature the wells were repeatedly washed between ten and 15 times with PBST (PBS + 0.1% Tween20): 100 μl of elution buffer (0.2 glycine-HCl, pH 2.2) was added and removed after 15 minutes then neutralized with 15 μl of Tris-HCl (pH 9.6). The eluant was titered and 50 μl added to 100 μl of a fresh overnight K91 culture. After at least 15 minutes at room temperature, the infected culture was added to 5 ml of LB and grown at 37 °C overnight. The following morning the cultures were pelleted and the supernatant filter sterilized by passage through a 0.45 μm syringe tip filter. The phage were precipitated with PEG solution (30% PEG 8000, 1.6 M NaCl) and resuspended in PBS. The panning was then repeated as before. In the third round a roughly equal number of cclac+ f1 phage were mixed with the second round eluant. The third round eluant was titred on K38*lacZ*$^-$ with Xgal and IPTG. Library phage form white plaques while the cclac+ phage form blue plaques. A white to blue plaque ratio of greater than 10:1 was usually indicative of a successful panning.

Phage sequencing and analysis by eluant phage titer

A dilution series of the final round eluant was plated on a lawn of K91. Individual plaques were picked, grown in rich media, and the phage containing supernatant filtered as above. single-stranded DNA was prepared with a Qiagen (Valencia, CA) qiaspin M13 kit as per manufacturer's instructions and submitted for sequencing at the Rockefeller University Sequencing Facility. The sequencing primer was obtained from New England Biolabs (Beverly, MA) and has the sequence 5′-CCCTCATAGTTAGCGTAACG-3′.

Individual phage clones were tested against various targets by the above panning procedure. When peptides were used to block, an additional step was added. After blocking, 100 μl of peptide in PBS was added to the wells and left at room temperature for one hour: 10^{10}-10^{11} phage were then added to the peptide solution and incubated and eluted as above. The cclac+ f1 were routinely included in all experiments as an indicator of non-specific background binding.

SDS-PAGE and Western blotting

All SDS-PAGE was performed with the buffer system of Laemmli on 12% (w/v) gels. Broad range standards from BioRad were used where appropriate. Typically gels were run at 90 V for one hour using a Biorad minigel system. Blotting was performed in a BioRad miniblot apparatus in transfer buffer (25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS, 20% (v/v) methanol) onto PVDF membrane for one hour at 90 V. Following blotting, the membranes were blocked in 3% BSA for at least two hours at room temperature. The membranes were then moved to the primary antibody solution in 1% BSA. For C1q detection, goat anti human C1q antisera (Sigma) was used at a dilution of 1:200. After two hours at room temperature with gentle shaking the membrane was washed with five changes of PBST and placed in a 1:500 dilution of anti-goat IgG HRP conjugate in 1% BSA. After one hour incubation as above the membrane was washed with three changes of PBST and two changes of PBS prior to being immersed in DAB development solution. For fibrinogen detection the blocked membrane was immersed in 1:200 dilution of anti-human fibrinogen HRP conjugate from Accurate Chemical (Westbury, NY) and incubated for two hours. The membrane was then washed and developed as above.

Acknowledgements

We thank the NIH Virology Training Grant CA09673 to Rockefeller University (B.T.M.) and the Sloan Foundation (D.S.T.) for support. We thank James Sullivan and Toby Rodman for preparing the serum fractions and Benjamin Schwarz, Aymeline Allix, and Choek Ho for technical assistance. We also thank Joshua Lederberg, Albert Libchaber, and Nicholas Chiorazzi for encouragement and insight during the course of this work, and Davorka Messmer and Kenji Adzuma for comments on the manuscript.

References

Adey, N. B., Mataragnon, A. H., Rider, J. E., Carter, J. M. & Kay, B. K. (1995). Characterization of phage that bind plastic from phage-displayed random peptide libraries. *Gene,* **156**, 27-31.

Bachman, B. J. (1987). Derivations of genotypes of some mutant deriviatives of *Escherichia coli* K-12. In Escherichia coli *and* Salmonella typhimurium *Cellular and Molecular Biology* (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M. & Umbarger, H. E., eds), 1st edit., vol. 2, pp. 1190-1219, American Society for Microbiology, Washington, DC.

Bale, M. D., Muller, M. F. & Ferry, J. D. (1985). Effects of fibrinogen-binding tetrapeptides on mechanical properties of fine fibrin clots. *Proc. Natl Acad. Sci. USA,* **82**, 1410-1413.

Bonnycastle, L. L., Mehroke, J. S., Rashed, M., Gong, X. & Scott, J. K. (1996). Probing the basis of antibody reactivity with a panel of constrained peptide libraries displayed by filamentous phage. *J. Mol. Biol.* **258**, 747-762.

Brenner, S. & Lerner, R. A. (1992). Encoded combinatorial chemistry. *Proc. Natl Acad. Sci. USA,* **89**, 5381-5383.

Cheng, X., Kay, B. K. & Juliano, R. L. (1996). Identification of a biologically significant DNA-binding peptide motif by use of a random phage display library. *Gene,* **171**, 1-8.

Cook, A. D., Davies, J. M., Myers, M. A., Mackay, I. R. & Rowley, M. J. (1998). Mimotopes identified by phage display for the monoclonal antibody CII-C1 to type II collagen. *J. Autoimmun.* **11**, 205-211.

Cortese, R., Felici, F., Galfre, G., Luzzago, A., Monaci, P. & Nicosia, A. (1994). Epitope discovery using peptide libraries displayed on phage. *Trends Biotechnol.* **12**, 262-267.

Cupples, C. G. & Miller, J. H. (1988). Effects of amino acid substitutions at the active site in *Escherichia coli* beta-galactosidase. *Genetics,* **120**, 637-644.

Everse, S. J., Spraggon, G., Veerapandian, L., Riley, M. & Doolittle, R. F. (1998). Crystal structure of fragment double-D from human fibrin with two different bound ligands. *Biochemistry,* **37**, 8637-8642.

Furlan, M., Rupp, C., Beck, E. A. & Svendsen, L. (1982). Effect of calcium and synthetic peptides on fibrin polymerization. *Thromb. Haemost.* **47**, 118-121.

Gebhardt, K., Lauvrak, V., Babaie, E., Eijsink, V. & Lindqvist, B. H. (1996). Adhesive peptides selected by phage display: characterization, applications and similarities with fibrinogen. *Pept. Res.* **9**, 269-278.

Germaschewski, V. & Murray, K. (1996). Identification of polyclonal serum specificities with phage-display libraries. *J. Virol. Methods,* **58**, 21-32.

Goodsell, D. S. (1998). *The Machinery of Life*, Copernicus Springer-Verlag, Inc., New York.

Green, R., Ellington, A. D. & Szostak, J. W. (1990). *In vitro* genetic analysis of the *Tetrahymena* self-splicing intron. *Nature,* **347**, 406-408.

Harfenist, E. J., Guccione, M. A., Packham, M. A. & Mustard, J. F. (1982). The use of the synthetic peptide, Gly-Pro-Arg-Pro, in the preparation of thrombin-degranulated rabbit platelets. *Blood,* **59**, 952-955.

Kinzler, K. W. & Vogelstein, B. (1989). Whole genome PCR: application to the identification of sequences bound by gene regulatory proteins. *Nucl. Acids Res.* **17**, 3645-3653.

Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. USA,* **82**, 488-492.

Kuyas, C. & Doolittle, R. F. (1986). Gly-Pro-Arg-Pro derivatives that bind to human plasma albumin and prevent fibrin formation. *Thromb. Res.* **43**, 485-490.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* **227**, 680-685.

Lauvrak, V., Brekke, O. H., Ihle, O. & Lindqvist, B. H. (1997). Identification and characterisation of C1q-binding phage displayed peptides [comment]. *Biol. Chem.* **378**, 1509-1519.

Lederberg, J. (1989). Replica plating and indirect selection of bacterial mutants: isolation of preadaptive mutants in bacteria by sib selection. *Genetics,* **121**, 395-399.

Mazzucchelli, L., Burritt, J. B., Jesaitis, A. J., Nusrat, A., Liang, T. W., Gewirtz, A. T., Schnell, F. J. & Parkos, C. A. (1999). Cell-specific peptide binding by human neutrophils [In Process Citation]. *Blood,* **93**, 1738-1748.

McConnell, S. J., Kendall, M. L., Reilly, T. M. & Hoess, R. H. (1994). Constrained peptide libraries as a tool for finding mimotopes. *Gene,* **151**, 115-118.

Messmer, B. T., Sullivan, J. J., Chiorazzi, N., Rodman, T. C. & Thaler, D. S. (1999). Two human neonatal IgM antibodies encoded by different variable-region genes bind the same linear peptide: evidence for a stereotyped repertoire of epitope recognition. *J. Immunol.* **162**, 2184-2192.

Morris, K. N., Jensen, K. B., Julin, C. M., Weil, M. & Gold, L. (1998). High affinity ligands from in vitro selection: complex targets. *Proc. Natl Acad. Sci. USA,* **95**, 2902-2907.

Oliphant, A. R. & Struhl, K. (1987). The use of random-sequence oligonucleotides for determining consensus sequences. *Methods Enzymol.* **155**, 568-582.

Pauling, L. (1945). Molecular structure and intermolecular forces. In *The Specificity of Serological Reactions* (Landsteiner, K., ed.), pp. 275-293, Harvard University Press, Cambridge, MA.

Peletskaya, E. N., Glinsky, G., Deutscher, S. L. & Quinn, T. P. (1996). Identification of peptide sequences that bind the Thomsen-Friedenreich cancer-associated glycoantigen from bacteriophage peptide display libraries. *Mol. Divers.* **2**, 13-18.

Pieczenik, G. (1999). Method and means for sorting and identifying biological information pat #:5866363, In *US Patent and Trademark Office (www. uspto. gov)*, USA.

Rajotte, D., Arap, W., Hagedorn, M., Koivunen, E., Pasqualini, R. & Ruoslahti, E. (1998). Molecular heterogeneity of the vascular endothelium revealed by in vivo phage display. *J. Clin. Invest.* **102**, 430-437.

Rebar, E. J., Greisman, H. A. & Pabo, C. O. (1996). Phage display methods for selecting zinc finger proteins with novel DNA- binding specificities. *Methods Enzymol.* **267**, 129-149.

Robertson, D. L. & Joyce, G. F. (1990). Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature,* **344**, 467-468.

Sahu, A., Kay, B. K. & Lambris, J. D. (1996). Inhibition of human complement by a C3-binding peptide isolated from a phage-displayed random peptide library. *J. Immunol.* **157**, 884-891.

Scott, J. K. & Smith, G. P. (1990). Searching for peptide ligands with an epitope library. *Science,* **249**, 386-390.

Siegel, D. L., Chang, T. Y., Russell, S. L. & Bunya, V. Y. (1997). Isolation of cell surface-specific human monoclonal antibodies using phage display and magnetically-activated cell sorting: applications in immunohematology. *J. Immunol. Methods,* **206**, 73-85.

Smith, G. P. & Scott, J. K. (1993). Libraries of peptides and proteins displayed on filamentous phage. *Methods Enzymol.* **217**, 228-257.

Smith, G. P., Petrenko, V. A. & Matthews, L. J. (1998). Cross-linked filamentous phage as an affinity matrix. *J. Immunol. Methods,* **215**, 151-161.

Stephen, C. W., Helminen, P. & Lane, D. P. (1995). Characterisation of epitopes on human p53 using phage-displayed peptide libraries: insights into antibody-peptide interactions. *J. Mol. Biol.* **248**, 58-78.

Szardenings, M., Tornroth, S., Mutulis, F., Muceniece, R., Keinanen, K., Kuusinen, A. & Wikberg, J. E. (1997). Phage display selection on whole cells yields a peptide specific for melanocortin receptor 1. *J. Biol. Chem.* **272**, 27943-27948.

Szostak, J. W. (1992). *In vitro* genetics. *Trends Biochem. Sci.* **17**, 89-93.

Tchernychev, B., Cabilly, S. & Wilchek, M. (1997). The epitopes for natural polyreactive antibodies are rich in proline. *Proc. Natl Acad. Sci. USA,* **94**, 6335-6339.

Tuerk, C. & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science,* **249**, 505-510.

Van Ewijk, W., de Kruif, J., Germeraad, W. T., Berendes, P., Ropke, C., Platenburg, P. P. & Logtenberg, T. (1997). Subtractive isolation of phage-displayed single-chain antibodies to thymic stromal cells by using intact thymic fragments. *Proc. Natl Acad. Sci. USA,* **94**, 3903-3908.

Vant-Hull, B., Payano-Baez, A., Davis, R. H. & Gold, L. (1998). The mathematics of SELEX against complex targets. *J. Mol. Biol.* **278**, 579-597.

Yanofsky, S. D., Baldwin, D. N., Butler, J. H., Holden, F. R., Jacobs, J. W., Balasubramanian, P., Chinn, J. P., Cwirla, S. E., Peters-Bhatt, E., Whitehorn, E. A., Tate, E. H., Akeson, A., Bowlin, T. L., Dower, W. J. & Barrett, R. W. (1996). High affinity type I interleukin 1 receptor antagonists discovered by screening recombinant peptide libraries. *Proc. Natl Acad. Sci USA,* **93**, 7381-7386.

Edited by M. Gottesman

(Received 2 September 1999; received in revised form 20 December 1999; accepted 20 December 1999)