periodically admitted to the system. After ethylene uptake had ceased, the solution was filtered and the polymer was washed with CH_2Cl_2 and then dried under vacuum. See Table V for the results of the GPC analysis. Ethylene-uptake rate measurements were carried out in the same apparatus. CH_2Cl_2 (49 mL) was placed in a 100-mL Schlenck flask in the glovebox. The flask was attached to the polymerization apparatus. The system was evacuated and then back-filled with ethylene and allowed to equilibrate to the desired initial pressure to insure saturation of the solvent with ethylene. The solvent was stirred vigorously with a magnetic stirrer. For temperature control, the flask was immersed in a water bath (19 °C), an ice water bath (0 °C), or an ethanol bath cooled by a Neslab cryocool cc-60II immersion cooler (10, -7, and -14 °C). **14** was dissolved in 1 mL of CH_2Cl_2 and this solution injected into the reaction flask with a gas-tight syringe. The pressure was periodically recorded, and the moles of ethylene taken up by the solution were calculated and plotted as a function of time. The rates of polymerization were determined from the initial slopes of these curves.

Reaction of 16 with Ethylene. Uptake Measurements and Analysis of the Products. 16 97 mg, 0.275 mmol) was placed in a 25-mL round-bottom flask. The flask was evacuated, and 3 mL of C_6D_6 was vacuum transferred into the flask; 8.6 equiv of ethylene (2.363 mmol; 651 Torr in 67 mL at 23 °C) was condensed in. The solution was warmed to room temperature and stirred for 1 week. The flask was then attached to a Toepler pump, and the gas remaining in the flask was collected and analyzed by GC: 1.661 mmol of ethylene remained; 0.051 mmol of methane, 0.0068 mmol of ethane, 0.0050 mmol of propene, and 0.061 mmol of butene were formed. The solution was then treated with 0.98 equiv of gaseous HCl (0.270 mmol; 74 Torr in 67 mL at 23 °C). The gases were once again collected in the Toepler pump and analyzed by GC: 0.040 mol of ethylene remained; 0.135 mmol of methane and 0.022 of mmol ethane were produced. Thus, 0.135 mmol of **16** reacted with 0.666 mmol of ethylene. The volatiles remaining in the solution were vacuum transferred and analyzed by GC/MS and ^1H NMR. Two GC fractions were analyzed: The first was butene. Mass spectrum, *m/e* (relative intensity): 37 (4.0), 38 (7.2), 39 (46.2), 40 (12.7), 41 (100), 42 (3.6), 43 (8.5), 50 (5.4), 53 (6.7), 55 (21.2), 56 (39.5), 57 (1.8), 58 (1.8). The second was hexene. Mass spectrum, *m/e* (relative intensity): 38 (4.4), 39 (45.6), 40 (13.8), 41 (100), 42 (69.4), 43 (50.2), 53 (7.7), 54 (5.8), 55 (58), 56 (90.7), 57 (4.9), 69 (24.1), 70 (1.5), 84 (29.6), 85 (2.3). ^1H NMR (C_6D_6): 0.81 (m, 20 H), 1.23 (m, 20 H), 1.58 (m, 2 H), 1.97 (t, 5 H), 3.99 (s, 30 H), 4.95 (m, 2 H), 5.05 (m, 1 H), 5.42 (m, 1 H), 5.75 (m, 1 H) ppm.

Magnetic Measurements. The magnetic susceptibility of polycrystalline samples of **10** at room temperature and of **7** and **16** over the temperature range 3-300 K was measured with a Faraday balance. Variable temperature control was obtained with a Janis helium flow cryostat. The samples were loaded into high-purity quartz ampules in a drybox, which were evacuated and then sealed. The diamagnetic force due to the sample holder was subtracted over the entire temperature range. The susceptibility was measured at 10 field strengths to check for ferromagnetic impurities (Honda-Owens method). Calibration was performed with $HgCo(SCN)_4$ (16.44×10^{-6} emu/g at 298 K). The data were corrected for diamagnetism with use of Pascal constants.

Acknowledgment. We thank Drs. S. J. Carter and D. L. Beach of Chevron Chemical Co. for helpful discussions and polymer characterization data. This research was supported by the National Science Foundation (Grant CHE-8820354).

Supplementary Material Available: Tables of atomic positional and thermal parameters for **10**, **14**, **16**, and **17** (10 pages). Ordering information is given on any current masthead page.

Reaction Classification and Retrieval. A Linkage between Synthesis Generation and Reaction Databases

James B. Hendrickson* and Todd M. Miller

Contribution from Edison Chemical Laboratories, Brandeis University, Waltham, Massachusetts 02254-9110. *Received January 8, 1990*

Abstract: A program is described that translates widely used reaction databases such as REACCS and SYNLIB into the same terms as are used for synthesis design by the SYNGEN program. This therefore makes it possible to look up in these databases all reactions that are of the same family as any reaction generated by SYNGEN. Of the computer-accessible reactions in REACCS, about three-fourths of the constructions may be classified and retrieved by SYNGEN, constituting a precedent library of ~19000 reactions. A further 9000 are retrievable from SYNLIB.

For a number of reasons, we elected to develop our program, SYNGEN, for synthesis design from a basis of chemical logic and not from a database library of reactions contained in the program.[1] As a result, the output from SYNGEN is presented as "theoretically reasonable" construction reactions but without literature precedent. Now, however, several extensive collections of reactions are available as computer-searchable databases.[2] The extent of overlap among these databases is reported to be small,[3] and so we have developed a linkage program to locate in these databases the literature precedents for our generated reactions. This means that the synthetic routes that SYNGEN creates for any target were previously just "paper chemistry", i.e., not tied to specific literature examples. Hence, the laboratory chemist could not assess realistically their practical efficacy in his particular case. Now, with an interface to commercial databases of reactions, however, it is possible for the SYNGEN user to access the literature for concrete examples of closely matching molecules to locate the details required.

In order to create such an interface, we specifically focus first on the *reaction* rather than the *compounds* in order to find all parallel cases. Subsequently, we narrow the definition to get closer matching of the detailed substitution around the reaction center. We examine first the *net structural change* in the reaction rather than structures or substructures of reactants and products. To this end we use a layered, or taxonomic, set of reaction classifications, from the general to the particular.[4]

In effect, we classify successive layers of definition for sorting and retrieving reactions, logically analogous to the system of Beilstein for classifying molecular structures. This classification

(1) (a) Hendrickson, J. B.; Toczko, A. G. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 137. (b) Hendrickson, J. B. *Acc. Chem. Res.* **1986**, *19*, 274. (c) Hendrickson, J. B.; Toczko, A. G. *Pure Appl. Chem.* **1988**, *60*, 1563; **1989**, *61*, 589. (d) Hendrickson, J. B.; Bernstein, Z.; Miller, T. M.; Parks, C.; Toczko, A. C., In *Expert System Applications in Chemisry*; Hohne, B., Pierce, T., Eds.; ACS Symposium Series 408; American Chemical Society: Washington, DC, 1989.

(2) The databases examined were REACCS from Molecular Design Ltd., San Leandro, CA, and SYNLIB from Distrubuted Chemical Graphics, Inc., Meadowbrook, PA.

(3) Borkent, J. H.; Oukes, F.; Noordik, J. H. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 148-150.

(4) Hendrickson, J. B. *J. Chem. Inf. Comput. Sci.* **1979**, *19*, 129.

0002-7863/91/1513-902$02.50/0

is overlaid on the reaction database to afford rapid searches, first to retrieve all the generally similar reactions and then to prune these down in matching detail of skeletal substitution as finely as desired. The generation of reactions in SYNGEN is based on this characterization of reactions, and so the successful linkage with literature databases also validates this approach to reaction classification.

Previous work on the derivation of reaction types by use of databases by Weise[5] used specific "group keys" (Gruppenschlüssel) to characterize structures involved in reactions from the SPRESI[6] (Speicherung und Recherche Strukturchemischer Informationen) database. His SYNKL program reportedly was able to classify correctly 17331 reactions from 25000,[7] producing 8611 individual reaction types, which were then used as an internal database to generate reactions in the AHMOS synthesis program. The Weise program differs from ours in being based on structures instead of reactions, and no attempt was made to define all possible reactions or to overlay major defined reaction classes on those types that were found in the database.

Characterization of Reactions

The basis for our classification is the simple but rigorous digital characterization of structures and reactions[8] developed for SYNGEN. In this system, a molecule is basically understood as a *skeleton*, or framework, of linked carbon atoms with *functionality* seen as reactive attachments on those skeletal carbons, i.e., as π-bonds or attached heteroatoms. The description of a molecule then consists of describing four synthetically important kinds of attachments, or bonds, on every carbon: H for bond to hydrogen, or electropositive atom; R for σ-bond to carbon; Π for π-bond to carbon; and Z for any bond (π or σ) to an electronegative heteroatom.

Each carbon in the skeleton is then described by the *numbers* of each kind of bond: h, σ, π, and z, respectively, and these must add up to 4. The functionality on any carbon is then broadly describable with two digits, z (0–4) and π (0–2). Hence, an entire molecule, with its skeleton carbons canonically numbered,[9] is simply annotated with a $z\pi$-list of two digits per carbon, listed in numerical order. The oxidation state at each carbon is $x_i = z_i - h_i$. This system provides a generalized or abstracted description of a molecular structure, coalescing many less significant distinctions. For closer discrimination, we have in SYNGEN further defined a subset of the z-value to impart the mechanistic function of any attached electronegative heteroatom;[1a] a similar expansion can be made for the electropositive atoms in the h-value, i.e., H, Si, B, and metals. The system has the advantage of rigor and simplicity, and incorporates all possible attachments on the skeletal carbons.

Reactions are described by their net structural change. For this we characterize a *unit reaction* as a unit exchange of these attachments[1a,3,4] on each involved carbon. These are described with two letters for each carbon, the first for the bond formed and the second for the bond broken. For the four kinds of attachment bonds there are therefore 16 unit reaction changes possible at each carbon involved, and these must describe every possible chemical change. Unit reactions at one carbon only will be reduction, HZ, oxidation, ZH, or substitutions, HH, ZZ, but unit reactions requiring more than one carbon are those that involve R or Π. A simple elimination involves two carbons, as in $\Pi H{\cdot}\Pi Z$, and the reverse, addition, reaction will be $H\Pi{\cdot}Z\Pi$. The construction reactions are RH, RZ, or RΠ at *each* of the two joining carbons, and their fragmentation counterparts are HR, ZR, or ΠR at each one; the RR change, characteristic of rearrangements, makes one C–C bond while breaking another and

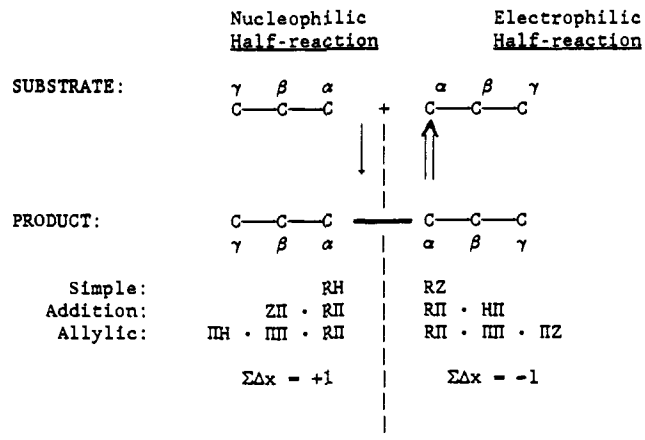


Figure 1. General characterization of construction reactions.

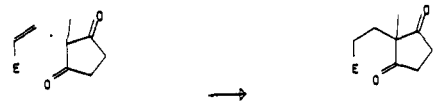
requires at least two other changing carbons to break and make the C–C bonds involved. The other unit reactions are *refunctionalizations* without alteration of the carbon skeleton, i.e., reactions with no change in R. The oxidation state change in any unit reaction can be quickly obtained from $\sum\Delta x$ between reactant and product.

The construction reactions are central to synthesis and to SYNGEN, and we explore here their classification and retrieval from commercial databases. A generalized summary of single construction reactions in Figure 1 shows a strand of up to three carbons, labeled α, β, and γ, out from the bond formed on either side and bearing the functionality that will change during construction. The full construction is seen as two half-reactions, each of which may be described by the possible unit reaction changes, shown in Figure 1. The six unit half-reactions include three oxidative ($\sum\Delta x = +1$), or nucleophilic, and three reductive ($\sum\Delta x = -1$), or electrophilic. These are categorized as simple substitution, on the α-carbon only, π-bond addition, requiring two carbons (α,β), and allylic substitutions, which require the full $\alpha\beta\gamma$-strand to change attachments. The allylic substitutions are vinylogues of simple substitutions, and we can also extend the additions to a vinylogous four-carbon strand ($\alpha\beta\gamma\delta$) in a similar way. Longer reactive strands are formally possible but very rarely seen.

Actual construction reactions are not all described by the net structural changes of Figure 1 because some are composite constructions, consisting of a construction–refunctionalization sequence in one step or operation. Previously discussed,[1a] these composites are recognized as SYNGEN: (1) prior reduction to an α-carbanion followed by RH nucleophilic construction; (2) elimination ($\Pi H{\cdot}\Pi Z$ or $\Pi Z.\Pi Z$) following construction; and (3) tautomerizations before or after construction. The first is exemplified by Grignard half-reactions, the second by the Wittig or dehydrative aldol reactions, and the third by common allylic prototropy. The overall net structural changes for these composite half-reactions have therefore been added to the list of six in Figure 1. With some further subdivision of large categories into chemically recognizable subheadings, we created a list of 25 construction half-reactions for SYNGEN: 16 nucleophiles and 9 electrophiles, described and labeled in Table I.

The two-character labels imply a letter first for nucleophiles and a number first for electrophiles, generally indicating the minimum necessary functionality level on the α-carbon of the reactant. The second character in the label is usually a number indicating the *span*, i.e., the number (1–3) of carbons exchanging attachments on the $\alpha\beta\gamma$-strand. The composites of the three categories above will be recognized as the following: (1) reductive carbanions R1, R2, R3, and RT; (2) elimination to double bond across the bond formed (α–α'), nucleophiles E1 and F1 and electrophiles 2E, or at the (α–β) bond, 2F; and (3) tautomerizations C2, A3, P3, and RT. In SYNGEN, the coupling of two half-reactions to a full construction demands one nucleophile and one electrophile for an overall isohypsic[8] ($\sum\Delta x = 0$) construction. The SYNGEN output of reactions is displayed[1] as the structures

(5) (a) Weise, A. *Z. Chem.* **1975**, *9*, 15. (b) Weise, A. *J. Prakt, Chem.* **1980**, *5*, 322.
(6) Sevjakova, L. A.; Sharnow, H. G.; Schildmann, G. *Informatik* **1977**, *24*, 17.
(7) Weise, A. *Z. Chem.* **1979**, *2*, 19.
(8) Hendrickson, J. B. *J. Am. Chem. Soc.* **1971**, *93*, 6487.
(9) Hendrickson, J. B.; Grier, D. L.; Toczko, A. G. *J. Chem. Inf. Comput. Sci.* **1983**, *23*, 171: **1984**, *24*, 195.



A1 Enolate
12 Conjugate addition

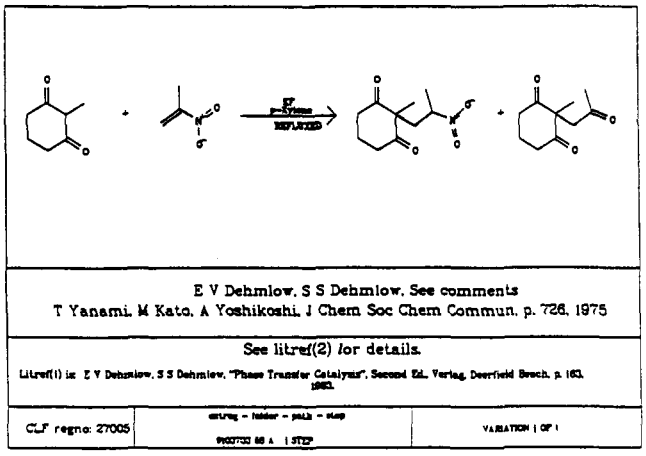


Figure 2. SYNGEN output and REACCS precedent (A1·12 reaction).

2E Carbonyl addition/elimination (alpha-alpha')
F1 Carbanion with Z-elimination (Wittig)

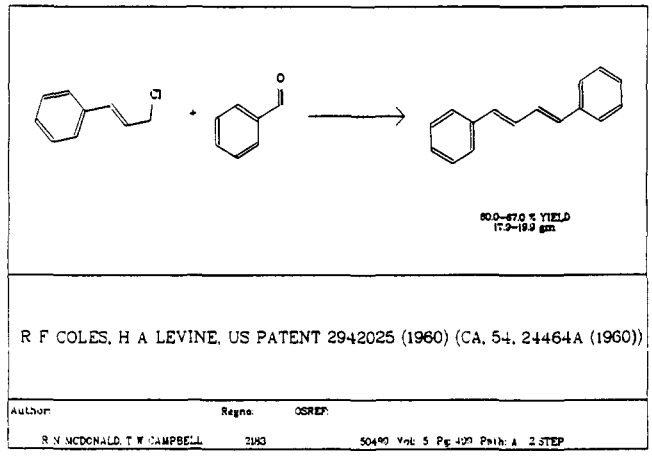


Figure 3. SYNGEN output and REACCS precedent (F1·2E reaction).

of the two reactants with the constructing bonds (1 or 2) marked, and these labels of the half-reaction pairs are shown in the following text. Two such reactions are shown at the top of Figures 2 and 3 for one construction and in Figure 5 for two (E, electron-withdrawing heteroatom group, such as $-NO_2$, $-SO_2R$, $-P^+R_3$, etc.).

It is these displayed reactions for which we will seek precedents in the reaction databases. The net structural change in each half-reaction can now be characterized by the change in the $z\pi$-list of the reactive $\alpha\beta\gamma$-strand from reactant to product. This $\Delta z\pi$-list is defined as $(z_\alpha\pi_\alpha z_\beta\pi_\beta z_\gamma\pi_\gamma)_{REAC} - (z_\alpha\pi_\alpha z_\beta\pi_\beta z_\gamma\pi_\gamma)_{PROD}$ for the minimal necessary reaction functionality, i.e., the parent reaction (Table I). In the computer, 2 bits is enough to record either z (0–3) or π (0–2), so the $\Delta z\pi$-list requires only 12 bits to describe a half-reaction family. The $\Delta z\pi$-lists are tabulated in Table II. Since some half-reactions have the same $\Delta z\pi$-list (cf., $\Delta z\pi$-list = 0), they may be distinguished by the value of $z\pi$ in the product.

Translation and Classification of Database Entries

The terms outlined above to characterize reaction change must now be overlaid on the databases to classify and presort their individual entries for rapid retrieval and matching with SYNGEN output reactions. Since the SYNGEN output consists solely of constructions, only constructions are developed here, but the same procedure will ultimately allow fragmentations and refunctionalizations to be cataloged analogously.

We examined the three common REACCS[10] databases: *Organic Syntheses* (ORGYSYN-89.1), *Theilheimer* (THEIL-88.1), and the Current Literature File (CLF-89.2); all are documented alike. The data available for our search consisted of connectivity tables of connected atoms and their connecting bond orders, for both reactants and product(s). Because some entries (Registry Numbers) display more than one product for a reaction, the total number of reactions, i.e., products, is larger than the number of entries; in many cases the two products are only stereoisomers.

In order to identify the net structural change in any reaction, we also require a correlation or mapping of each numbered product atom with its corresponding reactant atom number. The REACCS databases were originally compiled without this feature, but an extensive program for automatic atom–atom mapping has since (Version 7.1) been overlaid on these databases.[11] With correct atom mapping, a comparison of the atoms in reactant and product allows a direct tabulation of the bonds made and the bonds broken in any reaction. Detailed description is provided in the Experimental Section. This allows an initial breakdown of the total database into those reactions that construct one C–C σ-bond, those with two, and those with more C–C σ-bonds constructed. For purposes of SYNGEN comparison, all reactions that break C–C σ-bonds are set aside, i.e., fragmentations and skeletal rearrangements. Since SYNGEN now also recognizes heteroatoms N, S, and O as skeletal atoms,[1] the making of single C–N, C–S, and C–O bonds is also separately recorded. Other reactions with no change in the skeleton are refunctionalizations, i.e., redox, addition–elimination, etc. These are not used in SYNGEN but are important for other searches and are pursued in a later publication.[12]

This presorting of reaction categories is enumerated in Table III for each database. Of a total of 87716 REACCS reactions, there are 16753 single C–C constructions, somewhat less than one-fifth of the reactions. The proportion in the CLF (27%) is higher than in the other, older databases. This reflects an understandable increase in interest in these important reactions in the modern synthetic literature. The record shows a much smaller number of double and multiple constructions and of fragmentations and rearrangements. The refunctionalizations are not directly searched and so simply represent the remainder;[13] this total of 61916 reactions also includes 9347 reactions with unmapped atoms, which are inaccessible to our procedure. A common source of imperfect mapping occurs when reactants with skeletal carbons are written as text over the reaction arrow and so cannot be correlated. Many

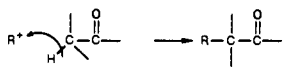
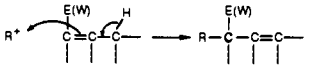
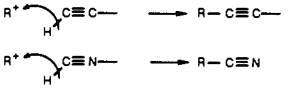
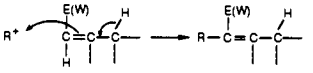
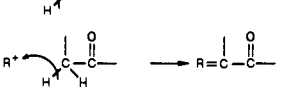
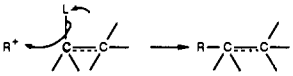
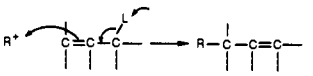
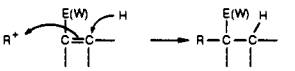
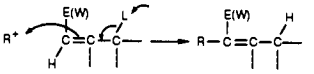
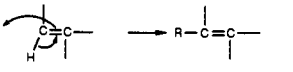
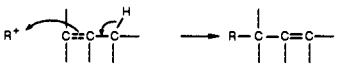
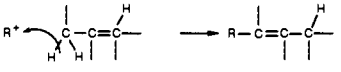
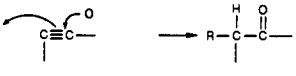
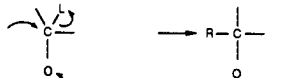
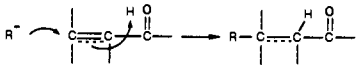
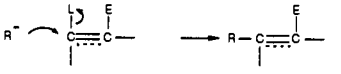
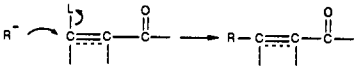
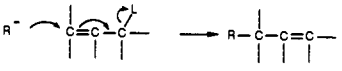
(10) Wipke, W. T.; Dill, J.; Hounshell, D.; Moock, T. E.; Nourse, J. G.; Grier, D. In *Modern Approaches to Chemical Reaction Searching*, Willett, P., Ed.; Gower: Aldershot, Hampshire, England, 1985; pp 92–117.

(11) Moock, T. E.; Nourse, J. G.; Grier, D. L.; Hounshell, W. D. In *Chemical Structures: The Universal Language of Chemistry*; Proceedings of the CSAC; Springer-Verlag, Springer-Verglag: New York, 1988; p 303.

(12) Hendrickson, J. B.; Miller, T. M. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, in press.

(13) The totals for the single C–N, C–S, and C–O bonds are also incorporated in the totals for refunctionalizations. In some entries with two "products", one may simply be unreacted starting material or the product of a reagent molecule and so of no importance. These cases are few, limited at most to the difference between entries and reactions (i.e., products) in Table III.

Table I. Construction Half-Reactions Used in SYNGEN

label	substrate[a] → product[a] $R^+ + \alpha\beta\gamma \rightarrow R - \alpha\beta\gamma$		label	substrate[a] → product[a] $R^+ + \alpha\beta\gamma \rightarrow R - \alpha\beta\gamma$	
	I. Activated Carbanion Nucleophiles				
A1		enolate (CO-stabilized carbanion)	F1		B1 with elimination of E
B1		hetero-stabilized carbanion	B3		allylic carbanion (stabilized)
C1		acetylene anion	P3		B3 tautomerized after construction
D1		cyanide anion	A3		B3 tautomerized carbanion construction
E1		A1 with elimination of H			
		B1 with elimination of H			
	II. Reductive Carbanion Nucleophiles				
R1		reductive carbanion	R3		reductive allylic carbanion
R2		reductive carbanion	RT		R3 tautomer
	III. Π-Nucleophiles				
P1		π-nucleophile: H-substitution	B3		allylic π-nucleophile
B2		π-nucleophile: RZ-addition	A3		B3 tautomer
C2		acetylene + H_2O addition			
	IV. Electrophiles				
I1		alkylation	I2		conjugate addition
21		carbonyl addition			
31		acylation	22		conjugate addition/ elimination
41		carboxylation			
2E		carbonyl addition/ α′-elimination	I3		allylic alkylation
2F		carbonyl addition/ β-elimination			

[a] Key: E, electron-withdrawing heteroatom group; W, carbonyl electron-withdrawing; L, leaving group.

of these may be constructions, as in the examples shown in Figure 4.

With the subset of single-construction reactions in hand, we proceed to classify them by their half-reactions. For each construction entry in the database, all $\alpha\beta\gamma$-strands out from each end of the constructed bond are identified from the connectivity table of the product. Then, the strand that changes its functionality, from reactant to product, is identified. This reactive strand is labeled ($\alpha\beta\gamma$), and its atom attachments are converted into z,π-terms in order to subtract the $z\pi$-list of product from that of reactant to obtain the $\Delta z\pi$-list identity (see Experimental Section). This is then matched against the $\Delta z\pi$-list numbers for the defined half-reactions of Table II. Both strands of a construction are so identified, and the half-reaction pair is assigned.

There are in Table II some ambiguous $\Delta z\pi$-lists, the same for both nucleophile and electrophile. These are the reductive constructions; i.e., the change for R1 is the same as that for I1 (alkylation)—both are represented overall by R–X → R–R′. The ambiguous pairs are R1 with I1, 21, 31, 41 and 22, R2 with I2, and R3 with I3. In these cases, the entry is recorded both ways, e.g., for —C=CX + CH_3X → —C=CCH_3, the assignment will be R1·I1 or 22·R1 depending on which substrate is read as the nucleophile (R1).

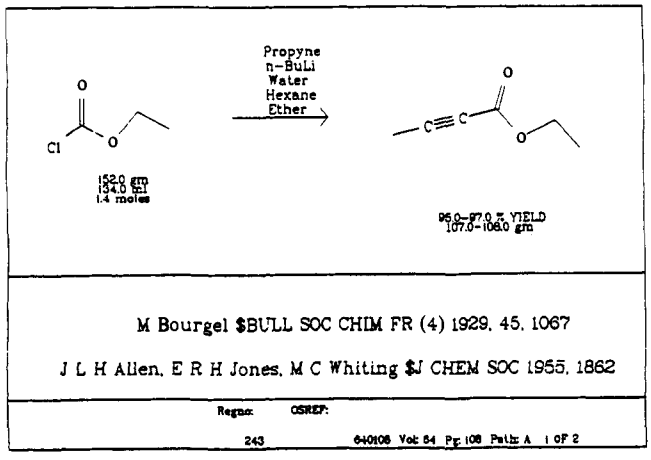
Assessment of Results

When the assignment is successfully made for a reaction, it is identified by its two half-reactions (Table I) and entered in a 9 × 16 matrix of electrophile × nucleophile half-reactions for rapid retrieval by SYNGEN. Thus, a simple aldol reaction is A1·21 and one with dehydration to an olefin across the constructed bond is E1·2E. The numbers of construction reactions classified in this matrix for each half-reaction pair are shown as Table IV. The ambiguous assignments described previously are kept in both categories for SYNGEN retrieval but only counted once in the

Table II. Characteristic $\Delta z\pi$-Lists for Half-Reactions

	$\Delta z\pi$-list (SUB-PROD)[a]		product $z\pi_\alpha$[b]		
nucleophile	bits	hex	bits	hex	electrophile
A1	0	0	0000	0	
B1	0	0	0100	4	
B1	0	0	1000	8	
C1	0	0	0010	2	
D1	0	0	1100	C	
P1[c]	0	0	0001	1	
E1	-000100000000	-100			
F1	001100000000	300			
	011100000000	700			2E
R1	010000000000	400	0000	0	11
R1	010000000000	400	0100	4	21
R1	010000000000	400	1000	8	31
	010000000000	400	1100	C	41
R1	010000000000	400	0001	1	22
R1	010000000000	400	0010	2	22
	011011110000	6F0			2F
B2	000011010000	0D0			
C2	000110100000	1A0			
R2	000100010000	110			12
A3	-000011111111	-0FF			
B3	000011111111	0FF			
R3	000100000011	103			13
RT	000000000100	004			

[a] The $\Delta z\pi$-list shown as 12 bits or as its hexadecimal equivalent for atoms α, β, and γ, with use of 4 bits for each atom $z\pi$. [b] Product $z\pi_\alpha$ (compare Table I), used only to distinguish ambiguities: 2 bits for z and 2 bits for π. [c] The P3 half-reaction of Table I has the same $\Delta z\pi$-list and is subsumed under P1 here.



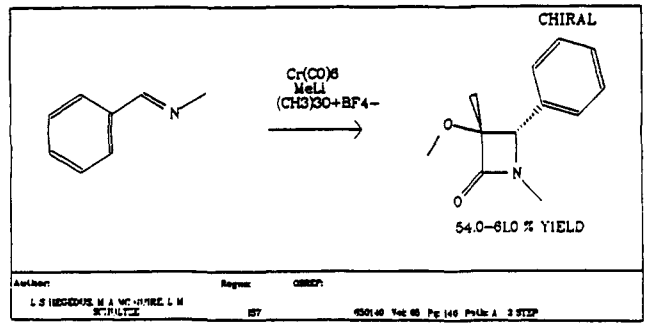


Figure 4. Unclassifiable REACCS entries.

enumerations of Tables III and IV. The *net structural changes* in the half-reactions in Table I are presented with an overlay of mechanistic arrows. The examples from the databases that are organized into the half-reactions matrix of Table IV are in fact virtually always found to be mechanistically consistent with these descriptions in Table I, although only the net structural change is actually required for identity. The matrix form of Table IV represents an efficient storage for fast retrieval of construction reactions. The output from SYNGEN shown at the top of Figure 2 was generated as an A1·12 construction. A representative precedent from that category of REACCS is shown below it. Figure 3 similarly represents a REACCS entry retrieved for the F1·2E reaction (above it) generated by SYNGEN.

The relative frequencies of the different half-reactions in REACCS are of interest to organic chemists since they imply a measure of the scope and reliability of any construction reaction that might be chosen for a synthesis. The most frequently found nucleophiles in Table IV are the enolate anion at 17% without subsequent elimination (A1) and another 8% with elimination (E1), together about one-fourth of all nucleophiles. A close second, nearly another one-fourth (23%), is R1, the reductive carbanion typical of Grignard reagents. The next most common (17%) is the π-nucleophile (P1), usually as the aromatic ring in aromatic substitution. Following these are the enol ether/enamine family (B2)[14] with 12% and the heteroatom-stabilized carbanions (B1) with 10%. Nucleophiles E1 and F1 construct a double bond, the former as the aldol reaction with dehydration (8%, above) and the latter the Wittig nucleophile half-reaction (4%). These are composite half-reactions that can only be used with a carbonyl electrophile followed by dehydration, i.e., 2E, hence, the open positions for them in Table IV. The least common half-reactions are the composites with tautomerization C2, A3, and RT, which have very few examples.

Among electrophiles, the ketone/aldehyde is the most common by far for simple carbonyl addition (21) and the composite addition–eliminations (2E and 2F), giving a total of over 5000 of the 12000 constructions, some 42%. Following this are the simple alkylations (11) with 16% and acylations (31) with 14%. The conjugate additions (12) are next with 11%, leaving only about one-sixth of the total for the other electrophiles. An examination of the three REACCS databases separately did not show any significant differences in the proportions of these several half-reactions.

Accessible REACCS entries that cannot be classified are those in which one or both of the $\Delta z\pi$-lists for the two half-reaction strands are different from the standard entries of Table II, or both are correct but of the same polarity, as A1·A1 or P1·B2. Paired nucleophiles like these constitute oxidative coupling ($\sum\Delta x = +2$), which is disallowed in SYNGEN and so not retained in the classification here, but paired electrophiles (reductive coupling, $\sum\Delta x = -2$) are included, implicit in the reductive nucleophiles R1, R2, R3, and RT, which have the same net structural change as their counterpart electrophiles.

The totals in Table III show that, of 16573 single C–C constructions, 12160 were classified correctly (and tabulated in Table IV). Of the remaining 4413 constructions, 470 were unclassifiable because of $z\pi$-changes at more than one β- or γ-carbon on one side of the construction. These are generally concurrent changes unrelated to the construction itself, as in an attendant aromatization. The 3943 constructions then remaining were tabulated by the $\Delta z\pi$-lists of their half-reactions, the most frequent 25 shown in Table V. The most frequent is $\Delta z\pi$-list = 0 for A1, etc. (Table I), seen either paired with another as oxidative coupling or paired with a nonstandard $\Delta z\pi$-list. For example, a common nonstandard $\Delta z\pi$-list is 100, interpreted as a loss of one π at the α-carbon and no change at β or γ. This surprising change is generally found to involve a four-carbon reactive strand, i.e., R + C=C—C=C + Z → R—C—C=C—C—Z. Since we only examine three carbons (α, β, γ) we miss the change of ZΠ at the fourth carbon and at β and γ $\Delta z\pi = 0$. We found virtually no other half-reactions with strands of four or longer. In these constructions, one half-reaction is usually a standard one (Table I) but the whole construction is unclassifiable because the partner half-reaction is not standard or two standard nucleophiles are paired in an oxidative construction.

(14) The B2 half-reaction is shown in Table I in its lowest oxidation state, essentially as a Friedel–Crafts nucleophile. The same net structural change in a higher state is the enol/enamine nucleophile; i.e., Z—C=C + R^+ → O=C—C—R.

Table III. Enumeration of Reaction Categories in REACCS and SYNLIB Databases

	ORGSYN	THEIL	CLF	totals REACCS	totals SYNLIB
C–C constructions*	729	6721	9303	16573	10364
double constructions	143	993	2450	3586	2252
multiple constructions	60	445	543	1048	263
C–C fragmentations and rearrangements	141	2247	2025	4413	4598
refunctionalizations	4317	37570	20029	61916	45424
single C–N formed*	325	4635	1778	6738	2645
single C–O formed*	60	905	508	1473	602
single C–S formed*	160	1901	1269	3330	1694
total reactions	5390	47976	34350	87716	62901
total entries	5018	46818	28449	80285	62901
*classified constructions					
single C–C	555	4758	6847	12160	6521
single C–N	214	2887	1135	4244	1356
single C–S	35	544	274	853	328
single C–O	93	892	565	1550	563
total classified				18854	8768
overall				27622	

Table IV. Classification of Single Construction Reactions in REACCS

E:	11	21	31	41	2E	22	12	13	2F	totals
Nu										
A1	540	567	334	75		77	265	162	12	2032
B1	359	384	120	30		46	140	99	11	1189
C1	52	40	28	12		53	6	2	0	193
E1					972					972
F1					523					523
D1	23	110	24	9		20	28	10	4	228
A3	0	3	7	0		1	2	0	3	16
R1	368	924	363	80		391	425	175	41	2767
R2	0	79	32	17		0	47	0	1	176
R3	0	134	9	9		0	17	2	1	172
RT	4	6	3	2		5	1	1	0	22
P1	305	288	684	148		240	207	24	173	2069
B2	247	614	152	64		52	112	189	2	1432
C2	4	5	1	0		0	3	3	0	16
B3	43	198	10	1		15	79	6	1	353
totals	1945	3352	1767	447	1495	900	1332	673	249	12,160

Table V. Frequency of Occurrence of $\Delta z\pi$-Lists for Unclassifiable Constructions in all REACCS Databases

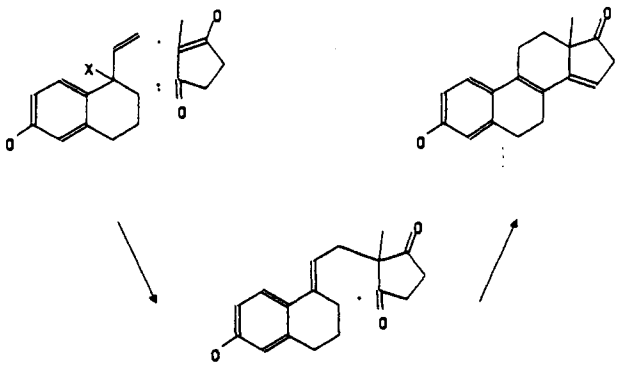
freq	$\Delta z\pi$-list[a]	half-reaction
1776	000	A1, B1, C1, D1, P1, P3
1114	400	R1, 11, 21, 31, 41, 22
539	700	2E
339	0D0	B2
329	800	
327	100	
321	-0D0	
300	300	F1
234	-100	E1
152	-400	
131	110	R2, 12
112	-030	
111	-001	
89	600	
88	003	
77	6F0	2F
77	040	
64	3C0	
64	330	
57	0FF	B3
53	6FD	
51	103	R3, 13
49	403	
44	3D0	
40	004	RT

[a] In hexadecimal numbers representing change at α, β, and γ, (4 bits each).

There are often entries in REACCS with multiple steps condensed as one reaction, and so with $\Delta z\pi$-lists different from those in Table II. However, a common practice in REACCS for these multistep situations is to record a separate entry for each of the steps and then to summarize the overall transformation in another entry. In these cases, the entry for the construction step will be located and correctly assigned while the entry for the overall transformation is then extraneous and may be safely ignored. In other cases, the intermediates in the multiple-step transformation cannot be isolated and so no simple construction is found. Thus, the traditional Skraup synthesis of quinolines from anilines and glycerol with strong acid (Entry No. 4885 in ORGSYN) incorporates a succession of six unit reactions (including an oxidation) and cannot exhibit a simple $\Delta z\pi$-list corresponding to any in Table II.

Even dropping these cases, however, we are left with about three-fourths of the accessible C–C construction reactions in REACCS that can be classified. Ths makes available an accessible library of some 12000 reactions for retrieval as literature precedents by SYNGEN.

When SYNGEN displays two constructions together as an annelation (Figure 5), one joining (intermolecular) construction and one cyclization, each may be separately examined in the REACCS precedents (for which we also annotate whether every construction is intermolecular or intramolecular). The first (B2·13) construction finds a simple model in the top REACCS entry in Figure 6, while the cyclization step is very closely matched by the REACCS analogue below it in Figure 6.



13 Allylic alkylation
B2 Pi-Nucleophile: Addition +Construction
A3 Allylic carbanion tautomerized
2F Carbonyl addition/elimination (alpha-beta)

Figure 5. SYNGEN output for an annelation.

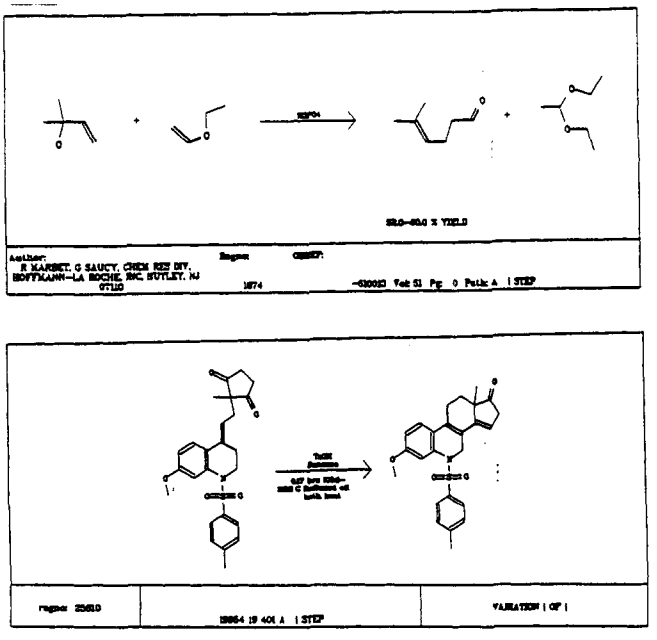


Figure 6. REACCS precedents for SYNGEN annelation.

Furthermore, in those cases for which the skeleton is taken as including hetereoatoms[18] there are also those reactions in Table II that create only single C–N, C–S, or C–O skeletal bonds, a total of 11 541 reactions in REACCS. These were also classified as half-reaction pairs just as with the C–C bond formation of Table IV and those classified are shown at the bottom of Table III; there are about 6600 that can be classified, i.e., only about 58%, rather less than the 73% of the C–C constructions classified. This lower proportion arises from a larger number of involved molecular species in which heteroatoms are not fully mapped.

New Half-Reactions

The frequent half-reactions with nonstandard $\Delta z\pi$-lists on $\alpha\beta\gamma$ proved interesting (Table V). Our original analysis shows that all possible unit reactions for construction (net structural change) are the six conversion of Figure 1, i.e., the nucleophiles RH (half-reactions A1, B1, C1, D1, and P1), RΠ·ZΠ (B2), and RΠ·ΠΠ·ΠH (B3) and the electrophiles RZ (11, 21, 31, 41, and 22), RΠ·HΠ (12), and RΠ·ΠΠ·ΠZ (13). The rest of the half-reactions in Table I are composites, a construction half-reaction combined with a refunctionalization reaction (before or after) to make a single operation with two mechanistic or unit-reaction steps. It is clear that more such composites are possible than the common ones selected for our standard SYNGEN set in Table I. These certainly constitute the most frequent cases (Tables IV and

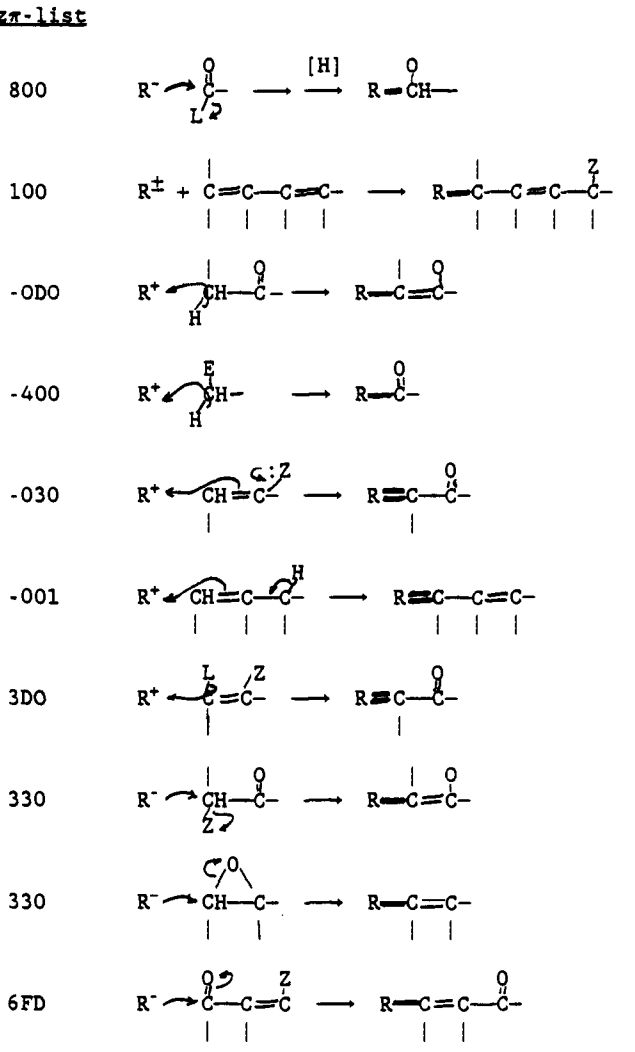


Figure 7. Potential new half-reactions from REACCS.

V) but some are quite uncommon, i.e., C2 and A3. All the other 23 half-reactions of Table I are found in the top 25 in the frequency poll of Table V, i.e., well-represented as halves of the unclassifiable full constructions.

The remaining 14 $\Delta z\pi$-lists in Table V that turn up often enough to demand attention were then examined as potentially important composite half-reactions that might usefully expand our standard set of Table I. The most common was $\Delta z\pi$-list = 800, representing $\Delta z_\alpha = -2$, which is commonly a construction to an electrophilic carboxyl derivative ($z = 3$) with subsequent reduction to $z = 1$, shown in Figure 7 (formatted as in Table I). However, it also includes diazo groups as electrophiles with $z = 2$ going to $z = 0$, as in $C{=}N_2 + R^- \rightarrow CH{-}R$. The next most common nonstandard $\Delta z\pi$-list is 100, discussed previoysly, and then –ODO, which represents a keto → enol change *after* construction, often associated with aromatization or a heterocyclization, as to an enol–lactone. It will be seen in Figure 7 to be the reverse of ODO, the change of an enol form to a ketone in the construction B2.[14]

The next new half-reaction is –400, shown in Figure 7 as an electron-withdrawing group (E = $-NO_2$, –SOR, etc.) stabilizing a carbanion as with B1 but subsequently oxidized to ketone, essentially an *umpolung*, or acyl anion equivalent. Figure 7 also collects from their REACCS frequencies three nucleophilic half-reactions that incorporate the subsequent elimination seen in E1, i.e., aldol nucleophile with dehydration. Two (–030 and 3DO) are enol equivalents of the same thing, the latter a reductive carbanion, and the third (–001) is basically an ene reaction nucleophile (B3) with further elimination, often occasioned by aromatization in the REACCS examples where it is found. The

$\Delta z\pi$-list of 330 represents an α-halo ketone electrophile that subsequently enolizes also.

The $\Delta z\pi$-list 003 involves addition to a quinone and reoxidation, and 3CO is the result of various constructions with oxidation, while 040 and 403 show no general pattern. Many of these less common composites, which end with a π-bond, are actually involved in an aromatization. The last two electrophiles in Figure 7 are especially interesting. The first, also a $z\pi$-list of 330, represents a subsequent dehydration after epoxide opening and turned up with some frequency. The last (6FD) is the synthetically attractive use of an enolized β-diketone to invert the ketone position on construction. With the exception of 100, all exhibit strands of three carbons or less like the standard half-reaction set of Table I, with which they are compared.

Treatment of the SYNLIB Database

In the case of the SYNLIB database (Version 3.0, July 1990), access was created in a somewhat different way. The format for an entry in SYNLIB is different from that in REACCS and allows only one reaction product per entry. When the reaction is entered into the database, the product structure is first described and stored with atom connectivity tables. The reactant structures are not stored as such, but rather are generated by application of a *manipulation table*, which is a record of the actions taken by the person entering the reaction to transform the target structure to substrate structure(s). Since the available operator commands for these manipulations allow adding and deleting atoms, the complete mapping of atoms from product to substrate may often be obscured and the atom–atom correspondences lost.

Since the procedure for identifying and classifying constructions had already been developed for REACCS, we created a file of SYNLIB entries reformatted as REACCS entries, complete with full atom mapping. From 62 901 SYNLIB entries this provided us with a subset of ~43 000 SYNLIB reactions converted to REACCS format for the further classification shown in Table III. From ~10 000 single constructions we could classify about 6500 constructions as two valid half-reactions, and these were similarly ordered into a 9 ×16 matrix like that of Table IV. As with REACCS, the number (45 424) listed for refunctionalizations includes the 19 660 entries that were not fully automapped. An initial attempt to classify SYNLIB using its own manipulation table for atom mapping afforded only about 3000 classified constructions. With the conversion to REACCS format and the use of the REACCS atom-mapping routine[11] (as provided by REACCS Version 7.1 for user entries), we were able to more than double the number of constructions classified.

Conclusion and Practical Use

In conclusion, we find that we are able to examine these two raw databases with a program interface based on unit reaction descriptions[4] and select out single and composite constructions of C–C, C–N, C–S, and C–O bonds that we can sort and rapidly access as literature precedents for SYNGEN. This procedure has provided us with a substantial literature database of ~28 000 references. Furthermore, the large proportion of finds lends confidence that the approach in ref 4 is a valid one and has allowed us to further classify the remaining refunctionalization reactions in a parallel way.[12]

In practice, the SYNGEN operator designates a generated construction in the output and simply requests a database search. This will immediately turn up a set of all examples in the database that have the same half-reaction pair, i.e., a set from the Table IV matrix. Since this set will often be very large, he may then apply a series of structural matching criteria, i.e., inter- vs intramolecular reaction or identity of σ-, z-, and/or π-values at the two α-(joining) atoms. If these do not provide adequate pruning, he may further elect identity of the β- and γ-atom values as well. The progress of the pruning is followed on the screen by the numbers of examples left after each selection, as illustrated by a typical screen in Figure 8 (for the case shown in Figure 2). When the number is small enough to evaluate, he may switch to seeing the particular precedents individually on screen, as in the examples of Figures 2, 3, 5, and 6. We find that these precedents are more completely and quickly located with our system than with available search methods in the REACCS or SYNLIB systems themselves.

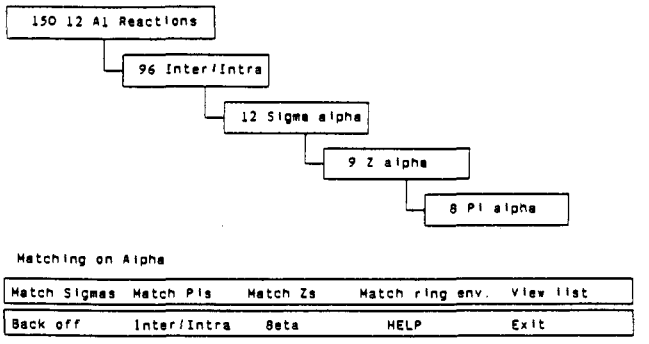


Figure 8. Pruning screen for matching reaction precedents.

Experimental Section

SYNLIB Conversion to REACCS Format. REACCS (Version 7.1) provides a Reaction Descriptor File (RD-file)[15] to move entries in and out of the standard (compressed) format in REACCS itself. A program was written to transform the SYNLIB product tables and manipulation tables into the RD-file format so that the SYNLIB database could be processed through this conversion to create a REACCS database of its entries. We found that the average entry in the SYNLIB database requires only about half as much space as an entry in the standard REACCS database. Once the SYNLIB database was transcribed into a REACCS RD-file, the automatic REACCS atom-mapping protocol[11] was applied to make a fully mapped RD-file. This was then used in the same way as the other (REACCS) databases for the subsequent analysis of its construction reactions.

Assignment of Atom Types and Recognition of Bonds Constructed. The following instructions are followed sequentially for each database entry. Atoms C, N, O, and S may be skeletal and are retained; hydrogens are converted to h-values for the skeletal atoms to which they are attached, and halogens and phosphorus are all reduced to z-values on their attached atoms. Metals are treated as reductively derived, hence as z-values on attached carbon. Starting materials are read into memory with a reserve copy made to be retrieved for multiple products or successive classification phases. For each atom number in the product, the substrate is scanned and the corresponding (mapped) atom in the substrate is located. Next, attachment lists for each atom (using only potentially skeletal atoms, C, N, O, and S) are created and compared, substrate to product. In correctly mapped pairs, the bonds made or broken are thus identified by the change in attachments of the atoms at each end of such bonds. The number and identity of such bonds are stored. Multiple C–C constructions and C–C fragmentations are omitted from further classification.

Sequential Examination of Construction Types. Since SYNGEN accommodates N, O, and S as skeletal atoms as well as carbon, we seek separate entries creating C–C, C–N, C–O, C–S, N–N, and N–S bonds. Each reaction is examined for classification four times, each with a different level of acceptance of heteroatoms as skeletal: C, only skeletal; C and N, skeletal; C, N, and S, skeletal; C, N, S, and O, skeletal. Heteroatoms not accepted as skeletal in any pass are reduced to z-values on the skeletal atoms to which they are attached. Thus, for the first iteration we keep all reactions making one C–C bond, even if C–N, C–O, etc., bonds may also be formed, and reject reactions making only C–N, C–S, etc., bonds as refunctionalizations. At the second iteration we retain in a second file single C–N or N–N bonds formed (with no C–C formed), ignoring C–S and C–O bonds formed. At the third iteration (and third file) we accept single C–S and also N–S bonds formed, etc. Each reaction will then be classified for half-reactions at the appropriate level of skeletal heteroatoms, and four matrixes are created from the four files (cf., Table IV for C–C constructions). The nature of the net structural change defining a construction half-reaction (i.e., Table II) is formally independent of the nature of the skeletal atoms at α, β, and γ (Table I). Therefore, the same classification into 25 half-reaction changes can be used with heteroatoms in the $\alpha\beta\gamma$-strand; some of these changes, however, will be disallowed by valence, as with Cl, for example, with α-N or S.

Construction Strand Identification. For a single construction the two α-atoms at each end of the constructed bond are now identified. The program next determines the least atom path between the two α-atoms, if any, in the substrates, by a recursive procedure. Intramolecular constructions (cyclizations) are identified as nonzero paths, and the path

(15) RD-file is a proporietary format of Molecular Design, Ltd.[2]

length is taken as the size of the ring formed. Next all possible β- and γ-atoms, out from each α-atom, are identified as the possible reaction strands. At any time the only strand atoms considered are those defined as skeletal for the current iteration level. β-Atoms are those attached to α (but not identical with the other α-atom in the product). γ-Atoms are those attached to each β but not identical with the other α- or any β-atom. All identified atoms are checked to be sure they are present in both substrate and product.

The h-, σ-, π-, and z-values of the $\alpha\beta\gamma$-atoms in each strand are recorded for both substrate and product, and for those that change in the reaction, the two $\Delta z\pi$-lists are calculated, one for each active $\alpha\beta\gamma$-strand. These are then identified as characterizing the two half-reactions that make up the construction of the bond (Table II). If there are changes in more than one β- or γ-atom out from an α-atom, the reaction is deleted as unclassifiable. When both strands are successfully characterized as half-reactions, the reaction entry number is recorded in the matrix array of Table IV; also recorded are the identity of the active strands, the σ-, z-, and π-values of the $\alpha\beta\gamma$-atoms of each strand, and the ring size for any cyclization.

The same procedure is followed iteratively for each level of skeletal atom identity. Thus, first only carbons are taken as skeletal atoms and N, O, and S are recorded as z-values on attached carbons. In the second iteration, nitrogen is also a skeletal atom and O and S atoms are recorded as functionality (z-values). An addition to an imine would be recorded as a 21 half-reaction in the first iteration but as a 12 addition with skeletal β-nitrogen in the second pass. The total reaction count is not increased by this duplication, but the reaction can be retrieved either as a carbonyl addition or as an imine addition. There are actually four matrixes created in this way for SYNGEN retrieval use, but only the C-C constructions are illustrated by Table IV.

Acknowledgment. We are grateful to the National Science Foundation (Grant CHE-8620066) and to the Eastman Kodak Co. for generous support of this work.

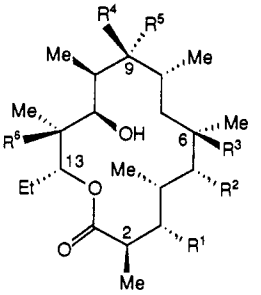
Total Synthesis of 9-Dihydroerythronolide B Derivatives and of Erythronolide B

Johann Mulzer,*,† Holger M. Kirstein,†,‡ Jürgen Buschmann,§ Christian Lehmann,§ and Peter Luger§,‖

Contribution from the Institut für Organische Chemie, Takustrasse 3, und Institut für Kristallographie, Takustrasse 6, der Freien Universität Berlin, D-1000 Berlin 33, FRG.
Received April 9, 1990

Abstract: A convergent total synthesis (22 steps on the longest linear route) of (-)-erythronolide B (**5**) and two 9-dihydro derivatives (**52** and **54**) thereof from (*R*)-2,3-*O*-isopropylideneglyceraldehyde (**20**) as the only source of chirality is described. A key step of the synthesis is the regio- and stereocontrolled coupling of the allyl sulfide anion **39** and ketone **26**, which can be directed to either α-adduct **40** or **41** by an appropriate choice of the conditions (Scheme V, Table II). From **40** and **41** the seco acids **47** and **49** are prepared, which are smoothly macrolactonized to **50** and **51** according to a modified Yamaguchi procedure. Hydroboration of **50** and **51** proceeds under macrocyclic stereocontrol to afford the 9-dihydroerythronolide B derivatives **52** and **54**, of which **54** is converted into **5** by a known oxidation-deketalization sequence.

The stereocontrolled total synthesis of erythromycin macrolides (**1-6**) has been an evergreen in organic chemistry for more than



1 R^1 = L-cladinosyl, R^2 = D-desosaminyl, R^3 = R^6 = OH, R^4=R^5=O
(Erythromycin A)
2 R^1 = L-cladinosyl, R^2 = D-desosaminyl, R^3 = OH, R^4=R^5=O, R^6 = H
(Erythromycin B)
3 R^1=R^2=R^3= R^6 = OH, R^4=R^5=O
(Erythronolide A)
4 R^1=R^2=R^3=R^4=R^6= OH, R^5 = H
(9S-Dihydroxyerythronolide A)
5 R^1=R^2=R^3= OH, R^4=R^5= O, R^6 = H
(Erythronolide B)
6 R^1=R^2= OH, R^3=R^6= H, R^4=R^5= O
(6-Deoxyerythronolide B)

one decade.[1] Whereas the diglycoside (erythromycin A, **1**) has been synthesized only once,[2] there are several syntheses of the aglycons erythronolide A (**3**)[3] erythronolide B (**5**)[4] and of 6-deoxyerythronolide B (**6**).[5] Additionally, a number of approaches to advanced synthetic intermediates has been reported, e.g., 9-(*S*)-dihydroerythronolide A (**4**)[6] and various seco acid derivatives

† Institut für Organische Chemie.
‡ Synthetic part of the work.
§ Institut für Kristallographie.
‖ X-ray analysis of compounds **46** and **52**.

(1) (a) Omura, S. *Macrolid Antibiotics: Chemistry, Biology and Practice*; Academic Press: Orlando, FL, 1984; p 127. (b) Quinkert, G. *Synform* **1988**, *6*, 1. (c) Nicolaou, K. C. *Tetrahedron* **1977**, *33*, 633. (d) Masamune, S.; Bates, G. S.; Corcoran, J. W. *Angew. Chem., Int. Ed. Engl.* **1977**, *16*, 585. (e) Paterson, I.; Mansuri, M. M. *Tetrahedron* **1985**, *41*, 3569. (f) Masamune, D.; Choy, W. *Aldrichimica Acta* **1982**, *15*, 47.

(2) Woodward, R. B.; Logusch, E.; Nambiar, K. B.; Sakan, K.; Ward, D. E.; Au-Yeung, D. W.; Balaram, P.; Browne, L. J.; Card, P. J.; Chen, C. H.; Chenvert, R. B.; Fliri, A.; Frobel, K.; Gais, H. J.; Garrat, D. G.; Hayakawa, K.; Heggie, W.; Hesson, D. P.; Hoppe, D.; Hoppe, I.; Hyatt, J. A.; Ikeda, D.; Jacobi, P. A.; Kim, K. S.; Kobuke, Y.; Kojima, K.; Krowicki, K.; Lee, V. J.; Leutert, T.; Malchenko, S.; Martens, J.; Matthews, R. S.; Ong, B. S.; Press, J. B.; Rajan Babu, T. V.; Rousseau, G.; Sauter, H. M.; Suzuki, M.; Tatsuta, K.; Tolbert, L. M.; Truesdale, E. A.; Uchida, I.; Ueda, Y.; Uyehara, T.; Vasella, A. T.; Vladuchick, W. C.; Wade, P. A.; Williams, R. M.; Wong, H. N. C. *J. Am. Chem. Soc.* **1981**, *103*, 3210, 3213, 3215.

(3) (a) Corey, E. J.; Hopkins, P. B.; Kim, S.; Yoo, S.; Nambiar, K. B.; Falck, J. R. *J. Am. Chem. Soc.* **1979**, *101*, 7131. (b) Kinoshita, M.; Ohsawa, N.; Gomi, S. *Carbohydr. Res.* **1982**, *109*, 5. (c) Kinoshita, M.; Arai, M.; Tomooka, K.; Nakata, M. *Tetrahedron Lett.* **1986**, *27*, 1811. (d) Kinoshita, M.; Arai, M.; Ohsawa, N.; Nakata, M. *Tetrahedron Lett.* **1986**, *27*, 1815. (e) Kochetkov, N. K.; Sviridov, A. F.; Ermolenko, M. S.; Yashunsky, D. V.; Borodkin, V. S. *Tetrahedron* **1989**, *45*, 5109.

(4) (a) Corey, E. J.; Trybulski, E. J.; Melvin, L. S.; Nicolaou, K. C.; Secrest, J. R.; Lett, R.; Sheldrake, P. W.; Falck, J. R.; Brunelle, D. J.; Haslanger, M. F.; Kim, S.; Yoo, S. *J. Am. Chem. Soc.* **1978**, *100*, 4618, 4620. (b) Kochetkov, N. K.; Sviridov, A. F.; Ermolenko, M. S.; Yashunsky, D. V.; Borodkin, V. S. *Tetrahedron Lett.* **1987**, *28*, 3835, 3839. (c) Reference 3e.

(5) (a) Masamune, S.; Hirama, M.; Mori, S.; Ali, Sk. A.; Garvey, D. S. *J. Am. Chem. Soc.* **1981**, *103*, 1568. (b) A formal synthesis of racemic **6** has been described: Miles, D. C.; Danishevsky, S. J.; Schulte, G. *J. Org. Chem.* **1990**, *55*, 1636.

0002-7863/91/1513-910$02.50/0