VOLUME 74, NUMBER 7 JOC Perspective April 3, 2009

© Copyright 2009 by the American Chemical Society

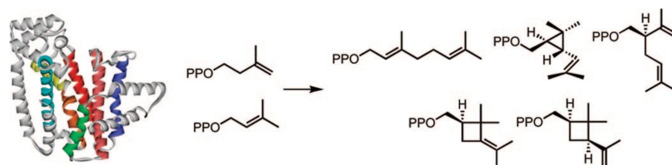
Bioorganic Chemistry. A Natural Reunion of the Physical and Life Sciences

C. Dale Poulter*

Department of Chemistry, University of Utah, 315 South 1400 East RM 2020, Salt Lake City, Utah 84112

poulter@chemistry.utah.edu

Received January 27, 2009



Organic substances were conceived as those found in living organisms. Although the definition was soon broadened to include all carbon-containing compounds, naturally occurring molecules have always held a special fascination for organic chemists. From these beginnings, molecules from nature were indespensible tools as generations of organic chemists developed new techniques for determining structures, analyzed the mechanisms of reactions, explored the effects conformation and stereochemistry on reactions, and found challenging new targets to synthesize. Only recently have organic chemists harnessed the powerful techniques of organic chemistry to study the functions of organic molecules in their biological hosts, the enzymes that synthesize molecules and the complex processes that occur in a cell. In this Perspective, I present a personal account of my entrée into bioorganic chemistry as a physical organic chemist and subsequent work to understand the chemical mechanisms of enzyme-catalyzed reactions, to develop techniques to identify and assign hydrogen bonds in tRNAs through NMR studies with isotopically labeled molecules, and to study how structure determines function in biosynthetic enzymes with proteins obtained by genetic engineering.

As its name suggests, organic chemistry emerged in the early nineteenth century as a branch of chemistry concerned with substances isolated from living organisms. The field soon expanded, however, to include carbon-containing molecules more generally, and chemists began to study the structures, physical properties, reactions, and chemical transformations of organic compounds, many of which were not obtained from nature. By the mid-twentieth century, most of the research in organic chemistry was not concerned with biological systems. Biology, on the other hand, focused largely on the morphologies and behaviors of organisms. Interpretation of biological phenomena at the molecular level was still in its infancy. Before that time, neither chemistry nor biology was sufficiently mature to nurture the other, and the research in each field was largely separate from the other. Thus, the great advances in both chemistry and biology before 1950 did not depend on insights from the other discipline.

Today, however, the former artificial boundaries between organic chemistry and biology have been blurred as scientists in each area are quick to adopt the knowledge and techniques of the other. Natural products chemists analyze newly sequenced genomes for clues to previously undiscovered biologically active molecules and new biosynthetic pathways. Biologists identify molecules that regulate signaling events during cellular development and govern intereactions among and between species. Synthetic organic chemists rely on biological assays to guide their design of molecules that bind tightly to enzymes and receptors. And in my own research and that of others, the techniques of physical organic chemistry and molecular biology are applied to enzymes and other large biomolecules to gain an

10.1021/jo900183c CCC: $40.75 © 2009 American Chemical Society
Published on Web 03/27/2009

understanding of their structures, the mechanisms of the reactions they catalyze, and their functions at levels previously accessible only for small molecules. All of this research could not have taken place without a "reunion" of organic chemistry and biology.

I am honored by the invitation from the Organic Division of the ACS to present a personal account of my research at the interface between chemistry and biology at the Centennial Symposium at the National ACS meeting in Philadelphia last year. That presentation serves as the template for this Perspective. I begin with a brief history that touches on the emergence of natural products, reaction mechanisms, synthesis, conformational analysis, and biosynthesis as subdisciplines of organic chemistry. In each of these areas, isoprenoid compounds, molecules synthesized by a biosynthetic pathway that been a major focus of my research program, were important in their development. What then follows is the sequence of events that led me into the field of bioorganic chemistry and the research they inspired.

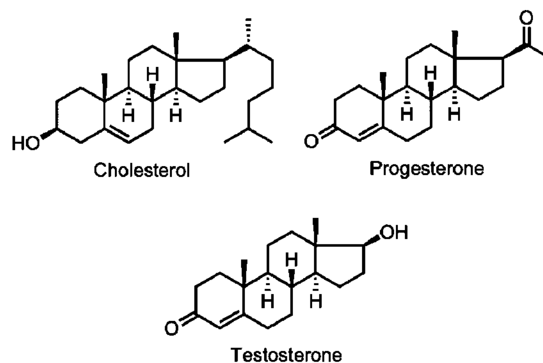
Historical Connections

Jon Berzelius was the first to define organic substances, although he only experimented briefly with organic molecules. In 1807, he was appointed Professor of Chemistry and Pharmacy at the Karolinska Institute and, in the same year, proposed that substances found in organisms, such as olive oil and sugar, be called "organic," while those characteristic of nonliving environments, such as water and salts, be called "inorganic."[1] During his career, Berzelius introduced terms such as isomerism, catalysis, protein, and polymer, all of which are found in today's organic chemistry textbooks. In 1836, he was presented with the Copley Medal of the Royal Society of London, an award that predates the Nobel Prize by 170 years and lists Michael Faraday, Joseph Priestly, James Cook, Charles Darwin, Louis Pasteur, Albert Einstein, James Watson, and Francis Crick among its receipents.

The definition of organic materials changed in a fundamental way in 1861 in a textbook written by August Kekule. He defined organic compounds as substances that contain carbon, although his definition conveniently overlooked compounds like calcium carbonate and sodium cyanide, which even today are considered "inorganic." Kekule, a brilliant theoretician, suggested that carbon was tetravalent and proposed correct structures for a variety of simple organic molecules, including his famous structure for benzene.

In 1879, Kekule convinced the administration at the University of Bonn to appoint Otto Wallach, a Privatdozent in his department, as Chair of the Department of Pharmacology.[2] Up to then, Wallach had dabbled in a variety of areas without a clear focus to his research program. Out of concern for his young colleague's career, Kekule showed Wallach a cabinet full of bottles of essential oils from a variety of plants and suggested that he could build a successful career by studying their contents. Over the years Wallach isolated and determined the structures of substances from the bottles, including a new group that he named "terpenes" from their presence in turpentine. The monoterpenes that Wallach worked with are simple by today's standards, but remember, the principal tools available to him were distillation, combustion analysis, chemical degradation, and a keen mind. In 1910, Wallach, a student of Friedrich Wohler who in turn was a student of Berzelius, was awarded the Nobel Prize in Chemistry "in recognition of his services to organic chemistry and the chemical industry by his pioneer work in the field of alicyclic compounds", a successful career indeed!

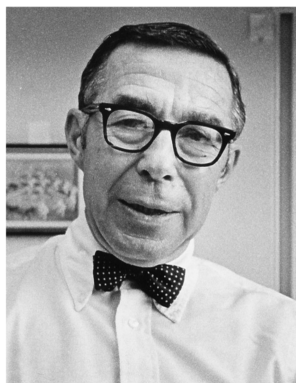
Wallach's pioneering work to establish the structures of natural occurring molecules was only the first of many instances where terpenoid or, more generally, isoprenoid molecules were associated with major developments in organic chemistry. Mechanistic organic chemistry and the recognition of carbocations as reactive intermediates grew from the discovery of the rearrangement of camphene hydrochloride to isobornyl chloride by Georg Wagner and subsequent studies by Hans Meerwein.[3] The elegant total synthesis of the natural enantiomers of the isoprenoid metabolites cholesterol, testosterone, and progesterone published by R. B. Woodward and his co-workers in 1952[4] was one of the early examples that inspired subsequent generations of organic chemists to synthesize molecules of ever increasing complexity. Many of the techniques now used to construct bonds, manipulate functional groups, and control stereochemistry were developed in the 1960s and 1970s as sesqui- and diterpenes were synthesized in the laboratories of Gilbert Stork, Clayton Heathcock, E. J. Corey, James Marshall, Samuel Danishefsky, and many others. Woodward and Corey received Nobel Prizes in Chemistry in 1965 and 1990, respectively, for their achievements in organic synthesis. The stereochemistry of reactions involving the tetracyclic sterol nucleus inspired Derek Barton to develop field of conformational analysis to first explain and then predict the stereochemical course of reactions.[5] He shared the 1969 Nobel Prize in Chemistry with Odd Hassel.



In the early 1950s, the development of increasingly sophisticated methods to detect and determine structures of organic molecules were applied to biological systems. Konrad Bloch[6] and Feodor Lynen[7] pioneered the use of radioisotopes in their studies that led to their discovery of the mevalonate pathway for biosynthesis of cholesterol from acetic acid. Their achievements were recognized by the 1964 Nobel Prize in Medicine. Although the connection with medicine may have been tenuous at the time, development of the statin family of drugs now used to treat heart disease can be traced directly to their work. Cholesterol was also the featured molecule in the groundbreaking work by John Cornforth, who used isotopes of hydrogen to break the stereochemical degeneracy of achiral tetrahedral carbon. He established the stereochemistry of the reactions for biosynthesis of squalene, a key precursor of cholesterol that has no chiral centers, from mevalonic acid.[8] Cornforth, "for his work on the stereochemistry of enzyme catalyzed reactions," shared the 1975 Nobel Prize in Chemistry with Vladimir Prelog.

As connections between genes and nucleic acids, enzymes and proteins, were established by geneticists and biochemists in the 1930s and 1940s, a group of chemists and physicists

became increasingly interested in the structures of these substances with the ultimate goal of understanding fundamental biological processes at the molecular level. Their work exploded on the scientific community in the 1950s with the structure of the α-helix in proteins by Pauling in 1951,[9] the double-helical structure of DNA in 1953 by Francis Crick and James Watson,[10] and the structure of the globular protein myoglobin in 1958 by John Kendrew.[11] Pauling was awarded the Nobel Prize in Chemistry in 1954. In 1962, Crick and Watson received the Noble Prize in Medicine. Kendrew, along with Max Perutz, received the Nobel Prize in Chemistry the same year.



William G. Dauben

Saul Weinstein

College and Graduate School

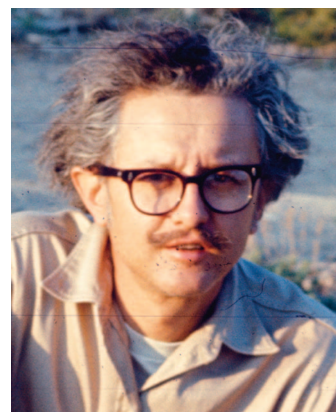
The advancements in the new field of molecular biology of the 1950s were not part of the chemistry curriculum at LSU when I enrolled as an undergraduate in 1960, nor were they in any of the textbooks in organic chemistry. There was little room for and no encouragement to to take courses in biology or biochemistry. The next four years were filled with mathematics, physics, German, and of course, chemistry, all of which were required for my B.S. degree. This curriculum was typical for chemistry departments across the country. When I enrolled in graduate school at Berkeley, only three courses were required for a Ph.D. degree, two in organic and one in physical chemistry. Graduate students interviewed for positions in research groups within a few weeks after first arriving and were expected to spend any time not devoted to classes or teaching in the laboratory. I was immersed in organic chemistry for the next 3 years. The Biochemistry Department, which I never visited, was about a 10-min walk away.

I did not have a clear idea, perhaps more to the point no idea at all, of what kind of research I wanted to do for my thesis when I interviewed with the organic faculty. Bill Dauben described a project involving organic photochemistry that sounded exciting, and I bit. In retrospect, I could have not made a better choice. Dauben was the perfect mentor for me. He was available for consultation and advice but left it to me to plan and execute my experiments. When Dauben appeared at the door of my laboratory, I was never sure if we were going to talk about research or baseball. The research project was also perfect for developing my knowledge and skills as an organic chemist. Photochemistry was just coming into prominence with the discovery of new reactions and new, sometimes exotic, products. My project, the photochemistry of s-trans dienes, involved a mixture of organic synthesis, isolation and identification of products, often on small scales, mechanistic organic chemistry, and photophysics, with a sufficient number of exciting surprises to more than compensate for the inevitable frustrations of research.

I began a two-year stay at UCLA with Saul Winstein in 1967, as an NIH Postdoctoral Fellow. Every day in Winstein's laboratory was interrupted at some point with "How goes it?" This was an invitation to talk about chemistry. In my case it was usually about experiments involving the direct observation of carbocations in "super acid" solvents and studying rearrangements not normally seen in nucleophilic solvents. Winstein was always eager to talk about what had just happened in the laboratory but, like Dauben, left the planning and execution of my research to me, although both Dauben and Winstein had high expectations about how the experiments were planned, conducted, and described in group meetings and reports. Saturday during basketball season was a prime time for postdocs to be in the laboratory. Winstein had season tickets at midcourt. If he was not going to a home game, he often gave his tickets to one of the postdocs in the laboratory on Saturday. And these were not ordinary tickets! The years from 1967 to 1969 marked the middle of UCLA's basketball dynasty, where they won 9 NCAA championship games in 10 years!

My Entrée into Bioorganic Chemistry

I started to look for an academic position in 1968, intending to establish a research program in the general area of mechanistic photochemistry. During the afternoon of my second day of interviews at the University of Utah, I was ushered into Bill Epstein's office. Epstein was a natural product chemist interested



Bill Epstein

in compounds found in plants native to Utah. He did not begin with the usual questions about my research ideas and teaching preferences but went to the black board and drew the two reactions shown in Scheme 1. He then said that he and Hans Rilling, a biochemist at Utah, had just determined the structure of an intermediate between farnesyl diphosphate and squalene, subsequently named presqualene diphosphate, that Rilling had discovered a few years earlier.[12] Epstein asked me to propose a mechanism for rearrangement of presqualene diphosphate to squalene. My answer was based on work with cyclopropylcarbinyl cations that Jack Roberts, Marjorie Caserio, and their colleagues had published several years before.[13,14] Epstein then asked me to predict the stereochemistry of the reaction. This was more difficult because I could see four possibilities, depending on how presqualene diphosphate and NADPH were positioned in the active site. I plunged ahead and proposed the

SCHEME 1. Reactions Catalyzed by Squalene Synthase

SCHEME 2. Mechanism for the Rearrangement of Presqualene Diphosphate to Squalene

mechanism shown in Scheme 2, unaware that the question could not be answered because the absolute stereochemistry of presqualene diphosphate had not been established. A few years later when George Popjak showed that (1*R*,2*R*,3*R*)-presqualene diphosphate is the natural enantiomer,[15] I was pleasantly surprised to find that the "most probable" of the four scenarios was the one that fit the data. The interview with Epstein and the many discussions we had after I joined the Chemistry Department at the University of Utah stimulated my interest about how enzymes catalyze the fundamental reactions nature uses to construct isoprenoid molecules, an interest that continues today. Bill Epstein is a master teacher, a valued mentor, and a friend who also introduced me to the pleasures and frustrations of fly fishing.

Modeling Enzyme-Catalyzed Reactions. Within a few months after moving to Utah as a new Assistant Professor, I began a series of studies with derivatives of (1*R*,3*R*)-chrysanthemol, a monoterpene model for (1*R*,2*R*,3*R*)-presqualene diphosphate, to see if the cyclopropylcarbinyl chemistry I proposed for the rearrangement of presqualene diphosphate to squalene, including stereochemistry, was feasible without enzyme catalysis. The eventual answer was yes, but not very efficiently.[16−18] When those experiments were carried out, there were no known natural monoterpenes with carbon skeletons corresponding to the tertiary cyclopropylcarbinyl cation and squalene, both of which we found in small amounts in the model studies (see Scheme 3). Later, Epstein and his students discovered examples of both in the essential oil of *Artemisia tridentata* ssp. spiciformis, also known as snowfield sagebrush, a plant common in the mountains of Utah.[19] The chrysanthemyl cation also rearranged to give two other carbon skeletons that corresponded to known monoterpenes. Thus, solvolysis of (1*R*,3*R*)-chrysanthemyl derivatives gave products with four of the eight "non-head-to-tail" carbon skeletons found in nature, and their absolute stereochemistries matched those of the related naturally occurring compounds. Since "non-head-to-tail" is not very specific, we introduced a way to describe the carbon skeletons by numbering the carbon atoms in individual isoprene units and specifying which atoms in one isoprene unit are joined to those in another.[16]

SCHEME 3. Products from Solvolysis of *N*-Methyl-4-chrysanthemyloxypyridinium Iodide and Naturally Occurring Irregular Monoterpenes with Related Carbon Skeletons (Blue)

Enzymes and Mechanisms. My work with model reactions drove home the realization that I needed to learn how to purify and assay enzymes if I wanted to study how they worked. Although research with enzymes was uncommon in chemistry departments in the early 1970s, there were a few chemists to look to for inspiration. John Cornforth was just completing his elegant stereochemical studies of cholesterol biosynthesis, where he distinguished among 16384 possibilities in establishing the stereochemistry for biosynthesis of squalene from isotopically labeled (*R*)-mevalonic acid by six enzymes (see Figure 1).[8] He concluded this study with the most challenging set of experiments of all, determining the absolute stereochemistry of the chiral methyl groups, labeled with hydrogen, deuterium, and tritium, in squalene.[8,20] Bill Jencks, a masterful scientist and a true gentleman, had just written a book, *Catalysis in Chemistry and Enzymology*, that was to become the classic reference in

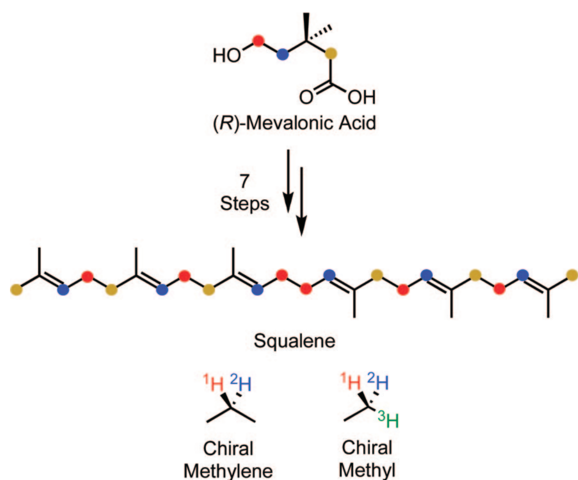


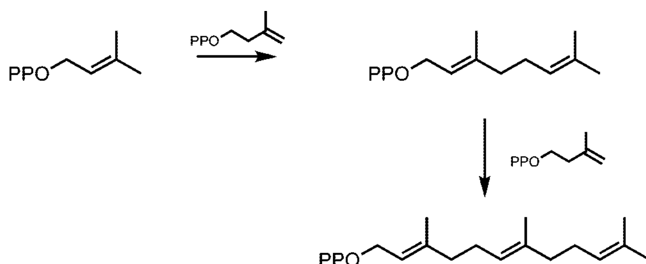
FIGURE 1. Chiral centers in squalene synthesized from labeled (*R*)-mevalonic acid.

the area.[21] Bob Abeles was using fundamental ideas from mechanistic organic chemistry to design compounds that were converted to potent irreversible inhibitors by the targeted enzyme.[22] He named these inhibitors "suicide substrates", explaining that the enzyme committed suicide when attempting to carry out its normal catalytic function. At an Enzyme, Coenzymes, Mechanisms Gordon Conference in the late 1970s where Abeles was speaking, Jeremy Knowles, never hesitant to inject a little humor into a discussion, raised the question whether the enzymes committed suicide using the inhibitors or were murdered by them. The question of whether enzymes or inhibitors were capable of having such intents, issues of importance in criminal trials, was not resolved.

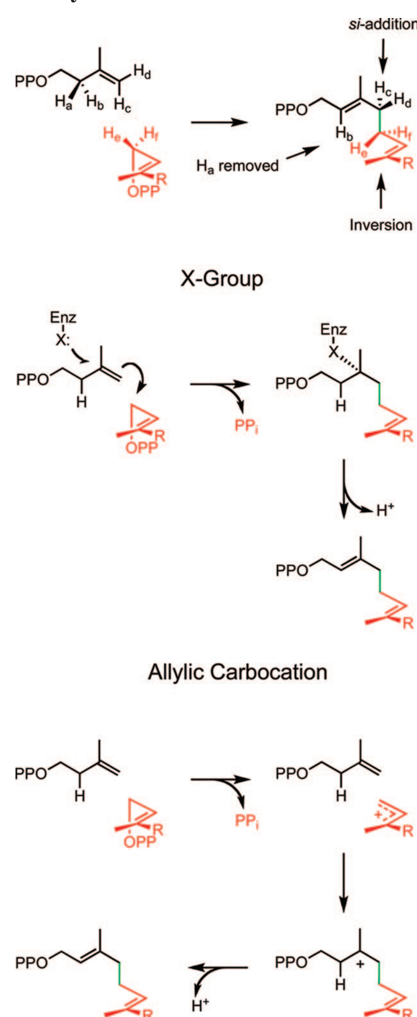
My training in enzymology was made possible by a Research Career Development Award from the National Institutes of Health, which allowed me to return to the laboratory, and by Hans Rilling, who had discovered presqualene diphosphate.[12] Rilling was a superb experimentalist who actively worked in the laboratory throughout his career. We worked side-by-side for a year in his laboratory in the University of Utah Hospital, where he introduced me to the pleasures of homogenizing chicken livers in a blender and taught me how to purify enzymes, assay their activities, and safely synthesize radioactive compounds. This experience gave me the confidence to bring the new techniques into my own group with the knowledge that I could help my students with problems that they encountered with their own projects. I followed the same approach when learning the techniques for cloning during a sabbatical leave several years later.

The Prenyl Transfer Reaction. The first enzyme I tackled was farnesyl diphosphate synthase. It catalyzes two rounds of chain elongation by adding the isoprene unit in dimethylallyl diphosphate to isopentenyl diphosphate to give geranyl diphosphate, followed by addition of the geranyl group to a second molecule of isopentenyl diphosphate to give farnesyl diphosphate (Scheme 4). Cornforth had observed that these reactions proceeded with inversion at the C(1) carbon of the allylic substrates and that the addition of the alkyl groups of dimethylallyl diphosphate and geranyl diphosphate to C(4) of isopentenyl diphosphate and elimination of the proton from C(2) occurred from the same side of the molecule. To account for the stereoselectivity of the addition−elimination reactions, he proposed the "X-group" mechanism, where formation of a carbon−carbon bond between the two substrates is synchronous with *trans* addition of a nucleophilic "X" group in the active site of the enzyme to C(3) of the isopentenyl residue, followed by a *trans* elimination of "X" and the proton at C(2) (see Scheme 5).[8]

SCHEME 4. Synthesis of Farnesyl Diphosphate from Dimethylallyl Diphosphate



SCHEME 5. Stereochemistry for Chain Elongation by Farnesyl Diphosphate Synthase: X-Group and Allylic Carbocation Alkylation Mechanisms



Although logical, it seemed to me that the intervention of an X-group was not required. The barrier for rotation about the C1−C2 bond in the dimethylallyl cation is sufficiently high that one would not expect it to compete with alkylation of the double bond in isopentenyl diphosphate. Thus, racemization at C(1) of the dimethylallyl unit would only be seen if the entire rigid allylic cation rotated within the active site of the enzyme. The observed stereochemistry could result from the orientation of

the substrates in the active site rather than a consequence of stereoelectronic features of the reaction. In addition, the X-group would have to perform a delicate juggling act by serving as a nucleophile in the addition step and a powerful leaving group in the elimination step. Application of "Occam's razor," a favorite argument used by Saul Winstein, favored a simple three-step sequence—formation of an electrophilic allylic carbocation, alkylation of the double bond in isopentenyl diphosphate by the allylic cation, and elimination of a proton to generate the C2−C3 double bond in the product—with the stereochemistry of the reaction dictated by how the substrates were bound in the active site. The limitations of using stereochemistry to support mechanistic arguments for enzyme-catalyzed reactions are recognized more clearly now than 40 years ago.

This hypothesis required experimental support. We reasoned analogues of isopentenyl diphosphate with fluorine at C2 would be suicide substrates for enzymes that catalyzed chain elongation by an X-group mechanism by blocking the final elimination step and leaving the X-group permanently alkylated. 2-Fluoro- and 2,2-difluoroisopentenyl diphosphate were competitive inhibitors for isopentenyl diphosphate, but neither compound irreversibly inhibited the enzyme.[23] In addition, 2-fluoroisopentenyl was an alternate substrate when incubated with geranyl diphosphate to give 2-fluorofarnesyl diphosphate.

We obtained evidence that the reaction was an electrophilic alkylation by measuring k_{cat} for the chain elongation by replacing hydrogen atoms in the methyl group at C(3) of the allylic substrate by fluorine, which should decrease in the rate of the reaction due to the destabilizing effect of the strongly electronegative fluorine atoms.[24,25] The most definitive results were obtained for the series of allylic compounds shown in Figure 2, where Y = CH_3, CH_2F, CHF_2, and CF_3. The effect of the

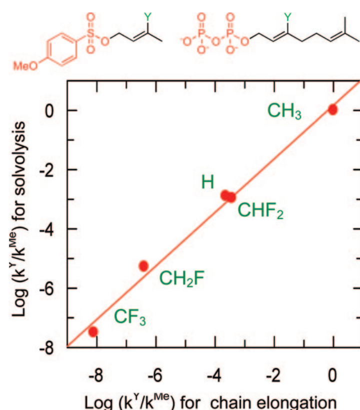


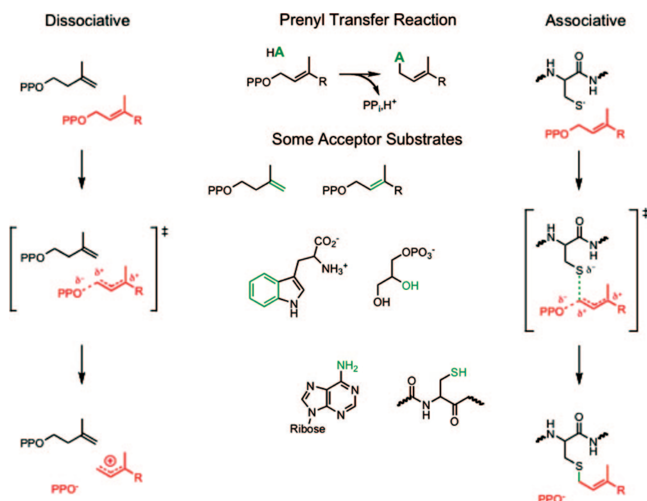
FIGURE 2. Hammett plot of relative rates for solvolysis of allylic methoxymethanesulfonates versus chain elongation of 3-substituted geranyl diphosphates by farnesyl diphosphate synthase, slope = 0.9.

fluorine was calibrated by measuring the solvolysis rates of aryl sulfonate derivatives of the allylic substrates and comparing these results with those of enzyme-catalyzed chain elongation for the corresponding diphosphate derivatives by avian farnesyl diphosphate synthase. A Hammett plot of the rates for solvolysis of the *p*-methoxybenzene sulfonate derivatives in 9:1 dioxane: water gave an excellent linear correlation with $\sigma^+ = -8.0$, while a plot of $\log(k^Y/k^{Me})_{solvolysis}$ versus $\log(k^Y/k^{Me})_{chain\ elongation}$ also gave an excellent linear correlation over 7 log units with a slope of 0.9.[26] Thus, the enzyme-catalyzed reaction is slightly more sensitive to the fluorinated substituents than the model reaction. The location of the fluorine atoms in the methyl group at C3 indicates that positive charge is delocalized to that location during the reaction. The linear log−log correlation between the rates of the model and enzyme-catalyzed reaction, with a slope near unity, is consistent with a dissociative electrophilic alkylation where the allylic cation is an intermediate that alkylates the double bond of isopentenyl diphosphate.

These experiments required substantial quantities of pure enzyme because of large amounts required to see turnover with the difluoro- and trifluoromethyl derivatives. Trips to the slaughterhouse to get fresh chicken livers became a "rite of passage" for new graduate students in the group, as did long days in the cold room for the 2 weeks required to convert 2 kg of livers to about 1 mg of pure enzyme.

Electrophilic alkylation of the double bond in isopentenyl diphosphate is only one example of a much larger family in the isoprenoid pathway that fall under the generic title of prenyl transfer reactions. They share the common feature that the allylic moiety from an electrophilic allylic diphosphate ester is added to an electron rich acceptor. In addition to carbon−carbon double bonds, acceptors include aromatic rings,[27] hydroxyl groups,[28] amino groups,[29] and sulfhydryl groups.[30] We have studied representative enzymes that process substrates with all of the different electron-rich acceptors. Although all prenyl transfer reactions are electrophilic alkylations, the timing of bond breaking and bond formation changes from a dissociative process for alkylation of carbon−carbon double bonds by farnesyl diphosphate synthase to an associative reaction for alkylation of the highly nucleophilic cysteine zinc thiolate by protein prenyltransferases (see Scheme 6). Thus, the prenyl

SCHEME 6. Prenyl-Transfer Reaction with Nucleophilic Moieties Shown in Green and Dissociative and Associative Mechanisms with Allylic Electrophiles Shown in Red



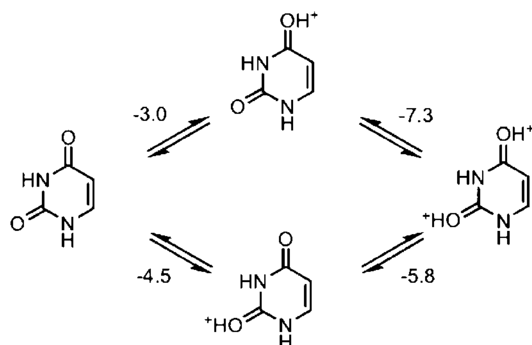
transfer enzymes belong to a "mechanistic superfamily" that catalyzes electrophilic alkylations.

NMR Studies of Isotopically Labeled tRNA. One of the great joys of for me as a professor at a research university is the freedom to follow my instincts wherever they lead, provided of course I can convince the reviewers of my proposals that the projects are worthwhile. I think that more often than not the end of the path one eventually takes during basic research is not at all obvious at the beginning. This was certainly the case when we began to study uracil. Although the 2,4-oxypyrimidines uracil and thymine can each potentially exist in one or more of six different tautomeric forms, the diketo

form is substantially more stable that the isomeric mono- and dienols. Had Watson and Crick not eventually realized that they were using the incorrect enolic tautomers of thymine, cytosine, and guanine in their models, they may well have lost their race with Linus Pauling's laboratory to determine the structure of DNA.[31]

Our interests in uracil were more mundane. We began with a study of the structure of the dication of uracil in "superacid" media[32] and measurements of the pK_a's for mono- and diprotonated uracil in strongly acidic solvents using ^{1}H NMR chemical shifts (Scheme 7).[33,34] It occurred to me that the

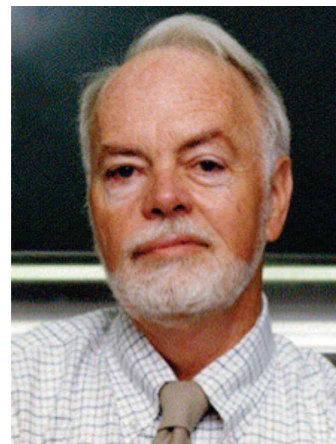
SCHEME 7. pK_a's for First and Second Protonations of Uracil



enhanced chemical shift dispersion seen for ^{13}C and ^{15}N might be advantageous for studying hydrogen-bonding interactions between nucleic acids. Given our interest in uracil, we initially decided to study hydrogen bonding between derivatives of uridine and adenonine. It had been reported that the ^{13}C resonances for C2 and C4 of the uracil moiety showed substantial downfield shifts upon hydrogen bonding with adenonine;[35] whereas, the ^{15}N resonance for N3 was insensitive.[36] This observation for nitrogen seemed to be at odds with the relatively small amount of chemical shift information available for that nucleus. At that time 100 MHz spectrometers were the highest field commonly available and pulsed-NMR was in its infancy. Sensitivity was an issue for the ^{13}C and ^{15}N measurements, so we decided to synthesize labeled versions of uracil for the experiments. The procedures we developed allowed us to label the individual carbon and nitrogen atoms in the uracil regioselectively.[37] NMR measurements and chemical shift assignments were straightforward with these compounds, and we demonstrated that ^{15}N NMR spectroscopy could be an important tool for studying hydrogen bonding in nucleic acids.[38]

^{15}N-Labeled tRNA. Beginning in the mid-1970s, I frequently talked to Jim McCloskey, a mass spectroscopist in the Medicinal Chemistry Department at Utah, about research. McCloskey, in collaboration with Susumu Nishimura at the National Cancer Institute in Tokyo, was identifying and characterizing modified bases in tRNA by mass spectrometry. About this time papers describing ^{1}H NMR studies of tRNAs began to appear in the literature.[39] The regions for the sugar and base protons were too complex to be of much value. However, the imino hydrogens involved in intramolecular hydrogen bonds that exchanged slowly with water were seen in the low field region from 10−15 ppm. Although most of these resonances overlapped, a few appeared as resolved peaks. As I read these papers, I kept thinking that many of the problems associated with overlapping resonances and assigning peaks to specific structural features in tRNA could be simplified by selectively labeling the molecules with ^{15}N.

The problem was how to obtain suitably labeled tRNA. I had filed this idea along with others in the "very interesting but



Jim McCloskey

impractical" category until a party at McCloskey's house during a visit by Nishimura. In the course of the evening, I talked to Nishimura about my thoughts for NMR experiments with labeled tRNAs. Nishimura supplied McCloskey with samples of pure tRNAs as part of their collaboration to identify new modified bases. He had large fermentors for growing *E. coli* and a strain that was auxotrophic for uracil. By the end of the party, we had agreed to collaborate. My group synthesized several grams of ^{15}N-labeled uracil and sent the material to Tokyo. Nishimura's group conducted large-scale incubations with his uracil-requiring auxotroph, purified the more abundant tRNAs, and send a few milligrams of the labeled tRNAs to us for the NMR studies. The labeling patterns for tRNA^{fMet} are shown in Figure 3.

The first samples arrived early 1982 and shortly afterward, with assistance from Ralph Hurd at Nicolet, we had ^{1}H spectra from his "state-of-the-art" 500 MHz FT instrument. Imino protons attached to ^{15}N were easily identified because of the ~90 Hz ^{1}H−^{15}N coupling constants (see Figure 4), which resulted in "triplet patterns" from a combination of the ^{1}H−^{15}N doublets and singlets from residual unlabeled tRNA. We were able to immediately confirm some assignments, show that others previously thought to be "reliable" were incorrect, and make several new assignments.[40] However, the real power of our approach was the ability to detect signals in congested regions of the spectrum. The ^{1}H resonances for protons in ^{1}H−^{15}N units were easily seen in a difference spectrum for tRNA^{fMet}, where FIDS with irradiation at frequencies corresponding to ^{15}N chemical shifts were subtracted from FIDS without irradiation (see Figure 5).[41] The signals for imino protons attached to unlabeled nitrogen atoms canceled, leaving only negative peaks flanked by less intense positive peaks for protons attached to ^{15}N when the single peaks in the decoupled spectra were subtracted from the trio of peaks in the coupled spectra. This technique was refined to obtain ^{15}N chemical shifts by collecting difference spectra as the decoupling frequency was changed in small increments at a reduced power setting. The difference decoupling technique could also be used to identify nearby protons from NOE enhancements. Thus, we were able to assign the imino protons in a GU wobble interaction that were previously buried in a congested region of the spectrum.[42]

The difference decoupling experiments were conducted at the newly established regional NSF NMR facility at Colorado State

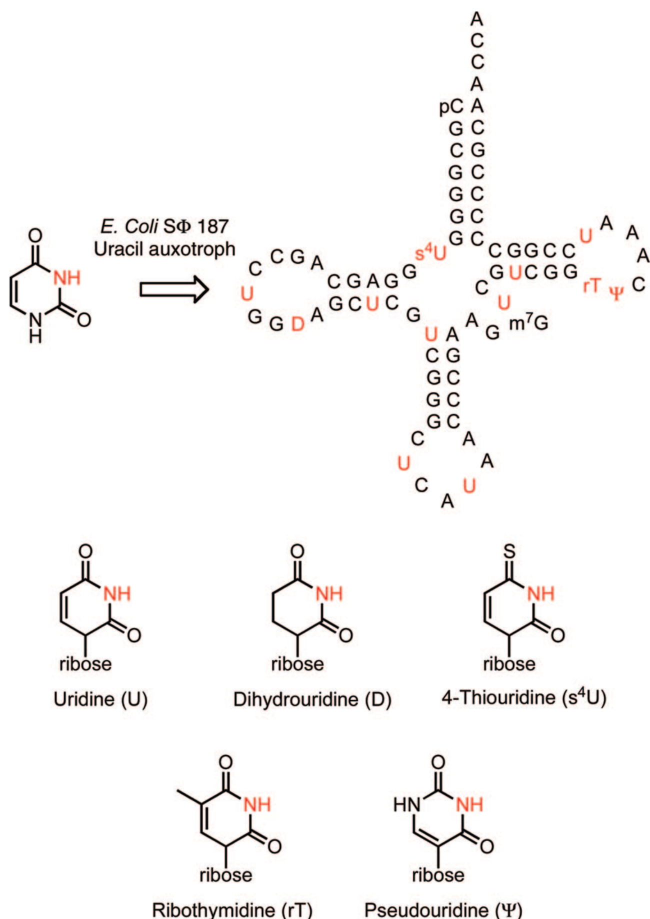


FIGURE 3. Sites of incorporation of labeled uracil into uridine and modified uridine bases in tRNA^{fMet}.

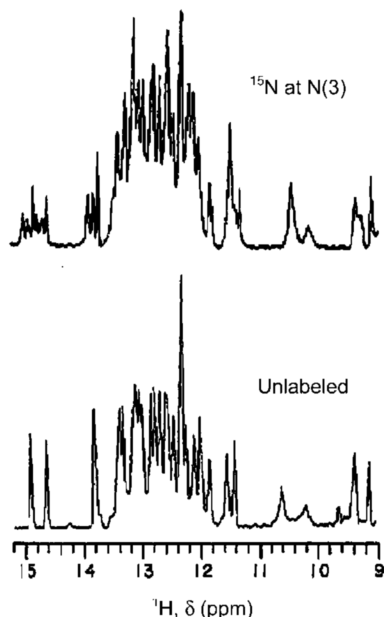


FIGURE 4. ^{1}H NMR spectra of ^{15}N-labeled and unlabeled *E. coli* tRNA^{fMet}.

University managed by Ralph Hurd and Ad Bax. Rich Griffey, a graduate student in my group, would take samples to the center and return with the 1-D spectra. The resolution of proton resonances in the difference decoupling experiments was limited by the magnitude of the ^{1}H$-$^{15}N coupling constants in the proton-nitrogen pairs. Bax encouraged Griffey to overcome this limitation by obtaining a ^{1}H$-$^{15}N 2-D spectrum. Two-dimensional chemical shift correlation spectroscopy had been reported several years earlier but not for a high molecular weight molecule. During the next few visits to Colorado State, Bax provided pulse sequences for Griffey to try without success, and he returned to Utah with more 1-D data. Then Griffey returned with the multiquantum 2-dimensional ^{1}H$-$^{15}N NMR spectrum of the labeled units in *E. coli* tRNA^{fMet} shown in Figure 6, which gave ^{1}H and ^{15}N chemical shifts at proton sensitivity as the only signals in the spectrum.[43] This was the first demonstration of the multi quantum heteronuclear 2-D correlation technique with a biopolymer, the procedure now widely used for structural analysis of proteins and nucleic acids in solution.

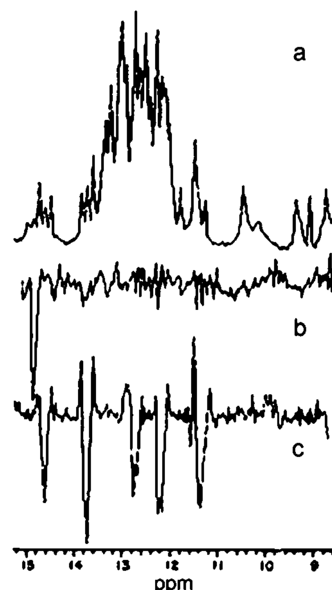


FIGURE 5. ^{1}H NMR spectrum for tRNA^{fMet} (part a). Difference-decoupled spectrum with irradiation at 36.487132 MHz (part b) and 36.486450 MHz (part b).

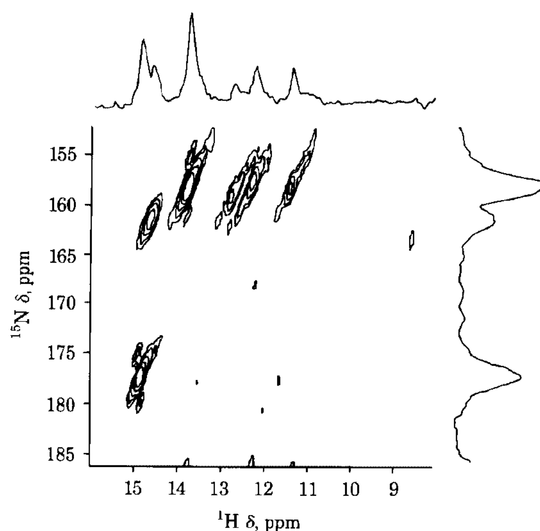


FIGURE 6. ^{1}H$-$^{15}N multiquantum 2D spectrum of *E. coli* tRNA^{fMet}.

Cloning. The NMR experiments succeeded beautifully, but obtaining "NMR quantities" of pure labeled tRNAs was a real

"tour de force." Our experiments consumed large amounts of labeled uracil, required large-scale fermentations and chromatographies, and were practical for only the most abundant tRNAs. During the 1987−88 academic year, I spent a sabbatical leave in Ray Gesteland's laboratory in the basement of the Biology building at the University of Utah learning how to cut and paste DNA. Only a few years earlier, Paul Berg, Walter Gilbert, and Fred Sanger had shared the Nobel Prize for their work with DNA restriction/ligation and sequencing. Their experiments had shown how to construct *E. coli* strains capable of "overproducing" proteins and nucleic acids. Gesteland paired me up with Nello Bossi, a gifted scientist and a great teacher, who at the time was a senior postdoctoral associate. Bossi was somewhat skeptical about me making much progress. I came in as a complete neophyte, with no background or laboratory expertise in molecular biology and "rusty" laboratory skills. What he discovered was that my laboratory experience with organic molecules was wonderful training for the skills needed for molecular biology. All an organic chemist needs to do is overcome the fear of carrying out reactions in water. A few restriction enzymes and DNA ligase had recently become commercially available, and Gesteland had just bought one of the first DNA synthesizers. These developments speeded up my work, and by the end of the year I had constructed a strain of *E. coli* that required uracil for growth and overproduced *E. coli* $tRNA^{Phe}$. We were now able to get multimilligram quantities of pure $tRNA^{Phe}$ from a 1-L fermentation using only a few milligrams of labeled uracil.[44] With the availability of synthetic DNA, PCR, rapid DNA sequencing, and better protocols for cloning and purification of DNA, the experiments took me a year to complete in 1988 can now be finished in a month or so by an inexperienced undergraduate student with proper guidance.

Ray Gesteland

Studies with Recombinant Enzymes. The carbon skeletons of isoprenoid molecules are formed from simple isoprenoid diphosphates during a "building" phase early in the pathway before they are adorned with a variety of different functional groups. There are eight different ways in which the individual isoprene units are attached by these building reactions (see Figure 7). Our model studies of the chrysanthemyl cation suggested that four of the structures, the 1′−1, 1′−3, c1′−1−2, and 2−1′−3 skeletons, are derived from the c1′−2−3 structure by rearrangement.[16−18] The 1′−4 skeleton is produced by condensation of an allylic diphosphate with isopentenyl diphosphate.[45] The c1′−2−3 skeleton is derived from condensation of two molecules of an allylic diphosphate, for example, the synthesis of presqualene diphosphate from farnesyl diphosphate.[46]

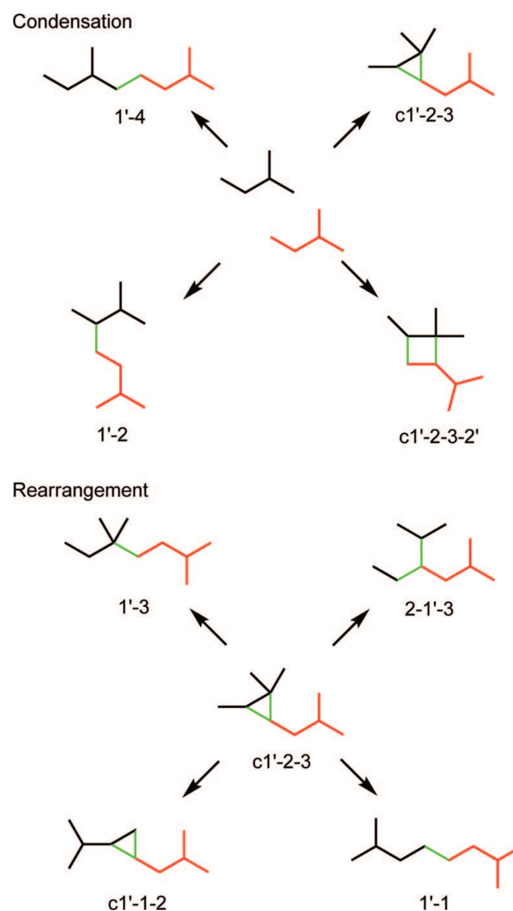


FIGURE 7. Isoprenoid carbon skeletons from joining two units (condensation) and from rearrangements of the c1′−2−3 skeleton.

We thought it likely that the 1′−2 and c1′−2−3−2′ skeletons are also formed by condensation of two allylic diphosphates. When I first started to study these reactions, the enzymes that synthesize molecules with 1′−4, 1′−1, and c1′−2−3 carbon skeletons were known. Those that catalyze synthesis of molecules with 1′-3, 2−1′−3, c1′−1−2, and c1′−2−3−2′ skeletons still remain to be identified! Although I will not discuss terpene cyclases in this Perspective, these enzymes catalyze intramolecular versions of prenyl transfer reactions to give products with 1′−2 and c1′-2−3 attachment patterns between isoprene units.

My experience with cloning in Ray Gesteland's laboratory proved to be invaluable for my work with biosynthetic enzymes. We were no longer restricted to proteins purified from natural sources! If a gene could be identified and cloned, the corresponding "recombinant" protein could usually be produced in "large" quantities and its amino acid sequence altered at will by site-directed mutagenesis. Weeks spent in the cold room to purify a protein became a thing of the past and it was now possible to study enzymes that were only found at very low levels in their host organisms. The first recombinant proteins we obtained were versions of two old friends, farnesyl diphosphate synthase[47] and isopentenyl diphosphate isomerase.[48] These were among the first biosynthetic enzymes to be produced by recombinant DNA technology. Since then, we have constructed overproducing microbial strains for enzymes that catalyze over 20 different biosynthetic reactions, variants of these enzymes from different organisms, and enzymes containing site-directed mutations. All of the proteins we work with now are products of cloning.

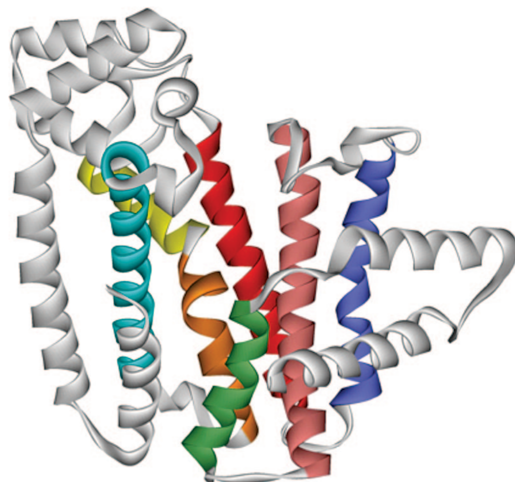


FIGURE 8. X-ray structure of farnesyl diphosphate synthase showing the six helicies containing highly conserved sequences that form the active site.

The 1′−4 Connection. In 1975, Hans Rilling had noticed that avian farnesyl diphosphate synthase precipitated from high concentrations of ammonium sulfate to give a mixture of crystalline and amorphous protein.[49] Both of us worked off and on without success to obtain crystalline enzyme suitable for X-ray crystallography by purification of the protein from chicken livers until Rilling retired almost 20 years later. When Mu Jing Yan, an exceptionally talented technician in my group, finally constructed an *E. coli* clone that overproduced avian farnesyl diphosphate synthase, availability of the protein no longer an issue. Shortly afterward, I entered into what proved to be a very productive collaboration with Larry Tarshis, a graduate student in Sacchettini's group, to obtain an X-ray structure of the protein. Within a year, we had a structure of the apo enzyme and structures of the enzyme complexed with geranyl diphosphate, farnesyl diphosphate, and geranylgeranyl diphosphate (see Figure 8).[50]

Farnesyl diphosphate synthase is a homodimer of all α-helical subunits. Five conserved regions found in the amino acid sequences of all of the enzymes that synthesize isoprenoid chains with *E*-double bonds—geranyl diphosphate, farnesyl diphosphate, geranylgeranyl diphosphate, and the longer chain diphosphate synthases—are located on six helices that cluster around the active site. These helices are the major structural features in a motif now called the "isoprenoid synthase" or "terpenoid synthase" fold.[51−53] In addition to the enzymes that synthesize chains with *E*-double bonds, the isoprenoid synthase fold was subsequently found in squalene synthase,[54] dehydrosqualene synthase,[55] and in mono-, sesqui-, and diterpene cyclases.[56]

The X-ray structures of farnesyl diphosphate synthase complexed with isoprenoid diphosphates suggested that the hydrocarbon chain of the product "grew" into a hydrophobic pocket located in the interior of the enzyme during chain elongation. The "floor" of the binding pocket in the avian enzyme consists of two π-stacked benzene moieties from phenylalanine residues at positions 112 and 113. Thus, it seemed reasonable that the depth of the pocket dictated the ultimate length of the growing hydrocarbon chain for the various chain elongation enzymes. This hypothesis was verified by constructing site-directed mutants of farnesyl diphosphate synthase that were selective for synthesis of C_{10}, C_{20}, C_{25}, or longer isoprenyl diphosphates.[57,58] Structural work with other isoprenoid chain elongations enzymes indicated that the mechanism used to regulate chain length in farnesyl diphosphate synthase is a general phenomenon for other members of the chain elongation family.[59]

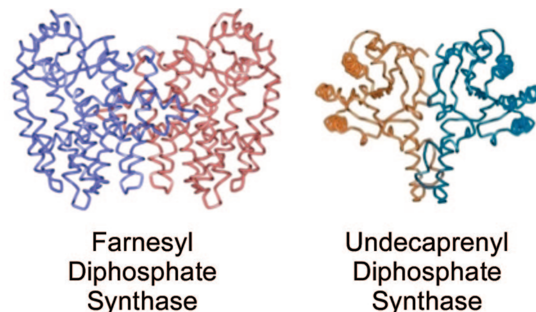


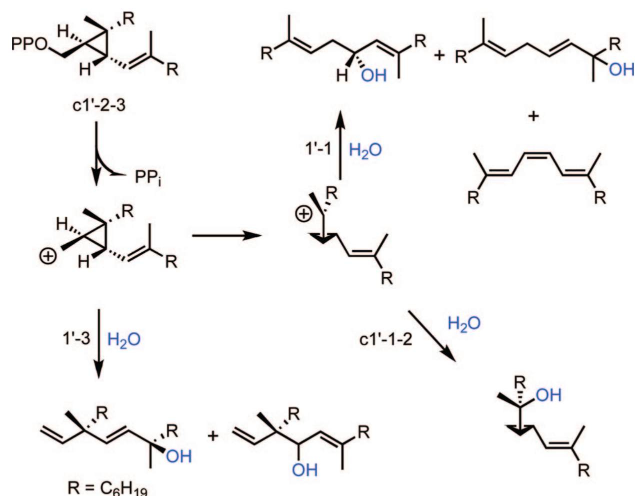
FIGURE 9. Crystal structures of farnesyl diphosphate synthase (*E*-selective) and undecaprenyl diphosphate synthase (*Z*-selective).

Farnesyl diphosphate synthase was discovered in the 1950s.[60] The enzyme appears to be required by all living cells. With the exception of a small group of parasitic bacteria with extremely small genomes,[61] genes coding for putative farnesyl diphosphate synthases have been found in all organisms studied thus far. The 1′−4 chain elongation enzymes can be divided into two groups, those that synthesize chains with *E*-double bonds and those that synthesize chains with *Z*-double bonds. Both groups of enzymes are homodimers that presumably catalyze chain elongation by the same electrophilic alkylation mechanism. The *Z*-isoprenoid diphosphate synthases typically synthesize a variety of long-chain molecules with 10 or more isoprene units, including rubber. These proteins are less well studied than their *E*-synthase counterparts. They belong to a different structural class than the *Z*-isoprenoid diphosphate synthases[62] (see Figure 9) and do not have the isoprenoid synthase fold found in all of the *E*-synthases.[50]

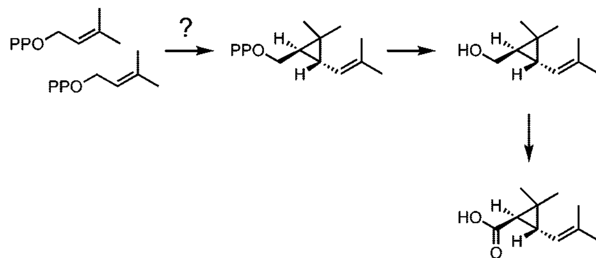
The 1′−1 Connection. Compounds with 1′−1 attachments are found at the branch points leading to the sterol, hopanoid, and carotenoid biosynthetic pathways. While metabolites derived from 1′−1 intermediates are not universally distributed, they are found in all members of the Eukaryotic and Archaeal kingdoms and in some Bacteria. As previously illustrated in Scheme 1, two molecules of farnesyl diphosphate are joined to give squalene.[46] Similar reactions convert geranylgeranyl diphosphate to phytoene during carotenoid biosynthesis. Squalene synthase and phytoene synthase are bifunctional enzymes. They catalyze condensation of their allylic isoprenoid substrates to give the c1′−2−3 intermediates, (1*R*,2*R*,3*R*)-presqualene diphosphate[63] and (1*R*,2*R*,3*R*)-prephytoene diphosphate,[64] respectively. Under normal catalytic conditions, these compounds are converted directly to squalene and phytoene without being released from the active site.[65]

Although our model work indicated that isoprenoid molecules with 1′−1, 1′−3, 2−1′−3, and c1′−1−2 skeletons are all synthesized by rearrangement of a c1′−2−3 cyclopropylcarbinyl cation, squalene synthase only gives 1′−1 products when incubated with farnesyl diphosphate and NADPH. Hans Rilling discovered that presqualene diphosphate was an intermediate in the synthesis of squalene by withholding NADPH from the normal incubation buffer and observing the formation of the c1′−2−3 cyclopropylcarbinyl diphosphate.[12] Presqualene diphosphate did not react under the conditions used in his experiments. Wild-type squalene synthase is a microsomal protein that is difficult to purify, and the preparations Rilling used in his experiments had low levels of squalene synthase activity. When the gene of yeast squalene synthase was isolated and character-

SCHEME 8. Solvolysis of Presqualene Diphosphate by Squalene Synthase ($R = C_{11}H_{19}$)



SCHEME 9. Proposed Role for Chrysanthemyl Diphosphate Synthase in Biosynthesis of Chrysanthemol and Chrysanthemic Acid



ized, the amino acid sequence of the encoded protein had a putative membrane-spanning α-helix at its C-terminus.[66] After much trial and error, we were able to obtain a moderately soluble version of the enzyme with essentially the same catalytic properties as the wild-type enzyme by deleting the amino acids that formed the helix.[67]

As previously seen by Rilling, farnesyl diphosphate was converted to presqualene diphosphate by the more active recombinant enzyme when incubated without NADPH. However, we also saw a slower, but clearly enzyme-catalyzed, formation of three new products with 1′−1 and 1′−3 carbon skeletons by "solvolysis" of presqualene diphosphate (Scheme 8)![68] The 1′−1 alcohol, hydroxysqualene, corresponds to the normal product of squalene synthase, except water has replaced the hydride from NADPH as the nucleophile that captures the putative c1′−1−2 carbocation. The allylic isomer of hydroxysqualene was also formed as a minor product. Dehydrosqualene, obtained by elimination of a proton from the c1′−1−2 cation, is a C_{30} analogue of phytoene. The carbon skeleton of the 1′−3 product, hydroxybotryococcene, is analogous to the naturally occurring C_{30} hydrocarbon botryococcene synthesized by the colonial photosynthetic green algae *Botryococcus branunii*. The oil sack that surrounds colonies of the algae can comprise up to 75% of the dry weight of the organism.[69] Feeding experiments with deuterium-labeled farnesol established that botryococcene is synthesized from two molecules of farnesyl diphosphate, probably with presqualene diphosphate as an intermediate.[70] The allylic isomer of hydroxybotryococcene was also a minor product from the incubation of presqualene diphosphate with recombinant squalene synthase.

In addition to its role as a cosubstrate, NADPH is an important architectural feature in the active site of the squalene synthase during the rearrangement of presqualene diphosphate to squalene. We thought that NADPH binding might facilitate the c1′−2−3 to c1′−1−2 cyclopropylcarbinyl−cyclopropylcarbinyl rearrangement and reduce the amount of 1′−3 products during the squalene synthase catalyzed solvolysis of presqualene diphosphate. We were pleasantly surprised to discover that there was a substantial change in distribution of products when $NADPH_3$, an inactive dihydro analogue of NADPH, was included in the buffer.[71] Hydroxysqualene and dehydrosqualene became the major products of the reaction, and the 1′−3 alcohols were not seen. However the real surprise of this experiment was formation of a new cyclopropylcarbinyl alcohol with the c1′−1−2 carbon skeleton whose absolute stereochemistry matched the prediction I had made to Bill Epstein when I interviewed for a job at the University of Utah 34 years earlier! This alcohol, formed by nucleophilic capture of the corresponding tertiary c1′−1−2 cyclopropylcarbinyl cation, represented the missing link in the mechanism I had originally proposed for the rearrangements required to convert presqualene diphosphate to squalene. I decided to name the alcohol "Rillingol" to honor Rilling's fundamental contributions to the enzymology of squalene synthase. His pioneering work made it possible for us to finally establish the mechanism of the c1′−2−3 to 1′−1 rearrangement, as well as the rearrangements that give 1′−3, 2−1′−3, and c1′−1−2 skeletons.

c1′−2−3, 1′−2, and c1′−2−3−2′ Connections. As our work on the rearrangements of presqualene diphosphate was drawing to a close, we turned to the c1′−2−3 cyclopropanation, the first of the two reactions catalyzed by the enzyme, with renewed interest. Earlier attempts to determine the mechanism of the reaction with squalene synthase were not definitive and we sought a system to study where the cyclopropanation reaction was decoupled from the rearrangements catalyzed by squalene synthase and phytoene synthase. The cyclopropanation enzyme responsible for synthesis of the c1′−2−3 monoterpenes chrysanthemol and chrysanthemic acid appeared to be appropriate alternative systems to study.

Chrysanthemic acid was discovered by Leopold Ruzicka during his experiments to determine the structures of the naturally occurring pyrethrin insecticides found in *Chrysanthemum cinerariaefolium* and related members of the Asteraceae family as a young assistant in Hermann Staudinger's laboratory at the Technische Hochschule at Karlsruhe.[72,73] Ruzicka formulated the "isoprene rule" and coined phrase "non-head-to-tail" to describe molecules where the isoprenoid units did not have the typical 1′−4 or "head-to-tail" attachments.[74] He shared the 1939 Nobel Prize in Chemistry "for his work with polymethylenes and higher terpenes" with Adolf Butenandt. Chrysanthemol was found by Bill Epstein in several species of sagebrush common in Utah.[19,75,76] We assumed that an unknown enzyme, chrysanthemyl diphosphate synthase, catalyzed formation of a c1′−2−3 monoterpene chrysanthemyl diphosphate as the first pathway specific step in biosynthesis of chrysanthemol and chrysanthemic acid.

Two plants appeared to be attractive sources for chrysanthemyl diphosphate synthase, *Chrysanthemum cinerariaefolium*, a well-established source of pyrethrins, and *Artemisia tridentata* ssp. spiciformis or snowfield sagebrush, a rather scraggly small bush that is common at elevations above 7000 ft in the mountains of Utah.[19] I would not have dared to tackle a major

project that involved biosynthetic plant proteins before the era of cloning. Plant enzymes, especially those in specialized biosynthetic pathways, are typically found in trace amounts and purification of substantial quantities of the proteins needed for mechanistic work is a daunting task. We used a sensitive radioisotope-based assay to follow enzyme activity to guide purification of the small amount of protein needed to obtain peptide sequences of fragments of the enzyme.[77] DNA containing open reading frames for the genes was isolated from cDNA libraries of *Chrysanthemum cinerariaefolium* and *Artemisia tridentata* ssp. spiciformis using nucleotide probes based on these sequences.[77,78]

The first of several serendipitous surprises in this project came when we obtained the sequences of the cDNA clones. I rather naively thought that the amino acid sequence for chrysanthemyl diphosphate synthase would probably resemble those of squalene synthase and phytoene synthase. To my surprise, the amino acid sequences for chrysanthemyl diphosphate synthase were similar to the *E*-double-bond chain elongation enzymes. In *A. tridentata* ssp. spiciformis, the sequences for chrysanthemyl diphosphate synthase and farnesyl diphosphate synthase were 75% identical and 96% similar!

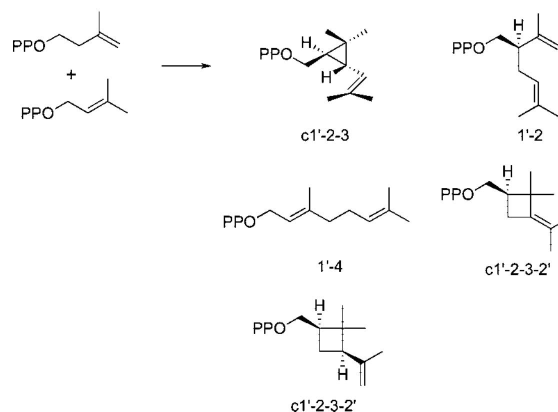
The creation of new biosynthetic enzymes is thought to occur by a series of changes in the genome that begins with duplication of a functional gene. The original function can be retained in one of the copies, while the other is "free" to acquire a new function through mutations in the original gene and by recombination with elements from another gene. Early during the process of acquiring a new function, the fledgling protein lacks the catalytic efficiency and specificity of its "mature" descendents. Those properties are acquired later in response to evolutionary pressures. Genetic considerations indicate that the gene for chrysanthemyl diphosphate synthase, a specialized enzyme found in a closely related family of plants, evolved perhaps as recently at 50 million years ago from the gene for farnesyl diphosphate synthase, an essential enzyme found in all organisms. Over time, the similarity between amino acid sequences for sibling proteins with different functions becomes less pronounced and may ultimately disappear, leaving only the characteristic "fold" in their tertiary structures as all that remains to trace their ancestry. This appears to be the case for squalene synthase and farnesyl diphosphate synthase where there is no discernible similarity between the amino acid sequences of the two enzymes. Yet when the crystal structure of human squalene synthase was published in 2000, the isoprenoid fold was seen in the helicies surrounding the active site.[54] At that point, it became clear that the squalene/phytoene synthases and the *E*-selective chain elongation enzymes shared a common ancestor long ago, presumable an ancient chain elongation enzyme.

The extraordinarily high level of sequence similarity between *A. tridentata* chrysanthemyl diphosphate synthase and farnesyl diphosphate synthase suggests that they shared a common ancestor rather recently. When we analyzed the products from incubation of chrysanthemyl diphosphate synthase with dimethylallyl diphosphate or a mixture of dimethylallyl diphosphate and isopentenyl diphosphate, we found an unexpected distribution of products. Incubations with dimethylallyl diphosphate gave (1*R*,3*R*)-chrysanthemyl diphosphate as expected. However, (*R*)-lavandulyl diphosphate, a 1′−2-branched monoterpene, formed 20% of the total products. The related alcohol, lavandulol, is a component of oil of lavender. Incubation of dimethylallyl diphosphate with isopentenyl diphosphate gave geranyl diphosphate as a third product! Thus, the catalytic machinery of chrysanthemyl diphosphate synthase synthesizes compounds with c1′−2−3, 1′−2, and 1′−4 attachments between two five-carbon isoprene units (see Scheme 10). Clearly, chrysanthemyl diphosphate synthase has not yet achieved the selectivity for cyclopropanation seen for squalene synthase and phytoene synthase and retains the capacity to catalyze chain elongation.

Given the high degree of similarity between sagebrush cyclopropanation and chain-elongation enzymes, we though it would be possible to construct catalytically active chimeric proteins that incorporated structural features from both. The same unique DNA restriction sites were introduced at identical locations in genes for both enzymes, which allowed us to replace any combination of the five conserved in one protein with those from the other.[79] These constructs were used to construct a series of catalytically active chimeric proteins where amino acids, beginning at the N-terminus of farnesyl diphosphate synthase, were replaced with increasing amounts of sequence from chrysanthemyl diphosphate synthase as shown in Figure 10. We discovered that the primary function of the chimeras changed smoothly from synthesis of farnesyl diphosphate to geranyl diphosphate to lavandulyl diphosphate to chrysanthemyl diphosphate. We were also surprised to find two previously unseen monoterpene diphosphates, (1*R*,3*R*)-planococcyl diphosphate and (*R*)-maconelliyl diphosphate, with c1′−2−3−2′ cyclobutane rings (see Scheme 10). The c299f and c243f chimeras synthesized monoterpenes representative of all four of the skeletons found in nature formed by joining two isoprenoid molecules.

The discovery that a protein with a single active site synthesizes isoprenoid moleules with 1′−4, c1′−2−3, 1′−2, and c1′−2−3−2′ carbon skeletons suggests that the chemical mechanisms for these reactions are similar. The mechanism shown in Scheme 11 for combining two molecules of dimethylallyl diphosphate to give c1′−2−3, 1′−2, and c1′−2−3−2′ compounds is based on the dissociative electrophilic alkylation mechanism for the 1′−4 condensation of isopentenyl diphosphate and an allylic diphosphate (see Scheme 5). The first step is formation of the dimethylallyl cation from one of the two molecules of dimethylallyl diphosphate bound in the active site. The highly electrophilic allylic cation alkylates the double bond in the second molecule of dimethylallyl diphosphate to give a protonated cyclopropane, which has two choices - to lose a proton to produce chrysanthemyl diphosphate or rearrange to a tertiary cation with a

SCHEME 10. Products Synthesized by Farnesyl Diphosphate, Chrysanthemyl Diphosphate, and Chimeras of the Two Enzymes When Incubated with Isopentenyl Diphosphate and Dimethylallyl Diphosphate[a]



[a] The carbon skeletons in these compounds represent the four structures formed in nature by combination of two smaller isoprenoid units.

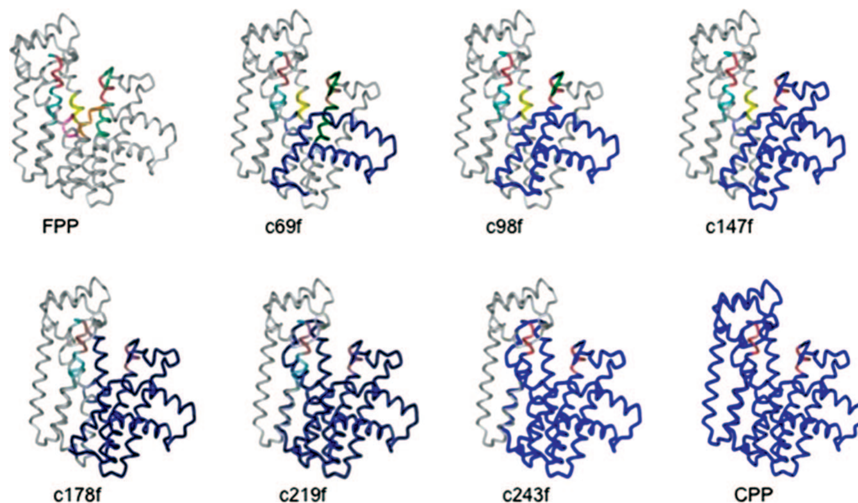
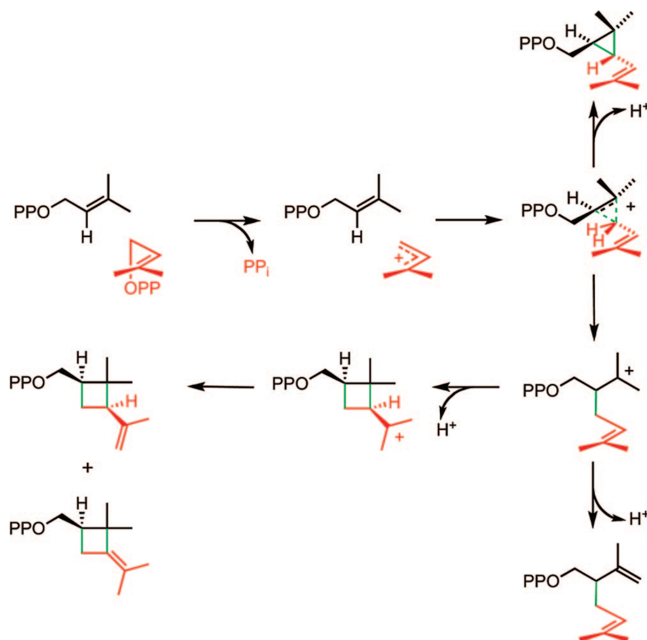


FIGURE 10. Farnesyl diphosphate synthase (FPP), chrysanthemyl diphosphate synthase (CPP), and chimeras constructed by replacing FPP sequence with CPP sequence starting at the N-terminus of FPP. The proteins are named according the number of amino acids in farnesyl diphosphate synthase that have been replaced by the corresponding amino acids from chrysanthemyl diphosphate synthase.

SCHEME 11. Dissociative Electrophilic Alkylation Mechanism for Biosynthesis of c1′−2−3, 1′−2, and c2−1′−2′−3 Isoprenoid Compounds



branched skeleton. In turn, the tertiary cation has two choices—to lose a proton to give lavandulyl diphosphate or rearrange to a tertiary cyclobutylcarbinyl cation. Finally, the tertiary cyclobutylcarbinyl cation loses a proton from a methyl group to give planococcyl diphosphate or from the cyclobutane ring to give maconelliyl diphosphate. Primary and secondary deuterium isotope effects on the formation of products at each of the partitioning steps support this sequence of events.[80]

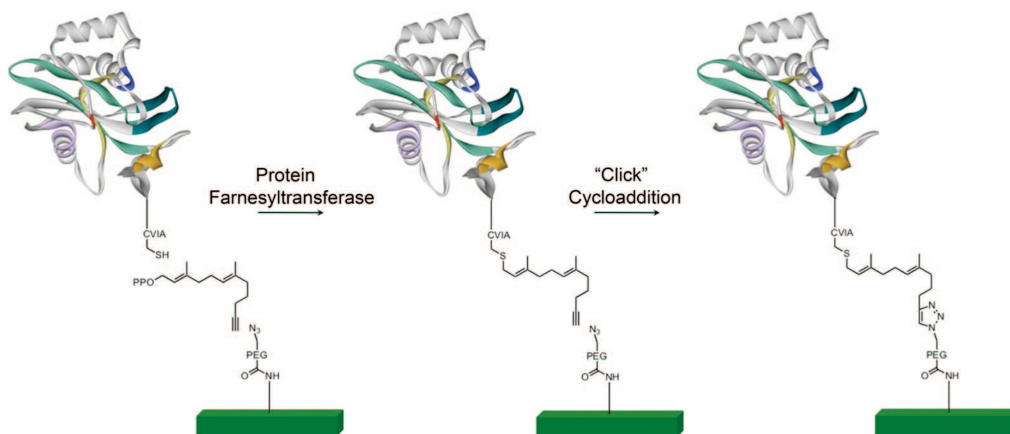
The idea of a common ancestor for the enzymes that synthesize c1′−2−3, 1′−2, and c1′−2−3−2′ isoprenoids is consistent with the stereochemistries of the natural products and the compounds synthesized by our chimeras. Their absolute stereochemistries are identical and are consistent with the mechanism shown in Scheme 11. The relative orientation of the two molecules of dimethylallyl diphosphate in the enzyme−substrate complex dictates the absolute stereochemistry of the chiral centers in the protonated cyclopropane intermediate and, ultimately, the absolute stereochemistry of the chiral centers in the 1′−2 and c1′−2−3−2′ skeletons. A similar orientation of the α-isoprene units in farnesyl diphosphate and geranylgeranyl diphosphate would lead to the same absolute stereochemistry for presqualene and prephytoene diphosphate. The stereeoselectivity of the cyclopropanation reaction is a consequence of the topology of an active site constructed from the isoprenoid fold, a motif that originally evolved to synthesize achiral molecules. The stereochemistry imprinted in the c1′−2−3 cyclopropane rings also dictates the absolute stereochemistries of the chiral centers in 2−1′−3 and c1′−1−2 skeletons, and at a more subtle level, the stereochemistry at methylene groups rendered chiral by deuterium substitution.

Reflections

Most of my research during my career at the University of Utah was not possible using the technology available to me when I first started. In my chosen field of biological chemistry, the ability to manipulate DNA to synthesize proteins in microorganisms and to alter the constituent amino acids in the proteins at will, coupled with advances in synthesis and analytical instruments, allowed me to create reality out of dreams. I had neither the backgrouond nor the vision to anticipate these developments when I entered graduate school 45 years ago. Scientific discoveries drive developments in technology, which in turn drive new scientific discoveries. This cycle will continue as long as science and technology are nourished. Because of the unpredictability of discovery and how these discoveries will impact developments in technology, especially those that occur in areas of science rather far removed from my area of expertise, I think it would be naïve for me to try to make specific predictions about advances too far into the future. I recall my amazement in the early 1970s at the “power” of my new HP hand-held calculator, which replaced my trusty Post Versalog slide rule, or in my ability to simulate complex NMR spectra from a program encoded in a large box of cards I carried to the campus computer center. Today my four-pound laptop is more powerful, has more storage, and is about 1000-fold less expensive than that computer. But in the shorter term, I can think several areas that offer great promise and will mention just a few.

Over the millennia, organisms have evolved to ability to efficiently synthesize a tremendous variety of complex organic

SCHEME 12. Regio- and Chemoselective Attachment of Proteins to Silica Surfaces



compounds, some of which have important uses in modern society. Traditionally, selective propagation of microbial strains or higher organisms has been used to enhance production of metabolites by their native hosts, often with dramatic improvements in yields. Modern molecular biology allows one to introduce genes encoding native and mutant biosynthetic enzymes from an organism into an unrelated surrogate host to synthesize novel molecules.

Metabolic engineering of complex biosynthetic pathways is still a rather new endeavor that faces a number of hurdles, including how to stably integrate large foreign gene clusters into the host's chromosome, how to control the activities of individual enzymes in the pathway through regulation of gene expression and turnover of mRNAs and enzymes, and how to ameliorate problems associated with the potential toxicity of the biosynthetic mRNA, enzymes and metabolites to the host. Cost is also a major consideration. It seems to me that in the long run the practice of using glucose as a feed stock for microbial fermentation of low value compounds will give way to cheaper sources of carbon, probably CO_2 fixed by miroorganisms capable of photosynthesis. In addition, while high value compounds can be synthesized economically in batches, problems associated the utilization or disposal of massive amounts of biomass from large-scale fermentations will require attention, perhaps by developing strains where the desired metabolites can be harvested continuously.

Many of the stunning developments in developmental biology of the past few years have relied heavily on imaging. A particularly elegant example is a recent report from Carolyn Bertozzi's laboratory where they substituted *N*-azidoacetylgalactosamine for *N*-acetylgalactosamine in the cell surface glycans of zebrafish embryos and visualized the glycans by attaching fluorescent tags using a copper-free version of the water-tolerant bio-orthogonal "click" reaction developed by Barry Sharpless.[81] The Bertozzi group was able to achieve spatial and temporal resolution of developmental changes in the embryos through the use of two- and three-color detection. New imaging procedures and related techniques to trace the fate of individual molecules in cells will facilitate our ability to understand at a molecular level how timing and duration of gene expression controls differentiation and development and how ultimately how to regulate these processes.

Analytical techniques for detecting molecules, with continuing developments in sensitivity and easy of use, has and will continue to have a profound impact on science and society. As one example, protein microarrays are becoming increasingly important as analytical devices in biology, medicine, and drug discovery. The stability and sensitivity of the arrays depend critically on how proteins are attached to the surface of a microchip. Typically, the protein molecules are immobilized on the surface by noncovalent absorption or are attached covalently by nucleophilic displacement reactions with electrophiles present on the surface. These reactions are often not regioselective for the array of nucleophiles on the protein surface and readouts with an array of heterogeneously immobilized proteins are typically less sensitive than with proteins attached regioselectively. We have approached this by developing a protein tagging protocol based on the protein farnesyltransferase reaction (see Scheme 12). Protein farnesyltransferase attaches a farnesyl unit to cysteine residues in C-terminal CaaX recognition motifs, where "C" is cysteine, "a" is an amino acid with a small aliphatic side chain, and "X" is alanine, serine, methionine, or glutamine. The CaaX motif is all that is needed for recognition by protein farnesyltransferase. The amino group of the cysteine can be attached to a hydrogen atom, a small peptide, a large enzyme, or a nonpeptidic unit so long as the substrate is soluble. We synthesized ω-terminal azide and alkyne analogues of farnesyl diphosphate that are excellent alternate substrates for protein farnesyl transferase.[82] Peptides and proteins modified with the analogues can be covalently linked to derivatized glass surfaces using the copper-catalyzed "click" reaction.[83] The attachment is regiospecfic and robust. This and related techniques for construction of protein arrays might lead to more durable and sensitive analytical devices.

Acknowledgment. Science is a social enterprise, and the community that nourished me is large. I have had the good fortune to share ideas and work with a large group of colleagues in industry and academia, only a few of whom are mentioned in this Perspective. I can only hope that my contributions to their careers approach what they have contributed to mine. One of my great joys as a scientist is watching my "academic children" leave the laboratory and establish their own independent careers. My style as a mentor has been to provide general direction but require that my students and postdoctorals "take charge" of their projects, intellectually and at the bench. All of the fundamental research that I publish from my laboratory is a collaborative effort involving undergraduates, graduates, and postdoctorals that it has been my pleasure to work with during the course of my career. I have benefited from their intellect, enthusiasm, and dedication. While only some of their names appear as coauthors in the publications cited in this Perspective, all have made enormous contributions.

The Institute for General Medical Sciences at the National Institutes of Health, through Grant Nos. GM21328, GM25521, and GM 32490, funded the projects I described. I could not have developed or sustained a research program at the interface of organic chemistry and biology without the generous, stable support provided for my training in biochemistry and molecular biology and for research program. Finally, I recall an award address delivered by Paul Gassman several years ago at an ACS meeting. He had two slides for acknowledgments. The first was the usual list of co-workers, collaborators, and funding agencies. The last slide was a picture of Paul's wife Gerda. I do not recall exactly what Paul said, but I do recall thinking how rarely we acknowlege our most important "collaborator". Susan and I were married in late August about a week before heading for Berkeley where I started graduate school and she finished her BS degree in Chemistry before earning a PhD with Clayton Heathcock. I am deeply grateful for the sacrifices, support, and wise council she has given me since then. Thanks go to Dennis Galloway and Michael Barnes of the University of California at Berkeley for the picture of William Dauben.

References

(1) Jorpes, J. E. *Jac. Berzelius: his life and work*, 1st Am. ed.; Univ. Calif. Press: Berkeley, 1970.
(2) *Nobel Lectures, Chemistry 1901−1921*; World Scientific: River Edge, NJ, 1999; pp 175−196.
(3) Meerwein, H.; van Emster, K. *Ber.* **1922**, *55*, 2500.
(4) Woodward, R. B.; Sondheimer, F.; Heusler, K.; McLamore, W. M. *J. Am. Chem. Soc.* **1952**, *74*, 4223–4251.
(5) Barton, D. H. R. *Science* **1970**, *169*, 539–544.
(6) Rilling, H. C.; Bloch, K. *J. Biol. Chem.* **1959**, *234*, 1424–1432.
(7) Knappe, J.; Ringelmann, E.; Lynen, F. *Biochem. Z.* **1959**, *332*, 195–213.
(8) Cornforth, J. W. *Angew. Chem., Int. Ed. Engl.* **1968**, *7*, 903–911.
(9) Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 205–211.
(10) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737–738.
(11) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. *Nature* **1958**, *181*, 662–666.
(12) Rilling, H. C. *J. Biol. Chem.* **1966**, *241*, 3233–3236.
(13) Caserio, M. C.; Graham, W. H.; Roberts, J. D. *Tetrahedron* **1960**, *11*, 171–182.
(14) Silver, M. S.; Caserio, M. C.; Rice, H. E.; Roberts, J. D. *J. Am. Chem. Soc.* **1961**, *83*, 3671–3678.
(15) Edmond, J.; Popjak, G.; Wong, S. M.; Williams, V. P. *J. Biol. Chem.* **1971**, *246*, 6254–6271.
(16) Poulter, C. D.; Marsh, L. L.; Hughes, J. M.; Argyle, J. C.; Satterwhite, D. M.; Goodfellow, R. J.; Moesinger, S. G. *J. Am. Chem. Soc.* **1977**, *99*, 3816–3823.
(17) Poulter, C. D.; Hughes, J. M. *J. Am. Chem. Soc.* **1977**, *99*, 3824–3829.
(18) Poulter, C. D.; Hughes, J. M. *J. Am. Chem. Soc.* **1977**, *99*, 3830–3837.
(19) Gunawardena, K.; Rivera, S. B.; Epstein, W. W. *Phytochemistry* **2002**, *59*, 197–203.
(20) Clifford, K.; Cornforth, J. W.; Mallaby, R.; Phillips, G. T. *Chem. Commun.* **1971**, *No. 24*, 1599–1600.
(21) Jencks, W. P. *Catalysis in Chemistry and Enzymology*; McGraw-Hill: New York, 1969.
(22) Abeles, R. H.; Maycock, A. L. *Acc. Chem. Res.* **1976**, *9*, 313–319.
(23) Poulter, C. D.; Argyle, J. C.; Mash, E. A. *J. Am. Chem. Soc.* **1977**, *99*, 957–959.
(24) Poulter, C. D.; Satterwhite, D. M. *Biochemistry* **1977**, *16*, 5470–5478.
(25) Poulter, C. D.; Argyle, J. C.; Mash, E. A. *J. Biol. Chem.* **1978**, *253*, 7227–7233.
(26) Poulter, C. D.; Wiggins, P. L.; Le, A. T. *J. Am. Chem. Soc.* **1981**, *103*, 3926–3927.
(27) Gebler, J. C.; Woodside, A. B.; Poulter, C. D. *J. Am. Chem. Soc.* **1992**, *114*, 7354–7360.
(28) Zhang, D.; Poulter, C. D. *J. Am. Chem. Soc.* **1993**, *115*, 1270–1277.
(29) Soderberg, T.; Poulter, C. D. *Biochemistry* **2001**, *40*, 1734–1740.
(30) Dolence, J. M.; Poulter, C. D. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5008–5011.
(31) Watson, J. D. *The Double Helix*; Atheneum: New York, 1968; pp 114−126.
(32) Poulter, C. D.; Anderson, R. B. *Tetrahedron Lett.* **1972**, *13*, 3823–3826.
(33) Frederick, G. D.; Poulter, C. D. *J. Am. Chem. Soc.* **1975**, *97*, 1797–1801.
(34) Poulter, C. D.; Frederick, G. D. *Tetrahedron Lett.* **1975**, *16*, 2171–2174.
(35) Iwahashi, H.; Kyohoku, Y. *J. Am. Chem. Soc.* **1977**, *99*, 7761–7765.
(36) Hawkes, G. E.; Randall, E. W.; Hull, W. E. *J. Chem. Soc., Perkin Trans. 2* **1977**, 1268–1275, No. 10.
(37) Roberts, J. L.; Poulter, C. D. *J. Org. Chem.* **1978**, *43*, 1547–1550.
(38) Poulter, C. D.; Livingston, C. L. *Tetrahedron Lett.* **1979**, *20*, 755–758.
(39) Reid, B. R.; Hurd, R. E. *Acc. Chem. Res.* **1977**, *10*, 396–402.
(40) Griffey, R. H.; Poulter, C. D.; Yamaizumi, Z.; Nishimura, S.; Hurd, R. E. *J. Am. Chem. Soc.* **1982**, *104*, 5810–5811.
(41) Griffey, R. H.; Poulter, C. D.; Yamaizumi, Z.; Nishimura, S.; Hawkins, B. L. *J. Am. Chem. Soc.* **1983**, *105*, 143–145.
(42) Griffey, R. H.; Poulter, C. D.; Yamaizumi, Z.; Nishimura, S.; Hurd, R. E. *J. Am. Chem. Soc.* **1982**, *104*, 5811–5813.
(43) Griffey, R. H.; Poulter, C. D.; Bax, A.; Hawkins, B. L.; Yamaizumi, Z.; Nishimura, S. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 5895–5897.
(44) Davis, D. R.; Poulter, C. D. *Biochemistry* **1991**, *30*, 4223–4231.
(45) Poulter, C. D.; Rilling, H. C.; In *Biosynthesis of Isoprenoid Compounds*; Porter, J. W., Spurgeon, S. L., Eds.; John Wiley & Sons, Inc.: New York, 1981; Vol. 1, pp 162−224.
(46) Poulter, C. D.; Rilling, H. C. In *Biosynthesis of isoprenoid compounds*; Porter, J. W., Spurgeon, S. L., Eds.; John Wiley & Sons, Inc.: New York, 1981; Vol. 1, pp 414−441.
(47) Anderson, M. S.; Yarger, J. G.; Burck, C. L.; Poulter, C. D. *J. Biol. Chem.* **1989**, *264*, 19176–19184.
(48) Anderson, M. S.; Muehlbacher, M.; Street, I. P.; Proffitt, J.; Poulter, C. D. *J. Biol. Chem.* **1989**, *264*, 19169–19175.
(49) Reed, B. D.; Rilling, H. C. *Biochemistry* **1975**, *14*, 50–54.
(50) Tarshis, L. C.; Yan, M.; Poulter, C. D.; Sacchettini, J. C. *Biochemistry* **1994**, *33*, 10871–10877.
(51) Lesburg, A.; Zhai, G.; Cane, D. E.; Christianson, D. W. *Science* **1997**, *277*, 1820–1824.
(52) Sacchettini, J. C.; Poulter, C. D. *Science* **1997**, *277*, 1788–1789.
(53) Christianson, D. W. *Curr. Opin. Chem. Biol.* **2008**, *12*, 141–150.
(54) Pandit, J.; Danley, D. E.; Schulte, G. K.; Mazzalupo, S.; Pauly, T. A.; Hayward, C. M.; Hamanaka, E. S.; Thompson, J. F.; Harwood, H. J., Jr. *J. Biol. Chem.* **2000**, *275*, 30610–30617.
(55) Liu, C. I.; Liu, G. Y.; Song, Y.; Yin, F.; Hensler, M. E.; Jeng, W. Y.; Nizet, V.; Wang, A. H.; Oldfield, E. *Science* **2008**, *319*, 1391–1394.
(56) Christianson, D. W. *Chem. Rev.* **2006**, *106*, 3412–3442.
(57) Fernandez, S. M.; Kellogg, B. A.; Poulter, C. D. *Biochemistry* **2000**, *39*, 15316–15321.
(58) Tarshis, L. C.; Proteau, P. J.; Kellogg, B. A.; Sacchettini, J. C.; Poulter, C. D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 15018–15023.
(59) Kellogg, B. A.; Poulter, C. D. *Curr. Opin. Chem. Biol.* **1997**, *1*, 570–578.
(60) Lynen, F.; Eggerer, H.; Henning, U.; Kessel, I. *Angew. Chem.* **1958**, *70*, 738–742.
(61) Perez-Brocal, V.; Gil, R.; Ramos, S.; Lamelas, A.; Postigo, M.; Michelena, J. M.; Silva, F. J.; Moya, A.; Latorre, A. *Science* **2006**, *314*, 312–313.
(62) Fujihashi, M.; Zhang, Y. W.; Higuchi, Y.; Li, X. Y.; Koyama, T.; Miki, K. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4337–4342.
(63) Popjak, G.; Edmond, J.; Wong, S. M. *J. Am. Chem. Soc.* **1973**, *95*, 2713–2714.
(64) Altman, L. J.; Laungani, D. R.; Rilling, H. C.; Vasak, J. *Chem. Soc., Chem. Commun.* **1977**, 860–861, No. 23.
(65) Radisky, E. S.; Poulter, C. D. *Biochemistry* **2000**, *39*, 1748–1760.
(66) Jennings, S. M.; Tsay, Y. H.; Fisch, T. M.; Robinson, G. W. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 6038–6042.
(67) Zhang, D.; Jennings, S. M.; Robinson, G. W.; Poulter, C. D. *Arch. Biochem. Biophys.* **1993**, *304*, 133–143.
(68) Jarstfer, M. B.; Zhang, D. L.; Poulter, C. D. *J. Am. Chem. Soc.* **2002**, *124*, 8834–8845.
(69) Maxwell, J. R.; Douglas, A. G.; Eglinton, G.; McCormick, A. *Phytochemistry* **1968**, *7*, 2157–2171.
(70) Huang, Z.; Poulter, C. D. *J. Am. Chem. Soc.* **1989**, *111*, 2713–2715.
(71) Blagg, B. S.; Jarstfer, M. B.; Rogers, D. H.; Poulter, C. D. *J. Am. Chem. Soc.* **2002**, *124*, 8846–8853.
(72) Staudinger, H.; Ruzicka, L *Helv. Chim. Acta* **1924**, *7*, 201–211.
(73) Staudinger, H.; Muntwyler, O.; Ruzicka, L.; Seibt, S. *Helv. Chim. Acta* **1924**, *7*, 390–406.
(74) Ruzicka, L. *Proc. Chem. Soc.* **1959**, Nov, 341−360.
(75) Alexander, K.; Epstein, W. W. *J. Org. Chem.* **1975**, *40*, 2576.
(76) Epstein, W. W.; Klobus, M. A.; Edison, A. S. *J. Org. Chem.* **1991**, *56*, 4451–4456.
(77) Rivera, S. B.; Swedlund, B. D.; King, G. J.; Bell, R. N.; Hussey, C. E., Jr.; Shattuck-Eidens, D. M.; Wrobel, W. M.; Peiser, G. D.; Poulter, C. D. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4373–4378.
(78) Hemmerlin, A.; Rivera, S. B.; Erickson, H. K.; Poulter, C. D. *J. Biol. Chem.* **2003**, *278*, 32132–32140.
(79) Thulasiram, H. V.; Erickson, H. K.; Poulter, C. D. *Science* **2007**, *316*, 73–76.
(80) Thulasiram, H. V.; Erickson, H. K.; Poulter, C. D. *J. Am. Chem. Soc.* **2008**, *130*, 1966–1971.
(81) Laughlin, S. T.; Baskin, J. M.; Amacher, S. L.; Bertozzi, C. R. *Science* **2008**, *320*, 664–667.
(82) Labadie, G. R.; Viswanathan, R.; Poulter, C. D. *J. Org. Chem.* **2007**, *72*, 9291–9297.
(83) Gauchet, C.; Labadie, G. R.; Poulter, C. D. *J. Am. Chem. Soc.* **2006**, *128*, 9274–9275.

JO900183C