Combinatorial and rapid screening approaches to homogeneous catalyst discovery and optimization

Robert H. Crabtree

Yale Chemistry, PO Box 208107, 225 Prospect St., New Haven, CT 06520-8107, USA

Received (in Bloomington, IN, USA) 3rd February 1999, Accepted 29th March 1999

Rapid screening and combinatorial chemistry are expected to influence the way homogeneous catalysts are discovered and developed.

Homogeneous catalyst discovery owes at least as much to empirical testing as to mechanistic understanding, because structure/function relationships are either obscure or not readily predictable for most catalysts. Once an initial lead catalyst has been identified—either in current experimental work in one’s own laboratory or in the literature—incremental changes in the ligand set, often guided by simple mechanistic considerations, are then made to bring the activity, selectivity or scope of the catalyst up to useful levels. Since each new potential catalyst is normally separately prepared and purified, then assayed, progress is relatively slow. In our own early studies, an iridium catalyst that is highly active for the hydrogenation of tri- and tetra-substituted alkenes was only found after a year, even though only 6–12 complexes were assayed. Activity proved to be poor in then-standard solvents and only when a weakly coordinating solvent, CH$_3$Cl, was used did high activity emerge. At that time CH$_3$Cl was considered inadvisable for homogeneous hydrogenation because of its oxidising character. This emphasizes the point that a new catalyst may need to be assayed under non-standard conditions, where rapid screening methods could prove useful.

With the advent of combinatorial chemistry, a new strategy is currently emerging that may accelerate the pace of homogeneous catalyst discovery and development. It combines parallel synthesis of a broad range of catalysts, the catalyst library, with a rapid and preferably parallel assay, the rapid screen. When perfected, this procedure is expected to greatly accelerate the rate at which catalysts are discovered or improved although the area is still in its earliest phases. Another potential advantage of making screening much easier is that complexes that do not appear promising candidates can also be assayed, leading to the possibility that unexpected classes of catalyst may be found. One recent example, discovered by conventional means, shows that an important class of catalyst for alkene polymerization long escaped discovery in spite of intense activity in the field, in this case because, being based on a late metal, it defied expectations based on what later proved to be oversimplified mechanistic arguments. For a decade or more in the development of alkene polymerization catalysis, it was thought that commercially useful polymer molecular weights in alkene polymerization would only be obtained with early metal catalysts. Late metals were considered to be too prone to fission and would therefore only form oligomers at best. In spite of intense activity in the polymerization area, highly active late metal complexes were only discovered very recently. They prove to have usefully different properties from their early metal counterparts. This example shows how the influence of inappropriately generalised mechanistic ideas can sometimes hold back advance in catalysis research by creating artificial conceptual barriers to innovation.

Combinatorial ideas have already made a strong impact in drug discovery, are becoming established in organic chemistry, and are now beginning to enter homogeneous catalysis. In addition to speeding catalyst discovery and development they are likely to change homogeneous catalysis in a number of ways. The convenience of polystyrene-bead-based combinatorial solid phase synthesis and on-bead testing of catalysts may emphasize the area of supported homogeneous catalysis. Solid phase organic synthesis (SPOS) allows certain types of reactions to be carried out particularly easily, including ones not possible in solution, so we may be led to envisage ligands not currently in the lexicon of the catalysis chemist.

Combinatorial chemistry

Combinatorial chemistry involves three steps: the rapid parallel synthesis of a library of many compounds of related structure, the parallel testing of these compounds for a desired property by an appropriate assay, and the identification of the compounds, called ‘hits’, that show the best desired properties. In this way, chemical structural diversity can be thought of as a multi-dimensional space probed by the combinatorial method. Once an initial hit has been identified in the initial broad library, a new library may be constructed that proves a smaller region of diversity space around the initial hit. Combinatorial chemistry
can therefore be considered as an artificially accelerated evolution process but with human rather than natural selection as its motor.

Merrifield’s solid phase synthesis of polypeptides (1963) on polystyrene beads was the first step on the road to the new ideas. In a Merrifield synthesis, the growing polypeptide chain is anchored to a polystyrene bead and extended one amino acid residue at a time using appropriate reagents; final cleavage gives the desired polypeptide. The method can readily be automated. The advantage of the solid phase is that the steps can be pushed to high yield using an excess of reagents and the growing polypeptide chain can be separated at each step simply by filtering the beads and washing. The availability of this procedure led to the concept of replacing a specific amino acid residue reactant, as is normally used in a specific chain-extension step, by a mixture of reactants, leading to a mixture of compounds in the final product. By having fixed amino acid residues at most sites but variable residues at a few sites, an unprecedentedly large degree of diversity could be obtained rather easily within a class of closely related polypeptides. Suitable procedures were then developed to assay for the desired ligand binding properties and identify which sequence is responsible for binding.

The next conceptual step was the recognition by Bunin and Ellman that the same sort of diversity could be created by SPOS for non-peptide organic compounds, where the substitutent groups on a central motif were permuted. The method was first applied to benzodiazepines, where four different substituents could be independently varied.

From this, it is a short step to the idea of creating a diverse library of potential ligands on a set of beads, binding a reactive metal complex to the supported ligand and assaying the resulting bead for catalytic activity. This and related ideas are being developed at the moment, both in academia and in such industrial companies as Symyx.

**Rapid screening of catalysts**

High throughput screening (HTS) in the pharmaceutical industry became so efficient by the early 1990s that it made the synthesis of drug candidates the rate determining step in the drug discovery process and directly encouraged the adoption of combinatorial synthesis. In the catalyst area, neither HTS methods nor combinatorial synthesis had been applied until very recently, so both have had to be developed together.

Assays reported to date can be either parallel or non-parallel: in a parallel assay, all the data are collected at one time, by a conventional method. It is clearly more efficient to screen a combinatorial catalyst library with a parallel HTS screen than with conventional GC or HPLC, although the availability of automatic sampling could modify this conclusion. Continuous assays have the advantage of allowing monitoring of a reaction in real time; others require some action to be taken to gather the data, such as taking a sample, in which case the method is discontinuous. Methods are also likely to vary in their quantitative value, going from a purely qualitative indication to a detailed quantitative analysis of all species present. They also vary very considerably in sophistication going from low to high tech. The former are generally easier to apply but the latter will be advantageous for automation. There will probably eventually be a hierarchy of such HTS assays, but the initial screen is likely often to be that for activity, because without sufficient activity, no catalyst, however selective, is likely to be useful.

Perhaps the simplest is our own continuous, parallel assay based on a reactive dye that bleaches when a catalytic reaction, such as hydrosilation, takes place. The assay is carried out in a glove box or glove box on a Teflon block drilled with 70 reaction wells. This allows us to continuously monitor a large set of catalysts until a ‘hit’ (the most active catalyst) is registered by the dye bleaching in one of the reaction wells. We needed a dye that would not have potentially interfering reactive groups, hence the choice of a ferrocenyl group as electron donor and of a pyridinium as acceptor; the bulky benzylic tail is needed to make the dye conveniently soluble. When the reactive C≡C or C≡N bond is saturated, the electronic connection between donor and acceptor is broken and the absorption coefficient of the material drops by a factor of ca. 100. The starting dye must absorb intensely [ε(EtOAc) = 12 600, 1a; 5200, 1b] so as to mask any catalyst color and to be a sensitive indicator. Some quantitative data can be obtained: an initial bleaching time, \( t_b \), corresponds to the first observable color change relative to a control well and a final time, \( t_f \), corresponds to complete bleaching. Other work indicated that \( t_f \) corresponds to ca. 40% dye conversion and \( t_i \) to ca. 95%. A long induction time therefore translates to a long \( t_f \) and high activity after induction to a short value of \( t_f - t_i \). Recording the data proved possible using a digital camera.

As an initial test, we applied the method to a library of conventional catalysts, some of which were known to be active while others had never been considered for hydrosilation. The results showed that such well known species as RhCl(PPh₃)₃ are very active catalysts for P–Ph₂SiH₂, but, unexpectedly, that among the most active of all was a palladacycle never previously tried for hydrosilation. For reasons that are not yet clear, the dyes are activated substrates and react much faster than conventional ones, like stilbene, but the relative order of activity of different catalysts seems to be preserved between dye-substrates and conventional ones. Controls demonstrated that catalyst, silane and substrate are all required for reaction to occur. Fig. 1 shows the result of a typical run.

![Fig. 1 Results of a typical run using the dyes 1a,b to screen the catalysts identified in the figure for hydrosilation with Ph₂SiH₂.

The heat output of a catalytic reaction has been used by several groups as a continuous, parallel assay for activity. For example, Morken and coworkers used an IR camera to record heat output. This was used to assay a simple polypeptide catalyst library on polystyrene beads that catalysed an exo-thermic ester hydrolysis in one case and in another the catalysts were homogeneous Mn, Cr and Co complexes for epoxide
opening. Each hit appeared as a bright spot on the image (Fig. 2) as the result of that bead having a temperature estimated to be ca. 1 °C above that of the reaction medium. Some differentiation of the spots by relative intensities was possible. Heterogeneous catalyst arrays can be analyzed in the same way.\textsuperscript{10}

Hartwig and coworkers\textsuperscript{11} have reported an interesting discontinuous fluorescence screen in which a series of forty conventional homogeneous Pd(dba)\textsubscript{2} + L (L = phosphine and diphosphine) catalysts were assayed for activity for coupling of an aryl halide with an alkene (Heck reaction). The aryl halide component was grafted onto cross-linked polystyrene and catalytically coupled with a soluble alkene bearing a powerfully fluorescent coumarin group (Fig. 3). At the end of the catalytic reaction the polymer beads were isolated by filtration and their fluorescence assayed visually as low, moderate or high. The illustrations were kindly provided by Professor Morken.

Asymmetric reactions pose special problems because a catalyst may be efficient only for one class of substrate but mediocre for others. To help solve this problem, Kagan\textsuperscript{15} has proposed a method in which the solvent can be useful. A study by Burgess\textsuperscript{16} illustrates the use of HTS to identify the best ligand/metal/solvent combination for a carbene insertion into a C–H bond [eqn. (3)]. In all, seven different metal ions, five chiral ligands and four solvents were studied using HPLC with an autosampler for determination of the diastereoisomer ratio. Each distinct combination was run in one well of a 96-well plate, allowing comparison in triplicate of seven catalysts at three different loadings. The data were good enough to allow the kinetics to be followed. As described in more detail below, the same method was also applied to assaying a small catalyst library on polystyrene beads.

Using an automated fluorescent plate reader allowed quantitative, parallel intensity data to be obtained in solution on a 96-well plate, allowing comparison in triplicate of seven catalysts at three different loadings. The data were good enough to allow the kinetics to be followed. As described in more detail below, the same method was also applied to assaying a small catalyst library on polystyrene beads.

Asymmetric reactions pose special problems because a catalyst may be efficient only for one class of substrate but mediocre for others. To help solve this problem, Kagan\textsuperscript{15} has proposed a method in which the method allows the products to be directly analyzed and selectivity data to be obtained. Catalyst libraries have also been evaluated by the conventional discontinuous, non-parallel methods of GC and HPLC.\textsuperscript{14}

Libraries of catalysts

A wide variety of methods\textsuperscript{2} have been used for implementing combinatorial strategies in other fields but we will only discuss methods that have been applied to transition metal chemistry. The types of libraries that can be considered are just as varied as the types of assays. Libraries, which may be soluble or polymer-bound, may consist of preformed catalysts or of ligands that are active without metal ions or are converted to the catalytically active form by loading a suitable metal-containing precursor \textit{in situ}. Catalysts can be highly solvent-sensitive, so varying the solvent can be useful. A study by Burgess et al.\textsuperscript{16} illustrates the use of HTS to identify the best ligand/metal/solvent combination for a carbene insertion into a C–H bond [eqn. (3)]. In all, seven different metal ions, five chiral ligands and four solvents were studied using HPLC with an autosampler for determination of the diastereoisomer ratio. Each distinct combination was run in one well of a 96-well microtitre plate. Silver ion, previously not generally considered as a suitable catalyst, was shown to be very effective.

Having described the principal known catalyst assays, we move to the types of catalyst libraries that have been considered to date. These can be soluble or polymer-bound and either conventional or combinatorial. Although using a rapid parallel
assay with a large combinatorial library is ultimately likely to be the most effective strategy for catalyst discovery, semiclassical approaches like conventional screening of combinatorial libraries or rapid screening of conventional catalysts are likely to be useful in certain situations.

A set of conventional homogeneous metal catalysts, such as was used in a study discussed above, constitutes one of the simplest soluble libraries, and can be useful when the best combination of metal and ligand type is sought in the initial stage of an investigation. Once these have been chosen, the natural next step would be to vary the ligand set using parallel synthesis. Solid phase organic synthesis (SPOS) being so widely used for polymer-bound library synthesis, many catalysts are likely to be most conveniently synthesized on beads. Extensive literature on SPOS is available,[17] but some discussion is appropriate here. Merrifield’s original 2% cross-linked styrene–divinylbenzene copolymer is one current recognized standard, because it swells in many organic solvents, thus allowing reagents access to the interior of the bead, yet it does not dissolve and has substantial thermal stability. The material is partially chloromethylated to form the base resin used in SPOS. The level of chloromethylation determines the maximum achievable loading of the resin. A linker is normally employed to provide the starting point for the organic synthetic steps proper and to introduce a spacer group so that the reactive functionality is held somewhat apart in space from the polymer itself. The organic synthesis then proceeds on the linker. High yields are required in each step if good overall yields are to be obtained. Many standard organic reactions have been found to go well on solid phase and an excess of reagents can often be used to drive these reactions to high yield. Finally, the product may be cleaved from the resin, if a cleavable linker has been chosen, or used directly on-bead, although full characterization may not be possible if this alternative is chosen. Other related resins, such as Wang resin,[2] have also proved useful in catalysis studies.

A small library of four purely organic peptide-like catalysts has been assayed for acyl transfer catalytic activity with Miller’s fluorescence assay [eqns. (1) and (2)].[12] In this case, the catalyst and the sensor were both grafted onto a Wang resin, each bead containing only a single type of catalyst. Once exposed to the substrates, acid was formed within the individual beads by the catalytic reaction. The resulting fluorescence [eqn. (2)] was detected visually with a microscope. Similar relative activities were observed between soluble and bead-bound versions of the same catalyst.

While not catalytic, Jacobsen’s[18] combinatorial syntheses of a peptide-like library having metal-ligating properties illustrates how standard[2] ‘split-and-pool’ techniques can be used to introduce diversity into a ligand library. In such a synthesis, different batches of beads receive different initial residues in the first step of a parallel synthesis, and are then pooled together. The beads are then split so each batch can be subjected to a synthesis step that attaches a different residue in the second position. The beads are once again pooled and the procedure repeated until all the residues have been added. The final pool contains beads that have every permutation of possible sequences. The Still[19] encoding procedure was used by which a covalently bond organic tagging compound was grafted onto each type of bead to allow the ligand present on that bead type to be identified via MS detection of the tag after cleavage. Finally, the beads were assayed for their ability to bind Ni(II) vs. Fe(III). Colorimetric detection of metal binding provided the required assay.

Solid phase synthesis of a ligand and loading a metal to make a catalyst precursor poses no particular difficulty, but cleavage of a reactive catalyst from the bead is not likely to be possible without degradation of the complex, so it is likely that the ligands will be characterized on an aliquot of the bead sample before metal binding and catalysts will normally be assayed on-bead. For a valid comparison of two bead-catalysts for activity, it is not necessary to know exactly how many of the purported active sites really are active, but all the beads in the study must have approximately the same proportion of active sites. That beads that prove to be most active contain active sites is certain, but it may be very hard to rule out the possibility that one of the inactive or barely active beads would have been extremely active if the desired complex had been formed as intended. At a severely practical level one could say that this does not matter—a catalyst that cannot be made is not of practical interest. On a theoretical level, however, such an outcome would severely perturb any structure–activity relationships one was hoping to extract and therefore lead to erroneous conclusions. From the point of view of improving our understanding of catalysis, these structure–activity relationships are likely to be the most important general points that emerge from combinatorial catalysis and it remains to be seen how reliable these will be. Direct comparison of bead catalyst data with data from the solution analogues will not necessarily give reliable information, because we will not know whether to ascribe any deviation to improper synthesis of the active catalyst on the bead or to catalyst–polymer interactions.

One technical problem that can be readily addressed is the possibility of slow diffusion of the catalytic reagents into and products out of the bead could mask activity differences. To distinguish such differences between catalysts we need the catalyst turnover, not the diffusion steps, to be rate-limiting. This is most readily achieved, we find, by limiting the loading of the resin. Instead of using commercially chloromethylated resin, we therefore have made our own; the appropriate level may need to be determined for each case. This is not too general. We also need a solvent that efficiently swells the polymer to allow rapid substrate access and product departure; the observed rate needs to reflect the rate of the catalytic and not the diffusion steps.

Assuming we use a suitable assay for picking the bead-bound catalyst with the properties that best correspond to what is needed, we are faced with a significant choice: do we try to make a homogeneous analogue of the successful catalyst or do we continue to use the bead-bound version in later work. The advantage of the homogeneous version is that it can be fully characterized and it is more amenable to full mechanistic studies. The successful ligand may be much harder to make by conventional methods than by SPOS, however, but this problem is unlikely to be serious. The most severe problem is that the activity and selectivity of the catalyst may well be very different in going from the bead-bound to the soluble forms; many examples of big differences in chemistry between soluble and polymer phase are known. As an example of a typical difference that might be expected to affect catalytic activity, a polymer-bound catalyst is unlikely to be as easily able to dimerize as a soluble version, and it may also interact with the linker or backbone polymer in some way. Not enough is known to be able to tell how big a problem this will prove to be, but it is already clear that solution and bead-bound catalysts do show significant differences. A significant dependence on the nature of the
polymer support and differences between bead-bound and the corresponding homogeneous catalyst in solution have already been seen for Jacobsen’s epoxidation catalyst.\textsuperscript{20} On the other hand, Hoveyda studied this specific problem and found generally good agreement between the selectivities of bead-bound and solution phase catalysts in the particular case studied.

We have prepared a diverse monophosphine library\textsuperscript{21} of the type \(P-\text{C}_6H_4PRR^\prime (P = \text{polymer})\) on cross linked polystyrene by the route shown in Fig. 4 and loaded it with \(\text{M(diene)}\)-

![Fig. 4 The synthetic route to the generalized phosphine library of ref. 21.](image)

(py)$_2]PF_6$ (\(M = \text{Rh}, \text{Ir} \); \(py = \text{pyridine}\)). The two resulting sublibraries were assayed for hydrosilation activity with dyes 1a,b following the prior\textsuperscript{8} procedure to find the best supported catalysts. The resulting activities followed a different phosphine dependence for Ir vs. Rh and for alkene dye vs. imine dye, so catalysts selective for either alkene or imine could be identified. Structure/activity trends were weakly marked. The best catalyst of all had \(M = \text{Ir}, \text{R} = \text{Rh} \) and \(R^\prime = \text{1-naphthyl} \), but only rapidly reduced the alkene dye and the Rh analogue was rather slow for both dyes, an unexpected combination given the other trends seen. A combinatorial search may be particularly useful in situations such as this where the existence of ill-defined trends makes it hard to design a good catalyst.

Limitations

These methods are likely to have a number of significant limitations. Not all catalytic problems will be readily susceptible to the rapid screening-combinatorial approach. Bead-bound catalysts may also differ strongly from their soluble analogues in certain cases, making correlations between the two types of systems very difficult. Bead bound catalysts are difficult to characterize, and so one may not always be able to tell for sure why any particular bead-catalyst is poorly active—it could be that the catalyst is inherently poorly active or that the intended complex was not properly formed on the bead or that a catalyst dimerization only possible in free solution is required for activity. Careful controls and reliable analysis of the materials for metal content will be needed. Different solvents are expected to swell the polymer to different extents, resulting in rate differences that are unrelated to what would be seen in solution. We typically check, for example, that the unmodified starting resin does not take up metal and act as a catalyst in the absence of covalently attached ligand groups and that typical high activity resins have similar metal analyses to medium and low-activity resins. Even more than in traditional catalyst work, care will be needed in interpreting the results of combinatorial studies.

Combinatorial methods are unlikely to displace traditional characterization and mechanistic work. Indeed, these methods may enrich mechanistic studies by making structure–activity relationships available over a much wider range of structures than is usual in traditional studies. They may also help correct mechanistic misconceptions based on examining too small an amount of data.

Combinatorial methods could potentially be applied to any inorganic, organometallic, or bioinorganic problem where a suitable assay can be devised. For example, functional modeling of enzyme active sites has proved very challenging by traditional approaches. They also seem suitable for determining what coordination or organometallic structures can bind, or selectively bind, particular classes of ligand; in many cases, these ligands could be covalently bound to dyes for easy visualization. This approach might also help in determining the resting state of catalysts. For example, in hydrosilation, the silane might be tagged with a red dye and the alkene with a blue one. In such a case, catalyst beads that bound silane but not alkene would be red, silane but not alkene, blue, and if both were bound the beads would be expected to appear purple.

Future developments

There are already initial indications that other areas of inorganic chemistry than homogeneous catalysis are also likely to benefit from combinatorial methods. Mallouk et al.\textsuperscript{22} have shown how they can be applied to finding electrochemical oxidation catalysts, for example, and they have also been used for materials synthesis, such as finding a superior luminescent material.\textsuperscript{23}

Although the area of combinatorial chemistry and rapid screening is very new as applied to inorganic chemistry, it has already made a significant impact. The challenge now is to develop methods to apply it to a variety of problems to see how widely they are applicable. Care will be needed to characterize the resulting materials and to check the reliability of the data obtained by comparison with traditional approaches on selected materials. There are so many different ways these ideas could be embodied, that the effort will require ingenuity combined with a close attention to practicality and care in interpretation.

Acknowledgements

I am grateful to the US Dept. of Energy, Catalytica Corp. and Mitsubishi Oil Co. for funding, Jennifer Loch, Alyssa White, Matthew Torres and Alan Cooper for unpublished data and the other coworkers mentioned in the citations for their work on the problem.

Notes and references


Paper 9/01022J