Fourier Transform Combinatorial Chemistry

Alan W. Schwabacher,*,[1] Yixing Shen, and Christopher W. Johnson

Department of Chemistry, University Wisconsin-Milwaukee Milwaukee, Wisconsin 53211

Received May 3, 1999

Chemical libraries are revolutionizing the discovery of useful compounds, including drugs, catalysts, and other materials.[2] Following the lead of biology, chemists have developed combinatorial approaches to synthesizing these large arrays of related molecules and screening them to identify the fittest members (e.g., the best inhibitor of a given enzyme, or the best catalyst for a given transformation). We report here a general scheme for library preparation that combines the two most valuable, but previously incompatible, features of existing schemes: full parallelism of synthesis and full library evaluation. Split-mix procedures[3] allow efficient preparation of massive libraries on beads of solid support; however, once the desirable components are functionally selected, their identities must be decoded individually.[2b,4] Spatial encoding on surfaces allows evaluation of all library members, although this has to date required elaborate instrumentation.[2c,5]

Our general scheme uses a one-dimensionally organized solid support in place of effectively zero-dimensional beads,[3] or two-dimensional surfaces.[2c,5] A linear solid support ("thread") allows fully parallel synthesis in a straightforward manner. The thread is wrapped around a cylinder in a single spiral layer (Figure 1A−C). Partitioning the cylinder lengthwise into regions defines each space between divisions as a distinct reaction vessel. Coupling a different reagent to each of the regions (one reagent per reaction vessel) creates repeating domains along the thread. Each domain bears a single species, and the identity of each species is specified by its distance from the end of the thread.

Before a second set of reagents can be coupled, the thread must be redivided into regions such that all reagent combinations will be equally represented. This is done by wrapping the thread around a cylinder of an appropriate diameter *different* from the first, followed by division into regions and coupling of the second set of reagents (Figure 1D−F). Repetition of this process allows use of as many steps as desired, each set of reagents on a cylinder of distinct diameter. The identity of each functional group varies periodically as a function of distance along the thread, with each cylinder providing a distinct period for the corresponding portion of the molecule.

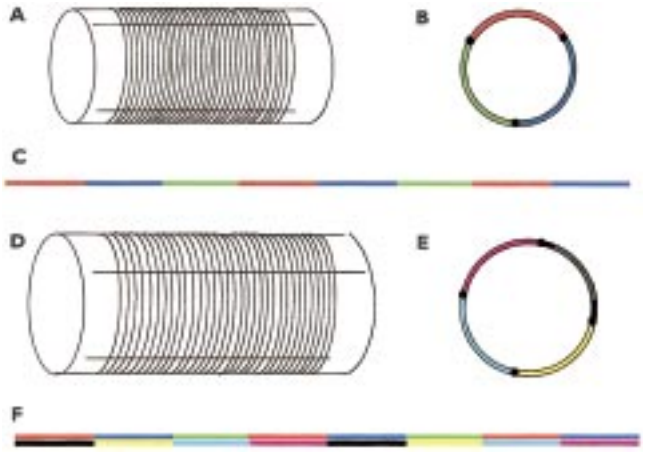

Figure 1. (A) Thread wrapped on a cylinder is partitioned lengthwise into regions. (B) Each region is exposed to a different reactant (red, blue, or green). (C) The solid support now bears repeating domains. (D) Support, wrapped on a cylinder of different size, (E) is redivided and exposed to a different set of reactants (black, yellow, cyan, and magenta). (F) All combinations are formed.

All combinations are equally represented, as long as appropriate cylinder ratios and sufficient thread length are used. Each compound is uniquely specified by its location on the thread, and no compound need appear more than once. This synthesis is analogous to schemes that use a monolithic solid support that is cut into portions at each diversity-generating step to ensure that all combinations are represented with no duplicates.[6] A fundamental distinction is that in our scheme the subdivision takes place in such a way that the support is not physically fragmented, so spatial information on species identity is retained.

We have implemented this strategy using peptide synthesis on a cotton thread, and evaluated binding of a small library of peptides to streptavidin. Peptide library sequences were Ac-X_2-X_1-Pro-Gln-Phe-Ala-Ala-Ala-linker thread, where X_1 and X_2 are specified in Figure 2A. The sequences include the known HPQ streptavidin-binding motif.[7] Standard FMOC peptide synthetic protocols, similar to those previously adapted to a cellulose solid support,[2c] gave satisfactory results. After exposure of the thread library to a fluorescent-labeled streptavidin, the thread was drawn through a focused beam in a fluorescence spectrophotometer (Figure 2B) to sequentially evaluate the binding of each thread-bound peptide (Figure 3A). Solution-phase assays could obviously be used if desired.

The sequence of the one-dimensional array of compounds on the thread is particularly conducive to data analysis and summary. One need store only the activity data in order, since the order contains much information in a computationally useful way. The nature of each functional group cycles along the sequence of compounds on the thread, with a period determined by the circumference of the cylinder used to append that group. If the activity (in our case, affinity for streptavidin) depends to a significant degree upon a particular portion of the molecule, the activity data will be periodic. The period will specify the portion of the molecule important for activity. Fourier transformation of the detector signal[8] provides a library spectrum (Figure 2D) consisting of peaks representing cylinder frequencies and their

(1) Correspondence should be addressed to Dr. Alan W. Schwabacher, awschwab@uwm.edu.

(2) (a) Geysen, H. M.; Rodda, S. J.; Mason, T. J. *Mol. Immunol.* **1986**, *23*, 709−715. (b) Houghton, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84−86. (c) Frank, R. *Tetrahedron* **1992**, *48*, 9217−9232. (d) Bunin, B. A.; Plunkett, M. J.; Ellman, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4708−4712. (e) Thompson, L. A.; Ellman, J. A. *Chem. Rev.* **1996**, *96*, 555−600. (f) Keating, T. A.; Armstrong, R. W. *J. Org. Chem.* **1996**, *61*, 8935−8939. (g) Nefzi, A.; Ostresh, J. M.; Houghten, R. A. *Chem. Rev.* **1997**, *97*, 449−472. (h) Gennari, C.; Nestler, H. P.; Piarulli, U.; Salom, B. *Liebigs Ann./Recueil* **1997**, 637−647. (i) Reddington, E.; Sapienza, A.; Gurau, B.; Viswanathan, R.; Sarangapani, S.; Smotkin, E.; Mallouk, T. E. *Science* **1998**, *280*, 1735−1737.

(3) Furka, A.; Sebestyen, M.; Asgedom, M.; Dibo, G. *Int. J. Pept. Protein Res.* **1991**, *37*, 487−493. Lam, K. S.; Lebl, M.; Krchnak, V. *Chem. Res.* **1997**, *97*, 411−448.

(4) Brenner, S.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5381−5383. Ohlmeyer, M. H. J.; Swanson, R. N.; Dillard, L. W.; Reader, J. C.; Saouline, G.; Kobayashi, R.; Wigler, M.; Still, W. C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 10922−10926. Czarnik, A. W. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12738−12739.

(5) Fodor, S. P. A.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D. *Science* **1991**, *251*, 767−773. Pirrung, M. C. *Chem. Rev.* **1997**, *97*, 473−488.

(6) Stankova, M.; Wade, S.; Lam, K. S.; Lebl, M. *Peptide Res.* **1994**, *7*, 292. Terret, N. K.; Gardner, M.; Gordon, D. W.; Kobylecki, R. J.; Steele, J. *Chem. Eur. J.* **1997**, *3*, 1917−1920.

(7) Devlin, J. J.; Panganiban, L. C.; Devlin, P. E. *Science* **1990**, *249*, 404−406. Schmidt, T. G. M.; Koepke, J.; Frank, R.; Skerra, A. *J. Mol. Biol.* **1996**, *255*, 753−766. Weber, P. C.; Pantoliano, M. W.; Salemme, F. R. *J. Am. Chem. Soc.* **1992**, *114*, 3197−3200.

10.1021/ja991452i CCC: $18.00
Published on Web 09/01/1999

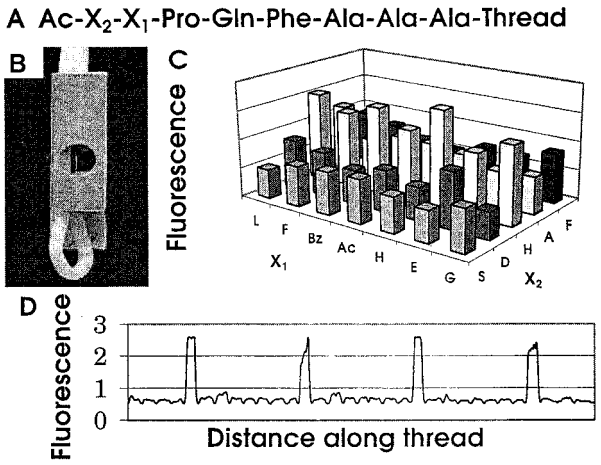

Figure 2. (A) Sequences of the peptides studied. X_1 cycles through His, Ser, Asp, Ala, Phe and X_2 cycles through Leu, Phe, Bz, Ac, His, Glu, Gly in that order. Phe in position X_2 was unacetylated. Standard FMOC peptide synthetic protocols were followed, similar to those previously adapted to a cellulose solid support.[2c] The linker is $NH_2(CH_2CH_2O)_3$-$CH_2CH_2NHCO_2$-cellulose. For couplings of X_1 and X_2 on cylinders, paraffin was applied with a modified hot glue gun to form an impermeable barrier separating the regions. (B) Cell that fits into a spectrometer for fluorescence measurements on thread. Teflon tubes direct the thread library from an audiotape cassette to the cell. Lenses focus exciting light onto thread, and collimate emitted light. (C) Traditional representation of peptide library binding, derived from data in Figure 3A. X_1 = Histidine is best.[7] (D) Measurement of fluorescein-conjugated streptavidin binding to biotin regions on a thread, obtained by driving thread through the cell with an audiotape player. Large peaks correspond to biotin regions, small peaks two samples later correspond to low concentration of biotin; others are negative controls. Black markings delineate samples because of background fluorescence.

harmonics.[9] The magnitude of a given frequency domain signal indicates the significance to activity, averaged over the library, of variation of the molecule in the region appended on the corresponding cylinder. If combinations of groups are required for activity, cycling will be seen with a period corresponding to the product of the cylinder sizes involved for the two groups. Thus trends in the entire library data are immediately apparent from the FT spectrum of the library, regardless of the number of data dimensions.

A "fitness profile" summarizes how well each specific functional group at a given molecular position contributes to activity. This library-averaged trend is easily extracted from the FT spectrum as follows. Only the frequency domain peak corresponding to that molecular region and its harmonics are placed in a small array. Reverse FT to the time domain gives the fitness profile for that position (Figure 3C,D). If two positions are involved, display of their cross-correlation, analogous to Figure 2C, is achieved by diagonally filling a rectangular array.

The design and analysis of combinatorial libraries made by any technique can benefit from these ideas. Our linear organization of library members provides the advantages of exceedingly compact information storage and visualization. A subset of a

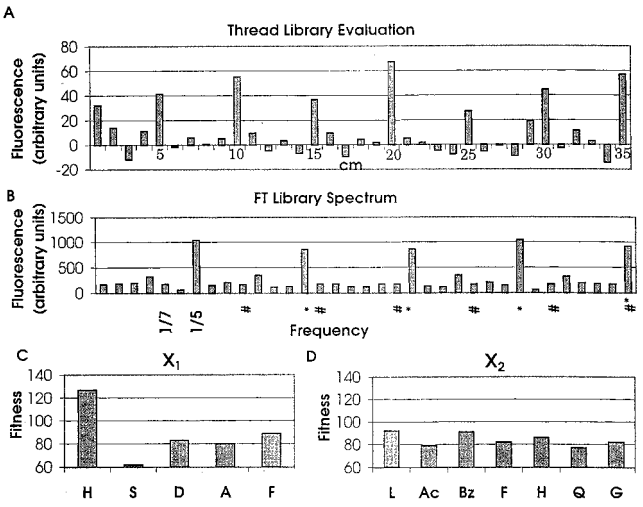

Figure 3. (A) Fluorescence evaluation of streptavidin binding to the thread library. After exposure of the thread library to fluorescein-labeled streptavidin, binding was measured by drawing the thread at a constant rate (by means of an audiotape player) through a focused beam of light in a fluorescence spectrometer. Values are triplicate averages, with 45 units of background fluorescence subtracted. (B) FT spectrum of the library obtained by Fourier transformation of data in (A). Peaks marked 1/5 and 1/7 indicate the significance of groups installed using cylinders of 5 or 7 compounds, respectively. Peaks marked * or # are harmonics of 1/5 or 1/7 peaks, respectively. (C and D) "Fitness profiles" for the X_1 and X_2 positions, respectively. See text.

library will be optimally substitutionally diverse if it can be represented as a contiguous portion of such a linear sequence, since every variable group changes at each step as one proceeds along the sequence of compounds.

The FT spectrum of a library is convenient for identifying general features important for binding, and may also provide a useful measure of the extent to which activity variation in a library is describable by trends. Some validity of trends (independence of variables) is demonstrated by the success of mixture deconvolution schemes,[2a,b] but idiosyncrasies (nonlinear effects) are also well documented.[10]

In summary, we present a new paradigm for preparation and evaluation of diverse collections of molecules. The simple synthetic procedure we present potentially approaches the ideal of fully parallel synthesis of massive libraries with full library analysis. In this era, it is a challenge to make use of the enormous quantity of data collected. Linear data organization with FT analysis stands as a unique approach to the efficient storage and use of the large quantities of library data acquired to assist the development of new drugs, catalysts, and materials.

Acknowledgment. This work was funded by National Science Foundation grant CHE-9726030.

Supporting Information Available: Experimental procedures for preparation of the thread solid support, peptide library preparation, and streptavidin binding assay (/PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA991452I

(8) Press: W. H.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes in C*, 2nd ed.; Cambridge University Press: Cambridge, 1993. Frigo, M.; Johnson, S. G. http://www.fftw.org.

(9) All harmonics are seen because these are repeated spikes, not sine waves.

(10) Liang, R.; Yan, L.; Loebach, J.; Ge, M.; Uozumi, Y.; Sekanina, K.; Horan, N.; Gildersleeve, J.; Thompson, C.; Smith, A.; Biswas, K.; Still, W. C.; Kahne, D. *Science* **1996**, *274*, 1520.