CONTOUR, a General Contour-Plotting Program for IBM-Compatible Microcomputers, and Its Application to Peptides

RICHARD B. DAVIDSON and S. SCOTT ZIMMERMAN*

Department of Chemistry and Graduate Section of Biochemistry, Brigham Young University, Provo, Utah 84602

Received December 5, 1989

A computer program, CONTOUR, was developed for displaying contour plots on an IBM-compatible microcomputer. The program interpolates between abscissa and ordinate data values to display lines of constant elevation on a rectangular grid. The contour plot can be printed on a dot-matrix or laser printer using the MS-DOS Print Screen feature or using a graphics "capture" program. CONTOUR was used to create conformational energy contour plots of *N*-formylalanine, *N*-acetylalanine, and *N*-acetyl-*N'*-methylalaninamide.

INTRODUCTION

Computer programs that generate contour plots have been available for several years. MMAP,[1] written in FORTRAN for the IBM 370/158, generates plotting instructions for a CALCOMP plotter. MMAP requires elevation data at the intersections of grid lines that form a parallelogram tiling of the plotting space. Since the program does not require a rectangular grid, MMAP can display, for example, contours of continuous functions defined on the face of unit cells of crystals.

Another contour-plotting program, written by A. Preusser[2] in ANSI standard FORTRAN X 3.9-1978 for the Control Data Cyber 175 computer, features the ability to produce contours from elevation data at irregularly spaced points. The plots are displayed on a CALCOMP plotter. The literature also includes other plotting programs, for example, by Bergman[3] and by Thalmann.[4]

None of these published programs, however, include the following features:

(1) ability to run on an IBM-compatible microcomputer;

(2) a user-friendly, menu-driven interface; and

(3) ability to display contour plots on a graphics monitor and to send the output to a dot-matrix or laser printer.

We have developed CONTOUR to fill these requirements. CONTOUR runs on any computer that is 100% compatible with IBM microcomputers using MS- or PC-DOS; provides an easy-to-use, menu-driven, interactive interface; displays output on all the popular graphics displays (CGA, EGA, VGA, Hercules, etc.); and allows printing of the contour plots on a dot-matrix or laser printer.

We have used CONTOUR to analyze conformational energies of small peptides as part of an ongoing research effort to understand the relationship between structure and function in these molecules.[5,6]

DESCRIPTION OF CONTOUR

CONTOUR was written in the C programming language, using Borland International's Turbo C, and was designed to run under MS-DOS version 2.0 or later on an IBM-compatible microcomputer equipped with a video graphics display adapter. CONTOUR consists of 2200 lines of code and comments, requires 112 kB (kilobytes) of memory, and supports most of the popular graphic displays, including the IBM color graphics adapter (CGA), multicolor graphics array (MCGA), enhanced graphics adapter (EGA), video graphics array (VGA), Hercules graphics adapter, AT&T 400-line graphics adapter, 3270 PC graphics adapter, and IBM 8514 graphics adapters.

CONTOUR employs pull-down menus and extensive error checking for ease of use. The program can generate and print a contour plot on a typical dot-matrix printer in less than 4 min.

Data Input. The elevation data must be presented in FORTRAN E or F format, one value per line beginning with the value associated with the least abscissa and least ordinate point and continuing from there sequentially with the abscissa values changing most rapidly. Only values of the dependent variable are entered. The program infers the values of the independent variables from the order of the input data and from information describing the plotting space that the user provides in the menus. The grid is assumed to be rectangular.

Algorithm. The interpolation algorithm is based on the four-point bivariate formula given by Abramowitz and Stegun.[7] This formula gives a hyperbola or hyperbolae for the curve(s) of intersection of a plane of constant value with the interpolated surface. However, interpolated points interior to a rectangle, R, formed by pairs of adjacent values of the independent variables are not used to make the display. Rather, the points of intersection of the hyperbolae with the boundaries of R are joined by straight lines, and these lines make up the contours that are displayed. The visual results, in the case of the peptide conformational energy plots, are, nevertheless, satisfactory when the data are presented at 10° intervals in each coordinate. The interpolation subroutine has been documented within the source code listing in case users wish to modify the algorithm. The interpolation routine returns a code number indicating the type of curve (straight line or hyperbola) resulting from the intersection (if any) of the interpolated surface and the constant value plane within R, together with the coefficients of the equation of the curve. The calling program can then use this information to generate as many points on the curve as desired.

Execution Time. The time required to generate a contour plot on the video screen depends upon the type of graphics (low-resolution CGA is faster than high-resolution EGA), processor speed (a 4.77-MHz IBM PC is obviously slower than a 10-MHz IBM AT), and the presence or absence of a math coprocessor. For example, the time required to generate the plot shown in Figure 2 is 14 min 35 s on a computer with CGA graphics and an 8088 microprocessor running at 4.77 MHz without a math coprocessor, but the same plot requires only 62 s on the same type of computer system *with* a math coprocessor. The time is 2 min 12 s on an 80286 EGA machine running at 10 MHz without a math coprocessor; the time is only 22 s on an 80286 VGA machine *with* a math coprocessor. The program contains code to automatically support an 80*x*87 math coprocessor if one is available.

Printing Contour Plots. Because of the wide range of printers available for MS-DOS machines, CONTOUR does not

* To whom correspondence should be addressed.

0095-2338/90/1630-0174$02.50/0

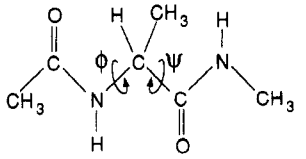

Figure 1. Structure of *N*-acetyl-*N'*-methylalaninamide, showing the dihedral angles ϕ and ψ, which were varied at intervals of 10° to determine the conformation energies for the contour plots.

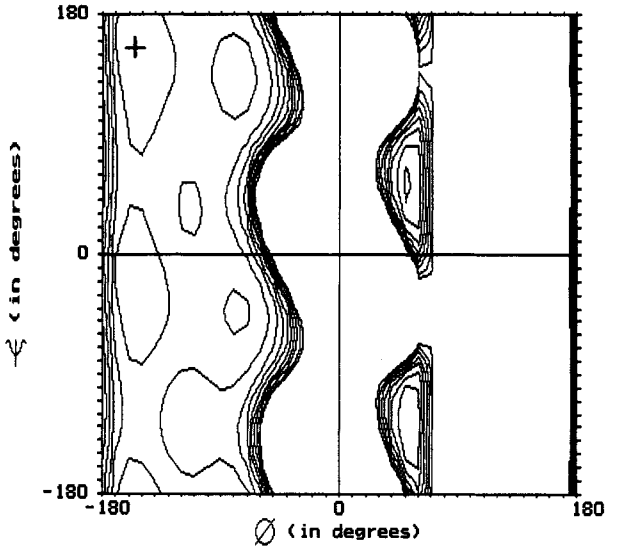

Figure 2. Conformational energy contour plot of *N*-formylalanine generated with CONTOUR, showing lines of constant conformational energy (in kilocalories per mole) above the global minimum. The global minimum is at $(\phi, \psi) = (-155.341°, 156.646°)$, marked by the +, having a total conformational energy of $E_0 = -3.665$ kcal/mol. The dihedral angle χ of the side-chain methyl group was held fixed at 58.843°. The graphics image of the plot was converted to a printable file with GRAB, a program published by WordPerfect Corp., and printed from WordPerfect on a laser printer.

contain printer drivers. Printing a contour plot can be effected, however, by using the MS-DOS Print Screen feature and GRAPHICS program, as described in the DOS user's manual. Alternatively, a graphics "capture" program, many of which are available commercially (e.g., the GRAB program that comes with WordPerfect version 5.0 and 5.1), can be used to capture the graphics screen image of CONTOUR, save the image in a file, and print the image to a dot-matrix or laser printer.

In some instances the quality of the printed contour plots may not be of publishable quality. The main purpose of CONTOUR, however, is to produce contour plots accurately and quickly on readily available computer equipment. The printed plots may require touch-up or retracing.

Availability. The CONTOUR executable code, the Borland International graphics drivers, the source code, documentation, and sample data are available through the Quantum Chemistry Program Exchange (QCPE), Department of Chemistry, Indiana University, Bloomington, IN 47405. Write QCPE for the catalog number, cost, and ordering procedure.

ENERGY CALCULATIONS

Calculations of the conformational energies of the small, blocked peptides were carried out by using PepCAD, a computer program described by Feller and Zimmerman.[6] We used the following procedure to create the contour maps:

(1) The conformation of all methyl groups was allowed to vary to minimize the energy at each point of a 37 × 37 grid covering ϕ–ψ space at 10° intervals in each dimension. All other dihedral angles were kept fixed during minimization. See Figure 1 for a diagram of *N*-acetyl-*N'*-methylalaninamide showing the ϕ–ψ dihedral angles. Although it may seem unnecessary to calculate energy values for ϕ, ψ at both −180° and +180°, CONTOUR requires these duplicate data because it is designed for general contour plotting and has no way of knowing that the data are duplicated.

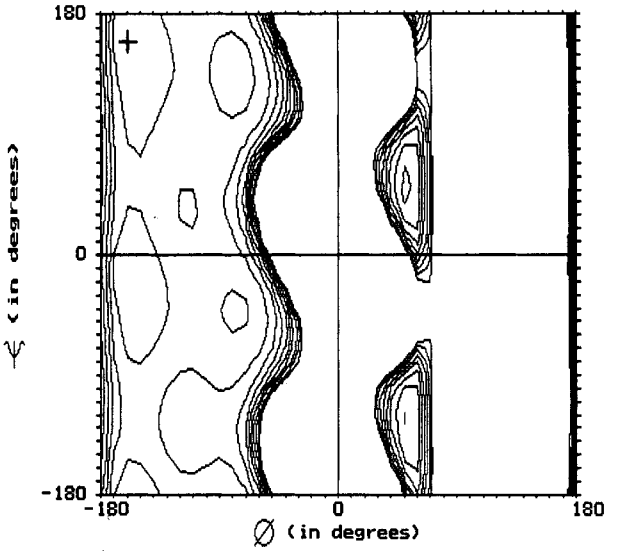

Figure 3. Conformational energy contour plot of *N*-acetylalanine. The global minimum is at $(\phi, \psi) = (-155.890°, 155.370°)$, marked by the +, having a total energy of −7.3979 kcal/mol. The dihedral angle of the acetyl methyl and alanine side-chain methyl groups were held fixed at 55.766° and 58.843°, respectively.

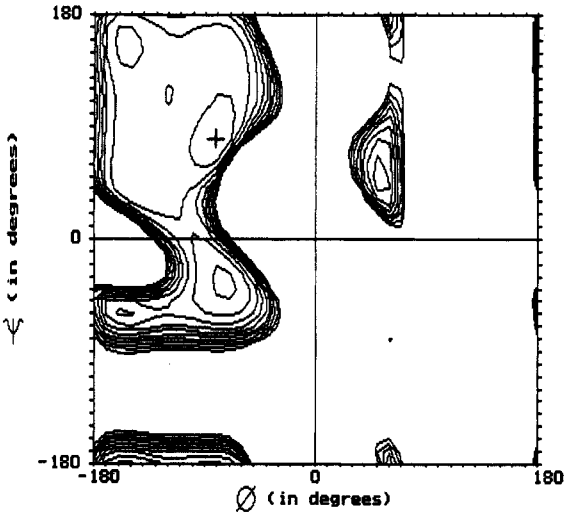

Figure 4. Conformational energy contour plot of *N*-acetyl-*N'*-methylalaninamide. The global minimum is at $(\phi, \psi) = (-79.990°, 75.670°)$, marked with a +, having a total energy of −5.108 kcal/mol. The dihedral angles of the acetyl methyl, the alanine side-chain methyl, and the amide methyl groups were held fixed at 62.640°, 60.920°, and 59.920°, respectively.

(2) At each point where a relative minimum energy occurred that was within 1 kcal/mol of the lowest minimum, the methyl group dihedral angles and the ϕ–ψ angles were allowed to vary during minimization of the total conformational energy. The resulting conformation of lowest energy was then considered the *global* minimum (the lowest point in the conformational space of the peptide).

(3) Finally, PepCAD was executed over the 37 × 37 point grid a second time with the configuration of the methyl group(s) fixed in the conformation of the global minimum found in step 2. Contour plots of the resulting conformational energies are shown in Figures 2–4.

We used CONTOUR to generate conformational energy contour plots of three simple peptides: *N*-formylalanine, *N*-acetylalanine, and *N*-acetyl-*N'*-methylalaninamide, the

latter of which is shown in Figure 1.

RESULTS AND DISCUSSION

A contour plot of the configurational energy levels in *N*-acetyl-*N'*-methylalaninamide has been previously published.[8,9] Since that plot was made, the parameters used in PepCAD have been modified to give results more closely approximating empirically observed energies.[10,11]

In the earlier work,[8,9] the dihedral angle of the side-chain methyl group was fixed at 60°. For the plot in Figure 4, the angles of the methyl groups were as follows: acetyl blocking group, 62.64°; alanine side chain, 60.92°; and *N'*-methyl group, 59.92°. The present results closely resemble the earlier ones except that Figure 4 shows a lower energy barrier by about 1 kcal/mol at the saddle point located at $\phi, \psi = -100°, 0°$.

For both Figures 2 and 3, the dihedral angle of the side-chain methyl group in the minimum-energy conformation was found to be within 0.002° of 58.845°, the value found for the global minimum conformation.

The methyl group on the acetyl blocking group has been shown to contribute little to the conformational energy properties.[5] The similarity of the contour plot of *N*-formylalanine and *N*-acetylalanine (Figures 2 and 3) also shows that the methyl group of the acetyl blocking group has little effect on the conformational properties of peptides with a free carboxyl group.

The contour plot of these molecules shows that about half the ϕ–ψ conformational space is "forbidden", having a conformational energy greater than 11 kcal/mol above the global minimum. One of these forbidden regions exhibits linear boundaries between $\phi = 71°$ and $\phi = 175°$. These boundaries are caused by the steric repulsions between the carbonyl oxygen of the N-terminals blocking group and the side-chain methyl group of alanine. The barriers are linear (independent of ψ) because the distance between the oxygen and the methyl group depends only on ϕ, not on ψ.

The other forbidden region has boundaries that vary roughly sinusoidally as a function of ψ, with a period of about 180°. One boundary oscillates about the line with $\phi = -43°$, and the other oscillates about the line with $\phi = +43°$.

The positions of the convex and concave features of the energy plot exhibit approximately the same periodicity in ψ as do the sinusoidal boundaries. This periodicity conforms to the varying distance between the oxygen atoms of the carboxyl terminus and the carbonyl oxygen atom of the amino-terminal formyl (or acetyl) blocking group. These oxygen atoms carry negative partial charges and therefore repel each other. The periodicity of the energy levels manifested in the contour diagram is consistent with the electrostatic repulsion that would be expected as the oxygen atoms approach and recede from each other while ψ is varied. The small irregularities observed in the periodicity are expected since the carboxyl oxygen atoms are not identical, one of them bearing a proton.

The most significant region of low energy for the residues with a free carboxyl group lies between $\phi = -180°$ and $\phi = -43°$ along the left side of the contour plot (see Figure 2). The global minimum occurs in this region, near $\phi, \psi = -155°, 157°$ for *N*-formylalanine (Figure 2) and $\phi, \psi = -156, +155$ for *N*-acetylalanine (Figure 3). An additional band of low energy exists between the sinusoidal boundary at $\phi = +43°$ and the linear boundary at $\phi = 72°$, with two local minima in this region, one near $\phi, \psi = 50°, -125°$ and one near $\phi, \psi = 50°, 55°$. These conformations have low probabilities of occurrence, however, because they are found at relatively high conformational energies (3 kcal/mol above the global minimum) and in relatively narrow energy wells (i.e., have low entropies).

In the contour plots of *N*-formylalanine (Figure 3), the sinusoidal boundaries at $\phi = -43°$ and $\phi = +43°$ are at the bases of two high-energy peaks having elliptical cross sections centered at $\phi, \psi = 0°, 0°$ and $\phi, \psi = 0°, 180°$. Saddle points with energies between 20 and 30 kcal/mol separate the two bands of relatively low energy.

A comparison of Figures 3 and 4 shows the profound effect of an *N'*-methylamide group on the conformational properties of an alanine single residue relative to the peptide with a free carboxyl group. With an *N'*-methyl group present, an additional forbidden region is superimposed on and cuts across the entire plot at right angles to the two forbidden regions that existed before the substitution (compare Figures 2 and 3 with Figure 4). The 11 kcal/mol boundaries of this new forbidden region lie approximately between $\psi = -150°$ and $\psi = -100°$. Another elliptical forbidden zone appears centered at $\phi, \psi = -180°, 0°$. These additional forbidden zones destroy the periodicity evident in the plots of the N-blocked residues having a free carboxyl group. The low energy surface features in the vicinity of the global minimum are altered as well, and the position of the global minimum is shifted by 70–80° in both ϕ and ψ. We conclude that the appearance of the new forbidden zones is a result of steric repulsions between the methyl group of the alanine side chain and the methyl group of the *N*-methyl blocking group, because when free rotation of these methyl groups is allowed at each ϕ, ψ point of the calculation, the high-energy barrier is markedly lowered (data not shown).

The contour plots (Figures 2–4) and the conformational energies determined by PepCAD are based upon a model in which bond lengths and bond angles remain fixed and only the ϕ, ψ dihedral angles are allowed to vary. If the bond lengths and bond angles were allowed to vary during energy calculations, the general nature of the contour plots would probably not change but the specific values of the relative energy may change to some degree. In particular, the barriers between lower energy minima would probably be lower if full energy relaxation were allowed. Previous results suggest that the approximation of fixed geometries yields reasonable results.[9]

CONCLUSION

The computer program CONTOUR has proven to be a valuable tool for examining energy profiles of peptide molecules. Because of its generality, CONTOUR should also prove useful to researchers in other disciplines.

Registry No. *N*-Formylalanine, 10512-86-4; *N*-acetylalanine, 97-69-8; *N*-acetyl-*N'*-methylalaninamide, 19701-83-8.

REFERENCES

(1) Faber, D. H.; Rutten-Keulemans, E. W. M.; Altona, C. Computer Plotting of Contour Maps: An Improved Method. *Comput. Chem.* **1979**, *3*, 51.
(2) Preusser, A. ALGORITHM 626 TRICP: A Contour Plot Program for Triangular Meshes. *ACM Trans. Math. Software* **1984**, *10*, 473.
(3) Bergmann, H. W.; Kern, R. J. A New Computer Program for Contour Maps of Texture Functions. *Appl. Crystallogr.* **1977**, *9*, 254.
(4) Thalmann, N.; Weber, J. A New Computer Program for Generating Three-Dimensional Plots of Electronic Densities and Related Contour Levels. *Chimia* **1977**, *31*, 361.
(5) Feller, D. C.; Delmoe, E. F.; Zimmerman, S. S. Conformational Energy Analysis of Peptides Using Microcomputers: Description of PepCAD and Analysis of *N*-Formyl-*N'*-Methylalanineamide. *Comput. Chem.* **1989**, *13*, 337. PepCAD is available through QCPE, Department of Chemistry, Indiana University, Bloomington, IN 47405.
(6) Feller, D. C.; Zimmerman, S. S. Conformational Energy Analysis of the Chemotactic Tripeptide Formyl-Met-Leu-Phe and Three Analogs. *Int. J. Pept. Protein Res.* **1989**, *34*, 229.
(7) Abramowitz, M.; Stegun, I. A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*; National Bureau of Standards Applied Mathematics Series 55; U.S. Government Printing Office: Washington, DC, 1965; p 882.
(8) Zimmerman, S. S.; Scheraga, H. A. Local Interactions in Bends of Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 4126.
(9) Zimmerman, S. S. Theoretical Methods in the Analysis of Peptide Conformation. In *The Peptides: Analysis, Synthesis, Biology*; Hruby, V. J., Ed.; Academic Press: New York, 1985; Vol. 7, p 165.

(10) Vásquez, M.; Némethy, G.; Scheraga, H. A. Computed Conformational States of the 20 Naturally Occurring Amino Acid Residues and of the Prototype Residue α-Aminobutyric Acid. *Macromolecules* **1977**, *10*, 1.

(11) Sippl, M. J.; Némethy, G.; Scheraga, H. A. Intermolecular Potentials from Crystal Data. 6. Determination of Empirical Potentials for O-H···O=C Hydrogen Bonds from Packing Configurations. *J. Phys. Chem.* **1984**, *88*, 6231.

Perspectives and Criteria for Chemical Information Instruction

ARLEEN N. SOMERVILLE

University of Rochester, Rochester, New York 14627

Received December 27, 1989

A review of Committee on Professional Training (CPT) guidelines since 1960 shows increasing emphasis on adequate library resources and chemical information instruction. Results of a survey are discussed and show that chemistry departments face major obstacles in implementing information instruction. Suggestions are offered for meeting these difficulties that challenge faculty, administrators, librarians, and CPT to meet the CPT's information and library guidelines.

INTRODUCTION

The American Chemical Society (ACS) established the Committee on Professional Training (CPT) in 1936 to help improve the quality of chemical education. Major activities of CPT include development of guidelines to ensure high-quality undergraduate programs for students planning careers in chemistry and evaluation of programs designed to meet these guidelines. The 50th anniversary of CPT's establishment was celebrated with a symposium, held at the April 1986 ACS National Meeting, that featured speakers who reviewed the past, present, and future of CPT from overall and specific topic perspectives. Chemical information was one of the specific topics represented. A 23-page booklet briefly summarizes all talks presented.[1] This paper is an expansion of the talk presented at the Symposium and updates the information to the present.

Chemical information instruction should be a vital part of every chemistry student's education. Yet, many faculty and librarians see a need to improve the extent and quality of instruction. A 1985 review of past activities aimed at providing information instruction shows considerable interest but sizeable gaps.[2] This paper explores in detail difficulties faced by chemistry departments based on an extensive survey and suggests how faculty, administrators, librarians, and CPT can improve compliance with CPT guidelines.

CPT GUIDELINES

An examination of the Committee on Professional Training guidelines over the past 25 years shows a gradually increasing emphasis on library resources and chemical information. Only a few sentences refer to library facilities in 1960 and 1962.[3] By 1965, the guidelines had become somewhat more specific and stated that "the library should carry *Beilstein* and *Chemical Abstracts*" and major journals.[4] In 1972 and 1977, the guidelines declared that "if library holdings do not include *Beilstein* and *Chemical Abstracts*, particularly, the Committee will seek concrete evidence"[5,6] that "the students learn to use these important references".[6] The alternative of searching computerized files instead of printed *Chemical Abstracts* was first proposed in 1977.

The 1983 guidelines include the most significant advances in support of chemical information.[7] Subscriptions to *Chemical Abstracts* would be expected with at least volume indexes, except in rare instances. In those cases the Committee required evidence that students have ready access to *Chemical Abstracts* in neighboring institutions or industrial libraries. The core curriculum should include a systematic use of the chemical literature. A new section discussed chemical literature and information retrieval and, for the first time, called for formal instruction "imparted either through a separate course or integrated into courses". It contended that adequate teaching of the use of the major sources such as *Chemical Abstracts*, *Beilstein*, *Gmelin*, and *Science Citation Index* would generally require formal lectures. Development of computer searching skills for bibliographic and numeric data bases was considered "highly desirable", including online practice.

The Appendix[8] to the 1983 Committee on Professional Training guidelines appeared several months later and includes a chemical retrieval section that lists the competencies that students should master during their undergraduate years and suggests ways in which the library skills may be taught. It specifies that students should acquire a demonstrable, basic understanding of the following: *Chemical Abstracts*; other major secondary works such as *Science Citation Index*, *Beilstein*, *Index Chemicus*, and *Current Contents*; standard reference works such as *Mellor*, *Gmelin*, and *Landolt-Boernstein*; and primary literature sources. To use the primary and secondary literature effectively, students should be familiar with the organization of the chemistry library and with techniques of manual and online literature searching. Specifically, students should be able to

- locate chemical and physical properties of substances, including spectra;
- locate references for the synthesis or reactions of substances or classes of substances;
- locate references to a desired type of chemical transformation;
- identify the Chemical Abstracts Service Registry Number for substances;
- complete a comprehensive subject search;
- compile a complete bibliography of an author's publications;
- locate review papers on a subject;
- know the importance of patents and be able to search for patents on a subject;
- know about the availability and contents of relevant computer databases.

Such proficiencies should be taught through formal instruction, either in a chemical information course or integrated into other chemistry courses. Coordination of this teaching effort is essential and can be achieved by designating a faculty member, a librarian, or a professor–librarian team as coordinator(s). Other suggested ways to aid information instruction included the following: sources that provide curricular materials; a list of recommended books and journals;[9] reminder of Chemical

0095-2338/90/1630-0177$02.50/0