Generation and screening of an oligonucleotide-encoded synthetic peptide library

(encoded synthetic libraries/combinatorial chemistry/peptide diversity)

MICHAEL C. NEEDELS, DAVID G. JONES, EMILY H. TATE, GREGORY L. HEINKEL, LYNN M. KOCHERSPERGER, WILLIAM J. DOWER, RONALD W. BARRETT, AND MARK A. GALLOP*

Affymax Research Institute, 4001 Miranda Avenue, Palo Alto, CA 94304

Communicated by Peter G. Schultz, August 2, 1993

ABSTRACT We have prepared a library of $\approx 10^6$ different peptide sequences on small, spherical (10-μm diameter) beads by the combinatorial chemical coupling of both L- and D-amino acid building blocks. To each bead is covalently attached many copies of a single peptide sequence and, additionally, copies of a unique single-stranded oligonucleotide that codes for that peptide sequence. The oligonucleotide tags are synthesized through a parallel combinatorial procedure that effectively records the process by which the encoded peptide sequence is assembled. The collection of beads was screened for binding to a fluorescently labeled anti-peptide antibody using a fluorescence-activated cell sorting instrument. Those beads to which the antibody bound tightly were isolated by fluorescence-activated sorting, and the oligonucleotide identifiers attached to individual sorted beads were amplified by the PCR. Sequences of the amplified DNAs were determined to reveal the identity of peptide sequences that bound to the antibody with high affinity. By combining the capacity for information storage in an oligonucleotide code with the tremendous level of amplification possible through the PCR, we have devised a means for specifying the identity of each member of a vast library of molecules synthesized from both natural and unnatural chemical building blocks. In addition, we have shown that the use of flow cytometry instrumentation permits facile isolation of individual beads that bear high-affinity ligands for biological receptors.

Ligands for macromolecular receptors can be identified by screening diverse collections of peptides produced through either molecular biological or synthetic chemical techniques. Recombinant peptide libraries have been generated by inserting degenerate oligonucleotides into genes encoding capsid proteins of filamentous bacteriophage (1–3) or the DNA-binding protein Lac I (4). These random libraries may contain >10^9 different peptides, each fused to a larger protein sequence that is physically linked to the genetic material encoding it. Chemical approaches to generating peptide libraries are not limited to combinatorial syntheses using just the 20 genetically coded amino acids. By expanding the building block set to include unnatural amino acids, the accessible sequence diversity is dramatically increased. In several of the strategies described for creating synthetic peptide libraries, the different peptides are tethered to a solid support in a spatially segregated manner (5, 6). Large libraries of soluble peptides have been prepared as peptide pools using the ''tea-bag'' method of multiple synthesis. Active peptides within these degenerate mixtures are identified through an iterative process of screening and sub-library resynthesis (7, 8).

Using the split-synthesis protocol pioneered by Furka *et al.* (9), Lam *et al.* (10) have prepared libraries containing $\approx 10^6$ peptides attached to 100- to 200-μm-diameter resin beads. The bead library is screened by incubation with a labeled receptor: beads binding to the receptor are identified by visual inspection and are selected with the aid of a micro-manipulator. Each bead contains 50–200 pmol of a single peptide sequence that may be determined directly either by Edman degradation or mass spectrometry analysis. In principle, one could create libraries of greater diversity using this approach by reducing the bead dimensions. The sensitivity of peptide-sequencing techniques is limited to ≈ 1 pmol, however, clearly limiting the scope of direct peptide-sequencing analysis. Moreover, neither analytical method provides for straightforward and unambiguous sequence analysis when the library building block set is expanded to include D- or other nonnatural amino acids.

We recognized that the products of a combinatorial peptide synthesis on resin beads could be explicitly specified if it were possible to attach an identifier tag to the beads coincident with each amino acid-coupling step in the synthesis. Each tag would then convey which amino acid monomer was coupled in a particular step of the synthesis, and the overall sequence of a peptide on any bead could be deduced by reading the set of tags on that bead. We now describe the use of single-stranded oligonucleotides to encode a combinatorial synthesis on 10-μm-diameter polystyrene beads. Peptides and nucleotides are assembled in parallel, alternating syntheses so that each bead bears many copies of both a single peptide sequence and a unique oligonucleotide identifier tag. We have generated an encoded synthetic library of some 8.2 $\times 10^5$ heptapeptides and screened it for binding to an anti-dynorphin B monoclonal antibody (mAb) D32.39 (4), using a fluorescence-activated cell sorting (FACS) instrument to select individual beads that strongly bind the antibody.

MATERIALS AND METHODS

Reagents and General Methods. The monodisperse 10-μm-diameter bead material used in this work was a custom-synthesized macroporous styrene/divinylbenzene copolymer functionalized with a 1,12-diaminododecane linker. All protected amino acids were obtained from Bachem.

Parallel Synthesis of a 69-Base Oligonucleotide and the Peptide H-Arg-Gln-Phe-Lys-Val-Val-Thr-NH$_2$ (RQFKVVT). The C-terminal seven amino acid fragment of the opioid peptide dynorphin B was synthesized in parallel with a 69-mer oligodeoxyribonucleotide (ST08) on 10-μm-diameter beads. The sequence of ST08 was 5'-ATC CAA TCT CTC CAC ATC TCT ATA CTA TCA TCA CC [TA TC CT AT TT

Abbreviations: Fmoc, 9-fluorenylmethoxycarbonyl; DMT, dimethoxytrityl; mAb, monoclonal antibody; FACS, fluorescence-

TT AC] CTC ACT CAC TTC CAT TCC AC-3'. Underlined portions of this sequence correspond to PCR-priming sites, while the region in italics is homologous to the primer used for sequencing this template. The 14-base sequence enclosed in brackets represents the coding region of the template.

The beads were first treated with a mixture of succinimidyl 4-O-DMT-oxybutyrate (where DMT is dimethoxytrityl; Molecular Probes) and the 1-oxybenzotriazole ester of either N-Fmoc-2,4-dimethoxy-4'-(carboxymethyloxy)benzhydrylamine (i.e., the acid-cleavable Knorr carboxamide linker, where Fmoc is 9-fluorenylmethoxycarbonyl) or N-Fmoc-Thr(*tert*-butyl)-OH (for noncleavable experiments). The ratio of Fmoc-protected amino groups to DMT-protected hydroxyl residues on the beads was determined spectrophotometrically to be $\approx$20:1. The beads were subjected to 20 cycles of oligonucleotide synthesis on an automated synthesizer using 3'-O-methyl-N,N-diisopropyl phosphoramidites of the following nucleosides: N^6-benzoyl-5'-O-DMT-(7-deaza)-2'-deoxyadenosine (Berry and Associates, Ann Arbor, MI), N^4-benzoyl-5'-O-DMT-2'-deoxycytidine, and 5'-O-DMT-thymidine (Glen Research, Sterling, VA). The beads were then removed from the instrument and treated for 5 min with 10% (vol/vol) piperidine in dimethylformamide to remove the Fmoc protecting group. After coupling the first amino acid residue [N-Fmoc-Thr(*tert*-butyl)-OH], the beads were treated with a tetrahydrofuran solution of acetic anhydride and 1-methylimidazole to cap any unreacted amines. All peptide coupling reactions were run for 20 min and contained 0.11 M Fmoc-amino acid, 0.1 M O-(benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate, 0.1 M 1-hydroxybenzotriazole, and 0.3 M diisopropylethylamine in dimethylformamide/CH_2Cl_2, 1:1. The beads were then subjected to two cycles of nucleotide addition on the synthesizer (detritylation with trichloroacetic acid; tetrazole-catalyzed phosphitylation; capping with acetic anhydride; oxidation with iodine in acetonitrile/water). Sequential steps of amino acid coupling and dinucleotide addition were repeated until synthesis of the peptide sequence RQFKVVT and construction of the oligonucleotide coding region had been completed. After an additional 35 cycles of oligonucleotide synthesis, the beads were treated sequentially with piperidine/dimethylformamide, 1:9 for 8 min; thiophenol/triethylamine/dioxane, 1:2:2, for 4 hr; ethylenediamine/ethanol, 1:1, for 5 hr at 55°C; and trifluoroacetic acid/water, 20:1, for 1 hr to fully deprotect both the peptide and oligonucleotide chains. In experiments using the acid-cleavable linker, the supernatant from the trifluoroacetic acid deprotection reaction was concentrated *in vacuo*, and the isolated crude peptide was then analyzed by HPLC using a Rainin reverse-phase C_{18} column.

Construction of an Encoded Library. The parallel synthesis chemistry outlined above was used in the construction of the library. The sites of peptide synthesis were differentiated from DNA-synthesis sites in this experiment by coupling to all beads a mixture of N-Fmoc-Thr(*tert*-butyl)-oxybenzotriazole and succinimidyl 4-O-DMT-oxybutyrate, as has been described. Sequences of oligonucleotide tags in the library deviated from ST08 only within the coding region. The 3'-conserved region of the oligonucleotide ST08 was first synthesized on a total bead mass of 35 mg ($\approx$1.75 × 10^8 beads). The Fmoc protecting group was removed, and the bead mass was divided into seven equal parts. To each aliquot was coupled one of seven different α-N-Fmoc-protected amino acids (side-chain protecting groups are shown in parenthesis): Arg(N^G-2,2,5,7,8-pentamethylchroman-6-sulfonyl), Gln(trityl), Phe, Lys(*tert*-butoxycarbonyl), Val, D-Val, and Thr(*tert*-butyl). Each part was then subjected to two rounds of automated oligonucleotide synthesis. The respective sequences of the appended dinucleotides that

pooled and mixed thoroughly; the entire bead mass was then subjected to Fmoc deprotection. This cycle of bead partitioning, peptide coupling, oligonucleotide-dimer synthesis, bead recombination, and Fmoc removal was repeated for a total of seven times. The final Fmoc protecting group was not removed. Rather, the pooled bead mass was subjected to 35 cycles of oligonucleotide synthesis. The library was then fully deprotected as described above.

FACS Analysis of Antibody Binding to Beads. A portion of the library (typically 0.5–2 mg of beads) was suspended in blocking buffer [phosphate-buffered saline (PBS)/1% bovine serum albumin/0.05% Tween-20] and incubated at room temperature for 1 hr. The beads were pelleted by centrifugation and resuspended in a solution of mAb D32.39 (10 μg/ml in blocking buffer) (4). The suspension was incubated on ice for 30 min, pelleted by centrifugation, and washed with blocking buffer. Beads were then suspended in a solution of phycoerythrin-conjugated goat anti-mouse antibody (Molecular Probes) for 20 min on ice. The beads were washed in blocking buffer and diluted in PBS for delivery into the FACS instrument (Becton Dickinson FACStar^{Plus}). Beads that had bound mAb D32.39 were identified by their acquired fluorescence. Individual beads from both the most brightly stained 0.17% of the library and from the region having the lowest fluorescence ($\approx$98%) were sorted into PCR microcentrifuge vials.

PCR of Bead-Bound Template and Sequencing of PCR Product. PCRs consisting of 45 amplification cycles were done with *Taq* polymerase (Perkin–Elmer) according to the manufacturer's instructions. The reactions contained dUTP and uracil DNA glycosidase (GIBCO/BRL) to prevent carryover contamination with soluble product from previous amplifications (11). Biotinylated PCR product from individual reactions was isolated with streptavidin-coated magnetic beads (Dynal, Great Neck, NY). After alkaline elution of the nonbiotinylated strand and washing, each bead sample was treated with sequencing mixture. Dideoxynucleotide sequencing was done by using the primer 5'-ATC TCT ATA CTA TCA-3' (SP15) and *Bst* polymerase (Bio-Rad), according to the manufacturer's instructions, except that a 1:100 ratio of deoxy- to dideoxynucleotide triphosphates (Pharmacia) was used.

Determination of Peptide-Binding Affinities. Binding affinities of various peptides for mAb D32.39 were measured in a competition binding experiment. A tracer peptide (LRRASLGGGRRQFKVVT; 50 pM) containing the known epitope for mAb D32.39 fused to a consensus substrate sequence for cAMP-dependent protein kinase was radiolabeled to high specific activity with [γ-^{33}P]ATP (12) and mixed with various concentrations of the peptide of interest (10 μM to 1 pM). The peptide mixtures were added to polystyrene wells coated with mAb D32.39 (0.1 μg/ml). Samples were incubated 2 hr at 4°C, the wells were washed with PBS, and the radioactivity associated with each well was counted and used to generate a competitive binding curve. Under the conditions of the assay the IC_{50} should be close to the dissociation constant (K_d) for the peptide.

RESULTS

Establishing a practical bead-based oligonucleotide-encoded peptide library methodology demands that several key technical criteria be met. These criteria include (*i*) the development of mutually compatible chemistries for parallel assembly of peptides and oligonucleotides, (*ii*) the selection of bead material with appropriate physical characteristics, (*iii*) the facile isolation of small beads bearing ligands that bind a target receptor, and (*iv*) successful PCR amplification and

from a macroporous styrene/divinylbenzene copolymer and derivatized with a dodecylamine linker are generally satisfactory for this work. The amino group loading of these beads was estimated to be ≈100 μmol/g by exhaustive acylation with Fmoc-glycine, followed by piperidine cleavage of the Fmoc group and spectrophotometric quantitation of the released piperidine–dibenzofulvene adduct (ε_{302} = 7800 liter·mol^{-1}·cm^{-1}). With 5×10^9 beads per g, this corresponds to a maximum peptide loading of ≈20 fmol per bead. Acylation of the beads with a mixture of an appropriately protected amino acid and an ω-hydroxy acid provided orthogonally differentiated amino and hydroxyl groups from which the peptide and nucleotide chains, respectively, could be extended. The average stoichiometry of peptide to oligonucleotide per bead is controlled by varying the ratio of amino and hydroxy acids coupled to the initial bead mass. Test peptide syntheses (5-mers to 12-mers) on these beads equipped with a trifluoroacetic acid-cleavable Knorr linker (13) using standard Fmoc chemistry were found to proceed with high fidelity that was indistinguishable from syntheses performed on conventional peptide synthesis resin, as determined by HPLC analysis of the crude cleaved peptide carboxamides (data not shown).

Parallel Synthesis of Peptides and Oligonucleotides. Parallel synthesis strategies require (*i*) the use of a set of protecting groups on the amino acids and nucleotide building blocks that are mutually orthogonal and (*ii*) that each of the polymer chains be stable to the reagents used in the synthesis and deprotection of the second chain. Although, in principle, a variety of protection/deprotection schemes could be envisaged, we preferred to use Fmoc/*tert*-butyl protection on the peptide building blocks because of the extensive commercial availability of natural and unnatural amino acids protected in this manner. However, the *tert*-butyl-based peptide side-chain protecting groups require treatment with strong acid (typically trifluoroacetic acid) for removal, conditions that lead to rapid depurination of oligonucleotides containing either 2′-deoxyadenosine (dA) or 2′-deoxyguanosine (dG) (14). This problem has been circumvented by using 7-deaza-2′-deoxyadenosine (c^{7}dA) for dA in the template oligonucleotide tag. The glycosidic bonds of deazapurine nucleosides are resistant to acid-catalyzed hydrolysis (15), and oligonucleotides incorporating these monomers are faithfully copied by thermostable polymerases used in the PCR (16, 17). Although not used in this work, acid-resistant guanosine analogs could also be incorporated into the template DNA.

5′-*O*-dimethoxytrityl 2′-deoxynucleoside 3′-(*O*-methyl-*N,N*-diisopropyl)phosphoramidites were used in all parallel syntheses. The reagent (I$_2$/collidine/H$_2$O/acetonitrile) used to convert the nucleotide phosphite intermediates to phosphotriesters in the DNA-synthesis protocol was not found to adversely affect either the readily oxidized residues tryptophan and methionine or any of the other protected amino acids used in this work (data not shown). Complete removal of the 5′-*O*-DMT group from the growing oligonucleotide chain was achieved in ≈40 sec using 1% trichloroacetic acid in dichloromethane, whereas all of the acid-labile side-chain protecting groups used conventionally in Fmoc/*tert*-butyl chemistry were inert to treatment with 1% trichloroacetic acid for 1 hr. Quantitative deprotection of the α-amino residues required 5- to 10-min treatment with piperidine/dimethylformamide (10% vol/vol) and also resulted in partial demethylation of the protected polynucleotide phosphotriesters ($t_{1/2} \approx 45$ min). Control experiments indicated that any aberrant phosphitylation of the resulting phosphodiester species during subsequent nucleotide chain elongation was reversed by the final oligonucleotide deprotection steps, as noted by other workers (18). At the completion of the parallel

olic ethylenediamine (debenzoylation of protected cytidine and 7-deaza-adenine residues). These mild, anhydrous aminolysis conditions did not adversely affect protected peptide sequences (19), which were deblocked using trifluoroacetic acid under standard conditions.

The carboxyl-terminal region of opioid peptide dynorphin B (YGGFLRRQFKVVT) has been previously shown to represent the epitope of anti-dynorphin B mAb D32.39 (4): the soluble heptapeptide RQFKVVT binds mAb D32.39 with high affinity ($K_d \approx 0.5$ nM). To test the efficacy of our chemical methods, a parallel synthesis of this peptide and a 69-base oligodeoxyribonucleotide was performed on orthogonally differentiated beads bearing an acid-cleavable Fmoc-protected carboxamide (Knorr) linker. The beads were exposed to full oligonucleotide and then peptide-deprotection conditions, and the trifluoroacetic acid supernatant containing the cleaved peptide was analyzed by reverse-phase HPLC. Fig. 1a shows that the crude peptide from the parallel synthesis consists of a single major component (coeluting with authentic RQFKVVT; data not shown) and that this crude product is not significantly different from that generated in a control peptide synthesis in which no oligonucleotide chemistry occurred (Fig. 1b). Fig. 2 demonstrates that the integrity of the DNA template containing c^{7}dA was also maintained through the course of the parallel synthesis chemistry.

Construction of a Large Encoded Combinatorial Library. An encoded library designed to contain 823,543 (7^7) different heptapeptides attached to 10-μm beads was constructed by a combinatorial synthesis using the seven amino acids arginine, glutamine, phenylalanine, lysine, valine, D-valine, and threonine. α-*N*-Fmoc-protected threonine and *O*-DMT-protected γ-oxybutyrate residues were first coupled to all the beads to provide the orthogonally differentiated amino and hydroxyl groups for this synthesis. Starting with a total bead mass of 35 mg (1.75×10^8 beads) ensured that each peptide sequence appeared on ≈200 different beads in the library. Peptide microsequencing analysis of an aliquot of the library confirmed that the seven amino acids were stochastically distributed among every position of the degenerate heptapeptide mixture (note that L-valine and D-valine are indistinguished in the Edman degradation procedure).

The binding of mAb D32.39 to control beads and to the bead library was analyzed by flow cytometry. Fig. 3a shows that beads carrying the positive control sequence RQFKVVT
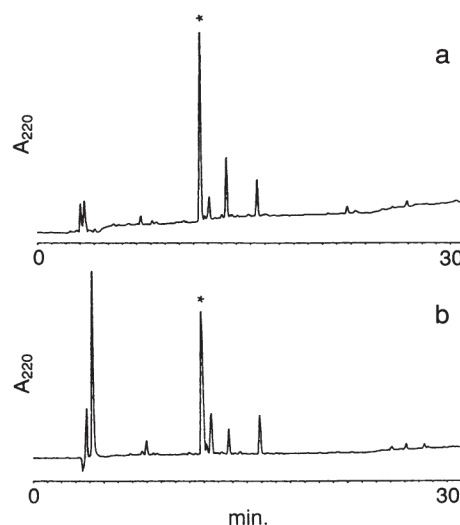

FIG. 1.   Reversed-phase HPLC chromatograms of crude peptide RQFKVVT. Asterisks mark peak corresponding to authentic mate-
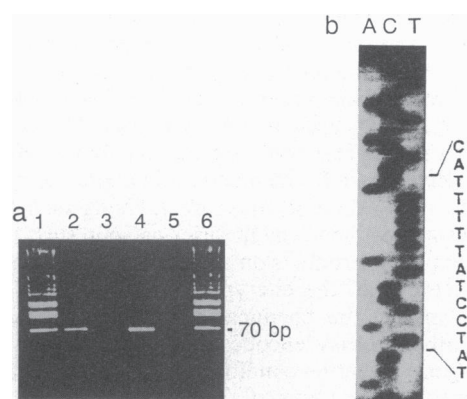
FIG. 2. Amplification and sequence analysis of oligonucleotide ST08 synthesized in parallel with peptide RQFKVVT. (*a*) Ethidium bromide-stained agarose gel electrophoresis of products from PCR amplifications of ST08 template attached to single sorted beads. Lanes: 1 and 6, DNA markers; 2, c^7dA-containing template from parallel synthesis; 3, dA-containing template after 1-hr treatment with 95% trifluoroacetic acid/5% H_2O; 4, untreated dA-containing template; 5, zero bead control. (*b*) Sequencing gel of the PCR amplification product from an individual bead; DNA sequence of the template coding region is shown at right.

FIG. 3. Flow cytometric analysis of binding of mAb D32.39 to 10-μm beads bearing peptide and oligonucleotide. Approximately 10^5 events are recorded in each experiment. Fluorescence intensity is shown on vertical axis. (*a*) A 1:1 mixture of underivatized (lower population) and RQFKVVT (upper population) beads as negative and positive controls. (*b*) Binding of mAb to library. (*c*) Specific binding to library blocked by preincubation of mAb with 10 μM RQFKVVT.

and a 69-mer oligonucleotide tag are strongly stained by the antibody, whereas blank beads are unstained. By contrast, only a small fraction of the encoded library bound mAb D32.39 (see Fig. 3*b*). Analysis of 10^5 events indicated that $\approx 2\%$ of the library stained above background levels. Significantly, this binding to mAb D32.39 was specific for the combining site, as it could be completely blocked by preincubating the mAb with soluble RQFKVVT peptide (Fig. 3*c*). Individual beads from the library having fluorescence intensities comparable with the positive control beads were sorted into microcentrifuge tubes for tag amplification by PCR (beads with fluorescence in the top 0.17% of the population were collected). Nucleotide sequences were obtained from 12 sorted beads, and the deduced peptide sequences are given in Table 1. Representative peptide sequences obtained from single beads having fluorescence that was not significantly above background are also tabulated for comparison.

These data are consistent with an earlier study showing that the preferred recognition sequence of mAb D32.39 is localized to the six-amino acid fragment RQFKVV of dynorphin B (4). Interestingly, D-valine appears best tolerated at positions outside the consensus motif. The range of affinities of peptides that were selected ($K_d \approx 0.3$–1400 nM) was not unexpected, given the design of the binding assay—i.e., bivalent primary antibody with labeled second antibody detection. We anticipate that manipulation of the binding valency (for example, directly labeled monovalent receptor)
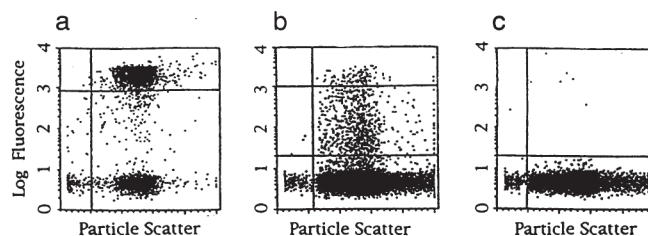
and the stringency of wash conditions will improve the capacity to isolate only the highest-affinity ligands.

DISCUSSION

We have developed chemistry to prepare a highly diverse oligonucleotide-encoded synthetic peptide library on microscopic beads by combinatorial synthesis. While this work was in progress, the concept of an oligonucleotide-encoded chemical synthesis was proposed independently by Brenner and Lerner (20). More recently, two other groups have shown that an L-amino acid peptide strand may be used to encode the combinatorial assembly of molecular structures that are not amenable to direct sequence analysis (21, 22). It seems likely that constraints on the sensitivity and throughput of the Edman procedure will ultimately restrict the scope of this peptide-coding approach to analyzing libraries of limited diversity.

Encoding a combinatorial synthetic procedure with oligonucleotides provides a mechanism for addressing the major limitations of ambiguity and sensitivity encountered in the direct structural analysis of minute quantities of ligands isolated from large libraries. The high capacity of DNA for information storage can be exploited to archive the precise details of a library's construction. In the example above, we used a "codon" structure of two contiguous nucleotides comprising three bases (c^7 dA, dC, and T), capable of encoding a synthesis incorporating up to $3^2 = 9$ amino acid building blocks (only seven were used in this library). If c^7 dG were also included in the coding template, then a combinatorial synthesis using 1000 different monomers could be accommodated by using a "codon" size of just 5 nt ($4^5 = 1024$).

A second outstanding advantage inherent in using an oligonucleotide-based coding scheme is the ability to achieve

Table 1. Amino acid sequences of peptides on beads that bind mAb D32.39

High fluorescence intensity				Low fluorescence intensity	
Sequence	K_d, nM	Sequence	K_d, nM	Sequence	K_d, mM
TFRQFKV(T)	0.29	QQFKVVQ(T)	370	QTvTvKK(T)	>1
TTRRFRV(T)	4.3	KQFKVTQ(T)	410	QQVQRQT(T)	>0.4
TVRQFKT(T)	8.8	TQFKVTK(T)	560	KTQvVQF(T)	ND
QvRQFKT(T)	16	TFRvFRV(T)	1400	QvTQvRV(T)	ND
RQFRTVQ(T)	76	FRRQFRV(T)	ND	FVVTVRV(T)	ND
KQFKVTK(T)	340	RQFKQVQ(T)	ND		
RQFKVVT	0.51	(positive control)			

A library of peptide-bearing beads was screened for antibody ligands by using an indirect fluorescence assay and FACS instrumentation. Sequences from the highly fluorescent beads are aligned to

tremendous levels of target amplification through the PCR. We are therefore able to work with tiny quantities of DNA template and, hence, to use solid supports of microscopic dimensions in our syntheses. This will facilitate the construction and screening of libraries that far exceed the diversity accessible through other tethered synthetic library techniques. Moreover, these libraries will employ manageable quantities of bead material that can therefore be assayed for receptor binding using practical volumes of biological reagents.

Standard FACS instrumentation permits bead (cell) fluorescence analysis rates of $\approx 10^4$ events per sec or, when operated in single bead cloning mode, sort rates that are 5- to 10-fold slower. In assaying very large libraries (e.g., $>>10^7$ beads) some form of affinity-selective prescreen could be used before individual bead isolation with the cell sorter. For example, receptor-coated sub-micron-sized superparamagnetic particles are frequently used to affinity-purify specific cells from large, mixed populations by magnetic activated sorting (23). It should be noted that to have a high probability of detecting very rare binding events, it is essential that each compound be present on many beads in the library. A practical upper limit for the size of an encoded library constructed from 10-μm particles, assuming a 100-fold redundancy, is probably 10^8–10^9 compounds synthesized on $\approx 10^{10}$–10^{11} beads. Even larger libraries might be prepared using smaller beads, but conventional cytometers are unlikely to detect or manipulate particles much less than ≈ 1 μm in diameter.

High-throughput screening of collections of chemically synthesized molecules and of natural products (such as microbial fermentation broths) has traditionally played a central role in the search for lead compounds for the development of new pharmacological agents. The remarkable surge of interest in combinatorial chemistry and the associated technologies for generating and evaluating molecular diversity represent significant milestones in the evolution of this paradigm of drug discovery (24). To date, peptide chemistry has been the principle vehicle for exploring the utility of combinatorial methods in ligand identification. This may be ascribed to the availability of a large and structurally diverse range of amino acid monomers, a relatively generic, high-yielding solid-phase coupling chemistry, and the synergy with biological approaches for generating recombinant peptide libraries. Moreover, the potent and specific biological activities of many low-molecular-weight peptides make these molecules attractive starting points for therapeutic drug discovery. Unfavorable pharmacodynamic properties, such as poor oral bioavailability and rapid clearance *in vivo*, have limited the more widespread development of peptidic compounds as drugs, however. This realization has recently inspired workers to extend the concepts of combinatorial organic synthesis beyond peptide chemistry to create libraries of known pharmacophores like benzodiazepines (25), as well as unnatural polymeric molecules, such as oligomeric N-substituted glycines ("peptoids") (26) and oligocarbamates (27). We anticipate that the coding strategy illustrated here could greatly facilitate the screening of large collections of compounds that cannot be directly sequenced and are produced by multistep synthesis.

We have shown here that oligonucleotide encryption provides a powerful mechanism for recording the structural identity of every member of a vast library of tethered peptides generated through a combinatorial synthesis. This technique should also be broadly applicable to encoding the combinatorial assembly of other nonpeptidic structures, providing the parallel synthetic schemes remain orthogonal and compatible. The net outcome of a combinatorial synthesis is unam-

proceed in very high yield to afford single products. This situation is approximated by peptide and DNA synthesis chemistries, and the resulting product structures are explicitly specified by the order of the building blocks and/or coupling reactions used in the synthesis. However, most synthetic organic reactions are more idiosyncratic, giving variable yields and frequently multiple products (such as regio- and stereoisomeric structures). Using such chemistry to synthesize combinatorial libraries on solid supports would yield a mixture of products on each bead in the library. In the most general case, the encryption of a synthesis may not uniquely specify the chemical structure of an associated entity. Rather, it may encode the exact synthetic protocol (e.g., reagents, reaction conditions, etc.) by which a member of the library was constructed. The library would be screened to identify "active recipes," which then could be reproduced on a preparative scale and fractionated (if necessary) to isolate the bioactive component(s). Encoded library technologies have considerable potential to expand the scope of combinatorial chemistry and its applications to drug discovery.

1. Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6378–6382.
2. Scott, J. K. & Smith, G. P. (1990) *Science* **249**, 386–390.
3. Devlin, J. J., Panganiban, L. C. & Devlin, P. E. (1990) *Science* **249**, 404–406.
4. Cull, M. G., Miller, J. F. & Schatz, P. J. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 1865–1869.
5. Geysen, H. M., Meloen, R. H. & Barteling, S. J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3998–4002.
6. Fodor, S. P. A., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T. & Solas, D. (1991) *Science* **251**, 767–773.
7. Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T. & Cuervo, J. H. (1991) *Nature (London)* **354**, 84–86.
8. Blake, J. & Litzi-Davis, L. (1992) *Bioconjugate Chem.* **3**, 510–513.
9. Furka, A., Sebestyen, F., Asgedom, M. & Dibo, G. (1991) *Int. J. Pept. Protein Res.* **37**, 487–493.
10. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. & Knapp, R. J. (1991) *Nature (London)* **354**, 82–84.
11. Longo, M., Berninger, M. & Hartley, J. (1990) *Gene* **93**, 125–128.
12. Li, B.-L., Langer, J. A., Schwartz, B. & Pestka, S. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 558–562.
13. Bernatowicz, M. S., Daniels, S. B. & Koster, H. (1989) *Tetrahedron Lett.* **30**, 4645–4648.
14. Capon, B. (1969) *Chem. Rev.* **69**, 407–498.
15. Scheit, K. H. (1980) *Nucleotide Analogs: Synthesis and Biological Function* (Wiley, New York), pp. 64–65.
16. McConlogue, L., Brown, M. A. D. & Innis, M. A. (1988) *Nucleic Acids Res.* **16**, 9869.
17. Barr, P. J., Thayer, R. M., Laybourn, P., Najarian, R. C., Seela, F. & Tolan, D. R. (1986) *BioTechniques* **4**, 428–432.
18. Lehmann, C., Xu, Y.-Z., Christodoulou, C., Tan, Z.-K. & Gait, M. J. (1989) *Nucleic Acids Res.* **17**, 2379–2390.
19. Juby, C. D., Richardson, C. D. & Brousseau, R. (1991) *Tetrahedron Lett.* **32**, 879–882.
20. Brenner, S. & Lerner, R. A. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 5181–5183.
21. Kerr, J. M., Banville, S. C. & Zuckermann, R. N. (1993) *J. Am. Chem. Soc.* **115**, 2529–2531.
22. Nikolaiev, V., Stierandova, A., Krchnak, V., Seligmann, B., Lam, K. S., Salmon, S. E. & Lebl, M. (1993) *Pept. Res.* **6**, 161–170.
23. Miltenyi, S., Muller, W., Weichel, W. & Radbruch, A. (1990) *Cytometry* **11**, 231–238.
24. Pavia, M. R., Sawyer, T. K. & Moos, W. H. (1993) *Bioorg. Med. Chem. Lett.* **3**, 387–396.
25. Bunin, B. A. & Ellman, J. A. (1992) *J. Am. Chem. Soc.* **114**, 10997–10998.
26. Simon, R. J., Kania, R. S., Zuckermann, R. N., Huebner, V. D., Jewell, D. A., Banville, S., Ng, S., Wang, L., Rosenberg, S., Marlowe, C. K., Spellmeyer, D. C., Tan, R., Frankel, A. D., Santi, D. V., Cohen, F. E. & Bartlett, P. A. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 9367–9371.
27. Cho, C. Y., Moran, E. J., Cherry, S., Stephans, J., Fodor, S. P. A., Adams, C., Sundaram, A., Jacobs, J. W. & Schultz, P. G. (1993)