between these two views; both are useful for discovering parts of the whole. I hope that students today are being trained with equal emphasis on analytic and numerical methodologies.

Prospectus

In summary, the prodigious growth in computing power is ushering in new approaches to complexity in many areas of science. Although the shift of methodology and aesthetics was foreseen by von Neumann over 30 years ago, the fulfillment of his vision is only beginning. For his vision to be realized, there are two major requirements. First, computers must continue their rapid rate of increase in speed so that more and more complex problems can be attacked on human time scales. Second, there must be much greater accessibility to the full range of computational tools that are needed so that a "critical mass" of scientists can work in each field of interest. Both of these requirements are likely to be met.

References and Notes

1. G. Birkoff, *SIAM Rev.* **25**, 1 (1983).
2. J. Glimm, in *Frontiers of Supercomputing*, N. Metropolis *et al.*, Eds. (Univ. of California Press, Berkeley, in press).
3. These colleagues include J. Centrella, B. DeWitt, P. Dykema, K. Eppley, C. Evans, J. Hawley, J. LeBlanc, M. Norman, J. Wilson, and K.-H. A. Winkler.
4. M. L. Norman, L. Smarr, K.-H. A. Winkler, M. D. Smith, *Astron. Astrophys.* **113**, 285 (1982); M. L. Norman, K.-H. A. Winkler, L. Smarr, in *Astrophysical Jets*, A. Ferrari and A. G. Pacholczyk, Eds. (Reidel, Dordrecht, 1982), p. 227; M. L. Norman, L. L. Smarr, K.-H. A. Winkler, in *Numerical Astrophysics: A Festschrift in Honor of James R. Wilson*, J. Centrella, J. LeBlanc, R. Bowers, Eds. (Jones and Bartlet, Portota Valley, Calif., 1985), p. 88; L. L. Smarr, M. L. Norman, K.-H. A. Winkler, *Physica D* **12**, 83 (1984); M. L. Norman, K.-H. A. Winkler, L. L. Smarr, in *Physics of Energy Transport in Extragalactic Radio Sources*, A. Bridle and J. Eilek, Eds., in preparation.
5. This work was performed while J. Hawley was a graduate student in the Astronomy Department at the University of Illinois. He is currently the Bantrell Fellow at the California Institute of Technology.
6. M. C. Begelman, R. D. Blandford, M. J. Rees, *Rev. Mod. Phys.* **56**, 255 (1984).
7. J. R. Wilson at the Lawrence Livermore National Laboratory was the first person to write a computer program that solved the general relativistic hydrodynamics equations for this problem; see J. R. Wilson, *Astrophys. J.* **173**, 431 (1972). Hawley, Wilson, and I wrote an updated version of the program and carefully calibrated it; see J. Hawley, L. L. Smarr, J. R. Wilson, *Astrophys. J.* **277**, 296 (1984); *Astrophys. J. Suppl. Ser.* **55**, 211 (1984).
8. J. Hawley, "A numerical study of nonspherical black hole accretion" [thesis, University of Illinois (1984)]; J. F. Hawley and L. L. Smarr, *AIP Conf. Proc.*, in press. Only black and white images are in these two sources.
9. In this discussion we assume that the black hole is nonrotating. Our program also handles the case of a rotating black hole. See (*8*) for details.
10. K.-H. A. Winkler developed the color-image algorithm for presenting the supercomputer solutions. J. Hawley adapted the approach and rewrote the algorithm for our project. The calculations were made on the Cray-1 supercomputer at the Max-Planck Institute for Physics and Astrophysics, Institute for Astrophysics.
11. B. Paczynski, *Astron. Astrophys.* **34**, 161 (1974).
12. M. M. Waldrop, *Science* **223**, 575 (1984).
13. The movie was made at the Max-Planck Institute for Physics and Astrophysics, Institute for Astrophysics.
14. N. J. Zabusky, *Phys. Today* **37**, 36 (July 1984); *J. Comp. Phys.* **43**, 195 (1981).
15. S. Wolfram, *Sci. Am.* **251**, 188 (September 1984).
16. W. H. Press, "Report by the Subcommittee on Computational Facilities for Theoretical Research to the Advisory Committee for Physics, Division of Physics" (National Science Foundation, Washington, D.C., 1981).

17. I thank J. Hawley for permission to use the unpublished color images in this article. My colleagues have been instrumental in the development of the ideas expressed in this article. In particular, I thank J. Centrella, C. Evans, J. Hawley, M. Norman, W. Press, R. Wilhelmson, K. Wilson, J. Wilson, K.-H. A. Winkler, and N. Zabusky. The hospitality and support of the Max-Planck Institute for Physics and Astrophysics, Institute for Astrophysics, and its director, R. Kippenhahn, are greatly appreciated. The images were made both at the Institute for Astrophysics and at the University of Illinois VAX and Image Processing Center. This work was partially supported by the National Science Foundation and the Alfred P. Sloan Foundation.

Computer-Assisted Analysis in Organic Synthesis

E. J. Corey, Alan K. Long, Stewart D. Rubenstein

The chemical synthesis of organic molecules has proceeded at an accelerating pace for more than a century and a half. Since the Wöhler synthesis of urea in 1828, organic synthesis has had an enormous impact on civilization and on the development of science itself. Advances in the understanding of chemical structure, chemical reactivity, stereochemistry, and biochemistry have been due in no small part to discoveries in organic synthesis. Yet all of the syntheses of the 19th century and most of those of the first half of this century were developed from a relatively primitive conceptual base. Many syntheses, especially in the 19th century, were discovered serendipitously (or opportunistically) in the sense that they were accomplished as unplanned results of exploratory studies of chemical interactions between different types of molecules. Other syntheses used a series of established reactions to convert a basic structure into a somewhat larger molecule. Such syntheses involved the successive attachment of functional or substituent groups through the use of replacement, condensation, or coupling reactions. Thus, the dyes alizarin (1869) and indigo (1890) were synthesized by elaboration of anthracene and aniline, respectively, and the alkaloid tropinone was made from cycloheptene (1901). In contrast to the vast majority of these early syntheses, which were based on the availability of starting materials that contained a major portion of the final atomic framework, a few syntheses emerged whose design depended on the knowledge of certain ring-forming reactions that could be used to build an atomic framework. Among the best examples of these are the syntheses of α-terpineol (W. H. Perkin, 1904), camphor (G. Komppa, 1903; W. H. Perkin, 1904), tropinone (R. Robinson, 1917), and equilinin (W. Bachmann, 1939) (*1*).

In the post–World War II period, synthesis attained a different level of sophistication partly as a result of the confluence of five stimuli: (i) the formulation of detailed electronic mechanisms for the fundamental organic reactions, (ii) the introduction of conformational analysis of organic structures and transition states based on stereochemical principles, (iii) the development of spectroscopic and other physical methods for structural analysis, (iv) the use of chromatographic methods of analysis and separation, and (v) the discovery and application of selective chemical reagents. As a result, the years 1945 to 1960 saw the accomplishment of a number of highly sophisticated syntheses of complex molecules, including vitamin A (O. Isler, 1949), cortisone (R. Woodward, R. Robinson, 1951), strychnine (R.

E. J. Corey is a professor, Alan K. Long is a research associate, and Stewart D. Rubenstein is a graduate research assistant in the Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138.

Woodward, 1954), cedrol (G. Stork, 1955), morphine (M. Gates, 1956), reserpine (R. Woodward, 1956), penicillin V (J. Sheehan, 1957), colchicine (A. Eschenmoser, 1959), and chlorophyll (R. Woodward, 1960) (*2*).

These achievements provided the impetus for further developments, which led to a great improvement in the power and elegance of synthetic planning. In the 1960's, general problem-solving strategies and methods that could be applied to the analysis of any complex synthetic problem were explicitly formulated, and the underlying principles of synthesis were defined in a way that made synthetic planning more logical, more systematic, and easier. The insights so gained had an impact on the teaching of organic synthesis as well as its practice. Even in the 1950's, synthesis was taught by the presentation of a series of illustrative (and generally unrelated) examples of actual syntheses. Chemists who learned synthesis in this inductive manner approached each problem in an ad hoc way. The intuitive search for clues to the solution of the problem at hand was not guided by effective and consciously applied general problem-solving techniques.

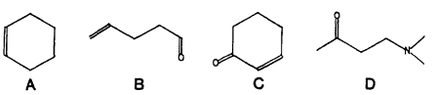
General Approach to the Analysis of Complex Synthetic Problems

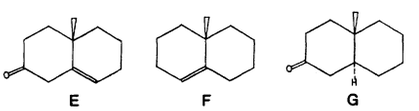
Defining a methodology that assists the chemist in the solution of any problem in organic synthesis is a more challenging task than developing a synthesis of a particular target structure. Such a methodology must be capable of, first, guiding the search for an effective strategy for initial simplification of a problem and, second, generating a large number of specific pathways of synthetic intermediates (or synthetic steps) that connect chemical starting materials with a target structure. A separate issue, which will not be addressed here, is the highly demanding task of reducing a synthetic plan to practice and demonstrating its validity and merit by experiment.

The importance of distinguishing between the two possible directions of a synthetic analysis was recognized about 20 years ago (*3–6*). The direction termed "synthetic" corresponds to the direction of execution of a chemical synthesis, and the direction referred to as "retrosynthetic" or "antithetic" is that in which a target structure can be transformed in a succession of steps through simpler structures into a starting point for a synthesis. The former direction of analysis lends itself to the design of "building-block" syntheses from easily definable or obvious starting materials. The latter, retrosynthetic analysis, is not only more powerful but is also suitable for complex problem solving (*5, 6*).

Retrosynthetic analysis depends on perception of the structural features of the reaction products (as opposed to those of the starting materials) and consideration of molecular changes in the retrosynthetic sense. To avoid confusion, special terms and a special graphic device are used to provide a distinction between the nomenclature appropriate to these analyses and that which is conventionally applied to synthesis in the direction of laboratory execution (Table 1) (*6*).

Basic retrons for the Carbo Diels-Alder, Claisen rearrangement, Robinson annulation, and Mannich transforms are represented as A, B, C, and D, respectively. All the structural elements in the retron for a transform must be present before the transform can be logically applied. The presence of part of a retron ("subretron" or "part-retron") within a target structure can, however, signal the possible application of a transform providing other transforms can be invoked to generate, by one or more steps, a precursor containing the full retron. Structures E, F, and G, for example, contain valid part-retrons for the Robinson annulation transform.





The purpose of retrosynthetic analysis is to generate a tree of synthetic intermediates that terminates with a number of alternative simple starting materials. The manner in which the retrosynthetic tree is generated will be a function of the strategies, search techniques, and chemical information used. Because of the great richness of information in organic chemistry, there is a danger of generating retrosynthetic trees which are too large to be manageable and which include numerous dead ends. Prevention of this combinatorial explosion requires the use of simplification techniques, search strategies, screening procedures, and logical analysis.

One way that synthetic problems can be simplified involves a type of substructure matching in which subunits of a target molecule are recognized as potential "synthons," or building blocks. Such identification not only provides a strong control over tree branching but also permits the use of a bidirectional search, in which trees of chemical intermediates are simultaneously grown in the retrosynthetic and synthetic directions (*6*). The elaboration of one tree is guided and facilitated by the development of the other. Reiterative cycles of perception and analysis are applied to the structures being examined and to the body of relevant chemical information, as they are in all phases of synthetic analysis.

A number of other key elements for the initial simplification of complex synthetic problems can be clearly identified, including: (i) identification of especially appropriate strategies for retrosynthetic analysis, (ii) use of available information on the chemistry of the target molecule, (iii) analysis of the target molecule in terms of its relation to other molecules or structures, either previously synthesized or not, (iv) characterization of the problem in terms of unusual features in the target structure, and (v) discovery of two or more high level strategies that can be used concurrently.

Strategies in Retrosynthetic Analysis

Since the success of retrosynthetic analysis depends on the selection and application of key transforms, it is necessary that the large body of knowledge of organic chemistry be appropriately

Summary. The planning of alternative routes for the synthesis of complex organic molecules has been facilitated by the formulation of guiding strategies that can be applied to a broad range of problems. Analysis of organic synthesis can be carried out in the retrosynthetic direction, opposite to the actual process of chemical synthesis, or bidirectionally, that is, as a combined retrosynthetic and synthetic search. An interactive computer program is described which utilizes the general strategies of retrosynthetic analysis and an appropriate database to generate pathways of chemical intermediates for chemical synthesis of a particular target structure. Computer graphics and standard chemical structures are used for man-machine communication.

Table 1. Terminology of synthetic and retrosynthetic analysis.

Synthetic analysis	Retrosynthetic analysis
Direction of laboratory execution is "synthetic."	Direction of logical analysis is "retrosynthetic" or "antithetic."
Represented with a single arrow (→).	Represented with a double arrow (⇒).
Process is a "reaction."	Process is a "transform," given the name of the corresponding reaction.
Reaction is keyed by the presence of specific reactive groups in the starting material(s) that "activate" the substance to undergo a synthetic step.	Transform is keyed by the presence of critical subunit, or "retron," that "actuates" the application of the transform to that target structure.

structured and made accessible. For each transform it is essential to know not only the keying features (full retrons, part-retrons, and so on) but also the factors that affect application of the transform in the context of the structural situation. Two of the most important of these factors are the power of the transform to simplify the structure in the retrosynthetic direction and the function of the transform, that is, its effect on the molecular skeleton, functional groups, or stereocenters. The most powerful synthetic reactions are those that reliably increase molecular complexity. The corresponding simplifying transforms, for example the Diels-Alder transform, are so powerful that their application to a particular target structure may be taken as a strategic goal of retrosynthetic analysis ("transform goal" or "T-goal"). A list of powerful simplifying transforms is shown in Table 2.

Transforms may decrease or increase molecular complexity and for this reason may play quite different roles in retrosynthetic analysis. The change effected by a transform may involve connection, disconnection, rearrangement, interchange, transposition, removal, addition, inversion, or other modifications. Transforms must be multiclassified according to these functions so that they may be made accessible for a variety of strategic reasons. Most transforms are much less powerful than those in Table 2, but they may be indispensable as "subgoals" to set up retrons for the more powerful transforms. Knowledge of the power and functions of the transforms in the database is important for the selection and execution of retrosynthetic strategies.

Five Types of Strategies for Retrosynthetic Analysis

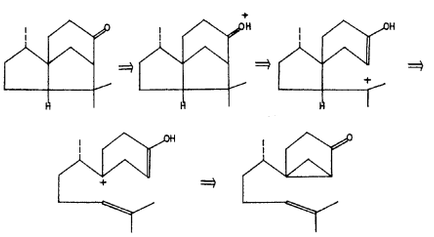
Transform-based strategies. The identification of a powerful simplifying transform that is well suited to a particular target molecule can produce a line of analysis resulting in an excellent solution. Normally, the retron required for direct application of that transform will not be present, and a multistep search or look-ahead will be needed for the successful application of such a transform. In most cases there are several potential locations for the retron, and all possible mappings need to be explored. For example, six [4 + 2] disconnections must be examined in rigorous attempts to apply the Diels-Alder transform to a particular six-membered ring (*6, 7*). Sometimes two or three powerful simplifying transforms can be applied successively to a target structure. In such situations quite elegant solutions can result. The recognition of such cases, though obviously not simple, is greatly facilitated by the purposeful attempt to apply this type of strategy. Another strategy is to repeatedly apply the same simplifying transform (*8*).

The key transforms for which a long-range search is worthwhile, as mentioned above, correspond to powerful synthetic reactions—those in which rings or ring pairs are formed, those which are stereocontrolled and position selective, those which bring together two or three major subunits, or those which can be used efficiently and economically on a large scale. There are effective techniques for selecting such transforms and for systematically and exhaustively carrying out the necessary long-range search processes. One such technique is to generate a structure in which the necessary retron (or part-retron) has been mapped onto the target in a particular way and then to proceed with a bidirectional search. This procedure would be repeated for all possible mappings.

Table 2. Some powerful simplifying transforms.

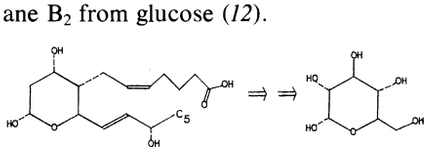
Carbo Diels-Alder
Quinone Diels-Alder
Hetero Diels-Alder
Robinson annulation
Position-selective partial aromatic reduction
Cation π-cyclization
Radical π-cyclization
Aldol cyclization
Sila-acyloin cyclization
Internal S_N2 cyclization
Internal nucleophilic acylation
Internal ene reaction
Internal cycloaddition: [4 + 2], [3 + 2], [4 + 3], [2 + 2], [2 + 1]
Pericyclic cation or anion closure
Sigmatropic rearrangements
Photocyclizations
Enantioselective π-addition
Diastereoselective π-addition
Fischer indole, Knorr pyrrole, and so on

"Mechanistic transforms" are another category of transform-based strategies. They are simply the retrosynthetic version of step-by-step electronic mechanisms. In applications of a mechanistic transform some structural objective is required for guidance (for example, the breaking of a certain bond or the removal of a certain obstacle to T-goal realization). The target is converted to a reactive intermediate or a synthetic equivalent from which the target would result synthetically, and then other reactive intermediates are generated mechanistically until a stable precursor is produced. An example of molecular simplification by mechanistic transform application is illustrated below (*9*).



Transform-based (T-goal) strategies are used effectively in the design of almost all complex contemporary syntheses (*10*).

Structure-goal (S-goal) strategies. The identification of a potential starting material, building block, retron-containing subunit, or initiating chiral element (*6, 11*) can lead to an S-goal from which a coordinated, bidirectional search would generate a synthetic sequence. An illustration is provided by the retrosynthetically designed synthesis of thromboxane B_2 from glucose (*12*).



Topological strategies. These strategies are used to identify one or more bonds whose disconnection can lead to major molecular simplification. These bonds may be in bridged- or fused-ring cyclic systems, or they may appear as connectors to appendages at rings, functional groups, or stereocenters (*13*). The recognition of such strategic disconnections is made possible by the application

of heuristic criteria derived from previous synthetic success. In the same way, it is possible to recognize one particular ring out of several in a polycyclic structure as a critical or strategic ring for early retrosynthetic disassembly. Conversely, it is possible to identify one or more rings which should be preserved intact during retrosynthetic simplification. Such topological strategies have been formulated in explicit terms (*13*) and applied to synthetic design (*14–16*).

Stereochemical strategies. These are heuristically derived procedures for reducing stereochemical complexity in the retrosynthetic direction. These procedures are used to remove stereocenters by applying stereospecific transforms or by taking advantage of steric screening or proximity between functional groups in the target structure. Stereocenters may be removed one at a time, in pairs, or in groups. In addition, stereochemical strategies can be used for directing the search toward enantio- or diastereoselective transforms (*17*). Stereoselective strategies may also guide the application of connective (ring-forming) transforms as subgoals that allow more effective retrosynthetic removal of stereocenters. A proximity between functional groups may allow simplification of the target by the application of functional group removal (FGR) transforms (*18*).

Functional group–oriented strategies. Many such strategies have been derived from synthetic experience. For example, one or more functional groups together with an interconnecting "atom path" (and possibly a stereochemical relationship) can often be used to key the application of a simplifying transform either directly or through a sequence of subgoals (*19, 20*). The same approach can be used to select tactical combinations of transforms that are frequently used in tandem in synthesis. The systematic interconversion, addition, or removal of functionality can pave the way for further retrosynthetic molecular simplification. Functional group modification may also be used to reduce the reactivity of the target or an intermediate toward one or more reagents in a synthetic sequence. One strategy to accomplish this is the use of internal protection; another is the replacement of a highly reactive unit by a less reactive synthetic equivalent (*21*).

The use of two or more strategies concurrently constitutes a powerful approach to the development of especially simple and elegant synthetic routes. The topological and stereochemical strategies lend themselves particularly well to concurrent application with other strategies (*15*).

Objectives of Computer-Assisted Synthetic Analysis

During the last 15 years, the concepts of retrosynthetic analysis have been used at Harvard as guidelines for the construction of a series of computer programs designed to help chemists plan synthetic routes. All of these programs have incorporated strategies that emulate the most effective problem-solving approaches of synthetic chemists, and the emphasis has been on assisting chemists with complex, rather than routine, synthetic problems. Since the number of possible approaches to the synthesis of a complicated molecule is extremely large, computer assistance would permit a more exhaustive analysis of any problem than an individual chemist would be likely to attempt. Further, because of the enormous body of chemical information needed in synthetic analysis, a large computerized database seemed highly advantageous. We have demonstrated that a machine can produce nonobvious, provocative, and useful solutions to synthetic problems. However, the task of developing a powerful problem-solving program that would be indispensable to the practicing chemist has proved to be both large and challenging.

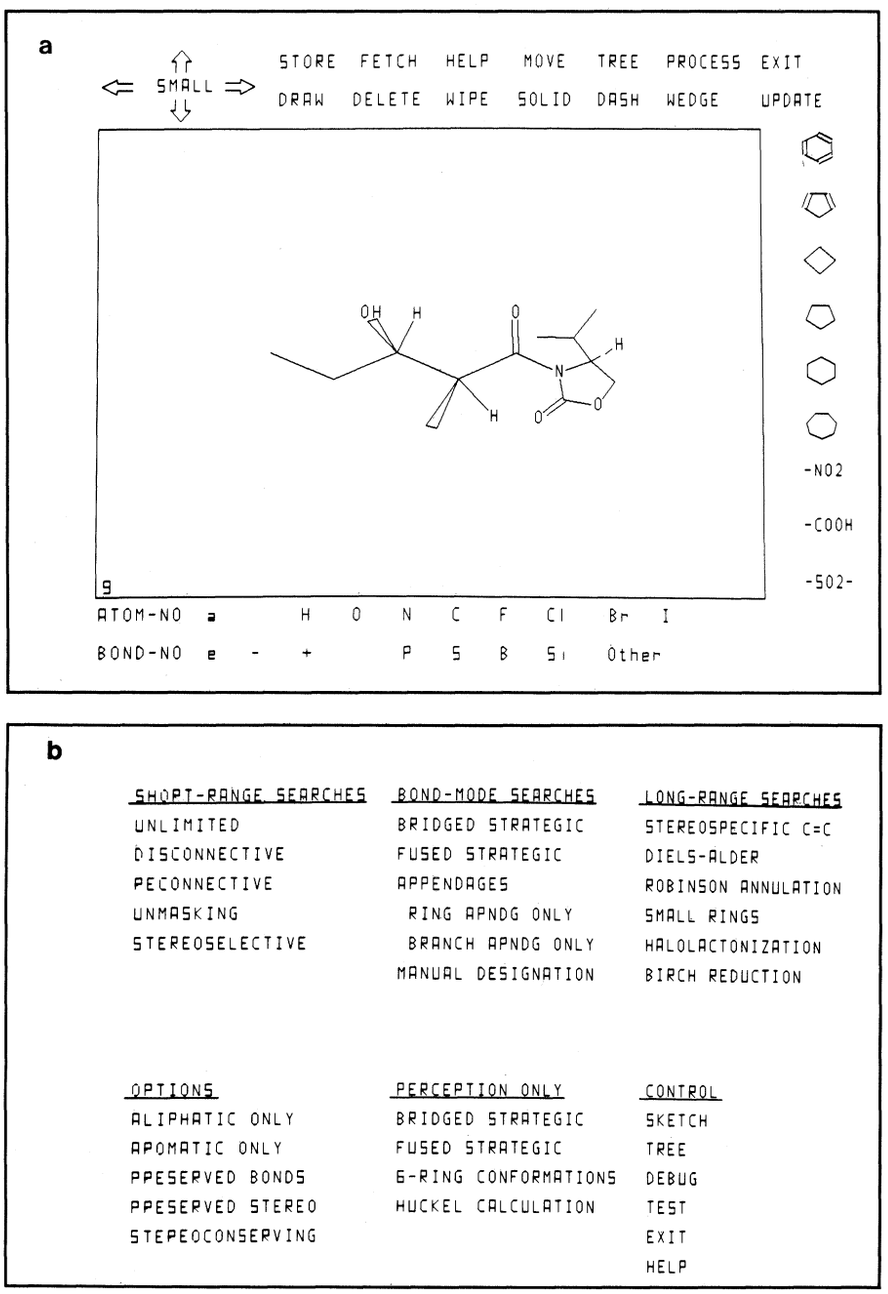


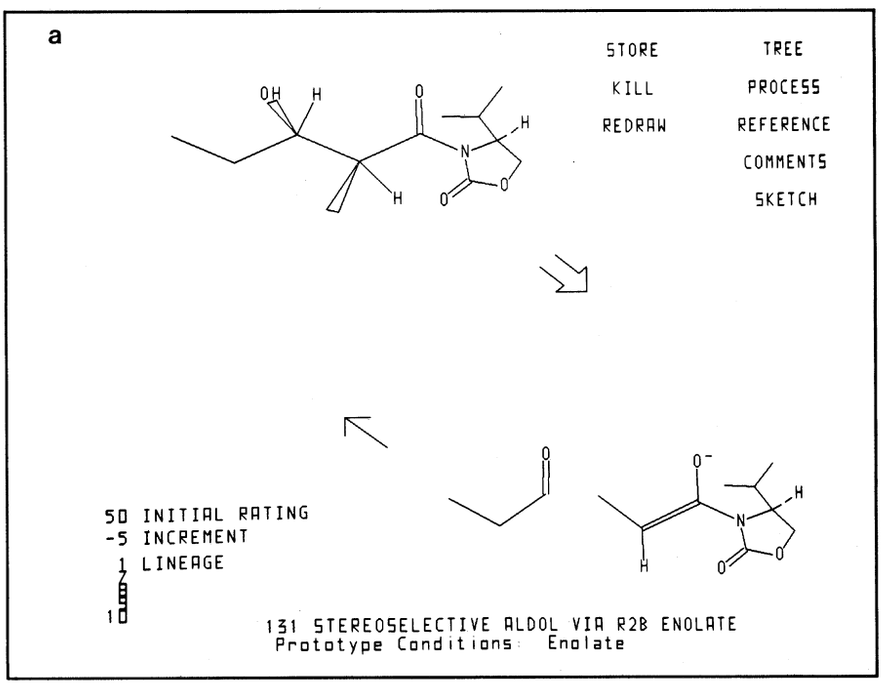
Fig. 1. (a) LHASA sketch pad display. The words and symbols around the edge of the drawing area are "buttons" that may be selected graphically by the user to switch the program from one input mode to another. (b) LHASA processing menu. The various tactics available to the chemist appear as "buttons" on this graphical display.

LHASA: Graphics and Flow of Control

The interactive program for computer-assisted synthetic analysis currently under development at Harvard is LHASA-11, a program written mainly in Fortran and implemented on Digital Equipment Corporation's VAX-11 computers. Like its predecessor programs OCSS (*5*), LHASA-1 (*6, 19, 22–24*), and LHASA-10 (*25*), LHASA-11 (*26*) (or just LHASA) is designed to communicate with a chemist entirely through the use of interactive computer graphics. The chemist "draws in" a synthetic target with a magnetic tablet and stylus, a mouse, a light pen, or some other graphical input device. As the chemist draws, LHASA displays a picture of the target structure on a CRT (cathode-ray tube) screen (Fig. 1a). The chemist then selects a synthetic strategy and one or more substrategies, or tactics, from a menu of processing options (Fig. 1b). LHASA suggests synthetic routes to the target that satisfy the goals of the selected strategy, displaying the resulting precursor structures on the CRT screen (Fig. 2a). When all the precursors have been generated and displayed, the analysis is summarized in a retrosynthetic tree (Fig. 2b), from which the chemist can choose any structure for further simplification.

In this interactive retrosynthetic analysis, the chemist takes responsibility for choosing the strategies and tactics and for deciding which precursors should be submitted to LHASA for further simplification. LHASA is responsible for selecting the actual transforms to be used, within the scope of the tactic chosen by the chemist, for evaluating these transforms in the context of the target structure, and for displaying the precursors that result from these transforms.

The process of deciding which transforms to display is central to the operation of the LHASA program. LHASA is not designed to invent chemistry that has never been performed in the laboratory. Instead, it relies on a database of known reactions (about 1100 at present). For LHASA to be able to decide which transforms are appropriate for a given synthetic target, the database must contain detailed chemical information about the corresponding reactions. In addition, the program must have a way of reading this chemical information and of evaluating it in the framework of what it knows about the target structure.



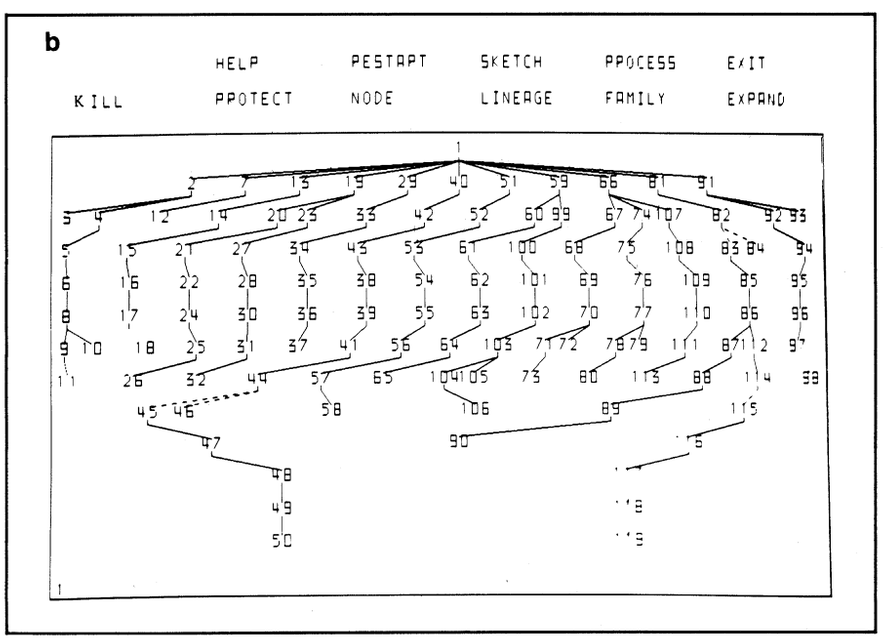


Fig. 2. (a) LHASA precursor display. LHASA suggests reactions to the chemist, showing the target at the upper left and the precursor at lower right. (b) LHASA retrosynthetic-tree display. The target structure drawn in by the chemist appears as node 1, with the various retrosynthetic precursors as higher numbered nodes.

Perception

As the chemist draws in the target structure, LHASA stores information about the position, type, and connectivity of each atom and bond in the molecule. Before any strategies are executed, the program uses this basic information about atoms and bonds to assemble a body of chemical knowledge about the target. This process is similar to a chemist's perception of a structure and is likewise called "perception." The program first looks at the various atom and bond types and checks the valency for each atom. Next, LHASA perceives functional groups (*23, 27*) and synthetically significant rings (*23, 24, 28*). The program then proceeds with higher level perception, looking for unstable and highly reactive arrangements of functionality, mutually incompatible functional groups, aromatic rings, and unstable tautomers. Fused and bridged rings are identified, and situations involving a very high degree of angle strain are

pointed out to the chemist. Carbon-carbon double bonds are classified according to their substitution pattern and stereochemistry. Absolute stereochemistry at sp^3 centers is evaluated, and ring-fusion stereochemistry is assigned. Structural appendages attached to rings and appendages emanating from functional group origins and other branch points are identified (*29*), and sets of appendage atoms on the two faces of each ring and ring-fusion composite are filled. All of this information is stored by LHASA to be used for reference during the process of transform evaluation. In addition, the target structure is assigned a numerical "name" by means of a canonicalization method similar to the Morgan algorithm (*30*). Each precursor structure generated by the program is similarly named for comparison.

Transform Evaluation

LHASA uses this stored perception information to answer chemical questions about the target structure in order to decide which transforms in the database are appropriate. In fact, a transform entry in the database consists of a series of questions about a generalized target structure. These questions reflect the scope and limitations of the reaction corresponding to that transform.

The entire transform database is written in a chemical English computer language called CHMTRN (for chemistry translator), which is designed especially for LHASA. CHMTRN can be easily understood by a synthetic chemist but is also readily compiled (translated) into a form that can be read by the LHASA Fortran modules. Use of CHMTRN makes it easy for chemists to understand, modify, and expand the information about reactions and serves to separate the chemical information from the Fortran executive CODE.

The process of transform evaluation, in which LHASA decides which transforms to disregard and which ones to show to the chemist, consists of a line-by-line interpretation by the program of the CHMTRN questions in the database. A sample transform entry is shown in Fig. 3. Preliminary, or "header," information in the transform entry gives the number and name of the transform, a descriptive diagram of the retrosynthetic step, some literature references, the length of the atom path joining the two "keying" functional groups, a utility rating for the transform, the names of the keying groups, and some labels used for cross-referencing and prescreening. Next come the scope and limitations questions, or "qualifiers," which either eliminate the transform or modify the initial rating, and the "conditions" statements, which define prototypical reaction conditions so that LHASA can assess the possibility of functional-group interference during the transform. Finally, the transform has a "mechanism" section that contains the actual commands for converting the target to the precursor structure. Lines preceded by three dots (. . .) are comment lines, included for the benefit of the chemist and not read by LHASA.

Retrosynthetic Strategies

As mentioned above, LHASA has been designed to emulate the thinking of a synthetic chemist. Nowhere is this more evident than in the strategies that the program uses to narrow the scope of the search for transforms. Several of the important retrosynthetic strategies we described above have been implemented in LHASA. The first of these is the simple *functional-group strategy*, in which the program searches in the target structure for the functional group (or pair of functional groups) that would result from the performance of the reactions corresponding to the transforms in the database. The functional group (or pair of groups connected by a path of bonds) for each transform is the retron for that transform. Approximately 25 percent of the transforms in the database are cross-referenced in such a way as to make them accessible to the functional-group strategy.

Possible synthetic routes suggested by LHASA for the synthesis of porantherine (*15*) (**1**) are shown in Fig. 4. These sequences illustrate the considerable retrosynthetic simplification that can often be realized by the application of the functional-group strategy. In processing porantherine, LHASA found that the exact retron for the Mannich transform (a carbonyl and an amine whose carbon origins are separated by a path of two carbon-carbon bonds) did not exist in the target structure. Accordingly, when the Mannich transform was requested, the program performed a nonsimplifying, functional group interchange (FGI) step as a "subgoal" to the goal disconnection, generating structure **2**. As with the Mannich transform leading to structure

```
        TRANSFORM 117
        NAME MICHAEL ADDITION OF HETERO NUCLEOPHILE
...HET-C-C-W => HET-H + C=C-W
        ...MARCH 585; HOUSE 596; B+P 468
        ...ORG. RXNS. VOL.5, 79-135 (1949)
        ...BULL. SOC. CHEM. FR. 254,325 (1962)
        ...PATH 2 BONDS
        RATING 50        ...Old rating 40
        GROUP*1 MUST BE KETONE OR CYANO OR ESTER OR ACID
           OR LACTONE OR AMIDE*3 OR AMIDE*2 OR AMIDE*1
           OR LACTAM OR VINYLW OR ALDEHYDE
        GROUP*2 MUST BE ETHER OR AMINE*1 OR AMINE*2 OR AMINE*3
           OR SULFIDE OR THIOL
        STUDENT
        REMOVES*STEREO CARBON2*1 ATOM*2
        BROKEN*BONDS BOND2*1
        ...
        KILL IF NO HYDROGEN ON ATOM*2
                ...REQUIRED FOR REACTION
        KILL IF MULTIPLE BOND ON ATOM*2 OFFPATH OR: ON ATOM*3 OFFPATH
                ...WOULD PRODUCE ALLENIC PRECURSOR
        IF BOND2*1 IS NOT IN A RING OF SIZE 5 THROUGH 7  &
           THEN KILL IF BOND2*1 IS IN A RING
        SUBTRACT 15 IF LEAVING GROUP ON ATOM*3
                ...POSSIBLE ELIMINATION
        ADD 15 IF ANOTHER WITHDRAWING BOND ON ATOM*2
                ...EASIER ADDITION
        SUBTRACT 15 FOR EACH WITHDRAWING BOND ON ATOM*3
                ...UNDESIRED MICHAEL POSSIBLE
        SUBTRACT 10 IF ATOM*3 IS A TERTIARY*CENTER
        IF NOT OLEFIN ON BOND*2 THEN KILL IF ATOM*2 IS NOT ENOLIZABLE
                ...STABLE ENOL PROVIDES DRIVING FORCE
        IF SECOND GROUP IS ETHER THEN CONDITIONS NaOR
        IF SECOND GROUP IS AMINE THEN CONDITIONS RNH2
        IF SECOND GROUP IS SULFIDE OR: THIOL THEN  &
           CONDITIONS NaSR
        ....
           BREAK BOND2*1
           JOIN ATOM*2 AND ATOM*3
        ....
```

Fig. 3. A sample transform entry from the LHASA database. The "header" information appears above the isolated set of three dots (. . .), and the transform "mechanism" is enclosed between the two sets of four dots (. . . .) at the end.

3, LHASA evaluated this FGI transform by reading the CHMTRN qualifiers associated with its entry in the database. A second invocation of the functional-group strategy, with structure **3** as the target, produced another Mannich disconnection, this time without the need for preliminary subgoals. The aldehyde in precursor structure **4** is displayed enclosed in a solid box to alert the chemist to a problem with interfering functionality. LHASA not only can identify these situations (*21*) but also can suggest protective groups for the interfering functional groups (*31*). Invocation of the same strategy on structure **4** led, by way of another FGI, to a number of interesting precursors, among them structures **6** to **9**.

Often, the retron that a chemist recognizes when performing a retrosynthetic analysis is too complex to be described by simple functional groups and the linkages between them. Since most of the retrons for a wide range of heterocyclic and aromatic chemistry fall into this category, modules have been added to LHASA (*32*) that key transforms on the basis of arbitrary patterns of atoms and bonds. This keying method has allowed the addition of many multicenter transforms, which now constitute approximately 50 percent of the database. These multicenter transforms include, for example, those which disconnect heteroaromatic rings (pyrrole, triazole, indole, and so on) into precursor fragments.

Several examples of the *transform-based strategy* have also been implemented in LHASA. The CHMTRN search tables corresponding to reactions like the Carbo Diels-Alder (*7*), Robinson annulation (*33*), Birch reduction, Quinone Diels-Alder, and Halolactonization (*34*) have been endowed with considerable power to request subgoal transforms to rectify differences between the target structure and the goal retrons. Retrosynthetic sequences of up to 25 steps can be generated by these aptly named "long-range" searches, which now constitute the remaining 25 percent of the CHMTRN database. Although several hundred such powerful transforms will eventually be included in the database, this inclusion will follow current work on the simplification of long-range search procedures.

Three of the 15 sequences suggested by the Robinson annulation search in LHASA (*33*) for the synthesis of valeranone (**10**) (*35, 36*) are shown in Fig. 5. The first route suggested is an excellent

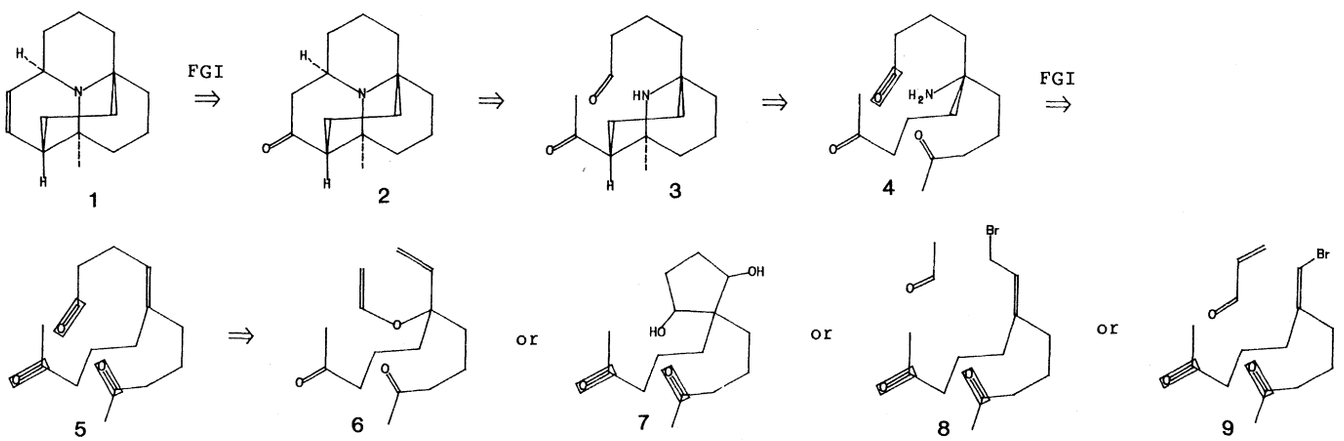


Fig. 4. Retrosynthetic pathways generated for porantherine (*15*) by the functional group–based search in LHASA. Boxed bonds indicate interfering functionality (*21, 29*). Functional group–interchange subgoal steps are marked FGI.

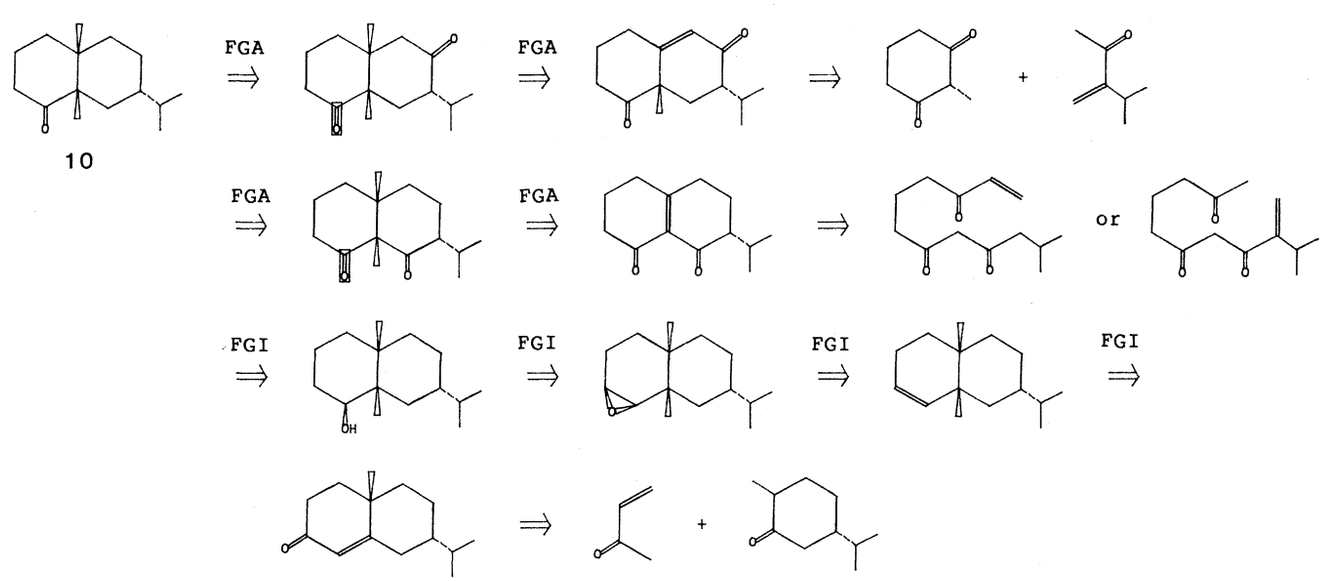


Fig. 5. Three sample sequences suggested by the Robinson annulation module in LHASA for the synthesis of valeranone (*33, 34*). Boxed bonds indicate interfering functionality (*21, 29*). Functional group–addition and functional group–interchange subgoal steps are marked FGA and FGI, respectively.

illustration of the power of the Robinson annulation search. By considering both six-membered rings in the target and all the possible orientations of each ring, LHASA has found a short, stereocontrolled route from readily available starting materials. A single, easily solved problem of functional group interference has also been pointed out. The second route suggested by the program shows that intramolecular possibilities should not be overlooked. The third sequence provides an insight into some of the stereochemical capabilities in LHASA. In the first FGI step (interchange of ketone for hydroxyl), the program automatically generates both epimeric alcohols. For the second step, however, the conformational-analysis module (*37*) predicts that only the β-epimer could result from diaxial opening of the desired epoxide. Hence, the α-alcohol is discarded and only the β-epoxide is displayed to the chemist.

Topological strategies are well represented in LHASA as well (*13*). LHASA has special perception modules that can recognize strategic bonds in bridged and fused cyclic systems, and the program can be asked to concentrate only on transforms that will disconnect one or more of these bonds. A chemist can also manually designate one or more bonds as strategic or, conversely, preserve a bond (or bonds) from disconnection. A recently implemented variation on the strategic-bond search in LHASA uses an algorithm that identifies pairs of ring bonds that can be disconnected by intramolecular cycloaddition transforms (*38*). Some examples of the bonds that LHASA finds strategic in a few sample targets are shown in Fig. 6.

Recent developments in the synthetic methodology for acyclic diastereoselection (*17*) have prompted the addition of a considerable amount of new chemistry to the LHASA database. Transforms for stereoselective aldol condensation (*39*), asymmetric epoxidation (*40*), various chelation-controlled carbonyl additions (*41*), and a number of other reactions are the foundation for the various *stereochemical strategies* in LHASA (*42*). Selection of transforms based on the relationships between stereocenters often leads to particularly effective synthetic plans. The stereoselective strategy in LHASA restricts the search for transforms to those designated as stereospecific or highly stereoselective. Conversely, when stereosimplification is not desired, the "stereoconserving" mode may be used to prevent the retrosynthetic removal of stereochemical information from a target. By analogy with strategic

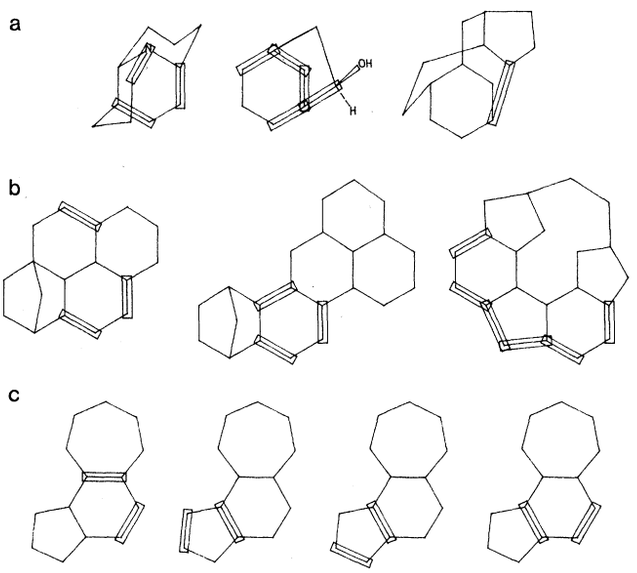


Fig. 6. (a) Bridged strategic bonds (*13*) identified by LHASA for some sample ring systems. Retrosynthetic disconnection of any of the boxed bonds will reduce the complex polycyclic bridged structures to simple fused or monocyclic systems. (b) Fusion strategic bonds identified by LHASA for some sample ring systems. Retrosynthetic disconnection of any of the boxed bonds will result in strategic simplifications of the polycyclic fused structures shown. (c) Strategic bond pairs for intramolecular cycloadditions in a sample ring system.

bonds, it is also possible for the user to designate stereocenters in the target for preservation or removal. Use of this last option will become less important as heuristic procedures are developed for identification of stereochemically critical target substructures. Results from the new stereochemical strategy are promising, particularly for acyclic target structures. A sample sequence generated by LHASA for the synthesis of erythronolide B seco acid (**11**) is shown in Fig. 7.

New Methodology for Transform Selection

In LHASA's initial stages, the program benefited from the rigid separation of transforms into categories based on their functions and the methods used to key them. This separation allowed straightforward implementation of program modules for each category of transform. However, one limitation of this scheme quickly became apparent. Sometimes, for example, a goal transform required a subgoal that was not a simple functional group interchange, addition, or removal. A temporary solution for this limitation was used in the long-range, transform-based search tables. The three types of functional group–oriented subgoals were used where possible, but other transforms, called "composite subgoals," were included as necessary. These composite subgoals frequently duplicated chemistry already present as goal transforms in other parts of the database.

A new method for describing transforms is being implemented to avoid this duplication and to increase the subgoal power of all the goal transforms in LHASA. For a transform to be used as a subgoal as well as a goal, the transform entry in the database must specify both the retron in the target and the resultant substructure in the precursor. In such a system, hypothetical sequences of subgoal steps can be grown before any transforms are actually tried by the program. Accordingly, the LHASA database is being modified to include pictorial representations of each transform (compare Fig. 8 with Fig. 3). These target and precursor patterns are perceived by a parsing algorithm that extracts keying information (for example, functional groups, broken bonds) as well as subgoal information (changes in bond orders, attachment of new atoms, and so on) and stores it for later use. The parsing algorithm is relatively slow, but can be run off-line (outside of LHASA) and is only executed once each time a transform is written or modified.

New Methods for the Generation of Subgoal Sequences

In practice, there is rarely an exact match between the target structure and the retron for a transform. Occasionally, a single subgoal will rectify the mismatch and allow the transform to be performed as a goal. More often, however, several subgoals are required at different sites in the target to correct the mismatch. With the diversity of chemistry available, many subgoal sequences can be devised to convert the target to the retron for the goal transform. A general method for locating the best subgoal sequences would increase the power of any high-level strategy.

The new methodology for generating

subgoal sequences with LHASA draws heavily upon techniques more familiar to the computer scientist than the chemist. The overall process is based on the principles of "heuristic search" (*43*), its objective being to select the best subgoal sequence from many possibilities without generating all of the sequences. First, all applicable mappings between the target structure and the retron for the desired goal transform are obtained. An "estimation function" is then applied to each hypothetical mapping to assess the difficulty of converting the target substructure to the goal retron. This estimation function yields a number that approximates the "chemical distance" (*44, 45*) between target and goal retron for each mapping (the chemical distance is a function of the number and the individual utility ratings of the interconnecting transforms). Next, transforms for the first synthetic subgoal step are chosen for the best of these mappings; this choice is aided by a database format in which both target and precursor patterns are specified for each transform. Each of the resulting partial sequences is assigned a merit value, which is based on the initial estimation function for the mapping modified by the utility of the particular transform chosen and by the degree to which the transform succeeds in fulfilling the subgoal request. The best partial sequence is then elaborated one step further in the synthetic direction. This evaluation and elaboration process is repeated until a complete sequence is hypothesized.

The program then performs the first retrosynthetic step from the target. At this stage, the program further evaluates the merit of the sequence by taking into account the applicability of the subgoal transform to the target, as evaluated by the CHMTRN scope and limitations qualifiers, and the degree of functional group interference encountered. If the sequence is still found to be better than any other partial sequence, the next retrosynthetic step is performed. The sequence is evaluated in this manner until either a step fails or the sequence is completed. If the estimate of the merit of the partial sequences is reasonably accurate, the first sequence completed will be the best one possible.

Applications of the New Subgoal Methodology

Currently, a LHASA executive routine chooses a set of transforms based on the requirements of the selected strategy and on the restrictions imposed by the chemist (Fig. 1b). For each of these transforms, the routine attempts to match the retron for the transform with a corresponding substructure in the target. Mismatches in functionality can be rectified by invoking one or more subgoal

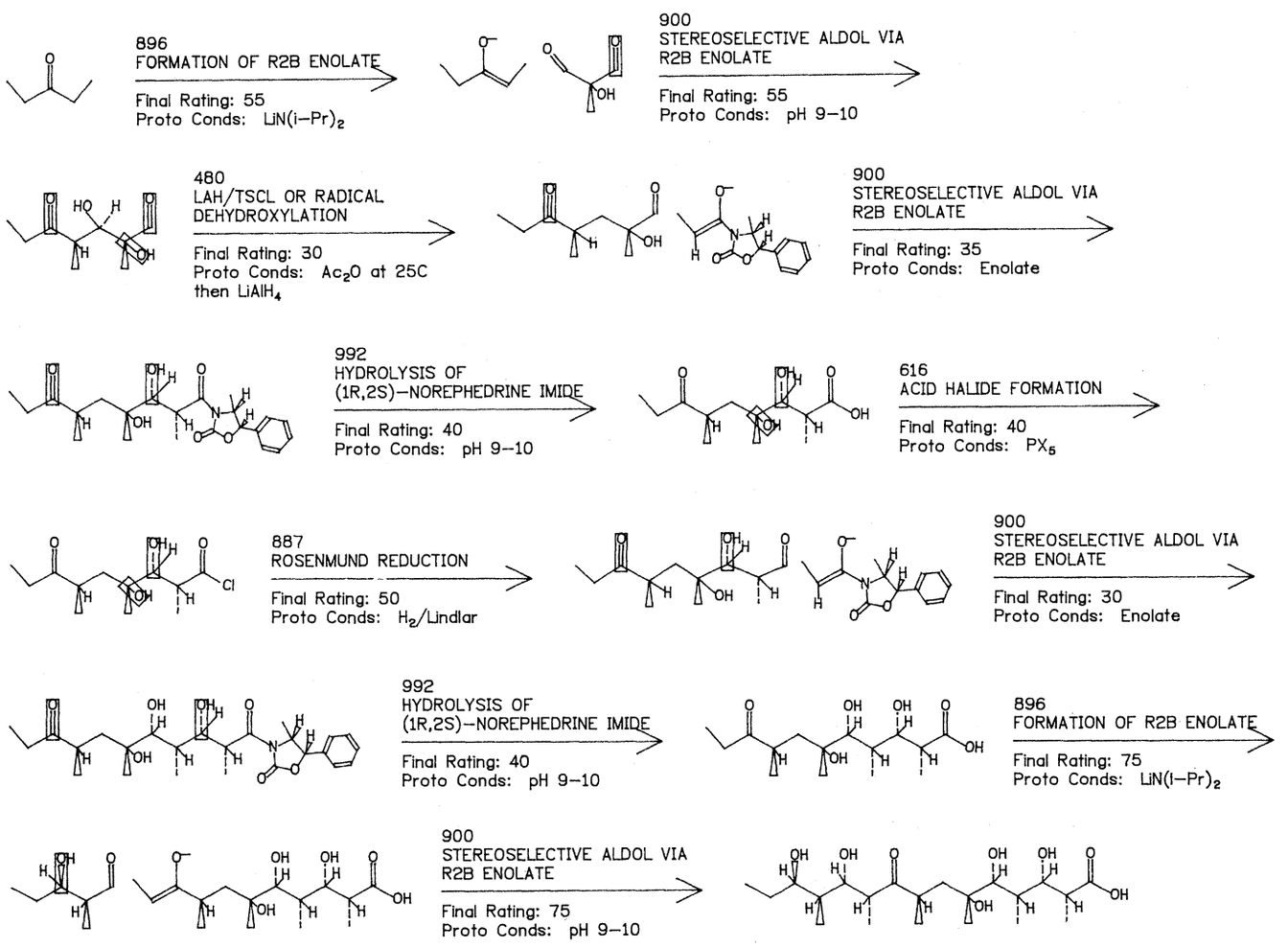


Fig. 7. Sample route for the synthesis of erythronolide B seco acid generated by the stereochemical strategy in LHASA. Boxed bonds indicate interfering functionality (*21, 29*). The sequence is presented in the synthetic direction, exactly as output by the off-line TREEPLOT program on a Hewlett-Packard 7470A Pen Plotter.

executive routines [FGI, FGA, FGR, sequential FGI (*20*), and so forth], but more complicated mismatches can only be handled by the few transforms for which long-range search tables have been written. The implementation of a new strategy requires considerable effort, either in Fortran or CHMTRN or both.

With the new methodology, the full range of functional group–keyed LHASA transforms can be used to rectify a mismatch between target and retron. The goal executive routine invokes a subgoal executive routine for which the inputs are the retron for the goal transform and a mapping between that substructure and the target molecule. Every strategy in LHASA can make use of this general subgoal algorithm. All that is required is that the strategy be able to identify an S-goal, to generate that S-goal from the target structure by calling the subgoal executive routine, and, if appropriate, to perform the transforms to convert the S-goal to the goal precursor.

One application of the new subgoal methodology is in transform-based strategies. In these strategies, a search can be constructed around a small set of S-goals, each of which can be readily converted retrosynthetically to the goal precursor. For example, ten such S-goals have been identified for halolactonization (*33*). Currently, the subgoal sequences to convert the target to such S-goals are rated by a process known as "prior procedure evaluation," or PPE (*33, 34*). This process is quite complex and must be reimplemented for each key reaction. With the new methodology, the algorithm to rank the subgoals will obviate the need for PPE and make it much easier to design long-range searches.

Another strategy in which the new subgoal methodology will see immediate application is the "starting material–oriented" search (*46*). In this newest of LHASA strategies, a starting material may be drawn in by the chemist or selected by the program from a pool of commercially available compounds. The objective is simply generation of the S-goal—no T-goal exists. The first step in such a search is the generation of an atom-by-atom mapping between target and starting material (*45*). When a mapping has been established, the problem of converting target to S-goal becomes similar to that encountered in a long-range search; the only difference is that the number of mismatches is usually greater owing to the necessity of matching an entire structure rather than a smaller substructure. Considerable progress has been made in this area, especially for aromatic compounds (*47*), and the methodology is currently being expanded (*48*).

```
        TRANSFORM 117
...NAME MICHAEL ADDITION OF HETERO NUCLEOPHILE
...MARCH 585; HOUSE 596; B+P 468
...ORG. RXNS. VOL.5, 79-135 (1949)
...BULL. SOC. CHEM. FR. 254,325 (1962)
...RATING 40
...
... W OR,NH2,NHR,NR2,S      W
... | |                      |
... C-C                  => C=C
... |
... H
...
...RATING 40
...
... W NH2,NHR,NR2      W
... | |                |
... C=C             => C#C
... |
... H
...
```

Fig. 8. A sample transform header from the new, unified LHASA database. All keying and subgoal information is extracted from the descriptive picture of the transform.

Problems for the Future

The new techniques for the generation of subgoals have opened up a number of promising areas for further work. One of these areas is the automatic evaluation of the various tactics possible within a selected strategy. The same heuristic criteria that are being developed for guiding the search for subgoal sequences may also be applied to the initial subgoal requests for different tactics.

For example, several ring-forming transforms might be considered for constructing a cyclic target structure. In turn, each transform might be considered in several orientations with respect to the target ring system. A rating could be assigned to each orientation of each transform on the basis of the inherent usefulness and reliability of the transform and on the simplification to be expected from its application. The final rating for a sequence leading to the goal transform would also depend on the subgoal transforms chosen for the sequence. Inclusion of the estimated ratings for the projected subgoal steps in the final rating would allow the best sequences, and thus the best tactics, to be found rapidly.

In fact, subgoal searches for several S-goals may be carried out in parallel. Since the rating of a partial sequence becomes more accurate as the program fills in the details of the sequence, by considering several requests in parallel, the best combination of an S-goal and a sequence may be obtained. For example, several starting materials may be chosen, either by the chemist or by a database screening program, and subgoal requests may be made for each of them. The subgoal search would begin by choosing subgoal steps for the best of these. However, if a good sequence were not found quickly, the program would proceed to evaluate a starting material that was initially rated as less promising. In this way, the program would find the best sequences for any of a group of starting materials. Currently, the algorithm is being implemented on a single processor; however, it is designed to be implemented on a network of parallel processors as well.

A number of long-term goals have developed in this research. One is to increase the involvement of the LHASA program in the strategy selection process. At present, strategy selection is left up to the chemist. Machine selection of strategies would involve identification of substructures that suggest particular synthetic approaches, ordering the chosen strategies according to various heuristic criteria, presentation of these strategic choices to the chemist for verification, and, finally, execution of the best strategy. New perceptual capabilities would have to be added to the program for it to be able to perform these tasks, but the heuristic method developed for choosing and ordering strategies would also find applications in other areas of organic synthesis.

A logical extension of automated strategy selection would be the concurrent use of two or more strategies. This approach could lead to particularly elegant plans for syntheses, depending of course on the range of strategies available and on the heuristic method developed to guide the selection process. Again, the heuristic method for this type of search would be useful outside the field of computer-assisted synthetic analysis.

The detailed stereochemical and mechanistic evaluation of transforms provides another area for further work. LHASA is already capable of sophisticated stereochemical analyses (*7, 8, 33, 36*), but the ability to make accurate predictions about molecular shape and reactivity in an interactive environment could stimulate development of new techniques (*37*). For example, the stereochemical course of certain reactions can be described by a "steric-approach-control" model. The stereochemistry of each transform could be calculated from such a model (*48, 49*), but in an interactive environment it would be more efficiently arrived at by "table look-up" procedures (*50*). To this end, LHASA transforms are to be categorized by mechanism to allow for greater standard-

ization of the database and to assist transform writers in deciding what structural features would help or hinder each reaction. The coordinated use of LHASA with programs that can perform detailed mechanistic evaluation of reactions in the synthetic direction [for example, CAMEO (*51*)] may prove useful in this effort.

Until recently, the database for LHASA has been kept quite small but, as the usefulness of the program has increased, so has the demand for a broader selection of transforms. The task of expanding the database is monumental, not only because of the number of organic reactions in the literature but also because the LHASA database contains more than just simple descriptions of reactants, products, and reaction conditions. For LHASA to assess correctly the applicability of its transforms to an arbitrary target structure, each transform entry must include information about yields, negative results, reaction mechanisms, alternative reaction pathways, adverse structural features, rate-enhancing properties, and so on. As shown in Fig. 3, this information must be input not in tabular form but rather in the form of scope and limitations queries about the target structure. Since designing and debugging these queries requires considerable time and effort, any method for streamlining the transform-writing process would represent a significant advance. In particular, a separate graphically driven program to assist the chemist in writing transforms is in the planning stages. This program would prompt the user for all the important features of a transform entry and would also suggest questions that might be appropriate for LHASA to ask about a target structure on the basis of the reaction mechanisms designated by the chemist. Like LHASA, this program would be interactive, processing the chemist's input and countering with suggestions of its own.

LHASA is only one of many computer tools that are starting to be used routinely by organic chemists. Another type of program that is rapidly gaining popularity is the reaction retrieval database system (for example, REACCS, SYNLIB, or ORAC) (*52*). Although one cannot write a LHASA transform with only the information in one of these databases, this type of system could prove useful for identifying new reactions to add to LHASA and for checking LHASA transforms for generality.

The task of expanding the LHASA database is of such magnitude that extensive collaboration among groups of chemists worldwide will be necessary. In fact, two consortia of chemical and pharmaceutical companies are already involved, LHASA UK in the United Kingdom and CASAG (the Computer-Assisted Synthetic Analysis Group) in the United States. Since this article has dealt primarily with developments in computer-assisted synthetic analysis at Harvard and the University of Leeds, it is important to mention the programs written by some of the other groups in the field. The SECS (*53*), CASP (*54*), and SYNCHEM II (*55*) programs are similar to LHASA in that they rely on databases of known reactions, whereas the EROS (*56*) and SYNGEN (*44*) programs base their synthetic suggestions on more simple bond-making and bond-breaking processes. As mentioned above, the CAMEO (*51*) program works in the synthetic direction and will be a useful adjunct to retrosynthetic analysis programs.

References and Notes

1. I. Fleming, *Selected Organic Syntheses* (Wiley, New York, 1973).
2. N. Anand, J. S. Bindra, S. Ranganathan, *Art in Organic Synthesis* (Holden-Day, San Francisco, 1970).
3. E. J. Corey *et al.*, *J. Am. Chem. Soc.* **86**, 478 (1964).
4. E. J. Corey, *Pure Appl. Chem.* **14**, 19 (1967).
5. _____ and W. T. Wipke, *Science* **166**, 178 (1969).
6. E. J. Corey, *Quart. Rev. Chem. Soc.* **25**, 455 (1971).
7. E. J. Corey, W. J. Howe, D. A. Pensak, *J. Am. Chem. Soc.* **96**, 7724 (1974).
8. E. J. Corey and A. K. Long, *J. Org. Chem.* **43**, 2208 (1978).
9. E. J. Corey and R. D. Balanson, *Tetrahedron Lett.* (1973), p. 3153.
10. F. Serratosa, *Heurisko: Introduction a la sintesis organica* (Alhambra, Madrid, 1975).
11. S. Hanessian, *Total Synthesis of Natural Products: The Chiron Approach* (Pergamon, Oxford, 1983).
12. E. J. Corey, M. Shibasaki, J. Knolle, *Tetrahedron Lett.* (1977), p. 1625.
13. E. J. Corey *et al.*, *J. Am. Chem. Soc.* **97**, 6116 (1975).
14. See, for example, *Nachr. Chem. Tech.* **31**, 462 (1983).
15. E. J. Corey and R. D. Balanson, *J. Am. Chem. Soc.* **96**, 6516 (1976).
16. A. G. Schultz *et al.*, *ibid.* **101**, 1276 (1979).
17. See, for example, J. D. Morrison, Ed., *Asymmetric Synthesis* (Academic Press, Orlando, Fla., 1982–1984), vol. 1–3.
18. E. J. Corey, J. F. Arnett, G. N. Widiger, *J. Am. Chem. Soc.* **97**, 430 (1975).
19. E. J. Corey, R. D. Cramer III, W. J. Howe, *ibid.* **94**, 440 (1972).
20. E. J. Corey and W. L. Jorgensen, *ibid.* **98**, 203 (1976).
21. E. J. Corey, H. W. Orf, D. A. Pensak, *ibid.*, p. 210.
22. E. J. Corey *et al.*, *ibid.* **94**, 421 (1972).
23. E. J. Corey *et al.*, *ibid.*, p. 431.
24. E. J. Corey and G. A. Petersson, *ibid.*, p. 460.
25. D. A. Pensak and E. J. Corey, *ACS Symp. Ser.* **61**, 1 (1977).
26. A. K. Long, S. D. Rubenstein, L. J. Joncas, *Chem. Eng. News* **61** (19), 22 (1983).
27. See (*25*), p. 11. The previously described functional-group perception algorithm was recently replaced (A. P. Johnson and P. L. M. Hoyle, unpublished results) by a pattern-based method that will allow recognition of many more group types.
28. See (*25*), p. 8.
29. E. J. Corey and W. L. Jorgensen, *J. Am. Chem. Soc.* **98**, 189 (1976).
30. E. J. Corey and S. D. Rubenstein, unpublished results.
31. E. J. Corey *et al.*, *J. Org. Chem.*, in press.
32. Pattern chemistry in LHASA was originally implemented by J. S. Dixon of Lederle Laboratories. Major developmental work in this area has been carried out at the University of Leeds, England, by C. Marshall, G. Hopkinson, and P. Hoyle, under the direction of A. P. Johnson. Contributions to the database for pattern chemistry have come mostly from H. Tumber, G. Hopkinson, and J. Mann at the University of Leeds and from the LHASA United Kingdom consortium of British chemical companies.
33. E. J. Corey, A. P. Johnson, A. K. Long, *J. Org. Chem.* **45**, 2051 (1980).
34. E. J. Corey *et al.*, *J. Chem. Inf. Comput. Sci.* **20**, 221 (1980).
35. J. A. Marshall, G. L. Bundy, W. I. Fanta, *J. Org. Chem.* **33**, 3913 (1966).
36. E. Wenkert and D. A. Berges, *J. Am. Chem. Soc.* **89**, 2507 (1967).
37. E. J. Corey and N. F. Feiner, *J. Org. Chem.* **45**, 757, 765 (1980).
38. E. J. Corey and G. E. Walker, unpublished results.
39. D. A. Evans, in *Proceedings of the Robert A. Welch Foundation Conferences on Chemical Research* (R. A. Welch Foundation, Houston, Texas, 1983), pp. 13–55.
40. T. Katsuki and K. B. Sharpless, *J. Am. Chem. Soc.* **102**, 5974 (1980).
41. R. Bernardi, C. Fuganti, P. Grasselli, *Tetrahedron Lett.* (1981), p. 4021; W. C. Still and J. H. McDonald III, *ibid.* (1980), p. 1031.
42. E. J. Corey, G. I. Lotto, G. Sello, unpublished results.
43. J. Pearl, *Heuristics: Intelligent Search Strategies for Computer Problem Solving* (Addison-Wesley, Reading, Mass., 1984).
44. J. B. Hendrickson and E. Braun-Keller, *J. Comput. Chem.* **1**, 323 (1980).
45. C. Marshall, "Computer-assisted design of organic syntheses," thesis, University of Leeds, (1984).
46. W. T. Wipke and D. Rogers, *J. Chem. Inf. Comput. Sci.* **24**, 71 (1984).
47. A. P. Johnson, *Chem. Br.* **21**, 59 (1985).
48. E. J. Corey and C. Marshall, unpublished results.
49. W. T. Wipke and P. Gund, *J. Am. Chem. Soc.* **96**, 299 (1974).
50. A steric-approach control module for LHASA was devised several years ago, but awaits implementation.
51. C. E. Peishoff and W. L. Jorgensen, *J. Org. Chem.* **48**, 1970 (1983).
52. These software packages are available from Molecular Design, Ltd., Beckman Instruments, and the University of Leeds, respectively.
53. W. T. Wipke, G. I. Ouchi, S. Krishnan, *Artif. Intell.* **11**, 173 (1978).
54. CASP is an offshoot of SECS and is the property of a consortium of Swiss and West German pharmaceutical companies.
55. K. K. Agarwal, D. L. Larsen, H. L. Gelernter, *Comput. Chem.* **2**, 75 (1978).
56. J. Gasteiger and C. Jochum, *Top. Curr. Chem.* **74**, 93 (1978).

57. We thank the past members of the LHASA research groups at Harvard and our co-workers G. Lotto, C. Marshall, and G. Walker. We also thank our collaborative research group at the University of Leeds, England, under the direction of A. P. Johnson. This work is supported by the National Institutes of Health and by our industrial collaborators.