

Epitope discovery using peptide libraries displayed on phage

Riccardo Cortese, Franco Felici, Giovanni Galfre, Alessandra Luzzago, Paolo Monaci and Alfredo Nicosia

Peptides displayed on phage, which mimic continuous and discontinuous epitopes, can be selected using purified antibodies or preparations of polyclonal serum. This review describes recent advances in this field, discusses the application of phage-display technology to the diagnosis of human diseases, and presents new ideas for the preparation of vaccines directed against specific epitopes on a pathogen.

In the past few years, rapid advances have added new dimensions to the experimental strategies used to identify and characterize the binding of ligands to receptors. These new technologies are based on a common theme: the preparation of mixtures (repertoires) of very large numbers of molecules (peptides, oligonucleotides or others) that have randomized sequences or structures. Methods have been devised to identify member(s) of the repertoire that have specific binding properties. Similar concepts have directed chemists towards the creation of 'chemical' repertoires, and molecular biologists towards the creation of 'genetic' repertoires. Among the 'genetic' repertoires, the system that has received most attention, primarily because of its simplicity and versatility, is the so-called 'phage display of peptides and proteins', which is based on the properties of filamentous phage M13, fd and f1.

Filamentous phage display

In 1982, Dulbecco¹ suggested that immunogenic peptides derived from well-characterized epitopes of pathogenic agents could be fused to the coat proteins of phage lambda or of other viruses. The resulting viral particles would 'display' the foreign peptide on their surface and could be used as the principal components of cell-free vaccines. In 1985, Smith² showed that the genome of phage fd could easily be manipulated to obtain phage particles displaying foreign epitopes (phagotopes) on their surface. These displayed epitopes would be recognized by the corresponding antibodies. This observation stimulated further research in this area, and has culminated in phage-display technology being used in a variety of applications (reviewed in Ref. 3).

Filamentous phage are the preferred vectors for the construction of 'genetic' repertoires because of the following properties:

- (1) Insertions of foreign sequences at the amino-terminus of either the pIII or, in certain cases, the pVIII coat protein are, in general, well tolerated; the foreign sequence is displayed on the surface of the viral particle and can bind to molecules such as antibodies and receptors³.
- (2) Even foreign sequences that interfere with the life cycle of the virus can be displayed by using the so-called 'two genes' or 'phagemid' system, in which chimaeric phage bearing both wild-type and recombinant coat proteins are produced³.
- (3) By using appropriate 'selector' molecules immobilized on a solid support, it is possible to affinity purify phage displaying the binding protein or peptides from a mixture of $>10^9$ different phage particles. The selected phage can then be amplified by re-infection of *Escherichia coli* cells. The sequence of the selected peptides can be deduced from the sequence of the DNA encapsidated in the same phage particles³.

Libraries have been constructed by cloning random oligonucleotide sequences in frame with the genes encoding the capsid proteins pVIII or pIII [known as random peptide libraries (RPLs)]; in this way, each phage displays a different peptide on its surface. There are five copies of pIII and ~ 2700 copies of pVIII on the coat of filamentous phage. Peptides of up to a few hundred residues in length can be displayed on pIII (Ref. 3); by contrast, the maximum length of peptide that can be displayed on pVIII is much shorter, probably no more than eight or nine residues (G. Cesareni, pers. commun.). However, if the two-gene system is used, the length of peptide that can be displayed on pVIII can be increased considerably to 100 or more residues^{4,5}. The two-gene system uses a mosaic of

R. Cortese, F. Felici, G. Galfre, A. Luzzago, P. Monaci and A. Nicosia are at the Istituto di Ricerche di Biologia Molecolare P. Angeletti, Via Pontina Km. 30.600, 00040 Pomezia (Roma), Italy.

Table 1. Phagotopes mimicking linear epitopes

Selector mAb ^a	α -IL1 β	PAb 240	3-E7	M33	447-52D
Antigen	h-interleukin 1 β	p 53	β -endorphin	Myohemerythrin	HIV1-V3 loop
Epitope	VQGEESNDK	FRHSVV	YGGF	DFGLEKI	HIGPGRAF
Phagotopes	LFDRE S NDW ^b DKSS S NDSTIA RDLHS S NDVV S NDGVWAIP YTD S NERTT YRDS S NEHQP	WRHSV ^c LRHSVV YRHSVI LRHSVI	YGGFPD ^c YGGLGL YGGLGI YGGLGR YGGLNV YGGLRA YGGLLE YGGIAS YGGIAV	DFLEWL ^c DFLEML DFLERI DFLEQL DFLVQL DFLHFI CEFLEC CRFLEC DFMEWL	VRHM GPGR GMVLGIT ^d LLRTIM IGPGR LLHS GILRQPLL IGPGR AP LQLL IGPGR TVGKIR ARGWGGVFY GPGR GE NRY GPGR VGFVRSQP SCMTSVLV GPGR QDN ILF GPGR CSVDAVSG PYSDDL ^{LL} SKYF GPGR

^a Data on monoclonal antibodies (mAbs) α -IL-1 β from Ref. 8; on PAb240 from Ref. 9; on 3-E7 from Ref. 10; on M33 from Ref. 11 and on 447-52D from Ref. 12.
^b Nonapeptide library displayed as pVIII-fusion.
^c Hexapeptide library displayed as pIII-fusion.
^d Pentadecapeptide library displayed as pIII-fusion.

wild-type pVIII molecules from a helper phage and recombinant pVIII molecules with the foreign peptide displayed on each phage particle. The ratio of recombinant to wild-type pVIII varies from 1:1 to 1:1000, depending on the specific foreign peptide and on the experimental conditions (R. Cortese *et al.*, unpublished).

The display of known proteins or natural antibody repertoires has recently been reviewed⁶. This article describes the use of RPLs to identify epitopes that are useful as diagnostic reagents or as immunogens.

Phagotopes mimicking natural epitopes

The use of RPLs has enabled epitope mapping using monoclonal antibodies (mAbs) as selector molecules. Typically, this technique generates families of phagotopes that have easily recognizable consensus sequences, often matching the sequence found in the natural antigens (see Table 1).

In an earlier review of epitope mapping using RPLs and mAbs, it was reported that approximately half of all the mAbs used failed to select specific phagotopes^{3,7}. The main cause of this problem was attributed to the difficulty of mimicking discontinuous epitopes with linear, unconstrained peptides. Moreover, it was suggested that, in many cases, binding, and thus selection, might be impossible with a freely moving linear peptide if the loss of entropy were not compensated by the enthalpy change on binding. To overcome these problems, 'constrained' peptide libraries have been constructed, in which the potential conformation of the displayed peptide is limited by the structural constraints built around it. For example, flanking the random sequence in the RPL by two cysteine residues forces, in most cases, the displayed peptide into a loop conformation^{6,13,14}. Other constraints, such as elements of secondary structure, have also been introduced into RPLs (Ref. 15). At present, it is not altogether clear to what extent

the specific constraints applied to the peptides have facilitated the successful isolation of phagotopes.

Specific epitopes can be selected from an RPL, only if the library contains phagotopes whose avidity-affinity for the selecting antibody is higher than the threshold determined by the stringency of the selection conditions. However, a number of things can be done to increase the likelihood of finding particular phagotopes. The most obvious is to increase the complexity of the library as much as possible. The next is to optimize the stringency of the selection procedure, so that phagotopes that bind specifically, but with low affinity, are identified. Nevertheless, regardless of the type of library used, we have been more successful in isolating specific phagotopes from the pVIII system, in which each phagotome is displayed in hundreds of copies, than from the pIII system, in which a maximum of five copies of each epitope is displayed (R. Cortese *et al.*, unpublished).

The differences between the two display systems could be exploited to implement a strategy of *in vitro* maturation of epitopes. Low-affinity phagotopes can first be selected from pVIII libraries, thus maximizing the probability of finding them. The epitopes selected in this first screen could then be mutagenized and transferred onto pIII phagemids displaying a single copy of the foreign peptide in order to identify the highest-affinity phagotopes.

Phagotopes mimicking discontinuous B-cell epitopes

Classical epitope mapping requires a set of overlapping synthetic peptides spanning a region of the antigen, or a set of deletions obtained by genetic manipulation of the antigen-coding gene. If a peptide was found to react with the antibody, the epitope was declared to be linear or continuous. By contrast, if a reacting peptide could not be identified, by default, the epitope was declared to be discontinuous, i.e. it

Selector mAb ^a	1B7	MGr-2	5.5	H107
Antigen ^b	Bp toxin	HER2/neu	AcChoR	h-H ferritin
Phagotopes	GRRPNPPPIP^c GRAPNPNLP GRSPGSHMA	YTPWMLPE^c SPPWMLPSV TPNWMLQNL TPRWMLSRE TPTWMLARW	DLVWLL^e DIVWLL LIEWLL	YDGSYRWAP^d YDAQHHWTR YNAFLNWVP
	FLRFHPHWL ^d RGTSPMDAT SRTGRCGHG	GPLDSLFAQ^d GPISALFAS GPIHALFLD GPISSLFGD	RLPGSFTY HIPGSIFHL RTTGSVFHL RWOGSAFAY YPYGSTFGL	

Sometimes, the consensus sequence of selected phagotopes cannot be found in the sequence of the natural antigen. In these cases, it is possible that the phagotopes are mimics of discontinuous stretches of natural epitopes. For instance, Felici *et al.*¹⁸ selected

By using mAb MGr2, which is directed against the extracellular domains of the HER2/neu receptor, a series of phagotopes was isolated (Felici and Orlandi, pers. commun.). Their consensus sequence (see Table 2) cannot be found, even at very low stringency, in the amino acid sequence of the neu polypeptide chain. However, using one of the phagotopes as an immunogen elicited an antibody response against HER2/neu.

More-compelling evidence, although still indirect, in support of the hypothesis that phagotopes can mimic discontinuous epitopes, has been reported by Luzzago *et al.*¹³ These authors used mAb H107, which is specific for the human ferritin H subunit, the three-dimensional structure of which is known²¹. Mutants were identified that failed to react with mAb H107, but which were assembled efficiently into the 24-subunit shell of ferritin, indicating that the mutation only had a 'local effect', and probably did not disturb the overall structure of the ferritin molecule. Mutations with these properties mapped to two distinct regions several residues away from each other, but located very closely in space, thus confirming the hypothesis that the epitope recognized by mAb H107 is composed of two discontinuous regions. Phagotopes selected with mAb H107 could be classified into two groups (see Table 2 for their consensus sequences). The amino acids that comprise the consensus sequences of the phagotopes are also very closely related in space in the natural epitope (see Fig. 1). The most probable interpretation of these data is that the two families of phagotopes are mimotopes of two different parts of the mAb H107 epitope. If this interpretation is correct, then phagotopes of both classes mimic regions that contain structural elements on either side of the discontinuity of the natural epitope.

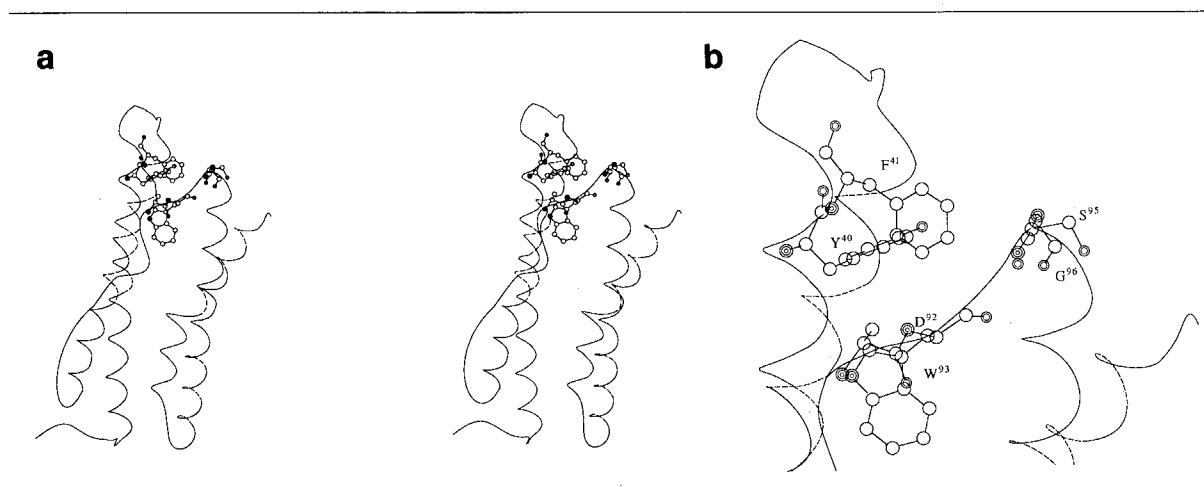


Figure 1

Structure of the H-ferritin epitope recognized by monoclonal antibody (mAb) H107. (a) Stereo-pair images of H-ferritin monomer. Amino acids involved in binding mAb H107 are shown. Side chains of the relevant amino acids are indicated. (b) Enlarged image of the epitope region of mAb H107.

Phagotopes mimicking non-peptidic epitopes

Phagotopes whose peptide sequences are, apparently, structural analogues of non-proteinaceous structures have also been isolated. Devlin *et al.*²² and Kay *et al.*²³ have isolated phagotopes that react with streptavidin. These molecules are likely to be analogues or mimics of biotin, on the basis of binding competition experiments. In both cases, most of the phagotopes that bind streptavidin display the consensus sequence HPQ. Weber *et al.*²⁴ have determined the crystal structure of the complex between streptavidin and the heptapeptide FSHPQNT, which contains the HPQ consensus sequence. A major factor stabilizing the streptavidin–heptapeptide interaction is a network of hydrogen bonds that includes: the side chains of histidine, proline and glutamine residues from the peptide; Asp126, Thr90 and Ser27 of streptavidin; as well as a water molecule. The peptide and biotin complexes with streptavidin form an analogous hydrogen-bond network in the crystal structure. However, the peptide appears to use only a subset of the biotin–interaction sites at the binding pocket.

Other examples of phagotopes mimicking non-peptidic epitopes have been reported by Scott *et al.*²⁵, Oldenburg *et al.*²⁶ and Hoess *et al.*²⁷. These authors have used either the carbohydrate-binding lectin concanavalin A (Con A) (Refs 25, 26) or mAb B3, which binds to the carbohydrate antigen Lewis^x (Ref. 27) as a selector molecule. Peptides that bind Con A display a common consensus sequence containing the amino acids YPY. Binding was specific, Ca²⁺-dependent and exhibited comparable affinity with methyl- α -D-mannopyranoside, the natural binder. The phagotope binding to mAb B3 has the sequence APWLYGPA, but mutagenesis experiments indicated that PWLY plays a crucial role. It is interesting to note that natural protein inhibitors of carbohydrate-binding proteins also contain aromatic-rich sequences that are apparently involved in the formation of the complex with carbohydrate²⁷. It has been suggested that the hydroxyl

group of Tyr might mimic the hydrogen-forming oxygen atoms of the sugar, whereas the hydrophobic aromatic ring might mimic the carbon atoms in the sugars^{7,25–27}. There are no reports, to date, on the use of carbohydrate-mimicking phagotopes as immunogens.

Phagotopes mimicking T-cell epitopes

T lymphocytes recognize peptide fragments of antigens in association with molecules of the class I or class II major histocompatibility complex (MHC), which are present on the surface of antigen-presenting cells. As the peptides that bind to class I molecules are only eight or nine amino acids long, by purifying natural complexes of peptides and class I molecules, it has been possible to identify many consensus sequences²⁸. However, similar work on the binding of peptides to class II molecules cannot easily be performed because, while only eight or nine residues are involved in binding, the peptides purified from natural complexes are much longer and heterogeneous in length²⁸.

A particularly effective way of using RPLs has been implemented by Sinigaglia and co-workers^{28,29}, who used purified MHC class II molecules. The authors used three closely related polymorphic class II molecules and compared the sequences of the corresponding three families of epitopes. All three families of phagotopes were found to have common features, as well as significant differences. The sequences can be aligned to show a conserved Tyr residue at position 1 and a Met residue at position 4 of the sequence. Furthermore, within each family, there appear to be allele-specific conserved residues at position 6. The authors proposed that Tyr and Met are the anchor residues, which fit into two pockets that are common to the three allelic class II molecules used in this study. By contrast, they proposed that residues at position 6 in the phagotope sequence correspond to a site on the class II molecule that is responsible for allele-specific binding. This information should be invaluable in the

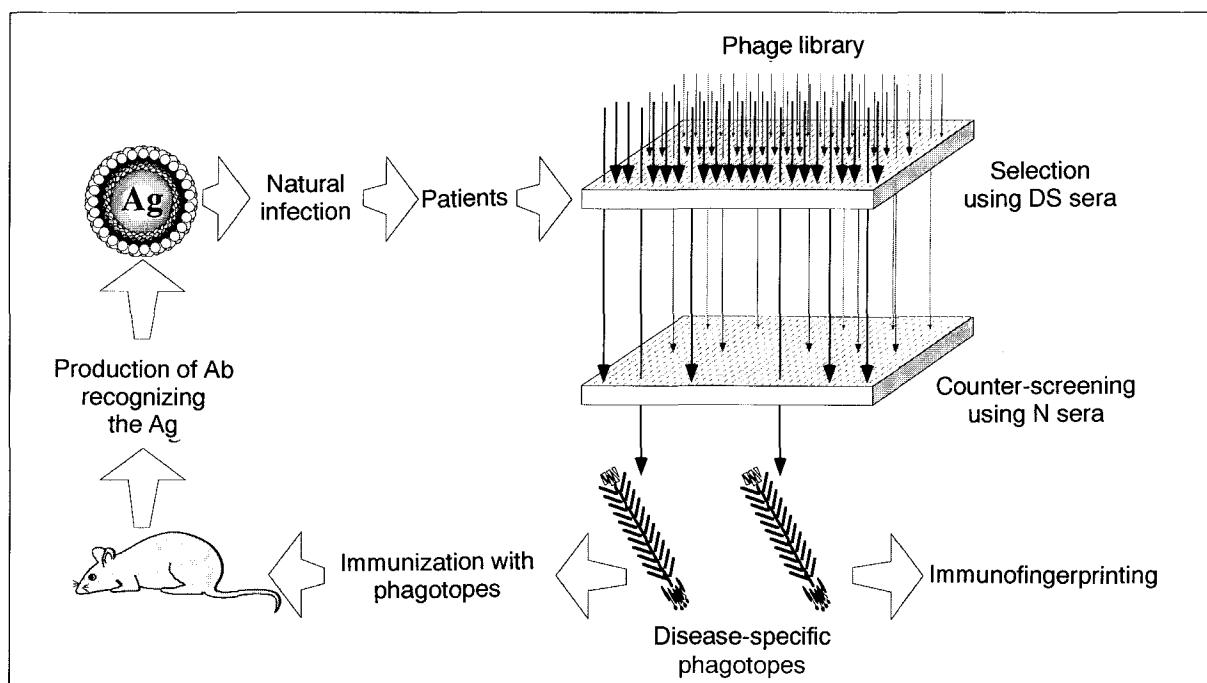


Figure 2

Strategy for identifying disease-specific phagotopes eliciting antibodies against the original antigen.

study of autoimmune diseases, as well as in the design of vaccines.

Selection of disease-specific phagotopes

In the previous examples, phagotopes were selected using mAbs or well-defined selector molecules. However, the technology can be applied to situations in which the selector molecules are not known, or are undefined; for example, to the polyclonal antibodies present in serum. This more-general application of RPLs opens new avenues for the identification and characterization of reagents that might prove useful in the diagnosis of human diseases, and for the preparation of novel vaccines.

Sera from patients suffering from an infectious disease often contain a large number of different antibodies directed against the infectious agent. Characterizing the nature, specificity and affinity of these antibodies is usually possible only if the natural antigen, either as a living microorganism, or as the product of a cloned gene, is available and can be used as a reagent. Likewise, in other diseases (for example, autoimmune diseases), the characterization of the immunological disorder is essentially confined to those antibodies for which a self-antigen is identified.

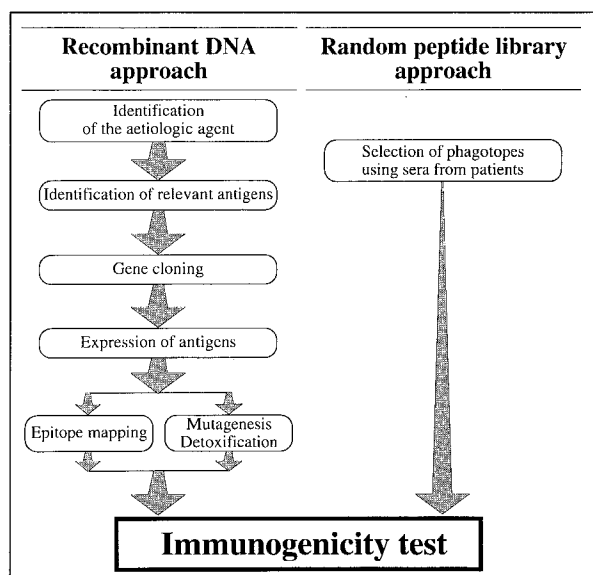
In those cases where the natural antigen is either unknown or not easily available, it is now possible to use RPLs. For this, a set of serum samples from patients affected by the disease under study (Disease Specific or DS sera), and a set of sera from normal individuals (Normal or N sera) are required. The objective is to identify disease-specific (DS) phagotopes, which are defined as those reacting with some or all DS sera and with none of the N sera, from the RPLs (see Fig. 2).

A protocol for this procedure was developed, using sera from an individual immunized with a commercial preparation of human hepatitis B virus (HBV) vaccine containing the purified recombinant HBV surface antigen (HBsAg) as a model system^{30,31}. The sera from immunized individuals were used as models of DS sera, whereas the sera from non-immunized individuals were used as models of N sera. Even though the nature of the antigen in this model case is known, this information is not required to isolate DS phagotopes.

By affinity selection of the RPL with a first DS serum sample, and screening the selected phage by immunoblotting with a second DS serum sample, phagotopes reacting with antibodies present in both samples were identified. The DS phagotopes present in this overlapping subset of phagotopes were identified by further screening with many different samples of DS sera, and counterscreening with samples of N sera. Finally, we obtained phagotopes that were shown to be antigenic mimics of HBsAg by binding competition experiments.

The results described in the previous section demonstrate that it is possible to identify peptides reacting with DS antibodies, without having any information on the natural antigen that elicited the production of DS antibodies in the first place.

In principle, the strategy used with the HBsAg model could be applied to any other disease. In the case of infectious diseases, a large set of DS phagotopes could be used to analyse the DS-specific humoral immune response of each patient, with a very high degree of resolution, thus allowing the determination of a type of DS 'immunofingerprint' for each individual. The application of this strategy to diseases for

**Figure 3**

Preparation of acellular vaccines: a comparison between the recombinant DNA strategy and the random peptide library approach.

which the cause is not known might lead to the isolation of a set of DS phagotopes, representing the 'immunofingerprint' of the unknown aetiological agent. The search for such an agent could be facilitated by the use of DS phagotopes as reagents.

A particularly attractive possibility is to apply these techniques to the study of autoimmune diseases. The hypothesis that these diseases are caused by viruses is very popular, but lacks any convincing evidence. Investigations in this field are largely confined to the identification of T- and B-cell self-antigens. Random peptide libraries could be used to identify phagotopes that react with DS antibodies, without any prejudice on the nature of the natural antigen. This will increase the chance of identifying anti-virus antibodies that are not necessarily self-antibodies and, of course, will broaden the available range of diagnostic and prognostic reagents.

Vaccines that 'fit' disease-specific antibodies

Phagotopes that react specifically with serum samples from HBsAg-vaccinated individuals were shown to be good immunogens in mice or rabbits, eliciting the production of antibodies recognizing HBsAg (Ref. 31). This observation suggests a novel way of preparing vaccines. This new strategy would be based on the identification, from RPLs, of those epitopes that 'fit' the DS antibodies present in the sera of the infected individual. This would be independent of any knowledge and/or availability of the natural antigens. In principle, this new strategy might simplify dramatically the identification, cloning and expression of the recombinant immunogen (Fig. 3).

A new generation of vaccines could simply consist of a cocktail of phagotopes, perhaps including phagotopes that mimic the carbohydrate moiety of glycoproteins or complex epitopes, selected using

sera from infected patients, preferably from those who have successfully overcome the disease and therefore would probably possess protecting antibodies.

The phage itself is an excellent immunogen, even when injected without adjuvant, eliciting a T-cell-dependent response against the displayed epitope^{32,33}. The identification of allele-specific T-cell phagotopes²⁹ adds another dimension to this strategy: vaccines could comprise phage particles which, at the same time, present B-cell phagotopes that 'fit' natural protective human antibodies and allele-specific class II T-cell phagotopes.

References

- 1 Dulbecco, R. (1986) *US Patent No. 4 593 002*
- 2 Smith, G. P. (1985) *Science* 228, 1315-1317
- 3 Smith, J. P. (1991) *Curr. Opin. Biotechnol.* 2, 668-673
- 4 Kang, A. S., Barbas, C. F., Janda, K. D., Benkovic, S. J. and Lerner, R. A. (1991) *Proc. Natl Acad. Sci. USA* 88, 4363-4366
- 5 Light, J. and Lerner, R. A. (1992) *Bioorganic Med. Chem. Lett.* 2, 1073-1078
- 6 Scott, J. K. (1992) *Trends Biochem. Sci.* 17, 241-245
- 7 Hoess, R. H. (1993) *Curr. Opin. Struct. Biol.* 3, 572-579
- 8 Felici, F., Castagnoli, L., Musacchio, A., Jappelli, R. and Cesareni, G. (1991) *J. Mol. Biol.* 222, 301-310
- 9 Stephen, C. W. and Lane, D. P. (1992) *J. Mol. Biol.* 225, 577-583
- 10 Cwirla, S. E., Peters, E. A., Barrett, R. W. and Dower, W. J. (1990) *Proc. Natl Acad. Sci. USA* 87, 6378-6382
- 11 Scott, J. K. and Smith, G. P. (1990) *Science* 249, 386-390
- 12 Keller, P. M. et al. (1993) *Virology* 193, 709-716
- 13 Luzzago, A., Felici, F., Tramontano, A., Pessi, A. and Cortese, R. (1993) *Gene* 128, 51-56
- 14 McLafferty, M. A., Kent, R. B., Ladner, R. C. and Markland, W. (1993) *Gene* 128, 29-30
- 15 Venturini, S., Martin, F. and Sollazzo, M. *Prot. Peptide Lett.* (in press)
- 16 Chothia, C. (1991) *Curr. Opin. Struct. Biol.* 1, 53-59
- 17 Cerino, A., Boender, P., La Monica, N., Rosa, C., Habets, W. and Mondelli, M. U. (1993) *J. Immunology* 151, 7005-7015
- 18 Felici, F., Luzzago, A., Folgori, A. and Cortese, R. (1993) *Gene* 128, 21-27
- 19 Bartoloni, A. et al. (1988) *Bio/Technology* 6, 83-86
- 20 Balass, M., Heldman, Y., Cabilly, S., Givol, D., Katchalski-Katzir, E. and Fuchs, S. (1994) *Proc. Natl Acad. Sci. USA* 90, 10638-10642
- 21 Lawson, D. M. et al. (1991) *Nature* 349, 541-544
- 22 Devlin, J. J., Panganiban, L. C. and Devlin, P. E. (1990) *Science* 249, 404-406
- 23 Kay, B. K., Adey, N. B., He, Y.-S., Manfredi, J. P., Mataragnon, A. H. and Fowlkes, D. M. (1993) *Gene* 128, 59-65
- 24 Weber, P. C., Pantoliano, M. W. and Thompson, L. D. (1992) *Biochemistry* 31, 9350-9354
- 25 Scott, J. K., Loganathan, D., Blaine Easley, R., Gong, X. and Goldstein, I. J. (1992) *Proc. Natl Acad. Sci. USA* 89, 5398-5402
- 26 Oldenburg, K. R., Loganathan, D., Goldstein, I. J., Schultz, P. G. and Gallop, M. A. (1992) *Proc. Natl Acad. Sci. USA* 89, 5393-5397
- 27 Hoess, R., Brinkmann, U., Handel, T. and Pastan, I. (1993) *Gene* 128, 43-49
- 28 Sinigaglia, F., Panina-Bordignon, P., Valsasini, P. and Hammer, J. in *Immunological Recognition of Peptides in Medicine and Biology* CRC (in press)
- 29 Hammer, J. et al. (1993) *Cell* 74, 197-203
- 30 Folgori, A. et al. (1994) in *Bacterial Protein Toxins*, Zbl. Bakt. Suppl. 24 (Freer et al, eds), pp. 415-426, Gustav Fischer
- 31 Folgori, A. et al. (1994) *EMBO J.* 13, 2236-2243
- 32 Greenwood, J., Willis, A. E. and Perham, R. N. (1991) *J. Mol. Biol.* 220, 821-827
- 33 Willis, A. E., Perham, R. N. and Wraith, D. (1993) *Gene* 128, 79-83