Pursuit of Noncovalent Interactions for Strategic Site-Selective Catalysis

Published as part of the Accounts of Chemical Research special issue “Holy Grails in Chemistry”.

F. Dean Toste,*§ Matthew S. Sigman,*† and Scott J. Miller*‡

§Department of Chemistry, University of California, Berkeley, California 94720, United States
†Department of Chemistry, University of Utah, 315 S. 1400 E., Salt Lake City, Utah 84112, United States
‡Department of Chemistry, Yale University, 225 Prospect Street, New Haven, Connecticut 06520, United States

ABSTRACT: Selective reactions on structures of high complexity can move beyond the mind’s eye and proof-of-principle. Enhanced understanding of noncovalent interactions and their interdependence, revealed through analysis of multiple parameters, should accelerate the discovery of efficient reactions in highly complex molecular environments.

The day will almost certainly come when chemists will inspect any molecule and be able to construct numerous analogs through a series of reactions that introduce changes, bond-by-bond, with a fantastic level of precision, in order to access the desired function. Control over all manner of selectivity, enantio-, diastereo-, regio-, and chemoselectivity, will be obligatory. Modern organic synthesis is headed in this direction. In fact, our field has taken significant steps toward developing such methods, catalysts, and reagents that facilitate control over reaction outcomes that can, in principle, produce several unique products. As the quintessential example, asymmetric reactions that control stereochemistry provide an inspirational look at how rapidly our field can evolve. Control over enantioselectivity for reactions that functionalize prochiral π-bonds, either reductively (Figure 1a) or oxidatively (Figure 1b), provide powerful examples. When the authors of this essay were born, catalysts for these types of reactions were quite rare. Today, we observe numerous asymmetric catalysts for these reaction types and, importantly, extraordinary growth in the breadth of reactions that organic/organometallic chemists can address creatively with numerous enantioselective catalyst types. Yet, the underbelly of this discipline is the vast resources and effort that are typically deployed for the identification of effective catalysts. While a series of privileged catalyst structures have emerged, “design” of catalysts for the immense number of interesting reactions remains highly challenging and correspondingly important intellectually and practically. Likely most practitioners in the field would agree that a high level of empiricism is required for catalyst optimization for virtually any catalytic enantioselective reaction. Thus, elucidation of the mechanistic underpinnings of enantioselectivity and generally improved catalyst performance are almost certainly “Holy Grails” for chemists who seek a much more rational foundation. Fortunately, in the modern physical organic study of catalytic asymmetric reactions, one can find a basis for optimism that our field is advancing along these lines wherein complex structure–function relationships can now be interrogated. One inescapable theme, now emerging repeatedly, is the interconnectivity of numerous factors. New ideas for considering functional group parametrization, as well as for developing multiple parameter analysis tools for complex reactions, are increasingly necessary to account for observations quantitatively (Figure 1c, vide infra).

The core of catalyst “design” is the fundamental understanding of the factors guiding observed outcomes. All manner of interactions, including covalent and noncovalent bonding interactions, contribute to the energies of transition states leading to divergent reaction outcomes. Yet, one consistent obstacle to obtaining such insight is the reality that the forces influencing reaction outputs such as enantioselectivity or site-selectivity, in particular when multifunctional catalysts are employed, are often noncovalent in nature. The energy increments associated with these interactions are typically modest (fractional kcal/mol thru ∼2–3 kcal/mol) and each modulates another in transition states, either stabilizing a competing pathway. The aggregate is a proverbial ensemble of transitions states leading to divergent outcomes, complicated in a factorial manner by the number of possible products. Additionally, the physical essence of many

Received: December 7, 2016
Published: March 21, 2017
noncovalent interactions remains controversial, and computational and experimental methods both depend on resolution of the nature of these forces.10,11 Perhaps due to these ambiguities, the emergence of noncovalent interactions as an explicit “design” principle in controlling reaction trajectories has been gradual, if not sluggish. Yet, it also holds tremendous promise, since these forces are attractive in nature. The achievement of selectivity by specific rate acceleration, rather than by inhibitory means and energetic destabilization suggested in archetypical, traditional steric type arguments for asymmetric induction, parallels the essence of catalysis and its capacity to promote reactions faster.12,13

As a paragon of what the future could look like, enzymes provide a hint.14 They have evolved to exhibit high levels of catalyst control and accelerated rates using noncovalent interactions. As a distinct illustration, site-selective reactions, those that select for the formation of one product when multiple products via the same reaction mechanism are possible, occur frequently in nature.15 Yet, site-selective catalysts are rare using nonezymatic, small molecule catalysts. As an oft-cited enzymatic example, cytochrome p450s are well-known to catalyze exceptionally site specific reactions on substrates with many C−H bonds, and indeed examples that precisely effect hydroxylation are favorites of scientists who endeavor to mimic these processes synthetically.16 The selective and staged oxidations that deliver Taxol from the parent terpenoid are particularly dramatic (Figure 2a).17 Yet, examples of rational variation of enzymes to achieve comprehensive diversification of a complex natural product through all possible C−H bond oxidation products are not yet known. Nonetheless, great advances are emerging at this fascinating research frontier through both enzymology and studies of small molecule catalysis.18,19

Applications of nonezymatic catalysts to the diversification of complex substrates might be said to be even more primitive. After all, success in these types of projects might be measured by the discovery of “n” different catalysts for the specific functionalization of, for example, each hydroxyl or amine present in a substrate like amphotericin (Figure 2b, green arrow).20,21 Might chemists also desire or dream of a comprehensive set of “m” catalysts for the site- and stereo-selective oxidation of each double bond present in a natural product containing a polyolefin (Figure 2b, red arrow)?22 Functional group-specific derivatization in a complex substrate, where many occurrences of the same functional group exist, surely represents a state-of-the-art challenge for those studying selective catalysis. The challenge intensifies with substrates that possess different functional groups that can react with common catalyst types.23 In any event, while progress is being made, comprehensive solutions to this type of “late stage functionalization” problem on a substrate of even modest complexity do not yet seem to be known. It is also a problem that requires the reordering of intrinsic functional group reactivity hierarchies, which is an energetically daunting challenge, also likely categorized as of a “Holy Grail” level of difficulty.

One representative area of chemistry that provides a window into the nature of the problem involves the catalyst-dependent modification of complex glycopeptides, like teicoplanin. The
site-selective modification of this target is an ongoing project in the Miller lab, where selective bromination and polyl alteration have been reported. As for the polyl, teicoplanin A₂₂ is a structure that contains 13 reactive hydroxyl groups. Employing a minimal protecting group strategy and a teicoplanin derivative, "(Allyl)₆-Teicoplanin A₂₂", that contains ten free hydroxyl groups, three distinct catalysts (a "red" one, a "blue" one, and a "green" one) were discovered that allow efficient modification (phosphorylation in this case) of three of the ten sites (Figure 3a). Encouraging aspects of this study included (a) the high selectivity obtained with the three catalysts, (b) the success of rational design based on noncovalent interactions between catalyst and substrate to achieve two of the catalysts, the "red" and "green" variants, and (c) the success of a combinatorial screening campaign to discover the third "blue" catalyst. Yet, the list of shortcomings of the study is longer and includes the following: (a) ten catalysts for the selective functionalization of the ten other hydroxyl groups remain unknown; (b) the three catalysts that were found require that teicoplanin itself be converted to a compound modified by several protecting groups; (c) rational design of catalysts for the other ten sites has, so far, been unsuccessful; (d) combinatorial libraries of catalysts, so far, have not yet delivered the other ten catalysts either. Nonetheless, the elucidation of a catalyst–substrate complex X-ray structure (Figure 3b), which reveals critical noncovalent associations that are clearly consistent with highly selective catalyst performance, provides a most optimistic sense that the overarching themes of the Holy Grail under discussion are attainable.

Even with molecules considerably smaller than teicoplanin, such as a relatively unbiased internal alkene, achieving high levels of site-specificity for the addition of an organometallic (to one end of the double bond or the other) is difficult. As an example, efforts in the Sigman group have focused on developing Heck-type reactions to achieve such site-selectivity for the migratory insertion step, which also occurs as part of an overall process delivering high enantioselectivity (Figure 3c). In this case, a remote biasing group is still required (in this case an alcohol); as this group is placed more distal from the reaction site, selectivity diminishes. Additional interesting observations, including electronic dependence of the boronic acid coupling partner, suggest that this process is even more multifaceted. It is humbling, but also inspiring, to recognize that challenges in lower complexity situations of this sort remain state-of-the-art hurdles to clear as part of extending these solutions to more complicated systems in the future.

What is required for the chemist to be able to develop rapidly panels of site-selective catalysts that can comprehensively diversify both simple and complex substrates of interest? Could it be that there will be a significant intersection of advances in mechanistic understanding of enantioselective reactions and the discovery of certain site-selective catalysts? Our thought is that there are common links between these seemingly disparate lines of inquiry in chemistry. On the other hand, our field has seen an eruption of new catalysts that are explicitly designed to take advantage of noncovalent

Figure 2. (a) Enzymatic hydrocarbon diversification to produce Taxol, but variants to hypothetical taxoid not yet known. (b) Generalized substrate diversification with catalysts, exemplified by amphotericin.
interactions at the heart of their mechanisms of action. At the same time, detailed studies of catalytic mechanisms involving data intensive inquiries of multiple catalyst and substrate parameters are now emerging that consistently point to the

Figure 3. (a) Identification of three distinct peptide-based catalysts that exhibit selectivity for individual hydroxyl groups of the complex glycopeptide teicoplanin. (b) X-ray crystallographic analysis revealing likely catalyst−substrate interactions. (c) Site selectivity dependency on migratory insertion in an enantioselective Heck arylation reaction.
operation of an interconnected array of noncovalent interactions that sum up to account for selectivity. All the while, the few cases of documented catalyst-dependent control of site-selectivity to deliver several substrate derivatives, including those derived from reversals of intrinsic functional group reactivity hierarchies, invariably point to the operation of subtle noncovalent interactions remote from bond-forming sites in the substrate molecule.

To achieve this goal requires an aggressive philosophy concerning empirical data collection, the use of this data, and not only the general goal of achieving a desired outcome but also an aspiration to understand why certain catalysts, catalyst/substrate combinations, and even solvent and additives are required at the culmination of an empirical optimization campaign. This is an approach that our groups have begun to integrate as a matter of course in probing the noncovalent interactions at play in the complex reactions we are studying.

Every data point—reaction yield, enantioselectivity, site-selectivity, mono/bis-/tris-functionalization ratio—can be construed as an invitation to physical organic analysis. Comprehensive, information-rich data sets create the potential for the whole to far exceed the sum of the parts. Compiling iterations of data sets, on a per substrate basis, is akin to exploring the scope of a reaction, with the potential for the extraction of emergent patterns that describe at a basic level why the reaction either performs well or not. This information can be construed as complex variants of a venerable physical organic chemistry experiment, the linear free energy relationship. If the outcomes can be correlated to parameters describing the structural permutations examined, several exciting possibilities including prediction of better outcomes to streamline development but also a hypothesis defining why the system performs in the manner it does. Reaction development, in this approach, takes the shape of the study of a complex system.

Early implementations of this approach are very promising. As a brief illustrative example, the Toste and Sigman teams collaborated to evaluate complex relationships of enantioselectivity to both substrate and catalyst structure in an intramolecular oxidative amination using chiral anion phase transfer catalysis (Figure 4a). As a first step, a combinatorial library was synthesized with an eye toward incorporating as much diversity as possible (Figure 4b). Subsequently, the entire data set was amassed followed by the collection of parameters describing both the substrate and catalyst. This is followed by collection of relevant, physical organic parameters, which are then statistically evaluated for correlation with the empirical data set. Complex models were initially discovered and
deconstructed to obtain mechanistic hypotheses allowing proposals to which types of noncovalent interactions were possibly responsible for asymmetric catalysis (Figure 4c).

Ultimately, this information was used to enable a virtual screen and validation of a better performing catalyst.

While this approach has now been applied to various reactions and contexts, the mechanistic models that result do not have the resolution that one generally sees depicted in today’s outputs of many studies of catalysis computationally. Advances in theory and computational stand to impact the future of catalyst design. Well-appreciated and clearly articulated challenges include the achievement of sufficient precision such that fractions of kcal/mol may be reliably assessed and compared throughout ensembles of transition states, often characterized by extreme conformational heterogeneity, in the face of the yet-to-be understood nature of certain noncovalent interactions. An exciting frontier is the integration of modern physical organic parametrization tools with theory to ultimately improve the resolution of understanding from kcal to cal, enabling sophisticated design.

Despite many challenges, an optimistic outlook emerges for this “Holy Grail” in the field of catalyst design. The interweaving of empirical screening data, statistical methods, transition state interrogation, and noncovalent interactions in the context of reaction development, where precise control of enantio- or regioselectivity is required, provides the backdrop for actual experiments and the accumulation of the necessary data. Against the odds, quite a few advances in these fields are emerging even today. Perhaps the ambition of total control over site-selectivity, in truly complex molecular settings with the aid of physical organic chemistry, theory, statistics, and analytical techniques can accelerate the discovery of this Holy Grail sooner than one might have thought possible just a few years ago.

### ACKNOWLEDGMENTS

F.D.T., M.S.S., and S.J.M. are grateful to the National Institute of General Medical Sciences of the National Institutes of Health for support (Grant R01 GM121383). We also want to thank all of our past and current co-workers who have not only made seminal contributions to our programs but also have provided endless inspiration.

### REFERENCES


### AUTHORS INFORMATION

Corresponding Authors

*E-mail address: fdtoste@berkeley.edu.

*E-mail address: sigman@chem.utah.edu.

*E-mail address: scott.miller@yale.edu.

### ORCID

F. Dean Toste: 0000-0001-8018-2198

Matthew S. Sigman: 0000-0002-5746-8830

Scott J. Miller: 0000-0001-7817-1318

### Notes

The authors declare no competing financial interest.


