Isotope or mass encoding of combinatorial libraries

H Mario Geysen, Craig D Wagner, Wanda M Bodnar, Christopher J Markworth, Gillian J Parke, Frank J Schoenen, David S Wagner and Daniel S Kinder

Background: Combinatorial chemistry using solid-phase synthesis is a rapidly developing technology that can result in a significant reduction in the time required to find and optimize lead compounds. The application of this approach to traditional medicinal chemistry has led to the construction of libraries of small organic molecules on resin beads. A major difficulty in developing large combinatorial libraries is the lack of a facile encoding and decoding methodology to identify active compounds.

Results: Several encoding schemes are described which use the ability of mass spectrometry to ascertain isotopic distributions. Molecular tags are attached to resin beads in parallel or on the linker used for chemical library synthesis. The tags are encoded via a controlled ratio of a number of stable isotopes on the tagging molecules, and range from a single to a complex isotopic distribution.

Conclusions: A novel coding scheme is described that is useful for the generation of large encoded combinatorial libraries. The code can be cleaved after assay and analyzed by mass spectrometry in an automated fashion. An important element of the combinatorial discovery process is the ability to extract the structure–activity relationship (SAR) information made available by library screening. The speed and sensitivity of the mass-encoding scheme has the potential to determine the full SAR for a given library.

Address: Glaxo Wellcome, Inc., Research Triangle Park, NC 27709, USA.

Correspondence: H Mario Geysen
e-mail: hmg5823@glaxo.com

Key words: combinatorial chemistry, encoding, mass spectrometry, stable isotope

Received: **22 Jul 1996**
Accepted: **2 Aug 1996**

Chemistry & Biology August 1996, **3**:679–688

 ISSN 1074-5521

Introduction

The term 'combinatorial chemistry' describes a collection of techniques for the synthesis of arrays or libraries of related chemical entities, and relies extensively on encoding and decoding strategies for the identification of active compounds. Combinatorial chemistry has been aggressively adopted by the pharmaceutical industry for its lead compound discovery and optimization programs, as it can generate large numbers of compounds for evaluation in appropriate assays. In its simplest form, compounds are produced as discrete entities in a spatially addressable format in modest numbers [1,2]. At the other extreme, millions or more compounds are synthesized in pools requiring extensive iterations of testing and resynthesis to determine which compounds are active [3].

Coding strategies fall into two broad categories, each encompassing several variations. In one strategy, the library is synthesized as pools of compounds. After assay, active pools are resynthesized as discrete compounds or as pools with ever smaller numbers of compounds until the active compounds are identified [3]. In the second strategy a different compound is chemically synthesized on every bead (resin particles) [4]. Either assays are performed with the compound still attached to the bead or the compound is cleaved from the bead, while its relationship to the bead is retained. The structures of active compounds are identified by 'reading' a code from the bead or by direct sequencing of the compound itself. A recent innovation imbeds a microcircuit code transmitter into a macro version of a solid support suitable for chemical synthesis, at the expense, however, of the numbers of pellets that can be used at an acceptable scale of synthesis [5].

Peptide and DNA/RNA libraries capitalize on well established direct sequencing techniques for identification of the desired compound. Several innovative strategies have been described for encoding bead libraries of small organic molecules with a moiety that can be sequenced. These include codes assembled from amino acids [6], nucleotides [7,8], halogenated benzenes [9] and secondary amides [10], but these techniques all require that the necessary chemistry to assemble the code should be orthogonal to that required for synthesis of the library. This represents an additional investment at the chemistry development stage, and requires added steps during library synthesis and often a numerically limiting procedure to read the code.

An essential component of any library synthesized in a solid-phase format is the chemical entity (link) on which synthesis is carried out. We describe a novel coding

strategy which is 'embedded' in this unavoidable synthetic overhead, and does not require adding a code element at each step of the synthesis. Furthermore, the code is read, after cleavage from the bead, using a simple and rapid mass spectroscopic procedure. Alternatively, the code can, in principle, be read directly using nuclear magnetic resonance spectroscopy (NMR), while still attached to the bead.

Results

Isotopic or mass encoding of bead libraries

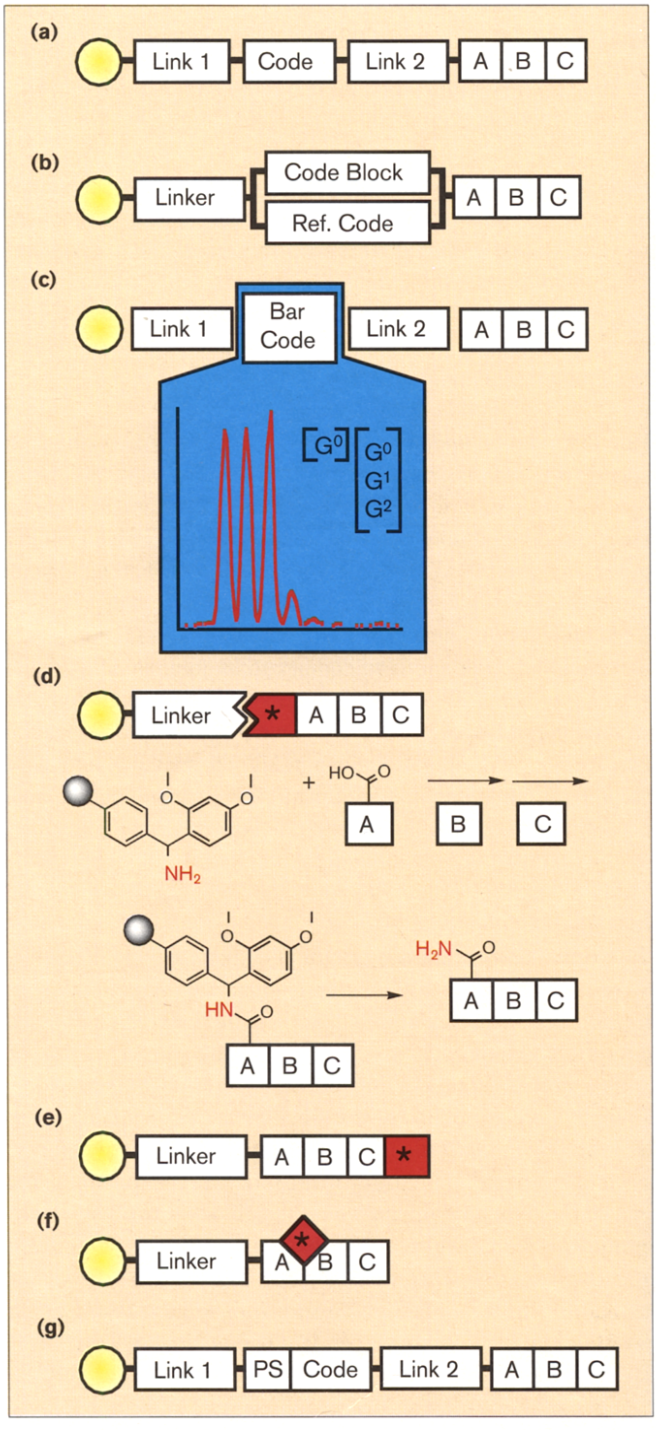
Mass spectrometry (MS) is a rapid and sensitive analytical tool routinely used to characterize and quantitate compounds. Electrospray ionization and matrix assisted laser desorption MS have been used to identify compounds from a single bead [11], such as those used in library synthesis. Although the absolute sensitivity of electrospray instruments can be a limiting factor when using the very smallest beads (10μm), rapid code-reading times of typically <1 min are readily achieved using beads of >100 μm. Chemically pure small molecules usually only appear once in the spectrum as a singly charged ion. Any isotopic compositional variation, as occurs naturally for carbon, results in a very predictable quantized mass shift, which is independent of the molecule's tendency to ionize. The ability to assign mass with an accuracy of 0.1 atomic mass unit (AMU) and relative intensity with an error level of ~3 % under ideal conditions and 5–10% in the worst case enables a novel encoding strategy using isotopically different mass units. While in principle any number of codes can be generated, it is preferable to minimize the size of, and thus chemical investment in, the encoding unit. For greatest sensitivity, the code should be present in the smallest number of peaks, preferably closely spaced in the spectrum. With these caveats in mind, the four basic encoding strategies described below are easily implemented, with the option of increasing the number of available codes by selective combinations of these four.

Single peak positional encoding

In this format (Fig. 1a), link 2 is initially cleaved to allow the compound to be tested for activity, and, when the compound tests positive, link one is cleaved and the coding block submitted for MS analysis. The mass of the coding block is engineered such that it appears in a convenient part of the spectrum, and it is used to represent the first monomer or building block used in the synthesis of a particular compound. An example in which ten distinct codes are generated is shown in Figure 2. Here the code block comprises a peptide spacer made up from the commercially available isotopic versions of glycine and alanine.

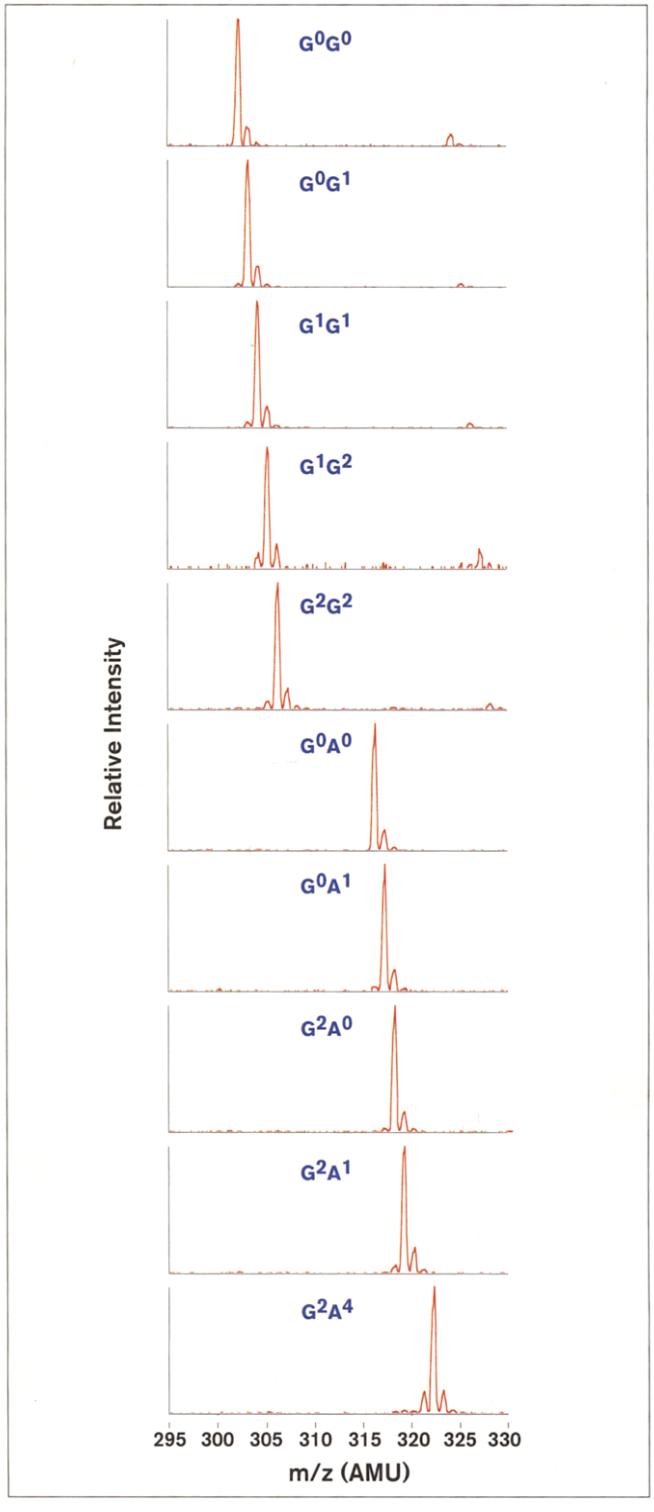
In a practical implementation (Fig. 3), to encode a library in which three monomers are chemically combined, the code block defines the identity of the first monomer, while, for a

Figure 1



Seven encoding options that can be read by mass spectrometry. **(a)** Single peak positional encoding relies on the absolute mass of the code component. **(b)** Double peak positional encoding uses a comparison of the mass of the code component with a reference component. **(c)** The bar code uses two (or more) isotopically mixed components. **(d–f)** Ratio encoding options in which the starred block represents an atom transferred into the ligand that is made up from a varying ratio of isotopes. **(g)** Peak splitting, where the splitting element shown adjacent to the code block is introduced as a mixture of two identical molecular entities differing only in isotopic composition in one or more of its atoms.

Figure 2



Single peak positional coding. Mass spectrum covering the mass range 295–330 AMU for 10 single peak positional codes derived from a spacer element comprising two residues (shown inset in each spectrum) taken from the set of readily available isotopic variants of glycine and alanine: G^0, G^1, G^2, A^0, A^1, A^4. Single letter codes are used for amino acids, with the accompanying number indicating the level of isotope doping used in terms of AMUs; for example, G^2 indicates glycine containing two ^{13}C atoms.

conventional split and recombine bead library synthesis, the third monomer is defined by the final pool from which the particular bead was derived. Determining the mass of the ligand and code block therefore allows the identity of the second monomer to be deduced by subtracting the now known monomer one and monomer three masses from the total ligand mass. This assumes no mass redundancies in the monomer two set used in the library synthesis.

The natural abundance of ^{13}C (1.1 %) results in side peaks at plus one AMU with 0.1–0.32 of the total intensity ascribed to the particular compound for molecules with 10–50 carbon atoms. In practice, the computer algorithm used to analyze the spectrum compensates for all the side peaks based on the actual or estimated number of carbon atoms present.

Double peak positional encoding

In this format (Fig. 1b), the code block is prepared to contain two like species, one of which is constant and acts as a reference mass, while the other has mass blocks similar to those described for the single peak positional encoding scheme. After cleavage of the link to the resin, ligand is generated with an attached constant reference mass increment as well as an attached variable code increment. The encoded information is now determined from the difference in mass of the coding unit and the reference mass unit. The advantage of this is that the code can be left attached to the ligand and read at the same time as the mass of the whole construct is determined by MS. An example of this encoding scheme is shown in Figure 4. This strategy gives two mass peaks that represent the ligand attached to either the reference code or the variable code. The difference in mass is used to determine the identity of monomer 1, while monomer 3 is known from the corresponding pool and therefore monomer 2 is inferred as described above.

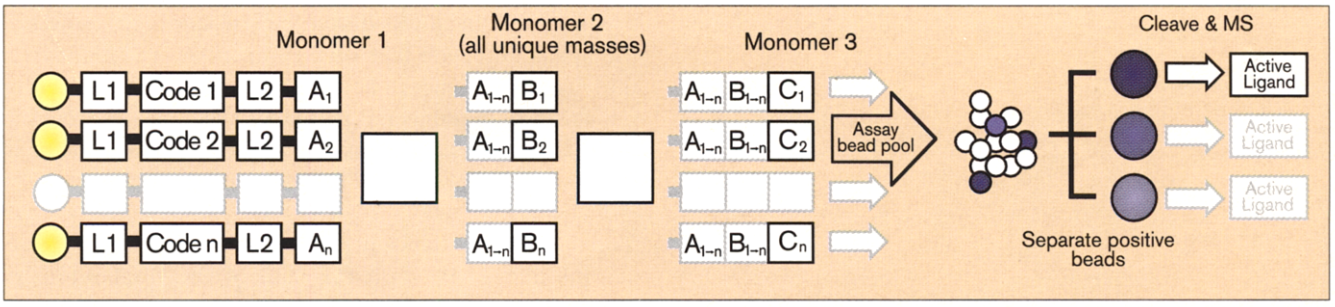
Bar code

This format (Fig. 1c) requires the use of mixed isotopically distinct code monomers at each step of code synthesis. The isotopically distinct species are present in equimolar concentrations and are therefore incorporated equally as their individual reactivity is not affected by the isotopic composition. A practical example in which 25 distinct MS patterns (codes) are generated over a 5-AMU region of the spectrum is shown in Figure 5. Decoding is as described for the single peak positional encoding scheme, with the added advantage that all the information is present over a very small mass range allowing for very rapid determination with good sensitivity in terms of total material needed to obtain the spectrum.

Ratio code

This coding scheme (Fig. 1d,e,f) uses a significantly different strategy from those described above. In its

Figure 3



Minimal library-encoding scheme. A basic split and recombine resin procedure [4] is depicted in which n A monomers are ligated to n linkers with their respective codes; these conjugates are then mixed and two more monomers added before assaying for activity. This yields a library of n^3 compounds that can be assayed and decoded.

simplest form the encoded information is present in just two peaks. The identity of the code can be determined with the code block either attached to or free from the ligand, depending on the number of links used in the complete construct. This technique has the added advantage that it can be implemented at the level of a single atom, and the code can be conveniently inserted whenever a common reagent is used which transfers at least one atom to the target compound or ligand.

At the encoding step, a mixture of a component common to all the library, such as the link or a reagent used during the synthesis, is prepared from different ratios of isotopically distinct but otherwise identical compounds. The resulting MS spectrum will show two peaks separated by the difference in mass of the two isotopes used, with the relative heights of the peaks (areas) determined by the ratio of the two species present. Ideally the isotopically distinct species used should differ by two or more AMUs as this would produce well separated peaks that would not be affected by the natural presence of ^{13}C (see above). To achieve this with the atoms that are most commonly used in such chemistry (C,H,N,O) typically requires that more than one atom be istopically substituted, as available stable isotopes of C, H and N differ by only one mass unit. Optimized rapid flow injection MS gives an accuracy in the peak area or height of ~3 %, so it is possible to define 21 codes based on 5 % increments in the ratio of the isotopically distinct unit used. An example in which a photocleavable link is made up using varying ratios of homogeneous ^{14}N and ^{15}N versions of the linker is shown in Figure 6.

Variations on a theme

The above examples represent a basic repertoire of encoding strategies using the fundamental aspects of a mass spectrometer and mass units ranging from single atoms to whole molecules in which the isotopic composition varies. In most cases the number of codes obtainable by this modest investment in the construction of the linking region is relatively small, of the order of 20–100, which is substantially less than the size of the combinatorial libraries that are commonly made. Selective combinations of the four basic strategies can, however, result in a significant increase in the number of available codes.

Peak splitting

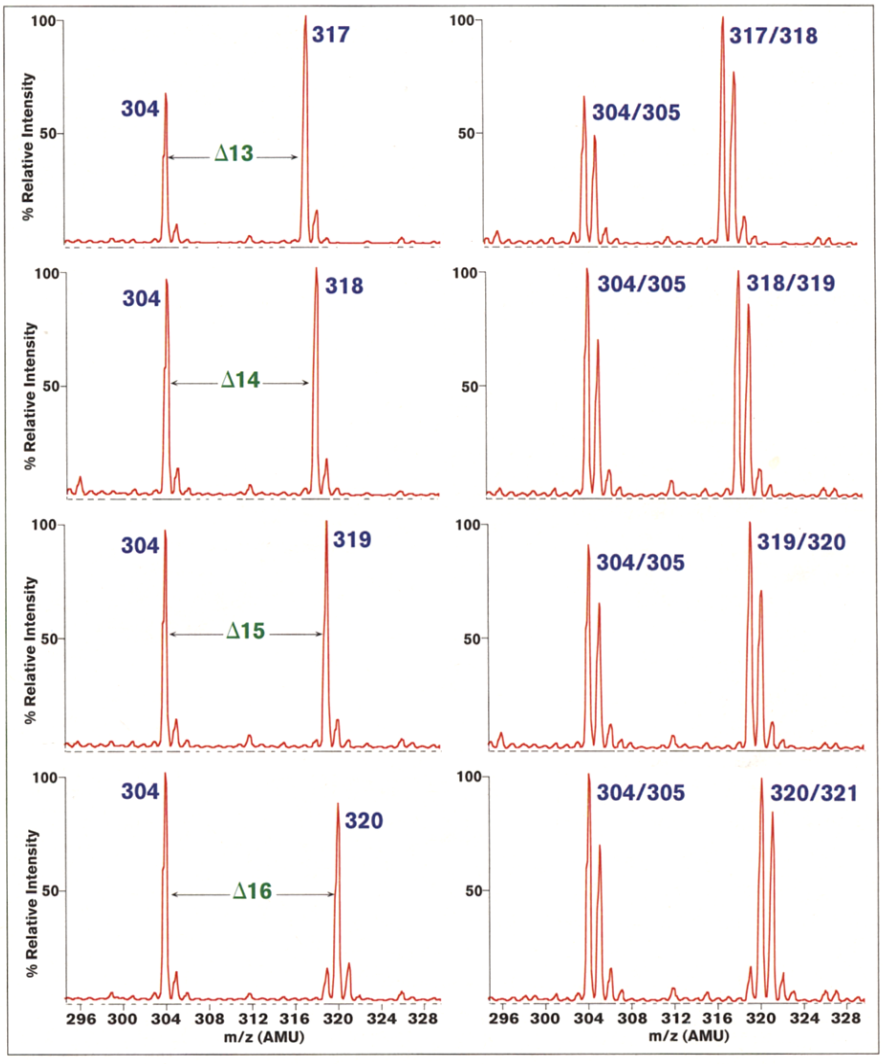
Addition of a peak-splitting element (Fig. 1g), consisting of an approximately 1:1 ratio of two otherwise identical entities that differ in mass, splits any peaks deriving from its inclusion into two. This makes it very easy to recognize in the spectrum (Fig. 4), and this method is particularly useful in monitoring the results of an individual synthesis or in highlighting code peaks for automated analysis. To achieve approximately equal peak intensity an equimolar concentration of the two components used should differ by 2 or more AMUs. In the case where the splitting results from a single atom, for example the $^{14}N/^{15}N$ mixed linker, the best ratio can be calculated from the formula $^{15}N/^{14}N = 1/(2 - 0.011 * n)$, where n is the average number of carbon atoms in the final compounds. The need to vary the ratio from 1:1 again results from the natural abundance of ^{13}C, and the significant number of carbon atoms in those compounds typically synthesized as part of combinatorial libraries.

Combination bar code and ratio coding

Of the many combinations possible, the one which most closely meets the requirements of keeping the code peaks localized to a small region of the MS is the combination of the ratio code with the bar code. Both are easily recognizable, and therefore lend themselves to easy decoding using an automated computer algorithm. The bar code is based on isotopically distinct glycine pairs, whereas the ratio portion is based on a glycine-alanine pair so that this part of the code is shifted away from the region in the spectrum occupied by the bar code. If the bar code is made up of five peaks by using glycines taken from the

Figure 4

Double peak codes. This encoding scheme results in two peaks in the mass spectrum; the code is the mass difference between the lower molecular weight peak (reference at 304), and the higher molecular weight peak. The position of this latter peak depends on the particular combination of isotopically distinct glycines or alanines used. Spectra on the left were derived from single beads using the scheme shown in Figure 1b. Spectra on the right were derived using the scheme shown in Fig. 1b, with a peak-splitting element (an equimolar mixture of G^0 and G^1) located in each arm to produce the double peaks. This enhancement to the generated code allows for easier automated code identification when analyzing the mass spectrum.

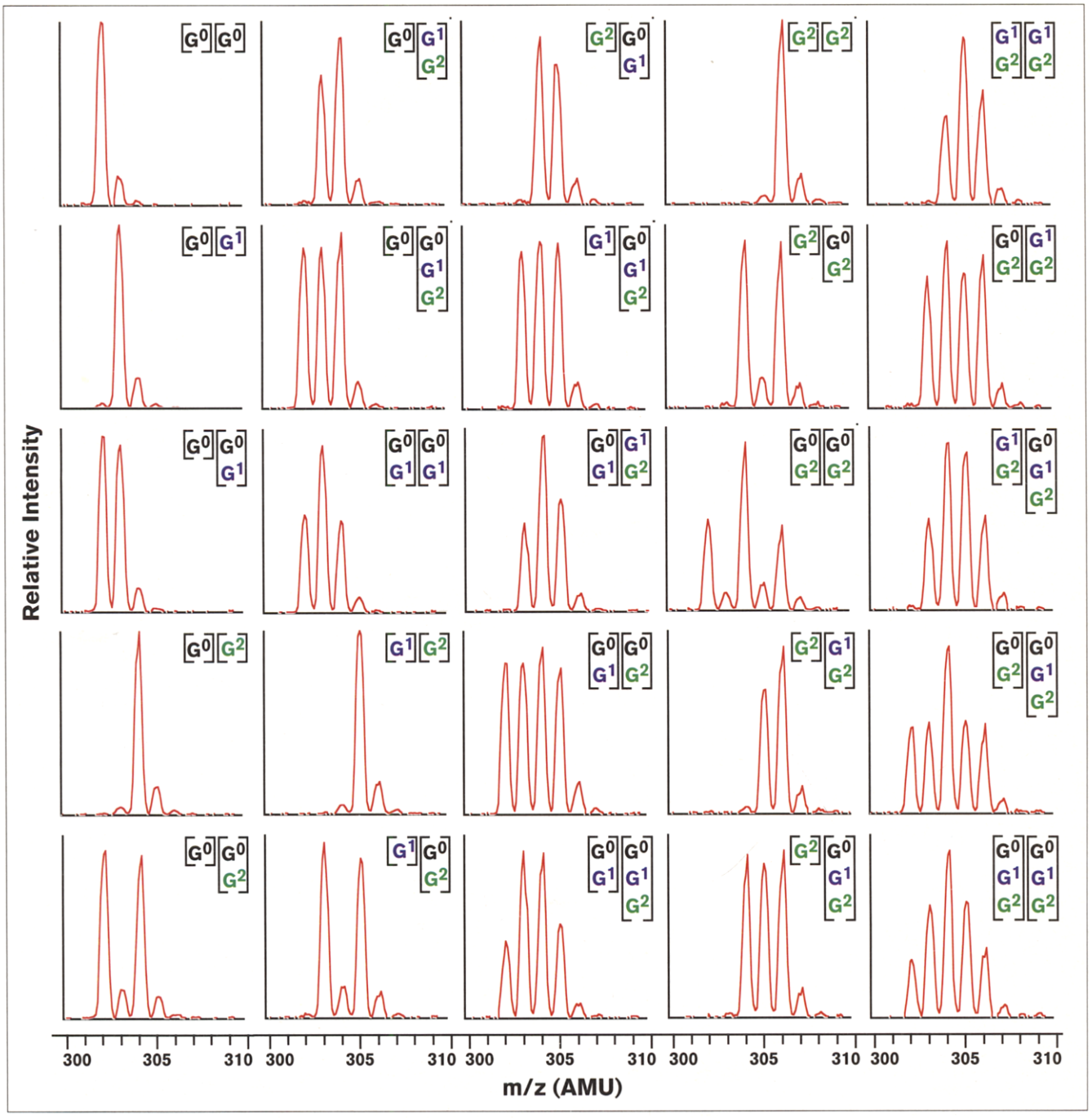


set [G^0,G^1,G^2] (as described in Fig. 5), and the ratio code of two peaks by using glycine mixtures in just 9 % increments taken from the set [G^0,G^1,G^2,G^3,G^4,G^5], 3750 distinct codes are realizable, as summarized in Figure 7. This number is easily increased if other permutations of isotopically distinct glycines are used to generate the bar portion of the code, or if the ratio increments are reduced to 5 % instead of the conservative 10 %.

With this number of codes an explicit coding scheme suitable for a three monomer library using 20 different monomers at each position, with or without mass redundancy, is readily implemented as shown in Figure 8. In this example the initial coupling of monomer A is in sets of 20 and therefore starts with 400 pools of uniquely coded resin, which, after combination, reduces to just 20 couplings for monomer B at the second step. This large number of initial pools is necessary if the identity of both monomer A and monomer B are to be specified by the initial label. Monomer C is coupled to 20 pools of beads after combining and resplitting all the resin, and the assay performed on each of the final pools without recombining.

A variation on this scheme starting with just 20 pools as per a more conventional split pool and recombine strategy would require a more elaborate orthogonal protection scheme to allow the synthesis of one of the code formats in conjunction with the coupling of monomer B at the second step. Alternatively, where a common reagent is used in conjunction with the second step in the library synthesis, a ratio code could be introduced directly without the requirement for an orthogonal protection scheme. Although encoding of monomer B may not be necessary for numerically small libraries where mass redundancy can be avoided, for larger libraries different compounds with equal mass

Figure 5



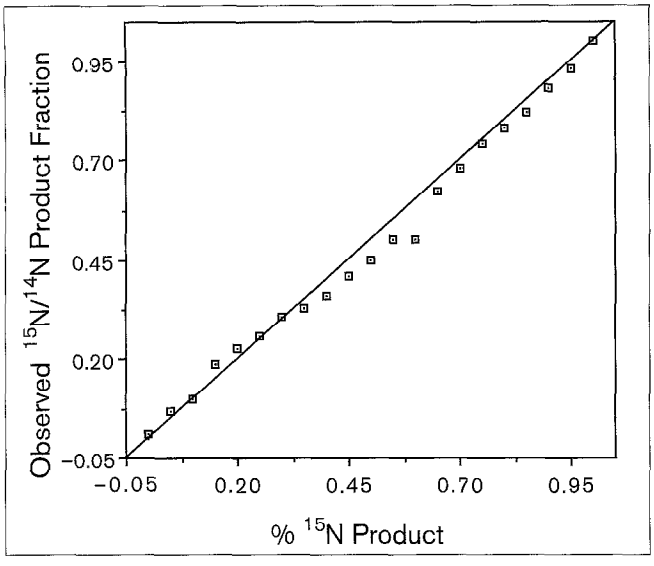
Bar codes can be used to generate 25 distinct mass spectrum patterns. The ratio in peak heights agree closely with those predicted. Where the ratio is not 1:1, as is the case for some of the codes, this is a result of unequal weightings when mixtures of isotopically distinct glycines are combined to form the dipeptide code blocks (actual usage shown inset in the spectrum). In the examples shown, the total encoding information falls over a 5 AMU range. The first seven codes equate to an inverted binary pattern.

will occur and a more elaborate encoding scheme will therefore be required.

Maximizing sensitivity

It has been assumed throughout that the encoded mass blocks were all readily detectable by MS. In general, introduction of a basic center such as an amine, easily accomplished by including lysine along with the code components, would ensure ionization of the code. For maximum sensitivity a work-up prior to submitting the code for MS analysis could involve production of a charged species, for example, reacting a basic nitrogen to a

Figure 6



Ratio codes. Observed fractional ^{15}N composition for the 21 codes resulting from 5 % incremental steps (100 to 0 %) in the ratio of N^{14}- to N^{15}- containing photocleavable linker as determined from the peak areas of the adjacent peaks in the mass spectrum.

quaternary center by alkylation, thus eliminating the ionization step and the need to acidify the sample [12]. This latter procedure could be used to effectively suppress the 'contamination' of the spectrum with non-code peaks, enhancing the ability to analyze or read the code. In cases where a compound library is synthesized from poorly ionizable members, the direct introduction of a ratio code into the molecules by the use of an ^{13}C-, ^{15}N- or ^{18}O-labeled linker or reagent is not effective.

NMR decoding

For ratio encoding schemes based on isotopes in which one of the pair has an odd number of protons in the nucleus, such as ^{1}H, ^{13}C, ^{15}N and ^{19}F, the code can be characterized by NMR. An example in which resin particles were encoded with varying ratios of a ^{13}C and ^{12}C acetic acid is shown in Figure 9. By integrating the signal at the chemical shift for the isotopically varied position against some reference atom environment, a ratio can be determined with an accuracy better than that observed using a mass spectrometer. In this case the naturally occurring ^{13}C present in the methylene groups of the resin construct was used as the reference signal. In principle, combining ratio encoding at several centers, each with a unique chemical shift, would allow the number of codes to be expanded to practical levels for the encoding of larger libraries. NMR techniques offer the potential of assaying a bead library using a direct binding assay [6], followed by decoding of the bead to identify the active compound in a non-destructive manner.

Discussion

Clearly the encoding schemes that have been discussed are most applicable to libraries that have one compound per bead, and therefore careful attention must be paid to match the particular strategy to the bead size and chemical loading used. In general, the larger bead sizes such as those in the 100 μm range (which have loadings of ~300 pmole per bead) have a large enough quantity of the

Figure 7

Combination coding scheme. The top of the figure shows a schematic drawing of the coding construct. The left panel shows examples of the 25 bar codes as illustrated in Figure 5, the center panel the 15 different glycine combinations that are each mixed in ten different ratios (right panel) to generate the ratio portion of the code. Combination of these elements gives a total of 3750 unique codes. As indicated in the schematic, automated reading of the ratio code would be simplified by incorporation of a peak-splitting element as described in the text by, for example, the use of an equimolar mixture of A^0 and A^2.

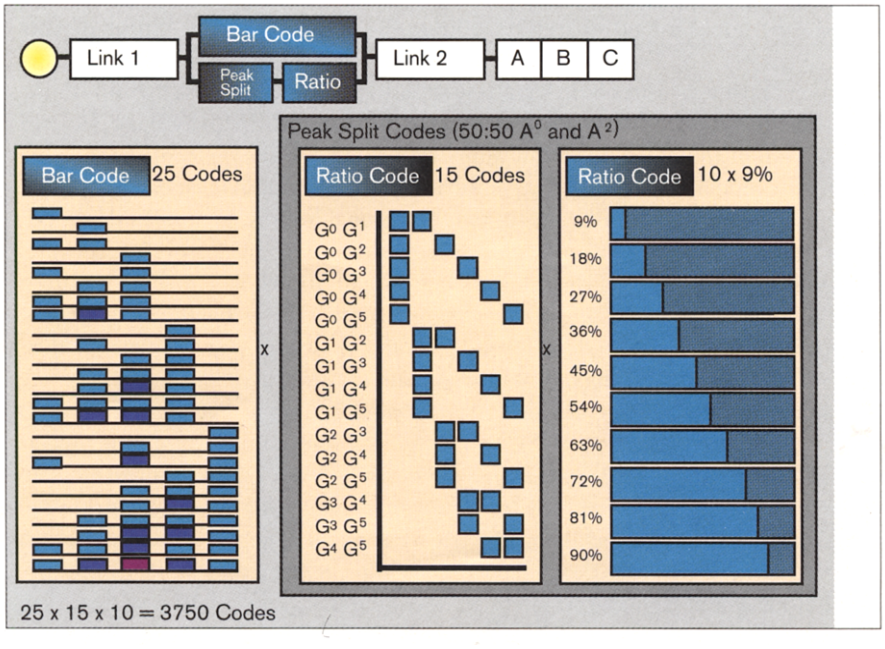
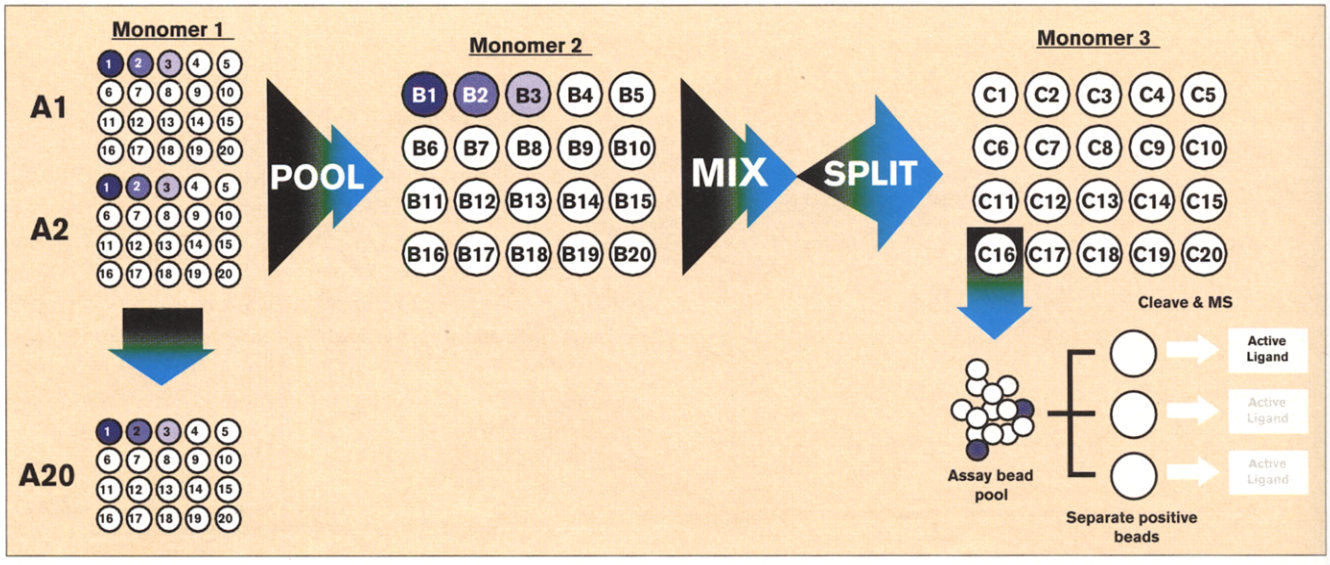


Figure 8



Complete library-coding scheme. Schematic drawing demonstrating an explicit encoding scheme for a 20^3-compound library. Each set of 20 wells (A1 through A20) contains a single, different monomer 1, but each position in the grid is marked with a different code that specifies the identity of monomer 2. After the coupling of monomer 1, all the beads from the wells in a particular position in the grid are pooled into the corresponding well in the B grid for the coupling of monomer two. This is followed by a conventional mix and split step, before coupling monomer 3 and assaying for active compounds followed by code removal to determine identity.

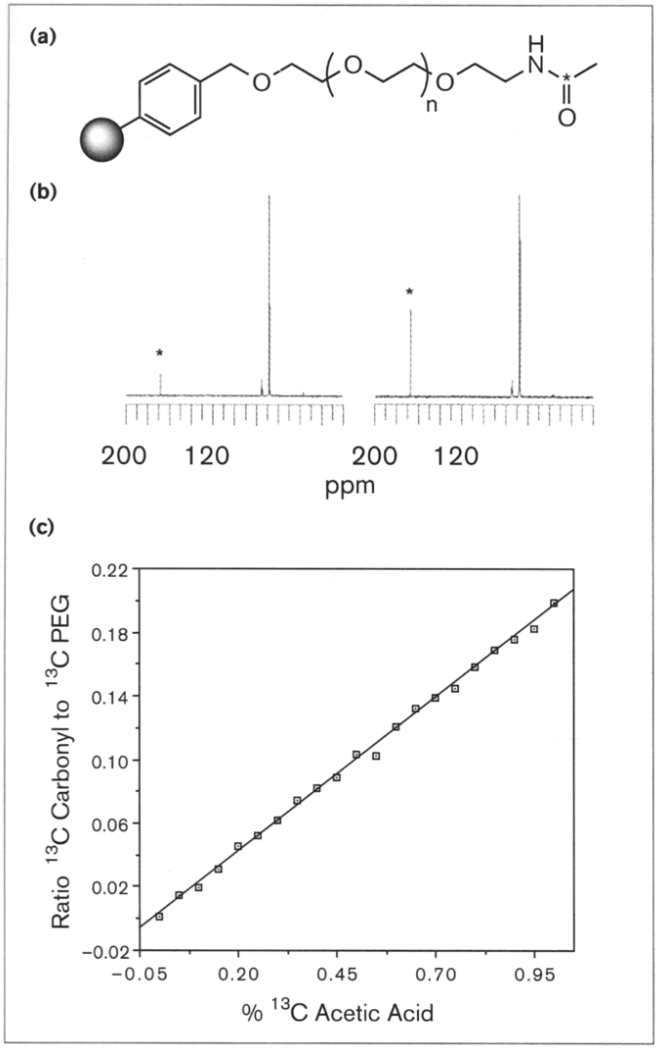
code components to allow the use of the more complex encoding strategies, with the code information distributed in several peaks. With small beads and low loadings, encoding information is restricted to as few peaks as possible to maintain adequate sensitivity in the mass spectrometer. In the future, it may be possible, with some optimization and equipment development, to read the assay result on a bead by bead basis in a fluorescence-activated cell sorter (FACS) and to route beads reading above a threshold for automated decoding.

Another feature of the coding strategies outlined here is that the code is usually superimposed onto the spacer/linker portion of the solid-phase medium, and therefore the initial synthetic investment is very similar to what is already minimally required. With the use of two orthogonal linkers, the synthesized library of compounds can be assayed in solution with the proviso that the encoded bead and the assay remain linked in some suitable manner. Alternatively, active compounds can be determined by a direct binding assay of the bead-bound ligand before reading of the code. Another important feature of this system is that reading of the code in a mass spectrometer only requires a cycle time of a few minutes. It remains to be seen what level of automation can be implemented to allow routine decoding to be carried out using NMR, and with what cycle time. Some recent publications demonstrate that, in principle, it may be possible to achieve single bead decoding in a flow-through context with reading times comparable to those of MS in the near future [13,14].

The structure–activity relationships (SARs) determined thus far have been limited by the lack of an encoding strategy for numerically large bead libraries that allows some approximation to complete decoding of the many active compounds usually present. This problem is compounded by the need to use a significant bead redundancy to allow for sampling variation at each of the resin pool splitting steps during the synthesis procedure (typically the number of beads must be >10 times the number of compounds to be synthesized). Assuming a high fidelity in the assay used, each active compound will show up more than once; this frequency should, on average, be equal to the bead redundancy used. With the more labor-intensive and time-consuming decoding schemes proposed to date, the number of beads eventually decoded is usually only a small percentage of the total, and therefore the potential SAR information inherent in the library is poorly realized. It is anticipated that the mass or isotopic encoding schemes described herein may rectify this situation.

The description above has emphasized the encoding of monomer usage as the chemical synthesis of the library progresses. A further application of such encoding would be to improve the efficiency of searches of reaction space as a first step in developing an appropriate chemistry for subsequent library generation. It is not uncommon to invest significant resources over some extended time to 'optimize' a particular chemistry on solid phase. Using a split and recombine strategy with a suitable encoding

Figure 9



Direct decoding of a ratio code by NMR spectroscopy of beads *in situ*. In **(a)** the position of the acetate used to implement the code is shown, with the ratio carbon atom as *. **(b)** Two examples of the NMR spectra from which ratios were determined by measuring the signal intensity of the level of ^{13}C (marked *) against the signal obtained from the naturally occurring ^{13}C methylene present in the polyethylene spacer of the resin, which was used as a reference. **(c)** A comparison of the ratio of ^{12}C- to ^{13}C-containing acetic acid that was coupled to the observed ratio of the isotopes on the beads.

scheme for the various parameters that define reaction space, such as solvent and catalyst concentration, it should be possible to significantly shorten the time needed to 'translate' solution-phase reactions to solid phase. Using a construct that keeps the code block attached to the compound(s) which are the subject of the chemical investigation, the code determines the applicable combinatorial conditions of synthesis for the observed product. To a first approximation, relative yields could be calculated from the normalized MS intensity, and then correlated with the procedure used to determine the optimal synthetic route for that compound.

Significance

Encoding and decoding strategies are central to combinatorial methods for the synthesis and subsequent analysis of chemical libraries. In all cases, the objective is to explore large parts of chemical space for a desired biological activity, and, in the process, to obtain a comprehensive SAR for the class of compounds synthesized. To date, only a sparse sample of the potential information inherent in numerically large chemical libraries has been obtained due to limitations in the available encoding and decoding schemes. Mass or isotopic encoding of bead libraries has the potential to more fully exploit combinatorial chemical methods for the drug discovery process.

Materials and methods

Reagents

Fmoc-protected amino acids were purchased from either Bachem Bioscience or Novabiochem. Isotopically labeled compounds were purchased from Cambridge Isotope Laboratories and were Fmoc protected as necessary. Fmoc Knorr linker was purchased from Novabiochem. Fmoc-3-(2-nitrophenyl)-3-aminopropionic acid (a photolabile linkage agent) was purchased from Universal Organics. Aminomethyl polystyrene resin (0.81 mmol g^{-1}) was purchased from Advanced ChemTech.

Chemistry

Linkers were attached to the resin manually in reaction shakers. The Knorr linker was attached using 3.5 equivalents of linker (6.12 g, 11.34 mmol) and 3.5 equivalents of PyBop (5.9 g, 11.34 mmol) in dimethylformamide (DMF). The solution was then activated with 10.5 equivalents of diisopropylethylamine (DIEA) (5.93 ml, 34.02 mmol) and the reaction run for 1 h with manual shaking every 5–10 min. The photolinker was attached using 3.5 equivalents of linker (4.89 g, 11.34 mmol) and 3.5 equivalents of PyBop (5.9 g, 11.34 mmol) in DMF. The solution was then activated with 10.5 equivalents of DIEA (5.93 ml, 34.02 mmol), and the reaction run for 1 h with manual shaking every 5–10 min.

Removal of the Fmoc protective group was accomplished using 25 % piperidine in DMF twice for 10 min, followed by thorough washing with DMF. Amino acids were coupled as 0.142 M solutions in N-methylpyrolidinone, using PyBOP/DIEA and *in situ* activation for 1.5 h at room temperature. Boc deprotection was performed with 25 % trifluoroacetic acid in dichloromethane for 30 min. Acetylations were performed with 33 % acetic anhydride in DMF for 30 min. All couplings were verified by ninhydrin color tests.

Cleavage from the Knorr linker was accomplished in 90 % trifluoroacetic acid, 3 % each water, phenol and thioanisole for 1.5 h. Samples were dried under a nitrogen stream, then dissolved in water and lyophilized twice. Photocleavage was performed in 2:1 water:tetrahydrofuran under a UV lamp for 12 h. Samples were lyophilized once before analysis.

Mass spectrometry

Electrospray ionization mass spectra were collected with a Sciex API-III mass spectrometer in the positive mode. Following chemical or photo-cleavage from either bulk resin or from a single bead, the analyte solution was diluted and an aliquot was flow injected into the mass spectrometer at 30 μl min^{-1} using a Gilson 232XL autosampler and a Hewlett Packard HP1090 pump (MeOH:water 1:1 with 1 % acetic acid). The ionspray needle was kept at −5000 V and the orifice potential at −60 V. Mass spectra were generated by averaging the spectra at 50 % of the total ion current (~25 scans for each sample).

NMR spectroscopy

Proton-decoupled ^{13}C spectra were obtained at 75 MHz using a Varian Unity 300 with line broadening of 1.0 Hz and a relaxation delay of 0.1 s. Spectra were obtained in $CDCl_3$ with tetramethylsilane as an internal standard. 21 x 0.1 g lots of Tentagel-NH_2 (0.26 mmol g^{-1}, 100–200 mesh) were coupled with ratios from 0 % to 100 % of 2-^{13}C to 2-^{12}C acetic acid, in 5 % steps. 0.05 g of each resin lot was suspended in a 5 mm NMR tube using $CDCl_3$. Solid phase ^{13}C spectra were collected over 720 transients, using an acquisition time of 1.4 s. Both the labeled ^{13}C spectra were collected over 720 transients, using an acquisition time of 1.4 s. Both the labeled ^{13}C acetic acid carbonyl and the naturally occurring ^{13}C polyethylene glycol peaks were integrated and a ratio of one to the other was calculated.

References

1. Bunin, B.A. & Ellman, J.A. (1992). A general and expedient method for the solid-phase synthesis of 1,4-benzodiazepine derivatives. *J. Am. Chem Soc.* **114**, 10997–10998.
2. Dewitt, S.H., Kiely, J.S., Stankovic, C.J., Schroeder, M.C., Cody, D.M.R. & Pavia, M.R. (1993). 'Diversomers': an approach to nonpeptide, nonoligomeric chemical diversity. *Proc. Natl. Acad. Sci. USA* **90**, 6909–6913.
3. Geysen, H.M., Rodda, S.J. & Mason, T.J. (1986). *A priori* delineation of a peptide which mimics a discontinuous antigenic determinant. *Immunology* **23**, 709–715.
4. Furka, A., Sebestyen, F., Asgedom, M. & Dibo, G. (1991). General method for rapid synthesis of multicomponent peptide mixtures. *Int. J. Pept. Protein Res.* **37**, 487–493.
5. Nicolaou, K.C., Xiao, X.-Y., Parandoosh, Z., Senyei, A. & Nova, M.P. (1995) Radiofrequency encoded combinatorial chemistry. *Angew. Chem. Intl. Ed. Eng.* **34**, 2289–2291.
6. Nikolaiev, V., *et al.*, & Lebl, M. (1993). Peptide encoding for structure determination of nonsequenceable polymers within libraries synthesized and tested on solid-phase supports. *Peptide Res.* **6**, 161–170.
7. Nielsen, J., Brenner, S. & Janda, K.D. (1993) Synthetic methods for the implementation of encoded combinatorial chemistry. *J. Am. Chem. Soc.* **115**, 9812–9813.
8. Brenner, S. & Lerner, R.A. (1992) Encoded combinatorial chemistry. *Proc. Natl. Acad. Sci. USA* **89**, 5381–5383.
9. Nestler, P.H., Bartlett, P.A. & Still, W.C. (1994) A general method for molecular tagging of encoded combinatorial chemistry libraries. *J Org. Chem.* **59**, 4723–4724.
10. Ni, Z.-J., *et al.*, & Gallop, M.A. (1996). Versatile approach to encoding combinatorial organic syntheses using chemically robust secondary amine tags. *J. Med. Chem.* **39**, 1601–1608.
11. Stankova, M., Issakova, O., Sepetov, N.F., Krchnak, V., Lam, K.S. & Lebl, M. (1994) Application of one-bead one-structure approach to identification of nonpeptidic ligands. *Drug Dev. Res.* **33**, 146–156.
12. Wagner, D.S., *et al.*, & Watson, J.T. (1991). Derivatization of peptides to enhance ionization efficiency and control fragmentation during analysis by FAB-CAD-MS/MS. *Biomed. Environ. Mass Spectrom.* **20**, 419–425.
13. Wu, N., Peck, T.L., Webb, A.G., Magin, R.L. & Sweedler, J.V. (1994). ^1H-NMR spectroscopy on the nanoliter scale for static and on-line measurements. *Anal. Chem.* **66**, 3849–3857.
14. Sarkar, S.K., Garigipati, R.S., Adams, J.L. & Keifer, P.A. (1996). An NMR method to identify nondestructively chemical compounds bound to a single solid-phase synthesis bead for combinatorial chemistry application. *J. Am. Chem. Soc.* **118**, 2305–2306.