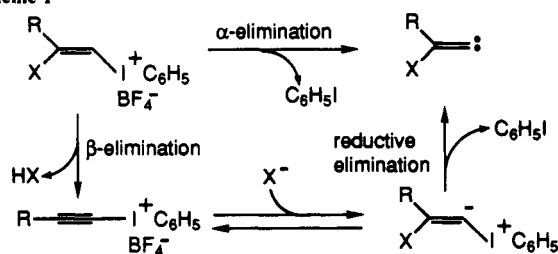
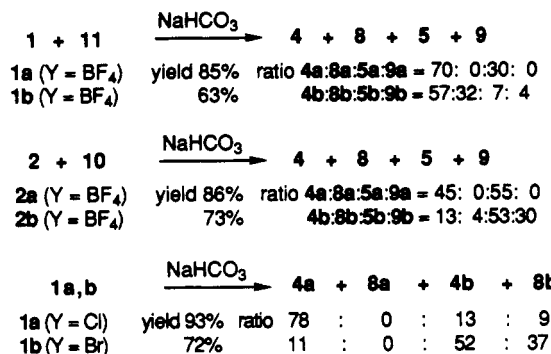


Scheme I



Scheme II

Table I. Reaction of (Z) - $(\beta$ -Halovinyl)iodonium Salts with Bases^a

substrate	reactn time, h	product, % yield ^b		
		alkyne	olefin	ratio ^c
1a (Y = Br)	4	4a, 98	8a, 0	100:0
1b (Y = Cl)	5	4b, 52	8b, 36	59:41
1b (Y = Cl)	4 ^d	4b, 59	8b, 41	59:41
1c (Y = Cl)	10 ^d		8c, 17 ^e	
2a (Y = Br)	5	5a, 95	9a, 0	100:0
2b (Y = Cl)	5	5b, 54	9b, 33	62:38
2b (Y = Cl)	19 ^d	5b, 41	9b, 25	62:38
3a (Y = Br)	5.5	6a, 77		
3b (Y = Cl)	6	6b, 92		

^a Unless otherwise noted, reactions were carried out using 1.2 equiv of NaHCO_3 at 0 °C in CH_2Cl_2 -MeOH-H₂O. ^b Determined by gas chromatography using an internal standard. ^c Ratios of 1,2-shift of halogens to 1,5-C-H insertions. ^d Reactions were carried out using 2-4 equiv of $n\text{-Bu}_4\text{NF}$ at room temperature in CH_2Cl_2 . ^e Isolated yield.

finding that 1,2-chlorine migration of alkylidene carbenes competes with 1,5-C-H insertion, make the comparison between the migratory aptitude of an α -phenyl group and α -halogen atoms very interesting. Because of the instability of $(\beta$ -halo- β -phenylvinyl)iodonium salts,¹⁰ (α -halo- α -phenylalkylidene)carbenes were directly generated from **12** through Michael type addition-reductive elimination sequences.^{8b} The reaction of [2-¹³C]-**12** (99% enriched) with LiX (X = Cl, Br, and I) in CH_2Cl_2 -MeOH at -78 °C afforded good yields of 1-halo-2-phenylacetylenes **7**. The ¹³C-enrichment at C-2 of **7** was found to be more than 98% by ¹³C NMR spectra. While there is no report concerning the migratory aptitude of α -halogen atoms of alkylidene carbenes,¹¹ these results clearly indicate that the rate of 1,2-migration of α -halogen atoms (I, Br, and Cl) of alkylidene carbenes is much greater than that of an α -phenyl group.

Acknowledgment. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan.

(9) For 1,2-migration of α -aryl groups of alkylidene carbenes, see: (a) Newman, M. S.; Kutner, A. *J. Am. Chem. Soc.* **1951**, *73*, 4199. (b) Stang, P. J.; Mangum, M. G.; Fox, D. P.; Haak, P. *J. Am. Chem. Soc.* **1974**, *96*, 4562. (c) Brown, R. F. C.; Eastwood, F. W.; Harrington, K. J.; McMullen, G. L. *Aust. J. Chem.* **1974**, *27*, 2393. (d) Stang, P. J.; Fox, D. P.; Collins, C. J.; Watson, C. R. *J. Org. Chem.* **1978**, *43*, 364. (e) Gilbert, J. C.; Weerasooriya, U. *J. Org. Chem.* **1982**, *47*, 1837.

(10) Beringer, F. M.; Galton, S. A. *J. Org. Chem.* **1965**, *30*, 1930.

(11) For 1,2-migration of halogen atoms of (α -haloalkyl)carbenes, see: (a) Bevan, W. I.; Haszeldine, R. N.; Young, J. C. *Chem. Ind. (London)* **1961**, 789. (b) Liu, M. T. H.; Bonneau, R. *J. Org. Chem.* **1992**, *57*, 2483.

Encoded Combinatorial Peptide Libraries Containing Non-Natural Amino Acids

Janice M. Kerr, Steven C. Banville, and Ronald N. Zuckermann*

Chiron Corporation, 4560 Horton Street
Emeryville, California 94608

Received December 22, 1992

Methods for the synthesis and screening of diverse peptide libraries have recently been developed for the rapid discovery of biologically-active lead compounds.¹⁻⁴ We describe here a method for encoding⁵ non-natural components⁶ in a diverse combinatorial library with standard amino acids by the parallel and alternating synthesis of two polymer chains: a binding ligand and a coding peptide. The use of a peptide tag allows the sequence of any isolated binding ligand to be identified by conventional peptide analyses and thus circumvents the problems sometimes associated with the analysis of novel biopolymers.

This combinatorial encoding strategy utilizes a resin-splitting peptide synthesis method^{7,8} to alternately synthesize a "binding" strand and a "coding" strand (Figure 1). Orthogonal protecting groups are used to allow for the individual and alternating extension of both polymer strands on each resin bead. Specifically, base-labile, N^α -[(9-fluorenylmethyl)oxy]carbonyl-protected (Fmoc-protected) monomers and acid-labile, N^α -[[2-(3,5-dimethoxyphenyl)prop-2-yl]oxy]carbonyl-protected (Ddz-protected)⁹ amino acids are used to synthesize the binding and coding strands, respectively. Although these two groups are orthogonal, the use of *tert*-butyl ester or trityl side chain protecting groups may require the use of an N^α protecting group with greater acid lability than Ddz.

The relationship between the binding and coding strands can assume a variety of configurations. The number of Fmoc monomers that can be represented depends on both the number of different Ddz amino acids used and on the length of the Ddz code. An efficient coding strategy requires the addition of Ddz amino acids only at a mixture position (where $n > 1$ in Figure 1). In the example presented here, four Ddz-protected amino acids were used in trimer sequences to allow for the representation of up to 64 non-natural monomers. Ddz-protected leucine, phenylalanine, glycine, and alanine were chosen as encoding monomers because they do not require side chain protection, and they give reproducibly strong signals upon Edman sequencing.

The isolation of receptor-binding ligands from a solution-phase or solid-phase encoded library can be performed by affinity selection⁴ or bead-staining techniques.² The identity of the binding sequence can then be determined by Edman sequencing of the coding strand. To avoid ambiguity, only the coding strand should be sequenced; the binding sequence must be acetylated or otherwise made nonsequenceable. Binding sequences of identical molecular weight can be distinguished by using different codes, allowing

(1) Fodor, S. P. A.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D. *Science* **1991**, *251*, 767-773.

(2) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmiarsky, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82-84.

(3) Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84-86.

(4) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C.; Santi, D. V. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4505-4509.

(5) Brenner, S.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5381-5383.

(6) 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.; Santi, D. V.; Cohen, F. E.; Bartlett, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9367-9371.

(7) Furka, A.; Sebastyen, M.; Asgedom, M.; Dibó, G. *Int. J. Pept. Protein Res.* **1991**, *37*, 487-493.

(8) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C. *Int. J. Pept. Protein Res.* **1992**, *40*, 498-507.

(9) Voss, C.; Birr, C. *Hoppe-Seyler's Z. Physiol. Chem.* **1981**, *362*, 717-725.

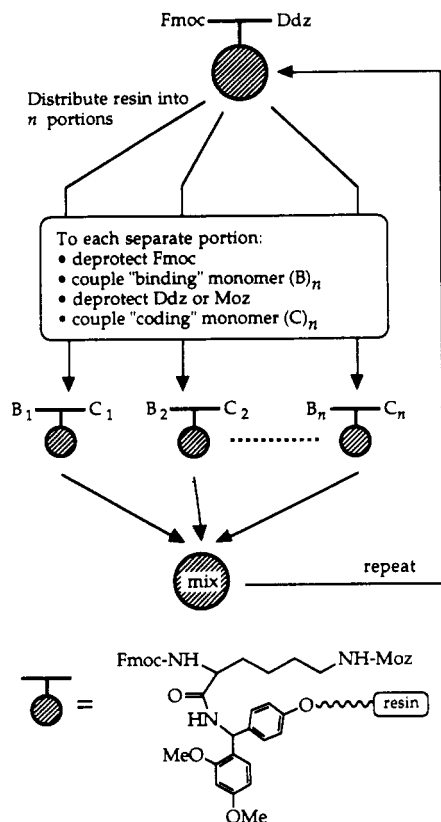


Figure 1. The synthesis of an encoded library consists of the following steps: (1) a bifunctional linker containing two orthogonally protected points of extension (N^{α} -Fmoc- N^{ϵ} -Moz-Lys-OH) is attached to polystyrene resin via an acid-labile linker, (2) the solid support is divided into n equal portions at a mixture position, (3) a unique N^{α} -Fmoc-protected non-natural monomer (B) is coupled to the "binding" strand, (4) a series of N^{ϵ} -Ddz-protected amino acids (C) are then coupled to the "coding" strand, and (5) the solid supports are recombined.

Table I. Composition of an Encoded Ligand Mixture: Ac-RAX₃HTTGX₂IX₁-NH₂

mixture position	non-natural monomer mixture components	amino acid code
X ₃	L-naphthylalanine L-homophenylalanine D-phenylalanine cyclohexyl-L-alanine	
X ₂	L-arginine D-arginine <i>N</i> -(3-guanidinopropyl)glycine <i>N</i> -(2-aminoethyl)glycine ornithine	FAL GLF LFL AGL GAF
X ₁	<i>N</i> - <i>sec</i> -butylglycine <i>N</i> -isobutylglycine <i>N</i> - <i>n</i> -butylglycine <i>N</i> - <i>n</i> -propylglycine <i>N</i> -cyclopropylglycine <i>N</i> -cyclohexylglycine cyclohexyl-L-alanine D-leucine D-valine norvaline	LGA LAF GFG AFA FGL ALF LGL FAF GAG ALG

verification by mass spectrometry.

To test this methodology, a 10mer epitope (Ac-RAFHTTGR_{II}-NH₂), known to bind to an anti-gp120 monoclonal antibody with submicromolar affinity, was substituted at three positions (designated by X₁, X₂, and X₃) with non-natural amino acids, N-substituted glycines, and conservative amino acid replacements (Table I). A total of 200 peptides were synthesized as four equimolar mixtures,¹⁰ each containing a unique X₃ residue. An

Table II. Competitive Inhibition of the Antibody-gp120 Interaction by the Four RAX₃HTTGX₂IX₁ Pools

identity of residue at X ₃	-log IC ₅₀ (M) ^a
L-naphthylalanine	7.5
L-homophenylalanine	6.2
D-phenylalanine	6.7
cyclohexyl-L-alanine	6.5

^aThese values represent the concentration of each mixture component.

Table III. Affinity Selected Components of X₃ = Naphthylalanine Pool

peak no.	coding strand sequence ^a		binding strand sequence		-log IC ₅₀ (M) of binding strand ^b
	X ₂	X ₁	X ₂	X ₁	
1	AGL	ALG	<i>N</i> -(2-aminoethyl)-glycine	norvaline	6.5
2 ^c	FAL	ALG	L-arginine	norvaline	7.4
3	FAL	GFG	L-arginine	<i>N</i> - <i>n</i> -butylglycine	7.0

^aDetermined by Edman sequencing. ^bThese decapeptides are N-acetylated and C-amidated. ^cThe identity of this compound was verified by mass spectrometry.

N^{α} -Fmoc-Lys(N^{ϵ} -4-methoxybenzylloxycarbonyl) (N^{ϵ} -Moz) linker¹¹ was attached to the resin (Figure 1) and the binding and coding strands were extended from the α and ϵ groups, respectively. An N-terminal phenylalanine was added to the coding strand to serve as an internal standard for Edman sequencing. Following completion of the synthesis, the peptides were cleaved from the resin to provide the solution-phase compounds as 1:1-binding/coding strand adducts (i.e., covalently linked via the lysine amino groups).

A competition binding assay was then used to determine the IC₅₀ values of these four pools (Table II). The pool with L-naphthylalanine at the X₃ position showed the greatest inhibition (-log IC₅₀ = 7.5) and likely contained the peptide(s) of greatest binding affinity. Isolation of the highest affinity peptide(s) in the X₃ = L-naphthylalanine pool was accomplished by affinity selection^{4,13} with the anti-gp120 antibody. Three major high-affinity components were observed and isolated by HPLC. Edman sequencing of the coding strand was performed to identify the substituents at X₁ and X₂ in the binding strand of these three isolated components (Table III).

The affinities of the decapeptide sequences that corresponded to the binding strands of the three isolated compounds were then confirmed by independent synthesis and assay. The binding decapeptide strand of the major affinity-selected product had the greatest affinity of all three peptides and had an affinity comparable to that of the original 10mer epitope. In order to determine the effect of the appended coding strand, the binding/coding adduct Ac-RABHTTGRIJK(ϵ -(FFALALG))A-NH₂ (where B = L-naphthylalanine, J = L-norvaline)—as originally selected—was also independently synthesized and assayed. The IC₅₀ value of the binding/coding adduct (-log IC₅₀ = 7.5) was similar to that of the binding decapeptide (-log IC₅₀ = 7.4), suggesting that, at least in this case, the coding sequences did not interfere significantly with the binding properties of the compounds in the mixture.

In conclusion, we have developed a method for encoding non-natural components in a combinatorial library with standard amino

(10) N^{α} -Fmoc groups were removed with 20% piperidine/DMF for 20 min. N^{ϵ} -Ddz or N^{ϵ} -Moz groups were removed with 5% trifluoroacetic acid/dichloromethane for 20 min, followed by neutralization with 2% *N,N*-diisopropylethylamine/DMF. Amino acids were coupled to the resin for 30 min as 0.3 M solutions in DMF containing 0.3 M 1-hydroxybenzotriazole and 0.3 M *N,N'*-diisopropylcarbodiimide. The resin-splitting algorithm was performed as described⁶ with the modifications shown in Figure 1.

(11) This compound was prepared from N^{α} -Fmoc-L-lysine and 4-methoxybenzylloxycarbonyl azide.¹²

(12) Wang, S. S.; Chen, S. T.; Wang, K. T.; Merrifield, R. B. *Int. J. Pept. Protein Res.* 1987, 30, 662-667.

(13) Kerr, J. M.; Banville, S. C.; Zuckermann, R. N. *Biol. Med. Chem. Lett.* 1993, 3, 463-468.

