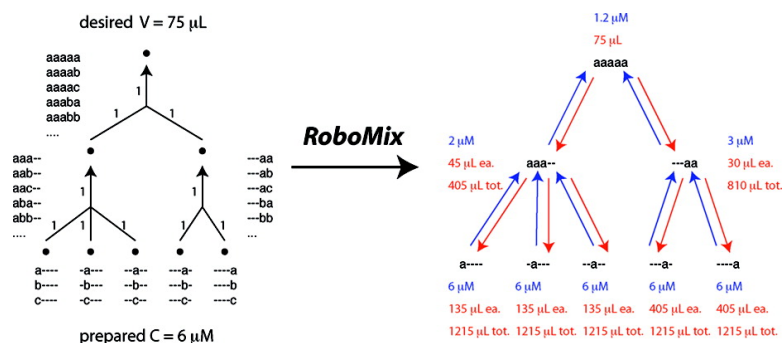


## Robotic Hierarchical Mixing for the Production of Combinatorial Libraries of Proteins and Small Molecules

Larisa V. Avramova, Jairav Desai, Shelley Weaver, Alan M. Friedman, and Chris Bailey-Kellogg

*J. Comb. Chem.*, **2008**, 10 (1), 63-68 • DOI: 10.1021/cc700106e • Publication Date (Web): 12 December 2007

Downloaded from <http://pubs.acs.org> on March 25, 2009



### More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)

# Robotic Hierarchical Mixing for the Production of Combinatorial Libraries of Proteins and Small Molecules

Larisa V. Avramova,<sup>†</sup> Jairav Desai,<sup>‡</sup> Shelley Weaver,<sup>‡</sup> Alan M. Friedman,<sup>†§</sup> and Chris Bailey-Kellogg<sup>\*‡</sup>

*Bindley Bioscience Center, Purdue University, West Lafayette, Indiana 47907, Department of Computer Science, Dartmouth College, Hanover, New Hampshire 03755, and Department of Biological Sciences, Purdue Cancer Center, and Markey Center for Structural Biology, Purdue University, West Lafayette, Indiana 47907*

*Received June 29, 2007*

We present a method to automatically plan a robotic process to mix individual combinations of reactants in individual reaction vessels (vials or wells in a multiwell plate), mixing any number of reactants in any desired stoichiometry, and ordering the mixing steps according to an arbitrarily complex treelike assembly protocol. This process enables the combinatorial generation of complete or partial product libraries in individual reaction vessels from intermediates formed in the presence of different sets of reactants. It can produce either libraries of chimeric genes constructed by ligation of fragments from different parent genes or libraries of chemical compounds constructed by convergent synthesis. Given concentrations of the input reactants and desired amounts or volumes of the products, our algorithm, RoboMix, computes the required reactant volumes and the resulting product concentrations, along with volumes and concentrations for all intermediate combinations. It outputs a sequence of robotic liquid transfer steps that ensures that each combination is correctly mixed even when individualized stoichiometries are employed and with any fractional yield for a product. It can also account for waste in robotic liquid handling and residual volume needed to ensure accurate aspiration. We demonstrate the effectiveness of the method in a test mixing dyes with different UV–vis absorption spectra, verifying the desired combinations spectroscopically.

## 1. Introduction

The ability of robotic liquid-handling systems to accept a series of commands to aspirate and dispense precise volumes into individual reaction vessels (vials or wells in a multiwell plate) makes it possible to carry out reactions between combinations of individually selected and apportioned reactants. The products of a reaction (with or without purification) can then form the intermediates for the assembly of more complex combinations in a hierarchical assembly tree (Figure 1). Such hierarchical assemblies are most valuable when either a limited number of reactants can be brought together in one step or where the reactive groups used would lead to undesired products if all the reactants were simply mixed together (or if the reactants were mixed in an incorrect order). Hierarchical mixing allows the order of the synthetic steps to be rigorously controlled, while robotic control enables the assembly of complex combinatorial libraries in which each product has a defined location. The defined location of individual products facilitates analysis of and adjustment for relative reactivities and provides for initial screening without the need for external coding or the identification of individual products.<sup>1–4</sup>

Once programmed properly, a robotic system can ensure that the correct combinations of reactants are mixed in the correct amounts, including allowance for desired stoichiometries, relative reactivities, and yields of the intermediates. Such robotic production of combinatorial libraries enables new approaches in combinatorial assembly of gene fragments for protein investigation and engineering<sup>5</sup> and in convergent synthesis of combinatorial libraries of druglike molecules and natural product variants.<sup>6</sup>

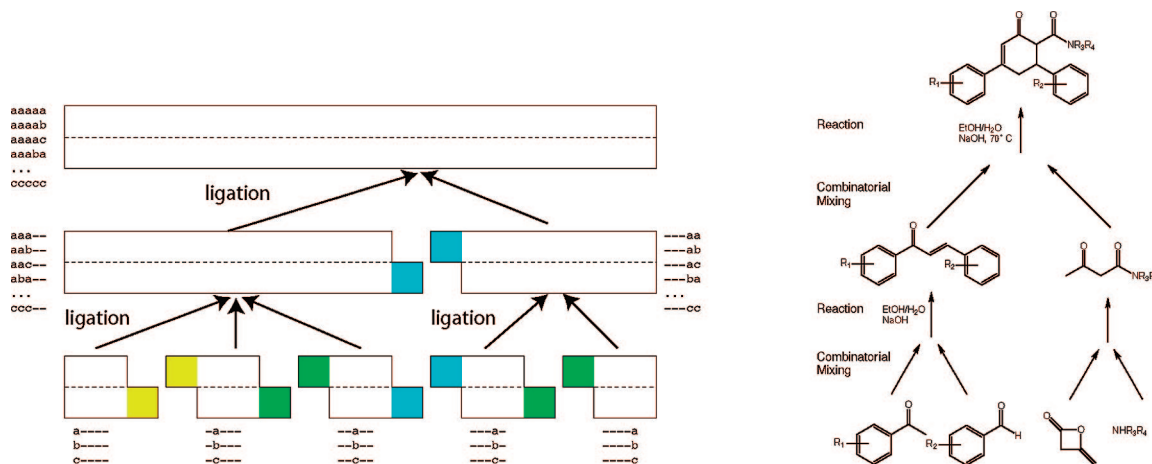
Protein engineering by site-directed recombination<sup>5,7–10</sup> generates a library of hybrids by mixing fragments of multiple homologous parent proteins at defined, intentionally selected breakpoints; the hybrids are subsequently tested for new or improved function. For example, given three parents and desired recombination breakpoints precisely dividing each parent into five gene fragments, we may wish to construct a library of all 3<sup>5</sup> hybrids. Several methods have been developed for manipulating DNA to generate libraries of recombinant genes. In SISDC,<sup>11</sup> restriction enzyme cleavage of genes with previously inserted tags containing the enzyme recognition sequence yields DNA fragments with single-stranded overhangs. Ligation of such fragments from multiple parents yields libraries potentially containing all combinations. SPLISO<sup>5</sup> directly combines individually prepared fragments at computationally planned overhangs. Individual preparation of the fragments allows for greater control of the combinations produced. In either case, to

\* To whom correspondence should be addressed. Phone: 603-646-3385. Fax: 603-646-1672. E-mail: cbk@cs.dartmouth.edu.

<sup>†</sup> Bindley Bioscience Center, Purdue University.

<sup>‡</sup> Department of Computer Science, Dartmouth College.

<sup>§</sup> Department of Biological Sciences, Purdue Cancer Center, and Markey Center for Structural Biology, Purdue University.



**Figure 1.** Combinatorial hierarchical mixing for gene assembly (left) and chemical synthesis (right). A set of variants is available for each reactant, and all products (resulting from each combination of variants) are to be produced under a desired stoichiometry and by the order specified in the tree structure. In the gene assembly example (left), hybrids are constructed by ligating five fragments chosen from three homologous parents. Colored blocks represent 5' nucleotide overhangs to be used in ligating DNA fragments together (upper and lower blocks of the same color are complementary). A combination (including the input reactants) is designated with a string of five letters, one for each reactant: a, b, or c for one of the variants or - for a reactant not included in the combination. Hierarchical assembly can avoid undesired ligations (e.g., between the second fragment and the last fragment) by producing intermediates in which the nucleotides that served as overhangs for lower-level fragment ligation are now internal. In the chemical synthesis example (right), adapted from refs 6 and 12, cyclohexenone derivatives are assembled hierarchically by producing a set of chalcones and a set of acetoacetamides first, before combining them. Here, too, alternative ordering of the reaction steps would yield different products.

ensure that the desired hybrids are constructed effectively (minimizing out-of-order ligation of fragments), the assembly may need to be organized hierarchically (Figure 1, left). In the example, we would first assemble the intermediates from the first three fragments separately from those for the final two fragments, before assembling the complete products. In a hierarchical assembly, the earlier steps hide some of the overhangs within the intermediates, making them unreactive.

Previous site-directed recombination projects have often used large-scale genetic selection to identify hybrids with desired activity.<sup>8,9</sup> However, this approach yields only data on hybrids with positive activity; negative data, which arises from noting which hybrids are missing from the selected pool, can only be evaluated probabilistically, typically under the assumption of unbiased hybrid generation. Alternatively, screening individual hybrids from libraries generated en masse requires substantial oversampling to ensure statistically complete coverage and is still susceptible to biased construction of hybrids. The robotic process described here enables separate construction and evaluation of each hybrid individually, thereby providing comprehensive positive and negative data for stability and activity after screening. Furthermore, knowing the composition of each hybrid from its location also allows a preliminary analysis of all the library members without extensive sequencing.

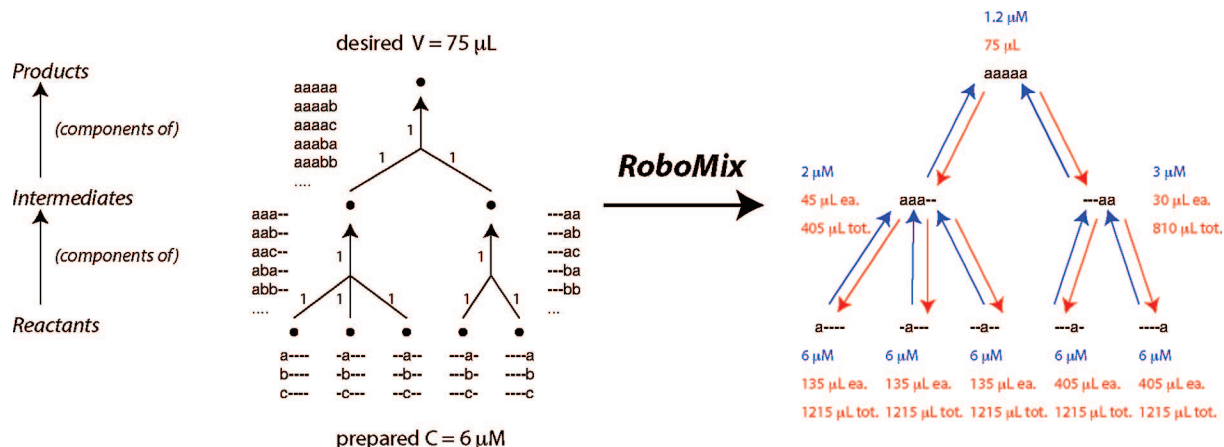
While protein engineering has recognized the value of hierarchical combinatorial assembly, chemical synthesis has historically preferred linear divergent synthesis.<sup>6</sup> Linear synthetic techniques, such as split-and-pool,<sup>13</sup> can generate great diversity but limit the later reactions to those compatible with the products of prior steps. In contrast, convergent synthesis<sup>12,14</sup> allows intermediate products to be made in steps that are incompatible with other intermediates, combining the intermediates later. An example of convergent synthesis from the group at Arqule is given in Figure 1, right.

Convergently synthesized libraries may be particularly valuable when the intermediates resemble fragments of natural products.<sup>14</sup>

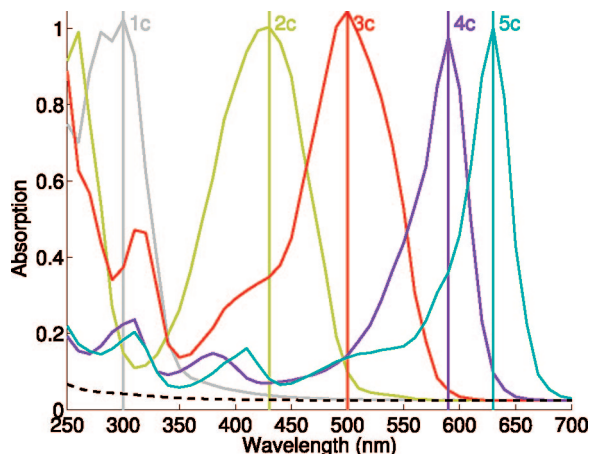
To effectively and automatically carry out such individual combinatorial hierarchical assemblies for proteins or small molecules, this paper presents a method, RoboMix, for computing a control program for a robotic liquid-handling system. Given the specified process, along with the concentrations of the reactants and either the desired volumes of the products (most useful for subsequent molecular biology of assembled gene products) or the desired amounts (most useful for chemical synthesis), RoboMix determines the required volumes of the reactants, the required volumes and resulting concentrations of the intermediates, and a sequence of robotic liquid-transfer steps (aspiration of a defined volume from one vessel, dispensing it to another) that ultimately yield all possible combinatorial products in defined locations. Extensions to our method also permit the synthesis of only a specified subset of the combinatorial products, enhancing the flexibility of library generation and the ability to focus on the most promising products.

## 2. Materials and Methods

**2.1. Computation of Concentrations and Volumes.** We are given sets of variants for the input reactants and wish to produce all possible combinations, each with a different set of choices for the reactants. (Defined subsets of combinations will be produced by exclusion from the default of all possible combinations.) Each combination is produced by following a set of hierarchical mixing steps, indicating the order in which to generate intermediates. We call components any combination of initial reactants and intermediates that will be mixed in a single step. An assembly tree (Figure 2, left) specifies the components and order, along with the input and



**Figure 2.** Computing transfer steps from hierarchical plan via RoboMix. (left) An example assembly tree has three steps: one step produces intermediate combinations from all variants for the first three reactants, combined in equal molar amounts to yield a molar amount of product (small numbers on the edges indicate stoichiometry); another produces intermediates from all combinations of the final two reactants, and the final step yields all products by combining components produced for the two intermediate steps. The program can handle such arbitrary tree structures. Also specified are the available concentrations of reactants (here, 6  $\mu$ M) and desired volumes of products (here, 75  $\mu$ L). (right) The resulting concentrations are first computed bottom-up (blue) by eq 1. Given the desired volume at the root (here 75  $\mu$ L), the required volumes are computed top-down (red) by eqs 3 and 4. The numbers shown assume no loss in liquid handling, no required residual volume, and 100% yield in the reaction; although the planning method can account for all such inefficiencies (see Figure 4 for the same calculation including such inefficiencies).



**Figure 3.** UV-vis absorption spectra for the five input dyes, each at the starting concentration for its c variant. The b variant is this concentration mixed 1:1 with water, while the a variant is water. The dashed line at the bottom is the spectrum of water.

output stoichiometry for the reaction at each node of the tree. Finally, we are given as inputs the concentrations of the reactants and the desired volumes (or amounts) of the products. The output gives a series of robotic transfer steps and all associated volumes and concentrations.

In summary, the computation proceeds recursively (Figure 2, right) following the tree structure: concentrations of components in each combination are first computed working up from the user-specified input concentrations of the reactants (leaves of the tree), and then corresponding volumes are computed working down from the desired volume or amount of the products (root of the tree). The volumes computed throughout the assembly are used for determination of both the transfer steps (including multiple steps of equal volume where the transfer volume is limited) and the size of reaction vessels to be placed on the robot deck for each step.

Consider a specific combination  $y$  (e.g., abacb, using the notation from Figure 1, left). Let  $M(y)$  be the set of

| From       | To        | Volume ( $\mu$ L) | Comment  |
|------------|-----------|-------------------|--|
| plate1 A1  | plate2 A1 | 165.1666667       | a---- $\rightarrow$ aaa--                          |
| plate1 A4  | plate2 A1 | 165.1666667       | --a-- $\rightarrow$ aaa--                          |
| plate1 A7  | plate2 A1 | 165.1666667       | ---ab $\rightarrow$ aaa--                          |
| plate1 A1  | plate2 A2 | 165.1666667       | a---- $\rightarrow$ aab--                          |
| plate1 A4  | plate2 A2 | 165.1666667       | --a-- $\rightarrow$ aab--                          |
| plate1 A8  | plate2 A2 | 165.1666667       | ---b- $\rightarrow$ aab--                          |
| ...        |           |                   |  |
| plate1 A10 | plate2 C4 | 156.8333333       | ---a- $\rightarrow$ ---aa (1 of 3 equal transfers) |
| plate1 A10 | plate2 C4 | 156.8333333       | ---a- $\rightarrow$ ---aa (2 of 3)                 |
| plate1 A10 | plate2 C4 | 156.8333333       | ---a- $\rightarrow$ ---aa (3 of 3)                 |
| plate1 B1  | plate2 C4 | 156.8333333       | ----a $\rightarrow$ ---aa (1 of 3)                 |
| plate1 B1  | plate2 C4 | 156.8333333       | ----a $\rightarrow$ ---aa (2 of 3)                 |
| plate1 B1  | plate2 C4 | 156.8333333       | ----a $\rightarrow$ ---aa (3 of 3)                 |
| plate1 A10 | plate2 C5 | 156.8333333       | ---a- $\rightarrow$ ---ab (1 of 3)                 |
| plate1 A10 | plate2 C5 | 156.8333333       | ---a- $\rightarrow$ ---ab (2 of 3)                 |
| plate1 A10 | plate2 C5 | 156.8333333       | ---a- $\rightarrow$ ---ab (3 of 3)                 |
| plate1 B2  | plate2 C5 | 156.8333333       | ----b $\rightarrow$ ---ab (1 of 3)                 |
| plate1 B2  | plate2 C5 | 156.8333333       | ----b $\rightarrow$ ---ab (2 of 3)                 |
| plate1 B2  | plate2 C5 | 156.8333333       | ----b $\rightarrow$ ---ab (3 of 3)                 |
| ...        |           |                   |  |
| plate2 A1  | plate3 A1 | 45                | aaa-- $\rightarrow$ aaaaa                          |
| plate2 C4  | plate3 A1 | 30                | ---aa $\rightarrow$ aaaaa                          |
| plate2 A1  | plate3 A2 | 45                | aaa-- $\rightarrow$ aaaaab                         |
| plate2 C5  | plate3 A2 | 30                | ---ab $\rightarrow$ aaaaab                         |
| ...        |           |                   |  |

**Figure 4.** Some of the transfer steps planned for the assembly tree of Figure 2, starting with equal concentrations of reactants, yielding 75  $\mu$ L of products, and accounting for 10% loss and 50  $\mu$ L residual volumes. The maximum transfer volume was set to 200  $\mu$ L, so all transfers in the second group were broken into multiple equal steps.

components mixed to produce  $y$  at a node in the assembly tree (e.g.,  $M(\text{abacb}) = \{\text{aba--}, \text{---cb}\}$  and  $M(\text{aba--}) = \{\text{a---}, \text{-b---}, \text{--a--}\}$ ). Assume that we know the desired volume  $V_y$  or desired molar amount (volume times concentration)  $V_y C_y$  of combination  $y$  and the available concentrations  $C_x, x \in M(y)$  of its components. (The recursive computation guarantees that these values have previously been calculated.) Given this information, we can then compute the resulting concentration  $C_y$  of  $y$ , and the required volumes  $V_{x \rightarrow y}$  contributed to  $y$  from each component  $x \in M(y)$ .

The user specifies the ideal stoichiometry of the reaction (ignoring for now any need for molar excess of certain components or less than 100% yield) as integer coefficients  $\alpha_y$  for the output and a set of integer coefficients  $\alpha_{x \rightarrow y}$  for

the inputs  $x \in M(y)$  at each node. Then output concentrations at each node can be computed from input concentrations by

$$\frac{1}{C_y} = \frac{1}{\alpha_{y \in M(y)}} \sum_{x \in M(y)} \frac{1}{\alpha_{x \rightarrow y} C_x} \quad (1)$$

In the special case where equal molar amounts of reactants mix to produce the same molar amount of a single product (with 100% yield), eq 1 can be simplified to

$$\frac{1}{C_y} = \sum_{x \in M(y)} \frac{1}{C_x} \quad (2)$$

By generalizing the  $\alpha$  coefficients used to express stoichiometry to nonintegers, we can also account for the need to add a molar excess of one or more of the components to ensure complete reaction (increasing  $\alpha_{x \rightarrow y}$ ), as well as account for reaction yields of less than 100% (reducing  $\alpha_y$ ).

After computing concentrations bottom-up for the whole tree, we compute volumes top-down based on the relationships between the amount of the output and the amounts of the inputs under the specified stoichiometry

$$\forall x \in M(y): V_{x \rightarrow y} = \frac{V_y C_y}{\alpha_y \alpha_{x \rightarrow y} C_x} \quad (3)$$

We employ eq 3 to compute  $V_{x \rightarrow y}$  given the concentration of  $x$  (input or computed bottom-up), the resulting concentration of  $y$  (computed bottom-up), and the desired volume of  $y$  (input or computed top-down). Note that at the root of the assembly tree, this equation naturally allows the user to give either the desired volume or the desired molar amount  $V_y C_y$ .

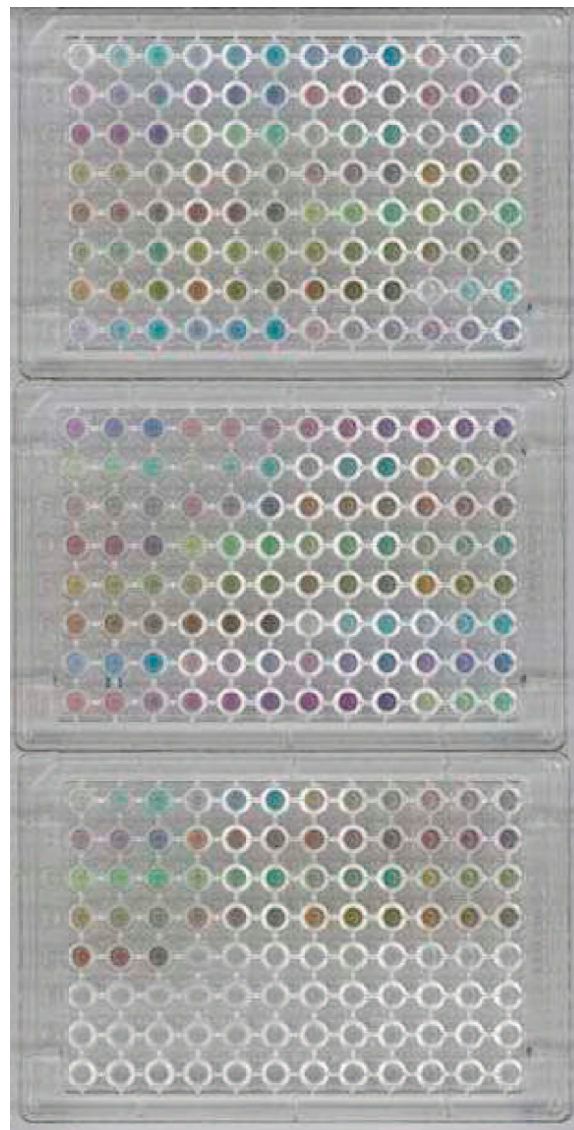
Each component participates in a number of combinations for its parent in the tree (e.g., aaa-- is in aaaaa, aaaab, aaaac, etc.). The total volume required for each component  $x$  is thus calculated by summing  $V_{x \rightarrow y}$  over all such combinations  $y$  for which  $x$  is a component, that is

$$V_x = \sum_{y: x \in M(y)} \frac{V_y C_y}{\alpha_y \alpha_{x \rightarrow y} C_x} \quad (4)$$

Like all liquid handling, robotic liquid handling is imperfect. Losses in each transfer typically result from liquid remaining with the tip after dispensing. The tip height calibration is also imperfect, requiring that some volume be left in the vessel to avoid the incorporation of air bubbles during aspiration with associated error in volume transferred. If we assume that a constant amount  $\lambda_1$  is left in the bottom of the vessel and a constant fraction  $\lambda_2$  of volume is lost in each transfer, then we can compute required volumes using a scaled, offset version of eq. 4

$$V_x = \lambda_1 + (1 + \lambda_2) \sum_{y: x \in M(y)} \frac{V_y C_y}{\alpha_y \alpha_{x \rightarrow y} C_x} \quad (5)$$

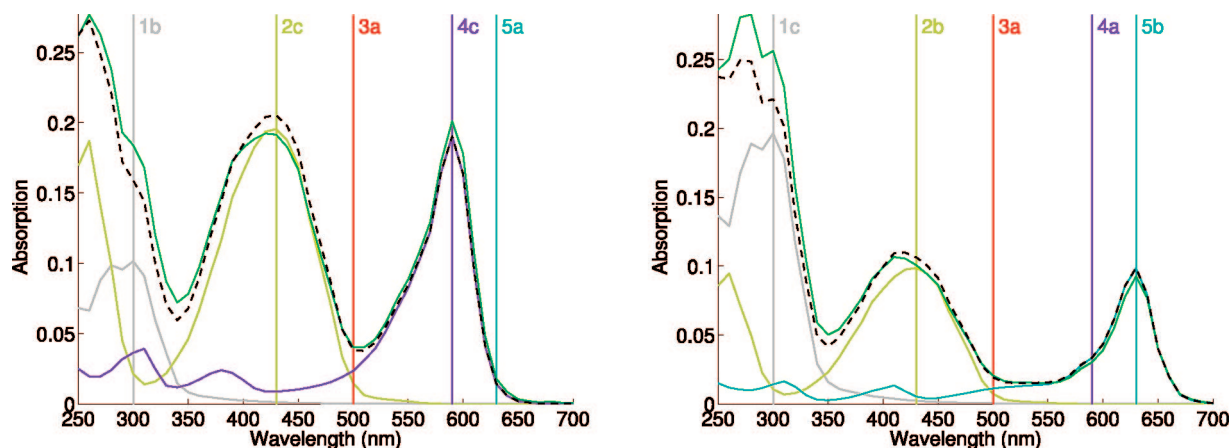
**2.2. Generation of the Control Program.** By default, RoboMix generates a control program for construction of a complete library (one that includes all products). However, it supports the specification of combinations to be excluded, so as to construct a partial library focused on the most useful products. Specific products (e.g., abcba) can be directly excluded. Furthermore, intermediates (e.g., aaa--) can also be excluded, resulting in exclusion of all products derived from them (e.g., aaaaa, aaaab, aaaac, etc.). Concentrations



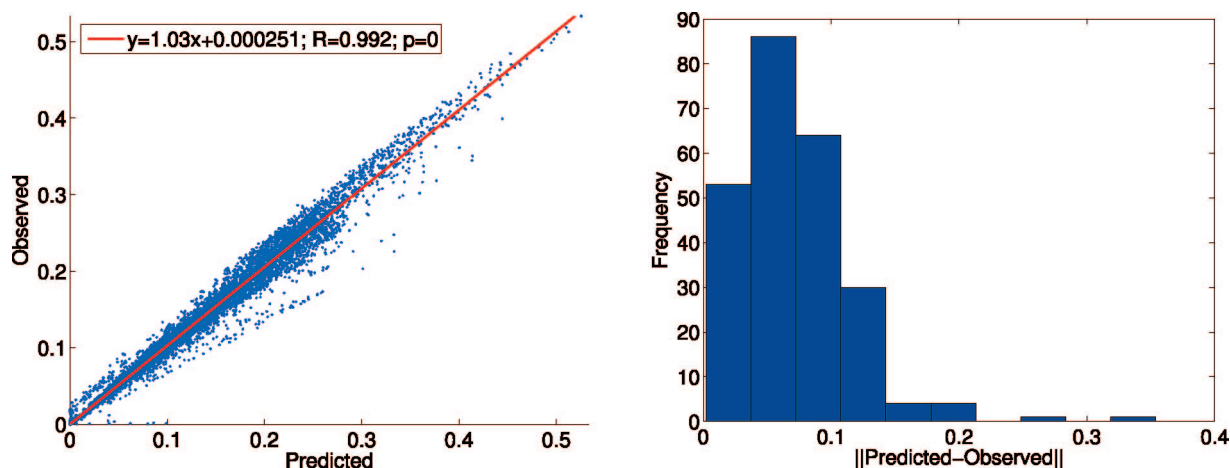
**Figure 5.** Plates with each well holding one product (combination of the three variants for each of the five reactants). The combinations are arranged left-right and top-bottom in order of ternary numbers, so that aaaaa is in well A1 of plate 1, aaaab is in well A2 of plate 1, ..., and ccccc is in the last well employed (E3) in plate 3. Smooth gradations of color visually indicate the progression of dye concentrations and absorption maxima throughout the set of mixtures.

and volumes are calculated to account for only the included intermediates and products.

After computing volumes and concentrations, RoboMix assigns vessels (plate locations) to each reactant, intermediate, and product, and outputs a linearized list of transfer steps with multiple equal volume transfers used to transfer volumes exceeding a specified maximum (e.g., a maximum tip volume). To facilitate common reaction steps and avoid optimization by internal robotic software, the list is partitioned by level in the assembly tree. Although the steps need to be segregated by level, the steps within each level can be done in parallel or in any order. After all the transfers for one level are completed, the common reagents can then be added and incubation conditions set by commands in the main robot control protocol. Since the assembly is organized by level, purification steps can be inserted, most conveniently



**Figure 6.** Water-subtracted observed spectrum (green solid line) vs predicted spectrum (black dashed line) and individual reactant contributions for two sample products: bcaca (left) and cbaab (right).



**Figure 7.** Prediction error over all products, as correlation between predicted and observed individual absorption datapoints (left) and histogram of Euclidean distances between predicted and observed spectra treated as vectors (right).

returning the purified intermediates in the same volume. If volumes/concentrations change after purification, then the program can be rerun with the new values.

**2.3. Implementation.** RoboMix has been implemented in platform-independent Python code. It produces both a detailed summary for the experimenters, as well as a comma-separated file of liquid transfer steps, which is suitable for multiple robotic platforms. The software can be freely obtained for academic use by request from the authors. A demonstration version with a graphical user interface is available at <http://www.cs.dartmouth.edu/~cbk/robomix/>.

**2.4. Test of Hierarchical Mixing.** Hierarchical mixing was tested using five different water-soluble and miscible dyes with distinct absorption spectra as easily measurable surrogate reactants: (1) imidazole (Sigma: CAS 288-32-4) with a peak at 300 nm, (2) a commercial yellow food dye (Kroger: FD&C Yellow #5, Tartrazine, E-102, CAS 1934-21-0) at 430 nm, (3) a commercial red food dye (Kroger: FD&C Red #40: Allura Red, E-120, CAS 25956-17-6) at 500 nm, (4) bromphenol blue (sodium salt, Mallinckrodt: CAS 62625-28-9) at 590 nm, and (5) a commercial blue food dye (Kroger: FD&C Blue #1 Brilliant Blue FCF, E-133, CAS 3844-45-9) at 630 nm. For each dye “reactant”, we prepared solutions of different concentrations to represent the different variants, such that the peak absorption for variant a was 0

(i.e., only water), for b it was 0.5, and for c, 1.0. Figure 3 presents the measured spectra for variant c of each of the reactants.

Hierarchical mixing was conducted on a Biomek FX (Beckman Coulter, Fullerton, CA) running Windows NT Biomek FX system software, version 2.5. A comma-separated file output by RoboMix provided the liquid-transfer steps for a protocol on the Biomek FX. Aspiration speed was set to 70  $\mu\text{L/s}$ , and dispensing speed was set to 70  $\mu\text{L/s}$  with sample mixing after dispensing. The starting materials were placed in deep 96-well blocks; intermediate mixtures were dispensed into deep 96-well blocks, while product mixtures were dispensed into UV-transparent half-area 96-well plates (Corning 3679). The mixing operations were completed in 2 h. After the samples were mixed, the product trays were placed in a Molecular Devices SpectroMax 384 P+ UV-vis plate reader running SoftMax Pro, version 4.8. The spectrum was read from 250 to 700 at 10 nm intervals without volume compensation.

Raw data output by SoftMax Pro were input into Matlab (version 7, MathWorks, Natick, MA). The absorbance spectrum from an equivalent well of water was subtracted from both the reactant and the product spectra. Predicted spectra were computed by taking linear combinations of the variant c spectra. Correlation between predicted and observed

absorption was calculated by treating the absorbance at each 10 nm as an independent datapoint. The total distance between predicted and observed spectra was calculated as the Euclidean distance between 46-element vectors containing every 10 nm datapoint as the elements.

### 3. Results and Discussion

To evaluate our method for combinatorial hierarchical mixing, we developed a plan for the tree in Figure 2, with five reactants (numbered 1–5), each with three different variants (a, b, or c). Each reactant was represented by a different water-soluble, miscible dye, with each variant a different concentration.

We ran our planning algorithm to determine a mixing process that would yield 75  $\mu\text{L}$  of each product, assuming equal concentrations of reactants and accounting for loss by setting  $\lambda_1 = 50 \mu\text{L}$  (residual volume) and  $\lambda_2 = 0.1$  (10% pipetting loss) in eq 5. A plan was computed practically instantaneously on a Macintosh G4 Powerbook. RoboMix determined that a starting volume of  $\sim 1685 \mu\text{L}$  for each choice of the first three reactants and  $\sim 1603 \mu\text{L}$  for each choice of the final two reactants is required. It then mixes  $\sim 165 \mu\text{L}$  of each choice of the first three reactants or  $\sim 471 \mu\text{L}$  of each choice of the last two. Finally, 45  $\mu\text{L}$  of each choice of the first intermediate is mixed with 30  $\mu\text{L}$  of each choice of the second to yield the product mixtures. We programmed a Biomek FX to conduct the planned list of liquid transfer steps; Figure 4 shows a subset of the list. The result was a set of  $3^5$  products arrayed in 96-well plates, displaying a range of colors reflecting the relative amounts of each dye (Figure 5).

By planning volumes for equal concentrations of the reactants and equal molar stoichiometry (eq 2), we expect the spectrum of a product to be an average of the spectra of the reactants; for example, bcaca should be the average of the b--- spectrum, the -c--- spectrum, and so forth. Figure 6 shows two examples of such combinations. We subtracted a background water spectrum from the reactant and product spectra. We then took the appropriate combinations of the water-subtracted reactant spectra and compared them against the corresponding water-subtracted product spectra.

We compared the predicted and observed spectra for all products. Figure 7, left shows the correlation between predicted and observed absorption from 250 to 700 nm, taken every 10 nm. With a slope of 1.03, an intercept of essentially 0, and a correlation coefficient  $R$  of 0.992 at a significance level exceeding machine precision, the predicted spectra are clearly very accurately reproduced. Finally, when the entire

spectrum of each product is considered, there is only a very small Euclidean distance between the predicted and observed spectra (Figure 7, right). The only significant outliers among 243 combinations are in products accaa (predicted-observed distance of  $\sim 0.35$ ) and bcbba ( $\sim 0.27$ ). These appear to be the result of a systematic error in liquid handling, since the associated transfer steps calculated by RoboMix were manually verified to be correct. We note that this type of analysis could also be used to calibrate errors in the liquid handling process. As a beginning to this analysis, we note that the combination of transfer and measurement error over the set of products is an average of 3.6% with an average absolute error of 8.9%. Overall, our results suggest that 621 robotic transfer steps were planned correctly and executed accurately.

**Acknowledgment.** We thank an anonymous reviewer for the suggestion to extend our method to produce only subsets of a complete combinatorial library. We gratefully acknowledge support for this work from the following sources: L.A., development support from the Bindley Bioscience Center; S.W., Dartmouth Women in Science Project (WISP); A.M.F. and C.B.K., a grant from NSF SEIII (IIS-0502801).

### References and Notes

- (1) Brenner, S.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5381–5383.
- (2) Smith, J. M.; Gard, J.; Cummings, W.; Kanizsai, A.; Krchňák, V. *J. Comb. Chem.* **1999**, *1*, 368–370.
- (3) Gartner, Z. J.; Tse, B. N.; Doyon, J. B.; Snyder, T. M.; Liu, D. R. *Science* **2004**, *305*, 1601–1605.
- (4) Halpin, D. R.; Harbury, P. B. *PLoS Biol.* **2004**, *2*, 1022–1030.
- (5) Saftalov, L.; Smith, P. A.; Friedman, A. M.; Bailey-Kellogg, C. *Proteins* **2006**, *64*, 629–642.
- (6) Beeler, A. B.; Schaus, S. E.; Porco, J. A., Jr. *Curr. Opin. Chem. Biol.* **2005**, *9*, 277–284.
- (7) Voigt, C. A.; Martinez, C.; Wang, Z. G.; Mayo, S. L.; Arnold, F. H. *Nat. Struct. Biol.* **2002**, *9*, 553–558.
- (8) Meyer, M. M.; Silberg, J. J.; Voigt, C. A.; Endelman, J. B.; Mayo, S. L.; Wang, Z. G.; Arnold, F. H. *Protein Sci.* **2003**, *12*, 1686–1693.
- (9) Otey, C. R.; Landwehr, M.; Endelman, J. B.; Hiraga, K.; Bloom, J. D.; Arnold, F. H. *PLoS Biol.* **2006**, *4*, e112.
- (10) Ye, X.; Friedman, A. M.; Bailey-Kellogg, C. *J. Comput. Biol.* **2007**, *14*, 777–790.
- (11) Hiraga, K.; Arnold, F. H. *J. Mol. Biol.* **2003**, *330*, 287–296.
- (12) Powers, D. G.; Casebier, D. S.; Fokas, D.; Ryan, W. J.; Troth, J. R.; Coffen, D. L. *Tetrahedron* **1998**, *54*, 4085–96.
- (13) Frank, R.; Heikens, W.; Heisterberg-Moutsis, G.; Blöcker, H. *Nucleic Acids Res.* **1983**, *11*, 4365–77.
- (14) Chen, C.; Li, X.; Lo, M. M.; Schreiber, S. L. *Angew. Chem., Int. Ed. Engl.* **2005**, *44*, 2249–52.

CC700106E