

## Complex synthetic chemical libraries indexed with molecular tags

(combinatorial chemistry/encoded libraries/peptides/antibody recognition)

MICHAEL H. J. OHLMEYER\*, ROBERT N. SWANSON†, LAWRENCE W. DILLARD\*, JOHN C. READER\*, GIGI ASOULINE†, RYUJI KOBAYASHI†, MICHAEL WIGLER†, AND W. CLARK STILL\*

\*Department of Chemistry, Columbia University, New York, NY 10027; and †Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724

Contributed by Michael Wigler, August 4, 1993

**ABSTRACT** Combinatorial methods of chemical synthesis allow the creation of molecular libraries having immense diversity. The utility of such libraries is dependent upon identifying the structures of the molecules so prepared. We describe the construction of a peptide combinatorial library, having 117,649 different members, synthesized on beads and indexed with inert chemical tags. These tags are used as a binary code to record the reaction history of each bead. The code can be read directly from a single bead by electron capture capillary gas chromatography. We demonstrate the correct selection of members of the library on the basis of binding to a monoclonal antibody.

With advances in the detailed knowledge of the molecular basis of biology and the development of techniques to isolate virtually any component of a biological system, it has become increasingly feasible to search for biologically active compounds by screening for natural or synthetic ligands to biological molecules of known importance. Ligands discovered in this manner can be useful agents if they mimic or block natural ligands, or if they interfere with the naturally occurring interactions of the biological target. They can also provide a starting point for the engineering of molecules with more desirable properties. Since the chance of finding valuable ligands will increase with the number of compounds screened, the success of the search will be best with massive libraries of compounds. Such libraries can have many sources. Plant and animal extracts, for example, provide a rich source of molecular diversity, though finding and identifying biologically active molecules at parts-per-million-to-billion levels can be problematic.

One of the most promising approaches to the synthesis of large collections of diverse molecules is known as combinatorial chemistry (1, 2), in which vast libraries of molecules having different chemical compositions are synthesized simultaneously. Combinatorial methods entail a series of chemical steps with multiple choices of chemical reagents for each step. The complexity, or number of members in a combinatorial library, is given by the product of the number of reagent choices for each step of the synthesis and can therefore be quite large. The challenge in using combinatorial libraries is the characterization of members of the library with particular desired properties. Several solutions to this problem have been described in the literature. Members of the library can be synthesized in *spatially segregated arrays*, but to date this has resulted only in relatively small libraries (3). Alternatively, in the *multivalent synthesis method*, a library of moderate complexity can be produced by pooling multiple choices of reagents during synthesis (4, 5). Once a given pool is shown to have an interesting property, it is resynthesized iteratively with lower and lower complexity until a single compound having the desired property is identified. The

multivalent method is not practical for construction of massive libraries because the concentration of any individual member of the library decreases with complexity. Moreover, cumbersome resyntheses are required to isolate individual compounds. Another approach, the *split synthesis method*, involves combinatorial synthesis on solid particles such as Merrifield synthesis beads (§, §, and refs. 6–10). Through a protocol of separating and mixing beads during the synthesis, each bead in the final library has a product from a single, specific reaction sequence chemically bound to it and that product is likely to differ from that bound to another bead. After selecting a particular bead having some desirable property, the identity of the attached compound is determined by analytical chemistry. Thus the split synthesis method can be employed only to synthesize compounds that can be readily elucidated by microscale sequencing, such as oligonucleotides and certain oligopeptides.

Yet another solution, the *cosynthesis method*, attempts to solve the structure elucidation problem by cosynthesizing a sequenceable tag that encodes the series of steps and reagents used in the synthesis of each library element. The tag and the corresponding library element are associated by a chemical bond. Once a library element is selected, the procedure used to synthesize it can be read by sequencing the tag. Oligonucleotide and oligopeptide tags have both been proposed (11, 12). The main problem with the cosynthesis method is that the tagging structures are chemically labile and incompatible with many of the reagents normally associated with synthetic organic chemistry. Additional limitations follow from the constraint of compatible protecting groups which allow the alternating cosynthesis of tag and library element. Moreover, the oligonucleotide or peptide tags may themselves associate selectively with biological receptors and confuse the assay.

We have devised an alternative method that is not plagued by these problems and that allows the construction of large chemically diverse libraries. As in the original split synthesis method, we synthesize library elements on microsphere beads (see also ref. 12). During each step of synthesis, however, we attach to the beads tagging molecules that encode both the step number and the chemical reagent used in that step. The array of tags used forms a binary record of the synthetic steps for each bead. Our tagging molecules are not sequentially connected, hence no cosynthesis is required. With only 20 such tags, we can uniquely encode  $2^{20} = 1,048,576$  different syntheses. In the following paragraphs, we describe the use of this method to prepare and study a

Abbreviations: DMF, dimethylformamide; EC, electron capture; mAb, monoclonal antibody.

†Furka, A., Sebestyen, M., Asgedom, M. & Dibo, G., 14th International Congress on Biochemistry, July 10–15, 1988, Prague, Czechoslovakia, Vol. 5, p. 47 (abstr.).

§Furka, A., Sebestyen, M., Asgedom, M. & Dibo, G., 10th International Symposium on Medical Chemistry, August 15–19, 1988, Budapest, Hungary, p. 288 (abstr.).

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.

chemically encoded combinatorial library of 117,649 peptides.

## MATERIALS AND METHODS

**Typical Tag Linker Preparation.** To a solution of 8-bromo-1-octanol (0.91 g, 4.35 mmol) and 2,4,6-trichlorophenol (1.03 g, 5.22 mmol) in dimethylformamide (DMF) (5 ml) was added cesium carbonate (1.70 g, 5.22 mmol), resulting in the evolution of gas and the precipitation of a white solid. The reaction mixture was stirred at 80°C for 2 hr. The mixture was diluted with toluene (50 ml), washed with 0.5 M NaOH (twice with 50 ml), 1 M HCl (twice with 50 ml), and water (50 ml), and the organic phase was dried (MgSO<sub>4</sub>). Removal of the solvent by evaporation gave 1.24 g (87% yield) of tag alcohol as a colorless oil.

The above tag alcohol (0.81 g, 2.5 mmol) was added to a 2 M solution of phosgene in toluene (15 ml) and stirred at room temperature for 1 hr. The excess phosgene and the toluene were removed by evaporation and the resulting crude chloroformate was dissolved in CH<sub>2</sub>Cl<sub>2</sub> (5 ml) and pyridine (0.61 ml, 7.5 mmol). *tert*-Butyl-4-(hydroxymethyl)-3-nitrobenzoate (13) (0.5 g, 1.98 mmol) was added and the reaction mixture was stirred at room temperature for 3 hr. The solution was diluted with ethyl acetate (75 ml) and poured into a separatory funnel. After washing with 1 M HCl (three times with 35 ml), saturated NaHCO<sub>3</sub> (twice with 35 ml), and saturated NaCl (35 ml), the organic phase was dried (MgSO<sub>4</sub>). The solvent was removed by evaporation and the residue was purified by chromatography on silica gel (5–7.5% ethyl acetate in petroleum ether), affording 0.95 g (79% yield) of the tag-linker *tert*-butyl ester as a clear oil.

Trifluoroacetic acid (3 ml) was added to a solution of the tag-linker *tert*-butyl ester (0.95 g, 1.57 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (30 ml) to deprotect the linker acid and the solution was stirred at room temperature for 7 hr. The mixture was then evaporated to dryness and the residue was redissolved in CH<sub>2</sub>Cl<sub>2</sub> (30 ml). The solution was washed with saturated NaCl (20 ml) and the organic phase was dried (MgSO<sub>4</sub>). Removal of the solvent by evaporation gave 0.75 g (87% yield) of the tag-linker acid (6B) as a pale yellow solid (see Fig. 1 and *Generation of a Large Encoded Library* below for nomenclature).

**Typical Encoded Library Synthesis Step.** *N*<sup>α</sup>-Fmoc-Glu(*t*Bu)-Glu(*t*Bu)-Asp(*t*Bu)-Leu-Gly<sub>4</sub>-NH-Merrifield resin was suspended in DMF (20 ml) and shaken for 2 min. After filtering, 1:1 diethylamine/DMF (40 ml) was added to remove the Fmoc (fluoren-9-ylmethoxycarbonyl) protecting groups and the resin was shaken for 1 hr. The resin was separated by filtration and washed with DMF (twice with 20 ml, 2 min each), 2:1 (vol/vol) dioxane/water (twice with 20 ml, 5 min each), DMF (three times with 20 ml, 2 min each), and CH<sub>2</sub>Cl<sub>2</sub> (three times with 20 ml, 2 min each) then dried under reduced pressure at 25°C.

Portions (150 mg) of the resin were placed in seven Merrifield vessels and suspended in CH<sub>2</sub>Cl<sub>2</sub> (5 ml). The appropriate linker-tag acids were activated as their acyl carbonates as follows (for the first coupling): 10A (6.6 mg, 0.0098 mmol) was dissolved in anhydrous ether (2 ml) and pyridine (10 μl) was added. Isobutyl chloroformate (1.3 μl, 0.0096 mmol) was added as a solution in anhydrous ether (0.1 ml). The resulting mixture was stirred at 25°C for 1 hr, during which time a fine white precipitate formed. The stirring was stopped and the precipitate was allowed to settle for 30 min. Solutions of the acylcarbonates of 9A and 8A were prepared in the same way. Aliquots (0.25 ml) of the supernatant solution of activated linker-tags were mixed to give the appropriate 3-bit binary tag codes as described in the text, and the appropriate coding mixtures of tags were added to each of the seven synthesis vessels. The vessels were shaken

in the dark for 12 hr, and then each was washed with CH<sub>2</sub>Cl<sub>2</sub> (four times with 10 ml, 2 min each). A solution of the symmetrical anhydride (ref. 14, pp. 80–83) of an *N*<sup>α</sup>-Fmoc amino acid in CH<sub>2</sub>Cl<sub>2</sub> (3 eq in 10 ml) was then added to the correspondingly coded batch of resin and shaken for 20 min. Five percent *N,N*-diisopropylethylamine in CH<sub>2</sub>Cl<sub>2</sub> (1 ml) was added and the mixture was shaken until the resin gave a negative Kaiser test. The resin batches were filtered, combined, and then washed with CH<sub>2</sub>Cl<sub>2</sub> (four times with 20 ml, 2 min each), isopropyl alcohol (twice with 20 ml, 2 min each), and CH<sub>2</sub>Cl<sub>2</sub> (four times with 20 ml, 2 min each). The next cycle of labeling/coupling was initiated by Fmoc deprotection as described above.

After Fmoc deprotection of the residues in the last position of the peptide, the side chain functionality was deprotected by suspending the resin in CH<sub>2</sub>Cl<sub>2</sub> (10 ml), adding thioanisole (2 ml), ethanedithiol (0.5 ml), and trifluoroacetic acid (10 ml) then shaking for 1 hr at 25°C. The resin was then washed with CH<sub>2</sub>Cl<sub>2</sub> (six times with 20 ml, 2 min each) and dried.

**Electron Capture (EC) Gas Chromatography (GC) Reading of Code.** A single selected synthesis bead was placed in a Pyrex capillary tube and washed with DMF (five times with 10 μl). The bead was then suspended in DMF (1 μL) and the capillary was sealed. The suspended bead was irradiated at 366 nm for 3 hr to release the tag alcohols, and the capillary tube was subsequently placed in a sand bath at 90°C for 2 hr. The tube was opened and bis(trimethylsilyl)acetamide (≈0.1 μl) was added to trimethylsilylate the tag alcohols. After centrifuging for 2 min, the tag solution above the bead (1 μl) was injected directly into an EC detection, capillary gas chromatograph for analysis.

**Antibody-Affinity Methods.** The anti-c-MYC monoclonal antibody (mAb) 9E10 has been described (15, 16). To test beads for binding to 9E10, beads were incubated in TBST (20 mM Tris-HCl, pH 7.5/500 mM NaCl/0.05% Tween-20) containing 1% bovine serum albumin (BSA) to block nonspecific protein binding sites. The beads were then centrifuged, resuspended in a 1:200 dilution of 9E10 ascites fluid in TBST/1% BSA, and incubated overnight at 4°C. Beads were subsequently washed three times in TBST and incubated for 90 min at room temperature in alkaline phosphatase-coupled goat antibodies against mouse IgG (Bio-Rad), diluted 1:3000 in TBST/1% BSA. After the beads had been washed twice in TBST and once in phosphatase buffer (100 mM Tris-HCl, pH 9.5/100 mM NaCl/5 mM MgCl<sub>2</sub>), they were incubated 1 hr at room temperature in phosphatase buffer containing 0.01 part each of AP color reagents A and B (Bio-Rad). To stop the reaction, the beads were washed twice in 20 mM NaEDTA, pH 7.4. Solution phase affinities between 9E10 and various peptides were determined by a modification of the competitive ELISA previously described (17), using a recombinant fusion protein containing at its amino terminus the antigenic peptide EQKLISEEDL, kindly provided by A. Polverino (Cold Spring Harbor Laboratory). The concentration of each peptide necessary to inhibit mAb binding by 50% (IC<sub>50</sub>) was determined in at least three independent assays.

## RESULTS AND DISCUSSION

**A Binary Encoding Scheme Using Chemically Inert Tags.** A simple binary code can be used to describe an organic synthesis; it is best illustrated by example. Imagine carrying out a combinatorial synthesis using any of seven different reagents in each of *N* steps. Such a combinatorial synthesis would yield 7<sup>*N*</sup> different final products. Let us designate the various reagents which can be used in any step as binary 001 (reagent 1), 010 (reagent 2), 011 (reagent 3), . . . 111 (reagent 7). We can now write a binary synthesis code describing any complete *N*-step synthesis using 3 × *N* binary digits. For example, if we used reagent 3 in the first step, the binary

numerical description would be "011." If we next used reagent 1 in the second step, the description would be "001 011." And if we finally used reagent 6 in the third step, we would obtain "110 001 011." This 9-bit binary *synthesis code* describes the synthesis and can be read from right to left in 3-bit blocks to decode the reagents used in each step of the synthesis. More bits per step could be used to encode more reagent designations. To represent such a synthesis code chemically, we use a set of distinguishable, sensitively detectable molecules as tags and the presence of a particular tag to represent a binary 1 for the corresponding bit. Using a set of nine tagging molecules, T9–T1, where T9 represents the leftmost binary bit and T1 represents the rightmost bit, the tag mixture containing only T9, T8, T4, T2, and T1 would represent the 110 001 011 synthesis code.

Various methods can be used to analyze minute quantities of organic tagging molecules. Capillary GC is a convenient analytical technique for separating and identifying tags. When tagging molecules with unusually electrophoric functionality (e.g., fluoro- or chlorocarbons) are used, the conjunction of EC detection (18) with GC separation enables us to analyze tags from a single 50- $\mu\text{m}$  microsynthesis bead without significant interference from contaminants. We employ a photocleavable linkage (see below) between the tag and the bead so that the tag may be selectively liberated from the bead.

**Generation of a Large Encoded Library.** We prepared a set of 18 GC-separable tags with linkers that allowed them to be attached to, and then detached from, the synthesis beads. The linker/tagging molecules we used are summarized in Fig. 1. By using different lengths of the hydrocarbon chain ( $N = 1$ –10) and three different aromatic electrophores (Ar), we were able to prepare more than 20 tagging molecules which were well-separated by capillary GC and selectively detected by EC at levels  $<1$  pmol. We designate these tags as  $N\text{Ar}$ , where  $N$  is the length of the hydrocarbon chain and Ar is the identity of the electrophore as shown in Fig. 1. Thus tag 2B has the structure with  $N = 2$  and  $\text{Ar} = \text{B}$ . The linker segment of these molecules incorporates a carboxylic acid for attachment to the synthesis beads and a photochemically labile *ortho*-nitrobenzylcarbonate for subsequent detachment of the tags.

To relate our tagging molecules to the binary bits of the synthesis code, we arrange them by their GC elution order. Thus the tag called T1 is retained the longest on our GC column and designates the rightmost bit of the binary synthesis code number. The next-longest retained tag we call T2, and so on. Using a 0.2 mm  $\times$  20 m methylsilicone capillary GC column with different combinations of tag chain lengths and electrophoric halobenzenes, we assembled 18 well-resolved tagging molecules (T1–T18) whose chemical compositions were, respectively, 10A, 9A, 8A, 7A, 6A, 5A, 4A, 3A, 6B, 2A, 5B, 1A, 4B, 3B, 2B, 1B, 2C, and 1C.

To test our encoding method in the context of screening for binding to a biological receptor, we synthesized an encoded combinatorial library of 117,649 peptides. In the standard single-letter codes for amino acids, this library had the sequence  $\text{H}_2\text{N-XXXXXXEEDLGGGG-Bead}$ , where the variable residue X was D, E, I, K, L, Q, or S. This library included  $\text{H}_2\text{N-EQKLISEEDLGGGG-Bead}$  among its se-

quences, and EQKLISEEDL is known to be bound by 9E10, a mAb directed against the human c-MYC protein (15, 16). The four glycines served as a molecular spacer to separate the epitope from the bead. Three binary bits were sufficient to represent the seven alternative reagents for each step. We assigned the following 3-bit, binary codes as follows: 001 = S (serine), 010 = I (isoleucine), 011 = K (lysine), 100 = L (leucine), 101 = Q (glutamine), 110 = E (glutamate), and 111 = D (aspartate).

We synthesized our library by first preparing the constant segment of the library ( $\text{H}_2\text{N-EEDLGGGG-Bead}$ ) on 1.05 g of 50- to 80- $\mu\text{m}$  Merrifield polystyrene synthesis beads, using standard solid-phase methods based on *t*Bu side-chain protection and Fmoc main-chain protection (ref. 14, pp. 80–83). After removing the N-terminal Fmoc protecting group, we divided the beads into seven 150-mg portions and processed each as described in *Materials and Methods*, attaching first the tags and then the corresponding amino acids to each portion. The tags were attached via their carboxylic acids to the synthesis beads by activating the linker carboxyl groups as mixed carbonic anhydrides and then adding an amount of activated tag corresponding to  $\approx 1\%$  of the free amino groups on the beads. In the process of this coupling, approximately 0.5% of the growing peptide chains were terminated for each tag added. The remaining free amino groups were then coupled in the usual way with the corresponding protected amino acids as their symmetrical anhydrides. After washing, the seven fractions were combined. After Fmoc deprotection, the beads were again divided into seven portions and processed as before except that in place of tags representing the first step (T1, T2, T3), tags representing the second step (T4, T5, T6) were used. At this point the library had  $7^2$  members. By repeating this procedure four more times using tags T7–T18 analogously, the entire encoded library of  $7^6 = 117,649$  different peptides was prepared.

Given any bead, the attached tags could be detached by UV irradiation and unambiguously decoded by ECGC (e.g., see Fig. 3). To verify that the codes corresponded to the actual peptide sequence present on the beads, two beads were picked at random, the tags present on each were released and read by ECGC, and then the peptide sequence present on each was determined by microsequencing. We observed complete concordance between the synthesis code and the peptide sequence in each case.

**Screening the Library with a mAb.** To pick out those members of our library that bound to the anti-c-MYC mAb, we mixed the bead library with the antibody and stained those beads that bound antibody by using alkaline phosphatase-coupled secondary antibodies (Fig. 2). When viewed under a low-power microscope, dark-staining beads could be easily distinguished from the vast majority of nearly colorless unstained beads and were individually picked by using a manual micropipetter. From two different antibody stainings of 30-mg samples of the peptide library, we picked out 40 dark-colored beads for decoding. We found that the stained beads had reaction histories leading to the presence of either the MYC epitope (EQKLISEEDL) or sequences that differed by one or two substitutions among the three N-terminal residues. In most cases, these sequences were found multiple

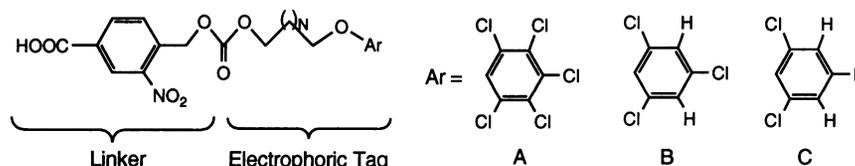


FIG. 1. Molecular tags which create a binary synthesis code.

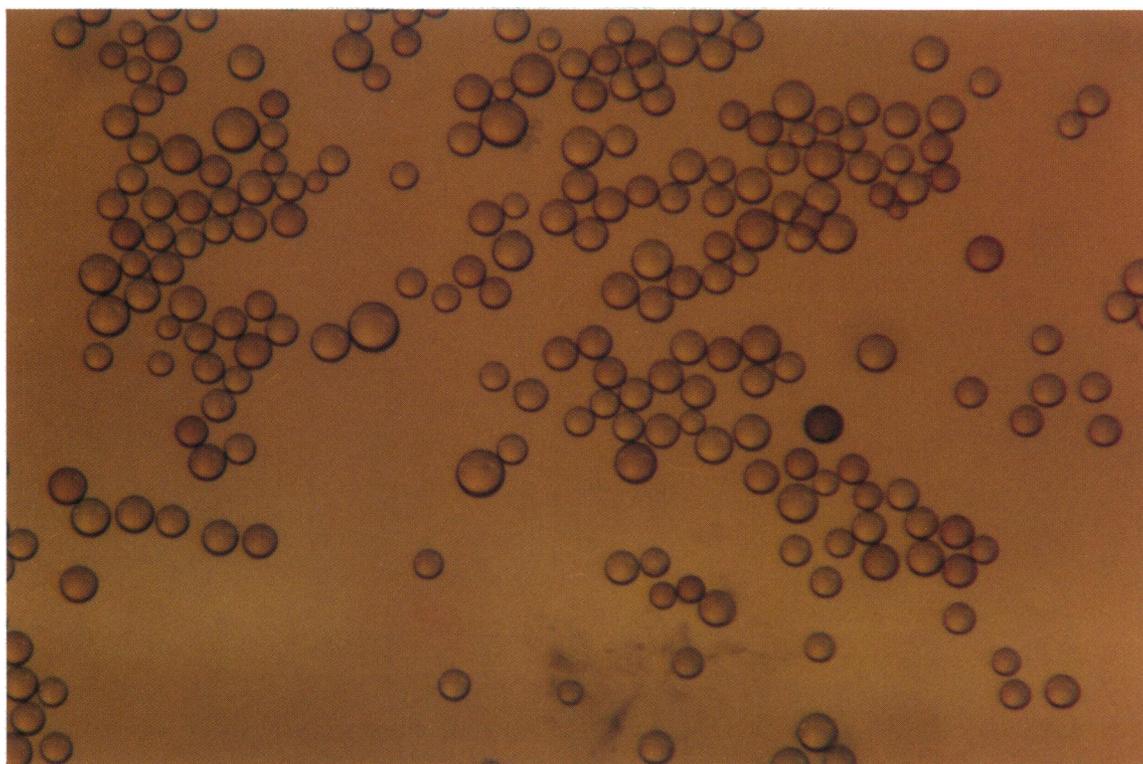


FIG. 2. Peptide library beads stained with mAb 9E10. The beads shown are approximately 50–80  $\mu\text{m}$  in diameter. A stained bead is visible in the middle, 4 cm from the right edge of the photograph.

times. The synthesis codes of the stained beads that were picked are listed in Table 1.

Reading the binary synthesis code of a single bead by using ECGC was a straightforward process, as the reader may see from an actual gas chromatogram of a stained EQKLIS bead taken from the antibody-binding experiment (Fig. 3). Peaks other than those of the tags (T18–T1) come from impurities in the solvent or silylating agent. Such impurity peaks occur at the same retention times ( $\pm 0.05$  min) in every chromatogram, and hence the synthesis codes of single beads can be read unambiguously in nearly all instances.

To test whether the peptides identified in our solid-phase binding experiment could also be bound to mAb 9E10 in free

solution, the affinities of the antibody for free peptides derived from five reactive (dark-staining) bead sequences were measured as  $\text{IC}_{50}$  values by using a competitive ELISA. From entries 1–5 of Table 1, it can be seen that all of the peptides derived from stained bead sequences are bound with high affinity. A control peptide derived from a nonstaining bead sequence (DKISSLEEDL) showed no detectable binding ( $\text{IC}_{50} > 500 \mu\text{M}$ ).

Since all of the peptides we identified on stained beads contained the sequence LIS, we decided to measure the affinity of the mAb 9E10 for MYC epitope-derived peptides that contained substitutions at these positions. As shown in entries 13–15 of Table 1, substitution of isoleucine for leu-

Table 1. Solution and solid-phase binding of peptide library elements to mAb 9E10

| Entry | Synthesis code*         | Sequence†      | Stain‡ | $\text{IC}_{50}$ , § $\mu\text{M}$ |
|-------|-------------------------|----------------|--------|------------------------------------|
| 1     | 110 101 011 100 010 001 | <b>E</b> QKLIS | Yes    | $1.31 \pm 0.05$                    |
| 2     | 100 101 011 100 010 001 | <b>L</b> QKLIS | Yes    | $1.36 \pm 0.51$                    |
| 3     | 101 101 011 100 010 001 | <b>Q</b> QKLIS | Yes    | $1.15 \pm 0.03$                    |
| 4     | 110 101 101 100 010 001 | <b>E</b> QQLIS | Yes    | $23.3 \pm 2.3$                     |
| 5     | 110 110 011 100 010 001 | <b>E</b> EKLIS | Yes    | $4.67 \pm 0.85$                    |
| 6     | 111 110 011 100 010 001 | <b>D</b> EKLIS | Yes    | ND                                 |
| 7     | 100 110 011 100 010 001 | <b>L</b> EKLIS | Yes    | ND                                 |
| 8     | 111 101 011 100 010 001 | <b>D</b> QKLIS | Yes    | ND                                 |
| 9     | 111 110 011 100 010 001 | <b>Q</b> EKLIS | Yes    | ND                                 |
| 10    | 110 111 011 100 010 001 | <b>E</b> DKLIS | Yes    | ND                                 |
| 11    | 100 101 101 100 010 001 | <b>L</b> QQLIS | Yes    | ND                                 |
| 12    | 111 011 010 001 001 100 | <b>D</b> KISSL | No     | >500                               |
| 13    | Bead not found          | <b>E</b> QKIIS | —      | $163.4 \pm 20.6$                   |
| 14    | Bead not found          | <b>E</b> QKLLS | —      | $40.1 \pm 5.0$                     |
| 15    | Bead not found          | <b>E</b> QKLID | —      | $166.2 \pm 38.8$                   |

\*Synthesis code of selected library beads.

†Amino acid residues indicated in boldface are those which differ from those found in the c-MYC protein epitope against which mAb 9E10 was raised (15, 16).

‡Staining with mAb 9E10.

§Peptides with the indicated sequence plus EEDL were synthesized and their  $\text{IC}_{50}$  values for binding were determined, except where indicated by ND.

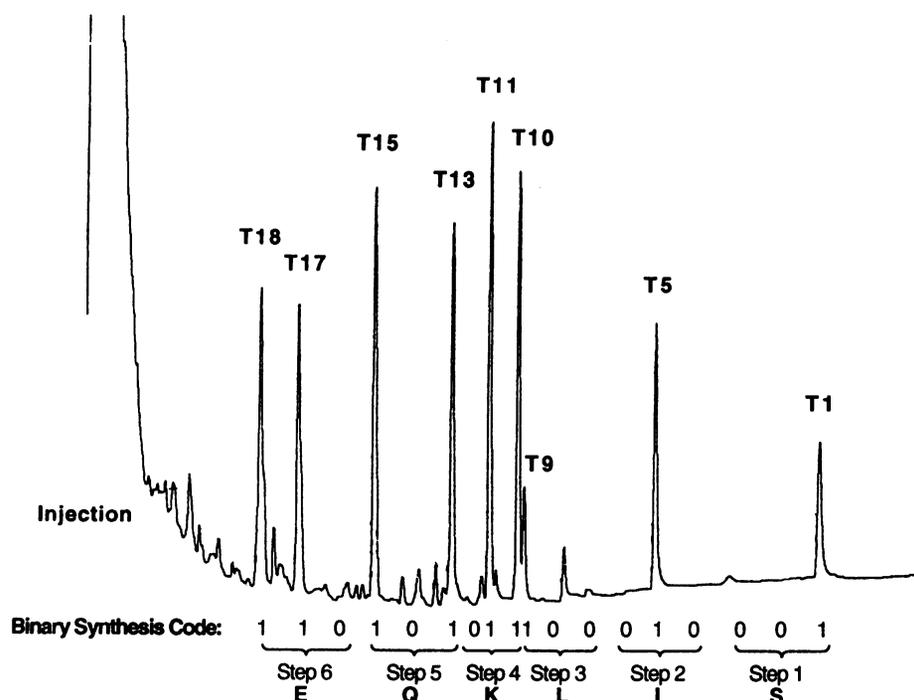


FIG. 3. Gas chromatogram of tags from EQKLISEEDLGGGG-Bead. The synthesis code from one stained synthesis bead was read by releasing and analyzing tags as described in the text.

cine, leucine for isoleucine, or aspartic acid for serine led to more weakly binding peptides having  $IC_{50}$  values of 40–166  $\mu$ M. We were surprised to find that the conservative substitutions of leucine and isoleucine produced such large changes in  $IC_{50}$ . Since none of these substitutions were found in a stained bead sequence (though such sequences should have been present in the library), and since the stained bead sequence EQQKLISEEDL ( $IC_{50} = 23 \mu$ M) was found, we have an indication of the minimum affinity required for detection of antibody binding to bead-supported peptides under the conditions described here.

The results above establish that a chemically inert, multiple tag labeling scheme can be useful for the practical generation of large encoded combinatorial libraries. While we demonstrate the method here by generating an encoded library of peptides, the method can be applied to other library types. The most exciting applications will likely involve the construction of similarly encoded small molecule organic libraries whose chemical elements cannot be sequenced the way peptides can. Studies of the detection limits of the molecular tags we have employed suggest that we can create encoded libraries having as many as  $10^9$  different members on approximately  $1 \text{ cm}^3$  of microsphere synthesis beads. Such complexity should be attainable by using 10- to 20- $\mu$ m beads, which can carry tags at the readily detectable level of 0.1 pmol. With the standard 50- to 80- $\mu$ m beads used in our current work, libraries having more than  $10^6$  members per  $\text{cm}^3$  are readily available. While the assays we used here to select beads were conducted with library members chemically bound to the synthesis beads, one can readily imagine schemes in which cleavable linkers between the beads and the library members allow off-bead solution-phase assays. Access to such encoded combinatorial libraries should provide a substantial benefit to those searching for new organic compounds having desirable properties.

The authors thank Richard Axel, Paul Bartlett, and David Beach for helpful discussions and Gilbert Stork for careful readings of the manuscript. This work was supported by National Science Foundation Grant CHE92 08254 (to W.C.S.) and grants from the National

Cancer Institute (to M.W.). M.W. is an American Cancer Society Professor.

- Jung, G. & Beck-Sickinger, A. G. (1992) *Angew. Chem. Int. Ed. Engl.* **31**, 367–383.
- Pavia, M. R., Sawyer, T. K. & Moos, W. H. (1993) *Bioorg. Med. Chem. Lett.* **3**, 387–396.
- Fodor, S. P., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T. & Solas, D. (1991) *Science* **251**, 767–773.
- Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T. & Cuervo, J. H. (1991) *Nature (London)* **354**, 84–86.
- Houghten, R. A. & Dooley, C. T. (1993) *Bioorg. Med. Chem. Lett.* **3**, 405–412.
- Furka, A., Sebestyen, M., Asgedom, M. & Dibo, G. (1991) *Int. J. Pept. Protein Res.* **37**, 487–493.
- Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. & Knapp, R. J. (1991) *Nature (London)* **354**, 82–84.
- Zuckermann, R. N., Kerr, J. M., Siani, M. A. & Banville, S. C. (1992) *Int. J. Pept. Protein Res.* **40**, 498–507.
- Sebestyen, F., Dibo, G., Kovacs, A. & Furka, A. (1993) *Bioorg. Med. Chem. Lett.* **3**, 413–418.
- Lam, K. S., Hruby, V. J., Lebl, M., Knapp, R. J., Kazmierski, W. M., Hersh, E. M. & Salmon, S. E. (1993) *Bioorg. Med. Chem. Lett.* **3**, 419–424.
- Brenner, S. & Lerner, R. A. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 5381–5383.
- Kerr, J. M., Banville, S. C. & Zuckermann, R. N. (1993) *J. Am. Chem. Soc.* **115**, 2529–2531.
- Barany, G. & Albericio, F. (1985) *J. Am. Chem. Soc.* **107**, 4936–4942.
- Stewart, J. M. & Young, J. D. (1984) *Solid Phase Peptide Synthesis* (Pierce Chem. Co., Rockford, IL), 2nd Ed., pp. 80–83.
- Evan, E. I., Lewis, G. K., Ramsay, G. & Bishop, M. (1985) *Mol. Cell Biol.* **5**, 3610–3616.
- Munro, S. & Pelham, H. R. B. (1987) *Cell* **48**, 899–907.
- Harlow, E. & Lane, D. (1988) *Antibodies: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 570–573.
- Grimrud, E. P. (1992) in *Detectors for Capillary Chromatography*, eds. Hill, H. H. & McMinn, D. G. (Wiley, New York), pp. 83–107.