CONAN (CONnectivity ANalysis): A Simple Approach in the Field of Computer-Aided Organic Synthesis. Example of the Taxane Framework

F. Barberis, R. Barone,* M. Arbelot, A. Baldy, and M. Chanon

Laboratoire AM3, case 561, URA CNRS 1411, Faculté des Sciences de St Jérôme, 13397 Marseille, Cedex 20, France

Received June 20, 1994[⊗]

CONAN, a program which combines two previous programs, SAS[1,2] and REKEST,[3] is described. The aim of CONAN is to assist the chemist in the search for strategies in synthesizing complex structures. The taxane skeleton has been analyzed.

INTRODUCTION

One of the major problems during the conception of a synthetic plan is to find a strategy. This often leads to discovering the critical steps of the synthesis. The writing of a strategy may be compared with the choice of which route to select to go from one town to another. The discovery of a strategy would consist on choosing the main towns to go through on the map. Then one would have to search for a route between each town.

In his paper describing the synthesis of longifolene, Corey first indicated a simple approach for searching the key-step of a synthesis.[4]

We have developed the SAS (simulated analytical synthesis) program according to this approach,[1,2] the aim of which was to display rapidly all the possible ways to dissect a structure. Applied to the ellipticine skeleton, it suggested several interesting dissections, one of them being realized experimentally later by another group.[5] This method was limited however, because it was not connected with actual reactions. We have developed another approach with the REKEST program.[3] REKEST allowed a better manipulation of the target skeleton by means of "pseudoreactions" which deleted and/or added bonds.

P. Wender's recent paper dealing with "connectivity analysis" illustrated with actual cases the interest of such an approach.[6] Since SAS and REKEST have been written for an Apple2 microcomputer which is presently less used, we decided to develop a new program including these two approaches with several improvements for a more recent microcomputer. In this paper we describe this new program called CONAN (CONnectivity ANalysis). It runs on an IBM/PC microcomputer.

CONAN

In this approach, the skeleton of the target is manipulated by means of pseudoreactions which simulate basic mechanisms of organic chemistry at a skeleton level. Compared with MARSEIL[7] or STRAKS[8] which search the key step by means of actual reactions this one is more schematic. The name of R. E. Howard's hero evocates also the crudeness aspect of this approach, the aim of which is to reveal potential precursors that even a good chemist could miss.

Figure 1 shows the pseudoreactions used in this version of the program. Reactions 1−3 simulate the SAS approach by deleting one or two bonds.

These reactions are in a separate file, and the user may code his own schemes. For example it is possible to create a file simulating the reactions forming a C−C bond in a way similar to Hendrickson's[9] or Moreau's[10] approaches. One may compare this code to Ugi and co-worker's one.[11] Our code has not the mathematical aspect of Munich's group: it is a "classical" one by the search of the substructures which describe the reactions in the connectivity table of the target. Since the program uses the matching of connectivity tables, it is possible to code reactions such as Diels−Alder, Cope, Michaël, etc. Stereochemistry is not coded, and there is no evaluation for discarding uninteresting solutions: the aim is not to find the best synthesis but to give a maximum of creative ideas to the chemist by a rough and quick cutting of the target. For large structures the number of solutions may be high. Since only the skeleton of the target is given, the structures are generally medium sized, and the number of solutions is not extreme. Nevertheless we plan to develop an evaluation module in terms of complexity or synthetic accessibility. Oddly, such an approach, which may be simulated by other retrosynthetic programs, has not been specifically described.

The developed interface is user friendly; the drawing of structures directly uses the mouse and is driven by menus. Also, the coding of reactions is done graphically: the user draws the substructures describing the reactions as in the MARSEIL[7] and STRAKS programs.[8]

The results obtained were interesting. Nevertheless a new improvement was necessary: when a precursor appears on the screen, we added an option which allows the user to analyze it directly in order to obtain its own precursors. The new precursors obtained cannot be analyzed. The procedure is limited for a two-step retroanalysis. For example, with the strategy which adds one bond, CONAN generates precursors which possess new rings, and it is worthwhile to have a rough analysis of them when they are on the screen. There are two advantages in this improvement. First, CONAN may reveal tandem reactions, which are typical cases of holosynthons[12] or multiple cyclizations.[13] Such reactions are particularly interesting for the construction of

[⊗] Abstract published in *Advance ACS Abstracts*, April 15, 1995.

0095-2338/95/1635-0467$09.00/0

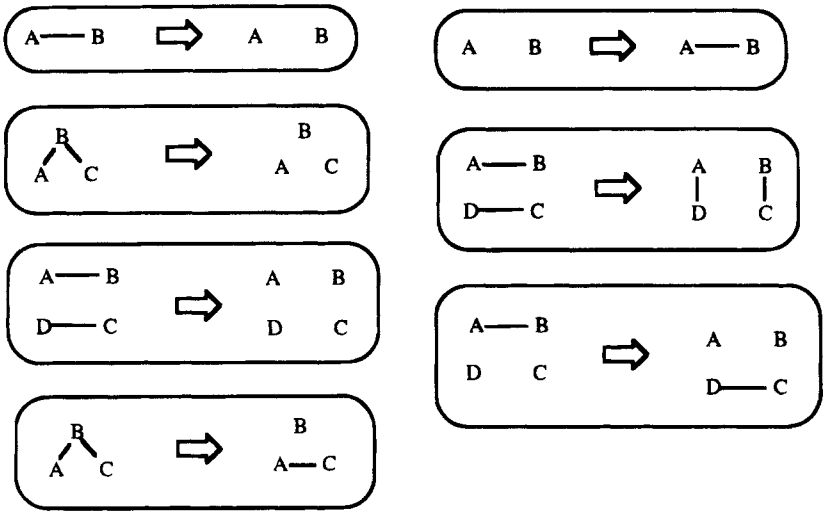


Figure 1. Pseudoreactions used in this approach. Letters A, B, C, and D stand for any atom.

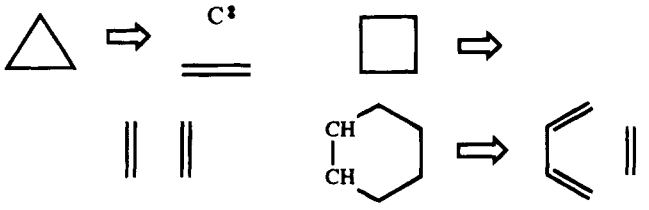


Figure 2. General schemes for the analysis of precursors.

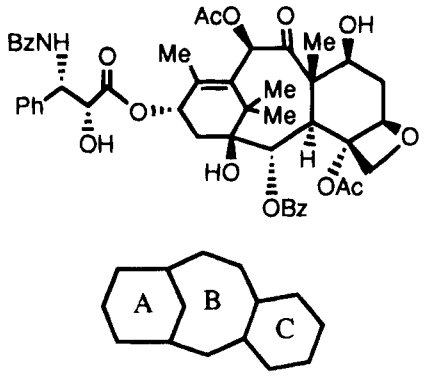


Figure 3. Structure of taxol and taxane skeleton.

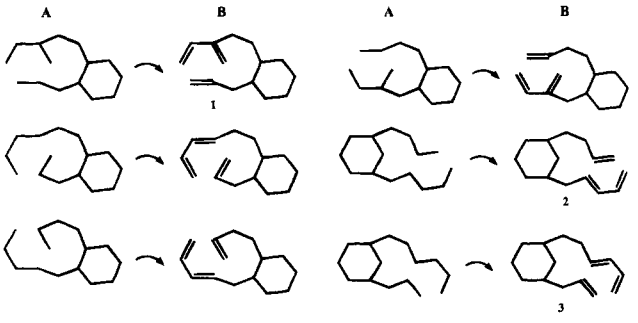


Figure 4. Precursors given by CONAN and the intramolecular Diels–Alder reactions suggested.

complex structures.[14] Second, CONAN allows both a breadth and depth manipulation of the retrosynthetic tree.

We have created two files of reactions, the first one includes the "reactions" of Figure 1, which give an idea for a general strategy, and in the second file we coded general schemes for the analysis of the precursors created in the first step such as [1 + 2], [2 + 2], or [4 + 2] cycloadditions

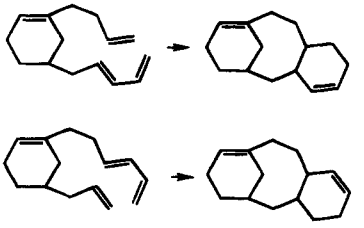


Figure 5. Suggestions for the formation of the bridgehead double bond in taxol.

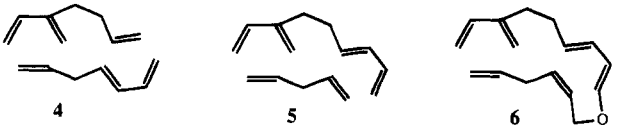


Figure 6. Tandem Diels–Alder reactions suggested by CONAN.

(Figure 2). It is of course possible, in the second step, to use the reactions of Figure 1.

RESULTS

In order to check the possibilities associated with this new program we tested it on the taxane skeleton which has aroused world-wide synthetic interest because of the anticancer activity exhibited by taxol[15] (Figure 3).

Figure 4 gathers solutions for the taxane skeleton in which two bonds have been deleted and which suggest intramolecular Diels–Alder reactions. Precursors given by the program are in column A. The Diels–Alder solutions which are suggested are in column B.

Solution **1** has been reported.[16] One difficulty in the synthesis of taxol is the formation of the bridgehead double bond. Solutions **2** and **3** suggest an interesting possibility: in these solutions the double bond could be already present in ring A and become bridgehead by the formation of ring C of taxane (Figure 5). A strategy similar to solution **2** was recently published.[17]

Analysis of precursor **1** for the formation of ring C and the analysis of **2** and **3** for the formation of ring A lead to solutions **4** and **5** (Figure 6).

Solution **5** suggests an efficient tandem Diels–Alder reaction. Such kind of tandem reaction has been successfully attempted by Krauss[18] (Scheme 1). It also suggests the

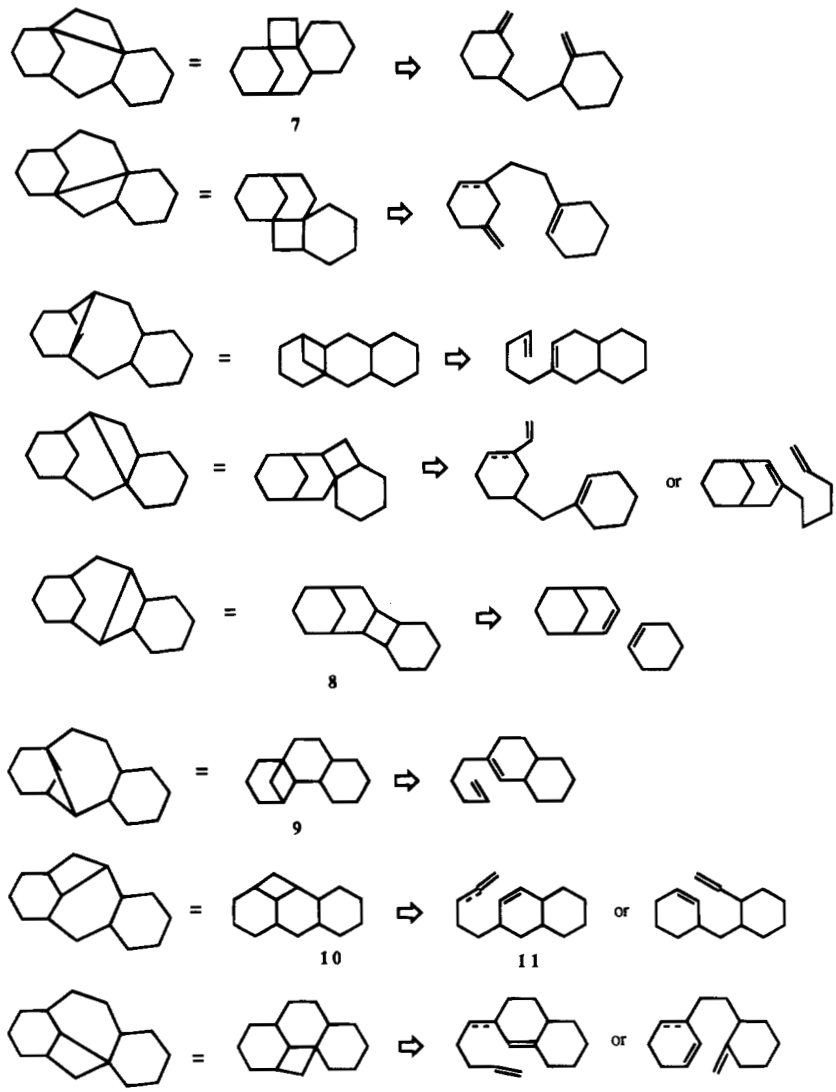
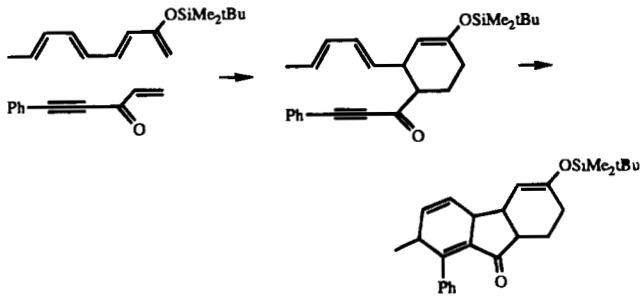
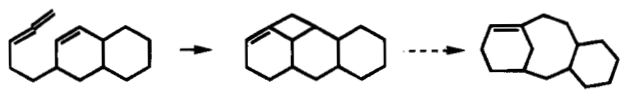


Figure 7. Strategies involving a [2 + 2] cycloaddition.

Scheme 1



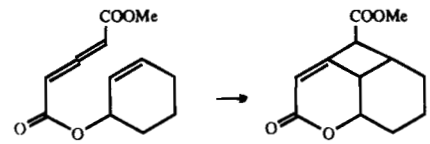
Scheme 2



possibility to build simultaneously the four rings of the taxol skeleton (structure **6**).

Figures 7−9 display solutions which add one bond in the retrosynthetic direction. By this strategy the precursors possess one more ring and are, therefore, more complex. Nevertheless this strategy may be interesting if the created precursors present an easy to build substructure. In the synthetic direction this approach corresponds to fragmentation reactions which provide interesting strategies for complex molecules.[19,20] In Figures 7 and 8 the dotted bonds indicate that, in this position, a double bond could already be present in order to form indirectly the bridgehead double bond of taxol.

Scheme 3



In Figure 7 are gathered the precursors in which the addition of a bond generates a four-membered ring and their own precursors by a [2 + 2] cycloaddition. Strategies involving structures **7** and **8** have been reported.[21,22] The precursor **9** suggests an approach developed by Wender from pinene by another strategy.[23] A strategy involving the sequence **11**−**10** was developed for the saturated framework.[24,25] However precursor **11** suggests an original approach for introducing the bridgehead double bond (Scheme 2).

A similar intramolecular [2 + 2] allene alkene cyclization has been described[26] (Scheme 3).

These intermediates involve six-membered rings, and [4 + 2] or [2 + 2 + 2] precursors are also possible but are not shown.

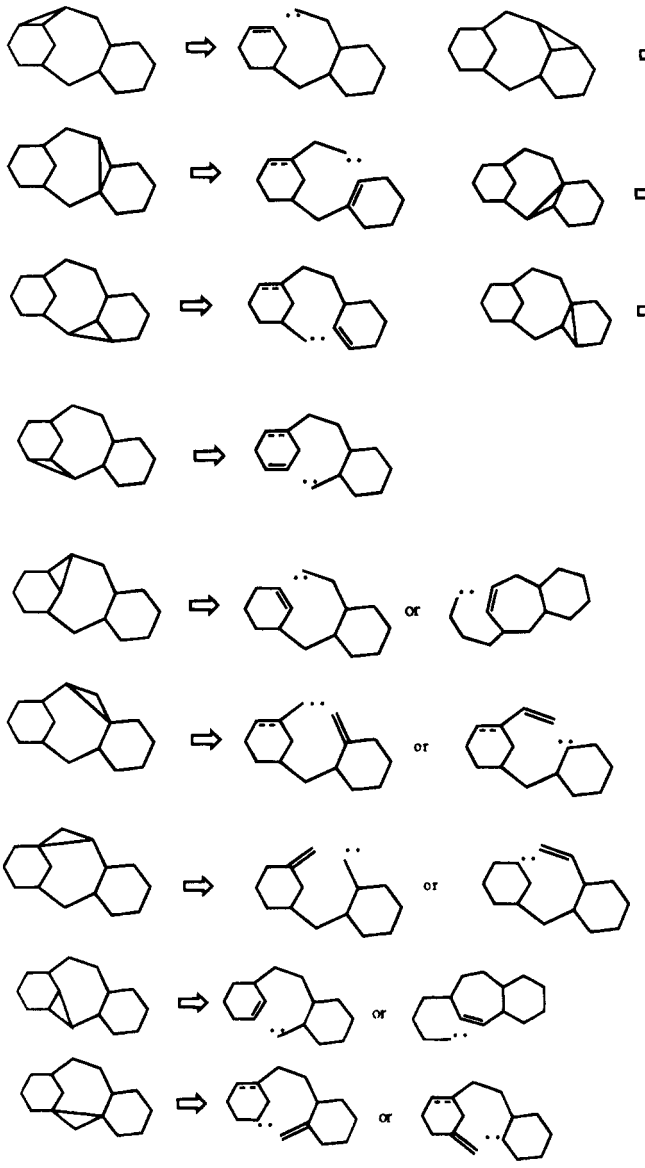


Figure 8. Solutions which suggest internal carbene double bond addition.

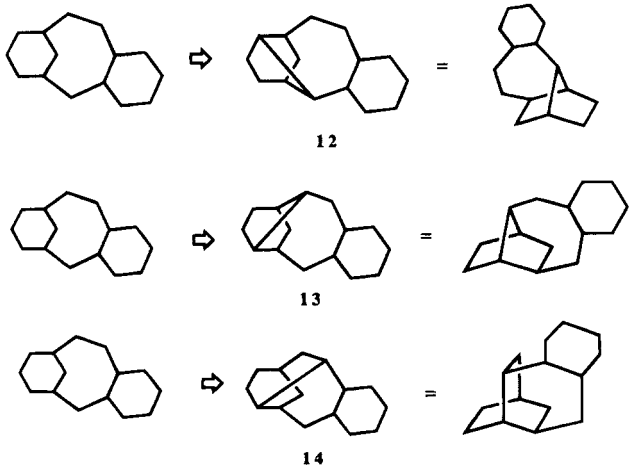


Figure 9. Solutions **12**–**14** present known substructures.

Solutions which also add one bond and in which a three-membered ring is created are gathered in Figure 8. Further analysis of these precursors leads to new precursors involving internal carbene addition on double bond. This strategy (intramolecular cyclopropanation followed by cyclopropyl ring opening) has been utilized in several syntheses.[27,28]

Three more solutions are depicted in Figure 9. Solutions **12**–**14** show attractive possibilities of conversion between known structures (longifolene, patchouli) and taxane. A strategy based on the fragmentation of the framework **14** (without ring C) was devised by Yamada et al. for the synthesis of the AB ring system.[29]

Finally, other solutions are shown in Figure 10. In these solutions two and three bonds were deleted. Some of them are interesting in terms of convergency or symmetry and correspond to known synthesis-like solutions: **15**–**18**,[15] **19**,[30] **20**,[31] and **21**.[32] Solutions **22** and **23** possess a simple skeleton. Solutions **24**–**29** suggest provocative syntheses from linear precursors. Solution **30** suggests a [2 + 2 + 2] cycloaddition.

CONCLUSION

The new program CONAN allows a rough two-step retroanalysis of the skeleton of a target. But even with a schematic approach, CONAN is able to suggest interesting tandem strategies. In the case of taxane CONAN was able not only to find known strategies but also to propose new ones.

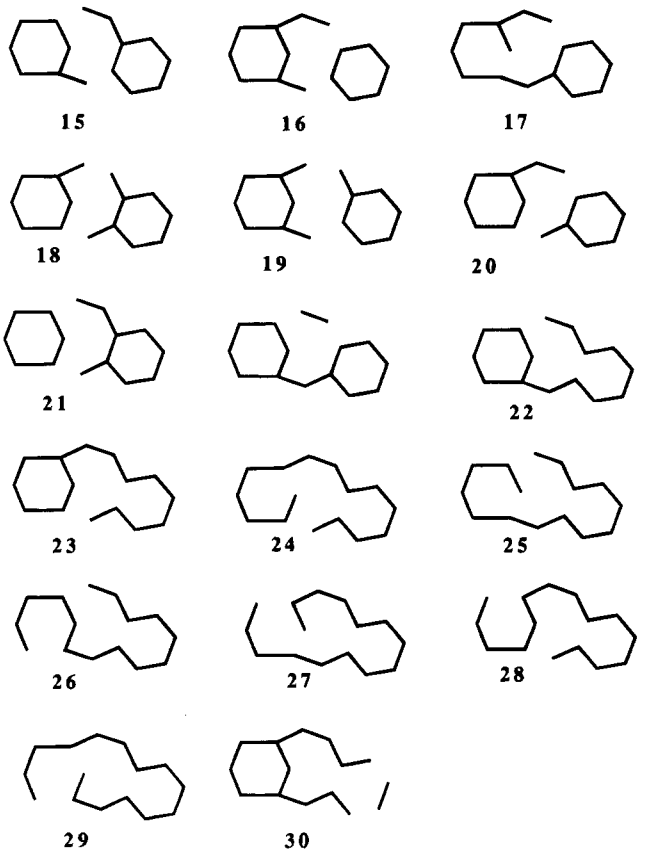


Figure 10. Other solutions involving the formation of two and three bonds.

CONAN is written in Fortran and runs on an IBM/PC and an IBM 3090 main frame computer. A Windows version is under development.

ACKNOWLEDGMENT

We thank Prof. P. Wender for stimulating discussions related to this subject.

REFERENCES AND NOTES

(1) Barone, R.; Chanon, M. *Comput. Chem.* **1979**, *3*, 83.
(2) Barone, R.; Chanon, M. *Heterocycles* **1981**, *16*, 1357.
(3) Barone, R.; Chanon, M. *Chimia* **1986**, *40(12)*, 436.
(4) Corey, E. J.; Ohno, M.; Mitra, R. B.; Vatakencherry, P. A. *J. Am. Chem. Soc.* **1964**, *86*, 478.
(5) Differding, E.; Ghosez, L. *Tetrahedron Lett.* **1985**, *26*, 1647.
(6) Wender, P. A.; Miller, B. L. *Organic Synthesis Theory and its Applications*; Hudlicky, T., Ed.; J.A.I. Press: 1993; Vol. 2, pp 27–66.
(7) (a) Mehta, G.; Barone, R.; Azario, P.; Barberis, F.; Arbelot, M.; Chanon, M. *Tetrahedron* **1992**, *48(41)*, 8953–8962. (b) Azario, P.; Barone, R.; Chanon, M. *J. Org. Chem.* **1988**, *53*, 720.
(8) Barone, R.; Barberis, F.; Arbelot, M.; Baldy, A.; Chanon, M. *Polish J. Chem.* **1994**, *68*, 2191–2198.
(9) Hendrickson, J. B.; Toczko, A. G. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 137.
(10) Moreau, G. *Nouv. J. Chim.* **1978**, *2(2)*, 187.
(11) Ugi, I.; Bauer, J.; Brandt, J.; Friedrich, J.; Gasteiger, J.; Jochum, C.; Schubert, W. *Angew. Chem., Int. Ed. Engl.* **1979**, *18*, 111.
(12) Barone, R.; Chanon, M. In *Computer Aids in Chemistry*; Vernin, G., Chanon, M., Eds.; Horwood: Chichester, 1986; p 69.
(13) Hendrickson, J. B.; Huang, P. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 145.
(14) Tietze, L. F.; Beifuss, U. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 131.
(15) (a) Nicolaou, K. C. *Angew. Chem., Int. Ed. Engl.* **1994**, *33(1)*, 15. (b) Wessjohann, L. *Angew. Chem., Int. Ed. Engl.* **1994**, *33(9)*, 959.
(16) (a) Shea, K. J.; Wise, S. *J. Am. Chem. Soc.* **1978**, *100*, 6519. (b) Brown, P. A.; Jenkins, P. R.; Fawcette, J.; Russel, D. R. *J. Chem. Soc., Chem. Commun.* **1984**, 253.
(17) (a) Lu, Y. F.; Fallis, A. G. *Tetrahedron Lett.* **1993**, *34(21)*, 3367. (b) Sakan, K.; Craven, B. M. *J. Am. Chem. Soc.* **1983**, *105*, 3732–3734.
(18) Krauss, G. A.; Taschner, M. J. *J. Am. Chem. Soc.* **1980**, *102*, 1974.
(19) Wender, P. A.; Manly, C. J. *J. Am. Chem. Soc.* **1990**, *112*, 8579 and references cited.
(20) Hendrickson, J. B. *J. Am. Chem. Soc.* **1986**, *108*, 6748.
(21) Kraus, G. A.; Thomas, P. J.; Hon, Y.-S. *J. Chem. Soc., Chem. Commun.* **1987**, 1849–1850.
(22) Blechert, S.; Muller, R.; Beitzel, M. *Tetrahedron* **1992**, *48*, 6953–6964.
(23) Wender, P. A.; Mucciaro, T. P. *J. Am. Chem. Soc.* **1992**, *114*, 5878.
(24) Begley, M. J.; Mellor, M.; Pattenden, G. *J. Chem. Soc., Perkin Trans. 1* **1983**, 1905–1912.
(25) Winkler, J. D.; Subrahmanyam, D.; Hsung, R. P. *Tetrahedron* **1992**, *48*, 7049–7056; **1993**, *49*, 291.
(26) Yoshida, M.; Hiromatsu, M.; Kanematsu, K. *J. Chem. Soc., Chem. Commun.* **1986**, 1168.
(27) Wong, H. N. C.; Hon, M. Y.; Tse, C. W.; Yip, Y. C.; Tanko, J.; Hudlicky, T. *Chem. Rev.* **1989**, *89*, 165–198.
(28) Padwa, A.; Krumpe, K. E. *Tetrahedron* **1992**, *48(26)*, 5385–5453.
(29) Nagaoka, H.; Ohsawa, K.; Takata, T.; Yamada, Y. *Tetrahedron Lett.* **1984**, *25*, 5389–5392.
(30) Sato, M.; Morihira, K.; Horiguchi, Y.; Kuwajima, I. *J. Org. Chem.* **1994**, *59*, 3165–3174.
(31) Swindell, C. S.; Chander, M. C.; Heerding, J. M.; Klimko, P. G.; Rahman, L. T.; Venkat Raman, J.; Venkataraman, H. *Tetrahedron Lett.* **1993**, *34*, 7005–7008.
(32) Swindell, C. S.; Fan, W.; Klimko, P. G. *Tetrahedron Lett.* **1994**, *35*, 4959–4962.

CI940323C