CHEMISTRY

Snap deconvolution: An informatics approach to high-throughput discovery of catalytic reactions

Konstantin Troshin[1,2] and John F. Hartwig[1,2]*

We present an approach to multidimensional high-throughput discovery of catalytic coupling reactions that integrates molecular design with automated analysis and interpretation of mass spectral data. We simultaneously assessed the reactivity of three pools of compounds that shared the same functional groups (halides, boronic acids, alkenes, and alkynes, among other groups) but carried inactive substituents having specifically designed differences in masses. The substituents were chosen such that the products from any class of reaction in multiple reaction sets would have unique differences in masses, thus allowing simultaneous identification of the products of all transformations in a set of reactants. In this way, we easily distinguished the products of new reactions from noise and known couplings. Using this method, we discovered an alkyne hydroallylation and a nickel-catalyzed variant of alkyne diarylation.

High-throughput experimentation (HTE) is one of the essential tools in drug discovery (*1*), but the potential of these methods to influence chemical reaction discovery has been limited (*2*). The most common application of HTE in reaction development is for rapid assessment of the effect of reaction parameters on yield or selectivity (*3*–*10*). In this context, HTE methods have been applied to search for conditions to improve known reactions, including one recent example of making them compatible with reactions on nanomoles of material for improving synthetic routes to druglike molecules (*11*). HTE also has been used recently to identify conditions for the late-stage functionalization of complex molecules, including a photoredox-based functionalization of biologically active heterocycles (*12*) and peptide-catalyzed site-selective modification of natural products (*13*).

The application of HTE to discover unknown classes of reactions has been more limited and often relies on customized use of analytical techniques such as colorimetry (*14*, *15*), microscopy (*16*), fluorescence (*17*, *18*), label-assisted matrix-assisted laser desorption/ionization–time of flight spectroscopy (MALDI-TOF) (*19*), self-assembled monolayer/MALDI mass spectrometry (SAMDI) (*20*), microfluidic reactors (*21*), immunoassays (*22*), and DNA-templated methods (*23*). Despite the value of these analytical techniques, their application requires one or more of the reactants to contain certain functional groups or markers, such as conjugated esters, halides (*14*, *15*), bifunctional linkers (*16*), fluorescence dyes, complex mass tags,

or DNA strains (*17*–*23*). These requirements limit the scope of reactants that can be used in this type of experimentation, thereby limiting the breadth of reactions that can be discovered. Strategies that do not require the introduction of functional tags into the reactants impose the least restriction on the scope of reactants and reaction conditions (*2*). The most straightforward of such strategies involves identification of products by mass spectrometry (MS).

MacMillan and co-workers reported a strategy by which potential reactants are mixed pairwise and subjected to various catalytic conditions and products are identified by gas chromatography–mass spectrometry (GC-MS) (*24*). Our group reported a different approach by which a mixture of 17 substrates with masses in a relatively narrow range were subjected to a catalyst and potential products were identified by GC-MS. Peaks with corresponding masses approximately double those of the cluster of masses of the reactants were then analyzed to determine the possible reaction products (*25*).

Conducting experiments with multiple prospective reactants in the same solution allows simultaneous evaluation of a large number of potential reactions, including multicomponent transformations that could not be observed in pairwise mixtures. However, the final mixtures are more challenging to analyze than those containing just two reactants. To assign products that were not readily identified from the mass spectrum, a deconvolution strategy was developed that involved several rounds of splitting the set of reactants into subsets and subjecting the mixtures of these subsets to the catalytic conditions to identify the one containing the reactants that formed the unidentified product. By this original deconvolution strategy, it was prohibitively time-consuming to identify the large number of peaks in the GC-MS

trace that could result from previously unreported reactions. Thus, a low-yielding but unknown process would likely be overlooked (*2*). For this reason, a method to identify reaction products in an automated fashion without the need for subsequent experiments is needed to make the HTE search protocol more practical and wide-ranging.

Here, we present such an automated method that we call "snap deconvolution" (SD). This method involves running three sets of reactions in parallel with mixtures of substrates containing the same functional groups, but different inert substituents, to vary their relative masses. The products are then identified by the unique differences in masses of the new components in the sets formed from substrates containing the same functional groups.

Formulation of the deconvolution strategy

Our strategy to identify products readily from mixtures of reactants is outlined in Fig. 1. Three sets of reactants (called the α, β, and γ sets) were purchased or prepared; each of the three sets contained the same array of functional groups, but each trio of reactants with the same functional groups contained different inert substituents having distinct masses (Fig. 1). For example, the α, β, and γ sets contained a terminal alkyne that was respectively appended to a decyl, nonyl, and undecyl group. For each of the three sets, a stock solution containing all reactants from that set was then added to the vials in a separate 96-well plate, each of which contained one metal precursor and one ligand. After the designated elapsed reaction time, the mixtures were analyzed by GC-MS.

For the SD strategy to be effective, the substituents and their masses were required to adhere to the following three criteria: (i) The combination of the differences in masses between the substituents in the β and α ($\Delta M_{\beta\alpha}$) set and between those in the γ and α sets ($\Delta M_{\gamma\alpha}$) (Fig. 1) must be distinct for each functional group. In addition, the combinations of mass differences corresponding to all possible binary and ternary reactions must be distinct. (ii) The substituents should exert the least possible influence on the reactivities of the corresponding functional groups. Thus, different masses were created in most cases by placing H, F, and D in the remote positions of the aromatic rings or varying numbers of CH_2 groups in aliphatic compounds. (iii) The reactants with these substituents should be commercially available or readily prepared.

A combination of functional groups that results in a product in the α set should then lead to analogous products in β and γ sets with a unique set of mass differences. Thus, the mass differences between the three products would identify the reactants that had been coupled. Even reactions that involve the incorporation of solvents, ligands, or other common fragments could be analyzed by this method because the additional component would be the same in each set, and, thus, the differences in masses would not be affected. The identity of the additional component would be clear from the difference between the mass of the product and the sum of

[1]Department of Chemistry, University of California, Berkeley, CA 94720, USA. [2]Chemical Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA.
*Corresponding author. Email: jhartwig@berkeley.edu

the masses of the components. An analogous scenario would reveal components of the reactants that might be lost during the reaction.

Design of substituents

Although just two sets of reactants are required to implement SD, our initial studies showed that experiments conducted with two sets of reactants are prone to false positives because the difference in masses of molecular ions for two peaks randomly selected from a GC-MS chromatogram of a complex mixture often matched that from potential reactions of two pairs of reactants in the two sets. This potential for false positives is drastically reduced by using more than two sets of reactants, because the potential that combinations of multiple differences in masses match those of random peaks in a chromatogram is much lower. However, creating a large number of sets of reactants would involve significant synthetic effort. As a compromise, we followed an experimental design in which three sets of reactants with the same functional groups, but distinct masses, were used. To ensure that substituents selected for the three sets of reactants adhered to the criteria outlined above and that the differences in masses of products formed from the α, β, and γ sets containing these substituents were unique, we developed a suite of MS Excel macros (26). The design of these macros is described in the supplementary text. This automated method was crucial to the implementation of our SD because it becomes prohibitively difficult to select a full set of substituents for a group of twelve or more reactants, as demonstrated in figs. S19 and S20.

To assess our expectation that the mass-tuning substituents would not affect the reactivity of the functional groups, we investigated how the yields (27) of three established reactions (Suzuki coupling, Heck reaction, and hydroarylation of alkynes) under four different catalytic conditions varied between the α, β, and γ sets. Figure S4 shows that the three product yields from the appropriate two reactants of the α, β, and γ sets alone with the catalysts were similar to each other. Furthermore, the product yields from components of the α, β, and γ sets were similar to each other when the reactions were conducted with a mixture of eight reactants (1 to 8) comprising the four needed for the established reactions and four additional ones added to model a mixture. Although the ligand effects on yields in some cases when the reactions were conducted with the two reactants alone were different from those when they were conducted in the mixture, the identities of the products were the same. Thus, these results validate our assertion that the substituents do not strongly influence the types of catalytic reactions the functional groups undergo and are appropriate for our design of the SD protocol.

Automated SD

To avoid laborious and prohibitively time-consuming manual analysis of complex chromatographic data (fig. S7), we developed MS Excel macros that analyze the raw GC-MS chromatograms and
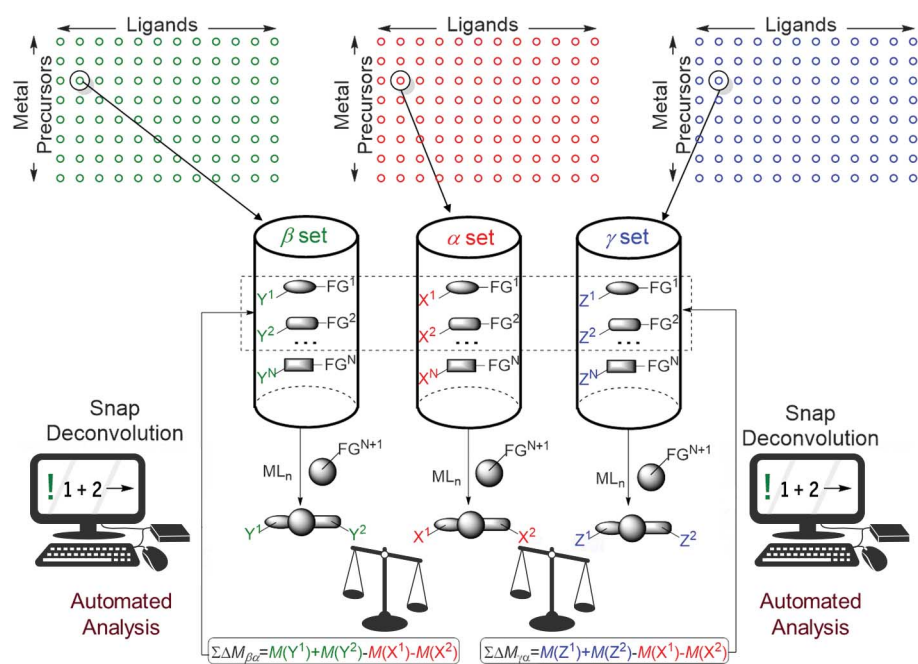

Fig. 1. The principle of SD. Schematic representation of experimental design and analysis.
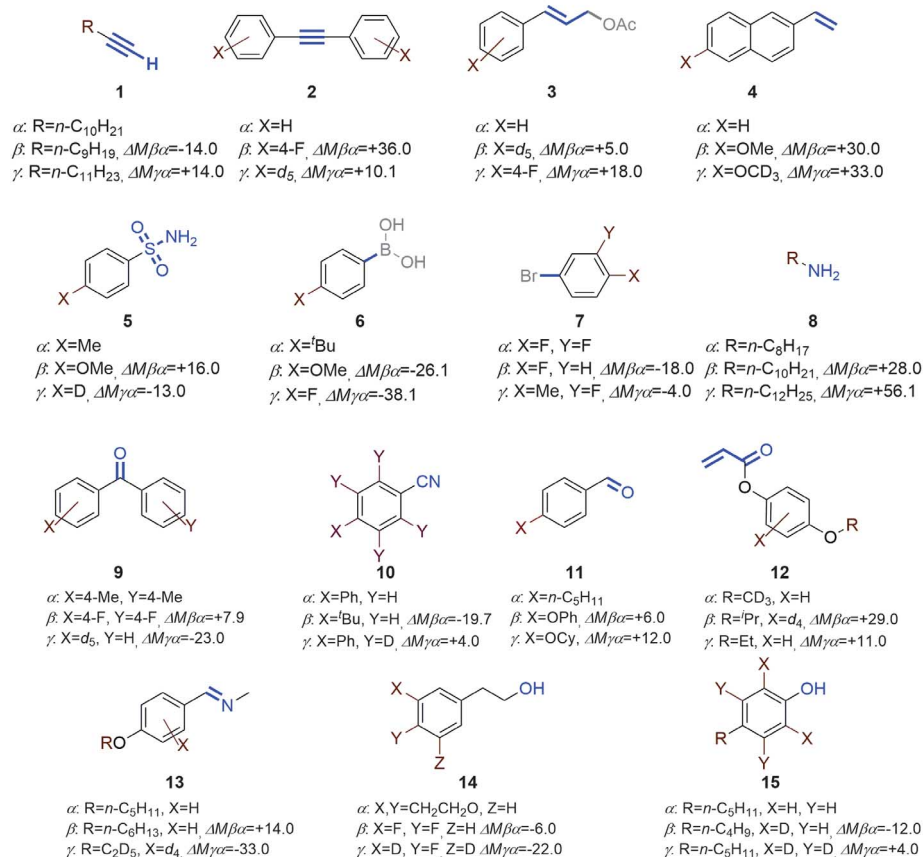

Fig. 2. Functional groups and substitution patterns. The functional groups expected to be absent in the final products are shown in gray.
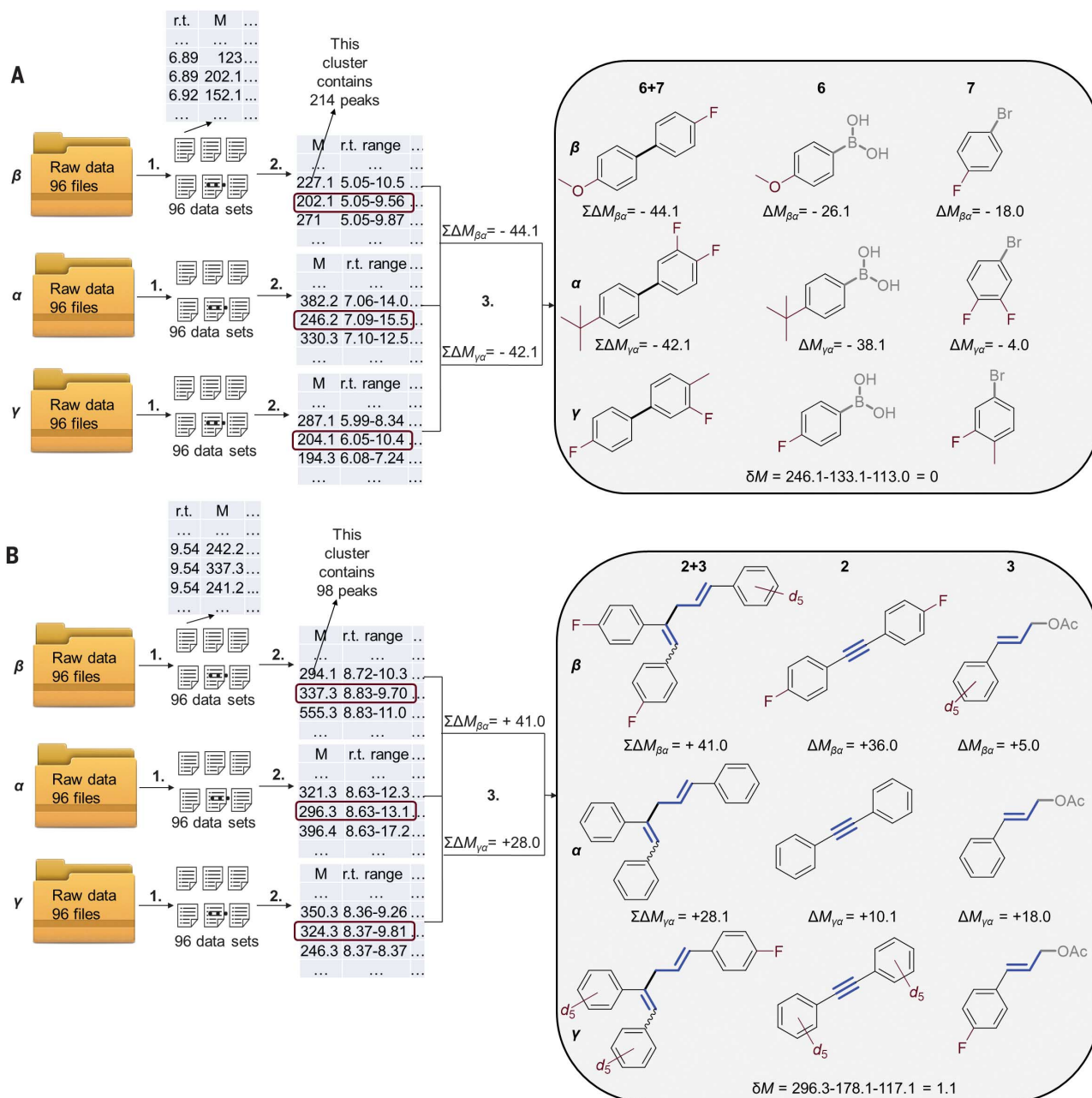
Fig. 3. Schematic representation of our SD. (**A**) Analysis that revealed the product of a Suzuki coupling. (**B**) Hydroallylation of alkynes revealed by the SD macros.

identify the products. The procedures for product identification were as follows: (i) Each chromatogram generated by the GC-MS software was converted to a set of data containing the retention time, molecular masses of the molecular ion and its fragments, and the absolute and relative (with respect to an internal standard) ion counts for the molecular ion. (ii) The resulting data from reactions in the same 96-well plate were then compared to each other, and molecular ions with the same masses formed from reactants in one set (e.g., the α set) were combined into clusters. Each cluster was labeled by the lowest and the highest retention time of molecular ions in this cluster. Clustering the data reduces the process-

ing time required by the module that performs the SD, as described below. (iii) The clusters containing the masses of products from the α set were then compared with those from the β and γ sets, and reactants that would generate those products were deduced. A specific set of reactants was assigned to the combination of three clusters (one from the α, one from the β, and one from the γ set) if they met two criteria. First, the intervals between the ranges of retention times must be less than a certain threshold (in this work, a threshold of 4 min was used) because products of similar structure with the substituents in this study should elute with similar retention times. Second, the value of the differ-

ences in masses of molecular ions must correspond to unique combinations of two or three reactants with the same functional groups in each set. For each pair or trio of reactants assigned to have undergone a reaction, a difference in mass between the experimentally observed molecular masses and the sum of masses of all the tentative reactants (δM) was calculated. If nonzero, this value corresponds to the mass of nontagged components (such as water) that would be introduced or lost during the reaction. To simplify the analysis, the masses of reactants likely to lose a functional group (shown in gray in Fig. 2) were introduced into the macro without these functional groups. For instance, the mass of 4-bromo-1,2-difluorobenzene
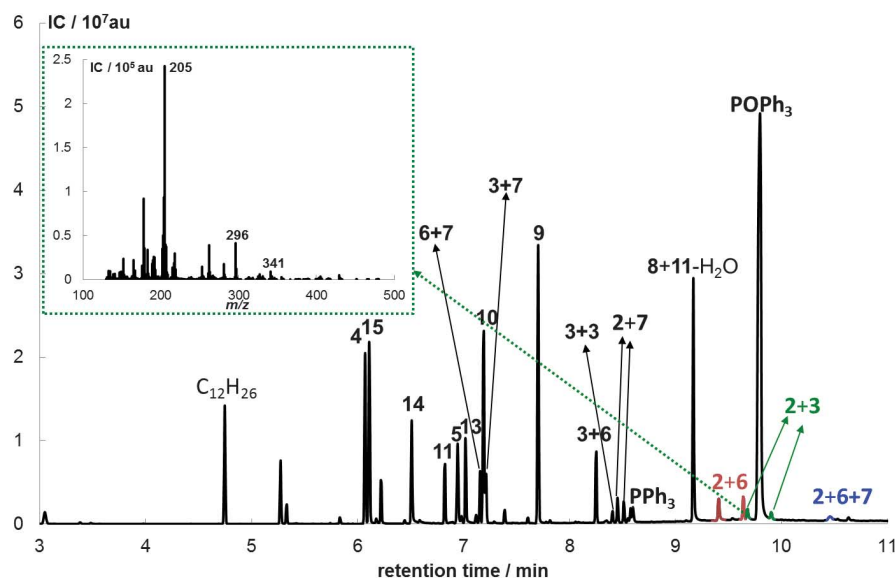
Fig. 4. Chromatogram acquired after heating a mixture of reactants 1 to 15α for 18 hours at 100°C in the presence of Ni(COD)₂ and PPh₃. (**Inset**) The MS trace for the peak of the product **2+3α** eluting at 9.68 min. The peaks for the isomers of the product **2+6α** are highlighted in red. The peaks for the isomers of the product **2+3α** are highlighted in green. The peak for the product **2+6+7α** is highlighted in blue.
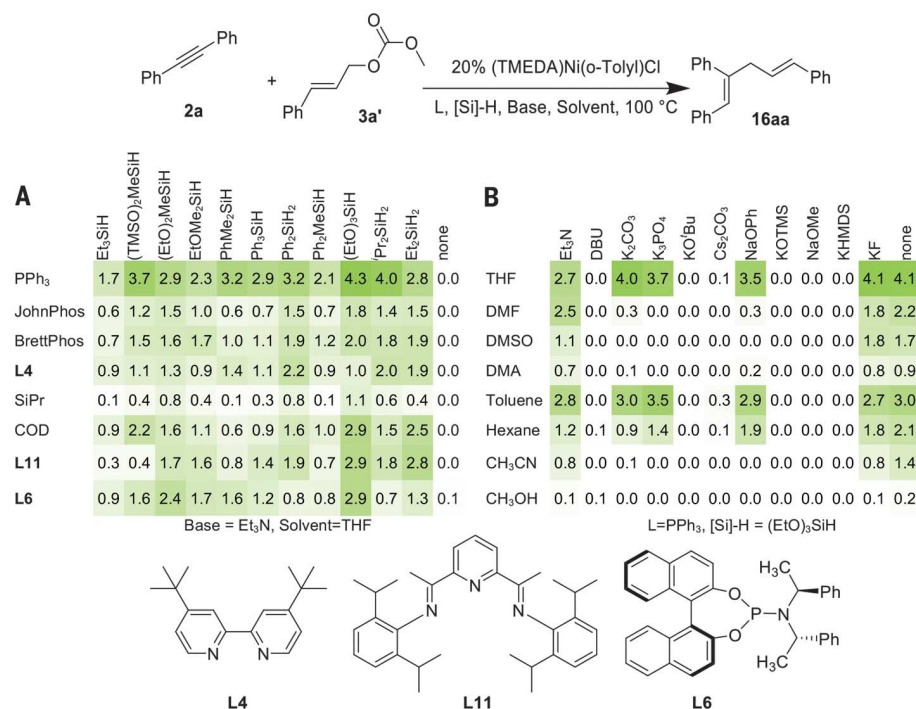

Fig. 5. Optimization of nickel-catalyzed alkyne hydroallylation. The effects of variations in (**A**) ligand and silane and (**B**) base and solvent on the relative ion counts of the molecular ion of the product (with respect to hexadecane).

or δM = –1, then the reaction occurred between tagged substrates with no gain or loss of components besides a hydride or a proton). (v) The selected reactions were then submitted to a detailed analysis by which the peaks within clusters were compared individually to identify which metal-ligand combinations enabled these reactions. If the metal-ligand combinations that resulted in a certain reaction in the α, β, and γ sets did not match, such a reaction was considered to be a false positive. If the difference in retention time of the tentative products in two sets exceeded a certain margin (4 min), then the proposed reaction was also considered to be a false positive. (vi) The detailed results were then analyzed with help of other MS Excel macros that present the collected data graphically or in tabular form (examples of such graphs are provided in figs. S8 and S9).

Steps (i) to (iii) of our SD are presented in Fig. 3A for the identification of the product from a Suzuki coupling (**6**+**7**), which was one of the positive controls of our system. During step (i) of the analysis, three groups of 96 sets of masses, intensities, and retention times (one list per reaction conducted in a 96-well plate) were generated. These sets of data were then combined into three collections of clusters of combined masses and retention times (525 clusters for the α set, 438 clusters for the β set, and 399 clusters for the γ set in this particular case). In step (iii) of the analysis, the cluster corresponding to molecular ions of 246.2 amu from the α set (336 peaks) was compared to that corresponding to a molecular ion of 202.1 amu from the β set (214 peaks) and that corresponding to a molecular ion of 204.1 amu from the γ set (373 peaks). This comparison resulted in values of $\Sigma\Delta M_{\beta\alpha}$ = 202.1 – 246.2 = –44.1 and $\Sigma\Delta M_{\gamma\alpha}$ = 204.1 – 246.2 = –42.1. If all peaks from the clusters were compared independently of each other, step (iii) would have been repeated 336 × 214 × 373 = 2.7 × 10⁷ times to identify the product of the Suzuki coupling alone. Instead, one evaluation was required to assign the combination of reactants **6** and **7** to these three clusters. This simple calculation demonstrates the importance of clustering the masses.

The experimental values of the sum of the differences in masses of the substituents ($\Sigma\Delta M$) matched those calculated for the combination of the arylboronic acids **6** and aryl bromides **7** ($\Sigma\Delta M_{\beta\alpha(calc)}$ = –26.1 – 18.0 = –44.1, $\Sigma\Delta M_{\gamma\alpha(calc)}$ = –38.1 – 4.0= –42.1). The difference between the mass of the experimentally observed product and the sum of masses of the reactants (δM) was 0 [both Br and B(OH)₂ were excluded from the calculation of the masses of the corresponding reactants, as noted above], indicating that no other reactants were involved in the reaction. In other words the product formed from the Suzuki coupling of an aryl halide with a boronic acid.

Application to reaction discovery

Having developed the SD system, we sought to assess its capabilities to identify a series of established reactions as positive controls (including the Suzuki coupling just used to explain the SD method) and to identify new reactions. Our initial

(**7α**) was entered as 113.0203 and not as 191.9836 because, in the vast majority of transition-metal–catalyzed reactions, the bromine would not be present in the products, and the mass of [4-(tert-butyl) phenyl]boronic acid (**6α**) was entered as 133.1017,

not as 178.1165, because the boronate group would likely be lost during the reaction. (iv) The most relevant transformations were then selected from the vast number of potential reactions, based on the values of δM (for instance, if δM = 0, δM = 1,

A

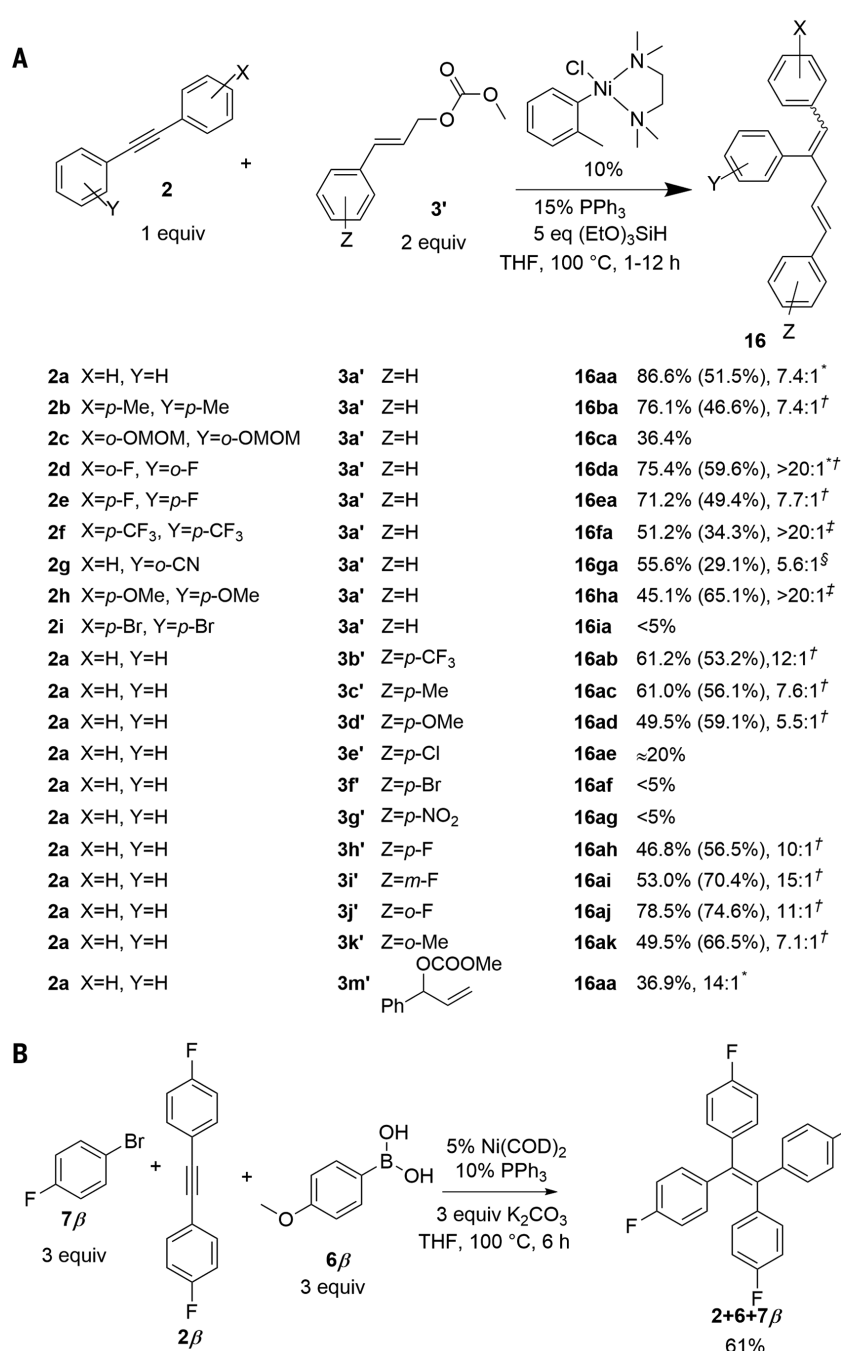|		
2a X=H, Y=H	**3a'** Z=H	**16aa** 86.6% (51.5%), 7.4:1*
2b X=p-Me, Y=p-Me	**3a'** Z=H	**16ba** 76.1% (46.6%), 7.4:1†
2c X=o-OMOM, Y=o-OMOM	**3a'** Z=H	**16ca** 36.4%
2d X=p-F, Y=o-F	**3a'** Z=H	**16da** 75.4% (59.6%), >20:1*†
2e X=p-F, Y=p-F	**3a'** Z=H	**16ea** 71.2% (49.4%), 7.7:1†
2f X=p-CF₃, Y=p-CF₃	**3a'** Z=H	**16fa** 51.2% (34.3%), >20:1‡
2g X=H, Y=o-CN	**3a'** Z=H	**16ga** 55.6% (29.1%), 5.6:1§
2h X=p-OMe, Y=p-OMe	**3a'** Z=H	**16ha** 45.1% (65.1%), >20:1‡
2i X=p-Br, Y=p-Br	**3a'** Z=H	**16ia** <5%
2a X=H, Y=H	**3b'** Z=p-CF₃	**16ab** 61.2% (53.2%),12:1†
2a X=H, Y=H	**3c'** Z=p-Me	**16ac** 61.0% (56.1%), 7.6:1†
2a X=H, Y=H	**3d'** Z=p-OMe	**16ad** 49.5% (59.1%), 5.5:1†
2a X=H, Y=H	**3e'** Z=p-Cl	**16ae** ≈20%
2a X=H, Y=H	**3f'** Z=p-Br	**16af** <5%
2a X=H, Y=H	**3g'** Z=p-NO₂	**16ag** <5%
2a X=H, Y=H	**3h'** Z=p-F	**16ah** 46.8% (56.5%), 10:1†
2a X=H, Y=H	**3i'** Z=m-F	**16ai** 53.0% (70.4%), 15:1†
2a X=H, Y=H	**3j'** Z=o-F	**16aj** 78.5% (74.6%), 11:1†
2a X=H, Y=H	**3k'** Z=o-Me	**16ak** 49.5% (66.5%), 7.1:1†
2a X=H, Y=H	**3m'** (OCOOMe, Ph)	**16aa** 36.9%, 14:1*

B

Fig. 6. Reactions discovered by SD. (A) Scope of hydroallylation of alkynes under conditions improved by additional high-throughput methods. The yields were determined by proton nuclear magnetic resonance (¹H NMR) spectroscopy. The yields shown in parentheses are isolated yields. The yield of **16ae** was estimated by GC using the response factor of **16aa**. The yield of the reaction between **2a** and **3m'** was determined by GC. *Z:E ratio based on GC data. †Z:E ratio based on ¹H NMR data for the product isolated by preparative high-performance liquid chromatography (HPLC). ‡Z:E ratio based on ¹H NMR data for the product isolated by column chromatography. §Z:E ratio based on ¹H NMR data for the crude mixture. (B) Nickel-catalyzed diarylation of alkynes. The yield was determined by GC.

studies focused on reactions catalyzed by complexes of earth-abundant metals. Eight first-row or early transition-metal complexes that could serve as catalyst precursors and 11 ligands (fig. S6) were included in a 96-well format that contained tetrahydro-furan (THF) solutions of 15 substrates in each well (Fig. 2, **1** to **15**). Three sets of these 96-well plates with the substrates of the α, β, and γ sets were assembled, sealed, and heated for 18 hours at 100°C.

The products from reactants included to undergo several known reactions were identified by our macros. These included products of Ni-catalyzed Suzuki (28) (**6**+**7**, Fig. 3A) and Heck (29) (**4**+**7**) couplings, hydroarylation of alkynes with boronic acids (**2**+**6**) or aryl bromides (**2**+**7**) (25), and allylation of boronic acids (**3**+**6**) (30) or aryl bromides (**3**+**7**) (31). In addition to these processes, the macro identified the product from a three-component reaction between aryl bromides, alkynes, and aryl boronic acids (**2**+**6**+**7**), which was previously reported with Pd (32, 33) but not with nickel. Finally, the products from an unknown process, the hydroallylation of internal alkynes (Fig. 3B, **2**+**3**), were identified. These products formed from a process catalyzed by several combinations of (COD)₂Ni and dative ligands (figs. S8 and S9 show the graphs generated by our macros for these reactions).

The importance of automated analysis for identification of these two previously unreported processes is illustrated by the data shown in Fig. 4 and figs. S10 to S14. Figure 4 shows the GC-MS chromatogram obtained after heating the mixture of reactants from the α set with a nickel catalyst. Several products were identified immediately by the SD method. As expected, many of these products result from known reactions. Others result from unknown processes. Some of the products could be identified readily from a single set of reactants from their masses alone; indeed, the intensities of the peaks shown in red are substantial, and two isomers of the vinylarene **2**+**6α** resulting from the combination of alkyne **2α** and boronic acid **6α** were readily identified from their strong molecular ion peaks in the mass spectrum (fig. S10A) and had been identified during our previous study on reactions within mixtures of reactants (25). However, the intensity of the peaks from the product **2**+**3α** (green on Fig. 4) and **2**+**6**+**7α** (blue on Fig. 4) are low in the chromatogram and even overlap those of other components in the mixture. Moreover, the inset shows that the molecular ion for **2**+**3α** is not clear in the mass spectrum. Nevertheless, the SD system showed that these peaks correspond to products of two unreported reactions, one involving two reactants and one involving three reactants in the mixture, and it revealed the identity of the products and reactants from the same experiment. Clearly the identification of a series of products formed from two and three components of the mixture by a manual analysis of analogous GC-MS traces of multiple 96-well plates containing a variety of metal precursors and ligands would be unfeasible.

To assess the validity of the assignment of the product of the new two-component coupling reaction, diphenylacetylene (**2a**) and cinnamyl acetate (**3α**) were subjected to the reaction conditions in the presence of Ni(COD)₂ in THF at 100°C. The GC-MS trace of the crude mixture contained a peak with the mass of the molecular ion (296.2 amu) corresponding to the molecular formula of the product and the same retention time as the peak of the tentatively assigned product for the α set. The mass of this product corresponds to that from

a formal hydroallylation of the alkyne, with Ni(0) serving as the reducing equivalent. To render the reaction catalytic and to provide a source of the required hydride, the reaction was run with triethylsilane. In this case, the product formed in a higher 20% yield. After evaluating various parameters by conventional HTE (3) (Fig. 5 and fig. S18) and individual experiments (see the supplementary text for details), we found that the hydroallylation of diphenyl acetylene conducted with [(TMEDA)Ni(o-tolyl)Cl] (34, 35) as the catalyst precursor, methyl carbonate as the leaving group, triethoxysilane as the reductant, THF as the solvent, and 100°C as the temperature gave the desired product in 87% yield by GC and 52% isolated yield. Additional experiments (table S28) assessing the effects of the types of catalytic reactions the functional groups undergo on its yield resulted in the conditions shown in Fig. 6A.

The scope of the hydroallylation of alkynes was assessed under the reaction conditions that we developed. As depicted in Fig. 6A, moderately electron-withdrawing and moderately electron-donating groups are tolerated in both the alkyne (2) and the allylic carbonate (3′). Strongly electron-donating (such as OMe or OMOM) or electron-withdrawing (such as CF₃ or CN) substituents in either 2 or 3′ led to a slight decrease of the reaction yields, whereas the nitro group present in (E)-methyl [3-(4-nitrophenyl)allyl] carbonate (3g′) completely suppressed the formation of 16 in the reaction with 2a. Substrates containing aryl bromides (2i, 3f′) also did not undergo the allylation process, and the substrate containing an aryl chloride (3e′) reacted in lower yield, presumably because of the known reactions of Ni(0) species with these functional groups (36).

A mixture of Z and E isomers (configuration of the double bond in the stilbene part of the molecule) was observed in most cases, but the Z isomer was always formed in larger amounts with a Z:E ratio of ≈7. When branched methyl (1-phenylallyl) carbonate (3m′) was used instead of 3a′ in reaction with 2a, the same product 16aa was formed, albeit in substantially lower yield. This formation of the same product from 3a′ and 3m′ indicates the intermediacy of a π-allyl nickel complex.

We could find no reports of nickel-catalyzed hydroallylation of alkynes. The closest reactions are a copper-catalyzed hydroallylation of trifluoromethyl-substituted alkynes (the CF₃ group was crucial for reactivity) with allylboronic acids (37), rather than allyl carbonates, Cu-catalyzed hydroallylation of vinylarenes (38), and alkynes (39), as well as allylation of alkynes via four-electron reductive coupling with α,β-unsaturated carbonyl compounds (40). The conditions described in (38) and (40) yielded only trace amounts of product 16 when conducted with the alkyne 2a and diethyl (1-phenylallyl) phosphate (41) or cinnamaldehyde (42), respectively, whereas the conditions described in (39) provided only 34% of 16aa using 2a and diethyl (1-phenylallyl) phosphate (43).

To assess the validity of our assignment of the product from three components in the reaction mixture, we conducted the reaction between the three reactants assigned by the SD to form the product from diarylation of an alkyne. Indeed, the reaction between diarylacetylene 2β, aryl bromide 7β, and aryl boronic acid 6β in the presence of 5% Ni(COD)₂, 10% of PPh₃, and K₂CO₃ as a base for 6 hours at 100°C gave the alkyne diarylation product 2+6+7β in 60% yield (by GC) without optimization (Fig. 6B). This result shows that the automated method for reaction discovery enables the identification of three-component processes, and the identification of this process and the hydroallylation reaction shows that products identified in small quantities in the mixture can signal processes that occur in good yield when run separately.

Outlook

The SD approach we report that combines chemical design with informatics enables about 75,000 possible reactions to be run and assayed for product identification within a few days using only a sealed 96-well plate, a GC-MS instrument (44), and an analytical suite of MS Excel macros. The methods and programs developed in this work can be easily extended. In our current investigations, we focused on reactions occurring between components containing two or three of the designed substituents. Our approach, however, is not limited to reactions between components containing different substituents. It is possible to use this experimental design to search for products resulting from additional components in the mixture. For example, products formed in a system containing an untagged reagent, such as CO, CO₂, a source of fluorine, or a source of a fluoroalkyl group, could be identified by selecting for products in each of the three sets with added masses corresponding to that of the additional reagent. (45–66)

REFERENCES AND NOTES

1. S. K. Doyle, M. S. Pop, H. L. Evans, A. N. Koehler, Curr. Opin. Chem. Biol. 30, 28–36 (2016).
2. K. D. Collins, T. Gensch, F. Glorius, Nat. Chem. 6, 859–871 (2014).
3. A. E. Metz, S. Berritt, S. D. Dreher, M. C. Kozlowski, Org. Lett. 14, 760–763 (2012).
4. S. M. Preshlock et al., J. Am. Chem. Soc. 135, 7572–7582 (2013).
5. A. Bellomo et al., Angew. Chem. Int. Ed. 51, 6912–6915 (2012).
6. J. R. Schmink, A. Bellomo, S. Berritt, Aldrichim Acta 46, 71–80 (2013).
7. S. D. Dreher, P. G. Dormer, D. L. Sandrock, G. A. Molander, J. Am. Chem. Soc. 130, 9257–9259 (2008).
8. G. A. Molander, S. R. Wisniewski, J. Am. Chem. Soc. 134, 16856–16868 (2012).
9. G. A. Molander, L. N. Cavalcanti, C. García-García, J. Org. Chem. 78, 6427–6439 (2013).
10. H. H. Jo, C.-Y. Lin, E. V. Anslyn, Acc. Chem. Res. 47, 2212–2221 (2014).
11. A. Buitrago Santanilla et al., Science 347, 49–53 (2015).
12. D. A. Dirocco et al., Angew. Chem. Int. Ed. 53, 4802–4806 (2014).
13. C. A. Lewis, K. E. Longcore, S. J. Miller, P. A. Wender, J. Nat. Prod. 72, 1864–1869 (2009).
14. J. A. Friest, S. Broussy, W. J. Chung, D. B. Berkowitz, Angew. Chem. Int. Ed. 50, 8895–8899 (2011).
15. M. S. Eom et al., Org. Lett. 18, 1720–1723 (2016).
16. P. Krattiger, C. McCarthy, A. Pfaltz, H. Wennemers, Angew. Chem. Int. Ed. 42, 1722–1724 (2003).
17. K. H. Shaughnessy, P. Kim, J. F. Hartwig, J. Am. Chem. Soc. 121, 2123–2132 (1999).
18. G. T. Copeland, S. J. Miller, J. Am. Chem. Soc. 121, 4306–4307 (1999).
19. J. R. Cabrera-Pardo, D. I. Chai, S. Liu, M. Mrksich, S. A. Kozmin, Nat. Chem. 5, 423–427 (2013).
20. A. B. Diagne, S. Li, G. A. Perkowski, M. Mrksich, R. J. Thomson, ACS Comb. Sci. 17, 658–662 (2015).
21. V. I. Martin, J. R. Goodell, O. J. Ingham, J. A. Porco Jr., A. B. Beeler, J. Org. Chem. 79, 3838–3846 (2014).
22. C. Créminon, F. Taran, Chem. Commun. 51, 7996–8009 (2015).
23. M. W. Kanan, M. M. Rozenman, K. Sakurai, T. M. Snyder, D. R. Liu, Nature 431, 545–549 (2004).
24. A. McNally, C. K. Prier, D. W. C. MacMillan, Science 334, 1114–1117 (2011).
25. D. W. Robbins, J. F. Hartwig, Science 333, 1423–1427 (2011).
26. Descriptions of all macros used in this work are available as part of the supplementary text. The macros themselves are also available as the supplementary materials.
27. The yields of the reaction were approximated by the relative ion counts (with respect to hexadecane) of the corresponding products in the GC-MS chromatograms of the reaction mixtures.
28. F.-S. Han, Chem. Soc. Rev. 42, 5270–5298 (2013).
29. T. M. Gøgsig et al., J. Am. Chem. Soc. 134, 443–452 (2012).
30. H. D. Srinivas, Q. Zhou, M. P. Watson, Org. Lett. 16, 3596–3599 (2014).
31. X. Cui et al., Org. Biomol. Chem. 11, 3094–3097 (2013).
32. C. Zhou, D. E. Emrich, R. C. Larock, Org. Lett. 5, 1579–1582 (2003).
33. A. Arcadi et al., J. Org. Chem. 78, 4490–4498 (2013).
34. J. D. Shields, E. E. Gray, A. G. Doyle, Org. Lett. 17, 2166–2169 (2015).
35. E. C. Hansen et al., Nat. Chem. 8, 1126–1130 (2016).
36. The reaction between diphenylacetylene 2a (X=H, Y=H) and allylic ester 3a' (Z=H) in the presence of the dihalo-substituted alkyne 2i (X=Br, Y=Br) failed to produce 16aa or 16ia after 1 hour of reaction time.
37. Y. Yamamoto, E. Ohkubo, M. Shibuya, Green Chem. 18, 4628–4632 (2016).
38. Y.-M. Wang, S. L. Buchwald, J. Am. Chem. Soc. 138, 5024–5027 (2016).
39. M. Mailig, A. Hazra, M. K. Armstrong, G. Lalic, J. Am. Chem. Soc. 139, 6969–6977 (2017).
40. D. P. Todd, B. B. Thompson, A. J. Nett, J. Montgomery, J. Am. Chem. Soc. 137, 12788–12791 (2015).
41. The reaction of 2a with diethyl (1-phenylallyl) phosphate using the conditions described in (38) resulted in 10% conversion of 2a and formation of trace quantities of 16aa and its isomers after 8 hours of reaction time (GC-MS). No improvement in the yield of 16aa was observed after an additional 10 hours of stirring.
42. The reaction of 2a with cinnamaldehyde using the conditions described in (40) produced (E)-(1,2-diphenylvinyl)triethylsilane as a major product as well as trace quantities of isomers of 16aa (GC-MS).
43. The reaction of 2a with diethyl (1-phenylallyl) phosphate using the conditions described in (39) resulted in formation of 16aa in 34% yield as well as formation of two isomers of 16aa in slightly lower quantities (detected by GC-MS). The same conditions afforded 18% of 16aa and a larger quantity of its isomer (detected by GC-MS) when cinnamyl diethyl phosphate was used as an allyl source.
44. Our method is not limited to analysis by GC-MS and could be used in combination with LC-MS data gained on products that are less volatile than those in the current study.

Snap deconvolution: An informatics approach to high-throughput discovery of catalytic reactions

Konstantin Troshin and John F. Hartwig

A triple search for coupling reactions

Coupling reactions are, in principle, good candidates for high-throughput discovery: Simply mix a diverse set of reagents and then look for products that combine two or three of their masses. In practice, however, numerous different products might have masses that are too similar to distinguish quickly. Troshin and Hartwig circumvented this problem by screening three pools of reagents in parallel that shared the same reactive functionality but differed in mass by carefully chosen increments. Specific products could then be identified in a noisy distribution by their distinctive expected mass differences across the three pools.

Science, this issue p. 175

View the article online
https://www.science.org/doi/10.1126/science.aan1568
Permissions
https://www.science.org/help/reprints-and-permissions

Use of this article is subject to the Terms of service