CHEMICAL REVIEWS

Volume 97, Number 7 November 1997

Introduction: Polyketide and Nonribosomal Polypeptide Biosynthesis. From Collie to *Coli*

It is now 100 years since J. N. Collie first advanced the notion that certain classes of aromatic natural products might be derived from simple two-carbon "CH_2-CO" building blocks by way of linear poly-β-keto intermediates that could undergo cyclization by well-precedented carbonyl condensation reactions. The proposal, which has proven to be fundamentally correct, if not in precise detail at least in overall concept, was all the more remarkable because it was an intrinsically *mechanistic* hypothesis in an age when notions of chemical mechanism, let alone their relevance to biological phenomena, were barely perceived. It would be some 20 years before Robinson's revolutionary synthesis of tropinone provided a mechanistic framework for the understanding of alkaloid biosynthesis, and more than 60 years until Ruzicka's mechanistic reformulation of the Biogenetic Isoprene Rule.

The discovery of the fungal antibiotic penicillin and the recognition of its unprecedented effectiveness in combating bacterial infection led to the development of the modern fermentation products industry after the Second World War, resulting ultimately in the discovery of literally thousands of microbial metabolites with a broad range of pharmaceutical applications. These now encompass not only antibacterials, antivirals, and antitumor compounds, but also products with immunosuppressant, antihypertensive, and antihypercholesterolemic properties as well, not to mention the recognition of potent mycotoxins such as the highly carcinogenic aflatoxins. The polyketides comprise a significant fraction, not only of the total number of microbial metabolites which have been identified with physiological activities, but of the much smaller number which have found the greatest commercial application, among which one finds erythromycin, monensin, mevinolin, FK-506, avermectin, tetracycline, and daunorubicin. At the same time, the nonribosomal peptide antibiotics have become increasingly important, with the discovery of vancomycin, actinomycin, and cyclosporin, among others of this class. Indeed, the β-lactam antibiotics penicillin and cephalosporin are themselves derivatives of a single nonribosomal tripeptide, L-δ-aminoadipyl-L-cysteinyl-D-valine (ACV).

Experimental investigation of the biosynthetic origins of polyketides and polypeptides, as well as all other classes of natural products, first became possible with the widespread availability of radioactive isotopes of carbon and hydrogen. Such studies were dramatically accelerated in the early 1970s with the introduction of NMR of many stable isotopes, particularly ^{13}C. The next 10−15 years witnessed a veritable explosion in the number of reported studies of polyketide and polypeptide biosynthesis, resulting in the identification of the fundamental precursors of these metabolites and the first detailed insights into the manner in which the individual building blocks were assembled and subsequently modified.

By the late 1980s, it was evident that the assembly of both aromatic and reduced polyketides bore a close mechanistic resemblance to the by then well-understood formation of saturated fatty acids. According to the generally accepted theories of polyketide biosynthesis, the assembly of a polyketide chain from simple activated derivatives of acetate, propionate, and, occasionally, butyrate, involved the repetitive application of a small number of reactions analogous to those of fatty acid biosynthesis: chain elongation or β-ketoacyl synthesis, followed by some or all of a sequence of ketoreduction, dehydration, and reduction. Among fungal metabolites, it was also evident that these four reactions were also supplemented by the selective addition of methyl groups catalyzed by methyl transferases utilizing *S*-adenosylmethionine. On the other hand, the biochemical basis for the programming of polyketide biosynthesis was totally obscure, with no simple biological analogy available to explain the choice or number of building blocks and the order of their assembly, the control of the level of reduction following each stage of polyketide chain elongation, or the pattern of cyclization or lactonization of the resulting chain. Similar problems confronted those working on polypeptide biosynthesis. Thus, while the primary building blocks for polypeptides were clearly not only the 20 common amino acids but a variety of novel nonproteinogenic amino acids as well, the factors governing the choice of amino acid substrate, controlling amino acid epimerization, determining the timing and number of N-methylations, and dictating the ultimate length of the resulting polypeptide chain, remained a mystery. Moreover, a major stumbling block to further progress in both areas was the absence of viable

S0009-2665(97)00097-6 CCC: $28.00

enzyme preparations which could mediate the cell-free formation of the fundamental polyketide and polypeptide skeletons.

The formation of polyketide and polypeptides can be contrasted with the manner in which DNA, RNA, and protein biosynthesis is programmed. In each of the latter cases, a single biochemical reaction, be it catalyzed by a DNA or RNA polymerase or a ribosomal peptide synthetase, is used to add nucleic acid or amino acid monomers to a growing biopolymer. In each case, the instructions are encoded in a template DNA or RNA strand which controls the introduction of the complementary DNA, RNA, or tRNA substrate with high fidelity. Interestingly, ribosomal peptide synthesis can be artificially redirected by loading tRNAs with unnatural amino acids. By contrast, the biosynthesis of complex oligosaccharides and carbohydrates does not involve such templates and more likely must rely on as yet poorly understood issues of molecular recognition and substrate specificity. Similarly, no such simple template-mediated program could be invoked to explain polyketide or nonribosomal peptide biosynthesis.

The key to untying the Gordian knot has come in the last decade with the breathtaking advances in microbial molecular genetics. The pioneering work of the Hopwood laboratory on *Streptomyces* genetics and the development by Hopwood and his collaborators of reliable methods for recombinant manipulations in *Actinomycetes* has opened the way not only to the cloning and sequencing of entire gene clusters responsible for natural product biosynthesis, but has led to the expression and characterization of enzymes mediating many of the key steps of polyketide and polypeptide biosynthesis. Some of these proteins have turned out to be among the largest and most complex enzymes ever discovered. The recognition of the modular nature of polypeptide synthases and many polyketide synthases has not only provided a sound theoretical framework for understanding the programming of the formation of these metabolites, but has already led to the rational modification of the parent proteins, resulting in the formation of "unnatural natural products" of novel structure. On the near horizon lies the possibility to understand and control the formation of a vast array of natural products and the promise of a profound understanding of the marvelous ways in which Nature carries out organic chemistry.

For many decades, the study of polyketide and polypeptide biosynthesis took place in almost entirely different scientific worlds, with little cross-fertilization between laboratories or individual researchers, and few direct lessons from one field to the other. All this has changed in the last few years with the explosion in the understanding of the molecular genetics and enzymology of these two major classes of natural products. Quite unexpectedly, some remarkable similarities have been recognized between the formation of these chemically unrelated families of metabolites, from the modular organization of the structural genes and derived proteins, to even the reliance on a common class of acyl pantothenyl carrier proteins, united by a recently discovered superfamily of acyl carrier protein synthetases. This thematic issue brings together some of the leading investigators in both areas. The introductory review by Hopwood summarizes many of the most important recent and historical advances in polyketide molecular genetics. Reviews by Gould, Fujii and Ebizuka, Hutchinson, and Minto and Townsend cover recent dramatic developments in the understanding of aromatic polyketide biosynthesis, while Katz, Khosla, Ikeda and Ōmura, and Staunton and Wilkinson review the breathtaking advances in the study of modular polyketide synthases. Finally, Baldwin, Byford, Shiau, and Schofield; Marahiel, Stachelhaus, and Mootz; and von Döhren and his colleagues at the Technische Universität in Berlin (Keller, Zocher, and Vater) lay out the beautiful story of the workings of peptide synthetases. Just as important, each review has laid out many of the unanswered questions in this area. The next several years should see rapid progress in the understanding of these complex biosynthetic systems, from the rational manipulation of biosynthetic genes to the solution of the three-dimensional structures of numerous biosynthetic proteins. Complementing this work will be studies of the control and regulation of natural product biosynthesis, expression of natural product biosynthetic pathways in heterologous hosts from *E. coli* to plants and animals, and dramatic increases in molecular diversity through combinatorial biosynthesis of entirely new metabolites of the polyketide, peptide, and hybrid polyketide−polypeptide classes. At the same time, we can expect fundamental advances in our understanding of the evolution of natural product pathways, providing a new union of the fields of genetics and biogenesis.

David E. Cane
Brown University
Guest Editor

CR970097G