Programmed Organic Synthesis

An account of a symposium on computer-assisted design of organic synthesis.

Karl Heusler

The silent explosion in scientific literature has made it virtually impossible for a single mind to keep track of all the developments taking place, even in a narrow field of specialization. The storage and retrieval of information has now become a science in its own right. Apart from the relatively simple function of providing a repository for known facts, the main purpose of a storage and retrieval system must be to supply the impetus for the advancement of knowledge and of technological know-how. To achieve this aim, the available information must be digested, evaluated, combined, correlated, and extrapolated to new fields. This has hitherto been the unique domain of the human mind; but just as the human memory, which is no longer capable of storing all the information relevant to a given field, has to a certain extent been supplemented or supplanted by the computer's memory, it has been suggested that the time has now come when human intelligence must be partly replaced or reinforced by artificial intelligence.

It was against this background that a group of scientists active in the fields of organic chemistry and "computer chemistry" met in April 1972, at the Ciba Foundation in London (*1*) to discuss the possibility of applying artificial intelligence to their particular area of specialization. Organic synthesis seemed an especially appropriate topic for such discussions because in chemistry, and particularly in organic chemistry, a system of information storage was developed at a very early stage. This system is based on the simple concept of the empirical formula, which is internationally recognized and has been in use for more than 80 years. When supplemented by structural characteristics, it permits the description of every chemical compound that has been prepared in fairly exact terms, and in a form that not only can be understood by chemists all over the world, but also lends itself quite readily to computerization. Although it may be somewhat cumbersome to use, it is a workable system, and it can still cope with the ever-increasing host of new compounds synthesized each year. Obviously, it cannot accommodate all the information available in the chemical literature, but it does provide a clue to where additional data can be found once a specific objective has been defined in structural terms.

Since such a system for the storage and retrieval of information, as operated, for instance, by Chemical Abstract Service, is already available to organic chemists, the discussions at the Ciba Foundation Symposium could therefore be devoted to more specific questions. One of the central issues dealt with was the analysis of the intricate processes of selection and rejection that take place in the mind of the chemist confronted with this information, and how he can best apply it to advance knowledge and understanding of chemistry and to satisfy the ever-growing needs of society for new products and new materials. Although it is possible to gain an insight into the principles governing chemical changes and reactions without preparing new substances, more often the immediate goal of organic chemistry is synthesis. How then does previous experience, as recorded in the literature, influence the chemist's decision about the way a synthetic problem should be tackled? Is it possible to approach the planning in a logical way? Is there some rational means of evaluating all the various possibilities? If so, can synthesis or the planning of synthesis be programmed? It was to questions of this nature that answers were sought at the symposium.

The Mechanisms of Problem-Solving

About 5 years ago, Corey and Wipke (*2*) offered tentative answers to some of these questions. They concluded that "chemists employ a variety of problem-solving techniques of varying sophistication, depending both on the chemist and on the problem." In parentheses it might be mentioned that such problem-solving techniques are not dealt with in chemistry textbooks and are only rarely taught systematically in university courses. Corey and Wipke distinguished three extreme methodologies "which differ with respect to analytical and logical sophistication and generality" and which they termed "direct associative," "intermediate," and "logic-centered" molecular synthesis. They classify peptide synthesis in the first category, because the "building blocks"—the natural amino acids—are obvious, and the strategy of arranging them in the most effective order can hardly be arrived at by logical deduction but rather on the basis of empirical knowledge.

The "intermediate" approach involves the recognition of a relation between a critical and a major unit in the structure to be synthesized. An example of this type of synthesis is the preparation of cortisone from naturally occurring steroids, such as diosgenin or bile acids.

At the other end of the scale we find "logic-centered complex molecular synthesis." Here the possible building blocks are not obvious from the structure of the target and are very numerous. A rational and penetrative analysis of the molecular structure of the target compound is therefore necessary to find the best precursors and starting materials for its synthesis.

The questions raised earlier concerning the problems involved in the strategy of synthesis only apply to this last category. It therefore seems desirable to try to estimate the importance of this category in an economical and scientific way. In this category we find large rather than small molecules and complicated rather than simple structures. Of the economically important chemical structures, only for a small number of compounds, such as certain drugs, insecticides, or special dyes, will "logic-centered complex molecular synthesis" be required. However, in these instances, it will probably be of paramount importance to find the best possible technically feasible synthesis. In this context, the questions mentioned above are of more than academic interest. They can be reformulated and put in more general terms: Can we still rely on only the intelligence and experience of individual chemists? Would it not be more economical to feed their collective wisdom into a computer-operated system that could be used by the chemical community for the benefit of mankind?

The group of chemists assembled in London in April 1972 was eminently quali-

---

The author is director of research of the Pharmaceuticals Division, CIBA-GEIGY Limited, Basle, Switzerland.

fied to give an authoritative opinion on this question. The consensus was that in the area of "programmed synthesis" there is a large gap between what is required and what is possible. Even if the system under discussion were developed further, it seems unlikely that this gap could be easily closed. The difficulty relates not only to the technical part (computer hardware and software) but also, and even more so, to the conceptual part; that is, it was felt that the definitions essential to the creation of such programs are lacking, and very probably will be for a long time. However, it is worthwhile to consider why these conclusions were reached and to isolate those parts of the problem that mainly evoked skepticism. None of the scientists present were able to define the system or the strategy adopted when a synthesis of a complex organic molecule is conceived. It was generally agreed that some of the "logic-centered analysis" described by Corey and Wipke (*2*) and recently by Gelernter *et al.* (*3*) enters into it, but this is certainly only a part of it. Corey himself stated at the meeting that "it is infuriating to force your mind to go through the binary tree of search, which must be the basis of every computerized system. [The "tree" is the quasi-genealogical succession of precursors of a target compound, which themselves become target compounds for further precursors.] Our mind tends to go in bigger leaps, to use more complex processes." And Woodward [see (*1*)] noted that the speed with which the human mind rejects hundreds of possibilities without even knowing is remarkable. Furthermore, as Prelog [see (*1*)] pointed out, chemists are able to recognize, for instance, seven centers of chirality "in no time," which, by a logical step-by-step analysis, would be very difficult. Is the "intuition" which allows the "leap" simply a consequence of many years of experience? To some extent, this is certainly the case, but of equal certainty experience alone does not endow us with this intuition. It also became clear, however, that the mechanisms underlying the conception of synthetic pathways to complex organic structures are different from those used in other sciences. In retrospect, Corey distinguished two different ways of solving scientific problems. The first is analogous to the way in which we learn to cope with problems of mathematics or chess. There are well-defined procedures, and after some practice we attain the ability to solve problems relatively quickly by applying them. The second is the way in which organic chemists learn the "art" of synthesis. The problems are posed, but there are no established procedures. Autodidactically and subconsciously, through trial and error, they acquire their own problem-solving methodology, their planning experience.

Whether behind all this empirical knowledge and intuition there is a system that could be formalized and documented remained an open question. If there is, it must be a highly complicated one; but if the human mind could master it, might it perhaps not, as Prelog intimated, learn with proper training to perform Fourier transforms and thus recognize molecular structures in x-ray diagrams directly?

Although no agreement was reached, it was clear that efforts to unveil some of the mechanisms of problem-solving in organic synthesis would certainly be useful and interesting. They could help to make problem-solving itself more efficient and add a new dimension to the work of the organic chemist. With this in mind, the progress already made toward programmed planning of organic synthesis appears worthy of closer consideration.

Programs for Design in Organic Synthesis

It has become a habit to measure progress in scientific technology in "generations," as, for instance, in the design of computers, nuclear accelerators, or atomic power stations, where each new, more sophisticated, and more expensive generation has made its already quite expensive precursor obsolete. Accepting and extending for the time being such categories, one can only conclude that the "zero" generation, namely, that of the individual organic chemist with his more or less refined trial-and-error program of planning organic synthesis, is not yet past. A number of first-generation synthetic-design programs have, however, already emerged; these are obviously still in their infancy, and it will be one of our tasks to evaluate their chances of becoming strong enough to compete successfully with the zero generation. Among the first-generation programs are the LHASA program (Logic and Heuristics Applied to Synthetic Analysis) elaborated by Corey (*4*), the Princeton Program by Wipke (*5*), Hendrickson's approach to systematic synthesis design (*6–9*), and the program by Gelernter (*3*) for the discovery of organic synthetic routes by computer, with the use of an IBM 360/67 system.

Further generations of synthesis programs have not yet emerged, but some basic ideas are being put forward, and Ugi (*10*) thinks there is good reason to believe that the not too distant future will bring a second and third generation of synthetic-design programs, which will succeed the purely empirical first-generation programs. Ugi believes that the second-generation programs will be based on mathematical models of chemistry in combination with selection rules based on empirical information, and that the third-generation programs will rely on the potential energy surfaces of multiatom systems.

In this article, I discuss only the first generation of programs, but certain comments also apply to the second generation. Although some of its operations (especially the target-structure analysis, as discussed below) are "logical structure-oriented"—that is, expressed in exact mathematical and operational terms, the first-generation programs are, in their synthesis-planning part, to a large extent "information-oriented." It therefore seems necessary to assess the value and the scope of the information on which this part of these programs is built.

Information for Program-Planning of Organic Synthesis

1) *Structural information.* Allusion has already been made to the importance of the information available in organic chemistry for any planning of organic synthesis, and to the availability of structural information such as that compiled in the *Chemical Abstracts*. The possibility of converting structural formulas into computer language, directly or by means of chemical names and a simplified notation such as the Wiswesser System (*11*), and of performing the reverse operation, in which the chemical formulas are visualized by optical display systems, is a remarkable demonstration of the power and potential of the use of the computer in chemistry. The topological code developed by the Chemical Abstract Service even makes it possible to recognize partial structures within complicated molecules and represents a recent extension of these systems. This information forms the easily accessible basis of programmed planning of organic synthesis. In principle, it allows the storage in a computer memory of every chemical structure, including all its relevant characteristics, such as its composition, bonding, identification of the spatial or bonding relations between functional groups, its stereochemistry, conformation, and chirality. Only the first three properties can at present be satisfactorily coded, an achievement which is nonetheless remarkable. This information has a clearly and rigorously defined physical basis in the atomic structures of the elements, which determine their bonding types, distances, and angles. Since chemical reactions are always concerned with the formation or breaking of one or several bonds, it is pos-

sible to analyze a target structure systematically by making or breaking bonds. Each of these operations produces a new structure, which in turn can be stored and analyzed in the same way. This procedure furnishes an increasing number of precursors for the target molecule.

The formalistic, systematic dissection of molecules in this manner ("retrosynthetic" or "antithetic" in Corey's terminology) can, in theory, reveal all the possible precursors and starting materials required for the synthesis of the target molecule. But these dissections are, of course, of limited value unless each theoretical "antithetic" cleavage has a real "synthetic" counterpart reaction, making it possible to correlate the theoretical precursor with an actual synthetic precursor. It is therefore necessary to supplement the structural information with a reaction catalog.

2) *Reaction catalog*. Since—as Corey pointed out (*12*)—"unlike macro-construction the micro-construction (i.e., synthesis) of a molecule from constituent units does not and cannot depend on the mechanical placement of these units in an appropriate arrangement, the atomic groups to be joined must find each other and bind together in the proper way without external intervention or delivery. Thus it is not sufficient that the constituent synthetic units required for a synthesis be available, they must also possess chemical properties or affinities which are so specific as to allow selective combination in only one of the many possible modes." This is what chemists call "reactions." The special "chemical properties or affinities" of the fragments, which allow the desired reaction to proceed, are determined by certain structural features in these fragments (such as activating groups, and leaving groups). Some of these features can still be recognized in the reaction product. In principle, reactions can therefore also be described in structural terms. If the structural features in a target molecule that are typical of the final product of a known reaction are identified, then by reference to a "reaction library" (the catalog of reactions stored in the computer memory and containing structural information on final products and precursors) the structure of the "reactive" fragments that go to make up the target molecule can be reconstituted. Whereas the structural information can be rigorously defined, coded, and analyzed and has a firm physical background, the fact that only part of the reaction information may be available (that is, the part recognizable within the structure of the target molecule) makes this type of correlation less reliable.

In fact, the above deductions are based on the tacit assumption that there are certain reactions that, when applied to clearly defined structures, will always succeed, although to a varying and probably unpredictable degree. Obviously, the yield of a given reaction in a specific case will, to a large extent, depend on the experimental conditions, which will have to be worked out in each case individually. But the generality of reactions is an axiom on which the system is built. This statement is not invalidated by the fact that in every approach to programmed synthesis it is recognized that occasionally "general reactions" cannot be applied to a specific starting material until suitable structural adaptations have been made, such as the introduction of protecting or activating groups. The difficulty still remains, however, that the differences between a desirable "general" reaction and an equally "general" but, for the purpose in hand, undesirable competitive reaction are often very slight indeed. The quantification of these differences is a problem to be dealt with below.

The technical difficulties involved in the computerization of such processes should not be underestimated. The computer experts present at the symposium in London repeatedly expressed the opinion that the elaboration of a fully consistent and complete system of information storage (either in the structure or in the reaction part) is a formidable task, and that such a system could only operate efficiently if control loops or similar circuits were built into it (refinements that have not been accomplished in any of the programs developed so far). Completeness is not likely to be reached in such a complicated area as organic chemistry. Some years ago, an attempt was made to compare the various types of indexing systems in a specialized field of aeronautics. Although the field was narrow and the literature limited and although a considerable effort went into the indexing, none of the systems proved satisfactory for adequate retrieval of the information.

If it is assumed that these difficulties will eventually be surmounted and some means will be found of storing the structural information as well as an appropriate reaction catalog in a computerized system, then the last step toward the implementation of the system is to match the two.

3) *The matching process*. Recognizing structures and envisaging the reactions by which these structures can be made is the basis of "synthesis." The process of conceiving a synthetic approach is strictly a step-by-step process. The target structure is identified, possible precursors are listed with the aid of the "reaction library," and these new structures are then considered as new final products, giving a further series of precursors. In this way, a "tree of synthesis" grows. It is quite clear, and especially so to all scientists who have approached the problem in this fashion, that the tree cannot be tended entirely by a machine, and its frequent inspection by the chemist is advisable. Corey himself prefers to use the term "computer-assisted design of organic synthesis" rather than "programmed synthesis." It is the chemist who determines when the tree should not be allowed to ramify any further; he decides which branches should be pruned because they are not likely to bear fruit. But—and this is one of the most crucial questions discussed at the symposium—how many of the decisions can be left to the machine and how many must be made by the chemist?

Evaluation

The yield must be evaluated qualitatively as well as quantitatively. In principle, every possibility emerging from the combination of the data stored in the structure-and-reaction library should be pursued. Any organic chemist given the structure of a chemical compound sufficiently complex to be a realistic candidate for computer-assisted synthesis can quickly jot down 10 to 20 single precursors that in a one-step reaction can be transformed into the target structure. If the same is done for each step in a ten-step synthesis, the number of starting structures will run to between 10^{10} to 10^{20}. If one includes synthetic detours temporarily introducing structures that appear neither in the target compound nor in the starting material, but are more than simple protecting groups (for example, structures designed to bring certain parts of a molecule into juxtaposition to facilitate a desired reaction—a device well known to the organic chemists), then the number of possibilities becomes virtually infinite. Pruning the tree therefore becomes an absolute necessity, but it must be done in such a way that all the useful synthetic schemes remain.

How efficiently this can be done will ultimately decide the usefulness of the whole system. Not surprisingly, it was this question of evaluation to which discussion at the 1972 conference in London always reverted. Obviously the zero-generation method, the chemist's mind, is the least well defined, yet it is often very successful. It combines experience with intuition, and semiempirical physicochemical rules and correlations in thermodynamics and kinetics with applied quantum mechanics, giving each criterion its appropriate weight, in a not strictly systematic but nevertheless sophisticated way. Since the weight and the relevance of each criterion in each instance and in each step is not rigorously defined by any mathematical terms, and since the weights change or are adapted year by year, month by month, or day by day, as new knowledge accumulates, the versatility of the zero-generation method still counts in its favor.

The first-generation method of Corey, Gelernter, and others, attempts to systematize experience by introducing factors intended to indicate whether a particular synthetic reaction will have a more or less reasonable chance of success. In Corey's system, each reaction is given a (target-independent) numerical rating, which is then coupled with a "set of conditional statements (qualifiers) which cause the basic rating to be increased or decreased by certain amounts if certain structural features are present in the particular target structure." Although the actual criteria, based on our present-day experimental and admittedly often rudimentary theoretical knowledge, according to which these ratings are allocated, are fairly rigidly built into the system and thus not rapidly adaptable to new knowledge, it is still the user's decision whether to apply them to his case. The user can choose for himself the "cutoff" value of the rating; that is, he can decide the value below which the rating must be before a particular branch of the synthetic tree is cut off. If, however, the user so chooses, he can still produce the unpruned tree, provided that the time required for its production by computer and the size of the tree does not prevent this.

The SYNCHEM system of Gelernter and collaborators (*3*), which in the authors' own words is still "in a rudimentary stage of development," also recognizes the dualism of solution generation and solution evaluation, as well as the ill-defined nature of the criteria for pruning the problem-solving tree of "unpromising and redundant subgoals." However, it sets itself rather ambitious aims, because the evaluation part is supposed to take into consideration the "probable yield, convenience, and reliability of a reaction," including also the number of distinct chemical processes (possibly requiring separation, purification, or concentration operations after each step) as well as the relative yields of competitive reactions and by-products. The problem of making quantitative rather than qualitative statements in these areas is formidable, and it seems difficult to share the opinion of the authors of SYNCHEM. One feature of the SYNCHEM system of particular interest to the chemist is the fact that it includes a "starting-material library" of substances that are commercially readily available. If the search leads to such a starting material, that particular branch of the tree stops growing.

A different approach has been put forward by Hendrickson. After proposing a system for organizing both structures and reactions in a common format (*6*), easily adaptable to computerization, he proceeded to a careful investigation of the scope of the problem (*7*), using graph theory as a basis for his calculations. The first evaluation was an analysis of the number of theoretical possibilities for the construction of the ring structure of some organic molecules, disregarding functional groups and stereochemistry. By way of illustration, it might be mentioned that the number of precursors deduced for possible syntheses of the camphor skeleton from three compounds (of which the classical synthesis by Komppa is one) involving three combination steps (that is, ignoring functional modifications between these steps) is 295—a rather large number, considering the restrictions imposed, but even so still not allowing for ring enlargements or contractions, nor for the construction of extra rings which are cleaved later on in the synthesis. For the construction of the camphor skeleton from single carbon "synthons" ("coal, air, and water") Hendrickson arrives at a number of about 40 million total routes (to be exact: 39,916,800). Other pertinent numerical examples relating to the hydroazulene, steroid, or tricyclane skeletons are to be found in Hendrickson's paper (*7*). Extension of these calculations leads to the conclusion, already put forward by Longuet-Higgins in London in 1972 [see (*1*)], that even for relatively simple organic molecules the total synthesis tree is prohibitively large and, in addition, ill-defined.

Hendrickson's proposal for dealing with this situation is threefold (*8*), his restrictions on the growth of the synthesis tree being independent of any empirical data. They are (i) a consolidation of the reaction catalog by generalizing minor variations in synthetic reactions into single units (whereby reactions are considered only as structural changes and not in mechanistic terms), (ii) isolated consideration of subtrees that are independent of each other, and (iii) the selection of "efficient sequences" according to such criteria as the preference of convergent over linear sequences, the lowest number of refunctionalization steps in relation to the number of construction steps, and the minimal number of construction steps.

Hendrickson's approach (*9*) avoids some of the pitfalls of the earlier approaches: it does not use unreliable empirical data but tries to express restrictions in exact mathematical terms. It remains to be seen whether these restrictions are really sufficient to limit the total of possible synthesis routes to a manageable number and whether this number is such that it offers advantages over the zero-generation method. However, the guidelines formulated by Hendrickson in respect to his third restriction criterion are certainly present in the mind of any organic chemist engaged in synthesis planning. It is the degree to which these criteria are followed that distinguishes elegant syntheses from dull and cumbersome ones, and differentiates the artist from the artisan in chemical synthesis.

At the other end of the spectrum of attempts at solving the problem of "planned synthesis" stands Ugi's approach (*10*). He stated that, in general, satisfactory approximations to a complete set of solutions of chemical problems such as synthesis could not be obtained with the aid of artificial intelligence, unless a unified (mathematical) theory of constitutional chemistry were to become available which could be used in a single and effective manner as a basis of computer programs. This statement remained unchallenged, but the majority of those present at the London meeting felt that we were much too far away from such a unified theory of chemistry to give serious consideration to any application based thereon. At a recent Roussel Round Table Conference in Paris, it was clearly stated that at present not even triatomic molecules can be satisfactorily described in mathematical-physical terms. It would be fanciful or at least extravagantly optimistic to hope that it might become possible to describe complex organic molecules of the type for which planned synthesis would be of interest in exact mathematical terms within the foreseeable future. Ugi's exercise, interesting as it may be, lacks the solid foundation on which to build a viable program. Even if that foundation did exist, how to deal with the large number of possibilities would still remain a problem.

Conclusions and Outlook

It is the purpose of this article to give an account of the situation regarding chemical synthesis and its interface with computer sciences. Chemical synthesis, to which modern life owes much in terms of comfort, health, and nutrition, but which has also upset the natural balances in the industrialized world, will remain an important field of endeavor for many scientists, whose sense of responsibility and soundness of judgment will determine the extent to which mankind may hope to benefit from future developments. But chemical synthesis also embodies an element of adventure and of art: it is a manifestation of human intellect. It therefore seems essential that we should consider how important this aspect really is and whether, in the future, synthesis could be more successfully directed by artificial intelligence and computers. Although not expressly formulated at the Ciba Foundation Symposium in 1972 by the panel of experts in chemical synthesis, these questions were in the minds of many of the participants.

The discussions showed that new developments are not going to change the scene in organic chemical synthesis within the next decade. There is no doubt, however, that an intensive effort will be made to systematize some areas of chemistry, and that such efforts will directly or indirectly stimulate activity in this field.

It is a fact that the analysis of synthesis planning has been neglected for a very long time. Yet such analyses are certainly of tremendous educational value, and Corey himself maintains that, even if the computerized systems do not become operative and useful, the insight they will afford into the obscure, intuitive ways in which synthetic schemes are conceived by the masters of organic synthesis will have a definite and positive influence on the quality of future synthetic work. Certain general principles, such as the concept of pseudosymmetry in chemical intermediates and the principle of convergent synthesis, although very difficult or even impossible to translate into mathematical language, nevertheless have a real meaning to organic chemists. The application of such principles to synthetic problems and the semisystematic evaluation of synthetic schemes will, it may be hoped, become part of a new educational system in organic chemistry. This immediate effect of the project discussed in this article is considered of such importance that the whole effort put into it will certainly have been well worthwhile.

One final point still remains to be stressed (as it was in a postsymposium letter by W. S. Johnson of Stanford University) and that is the importance of distinguishing between two different aspects of accomplishing a synthetic objective, namely between the planning and the execution of a synthetic scheme. As Johnson quite rightly puts it, "the present state of the art of synthesis is such that all well-planned synthetic schemes almost invariably fail to give the envisaged results at one, or more often, several stages. Thus a successful synthesis seldom follows and sometimes diverges dramatically from the original plan. It is this 'fallibility phenomenon' which renders organic synthesis at least as much of a creative challenge at the execution stage as at the planning stage." Thus, even if planning can be fully systematized, the second challenge will remain: that is, basically chemistry is and will always be an experimental science.

References and Notes

1. "Programmed Organic Synthesis," 24th Annual Lecture and Symposium, 13 and 14 April 1972, Ciba Foundation, London. The participants were E. J. Corey, C. Djerassi, A. Dreiding, A. Eschenmoser, J. Harley-Mason, K. Heusler, W. S. Johnson, H. C. Longuet-Higgins, L. N. Mander, J. Mathieu, O. S. Mills, D. A. Pensak, V. Prelog, R. A. Raphael, F. Sondheimer, D. W. Turner, I. Ugi, and R. B. Woodward.
2. E. J. Corey and W. T. Wipke, *Science* **166**, 178 (1969).
3. H. Gelernter *et al., Top. Curr. Chem.* **41**, 113 (1973).
4. E. J. Corey, W. T. Wipke, R. D. Cramer III, W. J. Howe, *J. Am. Chem. Soc.* **94**, 421 (1972).
5. W. T. Wipke and P. Gund, *ibid.* **96**, 299 (1974); W. T. Wipke, in *Proceedings of the Conference on Computers in Chemical Education and Research*, Northern Illinois University, 1971.
6. J. B. Hendrickson, *J. Am. Chem. Soc.* **93**, 6847, 6854 (1971).
7. ———, *ibid.*, in press.
8. ———, *ibid.*, in press.
9. ———, *Topics in Current Chemistry. A General Protocol of Systematic Synthesis Design* (Springer, New York, in press).
10. I. Ugi and P. Gillespie, *Angew. Chem.* **83**, 982, 990 (1971); *Angew. Chem. Int. Ed. Engl.* **10**, 914, 915 (1971); I. Ugi, *Trans. N.Y. Acad. Sci.* **34**, 416 (1972).
11. E. G. Smith, *The Wiswesser Line-Formula Chemical Notation* (McGraw-Hill, New York, 1968).
12. E. J. Corey, *Q. Rev. Chem. Soc. Lond.* **25**, 455 (1971).

Marine Biological Laboratory: Origins and Patrons

The 50th anniversary of the Main Building recalls that it is a creation of the National Research Council.

Detlev W. Bronk

Fifty years ago, in July 1925, the Main Building of the Marine Biological Laboratory at Woods Hole was dedicated. The events which led to that occasion are significant in the early history of the Carnegie Institution of Washington, the National Research Council, and philanthropic foundations that have had a profound influence in the furtherance of science.

The preceding quarter century had been a period of remarkable change and growth of the scientific endeavor in the United States. During that time two pioneer research institutions were created: the Carnegie Institution and the Rockefeller Institute for Medical Research. The Carnegie Corporation and the Rockefeller Foundation that were established in 1911 and 1913, respectively, enabled universities to expand their facilities for research. Barriers between sciences were being crossed in fields such as biochemistry and biophysics. The National Research Council (NRC) was organized by the National Academy of Sciences because, during the war of that period, the NRC had demonstrated its capacity for large service; it was being critically tested as a useful organization for the cooperative furtherance of science in times of peace. It was during this period and through these institutions that the Marine Biological Laboratory (MBL) developed the resources and status that made possible its phenomenal growth in this past half century.

The Woman's Education Association

Now that the role and rights of women in science are much discussed and the NRC has formed a committee on the Education and Employment of Women in Science and Engineering, it is timely to recall that some energetic, visionary women were largely responsible for creating the MBL. They were members of the Woman's Education Association of Boston, which was founded in 1871. At that time, when the teaching of science for women was in its infancy, the Association persuaded the Massachusetts Institute of Technology to admit women to courses in chemistry. One of

---

The author is President Emeritus of Rockefeller University, New York 10021. He was formerly president of the National Academy of Sciences.