

Enzyme-mediated spatial segregation on individual polymeric support beads: Application to generation and screening of encoded combinatorial libraries

(combinatorial chemistry/encoded libraries/peptides/proteolytic enzymes/solid-phase synthesis)

JOSEF VÁGNER*[†], GEORGE BARANY*[‡], KIT S. LAM[§], VIKTOR KRCHŇÁK[¶], NIKOLAI F. SEPETOV[¶], JAMES A. OSTREM[¶], PETER STROP[¶], AND MICHAL LEBL[¶]

*Department of Chemistry, University of Minnesota, 207 Pleasant Street S.E., Minneapolis, MN 55455; [§]Arizona Cancer Center and Department of Medicine, University of Arizona, College of Medicine, Tucson, AZ 85724; and [¶]Selectide Corporation, an Affiliate of Hoechst Marion Roussel, 1580 East Hanley Boulevard, Tucson, AZ 85737

Communicated by Bruce Merrifield, The Rockefeller University, New York, NY, April 3, 1996 (received for review December 18, 1995)

ABSTRACT Proteolysis of short N^α -protected peptide substrates bound to polyoxyethylene-polystyrene beads releases selectively free amino sites in the enzyme-accessible “surface” area. The substantial majority of functional sites in the “interior” of the polymeric support are not reached by the enzyme and remain uncleaved (protected). Subsequent synthesis with two classes of orthogonal protecting groups— N^α -*tert*-butyloxycarbonyl (Boc) and N^α -9-fluorenylmethoxycarbonyl (Fmoc)—allows generation of two structures on the same bead. The surface structure is available for receptor interactions, whereas the corresponding interior structure is used for coding. Coding structures are usually readily sequenceable peptides. This “shaving” methodology was illustrated by the preparation of a peptide-encoded model peptide combinatorial library containing 1.0×10^5 members at ≈ 6 -fold degeneracy. From this single library, good ligands were selected for three different receptors: anti- β -endorphin antibody, streptavidin, and thrombin, and the binding structures were deduced correctly by sequencing the coding peptides present on the same beads.

Evaluation of libraries of compounds generated by combinatorial chemistries has been appreciated recently as a promising approach to identify lead structures needed for accelerated drug discovery programs (1–5). The principles of the library approach were established originally on short peptide sequences assembled by repetitive cycles of solid-phase synthesis chemistry or displayed within the coat protein transcripts of appropriate partially randomized bacteriophage genes. Key to the success of these pioneer studies was the capability to readily microsequence linear unblocked peptides or oligonucleotides. The focus of recent library efforts has shifted away from peptides and toward peptidomimetic and/or nonpeptide structures, which may be more stable and potentially provide a more diverse sampling of conformational space (6–8). This expansion in target range is accompanied by new challenges for reliable assembly strategies and efficient determination of covalent structures of active principles recognized and selected by the biological screening.

Three general types of approaches have been described for determination of structures built by solid-phase synthesis: (i) direct spectroscopic analyses on microamounts of material bound to and/or released from polymeric supports (for examples see refs. 9–11); (ii) spatially addressable syntheses in which the structure of a compound is deduced from its position on an array (2, 12–14); and (iii) split synthesis procedures, whereby compounds are built up on solid-phase beads, each of

which has a unique history throughout the randomization steps and hence a unique structure—e.g., “one-bead–one-peptide” (4, 6). The appropriate synthetic strategies can be modified in a number of ways so that with each component introduced by a combinatorial step, a conjugate “tag” is added in a parallel step; these coding sequences or tags are read subsequently to decode the steps used for the construction of the structure on any given bead. Coding can be achieved by use of peptides (1, 15–17), (PCR-amplifiable) DNA (18, 19), or small electrophoric tags applied in a binary format (20).

In all coding approaches reported so far, both the coding and screening sequences are present contiguously throughout the beads, usually (but not always) at essentially equimolar levels. The present paper describes and documents a new set of concepts that permit the generation of encoded combinatorial libraries (Fig. 1). This approach, termed enzyme-mediated spatial segregation, or, more colloquially, “shaving,” relies on the physical separation of the coding and screening structures on individual polymer beads (1).**

PRINCIPLES AND DESIGN

The goal is to identify and exploit different physical characteristics of “surface” and “interior” areas of a given bead, in such a way that only the screening structure will be present on the surface (and hence available for possible biological interactions), whereas the coding structure (generally, a readily sequenceable linear peptide) is restricted exclusively to the interior. Experimental results presented herein (see Table 3)

Abbreviations: POE-PS, polyoxyethylene-polystyrene; Fmoc, N^α -9-fluorenylmethoxycarbonyl; Boc, N^α -*tert*-butyloxycarbonyl; DIPCIDI/HOBT, N,N' -diisopropylcarbodiimide/1-hydroxybenzotriazole; DMF, N,N -dimethylformamide; Ab, antibody; BCIP, 5-bromo-4-chloro-3-indoyl phosphate. Unless otherwise stated, amino acid symbols denote the L-configuration, and all solvent ratios and percentages are vol/vol.

[†]Present address: Novo Nordisk A/S, Novo Nordisk Park, DK-2760 Måløv, Denmark.

[‡]To whom reprint requests should be addressed. e-mail: barany@maroon.tc.umn.edu.

[¶]Present address: Houghten Pharmaceuticals, 3550 General Atomics Court, San Diego, CA 92121.

**Portions of this work were reported in preliminary form at the Third International Symposium on Innovation and Perspectives in Solid Phase Synthesis, University of Oxford, U.K., Aug. 31–Sept. 4, 1993 (see ref. 1) and the 35th Annual Medicinal Chemistry Symposium, University of Buffalo, NY, May 22–25, 1994.

^{††}TentaGel is a relatively hydrophilic commercial support developed by W. Rapp, in which ethylene oxide is polymerized onto initiator sites on preformed uniform polystyrene beads (see ref. 21). The final material is suitable both for solid-phase synthesis and for biological assays in aqueous media; the present studies used TentaGel with a diameter of 130 μ m, an estimated average molecular weight of pendant polyoxyethylene chains of 3000, and amino-functionalized to a loading of 0.21 mmol/g.

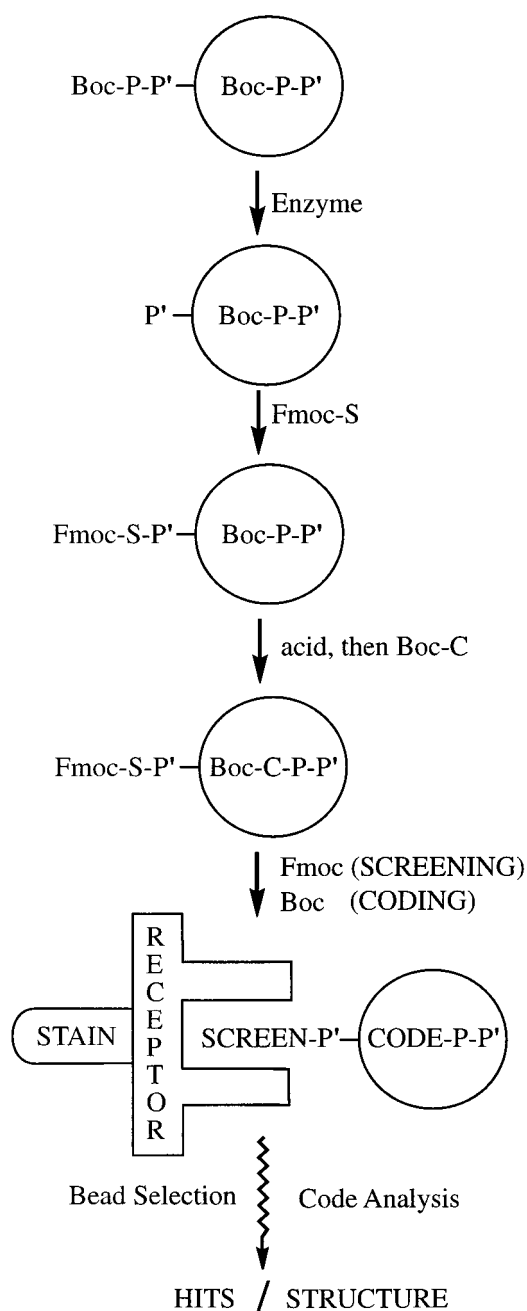


FIG. 1. Strategy and steps for enzyme-mediated spatial segregation. Large circle is POE-PS (TentaGel) polymeric support bead with interior and surface areas; P-P' is a substrate with a scissile bond between P and P'. S is the terminal residue of the screening structure, introduced by Fmoc chemistry, and C is the terminal residue of the coding structure, introduced by Boc chemistry. Subsequent cycles carried out in series or "zig-zag" fashion provide surface screening and interior coding structures. To avoid "false positive" results, it is imperative that coding structures be absent from the surface. In contrast, some screening structures can be tolerated within the interior areas, so long as sequencing carried out on beads selected for positive biological results gives unambiguous readouts of the coding structures.

show that, at most, 15% (and often substantially less) of the total functional sites in typical microporous polyoxyethylene-polystyrene (POE-PS≡TentaGel) beads^{††} used for solid-phase library work are receptor-accessible; this is consistent with the interpretation of Bayer and Rapp (21) that protein immobilization onto TentaGel is limited to the surface. Consequently, this plan dictates that the overwhelming majority of structures assembled on each bead are the interior coding

Table 1. Model peptides used in these studies

Peptide ligand	Receptor	Molecular mass	Refs.
Leu-His-Pro-Gln-Phe	Streptavidin (tetramer)	60 kDa	4, 23
Tyr-Gly-Gly-Phe-Leu	Anti- β -endorphin Ab	170 kDa	4
D-Phe-Pro-Arg-Pro-Gly	Thrombin	35 kDa	24

See *Materials and Methods* for details on assay conditions, staining procedures, and controls to determine when binding to the selected beads was specific.

sequences. A significant premise of the present work is that the required bead differentiation can be achieved by enzymatic reactions—e.g., proteolysis. Enzymes are macromolecular reagents (molecular mass, 20–50 kDa) that are not expected to penetrate beads to any appreciable extent during the time period that they act on their substrates; the susceptibility of a substrate to transformation by an enzyme parallels the later interaction of a screening structure with its biological receptor at the same physical sites on the beaded support.

The enzyme-mediated segregation procedure is carried out once, at the start of the library process (Fig. 1). Once two populations of differentially protected sites are established this way, each of them can be elaborated further by well-established orthogonal chemistries (22)—e.g., alternating *N*^α-9-fluorenylmethyloxycarbonyl (Fmoc; base-labile) and *tert*-butyloxycarbonyl (Boc; acid-labile) procedures for peptide chain assemblies (as shown in Fig. 1). Biological screening reveals active species on the surface areas, and sequencing of the beads thus identified gives unambiguously the corresponding interior coding structures.

As a prerequisite to the eventual application of the shaving concept to encoded combinatorial libraries, experiments were carried out in which beads were charged with two biologically active model peptides: a surface peptide that should interact specifically with its receptor, and an interior peptide that should not interact with its receptor despite being the predominant species on the bead (evidenced by analytical and sequencing data). For these studies, three model peptides were chosen; each is small, straightforward to construct, and involved in a specific, high-affinity binding interaction with a cognate macromolecular receptor (Table 1).

MATERIALS AND METHODS

Solid-Phase Peptide Synthesis. General materials, solvents, reagents, suppliers, and procedures have been described in previous publications from our laboratories (25). TentaGel-AM (130 μ m, 0.21 mmol/g; Rapp Polymere, Tübingen, Germany), extended by a β Ala-Gly- β Ala-Gly spacer and lacking a cleavable linker, was the starting support for shaving, synthesis, and library screening/selection/analysis work. Peptide chain assemblies, in continuous-flow syringe reactors (26), were by standard Boc and/or Fmoc chemistry, with *N,N'*-diisopropylcarbodiimide/1-hydroxybenzotriazole (DIPCDI/HOBt)-mediated couplings in *N,N*-dimethylformamide (DMF), and bromophenol blue monitoring (27). Boc removal was with CF₃COOH-CH₂Cl₂ (3:7) (2 plus 30 min) followed by *N*-methylmorpholine-CH₂Cl₂ (1:19) for neutralization; Fmoc removal was with piperidine-DMF (1:4) (2 plus 10 min). The protection schemes and final deprotection steps varied somewhat, depending on the synthetic objective, since for much of this work, two distinct peptides were assembled by orthogonal chemistries. Orthogonal Fmoc followed by Boc synthesis was carried out with 4-methoxy-2,4,6-trimethylbenzenesulfonyl, triphenylmethyl, and *tert*-butyl for Fmoc-Arg, His, and Tyr side-chain protection, respectively, followed by *p*-toluenesulfonyl, 2,4-dinitrophenyl, and H for Boc-Arg, His, and Tyr, respectively. At the end of the synthesis, sequential deprotection with thiophenol-DMF (1:9) for 1 h at 25°C; reagent K,

CF₃COOH–phenol–thioanisole–water–1,2–ethanedithiol (82.5:5:5:5:2.5), for 1 h at 25°C (if no 4-methoxy-2,4,6-trimethylbenzenesulfonyl was present) or 2 h at 50°C (if 4-methoxy-2,4,6-trimethylbenzenesulfonyl was present); piperidine–DMF (1:1) for 20 min at 25°C; and 5 mM aqueous HCl (to protonate free amino group) gave beads suitable for biological evaluation. Alternatively, orthogonal Boc followed by Fmoc was with toluenesulfonyl, Boc, and H for Boc-Arg, His, and Tyr, respectively, followed by 2,2,5,7,8-pentamethylchroman-6-sulfonyl, triphenylmethyl, and *tert*-butyl for Fmoc-Arg, His, and Tyr side-chain protection, respectively; final deprotection was achieved with reagent K and piperidine–DMF, either order, followed by HF–anisole (9:1) for 1 h at 0°C. Zig-zag syntheses, in which Fmoc and Boc chemistries were carried out alternatively, used the former of these strategies. Active peptides identified by biological screening (see below) and decoded according to the methods of this paper were resynthesized by standard Fmoc/*tert*-butyl chemistry on Rink-resin on TentaGel S.

Enzyme-Mediated Spatial Segregation (Shaving). Enzymes, all obtained from Sigma, were bovine pancreatic chymotrypsin type II, EC 3.4.21.1, molecular mass, 22 kDa; porcine pancreatic elastase type I, EC 3.4.21.36, molecular mass, 22 kDa; and porcine stomach mucosa pepsin A, EC 3.4.23.1, molecular mass, 35 kDa. The following buffers were used: 0.1 M ammonium carbonate (pH 7.8) or 0.1 M Tris/0.1 M CaCl₂ (pH 7.85) for chymotrypsin; 0.1 M Tris (pH 8.05) for elastase; and 0.1 M ammonium formate (pH 3) for pepsin. Synthesized N^α-protected substrate-resins were placed into a syringe reactor, washed first with DMF, MeOH, water, and buffer, and then combined with a solution of enzyme in buffer (concentration 0.1 to 1.0 mg enzyme per ml of buffer; 10 ml of enzyme solution used per g of substrate-resin). Proteolysis was allowed to proceed for 8 h at 37°C and was generally repeated twice more for 8 h each. Between incubations, substrate-resins were washed with buffer, water, *tert*-butanol–H₂O (1:1), water and buffer. Upon completion of the procedure, shaved substrate-resins were washed with buffer, water, MeOH, water, and 5 mM aqueous HCl. Direct biological testing followed, or, alternatively, the shaved resins were neutralized with *N*-methylmorpholine–DMF (1:19), washed with DMF, and acylated with the first protected amino acid designated for the surface.

Peptide-Encoded Peptide Library Synthesis. *Experiment A.* Starting with 0.6 g of Boc-Trp-Gly-βAla-Gly-βAla-Gly-POE-PS, shaving was carried out with chymotrypsin in 0.1 M (NH₄)₂CO₃ (pH 7.8) buffer, as described above. The appropriate Fmoc-amino acids (0.3 mmol) in DMF (1.5 ml) were activated externally by DIPCDI/HOBt and added immediately to each of nine syringe reactors containing equal portions of the original shaved support. Bromophenol blue tests confirmed that acylations were complete after 90-min couplings. Each reactor was washed separately with DMF and CH₂Cl₂, followed by Boc removal, neutralization, and washing with DMF. Next, the corresponding coding Boc-amino acid (see footnote to Table 5) was introduced by the DIPCDI/HOBt method. Resin beads were washed, pooled, Fmoc-deprotected (monitored by UV on the entire pool), and reappportioned into the nine reactors for the next library cycle. Upon completion of all cycles, final removal of N^α-Boc and all side-chain protecting groups was achieved by reagent K (30 min plus 16 h). The library was then washed with CH₂Cl₂ and DMF, neutralized with *N*-methylmorpholine–DMF (1:19), washed with DMF, MeOH, H₂O, and equilibrated in 5 mM aqueous HCl.

Experiment B. A similar experiment was started with 7.0 g of Boc-Ala-Gly-Val-Phe-Gly-βAla-Gly-βAla-Gly-POE-PS, shaved with chymotrypsin in 0.1 M Tris/0.1 M CaCl₂, pH 7.85 buffer.

Screening of Resin-Bound Peptides. For the procedures that follow, suppliers were: Boehringer Mannheim for anti-β-endorphin antibody (Ab) (clone 3E7); Pierce for streptavidin–alkaline phosphatase, 5-bromo-4-chloro-3-indoyl phosphate (BCIP), nitroblue tetrazolium chloride, and the *N*-hydroxy-succinimide ester of D-biotin (biotin-OSu); Bio-Rad for goat-anti-mouse alkaline phosphatase; and Enzyme Research Laboratories (South Bend, IN) for human thrombin, which was subsequently biotinylated with biotin-OSu. Assays could be carried out separately or in series on the same set of peptide-resin beads in a library; between biological testing, beads were recycled by washing with 8 M guanidinium hydrochloride (pH 1.0) three times for a total of 15 min, then with double-distilled water, and decolorized with DMF.

Streptavidin binding. Following the literature outline (23) with minor modifications, peptide-resin beads were first washed thoroughly with double-distilled water and then washed and coated with 0.05% gelatin to block nonspecific binding. There followed washings with 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 1.4 mM KH₂PO₄ (pH 7.2)/0.1% Tween-20 (PBS-T); 275 mM NaCl, 5.4 mM KCl, 8.6 mM Na₂HPO₄, 2.8 mM KH₂PO₄, pH 7.2/0.1% Tween-20/0.05% gelatin (2×PBS-GT); incubation for 1 h with 20 nM streptavidin–alkaline phosphatase in 2×PBS-GT; and again washings with PBS-T, 2×PBS-GT, and 137 mM NaCl/2.7 mM KCl/25 mM Tris base, pH 7.4 (TBS) buffers. The standard substrate BCIP was added [165 μg/ml in 0.2 M Tris·HCl, pH 8.4/0.2 M NaCl/2.4 mM MgCl₂ (BCIP) buffer], and beads were transferred to Petri dishes for color development (2 h). Those beads that showed a turquoise color were collected, recycled (8 M guanidinium hydrochloride, pH 1.0; H₂O; DMF), and recoated with gelatin. Next, 20 nM streptavidin–alkaline phosphatase in PBS-GT was added along with 0.1 M D-biotin. Under these circumstances (competition), about a third of the beads remained colorless after incubation with substrate. These colorless beads were selected, reincubated with 20 nM streptavidin–alkaline phosphatase followed by substrates, and shown to be colored again.

Anti-β-endorphin Ab binding. Following the literature outline (4), peptide-resin beads (directly after synthesis or recycled) were washed and blocked by gelatin coating, as described above. Incubation was carried out for 1 h with 5 ng of biotinylated anti-β-endorphin Ab per ml in 2×PBS-GT. Thorough washings, incubation with 20 nM streptavidin–alkaline phosphatase, and then addition of substrates for color development and selection were carried out as described above. Those beads that developed color were recycled and tested for specificity by incubation with 5 ng of anti-β-endorphin Ab per ml in PBS-GT buffer, followed by thorough washing and incubation with 100 ng goat-anti-mouse alkaline phosphatase per ml for 1 h. Beads were then washed with PBS-T, 2×PBS-GT, and TBS, and substrates were added. Those beads that developed color were concluded to contain specific-binding structures.

Thrombin binding. Peptide-resin beads were washed three times with 50 mM Hepes, pH 7.0/0.8 M NaCl/0.05% PEG-8000/0.02% NaN₃/0.25% Tween-20/0.1% BSA (HSBB-T) buffer and blocked by incubation with LSBB-T (HSBB-T diluted 1:3 with deionized H₂O)/0.05% gelatin for 15 min at 25°C. The buffer was removed, and beads were incubated with 10 volumes of 50 nM biotinylated-human thrombin and 5 nM streptavidin–alkaline phosphatase in LSBB-T for 1 h at 25°C. Beads were then washed three times each with LSBB-T and BCIP buffers, and finally incubated for 30–60 min with BCIP (165 μg/ml) in BCIP buffer. Staining was terminated by 10% aqueous acetic acid; after extensive washing with LSBB-T, stained beads were selected after microscopic inspection. To distinguish active site inhibitors from structures that bind thrombin at other sites, those beads stained with thrombin were stripped with 8 M guanidinium hydrochloride, decolorized with DMF, carefully washed and equilibrated with binding buffer, and reincubated with thrombin for 30 min in

the presence of 5 molar equivalents of the inhibitor D-Phe-Pro-Arg-chloromethyl ketone (obtained from Chemica Alta). Beads that did not stain after this step were reincubated with biotinylated human thrombin and streptavidin-alkaline phosphatase. The standard staining procedure was carried out, and beads that showed color were selected for sequencing.

RESULTS AND DISCUSSION

Selective Surface Modification by Enzymatic Shaving. For initial studies, several POE-PS resins loaded with a single peptide were incubated with chymotrypsin (Table 2) and then tested with the appropriate macromolecular receptor. Gratifyingly, it was possible to abolish all binding activity of beads bearing Tyr-Gly-Gly-Phe-Leu to the anti- β -endorphin Ab (Table 2, line 1), even though quantitative Edman degradation revealed that the loading of peptide on the beads was unaltered by the proteolysis step. On the other hand, chymotrypsin failed to completely remove Leu-His-Pro-Gln-Phe from the surface (Table 2, lines 2 and 3), consistent with this particular streptavidin-binding sequence being a poor substrate for the enzyme. A simple Tyr-Gly or Trp-Gly linker at the C terminus of the model pentapeptide created a much better chymotrypsin substrate, and the shaving procedure eliminated the interaction with streptavidin (Table 2, lines 4 and 5).

Subsequent studies reversed the experimental design to explore a range of proteolytic enzymes and address substrate optimization (Table 3). Substrates corresponding to the known cleavage specificities of the enzymes were established throughout the beads, and shaving followed to expose surface amino groups that were ultimately blocked by acetylation. Next, the Tyr-Gly-Gly-Phe-Leu ligand was assembled on the nominal interior areas; experiments were judged as successful when this interior peptide was shielded from the highly sensitive and specific binding interaction with its macromolecular receptor, the anti- β -endorphin Ab (see Table 3, lines 5, 6, 8, 9, and 12). Chymotryptic shaving saturates at ≈ 2 –2.5% of the total sites in the beads, whereas elastase or pepsin access 10–15% of the sites, and indeed shaving with the latter two enzymes at the 2.5–5% level fails to cleave all of the sites that may later interact with the Ab. *In toto*, these data suggest the dynamic nature of the shaving process.^{‡‡}

Further significant conclusions from these studies are that the absolute extents of shaving do not change with increased proteolysis times but are very much a function of the sequence of the substrate for the enzyme (e.g., Table 3, line 1 versus line 6; lines 7–9) and even an end group that is several residues removed from the cleavage site (e.g., Table 3, lines 10–12).

^{‡‡}A simplified model assumes that the 130- μ m beads are perfect spheres with a uniform distributions of sites. It follows that 1%, 2.5%, and 10% of the sites correspond respectively to surface shells of width 0.22, 0.54, and 2.2 μ m.

Published kinetic parameters on model substrates (Table 3, footnote *) reveal that key to achieving complete differentiation between surface and interior portions is a high k_{cat}/K_M ratio; a modest 3- to 5-fold improvement in catalytic efficiency is sufficient, whereas order of magnitude differences in K_M alone do not appear to be correlated.

Finally, the efficacy of shaving is a function of the sensitivity of the appropriate binding interaction with the macromolecular receptor (Table 1). To probe this parameter, calibration experiments were set up in which the amounts of pentapeptides synthesized on surface areas of the beads were limited. The protocol involved initial chymotryptic shaving, followed by coupling of defined mixtures of Fmoc-Gly and Boc-Gly (i.e., 1:0, 1:1, 1:9, and 1:99), and then Boc deprotection, acetylation, and synthesis by Fmoc chemistry. Even when Tyr-Gly-Gly-Phe-Leu was present on only 1% of the surface sites (corresponding to an estimated 2 μ mol/g, or $\leq 0.02\%$ of the total sites in the bead), it could be detected by the Ab interaction. The levels of peptide on any of these beads were too low to be detected by amino acid analysis or Edman degradation sequencing of single beads. Extrapolating this sensitivity level to the earlier studies, $>95\%$ enzymatic shaving of the surface would still leave ample surface sites bearing the putative interior peptide as judged by the anti- β -endorphin assay. In contrast, the Leu-His-Pro-Gln-Phe-streptavidin interaction was not observed at the 20 μ mol/g level (10% of the chymotrypsin-accessible surface sites), and the D-Phe-Pro-Arg-Pro-Gly-thrombin interaction was not observed at 60 μ mol/g (30% of surface). Results were the same regardless of whether shaved or bulk unsegregated POE-PS beads were used.

Spatial Segregation: One-Bead-Two-Peptide Models. The aforementioned results encouraged us to try chymotrypsin-mediated spatial segregation as the key step to establish surface/interior differentiation according to concepts explained earlier. POE-PS beads with a β -Ala-Gly- β -Ala-Gly spacer were extended with Boc-Gly followed by Boc-Trp, and then incubated with the enzyme. The resultant shaved resins were either “capped” by acetylation, or acylated by Fmoc-Leu, Phe, or Gly (the respective C-terminal residues of the three model peptides listed in Table 1). Subsequently, the necessary cycles of Fmoc chemistry were carried out to assemble the model peptides, which were presumed to be confined to the surface areas of the beads. After completion of the Fmoc syntheses, the N^α -protecting group of the majority of amino acyl-spacer chains on the resin was cleaved, and Boc chemistry was used to build a different model sequence, presumably in the interior areas. Final deprotection gave beads containing two peptides, suitable for analytical work and biological testing (Table 4).

The data (Table 4) show clearly that as a consequence of the shaving procedure: (i) the interior peptide structure was read out by sequential Edman degradation, and residues corresponding to the surface peptide were found only in trace

Table 2. Staining reaction of various peptide-beads, before and after shaving with chymotrypsin

Line no.	Peptide on bead	Binding to			
		Anti- β -endorphin		Streptavidin	
		Control	Shaved	Control	Shaved
1	Tyr \downarrow Gly-Gly-Phe \downarrow Leu	5+	0	0	ND
2	Leu-His-Pro-Gln-Phe \downarrow	ND	ND	3+	2+
3	Leu-His-Pro-Gln-Phe \downarrow Gly	ND	ND	ND	0 to 1+
4	Leu-His-Pro-Gln-Phe \downarrow Tyr \downarrow Gly	ND	ND	3+	0
5	Leu-His-Pro-Gln-Phe \downarrow Trp \downarrow Gly	ND	ND	3+	0

Sequence shown was assembled on β Ala-Gly- β Ala-Gly-POE-PS beads by Fmoc chemistry. Chymotryptic treatment at pH 7.8, 37°C three times for 8 h each was followed by biological testing (Table 1). ND, not determined. The ranking from 0 to 5+ designates the relative intensity of staining, from none to strongest. The likely scissile bond is indicated by the arrow (\downarrow). Subsequent to shaving, it was determined by quantitative Edman degradation sequence analysis that the average level of peptide on individual beads, 100–130 pmol, remained unaltered.

Table 3. Optimization of substrates for shaving with various enzymes

Line no.	Substrate	Enzyme	K_M ,* mM	k_{cat}/K_M ,* $s^{-1} M^{-1}$	Amount cleaved, %†	Binding to anti- β -endorphin
1	Boc-Phe \downarrow Gly	Chymotrypsin	0.21	905	<0.5	ND
2	Boc-Tyr \downarrow Gly	Chymotrypsin	1.13	300	~1	4+
3	Boc-Trp \downarrow Gly	Chymotrypsin	—	—	~1	3+
4	Boc-Gly-Gly-Phe \downarrow Leu-Gly	Chymotrypsin	0.67	3.1×10^4	1.8–2.0	1+
5	Boc-Ala-Ala-Pro-Phe \downarrow Gly	Chymotrypsin	0.06	1.0×10^6	2.0–2.5	0
6	Boc-Ala-Gly-Val-Phe \downarrow Gly	Chymotrypsin	1.13	1.23×10^5	2.1–2.3	0
7	Boc-Ala-Ala-Ala \downarrow Gly	Elastase	0.1	7.4×10^5	2.5	3+
8	Boc-Ala-Ala-Leu \downarrow Gly	Elastase	0.18	4.4×10^6	10.3	0
9	Boc-Ala-Ala-Nva \downarrow Gly	Elastase	0.08	6.8×10^6	13.4	0
10	Fmoc-Ala-Ala-Phe-Phe \downarrow Gly	Pepsin	—	—	2.4	3+
11	Z-Ala-Ala-Phe-Phe \downarrow Gly	Pepsin	0.04	7.05×10^6	5.3	2+
12	Boc-Ala-Ala-Phe-Phe \downarrow Gly	Pepsin	—	—	15.8	0

Sequence shown was assembled on β Ala-Gly- β Ala-Gly-POE-PS beads by Fmoc chemistry, except for the indicated *N*-terminal protected residue. Enzymatic treatment under conditions given in *Materials and Methods* was followed by acylation of the exposed surface amino groups with Fmoc-glycine, deblocking (see note †, below), and acetylation by 0.3 M *N*-acetylimidazole in DMF for 20 min. Next, the Boc group was removed, and assembly of Tyr-Gly-Gly-Phe-Leu was carried out by Fmoc chemistry on the nominal interior areas of the beads. This was confirmed by quantitative Edman degradation sequence analysis and amino acid analysis that the average level of the interior peptide was 100–130 pmol. Other conventions same as in Table 2.

*Kinetic parameters are for model compounds with the same sequence taken from the literature. Lines 1 and 2, *N*-acetyl-X-*p*-nitrophenylanilide for X = Phe (28) and X = Tyr (29); line 4, partial enkephalin sequence; compare to shaving data in Table 2, line 1 and to data on 3-carboxypropanoyl-Gly-Gly-Phe-*p*-nitrophenylanilide (28); line 5, 3-carboxypropanoyl-Ala-Ala-Pro-Phe-*p*-nitrophenylanilide (29); line 6, PEG-2000-Ala-Gly-Val-Phe-*p*-nitrophenylanilide (30); lines 7–9, Boc-Ala-Ala-Xaa-*S*-benzyl thioester for Xaa = Ala, Leu, and Nva (31); line 11, Z-Ala-Ala-Phe-Phe-3-(4-pyridyl)propyl-1-oxo ester (32).

†The amount of cleavage was estimated spectrophotometrically based on Fmoc release at the indicated stage of the protocol outlined in the legend above. Numbers in the 1% range approach the sensitivity limits of the UV technique.

amounts (1–2%, corresponding to the maximum degree of chymotryptic shaving described along with Table 3); (ii) the synthesized surface peptides were detected readily by their receptors, for all three model peptides; (iii) the synthesized interior peptides were not detected by the receptors in the cases of Leu-His-Pro-Gln-Phe and D-Phe-Pro-Arg-Pro-Gly, although the interaction between Tyr-Gly-Gly-Phe-Leu and the anti- β -endorphin Ab was so strong that it was observed (i.e., false positive) despite shaving (Table 4, lines 1 and 4; compare with Table 3, lines 2 and 3, and text discussion above about sensitivity; finally note that such false positives are not observed with better substrates—e.g., Table 3, lines 5 and 6); (iv) for most purposes, Trp-Gly is an adequate substrate for chymotryptic shaving, better than Tyr-Gly or Phe-Gly; (v) the original surface/interior screening/coding concept is supported in the Leu-His-Pro-Gln-Phe–streptavidin/D-Phe-Pro-

Arg-Pro-Gly–thrombin systems (either order; refer to Table 4, lines 6 and 7); and (vi) results were the same regardless of whether the two peptides were assembled in series or by a zig-zag strategy.

Peptide-Encoded Model Peptide Combinatorial Library.

The principles of enzyme-mediated segregation were applied to the construction of a model combinatorial library bearing two peptides assembled by orthogonal chemistries; the peptide present in vast molar excess throughout the interiors of the beads encoded the screening peptides present only on the surface (Table 5; see footnote for coding scheme). For each of five randomization steps, there were ten building blocks, meaning that the library would have 10^5 members and the same number of conjugate coding peptides on the same beads. On the scale of the experiment, with $\approx 6 \times 10^5$ beads, this reflected a 6-fold degeneracy—i.e., on average each sequence

Table 4. Staining reaction of various beads bearing two peptides assembled after shaving of Boc-Trp-Gly-spacer-POE-PS with chymotrypsin

Line no.	Peptide in bead area		Binding to		
	Surface	Interior	Anti- β -endorphin	Streptavidin	Thrombin
1	Ac-GG	YGGFLWG	4+	0	ND
2	Ac-GG	LHPQFWG	0	0	ND
3	Ac-GG	fPRPGWG	ND	ND	0
4	LHPQFG	YGGFLWG	3+	5+	ND
5	YGGFLG	LHPQFWG	4+	0	ND
6	fPRPGG	LHPQFWG	ND	0	3+
7	LHPQFG	fPRPGWG	ND	3+	0

The procedure for shaving Boc-Trp-Gly-spacer-POE-PS matches Table 3, line 3. Qualitatively similar results to all those reported here were obtained using Boc-Tyr-Gly-spacer-POE-PS matching Table 3, line 2 (complicated though by the possibility of peptide growth off the unprotected phenolic side-chain), and studies starting with Boc-Phe-Gly-spacer-POE-PS matching Table 3, line 1 were abandoned due to an insufficient level of enzymatic cleavage. In setting up this table, the one-letter code for amino acids has been used in order to conserve space. Surface and interior bead areas are defined operationally based on the experimental design described in the text. After shaving, Fmoc chemistry gave the surface peptide and Boc chemistry gave the interior peptide; these chemistries were applied either in series or by a zig-zag strategy. The washes after Fmoc deprotections were collected and quantitated spectrophotometrically as a gauge for the extent of shaving (found, 1–2%). After final deprotection, biological testing (Table 1) was carried out. Quantitative Edman degradation was carried out on the beads from lines 4 and 5. The readouts clearly gave the sequences corresponding to the peptides designated as interior (93–114 pmol at first cycle, 68–73 pmol at fifth cycle). At the same time, phenylthiohydantoin derivatives corresponding to readout of the designated surface peptides were absent or present in trace amounts (<1%); the only exception came with Phe (4–5 pmol) attributed to preview or lag of the major interior peptide. ND, not determined.

Table 5. Peptide-encoded model peptide combinatorial library

Receptor (natural motif)	Hits (interior sequence)	Found binding motifs (surface, as decoded)
Anti- β -endorphin Ab (YGGFL)	Nva-Gln-Gln-Pro-Xaa Xaa = Val, Gln, Met, Ile, Gly	Tyr-Gly-Gly-Phe-Xaa Xaa = Pro, Gly, D-Phe, Leu, His
Streptavidin (LHPOF)	Xaa-Gly-Val-Asn-Pro Xaa = Gln, Pro, Met	Xaa-His-Pro-Gln-Phe Xaa = Gly, Phe, D-Phe
Thrombin (fPRPG)	Ile-Gln-Ala-Val-Met Ala-Gln-Ala-Val-Met Nle-Ala-Pro-Val-Pro	Leu-Gly-Arg-Pro-D-Phe Arg-Gly-Arg-Pro-D-Phe Ile-Arg-Phe-Pro-Phe

As detailed in *Materials and Methods*, the Boc-Trp-Gly-spacer-POE-PS was shaved with chymotrypsin, and the screening/coding peptides were assembled by zig-zag Fmoc/Boc chemistries according to the following coding scheme; Arg by Ala; Gln by Asn; Gly by Gln; His by Gly; Pro by Val; Leu by Ile; Phe by Pro; D-Phe by Met; Ile by Nle; Tyr by Nva. Beads showing specific reactions were sequenced (middle part of table) and decoded (right part of table).

was expected to appear on six beads. After final deprotection, the library was screened in the same three systems used for the model studies (Table 1). Beads that showed staining indicative of positive interactions were selected, retested to confirm specificity, sequenced, and decoded (Table 5). Five beads selected by the anti- β -endorphin assay were decoded to show the repeating motif Tyr-Gly-Gly-Phe-Xaa (Xaa = Pro, Gly, D-Phe, Leu or His), duplicating the parent natural motif. Similarly, the streptavidin assay gave rise to 17 colored beads, three of which were sequenced and decoded to reveal the expected motif Xaa-His-Pro-Gln-Phe. Of 15 beads colorized in the thrombin assay, five were confirmed to be directed at the active site by competition with the inhibitor D-Phe-Pro-Arg-chloromethyl ketone. These active beads were analyzed and decoded; subsequently, the discovered peptide sequences Leu-Gly-Arg-Pro-D-Phe, Ile-Arg-Phe-Pro-Phe, and Arg-Gly-Arg-Pro-D-Phe were resynthesized and tested in solution assays with a chromogenic substrate (24). The K_i values for these three peptides were respectively 11.5, 10.2, and 5.7 μ M, all of which were stronger than the expected inhibitor D-Phe-Pro-Arg-Pro-Gly ($K_i = 52 \mu$ M).

CONCLUSIONS

We have defined an effective enzyme-mediated spatial segregation strategy that allows differentiation between surface and interior areas of POE-PS beaded supports. The enzymatic approach requires a good substrate for shaving and appropriate affinity of the macromolecular biological receptor to its ligand. The optimized shaving variation was demonstrated successfully on a peptide-encoded model peptide combinatorial library, and it holds considerable promise for nonpeptide libraries encoded by peptides. The expected recognition motifs reemerged from the model library in the anti- β -endorphin Ab and streptavidin systems, and led to the discovery of a new thrombin ligand Arg-Gly-Arg-Pro-D-Phe ($K_i = 5.7 \mu$ M), which binds with an order of magnitude greater affinity than the natural motif.

We thank Lydie Vágnerová, Farid Abdul-Latif, and Dasha Cabel for their dedicated and capable technical assistance.

- Vágner, J., Krchňák, V., Sepetov, N. F., Štrop, P., Lam, K. S., Barany, G. & Lebl, M. (1994) in *Innovation and Perspectives in Solid Phase Synthesis: Peptides, Proteins, and Nucleic Acids, Biological and Biomedical Applications 1994*, ed. Epton, R. (Mayflower Worldwide, Birmingham, U.K.), pp. 347–352.
- Geysen, H. M., Rodda, S. J. & Mason, T. J. (1986) *Mol. Immunol.* **7**, 709–715.
- Furka, A., Sebestyen, F., 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., Kazmieriski, W. M. & Knapp, R. J. (1991) *Nature (London)* **354**, 82–84.
- Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T. & Cuervo, J. H. (1991) *Nature (London)* **354**, 84–86.
- Lebl, M., Krchňák, V., Sepetov, N. F., Seligmann, B., Štrop, P., Felder, S. & Lam, K. S. (1995) *Biopolymers* **37**, 177–197.
- Gallop, M. A., Barrett, R. W., Dower, W. J., Fodor, S. P. A. & Gordon, E. M. (1994) *J. Med. Chem.* **37**, 1233–1251.
- Gordon, E. M., Barrett, R. W., Dower, W. J., Fodor, S. P. A. & Gallop, M. A. (1994) *J. Med. Chem.* **37**, 1385–1401.
- Zambias, R. A., Boulton, D. A. & Griffin, P. R. (1994) *Tetrahedron Lett.* **35**, 4283–4286.
- Youngquist, R. S., Fuentes, G. R., Lacey, M. P. & Keough, T. (1995) *J. Am. Chem. Soc.* **117**, 3900–3906.
- Fitch, W. L., Detre, G. & Holmes, C. P. (1994) *J. Org. Chem.* **59**, 7955–7956.
- Fodor, S. P. A., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T. & Solas D. (1991) *Science* **251**, 767–773.
- DeWitt, S. H., Kiely, J. S., Stankovic, C. J., Schroeder, M. C., Reynolds Cody, D. M. & Pavia, M. R. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 6909–6913.
- Frank, R. (1993) *Bioorg. Med. Chem. Lett.* **3**, 425–430.
- Kerr, J. M., Banville, S. C. & Zuckermann, R. N. (1993) *J. Am. Chem. Soc.* **115**, 2529–2531.
- Nikolaiev, V., Stierandová, A., Krchňák, V., Seligmann, B., Lam, K. S., Salmon, S. E. & Lebl, M. (1993) *Pept. Res.* **6**, 161–170.
- Krchňák, V., Weichsel, A. S., Cabel, D. & Lebl, M. (1995) *Pept. Res.* **8**, 198–205.
- Nielsen, J., Brenner, S. & Janda, K. D. (1993) *J. Am. Chem. Soc.* **115**, 9812–9813.
- Needels, M. C., Jones, D. G., Tate, E. H., Heinkel, G. L., Kochersperger, L. M., Dower, W. J., Barrett, R. W. & Gallop, M. A. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 10700–10704.
- Ohlmeyer, M. H. J., Swanson, R. N., Dillard, L. W., Reader, J. C., Asouline, G., Kobayashi, R., Wigler, M. & Still, W. C. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 10922–10926.
- Bayer, E. & Rapp, W. (1992) in *Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Application*, ed. Harris, J. M. (Plenum, New York), pp. 325–345.
- Barany, G. & Merrifield, R. B. (1979) in *The Peptides*, eds. Gross, E. & Meienhofer, J. (Academic, New York), Vol. 2, pp. 1–284.
- Lam, K. S. & Lebl, M. (1992) *Immunomethods* **1**, 11–15.
- Kelly, A. B., Maraganore, J. M., Bourdon, P., Hanson, S. R. & Harker, L. A. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 6040–6044.
- Fields, G. B., Tian, Z. & Barany, G. (1992) in *Synthetic Peptides: A User's Guide*, ed. Grant, G. A. (Freeman, New York), pp. 77–183.
- Krchňák, V. & Vágner, J. (1990) *Pept. Res.* **3**, 182–193.
- Krchňák, V., Vágner, J., Šafář, P. & Lebl, M. (1988) *Coll. Czech. Chem. Commun.* **53**, 2542–2548.
- Kasafírek, E. & Bartík, M. (1980) *Collect. Czech. Chem. Commun.* **45**, 442–451.
- DelMar, E. G., Largman, C., Brodrick, J. W. & Geokas, M. C. *Anal. Biochem.* **99**, 316–320.
- Ulbrich, K., Strohalm, J. & Kopeček, J. (1986) *Makromol. Chem.* **187**, 1131–1144.
- Kasafírek, E., Frič, P., Slab, J. & Mališ, F. (1976) *Eur. J. Biochem.* **69**, 1–13.
- Sachdev, G. P. & Fruton, J. S. (1970) *Biochemistry* **9**, 4465–4469.