Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 1865–1869, March 1992
Biochemistry

Screening for receptor ligands using large libraries of peptides linked to the C terminus of the *lac* repressor

(recombinant diversity/C-terminal peptides/DNA binding protein/affinity purification/antibody)

MILLARD G. CULL*, JEFF F. MILLER†, AND PETER J. SCHATZ*‡

*Affymax Research Institute, 4001 Miranda Avenue, Palo Alto, CA 94304; and †Department of Microbiology and Immunology, School of Medicine, and Molecular Biology Institute, University of California, Los Angeles, CA 90024

Communicated by Jonathan Beckwith, December 9, 1991 (received for review October 21, 1991)

ABSTRACT **We have constructed a large library of random peptides fused to the C terminus of the *lac* repressor. The DNA binding activity of the repressor protein physically links the peptides to the plasmid encoding them by binding to *lac* operator sequences on the plasmid. This linkage allows efficient enrichment for specific peptide ligands in the random population of peptides by affinity purification of the peptide–repressor–plasmid complexes with an immobilized receptor. After transformation of *Escherichia coli* with recovered plasmids, the library can be amplified for additional rounds of affinity enrichment or specific plasmids can be sequenced to determine the primary structure of the peptides. We used a monoclonal antibody specific for the peptide dynorphin B as a model receptor to screen a random dodecamer library. After only two rounds of enrichment, the majority of the plasmids in the selected population encoded fusion peptides that bound specifically to the antibody. These peptides contain a consensus sequence similar to a segment of dynorphin B (RQFKVV). This technique should be useful to find peptide ligands for a variety of biological receptors.**

The isolation of ligands that bind biological receptors is fundamental to the understanding of signal transduction and in the search for therapeutics. The ability to synthesize DNA chemically has made possible the construction of extremely large collections of nucleic acid and peptide sequences as potential ligands. Recently developed methods allow efficient screening of libraries for desired binding activities. For example, RNA molecules with the ability to bind a particular protein (1) or a dye (2) have been selected by alternate rounds of affinity selection and PCR amplification. A similar technique was used to determine the DNA sequences that bound a human transcription factor (3).

Application of efficient screening techniques to peptides requires the establishment of a physical or logical connection between each peptide and the nucleic acid that encoded it. After rounds of affinity enrichment, such a connection allows amplification and sequencing of the genetic material encoding interesting peptides. Three groups, building on the fusion phage approach of Parmley and Smith (4), have described such a system where the peptide is fused to the pIII coat protein of filamentous phage (5–7). In these systems, the peptide can be expressed at the N terminus of pIII or internal to the protein. The connection between peptide and the genetic material that codes for it is established because the fusion protein is part of the capsid enclosing the phage genomic DNA. Phage encoding peptide ligands for receptors of interest can be isolated from libraries of greater than 10^8 peptides after several rounds of affinity enrichment followed by phage growth. Other systems could be suggested for the construction of peptide libraries, including direct screening of nascent peptides on polysomes (1), display of peptides on other phage, or display of peptides directly on the surface of *Escherichia coli*. As in the filamentous phage system, all of these methods would rely on a physical association of the peptide with the nucleic acid that encoded it.

We have developed a system for generating peptide libraries that relies on a DNA binding protein to establish the connection between peptide and genetic material. Our technique, "peptides-on-plasmids," allows the construction of libraries of peptides linked to the genetic material encoding them through the DNA binding activity of the *lac* repressor protein (LacI). Random peptides are fused to the C terminus of the repressor by cloning degenerate oligonucleotides at the 3′ end of the repressor gene (*lacI*) present on a plasmid. The plasmid also contains *lac* repressor binding sites, so the fusions bind the same plasmid that encodes them. After cell lysis, libraries of peptide–LacI–plasmid complexes can be screened by repeated rounds of affinity enrichment and amplification after transformation of *E. coli* (see Fig. 1).

Although many DNA binding proteins could be used in the construction of such a library, we chose the *lac* repressor for several reasons. Structure–function relationships in LacI have been studied extensively through the construction of thousands of amino acid substitution variants of the protein (8). The repressor exists as a tetramer in its native form with two high-affinity DNA binding domains formed by the N termini of the subunits (9). The C-terminal domains form the dimer and tetramer contacts, but significantly, fusions of proteins as large as β-galactosidase can be made to the C terminus without eliminating the DNA binding activity of the repressor (10). Substitutions of other sequences, including eukaryotic nuclear localization signals, transcriptional activation domains, and nuclease domains, have been made at both the N and C termini of repressor without serious disruption of specific DNA binding (11–13). The high stability of the repressor–DNA complex permits its use in methods for identifying DNA binding proteins (14), for quantifying PCR-amplified DNA (15), and for cleavage of the *E. coli* and yeast genomes at a single site (16). Here we show its use in the construction of peptide ligand libraries with greater than 10^8 members.

MATERIALS AND METHODS

Bacterial Strains. The bacterial strains used were as follows: MC1061 [*araD139* Δ(*araABC-leu*)*7696 thr* Δ*lacX74 galU galK hsdR mcrB rpsL*(*strA*) *thi*], ARI20 [F′ *lac*$^+$ *pro*$^+$ *lacI*^{q}*L8 lacIam74* ‖ Δ(*lac-pro*) *thi rpsL*(*strA*) *recA*::*cat*], XL1-Blue (F′ *proAB lacI*^{q} *lacZ*Δ*M15* Tn*10* ‖ *recA1 endA1 gyrA96 thi hsdR17 supE44 relA1 lac*).

Plasmids. The plasmids were constructed in vector pBAD18 (Luz-Maria Guzman-Verduzco, Mike Carson, and

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: BSA, bovine serum albumin.

‡To whom reprint requests should be addressed.

Jon Beckwith, personal communication). The *lacI* gene was placed under the control of the *araB* promoter with a strong consensus ribosome binding site (17), an ATG start codon at the 5′ end, and cloning sites at the 3′ end. Two $lacO_s$ sites, separated by unrelated sequence (the human D_2 dopamine receptor gene; ref. 18), were added to produce pMC5 (see Fig. 2; the complete sequence is available from the authors). Plasmid pMC3 includes sequence coding for dynorphin B (YGGFLRQFKVVT) linked to the end of the sequence encoding the GADGA linker peptide in pMC5.

Construction of a Random Dodecamer Library. After synthesis and phosphorylation, 400 pmol of ON-332, 400 pmol of ON-369, and 400 pmol of ON-370 were annealed as shown in Fig. 2 in 25-μl of reaction buffer (10 mM Tris·HCl, pH 7.4/1 mM EDTA/100 mM NaCl), by heating to 65°C for 10 min and cooling for 30 min to room temperature. The annealed oligonucleotides were ligated to 64 μg of the large pMC5 *Sfi* I–*Hin*dIII fragment at a 4:1 molar ratio in a 3.2-ml ligation reaction mixture containing 5% (wt/vol) PEG 8000, 3200 units of *Hin*dIII, 194 Weiss units of T4 ligase (New England Biolabs), 1 mM ATP, 20 mM Tris·HCl (pH 7.5), 10 mM $MgCl_2$, 0.1 mM EDTA, bovine serum albumin (BSA; 50 μg/ml), and 2 mM dithiothreitol overnight at 15°C. After ethanol precipitation, 4 μg of the ligated DNA was introduced into MC1061 by electroporation (19), to yield 5.5×10^8 transformants, which were amplified 1000-fold by growth to an A_{600} of 1 unit in 1 liter of LB medium (10 g of tryptone/5 g of yeast extract/5 g of NaCl per liter) containing ampicillin at 100 μg/ml. The cells were concentrated by centrifugation at 5500 × *g* for 6 min, washed once in ice-cold 50 mM Tris·HCl, pH 7.6/10 mM EDTA/100 mM KCl, followed by a wash in ice-cold 10 mM Tris·HCl/0.1 mM EDTA/100 mM KCl. The final pellet was resuspended in 16 ml of HEG buffer (35 mM Hepes·KOH, pH 7.5/0.1 mM EDTA/100 mM sodium glutamate), distributed into 19 1-ml tubes, frozen on dry ice, and stored at −70°C.

Panning. One sample (1.0 ml) of the library was thawed on ice and added to 9 ml of lysis buffer [35 mM Hepes (pH 7.5 with KOH)/0.1 mM EDTA/100 mM sodium glutamate/5% (vol/vol) glycerol/BSA (0.3 mg/ml)/1 mM dithiothreitol/0.1 mM phenylmethylsulfonyl fluoride]. Lysozyme was added (0.3 ml at 10 mg/ml in HEG) and the mixture was incubated on ice for 1 hr. The lysate was centrifuged at 20,000 × *g* for 15 min, and the supernatant was concentrated using a Centriprep-100 concentrator (Amicon). The concentrated supernatant (≈0.5 ml) was washed with 10 ml of HEG buffer and concentrated as before. A sample (5%) of the total lysate was removed to determine the prepanned input of plasmid complexes. Half of the remaining concentrated lysate was added to D32.39 (20)-antibody-coated sheep anti-mouse (Fc)-coupled magnetic beads (10 μg of D32.39 added to 5 mg of Dynal beads for 1 hr at 25°C followed by six washes with HEG), and half was added to uncoated beads. After incubating the lysates with the beads at 0°C for 1 hr with shaking, the beads were washed three times with 5 ml of cold HEG/0.1% BSA and then three times with HEG using a MACS 0.6 tesla magnet (Miltenyi Biotec, Bergish Gladbach, Germany) to immobilize the beads. The plasmids were dissociated from the beads by phenol extraction, and after adding 20 μg of glycogen (Boehringer Mannheim), the DNA was precipitated with an equal volume of isopropanol. The DNA was resuspended in either 4 μl (panned DNA) or 400 μl (prepanned DNA) of H_2O. Strain MC1061 was transformed using 2 μl of each DNA solution to permit counts of recovered plasmids and amplification of the selected plasmids.

ELISA. Overnight cultures from single colonies were diluted 1:10 in 3 ml of LB ampicillin (100 μg/ml) and grown 1 hr at 37°C. The expression of the LacI–peptide fusions was induced by the addition of arabinose to 0.2% for 3 hr. The cells were lysed as described above in 1 ml of lysis buffer plus lysozyme and stored at −70°C. Thawed crude lysate was added to 2 wells of a 96-well plate (100 μl per well), and the plate was incubated at 37°C. After 45 min, 100 μl of 1% BSA/PBS (10 mM sodium phosphate, pH 7.4/120 mM NaCl/2.7 mM KCl) was added for an additional 15 min at 37°C, followed by three washes with PBS/0.05% Tween 20. Each well then was blocked with 1% BSA/PBS (200 μl per well) for 30 min at 37°C and the wells were washed as before. The primary antibody D32.39 (100 μl of antibody at 1 μg/ml in PBS/0.1% BSA) was added to each well; the plate was incubated at 25°C for 1 hr and then washed as before. The plate was developed with alkaline phosphatase-conjugated goat anti-rabbit antibody (GIBCO/BRL) in PBS/0.1% BSA (100 μl per well for 1 hr at 25°C) followed by a 6-min treatment with *p*-nitrophenyl phosphate (4 mg/ml) in 1 M diethanolamine hydrochloride, pH 9.8/0.24 mM $MgCl_2$ (200 μl per well, terminated with 2 M NaOH at 50 μl per well). The A_{405} was measured on a plate reader (Biomek, Beckman, Palo Alto, CA). The positive control for the ELISA was MC1061 transformed with pMC3, encoding the LacI–dynorphin B fusion. The negative controls were wells not coated with lysate. Background variability was calculated from the wells containing lysates from 16 colonies selected at random from the library, none of which scored significantly above the negative controls. Wells were scored as positive if the measured absorbance was at least two standard deviations above background.

DNA Sequencing. Double- and single-stranded plasmid DNA, isolated from strain XL1-Blue (Stratagene), was sequenced using Sequenase (United States Biochemical) according to the manufacturer's instructions.

RESULTS

Library Design. The design of the peptides-on-plasmids library is shown in Figs. 1 and 2. The library plasmid pMC5 contains two major functional elements in a vector that permits replication and selection in *E. coli*. The *lacI* gene is expressed under the control of the *araB* promoter and has a series of restriction enzyme sites at the 3′ end of the gene. Synthetic oligonucleotides cloned into these sites fuse the repressor protein to additional peptide sequence. Each member of a library is a plasmid encoding a peptide fused to the C terminus of repressor. Because the plasmid also contains *lacO* sequences (binding sites for repressor), fusion proteins will bind to the plasmids that encode them. After these complexes form intracellularly, the cells containing a library are lysed and the complexes are partially purified away from cell debris. In a process called panning, plasmid–peptide complexes that bind specifically to immobilized receptors are separated from nonbinding complexes, which are washed

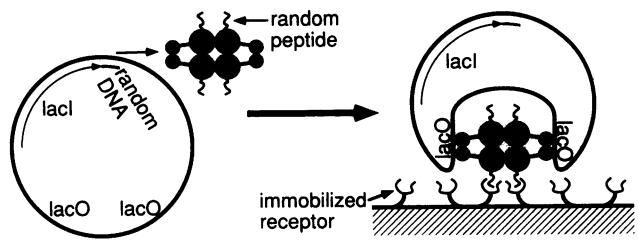

FIG. 1. Design of the peptides-on-plasmids library. The library plasmid carries the *lacI* gene with random coding sequence fused to its 3′ end. The *lac* repressor–peptide fusions produced by the hybrid genes bind to the *lacO* sites on the same plasmid that encodes them. After lysis of cells containing the random library, those plasmid–repressor–peptide complexes that specifically bind an immobilized receptor are enriched by affinity purification. Transformation of *E. coli* with recovered plasmids allows additional rounds of panning or sequencing of isolated clones.

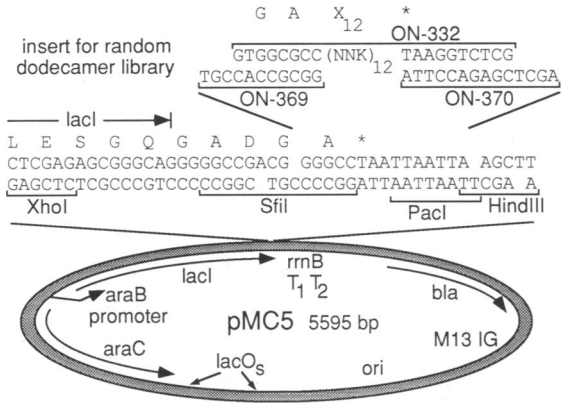

FIG. 2. Construction of the peptides-on-plasmids library. Hybridization of ON-332 to ON-369 and ON-370 produces a fragment with cohesive ends compatible with *Sfi* I/*Hin*dIII-digested plasmid pMC5. The ligation product adds sequence coding for 12 random amino acids to the end of *lacI* through a 6-codon linker. The library plasmid also contains the *rrnB* transcriptional terminator, the *bla* gene to permit selection on ampicillin, the M13 phage intragenic region to permit rescue of single-stranded DNA, a plasmid replication origin (*ori*), two $lacO_s$ sequences (with their centers spaced 326-base-pairs apart), and the *araC* gene to permit positive and negative regulation of the *araB* promoter driving expression of the *lacI* fusion gene.

away. Finally, the plasmid DNA is released from bound complexes and reintroduced into *E. coli* by transformation. By using this population of transformants, additional cycles of panning can be repeated to increase the proportion of peptides in the population that are specific for the receptor. The primary structure of the binding peptides can then be determined by sequencing the 3′ region of the *lacI* fusion gene.

For the panning process to succeed, the fusion protein must remain connected to the plasmid that encoded it throughout the procedure. Repressor binds to a single wild-type *lacO* with a dissociation constant of 1×10^{-13} M and a half-life of ≈30 min at 24°C (21). Two techniques have been used to increase the stability of the plasmid–repressor complex. The repressor tetramer has two DNA binding sites and exhibits strong cooperativity of binding to DNA molecules with two *lacO* sequences. A single tetramer can bind to suitably spaced sites on a plasmid, forming a loop of DNA between the two sites, and the resulting complex is stable for days (22–27). Plasmid pMC5 has two *lacO* sequences to take advantage of this strong cooperativity. The complex is further stabilized by using a symmetric variant of the *lacO* sequence, $lacO_s$, which has ≈10-fold higher affinity for repressor than the wild-type sequence (28, 29).

It is also important to control the ratio of repressors to plasmids so that plasmids are saturated with repressors, without a vast excess of repressor. Too little repressor could result in plasmids with free binding sites that might be filled by repressors from other cells in the population during cell lysis, thus scrambling the connection between the genetic information and the peptide ligand. Too much repressor could lead to titration of available receptor sites during panning by repressor molecules not bound to plasmid. To control this ratio, we have used the inducible *araB* promoter, which is regulated both positively and negatively by the AraC protein, also specified by the plasmid (30). This promoter can be catabolite-repressed (e.g., by adding glucose to the growth medium) and induced through the addition of L-arabinose to the medium.

A Test System. To test the peptides-on-plasmids system, we chose monoclonal antibody D32.39, which binds to dynorphin B, a 13-amino acid opioid peptide (20). As a positive control, we constructed plasmid pMC3 by attaching an oligonucleotide to the *lacI* gene, which encodes a short peptide linker (GADGA) followed by dynorphin B (YGGFLRRQFKVVT). Antibody D32.39 served two purposes. Initially, we used it, in combination with pMC3, as a reagent to measure expression of the *lacI*–dynorphin B fusion gene to determine the growth conditions needed to maintain an acceptable ratio of repressors to plasmids. The antibody served subsequently as a test receptor in panning experiments.

Western blot analysis with D32.39 revealed that the expression level of the *lacI* fusion gene could be controlled over a very wide range through changes in the growth medium. Growth in LB allowed detection of a faint band of the expected molecular weight, whereas addition of 0.2% glucose rendered this band undetectable. Growth in LB/0.2% L-arabinose led to the production of a very heavy band on a stained gel, representing >10% of the total cell protein, most of which was present in insoluble inclusion bodies (data not shown).

We estimated the *lacI* expression level necessary to fill the available binding sites in the plasmid by observing the behavior of strain ARI20 ($lacI^-$ $lacZYA^+$) transformed with pMC3 or pMC5 (encoding only the linker peptide GADGA). Since the *lacO* sites in the plasmids have higher affinity than those in the *lacZYA* operon, the available repressor should fill them first. Substantial repression of *lacZYA* should be observed only if there is an excess of repressor beyond the amount needed to fill the plasmid sites. As shown by color level on indicator plates and direct assays of β-galactosidase (31), the amount of repressor produced by pMC5 proved sufficient to fill the *lacO* sites and repress *lacZYA* in ARI20 >200-fold during growth in normal LB medium (2.4 units compared to 500 units from ARI20 transformed with the $lacI^-$ vector pBAD18). The repressor encoded by pMC3 was partially inactivated by the addition of the dynorphin B tail, allowing ≈10-fold higher expression of *lacZYA* (37 units). Because of the apparent excess production of repressor under these conditions, we used LB for all subsequent panning experiments.

Antibody D32.39 and the pMC3 complex served next as a receptor–ligand positive control in panning experiments to determine our ability to recover plasmids based on the sequence of the fusion peptide. The negative controls were pMC5, which encodes only the linker fusion peptide (GADGA), and pMC1, which encodes the dynorphin B peptide but lacks the *lacO* sequences carried by pMC3 and pMC5. We panned lysates of strains carrying each plasmid using the D32.39 antibody immobilized on polystyrene Petri dishes. After washing, we recovered plasmids from complexes bound to the plates by phenol extraction, followed by transformation of *E. coli*. Experiments with pure lysates demonstrated ≈100-fold more transformants recovered from pMC3 lysates compared to the negative controls. Experiments with mixed lysates revealed enrichment of pMC3 vs. controls among the population of recovered plasmids (data not shown). Experiments in which cells were mixed before lysis yielded similar results. We concluded that the plasmid–LacI–peptide complexes were sufficiently stable to allow enrichment of plasmids on the basis of the peptide that they encode.

Screening of a Random Dodecapeptide Library. To determine if the peptides-on-plasmids approach could be used to isolate additional ligands for a receptor, we constructed a random dodecapeptide library and screened it with the D32.39 antibody. We constructed the library in pMC5 using the half-site cloning strategy of Cwirla *et al.* (5). The random dodecamer peptide sequence, connected to the C terminus of the *lac* repressor through a linker peptide (GADGGA), is specified by a degenerate oligonucleotide population con-

taining 12 codons of the form NNK, where N is any base and K is G or T. There are 32 possible codons resulting from the NNK motif: 12 amino acids are encoded by unique codons, 5 amino acids are each encoded by 2 codons, 3 amino acids are each encoded by 3 codons, and only one of the three stop codons is possible. Transformation of strain MC1061 using 4 μg of pMC5 ligated to a 4-fold molar excess of annealed oligonucleotides yielded a test library of 5.5×10^8 clones.

We panned the test library against D32.39 antibody coupled to sheep anti-mouse-coated magnetic beads. At each round of panning, we added plasmid complexes to the beads sufficient to yield 10^{10}–10^{11} transformants (Table 1). After panning, the recovered plasmids yielded transformants ranging in number from $\approx 10^8$ in early rounds to $\approx 10^{11}$ in the fourth and final round (Table 1). Compared to the number of transformants from antibody-panned complexes, panning against unmodified polystyrene beads produced orders of magnitude fewer transformants (Table 1).

We used an ELISA to test MC1061 transformants from the second, third, and fourth rounds for D32.39-specific ligands. Of randomly picked colonies, 35 of 58 (60%) tested positive by ELISA: 11 of 20 from round two, 12 of 16 from round three, and 12 of 22 from round four. None of 16 random colonies from the unpanned library scored significantly above background. These data demonstrate the rapid enrichment of specific ligands: after only two rounds of panning, the majority of plasmids encoded peptides with affinity for the D32.39 antibody.

To determine the primary structure of the peptide ligands, we sequenced plasmids from both ELISA$^+$ and ELISA$^-$ colonies, which were obtained after panning. The translated peptide sequence for all ELISA$^+$ colonies examined defined the consensus shown in Fig. 3. The preferred recognition sequence for the D32.39 antibody apparently covers a 6-amino acid region of the dynorphin B peptide (RQFKVV). In the first position, arginine is invariant for all of the ELISA$^+$ clones. No strong bias was evident for residues in the second position. In the third position, however, aromatic residues and asparagine are favored. Phenylalanine, histidine, asparagine, tryptophan, and tyrosine, in order of frequency, account for 95% of the residues. The fourth position shows a strong bias for the positively charged residues lysine (69%) and arginine (21%). The fifth position is occupied almost exclusively by hydrophobic residues, most of which are valine (81%). Valine and threonine predominate in the sixth position (76%), with serine and isoleucine accounting for most of the remaining amino acids. It is clear from these sequences that the D32.39 antibody binds a specific amino acid motif. The absence of duplicate peptide sequences underscores the tolerance of this motif to conservative variations.

Of the ELISA$^-$ clones obtained after panning, greater than half showed peptide sequence similarity to the consensus motif (Fig. 3). None of 19 isolates sequenced from the unpanned library showed any such similarity. Some of these ELISA$^-$ sequences differ enough from the consensus that their affinity for the antibody may be insufficient to permit detection in the ELISA. There are, however, ELISA$^-$ sequences identical, in the five conserved residues of the consensus region, to ELISA$^+$ clones (e.g., sequences 28 and 57). There may be amino acids outside the consensus region that affect binding of the peptide to antibody, its susceptibility to *E. coli* proteases, or its availability in the ELISA.

Table 1. Panning of random dodecamer library against D32.39 antibody

Panning round	Number of transformants: Input	D32.39 beads	Uncoated beads
1	1.6×10^{10}	9×10^7	1.7×10^5
2	1.4×10^{11}	6.1×10^7	1.2×10^4
3	1.7×10^{11}	2.0×10^9	40
4	—	1.6×10^{11}	4×10^4

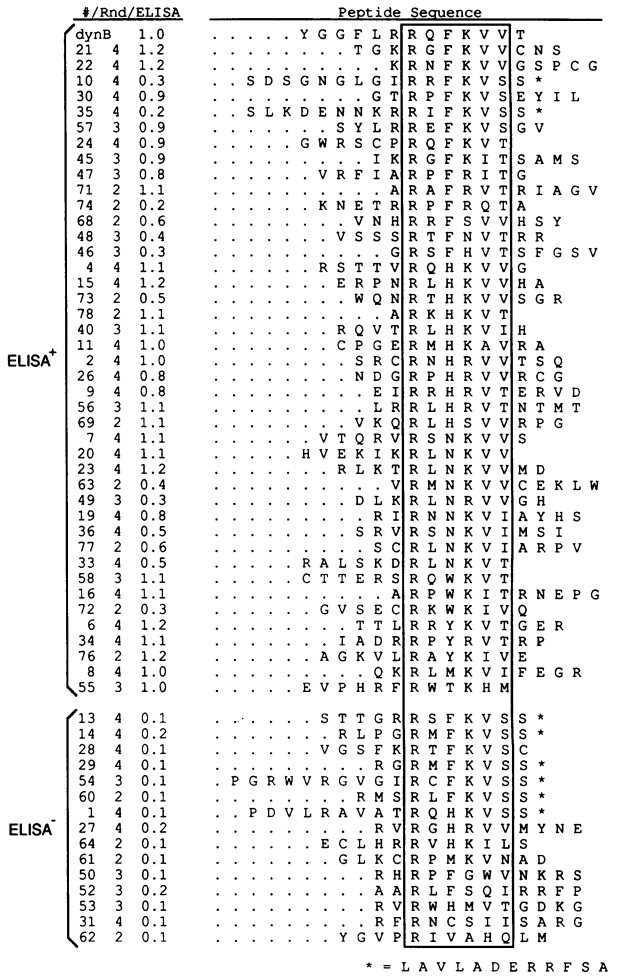

FIG. 3. ELISA$^+$ (*Upper*) and ELISA$^-$ (*Lower*) sequences isolated by panning with the D32.39 antibody. Each sequence is listed with a clone number, the panning round (Rnd) in which it was isolated, and the result of the ELISA with D32.39 antibody. The sequences are aligned to show the D32.39 epitope that they share (box). Sequences ending in * continue with the sequence shown (*Lower Right*) due to a frameshift caused by the presence of 35 instead of 36 bases in the degenerate coding region.

The requirements for enrichment of clones during panning are clearly less stringent than the requirements for scoring positive in the ELISA, perhaps due to the multivalent display of peptides on the plasmid complexes (see *Discussion*). The fact that even the ELISA$^-$ clones frequently have an obvious consensus sequence demonstrates the low background of nonspecific clones isolated using this technique. The isolation of a group of obviously related peptides from a library of $>10^8$ peptides shows the utility of the peptides-on-plasmids technique for isolating ligands for biological receptors.

DISCUSSION

We have used the DNA binding activity of *lac* repressor as a link between random peptides and the DNA encoding them to construct large peptide ligand libraries that can be efficiently screened. Plasmid–repressor–peptide complexes are isolated by panning against an immobilized receptor, the plasmids are amplified after transformation of *E. coli*, and the procedure is repeated to enrich for plasmids encoding peptides specific for the receptor. The repressor binds to the

library plasmid with sufficient avidity to allow panning of the library on immobilized receptor without problematic levels of dissociation. We have used this system to identify a series of related peptides that bind to a monoclonal antibody whose epitope was previously uncharacterized. The consensus sequence for binding is similar to sequences found using phage libraries (ref. 5; Ron Barrett, personal communication), using peptides fused to maltose binding protein (Charles Hart, personal communication), and using light-directed spatially addressable parallel chemical synthesis (ref. 32; Chris Holmes, personal communication).

Several features of peptides-on-plasmids libraries distinguish them from the phage libraries that have been described (5–7). The random peptides are displayed with a free C terminus instead of at the N terminus or internal to the carrier protein, leaving a different end of the peptides free to interact with receptors. Therefore, the use of both techniques in conjunction will increase the diversity of peptide structures that are available for receptor binding. The C-terminal mode of display also ensures that stop codons in the degenerate region, which occur with increasing frequency in longer degenerate oligonucleotides, shorten rather than destroy individual clones. The *lac* repressor fusions should allow the display of potential ligands with a wide range of sizes, given the observation that proteins as large as β-galactosidase can be fused to the repressor without eliminating DNA binding activity (10).

Characteristics of the carrier proteins impose different constraints on the two methods. The repressor fusions are cytoplasmic, unlike the phage fusions that are exported to the periplasm. Thus there is no need for peptides fused to repressor to be compatible with the protein export apparatus and the formation of an intact phage coat. The peptides need simply to be compatible with the formation of at least a repressor dimer, which is the smallest form of the protein that can bind DNA (33, 34). In addition, the use of both methods increases total available peptide diversity because the two types of libraries are exposed to different cellular compartments and thus to different sets of *E. coli* proteases.

The *lac* repressor displays multiple copies of the peptide on each library particle. Each repressor tetramer, in principle, displays four peptides that are available for binding to receptors. In addition, each plasmid monomer can bind up to two tetramers (if no loop is formed), and tandemly duplicated forms of the plasmid can display higher multiples of two tetramers. This multivalent display allows the isolation of ligands with moderate affinity ($K_d \approx 1 \times 10^{-6}$ M; ref. 5). Under multivalent conditions, these moderate-affinity ligands can obscure less-numerous high-affinity ligands. The higher-affinity ligands ($K_d \approx 1 \times 10^{-9}$ M) can be isolated from the population by panning with receptor immobilized at low density to ensure monovalent binding conditions (R. W. Barrett, S. E. Cwirla, M. S. Ackerman, A. E. Olson, E. A. Peters, and W. J. Dower, personal communication). For receptors whose normal ligands are not small peptides, this multivalency of display will be an advantage for identifying initial families of moderate-affinity ligands. Screening variants of the moderate-affinity peptides under monovalent conditions will allow selection of ligands with higher affinity.

In conclusion, we have developed a system that permits the construction of a wide variety of C-terminally exposed peptide libraries through a direct linkage of the peptide to the genetic material encoding it. This system should be useful to discover ligands for a variety of biological receptors.

We thank Ramesh Bhatt for excellent sequence analysis and Elizabeth Peters for oligonucleotide synthesis. We thank Luz-Maria Guzman-Verduzco, Mike Carson, Alan Derman, and Jon Beckwith for bacterial strains, plasmids, and advice. We thank Peter Markiewicz and Jeffrey H. Miller for strains and advice. We thank Bill Dower, Ron Barrett, Charles Hart, Bruce England, Steve Yanofsky, Erik Whitehorn, Steve Cwirla, and Martha Ackerman for advice and comments.

1. Tuerk, C. & Gold, L. (1990) *Science* **249**, 505–510.
2. Ellington, A. D. & Szostak, J. W. (1990) *Nature (London)* **346**, 818–822.
3. Thiesen, H.-J. & Bach, C. (1990) *Nucleic Acids Res.* **18**, 3203–3209.
4. Parmley, S. F. & Smith, G. P. (1988) *Gene* **73**, 305–318.
5. Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6378–6382.
6. Devlin, J. J., Panganiban, L. C. & Devlin, P. E. (1990) *Science* **249**, 404–406.
7. Scott, J. K. & Smith, G. P. (1990) *Science* **249**, 386–390.
8. Kleina, L. G. & Miller, J. H. (1990) *J. Mol. Biol.* **212**, 295–318.
9. Beyreuther, K. (1980) in *The Operon*, eds. Miller, J. H. & Reznikoff, W. S. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 123–154.
10. Müller-Hill, B. & Kania, J. (1974) *Nature (London)* **249**, 561–563.
11. Hu, M. C.-T. & Davidson, N. (1991) *Gene* **99**, 141–150.
12. Labow, M. A., Baim, S. B., Shenk, T. & Levine, A. J. (1990) *Mol. Cell. Biol.* **10**, 3343–3356.
13. Panayotatos, N., Fontaine, A. & Bäckman, S. (1989) *J. Biol. Chem.* **264**, 15066–15069.
14. Levens, D. & Howley, P. M. (1985) *Mol. Cell. Biol.* **5**, 2307–2315.
15. Lundeberg, J., Wahlberg, J. & Uhlén, M. (1991) *BioTechniques* **10**, 68–75.
16. Koob, M. & Szybalski, W. (1990) *Science* **250**, 271–273.
17. Gold, L. & Stormo, G. D. (1990) *Methods Enzymol.* **185**, 89–103.
18. England, B. P., Ackerman, M. S. & Barrett, R. W. (1991) *FEBS Lett.* **279**, 87–90.
19. Dower, W. J., Miller, J. F. & Ragsdale, C. W. (1988) *Nucleic Acids Res.* **16**, 6127–6145.
20. Barrett, R. W. & Goldstein, A. (1985) *Neuropeptides* **6**, 113–120.
21. Barkley, M. D. & Bourgeois, S. (1980) in *The Operon*, eds. Miller, J. H. & Reznikoff, W. S. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 177–220.
22. Besse, M., von Wilcken-Bergmann, B. & Müller-Hill, B. (1986) *EMBO J.* **5**, 1377–1381.
23. Flashner, Y. & Gralla, J. D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8968–8972.
24. Hsieh, W.-T., Whitson, P. A., Matthews, K. S. & Wells, R. D. (1987) *J. Biol. Chem.* **262**, 14583–14591.
25. Krämer, H., Niemöller, M., Amouyal, M., Revet, B., Wilcken-Bergmann, B. & Müller-Hill, B. (1987) *EMBO J.* **6**, 1481–1491.
26. Mossing, M. C. & Record, M. T., Jr. (1986) *Science* **233**, 889–892.
27. Whitson, P. A., Hsieh, W.-T., Wells, R. D. & Matthews, K. S. (1987) *J. Biol. Chem.* **262**, 14592–14599.
28. Sadler, J. R., Sasmor, H. & Betz, J. L. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6785–6789.
29. Simons, A., Tils, D., von Wilcken-Bergmann, B. & Müller-Hill, B. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1624–1628.
30. Lee, N. (1980) in *The Operon*, eds. Miller, J. H. & Reznikoff, W. S. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 389–409.
31. Miller, J. H. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
32. Fodor, S. P. A., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T. & Solas, D. (1991) *Science* **251**, 767–773.
33. Daly, T. J. & Matthews, K. S. (1986) *Biochemistry* **25**, 5474–5478.
34. Kania, J. & Brown, D. T. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3529–3533.