Creating Molecules: Evolution of Randomness

Laura L. Kiessling

An urge to synthesize molecules dates back to the days of alchemy. Among the motivations that chemists currently find compelling are understanding how the structure of a molecule influences its function, studying the stability and reactivity of a particular molecule, and creating new molecules with new properties. The scientific hypotheses that we can address by making new molecules have radically changed with our understanding of chemistry and biology. These changes are reflected in the way chemists design and build molecules. The construction of organic molecules began with serendipity. A period wherein chemists learned to manipulate and form molecules with control and precision followed. Now, with a confidence in their ability to generate specific molecules, scientists are introducing randomness into their creations. This article explores the evolution of strategy in the design and construction of new organic molecules. Although similar transformations in strategy are taking place in other areas of chemistry, it is in the generation and synthesis of organic molecules that the trends are most apparent.

Molecules, those that comprise our world and those that we have yet to discover, have definite three-dimensional structures.[1] The varied connectivities and geometries of its atomic components determine a molecule's properties. The interrelationship of molecular structure and molecular function has vital consequences in biology. As a result, molecules can have similar shapes but completely different biological activities.

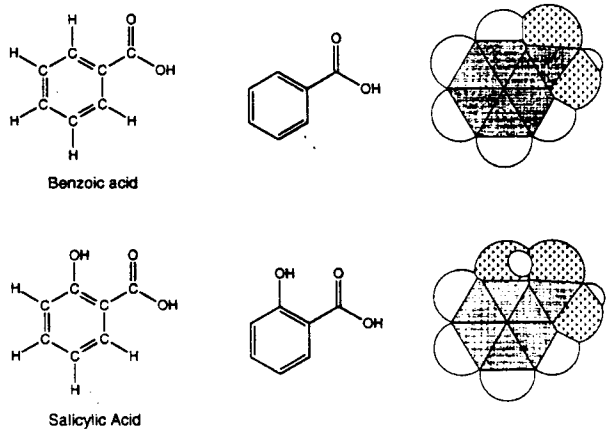
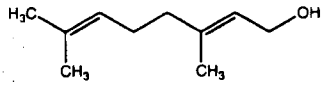


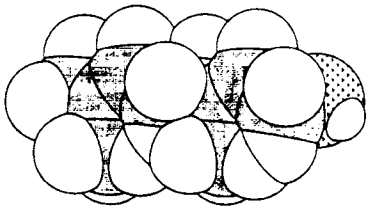
Figure 1. Different depictions of the two molecules benzoic acid and salicylic acid. Lines between atoms represent bonds. The left structures explicitly show each of the atoms. The center representation is a short-hand form with each vertex representing a C atom. H atoms are implied because each C atom has four bonds to it. The renderings on the right illustrate the space filled by the molecules.

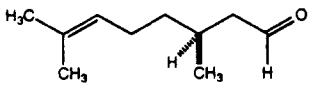
Nature provides us with molecules wherein her perspicacious alterations of composition and shape forge related molecules for very different purposes. Two compounds that differ by only one atom are benzoic acid and salicylic acid (Figure 1). The former is a food preservative, and the latter is derived from aspirin. A case wherein molecules differ by connectivity is seen by comparison of geraniol with citronellal (Figure 2A). Both molecules have the same composition: 10 carbon, 18 hydrogen, and 1 oxygen atom. However, as a consequence of their activities, which result from their different arrangements of atoms, people exploited these two molecules for different uses. Geraniol provides the rose scent for many perfumes, while endowing them with insect attractant activities; citronellal contributes to the function of citronella candles, in

which it is incorporated because of its ability to repel insects. Finally, there are examples where the most subtle variation of molecular structure affords substances with divergent properties, as is the case for the right- and left-handed forms of carvone (Figure 2B). The former smells and tastes like caraway, and the mirror image compound is used as spearmint flavoring. These examples delineate how the properties of two seemingly-related molecules can deviate: 1) they have similar shapes but differ by a single atom; 2) they have the same atoms connected differently; 3) they possess the same atoms, the same connectivities, but the atoms are attached in different spatial arrangements. Alternatively, the same physiological response can be elicited by very different substances.



geraniol





citronellal

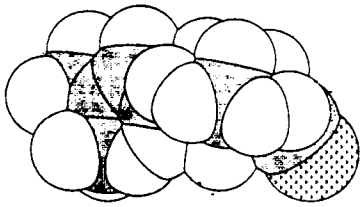
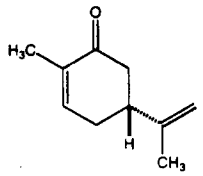
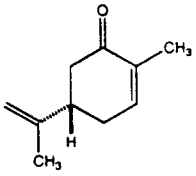


Figure 2A. Geraniol and citronellal have identical chemical compositions but the atoms are bonded differently. Although the molecules look somewhat similar, they have different biological properties. Heavy and dashed lines respectively indicate bonds coming out of and going into the plane of the paper.

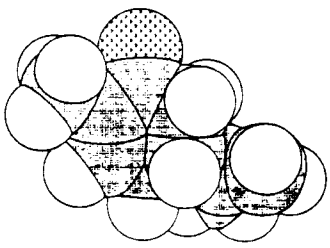
Given the interrelationship of molecular structure and biological function, the ability to create and synthesize molecules provides opportunities to study and shape our environment. In order to understand and exploit the properties that result from molecular structure, chemists must develop principles for generating particular molecules. The creation of molecules involves two distinct operations: the design of a target compound and its chemical synthesis. Sir J. W. Cornforth has defined chemical synthesis as "the intentional construction of molecules by means of chemical reactions."[2] In this article, I discuss how ideas and principles employed to synthesize new molecules have evolved.



R-(-)-Carvone
spearmint



S-(+)-Carvone
caraway



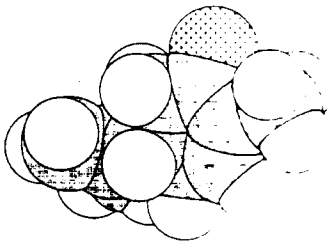


Figure 2B. R- and S-carvone are molecules that have different biological properties because they are mirror image isomers.

The chemical synthesis of organic molecules began auspiciously with urea, a component of urine. In 1828, Wöhler mixed two inorganic compounds, potassium cyanate and ammonium sulfate, in an effort to synthesize ammonium cyanate. Instead, he generated the organic compound urea (Figure 3).[3] This result was revolutionary. Wöhler's synthesis of urea repudiated the vitalism doctrine, which maintained that molecules that were found in living matter were endowed with a "vital force" so that they could not be generated from non-living materials. Even with Wöhler's result, this theory was not completely dismissed until Kolbe synthesized acetic acid, the active component of vinegar and bad wine. Still, it is striking that the molecules that make up a hand, a piece of tree bark, or a fly's eye are composed of atoms and can be manipulated by the same principles as metals and minerals.

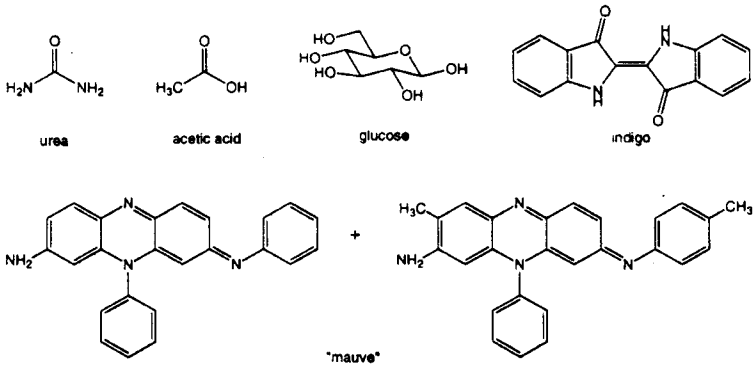


Fig. 3. Structures of molecules chemically synthesized in the nineteenth century

With the death of vitalism, nineteenth-century chemists began to explore the reactivity of molecules. Because little was known about what might happen when two different compounds were mixed—that is, what portions of the molecules might combine or split apart—most synthetic routes could not be planned. Wöhler eloquently communicated the difficulties in 1835: "Organic chemistry just now is enough to drive one mad. It gives me the impression of a primeval tropic forest, full of the most remarkable things, a monstrous and boundless thicket, with no way of escape, into which one may well dread to enter." Despite the limited knowledge of chemical reactions and the difficulties of isolating and characterizing organic molecules in the nineteenth century, a number of compounds were synthesized, more by prospecting and persistence than by design.[4] Notable achievements include Perkin's serendipitous synthesis of the first commercially successful non-natural dye, mauve, the generation of indigo, and Fischer's landmark synthesis of glucose in 1890 (Figure 3).

In comparison to acetic acid and urea, these molecules are structurally complex. At that time, naturally-occurring molecules were selected for synthesis and those who assembled them had little knowledge of the wide ranging reaction processes of organic molecules. Because most of these schemes for putting together molecules were unrefined, many interesting structures could not be synthesized and the deliberate creation of complex new molecules was generally impracticable. Consequently, a goal of practitioners of chemical synthesis became to demonstrate one's mastery of chemical architecture.

The first half of the twentieth century marked new developments in the creation of organic molecules. First, the discovery of many new reactions provided chemists with the armamentariam to synthesize complex structures. As a result, many complicated and biologically interesting molecules could be generated. In the 1950s, Robert Burns Woodward, one of the greatest molecule builders, suggested that the synthesis of molecules represented a new art. In an assessment of chemical synthesis, Woodward, said, "There is excitement, adventure, and challenge, and there can be great art, in organic synthesis. These alone should be enough."[5] These words captured the imaginations of a generation of organic chemists, who set out to prove that any molecule, no matter how complicated, could be made. In many ways, they succeeded. Chemists have synthesized ever more complex molecules until even a structure like palytoxin, a potent toxin from marine soft coral, was generated by chemical synthesis.[6]

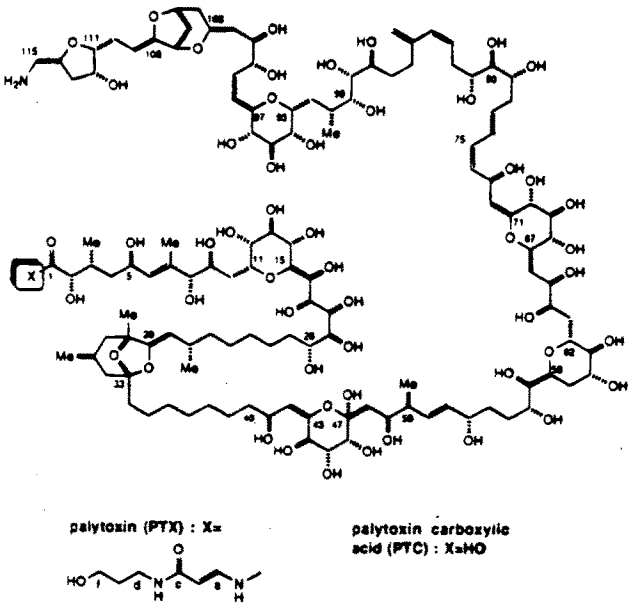


Figure 4. The natural product palytoxin, which Kishi and coworkers synthesized, is one of the most complex molecules every generated in a laboratory.

Since the time of Wöhler, chemists have struggled to gain control of molecular architecture. It is now widely believed that any specified target molecule can be synthesized; therefore, we have gained that control. What of this vision of synthesis as an art form? Those who make molecules can appreciate an unprecedented series of reactions just as an architect might admire a unique feature of a building. So perhaps synthesis qualifies as an art form, although the audience for each exhibition is small. If synthesis is to be widely appreciated, the product itself must be significant not solely the method by which it was obtained. The importance of synthesis in the creation of molecules with valuable properties is illustrated by the development of the pharmaceutical industry.

The beginnings of the pharmaceutical industry ushered in a new era in molecular design. In 1932 the synthesis of the first sulfanilamide agent, which was a pioneering drug that specifically targeted bacteria, prompted many companies to search for sulfa derivatives. By determining which analogs had anti-bactericidal properties, chemists uncovered the parts of the molecule essential for biological activity. The realization that molecules could be designed and synthesized with tailored properties altered the number of potential molecular targets. Nature was not alone in her ability to handcraft molecules with desirable activities. Chemists exploited her molecular designs as a springboard to create customized compounds. Efforts to create non-natural molecules with unique attributes drive much of the ongoing chemical research both in academia and industry. Consequently, creation of molecules involves both design and synthesis. Of these two, molecular design has emerged as a critical challenge.

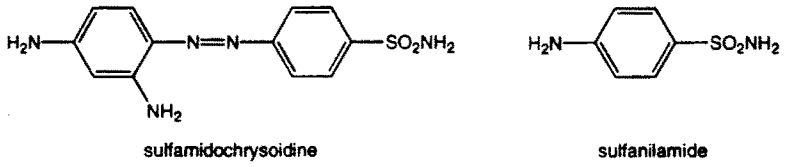


Figure 5. Sulfamidochrysoidine was the first selective anti-bacterial agent discovered. The compound was originally synthesized as a dye (red). Fourneau and coworkers then generated many variants, which revealed that the compound decomposed in vivo to sulfanilamide, the true bactericide.[7]

Chemists have learned to construct molecules with control of covalent bond formation, but what molecules should be constructed? Initially, Nature supplied the challenges for molecule synthesis, and chemists tried to mimic her ability to synthesize interesting, biologically active compounds. Yet, to create substances with new biological activities or specificities, we must produce molecules that do not naturally occur. Because we do not know how to specifically design each molecule to have the desired properties, we need an alternative strategy.

The new strategy is the addition of randomness into molecule creation. Scientists interested in biological systems first explored this possibility. They recognized that despite their sophistication, Nature is more efficient at creating molecules than were they. In contrast to the approach historically pursued by chemists, she does not focus on making a single molecule. For example, the immune system, which cannot know all its enemies, generates a multitude of different antibody molecules when faced with an invader. The immune system then increases production of the antibodies that most effectively combat the assault. Consequently, there is a need for randomness in biological processes. By appropriating Nature's tactics to achieve biological activity, that is to generate many molecules and choose the ones with the desired activity, scientists at the chemistry-biology interface developed synthetic plans with intellectual parallels to evolution. The new goal is not to make a single molecule with absolute control but rather to synthesize arrays of molecules from which to select for activity.[8] Scientists began deliberately to investigate methods to generate assortments of molecules in one operation. While most chemists were still refining their ability to synthesize target structures, those whose knowledge spanned chemistry and biology were exploring randomness as a molecular design consideration.

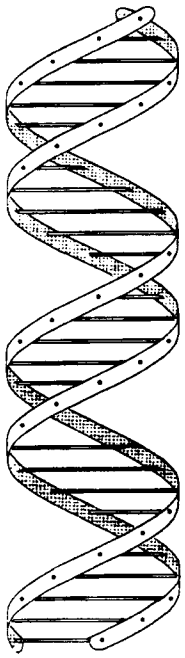


Figure 6. DNA duplex: ribbons indicate the DNA backbone held together by covalent bonds. Horizontal rods indicate non-covalent interactions between base pairs.

Biological approaches to molecular diversity emerged first. Chemical methods for the synthesis of DNA single strands were used to random generate collections of DNA molecules. DNA possesses the four nucleotide bases, A, G, C, and T, which serve as building blocks. By covalently linking 10 building blocks together, an assortment of 10^4 DNA molecules can be generated. Similarly large molecular assortments of peptides have been synthesized, which can be very diverse because there are 20 common amino acid building blocks. Random libraries of molecules have been used for a wide variety of applications, such as making pools of mutant proteins, elucidating the preferred DNA or peptide sequence for protein binding, and creating molecules that act as inhibitors of biological processes.

These applications underscore the value of randomness in molecule construction and the importance of designing synthetic routes that are amenable to it. For example, a critical unsolved problem in the area of biological diversity is the creation of carbohydrate libraries. A method for assembling different carbohydrates would allow scientists to define the biological roles of carbohydrates, which are less understood than are those of proteins or nucleic acids. Difficulties in making assortments of carbohydrates are a result of the idiosyncratic reactivites of individual saccharide units. Therefore, the problems associated with making carbohydrate libraries are rooted in the inability of chemists to assemble the units as effectively as Nature does. Hypothesizing that saccharide residues attached by non-natural linkages might have many of the same properties as carbohydrates, my research group is developing methods to fabricate carbohydrate analogs. Our strategy is to make arrays of carbohydrate-like molecules from which we can select derivatives with biological activity. By incorporating randomness into molecular design, we will explore completely new molecular arrangements and geometries.

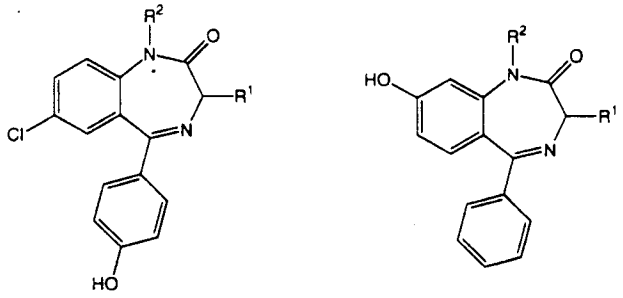


Figure 7. Ellman and coworkers synthesized a group of benzodiazepine molecules. Using the templates shown, 192 molecules were synthesized simultaneously by varying R^1 and R^2.

To investigate diversity in small molecule drug design, Ellman and coworkers created a molecular array or library of benzodiazepine derivatives.[9] The benzodiazepine class of molecules was chosen because they display a wide range of biological activities. For example, the commonly prescribed Valium and Librium exhibit anxiolytic, hypnotic, and muscle-relaxing activities. Medicinal chemists have continued to tinker with these structures to create molecules that inhibit a variety of proteins including enzymes from HIV[10] and those involved in cancer.[11] Consequently, the creation of a large assortment of benzodiazepines by one synthetic process has enormous ramifications for generating biologically relevant molecules. Ellman and coworkers synthesized simultaneously a collection of 192 molecules by varying the positions labeled R in Figure 7. They determined that 32 molecules in the library had inhibitory activity against the cholescystokinin receptor, a protein implicated in pancreatitis, ulcer conditions, and eating disorders. Therefore, in a single synthesis, Ellman and coworkers generated 32 molecules with biological activity.

Larger libraries of molecules built around different molecular templates are currently being fashioned. Many questions regarding the design of these molecular collages remain. What molecules should serve as the centerpiece? Should they be rigid or flexible? What kinds of covalent bonds can be formed? Future investigations should elucidate critical design and synthesis parameters.

The science of molecules, chemistry, is the field that creates new molecules that possess important properties. The synthesis and creation of organic molecules in the laboratory began with a serendipitous experiment in 1828. For over 160 years, we have worked to achieve maximum control in the construction of molecules. Most of these endeavors were aimed at becoming adept at chemically synthesizing naturally occurring molecules, especially those produced by organisms. Chemists have collected and refined the necessary tools for building such molecules. Now with these tools, we have looked toward Nature for new arenas in which to apply them. Biological systems have always taken advantage of randomness, but molecule-building chemists have fought against it. Only now are we beginning to interject entropy into our creations.

Notes

1 For an excellent introduction to structure and function of organic molecules, see P. W. Atkins, *Molecules* (New York: W. H. Freeman and Company, 1987), 197.

2 J. W. Cornforth, "The Trouble with Synthesis," *Aldrichimica Acta* 27:3 (1994), 71-7.

3 Wöhler's synthesis of urea raises the issue of how he determined what the product was. Interestingly, as a medical student Wöhler gained extensive experience isolating urea from dog urine; consequently, he was quite familiar with crystalline urea. For a discussion, see R. M. Roberts, *Serendipity: Accidental Discoveries in Science* (New York: John Wiley & Sons, 1989), 42-8.

4 I. Fleming, *Selected Organic Syntheses* (London: John Wiley & Sons, 1973), 227.

5 R. B. Woodward, "Synthesis," *Perspectives in Organic Chemistry*, ed. A. Todd (New York: Interscience Publishers, Inc., 1956), 158.

6 E. Suh and Y. Kishi, "Synthesis of Palytoxin from Palytoxin Carboxylic Acid," *The Journal of the American Chemical Society* 116 (1994), 11205-6.

7 W. H. Brock, *The Norton History of Chemistry*, ed. R. Porter, The Norton History of Science (New York: W. W. Norton, 1992), 620-39.

8 See M. A. Gallop et al., "Applications of Combinatorial Technologies to Drug Discovery. 1. Background and Peptide Combinatorial Libraries," *Journal of Medicinal Chemistry* 37 (1994), 1233-51 and E. M. Gordon et al., "Applications of Combinatorial Technologies to Drug Discovery. 2. Combinatorial Organic Synthesis, Library Screening Strategies and Future Directions," *Journal of Medicinal Chemistry* 37 (1994), 1385-401.

9 B. A. Bunin, M. J. Plunkett, and J. A. Ellman, "The combinatorial synthesis and chemical and biological evaluation of a 1,4-benzodiazepine library," *Proceedings of the National Academy of Sciences, USA* 91:11 (1994), 4708-12.

10 See R. Pauwels et al., "Potent and selective inhibition of HIV-1 replication in vitro by a novel series of TIBO derivatives," *Nature* 343 (1990), 470-4 and M.-C. Hsu et al., "Inhibition of HIV Replication in Acute and Chronic Infections in Vitro by a Tat Antagonist," *Science* 254 (1991), 1799-802.

11 G. L. James et al., "Benszdiazepine Peptidomimetics: Potent Inhibitors of Ras Farnesylation in Animal Cells," *Science* 260 (1993), 1937-42.