#2612 THE INSTITUTE OF PAPER CHEMISTRY
(Metal-Carbohydrates)
Project Reports (2)

Institute of Paper Science and Technology
Central Files

PROJECT REPORT FORM

Copies to: ✓Central Records
John W. Swanson
R. M. Leekley
Gerald F. Richards
Dale G. Williams
Editorial Department

PROJECT NO. 2612
COOPERATOR Institute of Paper Chemistr
REPORT NO. 3
DATE March 12, 1971
NOTE BOOK --
PAGE TO
SIGNED Gerald F. Richards
Gerald F. Richards

STRUCTURE STUDIES OF METAL-CARBOHYDRATE COMPLEXES

CRYSTALLOGRAPHIC COMPUTER PROGRAMS FOR THE IBM 360-44

SUMMARY

Project 2612 is concerned with the study of metal-carbohydrate complexes and the general development of x-ray diffraction as a research technique for staff and students at The Institute of Paper Chemistry. Because of the extensive calculations involved in single crystal x-ray diffraction research, an adequate computing facility is necessary for the proper use of this technique. The IBM 360-44 computer provides more than adequate computing for crystallography research and its "hands on" availability makes the computational aspects of crystallography much easier at IPC than at most other universities. This report is concerned with crystallographic programming systems which have been developed for the IBM 360-44 computer.

The programs described in this report have been obtained in two ways: programs written by other crystallographers willing to share their programs with outside users, and programs written at IPC by staff and students. The programs have been tested by utilizing diffraction data from known structures, primarily the data from α-D-xylose·$CaCl_2$·$3H_2O$(Project 2612, Project Report 2, September 3, 1968).

Summarized below are some of the program systems which are presently available for use on the IBM 360-44 computer:

1. NRC Crystallographic Programs for the IBM 360 System. This is a group of fourteen interrelated programs for data reduction, structure solution, and structure refinement which were obtained from crystallographers at the National Research Council, Ottawa, Canada.

2. Direct Acentric Phasing Programs. A series of programs obtained from various sources which have been tested and are presently in use. These programs provide for structure phasing directly from the intensity data. They are particularly applicable to all light atom, noncentrosymmetric structures.

3. The X-Ray 70 System. This system of programs was obtained from Dr. James M. Stewart, University of Maryland Computer Center, and it is a collection of programs accumulated by Stewart from many authors and integrated into one system for distribution to outside users. This set is not in use at this time.

4. Patterson Search Programs. These programs provide an alternative method for the solution of all light atom crystal structures when the conformation of part of the molecule under consideration is known.

5. Miscellaneous programs. These programs have been written at IPC and, in general, they complement the other programs described in this report. To a large extent, these programs are due to the efforts of two students, John A. Heitman and Robert A. Moran.

A complete self-consistent set of computer programs for the necessary calculations in x-ray crystallography has been prepared and tested on the IBM 360-44 computer.

INTRODUCTION

The primary objective of Project 2612 is the application of x-ray diffraction techniques to the study of metal-carbohydrate complexes. An important secondary objective is the development of the x-ray technique in such a manner that it can be utilized by students in thesis research. Project Report No. 2 summarizes these and other objectives, give an introduction to crystallography and the x-ray diffraction method, discusses the results of the crystal structure determination of α-D-xylose·$CaCl_2$·$3H_2O$, and presents the results of synthesis research involving the preparation of new metal-carbohydrate complexes.

During the month of September, 1968 The Institute of Paper Chemistry replaced the IBM-1620 computer with a new, faster, more versatile IBM Model 360-44. Because of the long, iterative calculations necessary in single crystal diffraction research, the large computer is necessary--especially as more difficult crystal structure problems are attempted which require more data and data analysis, such as the large carbohydrate structures being worked on by the students, Robert Moran and John Heitman. Newer methods of solving crystal structure problems also rely extensively on large computers, perhaps to some extent because computers are more readily available today. It has been necessary to develop an extensive system of programming which will take maximum advantage of the computer's capability, require a minimum of effort by the user, and assure a straightforward path (as straightforward as possible) to the solution of a crystal structure. The development of this programming system is now complete, although additions and improvements will always be necessary. This report is concerned exclusively with this programming system.

In order to test new computer programs, it was necessary to have available data from a known crystal structure. The data used most frequently was that from the known structure of α-D-xylose·$CaCl_2$·$3H_2O$ (Project 2612, Report 2, '68).

The procedure followed was to utilize the data from the α-D-xylose·$CaCl_2$·$3H_2O$ complex, which required refinement, in order to test and place in operation programs which were obtainable from outside sources. An attempt was made, as much as possible, to implement programs prior to the time they were needed by students.

In this report, the introduction and the first part of the experimental section contain a general discussion of computers, computer systems, and the sources of the programs. The last part of the report concerns specific programs, their purpose, where they are stored, code names, etc., and is designed primarily for the users of the programming system.

VARIOUS COMPUTERS

Common in the language of people who work with computers are the terms, hardware and software. Hardware refers to the actual computer and the various physical accessories necessary for the operation of the computer. Software is a term used to describe programs which are written in machine language and which operate in the core of the machine throughout any computer activity. These programs are designed by the manufacturer and are a major part of the overall system. They interpret (compile) the language used by the normal programmer in such a manner that it can be stored and used efficiently in the core of the machine. The software system directs allthe functions of the computer when properly instructed by the user.

There are many computers in operation today which range in price from several thousand to several million dollars. Among the more common computers are the UNIVAC 1107 and 1108; CDC 1604, 3600, 6400, and 6600; XDS sigma 5 and sigma 6; Burroughs 5500, IBM 7040, 7090, 7094, and the IBM system 360, models 40, 44,50, 65, and 75. The difference in price reflects, in general, a difference

in the size of the high-speed core of the machine and the rate at which a given operation can be executed in core. Obviously a person programming for use on a large computer (large core area) will take maximum advantage of the core storage in designing a program; the same calculation on a smaller computer would require greater use of storage devices (tapes, disks) so that a smaller segment of the calculation need be in core storage at any given time.

An IBM 360-44 computer has been installed and is in operation at The Institute of Paper Chemistry. This machine is of intermediate size and has a high calculation speed for the size of its core storage. It was designed by IBM specifically for scientific type calculations, and it is faster than the IBM-50 (about 30% faster based on crystallographic calculations made here and comparable calculations made on a Model 50). The Model 50, however, has features which make it attractive for business and accounting purposes.

LANGUAGE, PROGRAMS, AND SYSTEMS

A computer program is developed for a calculation by describing the calculation in a language which can be interpreted by the computer through the appropriate system (software) available to it. Several languages have developed over the past years, including various versions of FORTRAN and ALGOL. Computer manufacturers have agreed on one well-defined standard language, USA FORTRAN IV (1). However, it is still not possible to interchange programs for the same calculations which are run on machines manufactured by two different companies. It appears that the various computer manufacturers have taken some pains to frustrate the interchangeability of programs and evade the spirit of the agreed standards.

As mentioned previously, the FORTRAN statements which comprise the

program are compiled by the operating system (this system is a program itself stored in core) in a form which is recognizable by the machine. Obviously, the flexibility and power of a program depends on the flexibility and power of the system used to compile it.

The Institute of Paper Chemistry has three different operating systems, all of which can perform certain similar functions--but each of which possesses certain advantages. These systems are a remote control system (RAX), OS-360, and 44-PS. The RAX system allows the computer to be used by a number of users simultaneously and input of data by several modes, which include remote video terminals. This system is so extensive and complex, however, that it occupies a considerable amount of the core area of the machine and thus limits the core storage available for programmed calculations. This is the reason it is not practical for most crystallographic calculations. The OS-360 system is a very flexible system and is the one used on the larger 360 series machines Models 50, 65, 75. It does have certain features which make it attractive for accounting, retrieval, and business office work. The system which was designed particularly for the Model 44 machine is 44-PS. This system uses the least amount of core storage and takes maximum advantage of the machine's capability. It is thus the one being used for most crystallographic calculations.

EXPERIMENTAL

GENERAL APPROACH AND SOURCE OF PROGRAMS

In order to work effectively with the computer for specific calculations, it is necessary to understand quite well both the nature of the calculation and the operation of the computer. It is seldom possible to perform many calculations in a routine manner without encountering some difficulty

which requires intervention and some knowledge on the part of the user; therefore it is advantageous for the user to write his own programs. Several years time could be required to write all the necessary crystallographic computer programs and test them, so this was not practical. It was decided that a sensible course to follow was to acquire the best programs available from other crystallographers, those which were most compatible with the IBM 360-44; test them; make them operational on the IBM 360-44; and then use locally written programs to bridge any gaps which existed in the resulting system. Two students, John Heitman and Robert Moran, are primarily responsble for the locally written programs.

Considerable care was exercised in requesting programs from outside crystallographers. In general, requests were directed only to people who openly solicited requests in journal articles or by listing their programs in the International Union of Crystallographers Monograph (2). This was done because considerable effort and expense is involved in sending the programs, particularly when the program is sent on computer cards or magnetic tape. On the other hand, this type of programming is the major part of some scientist's research, supported by outside grants, and requests for their programs enhance their personal research.

NRC CRYSTALLOGRAPHIC PROGRAMS FOR THE IBM/360 SYSTEM (2)

A group of crystallographers from the Divisions of Pure Physics and Pure Chemistry, National Research Council, Ottawa, Canada (F. R. Ahmed, S. R. Hall, M. E. Pippy, and C. P. Saunderson) have prepared a series of programs in FORTRAN IV for the determination and refinement of crystal structures. These programs were written for an IBM 360-50 computer, but they were designed so that they would work on most IBM 360 systems. Descriptive write-ups,

FORTRAN listings, and FORTRAN source decks, along with some suggested changes, were obtained from Dr. F. R. Ahmed.

The FORTRAN source decks were obtained from Dr. Ahmed on a magnetic tape. From the magnetic tape card decks were prepared for each of the fourteen programs in the NRC system. Changes were made in the programs that were tested as recommended by Dr. Ahmed. The number assignment for the input/output units were changed so that they corresponded to the numbering for the IBM 360-44 units. Several of the programs (see Table I) were compiled and stored in the 44-PS disk library. The programs chosen for testing were those most urgently needed for the crystallographic work underway at The Institute of Paper Chemistry. The data used for testing was the data available from the α-D-xylose$\cdot CaCl_2 \cdot 3H_2O$ structure. Following closely the instructions given in the write-ups, input cards were prepared for each program and the program was tested until it appeared to operate satisfactorily. Programs in the NRC system are tabulated in Table I. These programs are stored in the 44-ps phase library under the name listed in Table I (i.e., NRC-2, etc.). Many of these programs have been subsequently used by the students, and those tested appear to be running correctly.

The programming system, its design, and its many important features are discussed in detail in the write-ups provided by Dr. Ahmed. Many of these write-ups have been duplicated and are available to the users of the system. The general layout of the system is similar to that designed by crystallography groups in Glasgow, Leeds, and Oxford. These are groups which Dr. Ahmed apparently had contact with before working at the National Research Council in

TABLE I

SUMMARY OF NRC SYSTEM

Program No.	Name	Approx. Size in Words	Test Status*
NRC -1	Gonionstat settings	3,296	NT
NRC-2A	"Picker Automatic" data processing	22,483	NT
NRC-2	Data reduction and tape generation	7,848	T
NRC-3	Absorption correction (3-circle goniostat geometry)	8,500	NT
NRC-4	Symbolic addition procedure Centrosymmetric	22,500	NT
NRC-8	Fourier	23,500	T
NRC-9	Differential Synthesis	27,685	NT
NRC-10	Structure-factor least-squares (block diagonal)	22,000	T
NRC-12	Scan of interatomic distances and angles	23,560	T
NRC-14	Error analyses and the agreement summary	3,373	T
NRC-21	The d-spacings	9,536	T
NRC-22	Mean Plane	5,900	T
NRC-23	Structure factor tables	13,800	NT
NRC-24	Projection onto a plane	--	NT

*NT--not tested.

T--tested.

Canada. The programs are intended to be self-consistant; that is, one program can follow another with a minimum amount of effort on the part of the user. In order to facilitate this continuity, standard formats are adopted for the data cards and for the lists stored on magnetic tape. All relevant data and tables for a given structure are stored on a magnetic tape which is assumed to be the input for all the main programs. A brief description of each of the programs follows.

NRC-1, Goniostat Settings. This program applies to an instrument which is not a part of the Institute's x-ray equipment. It has, therefore, not been compiled.

NRC-2A, "Picker Automatic" data processing. This is a data processing program which is also not applicable to the x-ray research at the Institute.

NRC-2, Data reduction and tape generation. At the start of a structure determination, this is the first program to be run. It arranges the data of a given structure into lists of standard formats which are recognized by the other NRC crystallographic programs, and it stores the lists on tape or disk. This data file serves as input to most of the other programs in the series. The program performs other operations such as the calculation of $\sin^2\theta$, interpolation on the scattering-factor curves, derivation of $|Fo|$ from the measured intensities, assignment of weights to the reflections, and the application of a sharpening function to Fo^2.

NRC-3, Absorption correction (3-circle goniostat geometry). This program is not needed at the present time.

NRC-4, Symbolic addition procedure (centrosymmetric). The purpose of this program is to estimate structure factor phases of centrosymmetric structures using symbolic addition methods. Since the structures being worked on here are noncentrosymmetric, this program has not been compiled.

NRC-8, Fourier. The Fourier program can be used to calculate 3-D Patterson, Fourier, and difference syntheses. It is designed such that Fourier summations can be output on maps which are either undistorted or distorted. The distorted map saves some computer time. This program can be run immediately following a structure factor calculation. All three of the above functions have been tested.

NRC-9, Differential Syntheses. The main purpose of this program is to refine the atomic positions by the differential synthesis method. This program complements the least-squares refinement. NRC-9 has only recently been released and it has not been checked.

NRC-10, Structure Factor Least Squares. This program calculates structure factors, refines the positional and thermal parameters, refines the occupation factors and the overall scale factor and estimates the standard deviations of the refined parameters. The refinement is carried out by means of a block-diagonal least-squares approximation.

NRC-12, Scan of interatomic distances and angles. Given a set of atomic coordinates, this program scans the given atoms for the intramolecular distances, derives the equivalent positions and scans for the intermolecular distances, prepares a summary of the coordination around each of the given atoms, calculates the angles between the bonds, and estimates the standard deviations of the calculated distances and angles.

NRC-14, Error Analysis and the Agreement Summary. The output from this program aids the user in assessing errors which may exist in the structure factor data. It also helps in making decisions relative to weighting schemes used in the least squares program (NRC-10).

NRC-21, The d-spacings. The purpose of this program is to generate indices and compute the d-spacings for all the nonequivalent reflections within a selected sphere, or part of a sphere, in reciprocal space.

NRC-22, Mean Plane. This program calculates the mean plane through a group of atoms, examines the planarity of the group by means of the χ^2-test, calculates the distances of any other atoms from the mean plane, and estimates the standard deviations of the calculated distances.

NRC-23, Structure Factor Tables. This program produces a listing of the structure factor table in the right proportions for photographing and publication in Acta Crystallographica.

NRC-24, Projection onto a plane. This is a new program which is an aid in the preparation of drawings for the purpose of illustrating a

structure. It projects the atoms of a molecule onto a given plane for preparation of a clinographic drawing of the structure.

As mentioned previously, the programs were tested using the data from the α-D-xylose$\cdot CaCl_3 \cdot 3H_2O$ structure. Approximately 1300 reflections were used in space group, $P2_1$. Some estimate of computer time requirement was obtained using this data. A three-dimensional Fourier could be summed in less than one-half hour, the exact time depending on the density of grid points. Likewise, a least-squares refinement cycle required about ten minutes. Based on some figures given by Ahmed for the IBM 360-50, it would appear that the Model 44, when used with the 44-ps system, is somewhat faster than the Model 50.

PROGRAMS SUPPLEMENTARY TO THE NRC SYSTEM

The design of the NRC programming system assumes that the data collection is carried out by means of a diffractometer, particularly the program NRC-2. Since data has been collected here at the Institute by means of photographic films, it has been necessary to write a series of programs to prepare the data for input into NRC-2. John Heitman has written a series of short programs for interfilm scaling (3). The following rather extensive programs were written by Bob Moran.

A program which evaluates the cell constants (along with the errors associated with each constant) from back-reflection Weissenberg data (4); a program to apply the Lorentz, polarization, and spot shape corrections to the observed intensity data(5); and programs to place the structure factors on one common scale and in a format for input to NRC-2 (6).

Direct Acentric Phasing Programs

The crystal structures presently being worked on at The Institute of Paper Chemistry are large noncentrosymmetric structures, which are the most difficult types of structures to solve. The best methods currently known for solving this type of structure are the direct-phasing methods introduced by the Karls (7). An attempt has been made to acquire programs which apply this method for the solution of crystal structures. These are summarized below.

E-Value Program

William E. Scott, a former student at The Institute of Paper Chemistry, has written a series of programs for the IBM-360 system for the application of symbolic addition methods to crystal structure determinations. One of these programs, a program to convert relative structure factors to normalized E's, has been compiled and tested. This conversion is made by the K-curve method (8). The program accepts the structure factors output from the L-P and interlayer scaling programs (discussed in previous section) and outputs two card decks, one deck containing the E-values sorted in order of descending magnitude, and the other containing the amplitudes F, Fabs, E, E^2-1, and F^2 with their associated h, k, l, in a format ready for input to NRC-2. A table containing the E-value statistics is output on the printer. A detailed write-up was obtained with the program.

PHASE

The program PHASE and its write-up was obtained from Dr. D. F. Koenig, Brookhaven National Laboratory, Upton, New York. PHASE is a semiautomatic symbolic addition and tangent refinement program. By means of control decks specified by the user, the program can be used to perform sequences of symbolic addition and tangent refinement phasing. The program was originally written by Dr. Koenig for use on a Control Data Corporation 6600 computer. John Heitman altered the program for use on the 360 system (9). These alterations were extensive, since the original version contained several machine-dependent features.

MULTAN

MULTAN is a series of three programs for the solution of noncentrosymmetric crystal structures. This system was provided by Peter Main, Department of Physics, University of York, York, England. The programs perform the following operations.

SIGMA 2 sets up all the phase relationships which satisfy the sigma-2 relationship.

CONVERGE determines the signs of reflections which are structure invariants, fixes the origin and enantiomorph, and assigns phases to a small number of other reflections which appear to lead to good phase determination.

FASTAN develops each set of phases produced by the previous program by means of the tangent formula.

The ideas upon which these programs are based have been published by G. Germain, P. Main, and M. M. Woolfson. Appendix I contains a reprint of this article.

MAGENC

Dr. Chun-Che Tsai wrote this program as part of his Ph. D. thesis work at the University of Indiana (10). The program MAGENC is similar in principle to MULTAN. However, MAGENC makes greater use of the sigma-2 formula in developing multiple phase sets and has more indicators for assessing the correctness of the phase sets. Besides a listing of the source decks, a portion of his thesis describing the program was also obtained from Dr. Tsai.

The programs PHASE, MULTAN, and MATENC have been compiled and stored in the 44-ps phase library. These programs have been used several times and they appear to work satisfactorily.

GAASA I-VI

FORTRAN listings of six programs have been obtained from three Swedish scientists, Ove Lindgren, Oliver Lindquist, and Jens Nyborg. These programs apply the symbolic addition procedure according to the conventional methods described by J. Karle and I. Karle (7). Several reflections are assigned symbols for their phases, and then after symbolic addition is applied, an attempt is made to evaluate the symbols giving the most probably phase possibilities. A more detailed description of this program is given in Appendix II. These programs have not been tested.

NRC-5

Recently, the crystallographic group at the National Research Council of Canada has released a new program, NRC-5, which is compatible with the NRC programs described previously in this report. The program was written by C. R. Huber and F. R. Brisse of the National Research Council, and are based to a large extent on a set written by S. R. Hall for the PDP-6 computer at the University of Western Australia. The NRC-5 program consists of five separate routines:

DP I Calculation of the overall isotropic and anisotropic temperature factors and overall scale factor.

DP II Calculation of normalized structure factors, and sorting of all $|E|$'s over a specified minimum in descending order of $|E|$.

DP III Search for all reflection triplets satisfying the sigma-2 relationship and with $|E|$'s above some specified threshold value.

DP IV Repetitive use of the tangent formula to extend and refine the structure factor phases for all reflections with $|E|$ above some specified minimum.

DP V Program for developing partial structure information.

These programs have been compiled and stored under the code names NRC 51-55. Preliminary tests indicate that these programs are operating satisfactorily.

Least-Squares Analysis of Structure Invariants

A variation of the symbolic addition methods for solving structures has been developed by Hauptman (11). This method requires the calculation of the structure invariants. Charles M. Weeks from the Medical Foundation of

Buffalo, Buffalo, New York has programmed this method for the CDC 6400 computer. These programs were obtained from Dr. Weeks. They have been revised for the IBM 360-44 by John Heitman and Robert Moran. There are four programs in this group which have the following names and functions.

(1) PROGRAM BUILD

Generates vector triples.

(2) PROGRAM TPROD

Computation of the structure invariants corresponding to these triples by means of the triple product relationship.

(3) PROGRAM MDKS

Computation of the structure invariants corresponding to the triples by means of the MDKS relationship .

(4) PROGRAM LSP

Derivation of phases from these invariants by least-squares analysis.

These programs have been compiled, tested, and stored in the 44-ps phase library under the above code names. They appear to be running correctly.

Since the direct acentric phasing methods are a relatively new approach to the solution of noncentrosymmetric structures, an attempt has been made to acquire several of these programs despite the obvious duplication. Even though a program is not used, its write-up often gives helpful hints and ideas for approaching a new problem.

Miscellaneous Programs

In order to coordinate the direct-phasing programs with the NRC programming system, it has been necessary to devise a series of small programs. These programs have been written by John Heitman. Also, some modifications were also made in NRC-8 and NRC-10 in order to facilitate the tangent formula recycle procedure (12). This integrated system was discussed by John Heitman in a progress report. Parts of this report are reproduced in Appendix III.

The X-RAY 70 System

Under the direction and influence of Dr. James M. Stewart from the Computer Science Center of the University of Maryland, a collection of interrelated FORTRAN programs for the structural analysis of crystals from diffraction data has been prepared and made available to outside crystallographers. Each program is written in a neutral subset of FORTRAN IV such that the program can be compiled and run interchangeably on most any modern computer. Many authors have contributed to the system. An author prepares an appropriately coded program related to the system and sends it to Dr. Stewart who arranges for the program to be tested on different computers. After adequate testing, the program is incorporated into the master system. Upon request to Dr. Stewart, along with a fee of $60.00 to cover costs, a copy of the master system can be obtained. The instructions, write-ups, and FORTRAN cards are all sent on a single IBM computer tape. This tape has been obtained from Dr. Stewart and some preliminary work has been done on implementing the system. At this time none of the programs have been run on the IBM 360-44. Table II summarizes the crystallographic programs available in this sytem.

TABLE II

PROGRAMS IN X-RAY 70 SYSTEM

Code Name	Description
GENERL	General description of the system
BLOKLS	Atomic parameter refinement by block-diagonal lease squares
BONDLA	Determination of contact and bond distances and angles with estimated errors
CRYLSQ	General crystallographic least squares program
DATCO3	Treatment of card controlled XRD-6 output
DATCO5	Treatment of diffractometer data
DATFIX	Preliminary data scaling, calculation of quasi-normalized structure factors (E), and estimation of overall temperature factor
DATRDN	Preparation of binary data file and preliminary treatment of data and symmetry
DIFPCH	Preparation of sorted diffractometer setting cards
DIFSET	Generation of settings for gathering data on an automated diffractometer
DUMCOP	Dump or copy the binary data file
FC	Structure factor calculation
FOURR	Fourier transformations to give Patterson, vector, electron density, difference, or E maps
LISTFC	Lists structure factors for publication
LOADAT	Load atomic parameters into the binary data file
LSQPL	Calculation of equations of 3-dimensional planes and lines
MODIFY	Binary data file editing and generation of pseudo data
NORMSF	Preliminary data scaling, calculation of quasi-normalized structure factors (E), and estimation of overall temperature factor
ORFLS	Atomic parameter refinement by full matrix least squares
PARAM	Least squares refinement of cell parameters from two Theta data
PEKPIK	Search Fourier maps for peaks
PHASE	Search of Sigma-two relationships for a set of possible phases
RLIST	Lists R values for various zones and other reflection classes.
SIGMA	Forms triples for phase determination (centric or acentric)
SIGMA2	Generation of Sigma-two relationships
TANGEN	Use of tangent formula to calculate phases
WRITEU	Generation of write-up
WTANAL	Analysis of least squares weighting schemes
WTLSSQ	Revision of least squares weighting schemes

This system is very sophisticated in its methods for controlling the programs in the core of the machine and in its ability to handle data files. It operates under its own FORTRAN subsystem called the "nucleus". The nucleus is a set of programs located in the core of the machine which recognize operation and program calling cards, take any necessary action upon them, and provide general file handling capabilities.

Many programs in this system perform the same functions as those in the NRC system. For this reason it has not been necessary to implement the system entirely.

Patterson Search Program

Patterson methods have traditionally been very effective for the solution of crystal structure which contains heavy atoms in the unit cell (i.e., Cu, Zn, Pt). For structures which contained only light atoms (i.e., O, C, N), these methods have been less effective. This has been due to the similarity in peak height for all peaks obtained in the Patterson calculation. Recently, (due to the availability of high-speed computers) these methods have been effectively used to solve all light atom, noncentrosymmetric structures (13). Good results have been obtained when the conformation of part of the molecule under consideration is known. The procedure is to input the geometry of the known part of the molecule, generate the Patterson vectors associated with it, then search the entire group of Patterson peaks from the structure for vector sets compatible with the input structure fragment. Output from the calculation are sets of coordinates corresponding to possible positions in the unit cell for the fragment and a reliability index for each set. Robert Moran has obtained a program to perform these calculations (14). It is currently being revised for use on the IBM 360-44.

RESULTS AND CONCLUSIONS

A complete self-consistent set of computer programs for the necessary calculations in x-ray crystallography has been prepared and tested on the IBM 360-44 computer. These programs encompass the most recent developments in acentric phasing and Patterson search techniques. Included in the program set are programs for data reduction, structure solution, structure refinement, and programs for manuscript preparation.

At the present time, for the convenience of the current users, the source decks for these programs are located in three rooms--Room 1225, the computer room, and Room 118 in the Krannert Graduate Research Center. Each program has been compiled and stored in the 44-ps phase library under the code name given in this report. A back-up tape containing the 44-ps library and these programs has been prepared by the computer center staff.

LITERATURE CITED

1. USA Standard FORTRAN (USA X3.9-1966). USA Standards Institute, New York, 1966.

2. IU Cr World List of Crystallographic Computer Programs, 1966.

3. Heitman, J. A., The Institute of Paper Chemistry, Progress Report No. 5, February 13, 1970.

4. Moran, R. A., The Institute of Paper Chemistry, Progress Report No. 3, May 16, 1969.

5. Moran, R. A., The Institute of Paper Chemistry, Progress Report No. 6, Feb. 13, 1970.

6. Moran, R. A., The Institute of Paper Chemistry, Progress Report No. 7, June 3, 1970.

7. Karle, J., and Karle, I., Acta Cryst. 21:849(1966).

8. Karle, J., and Hauptman, H., Acta Cryst.9:635(1956).

9. Heitman, J. A., The Institute of Paper Chemistry, Progress Report No. 6, May 8, 1970.

10. Tsai, Ph. D. Thesis, Part III, Indiana University, 1968.

11. Hauptman, H. In "Crystallographic Computing". F. R. Ahmed. Ed., Munksgaard, Copenhagen, 1970, p. 41.

12. Karle, J., Acta Cryst. B24:182 (1968).

13. Hornstra, J., In "Crystallographic Computing". F. R. Ahmed, Ed., Munksgaard, Copenhagen, 1970, p. 103.

14. Braum, P. B., Hornstra, J., and Leenhouts, J. I., Philips Res. Rep. 24:85 (1969).

Acta Cryst. (1970). B26, 274

On the Application of Phase Relationships to Complex Structures
II. Getting a Good Start

BY G. GERMAIN
Laboratoire de Chimie-Physique, Université de Louvain, 39 Schapenstraat, Louvain, Belgium

AND P. MAIN AND M. M. WOOLFSON
Physics Department, University of York, England

(*Received* 17 *September* 1969)

The most important element governing the success or otherwise of direct methods is the choice of the starting point from which new phases are developed. This paper describes three ways in which a good start to the phase determining process can be made: (i) Where, early in the process of determining new phases, a phase relationship of low reliability must be used then the phase determined by it is allowed to have alternative values, one of which should be close to the correct value. An application of one variant of this technique is described. (ii) A method is described for finding groups of reflexions very highly interrelated by phase relationships. It is possible to ensure, by using these reflexions as a starting point for phase development, that multiple indications for new phases are soon encountered. (iii) An algorithm is described for generating a schematic route for phase determination which indicates the best phases to use as a starting point for phase development including those which fix the origin and enantiomorph. A formula is derived for estimating the probable standard deviation of a phase, which would be given by a number of different phase relationships, before any phase information is known.

Introduction

Direct methods of solving crystal structures, by their very nature, might seem to be objective mathematical procedures the success of which should be independent of the skill of the user. Nevertheless there still appears to be a 'green fingers' element in operation and sometimes an unsuccessful application of direct methods to a structure by one individual has been followed by a successful application by someone else.

Perhaps the most important element influencing success or failure is the choice of starting point from which is begun the procedure of determining new signs or general phases. The usual techniques require a 'starting-set of phases' (SSP) to which are assigned definite phases or phase symbols and on this base the whole subsequent process depends. For some structures it has been found that there are 'lucky' SSP's which lead to structure solution while the great majority of SSP's do not. One way in which an 'unlucky' SSP may operate for centrosymmetric structures is for there to be an invalid sign relationship in the first few steps of sign determination which leads subsequently to a complete breakdown of the sign-determining process. A method of allowing for such a contingency has been discussed by Germain & Woolfson (1968).

We are now going to describe general approaches to the problem of getting a good starting point. Two types of procedure will be considered:

(1) Whenever a weak link develops in the chain of phase determination, leading to an uncertain new phase or a doubtful new sign, one proceeds thereafter with some alternative values of the new phase or sign.

(2) The strongest possible SSP is chosen so that weak links are eliminated or reduced in number and so that all new phases are determined by strong relationships. Since, for a given structure determination, the number of very strong phase relationships is limited this usually means that one must get multiple indications for new phases as soon as possible.

A combination of procedures (1) and (2) is likely to be useful for dealing with difficult structures. However the computer time required to allow for extreme situations can easily become excessive and solutions should, in the first place, normally be attempted with the minimum of effort by the most straightforward procedures. Only when the simple approach fails should sophisticated and time-consuming techniques be tried.

Forking in non-centrosymmetric structures

Let us consider a situation in which a number of SSP's have been set up in the way described by Germain & Woolfson (1968). Some reflexions, in general three, may be allocated specific phases to fix the origin of the cell, one reflexion may, depending on the space group, be allocated a specific phase or the alternatives ($\pi/4$, $3\pi/4$) to fix the enantiomorph and a number of others can be tried in all possible combinations of ($\pm\pi/4$, $\pm 3\pi/4$), (0, π) or ($\pm\pi/2$) depending on the reflexion type. New phase indications will be given by the well-known tangent formula

$$\tan(\varphi_I)_{\mathbf{h}} \simeq \frac{\sum_{\mathbf{h}'} |E_{\mathbf{h}} \cdot E_{\mathbf{h}-\mathbf{h}'}| \sin(\varphi_{\mathbf{h}'} + \varphi_{\mathbf{h}-\mathbf{h}'})}{\sum_{\mathbf{h}'} |E_{\mathbf{h}} \cdot E_{\mathbf{h}-\mathbf{h}'}| \cos(\varphi_{\mathbf{h}'} + \varphi_{\mathbf{h}-\mathbf{h}'})} = \frac{S_{\mathbf{h}}}{C_{\mathbf{h}}}, \quad (1)$$

where $(\varphi_I)_{\mathbf{h}}$ is the indicated value of the phase whose true value is $\varphi_{\mathbf{h}}$.

The strength of the indication can be estimated from

$$\alpha_{\mathbf{h}} = 2\sigma_3\sigma_2^{-3/2}\,|E_{\mathbf{h}}|\,(S_{\mathbf{h}}^2 + C_{\mathbf{h}}^2)^{1/2}\,, \qquad (2)$$

where the symbols have been defined by Karle & Karle (1966). If there is only one contributor to the indication for the new phase then we write

$$\alpha_{\mathbf{h}} = \kappa_{\mathbf{h},\mathbf{h}'} = 2\sigma_3\sigma_2^{-3/2}\,|E_{\mathbf{h}}E_{\mathbf{h}'}E_{\mathbf{h}-\mathbf{h}'}|\,. \qquad (3)$$

Karle & Karle gave a formula for the variance of a phase indication given by equation (1) in terms of the corresponding value of α. We have calculated the variance, V_α, and standard deviation, σ_α, for a range of values of α and these are given in Table 1.

On the basis of experience it is found that if a newly determined phase is initially within about 45° of its true value then it will in its turn lead to useful new information. The error in the phase will also probably be reduced at some later stage by tangent-formula refinement.

The probability distribution of the phase indication from equation (1) will, for small deviations, be approximately Gaussian and we shall assume in the following discussion that the statistics of the Gaussian distribution may be applied. It can be seen from Table 1 that if $\alpha \simeq 7{\cdot}0$ then 45° represents two standard deviations which means that there is a probability of 0·95 that the indication for the new phase is within 45° of the true value.

If we use the criterion that there must be a probability of 0·95 for the new phase being within 45° of its true value before it is acceptable, how do we proceed when the best new indication corresponds to an α less than 7·0? Let us consider the case $\alpha = 2{\cdot}3$. Since $\sigma_\alpha = 45°$ there is a probability of 0·95 that the true phase, φ_r, is within 90° of the indicated phase, φ_I. Thus if we take two alternative values for the new phase, $\varphi_I + \sigma_\alpha$ and $\varphi_I - \sigma_\alpha$, there is a 95% probability that one or other of them is within 45° of the true phase. For $\alpha < 2{\cdot}3$ one should have to take three alternative values for the new phase; thus for $\alpha = 1{\cdot}2$, corresponding to $\sigma_\alpha = 67\frac{1}{2}°$ there will be a 95% probability that one of the three values, φ_I, $\varphi_I + 4\sigma_\alpha/3$, $\varphi_I - 4\sigma_\alpha/3$ will be within 45° of φ_r. For $\alpha < 1{\cdot}2$ one would simply use the technique of allocating the trial phases to different quadrants *i.e.* $\pm\pi/4$, $\pm 3\pi/4$, which gives complete certainty that one of the allocated values is within 45°.

There are of course other criteria which could be used other than those given above. Thus one might wish to restrain the maximum probable error to less than 45° or to take a probability other than 0·95. Table 2 gives the α ranges for the numbers of alternatives required for various combinations of probability and maximum error.

Thus we can see that if we are working to the 0·90 probability limit and to an error limit of 35° then an α equal to 7·8 or more is required for a single alternative to satisfy the chosen limits while with α between 2·6 and 7·8 two alternatives would be necessary.

This provides us with the basis of a branching procedure which, whenever the phase determining process reaches a point where some probability criterion is not satisfied, takes two, three or four alternative routes. There are various ways in which such a process can be operated – for example one can specify that no route should involve more than two 4-branches, two 3-branches and three 2-branches; when this limit is

Table 1. *Variances and standard deviations for a range of* α

α	V_α (radians²)	σ_α (radians)	σ_α (degrees)	α	V_α (radians²)	σ_α (radians)	σ_α (degrees)
0·0	3·290	1·814	103·9	6·0	0·184	0·429	24·6
0·5	2·349	1·533	87·8	7·0	0·155	0·394	22·6
1·0	1·604	1·267	72·6	8·0	0·134	0·366	21·0
1·5	1·091	1·044	59·8	10·0	0·106	0·325	18·6
2·0	0·764	0·874	50·1	12·0	0·087	0·295	16·9
2·5	0·563	0·750	43·0	14·0	0·074	0·272	15·6
3·0	0·437	0·661	37·9	16·0	0·065	0·254	14·6
3·5	0·354	0·595	34·1	18·0	0·057	0·239	13·7
4·0	0·298	0·546	31·3	20·0	0·051	0·227	13·0
5·0	0·227	0·477	27·3				

Table 2. α *ranges for various combinations of probability and maximum error*

Probability of error less than	0·90				0·95				0·99			
	30°	35°	40°	45°	30°	35°	40°	45°	30°	35°	40°	45°
No. of alternatives	α	α	α	α	α	α	α	α	α	α	α	α
1	10·5	7·8	6·3	5·0	14·7	10·8	8·5	6·8	~25	18·0	14·5	11·2
2	3·2	2·6	2·1	$1{\cdot}7_5$	4·2	3·3	2·7	2·3	6·7	5·1	4·1	3·4
3	$1{\cdot}7_5$	$1{\cdot}3_5$	1·0	$0{\cdot}6_5$	2·3	1·8	1·4	$1{\cdot}1_5$	3·4	2·7	2·2	1·9
4												

reached the straight-through path is followed thereafter. This condition would require up to 1142 routes of phase determination to be explored, a formidable but not impossible task. It is our experience, gained with the quadrant-phase-fixing method that it is rarely necessary to follow through all the possible paths to the end as many soon reveal themselves as implausible by the poor quality of new phase generation.

More complex computer algorithms can be devised – for example alternative routes could be generated in such a way that no route has more than a certain total deviation from the main straight-through path – but in practice simple applications of branching procedures will be found to be the most useful.

A trial of a simple branching technique has been made with a known structure, the photolysis product solved by Karle, Karle & Estlin (1967). Four reflexions were selected to fix the origin and enantiomorph; these reflexions, their $|E|$ values and allocated phases (as given by the published solution) are given in Table 3.

From this beginning it is possible to develop five new phases, each being given by a single relationship. These reflexions, their $|E|$ values and indicated phases are also listed in Table 3 together with the pairs of reflexions which give them. It will be seen that the indicated phases agree quite well with the published phases, given in parentheses, although larger errors would be expected from the standard deviations associated with the phase relationships.

The strategy adopted in this trial was to allow up to two deviations of $\pm 40°$ in the five relationships displayed in Table 3. This gives the 51 possible SSP's which are illustrated in Table 4 where the symbols + and − indicate the sign of the 40° deviation. It should be noted that the deviations are in the *relationships* and not in the phases as given in Table 3. Thus if we follow through set 51 of deviations we find the five derived phases in Table 3 as 40, 170, 100, 170, 80°.

For each of the 51 SSP's there were developed 50 new phases by the application of equation (1) which was also used to refine the phases in the sense of giving them the greatest self-consistency. At this stage there was calculated for each SSP a figure of merit which, in the terminology of equation (1) was

$$Z=\sum_{\mathbf{h}} |E_{\mathbf{h}}|^2 (S_{\mathbf{h}}^2+C_{\mathbf{h}}^2) . \qquad (4)$$

Since the trial was with a known structure it was possible at this stage to look at the values of Z and the corresponding values of $\langle|\Delta\varphi|\rangle$, the mean deviation of phase from the true value. These are listed in order of magnitude of Z in Table 5.

This Table shows a number of extremely interesting features. Firstly pairs of sets, whose set numbers added together equal 52, have very similar values of Z (they usually differ in figures less significant than those given in the Table). An inspection of Table 4 reveals that these correspond to pairs of sets which have deviations in opposite directions. The reason for this is not understood and is probably a feature of the pattern of phase development in this structure. For the pairs of sets with the highest values of Z one set corresponds to a very satisfactory development of phases while the other gives almost random phases.

At this point the 16 sets with the lowest values of Z were discarded, which incidently included set 26, the no-deviation set, and phases were found for another 50 reflexions for each remaining set. The situation at that stage is shown in Table 6.

Another 50 phases were developed for the 14 sets with highest Z in Table 6 and the final Z values (see Table 7) just began to discriminate the good sets from the bad:

In fact the seven good sets were all very similar and could all be developed to give sufficient phases to give a complete solution to the structure. Just as a matter of interest set 26, the no-deviation set, was developed up to the 150 phase stage and gave $\langle|\Delta\varphi|\rangle=32{\cdot}7°$. It was inferior to some of the other sets but it might have just been good enough to reveal the structure.

While we have carried out preliminary experiments with branching procedures which indicate their high potentiality for coping with complex structures these have not yet been systematically incorporated into any of our phase-determining computer packages. We have, however, put rather more effort into evolving methods of finding favourable SSP's and this work will now be described.

The use of *n*-sets

A phase relationship may be expressed in the form

$$\cos(\varphi_{\mathbf{h}_1}+\varphi_{\mathbf{h}_2}+\varphi_{\mathbf{h}_3})=1-\delta^2 , \qquad (5)$$

where $\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3=0$ and δ^2 is, in general, small and has a probability distribution determined by the magnitude of $\kappa_{\mathbf{h},\mathbf{h}'}$. Knowledge of any two phases gives a

Table 3. *E values and allocated phases for the trial structure*

h	k	l	$\|E\|$	φ		Given by
4	0	7	2·29	90°		Origin and enantiomorph
0	1	15	2·61	270		
5	4	0	2·28	90		
1	0	13	1·69	270		
4	4	13	1·62	0	(342°)	(5 4 0) & (1 0 13)
8	4	6	2·07	90	(92)	(4 0 7) & (4 4 13)
7	4	7	2·04	180	(182)	(1 0 13) & (8 4 6)
6	4	6	2·45	90	(86)	(1 0 13) & (7 4 7)
2	4	13	2·42	0	(342)	(4 0 7) & (6 4 6)

Table 4. *Possible SSP's*

Relationship	Set number 1					2					3					4					5				
	1	2	3	4	5	0	9	8	7	6	0	9	8	7	6	0	9	8	7	6	0	9	8	7	6
1																									
						0	0	0	0	0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
						0	0	0	0	0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
						0	0	0	0	0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
						0	0	0	0	0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
						0	0	0	0	0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2																									
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3																									
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4																									
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5																									
	1	1	1	1	1	0	0	0	0	0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	1	1	1	1	1	0	0	0	0	0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	1	1	1	1	1	0	0	0	0	0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	1	1	1	1	1	0	0	0	0	0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	1	1	1	1	1	0	0	0	0	0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

probable value for the third. Such a relationship may be expressed in a generalized notation

$$R \equiv \{\varphi_{h_1}, \varphi_{h_2}, \varphi_{h_3}\}, \tag{6}$$

where containing the phases in the curled bracket means that they are related but does not give explicitly the exact form of the relationship.

Now let us consider a pair of relationships such as

$$R_1 \equiv \{\varphi_a, \varphi_b, \varphi_c\},$$
$$R_2 \equiv \{\varphi_a, \varphi_b, \varphi_d\}. \tag{7}$$

It is clear that if φ_a and φ_b are both known then φ_c and φ_d can be found. Other pairs of known phases can give knowledge about the other two, *e.g.* φ_a and φ_c. The feature of the pair of relationships R_1 and R_2 is that they contain two phases in common. If more relationships can be found which also contain this pair of phases then fixing two suitable phases will give phase indications for all the other phases contained in the relationships.

We can express the total information of the relationships (7) in the form

$$R_1 + R_2 \equiv [\varphi_a, \varphi_b]\, \varphi_c, \varphi_d, \tag{8}$$

and we say that φ_c and φ_d are both associated with the 2-set φ_a, φ_b. If there are more relationships associated with the same 2-set we should have

$$R_1 + R_2 + R_3 + \ldots = [\varphi_a, \varphi_b]\, \varphi_c, \varphi_d, \varphi_e, \ldots. \tag{9}$$

Fixing both components of the 2-set or one member of the 2-set and any one other phase leads to indications for all other phases, each given by a single relationship.

There are, of course, three 2-sets associated with each relationship and sets of relationships can be found with interlinked 2-sets of the form shown below:

$$\begin{aligned} R_1 &\equiv \{\varphi_a, \varphi_b, \varphi_c\} \\ R_2 &\equiv \{\varphi_a, \varphi_b, \varphi_d\} \\ R_3 &\equiv \{\varphi_a, \varphi_c, \varphi_e\} \\ R_4 &\equiv \{\varphi_b, \varphi_d, \varphi_f\}. \end{aligned} \tag{10}$$

The total information of relationships (10) may be represented by

$$R_1 + R_2 + R_3 + R_4 \equiv$$
$$[\varphi_a, \varphi_b]\, \varphi_c, \varphi_d \cup [\varphi_a, \varphi_c]\, \varphi_b, \varphi_e \cup [\varphi_b, \varphi_d]\, \varphi_a, \varphi_f \tag{11}$$

We have borrowed the terminology of set theory to indicate that the individual components on the right-hand side have some duplicated elements. If two phases are fixed, either a pair in a 2-set or one from a 2-set and one phase associated with that 2-set, then all the remaining phases can be found. Finding relationships linked by 2-sets gives a train of phase indications which can ensure the steady development of phases but there is only one relationship for each phase indication and it is unlikely that all the indications will be strong.

Of far more practical application is the exploitation of sets of relationships linked as

$$\begin{aligned} R_1 &\equiv \{\varphi_a, \varphi_b, \varphi_c\} \\ R_2 &\equiv \{\varphi_a, \varphi_d, \varphi_e\} \\ R_3 &\equiv \{\varphi_f, \varphi_b, \varphi_d\} \\ R_4 &\equiv \{\varphi_f, \varphi_c, \varphi_e\} \,. \end{aligned} \tag{12}$$

We may write

$$\begin{aligned} R_1 + R_2 &\equiv [\varphi_b, \varphi_c]\, \varphi_a + [\varphi_d, \varphi_e]\varphi_a \\ &\equiv [\varphi_b, \varphi_c, \varphi_d, \varphi_e]\, \varphi_a \end{aligned} \tag{13}$$

and φ_a is said to be associated with the 4-set $[\varphi_b, \varphi_c, \varphi_d, \varphi_e]$. If three of the five phases are known, either three from the 4-set or a correctly chosen two from the 4-set and φ_a, then the other two can be found from the pair of relationships.

The other pair of relationships involve the same 4-set and hence we may write

$$R_1 + R_2 + R_3 + R_4 \equiv [\varphi_b, \varphi_c, \varphi_d, \varphi_e]\, \varphi_a, \varphi_f \,. \tag{14}$$

In this situation a knowledge of three phases leaves three unknown phases and four relationships, and one of the phases is given by a double indication. For example, if φ_b, φ_c and φ_d are known, φ_a may be found from R_1, φ_e from R_2 and then φ_f is given by both R_3 and R_4.

The practical application of 4-sets is quite straightforward. One can frequently find many pairs of relationships with a common 4-set. If k such pairs of relationships are found then they involve $4+k$ phases and there are $2k$ relationships. If three phases are known – for example because they are among those used to fix the origin and enantiomorph – then two more phases are indicated with a single relationship and $k-1$ phases will have double indications.

This general principle for finding highly interrelated phases can be extended. The three relationships

$$\begin{aligned} R_1 &\equiv \{\varphi_a, \varphi_b, \varphi_c\} \\ R_2 &\equiv \{\varphi_a, \varphi_d, \varphi_e\} \\ R_3 &\equiv \{\varphi_a, \varphi_f, \varphi_g\} \end{aligned} \tag{15}$$

can be combined as

$$R_1 + R_2 + R_3 \equiv [\varphi_b, \varphi_c, \varphi_d, \varphi_e, \varphi_f, \varphi_g]\, \varphi_a \,. \tag{16}$$

If four phases are fixed, all from the 6-set or a correctly chosen three from the 6-set and φ_a, then the remaining three phases are indicated, each by one relationship. However if there are other groups of three relationships giving the same 6-set then multiple phase indications result. As an example we consider in addition to (15) the relationships

Table 5. *Z and* $\langle|\Delta\varphi|\rangle$ *values*

Set	Z	$\langle\|\Delta\varphi\|\rangle$	Set	Z	$\langle\|\Delta\varphi\|\rangle$	Set	Z	$\langle\|\Delta\varphi\|\rangle$
47	$2{\cdot}124\times10^5$	20·2°	22	$1{\cdot}053\times10^5$	39·4°	19	$0{\cdot}926\times10^5$	38·1
5	2·124	80·9	35	1·051	37·0	33	0·926	84·8
40	2·065	22·8	17	1·051	81·6	44	0·845	86·7
12	2·065	88·3	32	1·042	39·8	8	0·845	87·8
24	1·288	78·4	20	1·042	32·7	43	0·828	72·5
28	1·288	22·8	45	1·003	82·1	9	0·828	59·6
14	1·248	77·3	7	1·003	38·6	41	0·740	83·8
38	1·248	25·5	42	1·002	43·9	11	0·740	86·9
6	1·180	79·3	10	1·002	80·1	49	0·736	84·4
46	1·180	27·3	29	1·000	39·0	3	0·736	85·2
34	1·176	76·8	23	1·000	81·8	31	0·671	84·5
18	1·176	32·3	48	0·999	42·3	21	0·671	85·7
51	1·111	88·0	4	0·999	94·0	26	0·529	47·2
27	1·099	47·8	39	0·962	38·9	37	0·521	56·6
25	1·099	80·8	13	0·962	83·0	50	0·521	44·5
1	1·070	91·0	36	0·927	85·8	2	0·518	49·4
30	1·053	32·0	16	0·927	86·7	15	0·517	44·2

Table 6. *Z and* $\langle|\Delta\varphi|\rangle$ *values after discarding of the lowest* 16 *Z-value sets*

Set	Z	$\langle\|\Delta\varphi\|\rangle$	Set	Z	$\langle\|\Delta\varphi\|\rangle$	Set	Z	$\langle\|\Delta\varphi\|\rangle$
45	$6{\cdot}624\times10^5$	83·6°	40	$6{\cdot}190\times10^5$	19·1°	34	$5{\cdot}541\times10^5$	84·1°
28	6·617	16·5	42	6·172	23·9	48	5·290	56·1
17	6·607	83·3	5	6·061	81·3	4	5·156	85·9
7	6·577	17·7	19	6·005	21·2	36	4·828	90·7
10	6·572	82·3	23	6·001	81·4	51	4·806	92·6
24	6·550	81·1	12	5·993	83·0	13	4·710	96·4
35	6·537	18·5	30	5·958	19·1	27	4·566	34·8
14	6·443	82·2	18	5·904	27·0	1	4·539	95·2
38	6·431	17·1	32	5·904	81·6	25	4·348	91·8
46	6·325	17·4	20	5·898	19·5	16	4·317	95·0
47	6·310	18·2	22	5·892	79·5	39	4·131	89·7
6	6·261	80·7	29	5·890	26·8			

$$R_4 \equiv \{\varphi_h\ \varphi_b\ \varphi_e\} \qquad R_7 \equiv \{\varphi_i\ \varphi_b\ \varphi_g\}$$
$$R_5 \equiv \{\varphi_h\ \varphi_c\ \varphi_f\} \qquad R_8 \equiv \{\varphi_i\ \varphi_c\ \varphi_d\}$$
$$R_6 \equiv \{\varphi_h\ \varphi_d\ \varphi_g\} \qquad R_9 \equiv \{\varphi_i\ \varphi_e\ \varphi_f\} \tag{17}$$

which give altogether

$$\sum_{n=1}^{9} R_n \equiv [\varphi_b, \varphi_c, \varphi_d, \varphi_e, \varphi_f, \varphi_g]\ \varphi_a, \varphi_h, \varphi_i\,. \tag{18}$$

If four phases are known then two more will be found from single indications, two from double indications and one from a triple indication.

A general program has been written which, to a limited extent, exploits the n-set principle to find efficient SSP's. The process will be described with reference to data for the photolysis product, listed in Table 8.

(*a*) For each reflexion the quantity

$$\psi_{\mathbf{h}} = \sum_{\mathbf{h}'} \kappa_{\mathbf{h},\mathbf{h}'} \tag{19}$$

is computed where the summation is over all relationships in which the reflexion is involved.

(*b*) Four reflexions are chosen to fix the origin and enantiomorph in such a way that for the four reflexions $\sum\psi_{\mathbf{h}}$ is as large as possible. For the photolysis product these were found to be reflexions with code numbers 2, 29, 104 and 141.

A search is now made for the m other reflexions, where m is an input parameter of the program, which, with the original four, will give the most efficient SSP. Efficiency is judged by the criterion that the least value of α for a phase determination should be as high as possible.

For the present example we take $m=2$.

(*c*) The first listed reflexion was found which was not one of the four found in (*b*) and had a value of ψ greater than 0·2 $\psi_{\max}$, where $\psi_{\max}$ is the maximum ψ for all the reflexions. This reflexion, that with code number 1 in Table 8, we shall refer to as r_1.

(*d*) All the relationships in which the reflexion is involved are found. There are 36 such relationships for r_1 and the six strongest of these are given below

$$\begin{aligned}
\varphi_1+\varphi_{70}-\varphi_{74} &\simeq 0\,, & \kappa&=3{\cdot}20\\
\varphi_1+\varphi_{141}-\varphi_{144} &\simeq 0\,, & \kappa&=3{\cdot}02\\
\varphi_1-\varphi_{126}-\varphi_{127}+\pi &\simeq 0\,, & \kappa&=2{\cdot}76\\
\varphi_1-\varphi_{151}-\varphi_{152} &\simeq 0\,, & \kappa&=2{\cdot}70\\
\varphi_1+\varphi_{109}-\varphi_{113} &\simeq 0\,, & \kappa&=2{\cdot}61\\
\varphi_1+\varphi_{171}-\varphi_{175} &\simeq 0\,, & \kappa&=2{\cdot}23
\end{aligned}$$

From these relationships all the 6300 4-sets with which r_1 is linked are generated. If in any of these 4-sets two members are known [because they are two of the four fixed in (*b*)] then fixing the remaining two will give a double indication for the phase of r_1. In fact none of the 4-sets with which r_1 is associated had two known members so the next reflexion in Table 8 was found satisfying the conditions given in (*c*). This was the reflexion with code number 3, r_3.

One of the 4-sets with which r_3 was associated had two known members. This is derived from

$$\varphi_2-\varphi_3-\varphi_6\simeq 0\,, \quad \kappa=2{\cdot}70$$

and

$$\varphi_3-\varphi_{140}-\varphi_{141}+\pi\simeq 0\,, \quad \kappa=1{\cdot}52$$

giving the 4-set $[\varphi_2, \varphi_6, \varphi_{140}, \varphi_{141}]$.

If φ_6 and φ_{140} are fixed then φ_3 is given by a double indication.

(*e*) Starting with the seven phases, $\varphi_2, \varphi_3, \varphi_6, \varphi_{29}, \varphi_{104}, \varphi_{140}, \varphi_{141}$ the complete set of relationships is explored to see how a development of phases might progress. It is assumed that when a phase is given by two or more relationships the value of α is given by $\alpha_\kappa=\sum\kappa$, an assumption which will be dealt with in the following section. The process is continued until a point is reached where no new phase can be found or until the total value of $\sum\alpha_\kappa$ for all the reflexions which have been found equals $\frac{2}{3}\sum\kappa$ for all the relationships available.

The minimum value of α_κ is noted for the phase development from this set of seven phases.

(*f*) The reflexions following r_3 in Table 8 are explored in sequence. If they satisfy the conditions in (*c*) then they are dealt with in the way described for r_3. An updated record is kept of that pair of reflexions which, with the original four, gives the highest minimum value of α_κ and the value of $(\alpha_\kappa)_{\min}$.

For the application of the method to the photolysis product it was found that fixing φ_{127} and φ_{146} gave a minimum α_κ of 3·0 for the remainder of the phase determination. While the process may seem to be very time consuming there are various time-saving features – for example no schematic phase development is followed if the minimum α_κ is less than the highest minimum previously recorded. The process outlined above took two minutes on an IBM 7090 computer.

If a seven reflexion starting point was permitted then the minimum α_κ was about 4·0. In an attempt to simu-

Table 7. *Final Z values*

Set	Z	$\langle\lvert\Delta\varphi\rvert\rangle$	Set	Z	$\langle\lvert\Delta\varphi\rvert\rangle$	Set	Z	$\langle\lvert\Delta\varphi\rvert\rangle$
40	$1{\cdot}191\times10^6$	18·6°	46	$1{\cdot}191\times10^6$	18·4°	17	$1{\cdot}069\times10^6$	81·4°
7	1·191	18·5	47	1·191	18·5	10	1·062	84·5
35	1·191	18·6	24	1·122	81·1	45	1·062	84·5
38	1·191	18·5	14	1·122	81·1	42	1·018	41·6
28	1·191	18·5	6	1·111	81·4			

late a difficult structure 6-sets were used to find a starting point of ten reflexions which give a minimum α_κ equal to about 6·0.

In the above work it was assumed that α_κ gave an indication of the expected value of α. This would imply that where there were several indications for a given phase they were all completely consistent; since this is rarely true α_κ will normally be an over-estimate of α. We shall consider now how to get a better estimate of α from the individual values of κ.

Table 8. *Data for the photolysis product*

CODE	H	K	L	E	Ø	CODE	H	K	L	E	Ø	CODE	H	K	L	E	Ø
1	0	0	12	2.38	180	2	0	1	15	2.58	270	3	0	2	8	2.04	180
4	0	2	10	1.31	0	5	0	3	5	1.36	270	6	0	3	7	2.11	90
7	0	3	13	1.90	270	8	0	3	14	2.30	90	9	0	4	4	1.51	180
10	0	4	8	2.19	0	11	0	4	11	1.79	0	12	0	4	12	1.44	180
13	0	4	14	1.94	0	14	0	4	17	1.96	180	15	0	4	19	1.45	0
16	0	5	7	2.40	270	17	0	5	16	1.34	270	18	0	5	17	1.42	90
19	0	6	5	1.79	180	20	0	6	10	1.33	180	21	0	6	11	1.35	0
22	0	6	13	1.50	180	23	0	6	16	1.80	0	24	0	6	17	1.53	0
25	0	6	18	1.86	180	26	1	0	4	1.42	180	27	1	0	8	1.59	180
28	1	0	11	1.28	90	29	1	0	13	1.69	270	30	1	1	3	1.35	262
31	1	1	7	1.38	199	32	1	1	14	1.51	178	33	1	2	0	1.46	270
34	1	2	2	1.43	334	35	1	2	12	1.62	66	36	1	2	14	1.46	215
37	1	3	3	1.25	137	38	1	3	12	1.65	197	39	1	4	4	1.41	248
40	1	4	6	1.25	73	41	1	4	9	1.29	146	42	1	4	10	1.31	315
43	1	4	14	1.91	93	44	1	4	18	1.42	295	45	1	5	9	1.43	19
46	1	5	12	1.38	331	47	1	5	13	1.89	222	48	1	5	16	1.28	40
49	1	5	18	1.52	296	50	1	6	0	1.46	270	51	1	6	13	1.53	308
52	1	6	15	1.54	154	53	1	6	17	1.25	47	54	2	1	1	1.91	303
55	2	1	3	1.53	125	56	2	1	13	1.40	110	57	2	1	15	1.38	288
58	2	1	19	1.41	131	59	2	2	1	1.54	348	60	2	2	4	1.38	8
61	2	2	14	1.84	168	62	2	2	18	1.45	16	63	2	3	4	1.30	200
64	2	3	5	1.37	272	65	2	3	6	1.92	260	66	2	3	7	1.49	65
67	2	3	9	1.50	129	68	2	3	12	1.35	314	69	2	3	14	1.28	150
70	2	4	1	2.23	157	71	2	4	5	1.96	351	72	2	4	8	1.73	19
73	2	4	10	2.05	220	74	2	4	13	2.38	341	75	2	4	17	1.86	195
76	2	4	18	1.66	200	77	2	5	7	1.49	259	78	2	5	8	1.30	332
79	2	5	10	1.45	144	80	2	5	12	1.86	55	81	2	5	14	1.27	258
82	2	6	8	2.14	168	83	2	6	9	1.74	26	84	2	6	15	1.65	326
85	3	1	4	1.34	3	86	3	1	12	1.50	36	87	3	1	14	1.30	222
88	3	1	15	1.37	297	89	3	2	7	1.31	17	90	3	3	12	1.67	176
91	3	4	2	1.34	211	92	3	4	3	1.92	324	93	3	4	7	1.30	119
94	3	4	10	1.62	225	95	3	4	14	1.27	57	96	3	4	15	1.27	140
97	3	5	1	1.25	308	98	3	5	4	1.42	65	99	3	5	6	2.13	187
100	3	5	9	1.68	324	101	3	5	16	1.52	180	102	4	0	0	1.80	0
103	4	0	4	1.85	180	104	4	0	7	2.31	90	105	4	0	9	1.26	270
106	4	0	11	1.58	270	107	4	0	12	1.52	180	108	4	1	15	1.25	281
109	4	2	0	2.51	180	110	4	2	2	1.49	332	111	4	2	3	1.36	39
112	4	2	7	1.40	328	113	4	2	12	1.73	338	114	4	2	14	1.37	168
115	4	3	7	1.32	134	116	4	3	9	1.80	12	117	4	4	1	1.36	144
118	4	4	7	1.97	146	119	4	4	8	2.47	16	120	4	4	10	1.25	162
121	4	4	11	1.36	199	122	4	4	13	1.58	342	123	4	5	0	1.60	0
124	4	5	2	1.40	154	125	4	5	4	1.35	210	126	4	5	5	2.10	91
127	4	5	7	2.18	266	128	4	5	8	1.64	266	129	4	5	10	1.42	5[illegible]
130	4	5	14	1.41	343	131	4	6	3	1.27	321	132	4	6	5	1.78	212
133	4	6	6	1.34	318	134	4	6	8	1.98	152	135	4	6	9	2.02	39
136	4	6	10	1.31	31	137	4	6	12	1.31	208	138	5	0	8	1.46	180
139	5	2	6	1.86	183	140	5	2	8	1.35	322	141	5	4	0	2.27	90
142	5	4	5	1.28	316	143	5	4	7	1.44	215	144	5	4	12	2.21	228
145	5	4	13	1.61	180	146	5	4	14	2.11	29	147	5	4	15	1.45	347
148	5	5	0	1.50	270	149	5	5	1	1.73	110	150	5	5	2	1.55	127
151	5	5	5	1.68	45	152	5	5	7	2.67	234	153	5	5	12	1.67	7
154	5	6	4	1.50	98	155	5	6	11	1.48	145	156	5	6	14	1.32	190
157	6	1	8	1.43	27	158	6	2	0	1.27	180	159	6	2	1	1.37	329
160	6	2	5	1.53	94	161	6	2	7	1.77	265	162	6	3	5	1.30	49
163	6	3	13	1.43	9	164	6	3	14	1.53	115	165	6	4	1	1.91	156
166	6	4	2	1.45	302	167	6	4	4	1.87	199	168	6	4	6	2.40	86
169	6	4	8	1.48	312	170	6	4	12	1.63	144	171	6	5	0	1.93	0
172	6	5	7	1.51	278	173	6	5	10	1.42	328	174	6	5	11	2.07	63
175	6	5	12	1.92	165	176	6	5	13	1.73	260	177	6	6	2	1.47	126
178	6	6	3	1.40	94	179	6	6	4	1.33	333	180	7	3	1	1.35	89
181	7	3	10	1.33	48	182	7	4	5	1.51	36	183	7	4	7	1.98	182
184	7	4	10	1.29	125	185	7	5	3	1.34	263	186	7	5	9	1.44	254
187	7	6	3	1.38	92	188	7	6	5	1.49	238	189	8	0	7	1.72	90
190	8	2	5	1.56	87	191	8	2	7	1.85	289	192	8	3	3	1.45	10
193	8	3	5	1.44	224	194	8	4	1	1.45	111	195	8	4	3	1.43	239
196	8	4	4	1.59	266	197	8	4	6	1.97	92	198	8	5	0	2.97	0
199	8	5	1	1.38	318	200	8	5	2	2.36	187	201	8	6	2	1.49	85

Estimating α in the absence of phase information

The way in which the different components of the tangent formula combine to give an estimate of $\varphi_{\mathbf{h}}$ is shown in the form of a phase-vector diagram in Fig. 1. The amplitudes of the components of the diagram are of length $\kappa_{\mathbf{h},\mathbf{h}'}$ and they have phases $(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'})$. The resultant has length $\alpha_{\mathbf{h}}$ and its phase, $(\varphi_I)_{\mathbf{h}}$ is the indicated value of the phase whose true value is $\varphi_{\mathbf{h}}$.

Now we can also write

$$\alpha_{\mathbf{h}}=[\{\sum_{\mathbf{h}'}\kappa_{\mathbf{h},\mathbf{h}'}\sin(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\}^2+\{\sum_{\mathbf{h}'}\kappa_{\mathbf{h},\mathbf{h}'}\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\}^2]^{1/2}\,, \quad (20)$$

since this corresponds to a rotation of Fig. 1 through an angle $-\varphi_{\mathbf{h}}$, as is shown in Fig. 2, and clearly leaves the resultant unchanged in magnitude.

Cochran (1955) found the probability distribution for $\varphi_{\mathbf{h}}$ for a fixed $\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}$ and given value of $\kappa_{\mathbf{h},\mathbf{h}'}$.

This is

$$P(\varphi_{\mathbf{h}})=[2\pi I_0(\kappa_{\mathbf{h},\mathbf{h}'})]^{-1}\exp[\kappa_{\mathbf{h},\mathbf{h}'}\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})]\,, \quad (21)$$

where I_0 is a modified Bessel function.

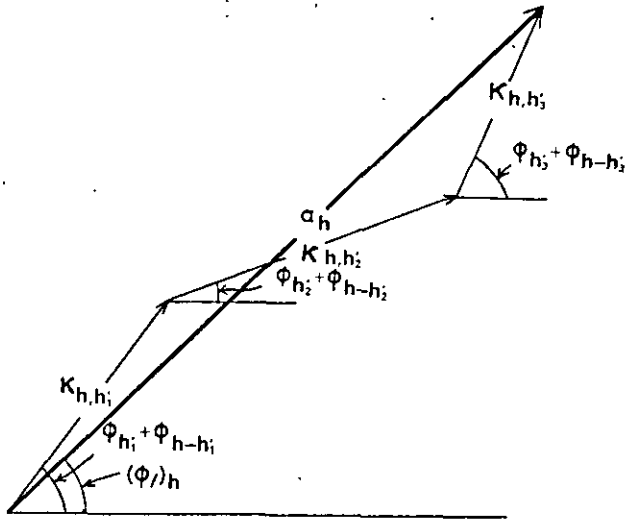


Fig. 1. The composition of different phase indications $\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}$ with amplitudes $\kappa_{\mathbf{h},\mathbf{h}'}$ to give a resultant of amplitude $\alpha_{\mathbf{h}}$ and phase $(\varphi_I)_{\mathbf{h}}$.

We can use this probability distribution to find the mean value of $\alpha_{\mathbf{h}}^2$ for a given set of $\kappa_{\mathbf{h},\mathbf{h}'}$. Thus, from equation (20) we find

$$\begin{aligned}\alpha^2=&\sum_{\mathbf{h}'}\kappa_{\mathbf{h},\mathbf{h}'}^2\\&+\sum_{\substack{\mathbf{h}'\\ \mathbf{h}'\neq\mathbf{h}''}}\sum_{\mathbf{h}''}\kappa_{\mathbf{h},\mathbf{h}'}\kappa_{\mathbf{h},\mathbf{h}''}\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\\&\qquad\cos(\varphi_{\mathbf{h}''}+\varphi_{\mathbf{h}-\mathbf{h}''}-\varphi_{\mathbf{h}})\\&+\sum_{\substack{\mathbf{h}'\\ \mathbf{h}'\neq\mathbf{h}''}}\sum_{\mathbf{h}''}\kappa_{\mathbf{h},\mathbf{h}'}\kappa_{\mathbf{h},\mathbf{h}''}\sin(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\\&\qquad\sin(\varphi_{\mathbf{h}''}+\varphi_{\mathbf{h}-\mathbf{h}''}-\varphi_{\mathbf{h}})\,. \quad (22)\end{aligned}$$

The expectation value of $\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})$ may be found from (21) and is

$$\langle\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'})\rangle=\int_0^{2\pi}P(\varphi_{\mathbf{h}})\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\,\mathrm{d}\varphi_{\mathbf{h}}\,. \quad (23)$$

The exponential term in equation (21) may be expanded as a sum of modified Bessel functions, *viz.*

$$\exp[\kappa_{\mathbf{h},\mathbf{h}'}\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})]=I_0(\kappa_{\mathbf{h},\mathbf{h}'})+2\sum_{m=1}^{\infty}I_m(\kappa_{\mathbf{h},\mathbf{h}'})\cos m(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\,. \quad (24)$$

Substituting from (24) into (23) and integrating we find

$$\langle\cos(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\rangle=\frac{I_1(\kappa_{\mathbf{h},\mathbf{h}'})}{I_0(\kappa_{\mathbf{h},\mathbf{h}'})}\,. \quad (25)$$

Similarly it can be found that

$$\langle\sin(\varphi_{\mathbf{h}'}+\varphi_{\mathbf{h}-\mathbf{h}'}-\varphi_{\mathbf{h}})\rangle=0\,. \quad (26)$$

Putting expectation values into equation (22) we find

$$\langle\alpha_{\mathbf{h}}^2\rangle=\sum_{\mathbf{h}}\kappa_{\mathbf{h},\mathbf{h}'}^2+2\sum_{\substack{\mathbf{h}'\\ \mathbf{h}'\neq\mathbf{h}''}}\sum_{\mathbf{h}''}\kappa_{\mathbf{h},\mathbf{h}'}\kappa_{\mathbf{h},\mathbf{h}''}\times\frac{I_1(\kappa_{\mathbf{h},\mathbf{h}'})}{I_0(\kappa_{\mathbf{h},\mathbf{h}'})}\,\frac{I_1(\kappa_{\mathbf{h},\mathbf{h}''})}{I_0(\kappa_{\mathbf{h},\mathbf{h}''})}\,. \quad (27)$$

This formula is easy to apply to find a value of $\langle\alpha_{\mathbf{h}}^2\rangle^{1/2}$ when a number of κ's are known. For use in a computer for a range of κ up to 5·0 it is found that

$$I_1(\kappa)/I_0(\kappa)\simeq 0{\cdot}5658\kappa-0{\cdot}1304\kappa^2+0{\cdot}0106\kappa^3 \quad (28)$$

gives a maximum error of about 4%. Some values are given in Table 9.

As an example of the application of equation (27) we consider a structure containing 100 equal atoms in

Table 9. *Values of $I_1(\kappa)/I_0(\kappa)$ for various κ*

κ	$I_1(\kappa)/I_0(\kappa)$	Series	κ	$I_1(\kappa)/I_0(\kappa)$	Series
0·0	0·000	0·000	2·5	0·765	0·765
0·5	0·243	0·252	3·0	0·810	0·810
1·0	0·446	0·446	3·5	0·841	0·837
1·5	0·596	0·591	4·0	0·863	0·855
2·0	0·698	0·695	4·5	0·880	0·871
			5·0	0·893	0·894

Table 10. *The first part of the computer output*

185	0.00						
121	0.00						
63	0.00						
53	0.00						
48	0.00						
41	0.00						
21	0.00						
20	0.00						
18	0.00						
17	0.00						
78	1.22						
		-54	135	1.22			
186	1.23						
		101	-104	-1.23			
98	1.25						
		-1	101	1.25			
101	1.22						
		-45	-104	-1.22			
45	1.24						
		104	150	-1.24			
37	1.27						
		2	-35	-1.27			
97	1.31						
		-54	-141	1.31			
173	1.33						
		-54	135	1.33			
44	1.34						
		103	146	1.34			
162	1.36						
		-35	-152	1.36			
15	1.37						
		106	119	-1.37			
96	1.41						
		1	92	1.41			
40	1.42						
		109	139	1.42			
180	1.42						
		-54	141	1.42			
131	1.48						
		1	-135	-1.48			
188	1.50						
		-70	139	-1.50			
49	1.56						
		106	152	-1.56			
158	1.63						
		82	119	1.63			
81	1.77						
		2	-70	-1.77			
181	1.99						
		2	-182	-1.26	35	-200	-1.23
4	2.02						
		-10	25	1.29	70	83	-1.23
91	2.04						
		1	-94	-1.25	-82	139	-1.29
142	2.05						
		29	-119	1.30	54	99	-1.26
199	2.06						
		-103	126	1.30	-104	128	-1.27
137	2.06						
		19	104	1.31	-54	174	-1.26
155	2.07						
		14	-139	-1.31	61	-92	1.27
93	2.11						
		-43	104	-1.39	141	189	1.23
95	2.12						
		-82	-139	1.23	104	183	-1.41
42	2.14						
		-92	104	-1.41	-103	146	1.24
164	2.15						
		-55	152	1.28	92	104	1.39
46	2.18						
		-90	109	1.40	104	151	-1.30
24	2.28						
		1	19	1.58	-60	74	1.22
148	2.28						
		-1	153	1.45	2	-147	1.36
153	1.68						
		38	109	1.68			
147	2.27						
		52	-109	1.36	88	-198	1.43
140	2.30						
		3	-141	-1.52	27	109	1.31

Table 11. *The final portion of the computer output*

118	6.07
65	5.08
8	5.25
56	5.09
59	5.43
126	5.58
1	4.17
57	6.19
74	5.71
123	6.43
108	6.72
102	6.51
189	6.24
168	6.42
70	4.60
161	5.24
61	4.93
66	5.29
3	5.22
119	4.51
77	4.77
134	4.80
10	5.64
16	4.75
157	5.84
82	5.11
165	3.92
72	2.89
54	0.00
109	4.86
6	2.81
191	0.00
127	3.63
198	0.00
141	0.00
29	0.00
2	0.00
104	0.00

54	65	1.75
-61	-161	-1.56
8	-82	2.02
-59	127	1.56
54	74	-2.24
157	168	1.69
1	54	1.54
-82	-126	1.53
-70	-109	2.09
168	-191	-1.66
1	-127	-2.64
-70	74	3.06
6	72	-1.22
56	119	-1.23
-82	-127	1.56
-74	102	1.24
16	-104	-2.15
-102	198	2.07
2	102	1.41
10	127	1.45
-10	119	2.36
-104	189	1.73
16	-198	-2.97
72	165	1.38
54	127	-2.42
-72	-104	-2.16
2	-157	1.58
66	-198	-1.90
2	-54	-2.20
66	127	-1.45
2	-72	-1.61
-54	119	1.70
2	-6	2.69
2	-127	3.37
2	-72	1.61
-54	134	1.37
2	-127	-2.70
2	-6	-2.89
54	72	1.92
6	165	-1.40
72	-198	1.78
-54	127	-2.16
54	-198	-2.63
6	54	1.69
-6	127	2.80
191	198	-2.81
104	-198	-3.63

56	-65	1.28
109	191	2.22
16	59	1.72
56	70	-1.54
72	126	1.23
-127	-168	-1.78
-109	165	1.79
54	168	2.33
168	168	1.66
-16	-72	-1.39
72	127	1.26
104	168	-3.20
54	-70	-1.65
127	189	1.46
2	-109	-1.96
-16	-119	-1.80
-16	127	2.28
61	161	1.36
-61	-134	2.12
127	-157	-1.69
3	-165	-1.67
6	77	-1.40
3	-54	-1.41
-77	-109	-1.35
-109	119	3.07
-54	77	1.70
2	-82	-2.00
10	109	2.64
2	-16	3.29
-54	82	2.38
54	104	1.53
-82	198	-2.20
-72	-109	2.25
72	-104	-1.85
54	127	-1.75
-104	191	2.60

59	82	1.57
59	126	-1.51
57	-70	1.52
74	123	-1.29
-119	161	1.63
-74	-157	1.73
16	82	-1.72
-77	-119	-1.23
-168	189	2.38
54	165	1.42
6	119	-1.58
16	134	-1.44
-70	165	1.86
70	-168	2.23
-70	104	-3.00
-134	161	1.90
61	-104	-1.82
-16	-66	-1.60
10	54	-1.51
10	-54	-1.51
-157	-165	1.59
54	191	-1.23
-165	191	1.83

the unit cell and $|E_h|=3{\cdot}0$ with three contributors for which $|E_{h'}E_{h-h'}|$ equals 3·0, 4·0 and 5·0. The corresponding values of κ are 1·8, 2·4 and 3·0 and the value of $\langle\alpha_h^2\rangle^{1/2}$ computed from equation (27) is 6·1 compared with the value 7·2 which would result if the three phase indications agreed precisely.

It is worthwhile to use values of $\langle\alpha_h^2\rangle^{1/2}$ rather than $\sum\kappa_{h,h'}$ in assessing the probable reliability of a new phase indication for the n-set method. This is particularly true when there are a large number of weak indications as otherwise the probable reliability of the phase indication can be greatly overestimated.

The convergence method

We shall now describe an alternative method of testing the effectiveness of using a particular set of origin and enantiomorph-fixing reflexions and of selecting other reflexions to get the best possible SSP. The method will be described by its application to the photolysis product data listed in Table 8. The origin and enantiomorph combination tested is that used for the n-set example, *i.e.* reflexions with code numbers 2, 29, 104 and 141.

The first step is to work out the value of $\langle\alpha^2\rangle^{1/2}$ for each of the reflexions from the κ's of the relationships in which it is involved. The lowest value of $\langle\alpha^2\rangle^{1/2}$ is sought and in fact it is found that for ten reflexions $\langle\alpha^2\rangle^{1/2}=0$. These reflexions are removed from the data and their computer-program code numbers and values of $\langle\alpha^2\rangle^{1/2}$ are listed, as shown in Table 10. Reflexion 78 is that with the next lowest value of $\langle\alpha^2\rangle^{1/2}$, equal to 1·22. This is listed in Table 10 with the value of $\langle\alpha^2\rangle^{1/2}$ and the single relationship which gives it is indicated by the numbers which follow, which indicate that

$$(\varphi_I)_{78}=-\varphi_{54}+\varphi_{135}\,,\quad \kappa=1{\cdot}22\,.$$

Reflexion 78 is now removed from the list of reflexions and the phase relationship is removed from the list of relationships. The values of $\langle\alpha^2\rangle^{1/2}$ for reflexions 54 and 135 are modified to allow for the removal of the phase relationship and once more the reflexion with the lowest value of $\langle\alpha^2\rangle^{1/2}$ is sought. As before this is listed, removed from the set of reflexions, all relationships in which it is involved removed from the set of relationships and new modified values of $\langle\alpha^2\rangle^{1/2}$ are found. This process is continued with the proviso that the origin and enantiomorph fixing reflexions are never removed from the system but remain to the very end. Table 10 shows a reproduction of the beginning of the computer output for the photolysis-product problem. When a negative κ is shown it implies that there is a π to be incorporated in the phase relationship. Thus, for the first reflexion with a double indication for phase, that with code number 181, we have

$$\varphi_{181}\simeq\varphi_2-\varphi_{182}+\pi\,,\quad \kappa=1{\cdot}26$$

and

$$\varphi_{181}\simeq\varphi_{35}-\varphi_{200}+\pi\,,\quad \kappa=1{\cdot}23$$

giving overall $\langle\alpha^2\rangle_{181}^{1/2}=1{\cdot}99$.

However, the interesting and most useful part of the computer output comes at the end, where the process 'converges' on the four starting reflexions, and this is reproduced in Table 11.

It should be noticed that, from the way this convergence tabulation is produced, the relationships indicating any particular phase involve only those reflexions which have not yet been eliminated from the Table and hence are listed somewhere below. The other characteristic of the convergence process is that at each stage it removes from the system that reflexion which is least well linked to the remainder and thus it tends to enrich the quality of the interrelationships of those which remain.

Let us now examine Table 11 to see what it tells us. Starting from the bottom there are the four reflexions which fix the origin and enantiomorph and for these $\langle\alpha^2\rangle^{1/2}=0$ as must be so since such reflexions cannot be interelated. There follows reflexion 198 which is not linked to any below and then we find reflexion 127

Table 12. *The number of reflexions which must be fixed for different values of* α

Minimum α	Number to fix	Code numbers
2·0	7	2, 29, 54, 104, 141, 191, 198
2·5	7	2, 29, 54, 104, 141, 191, 198
3·0	9	2, 6, 29, 54, 72, 104, 141, 191, 198
3·5	11	2, 6, 29, 54, 72, 99, 104, 141, 149, 191, 198
4·0	14	2, 6, 29, 31, 54, 72, 99, 104, 127, 141, 149, 165, 191, 198

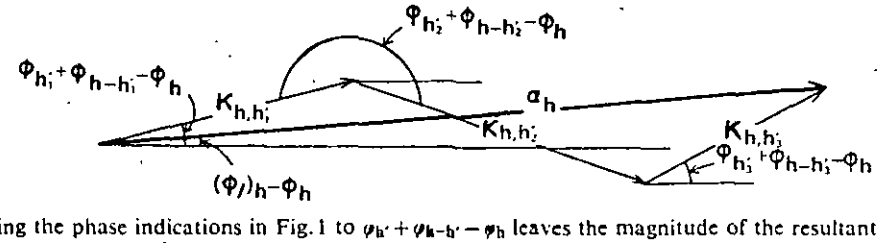


Fig. 2. Changing the phase indications in Fig. 1 to $\varphi_{h'}+\varphi_{h-h'}-\varphi_h$ leaves the magnitude of the resultant unchanged.

linked to 104 and 198 by a fairly strong phase relationship. The next reflexion 191 has no contributors and it is followed by reflexion 6 with one fairly strong phase indication. We then find the first double indication with a correspondingly large value of $\langle\alpha^2\rangle^{1/2}$, 4·86, and next reflexion 54 with no contributor. From this point onwards there are multiple indications with large values of $\langle\alpha^2\rangle^{1/2}$ and, if the phase development up to reflexion 72 has been satisfactory, it should stand a good chance of progressing successfully thereafter.

If it is decided, on the basis of a judgement of the complexity of the problem, that one should not rely on any α less than, say, 2·5 then it is straightforward from the tabulated values of $\langle\alpha^2\rangle^{1/2}$ to decide which reflexions should be incorporated in the starting set. With this α limit for the present problem one needs a starting set of 7 reflexions and the computer program outputs their code numbers. Of course one does not scan the whole convergence table or one will end up deciding that all the reflexions listed in Table 10 need to be incorporated in the starting set. In practice it is sufficient to examine the first 50 or so steps in the schematic phase development process.

In Table 12 are shown the number of reflexions which need to be fixed for values of α from 2·0 to 4·0 by steps of 0·5 for the photolysis product.

The convergence program is quite rapid; with 200 reflexions, 800 relationships and a medium-speed computer (70,000 operations per second) the whole process takes two minutes. Thus it is possible to run the program on a number of trial sets of origin and enantiomorph fixing reflexions. To assist the judgement of which set is to be preferred one may use as a figure-of merit

$$C=\sum_{r=1}^{100}\langle\alpha^2\rangle_r^{1/2},$$

where the summation is over the first 100 stages of phase development for which phases are not fixed.

It must be appreciated that the convergence mapping is a *schematic* phase-development scheme and the actual progress of phase development will be dictated by *computed* values of α and not the estimated values, $\langle\alpha^2\rangle^{1/2}$. In practice the first steps of an actual phase development usually follow closely the schematic scheme. When deviations do occur it is often due to having calculated values of α which are higher than the corresponding values of $\langle\alpha^2\rangle^{1/2}$ so that the actual path of phase development is better than that indicated in the convergence map.

It is possible to use the principle of convergence mapping in a more elaborate way in which one repeatedly obtains an updated estimate of which new reflexions should be fixed. To illustrate this let us suppose that a starting set is chosen from a convergence map for a structure with space group $P2_12_12_1$ and the map indicates that the four origin and enantiomorph fixing reflexions plus two others will enable some progress to be made with phase development. One of the two extra reflexions might, let us say, be a general one and tried in four quadrants, ($\pm\pi/4$, $3\pi/4$) while the other may be special reflexion $(0, \pi)$ or $(\pi/2, -\pi/2)$. The eight starting points can be used to find ten new phases and the group of ten will, in general, be different for each of the starting points. New convergence maps, one for each of the eight starting points, converging on the sixteen reflexions with known phase may indicate that for some starting points two more reflexions need to be fixed, for others only one or perhaps none at all.

There will obviously be a time penalty to pay for such a scheme but it might be worth while if one is dealing with a difficult structure.

Implementation

The techniques described in this paper are being incorporated in a non-centrosymmetric program package, written in FORTRAN conforming to the A.S.A. standard.

The package will be flexible so that, for example, an '*n*-sets' or 'convergence' routine can be used or can be bypassed. The system will offer an automatic route for structure determination which will not require any intervention on the part of the user although it will be possible for an educated user to exercise judgement and to intervene. For the automatic route, decision-making processes, such as convergence mapping, will be used to optimize as far as possible the probability of finding the correct solution.

The extensive collaboration between our laboratories has been made possible by the generous financial support of the North Atlantic Treaty Organization. Two of us (P.M. & M.M.W.) would like to express our gratitude for continued support by the Science Research Council. The success of our work has been assisted by the excellence of the computer facilities provided both at Louvain and at York, and also at Pittsburgh where one of us (P.M.) spent a period developing the *n*-set method. We have also been encouraged by the interest of Professor M. van Meerssche in our project.

References

COCHRAN, W. (1955). *Acta Cryst.* **8**, 473.
GERMAIN, G. & WOOLFSON, M. M. (1968). *Acta Cryst.* B**24**, 91.
KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849.
KARLE, I. L., KARLE, J. & ESTLIN, J. A. (1967). *Acta Cryst.* **23**, 494.

1. Hallgren, B. and Larsson, S. *J. Lipid Res.* **3** (1962) 39; **3** (1962) 331.
2. Tsujimoto, M. and Toyama, Y. *Chem. Umschau* 27, 29, 35, 43, (1922).
3. Holman, R. T., Lundberg, W. O and Malkin, T. *Progress in the chemistry of fats and other lipids* **5** (1958) 118.
4. Brohult, A. *Acta Radiol. Suppl.* **223** (1963).
5. Brohult, J. *Opuscula Medica Suppl.* XII (1969).
6. Brohult, A., Brohult, J. and Brohult, S. *To be published.*

Received February 7, 1970.

A General Set of Computer Programs for the Determination of Crystal Structures by Means of Symbolic Addition Methods

OVE LINDGREN, OLIVER LINDQVIST and JENS NYBORG

Department of Inorganic Chemistry, Chalmers University of Technology and the University of Göteborg, P.O. Box, S-402 20 Göteborg 5, Sweden

A set of computer programs (GAASA I–VI) for solving crystal structures by "Direct Methods", mainly according to the procedure of symbolic addition, using the Σ_2 relation, developed by Hauptmann and Karle,[1,2] has been written. All space groups, centrosymmetric as well as non-centrosymmetric, are included.

The programs have been written in FORTRAN IV for the IBM-360 system. Each of the six parts is compiled separately, and all data are stored between the different programs on disc or magnetic tape. The core memory requirement is at present 143 000 bytes, which, technically, permits the treatment of structures with 15 000–20 000 reflections in the limiting sphere. For more moderate structures, the core memory requirement could easily be lowered to about 80 000 bytes, but below this limit, considerable rewriting of the programs would be necessary.

The required input is a list of independent, observed structure factors ($h,k,l,|F_o|$, $\sin\theta/\lambda$), preferably stored on disc or magnetic tape. Each of the six subprograms requires, moreover, a number of control cards which define the structural problem.

The course of the phase determination is as follows:

GAASA I — Absolute scale and overall temperature factor according to Wilson[3] are evaluated.

GAASA II — $|E|$ values are calculated from $|F_o|_{abs}$, and $|E|$ statistics are performed. The reflections are sorted in decreasing order of $|E|$.

GAASA III — All symmetry dependent reflections are generated. A list of pairs according to the Σ_2 relation is prepared for each $|E|$ value. Only $|E|$ values greater than a limit, chosen by the user, are taken into account.

GAASA IV — Symbols are assigned for phases or signs of a number of reflections (*e.g.* 10) with large $|E|$ values. Symbolic addition is applied in a cyclic procedure, giving symbols for the phases of new reflections and equations between the symbols, which are subsequently solved. When a sufficient number of symbols have been evaluated, origin determining signs or phases are inserted for the remaining symbols.

GAASA V — Centrosymmetric case. The set of E values are expanded using the Σ_2 summation.

GAASA VI — Non-centrosymmetric case. The phase assignment of the largest $|E|$ values is first performed according to formula *2.9* given by Karle and Karle[2] and the phases are refined using the tangent formula.[4] $|E|_{calc}$ values are then evaluated.

From *GAASA V* or *VI* a list of h, k, l, A, and B ($A=|E|\cos\alpha$, $B=|E|\sin\alpha$) are stored on disc for Fourier summation.

The first tests of this set of programs were performed on two small, already known structures:

(1) Catechol, $C_6H_4(OH)_2$, with one molecule in the asymmetric unit of $P2_1/a$.[5] Signs were evaluated for 109 of 111 reflections with $|E|>1.5$, and all determinations were correct. All heavy atoms were very well resolved in E maps summed with these 109 reflections (11 % of all observed and non-observed reflections), and no additional significant maxima appeared. The execution time for *GAASA I–IV* for this structure is less than 15 min for an IBM 360/50 computer.

(2) Cytosine, $C_4H_5N_3O$, with one molecule in the asymmetric unit of $P2_12_12_1$.[6] 133 reflections (20 %) with $|E| > 1.2$ were assigned phases. A comparison with the refined data gave a mean deviation of 14° and a maximum deviation of 64° in α. Most satisfactory E maps with no spurious peaks were also obtained in this case.

Other known structures which have successfully been handled are the ferrocene derivative, $Fe_6C_{70}O_2H_{68}$,[8] and N-phenyl-N'-benzoylselenourea, $C_6H_5 \cdot NH \cdot CSe \cdot NH \cdot CO \cdot C_6H_5$,[9] with space groups $P\bar{1}$ and $P2_1/c$, respectively. The first unknown structure which was solved with this program set was $K_2CS_3 \cdot H_2O$.[10]

When coding *GAASA I* and *II*, extensive use was made of *SAP 1* and *2* written by Hall.[11]

The FORTRAN IV listings and a short description of the card input for the program set can be obtained from the authors. A discussion of the methods applied in the programs, and especially in *GAASA IV*, will be published.[7]

Acknowledgement. This work has been supported by the *Swedish Natural Science Research Council* (Contract No. 2318) and by *Chalmers University of Technology*, which has supplied a grant to cover the costs of the computer work at the *Gothenburg Universities' Computing Centre.*

1. Hauptmann, H. and Karle, J. *A.C.A. Monograph No. 3*, The Letter shop, Wilmington 1953.
2. Karle, J. and Karle, I. L. *Acta Cryst.* **21** (1966) 849.
3. Wilson, A. J. C. *Nature* **150** (1942) 151.
4. Karle, J. and Hauptmann, H. *Acta Cryst.* **9** (1956) 635.
5. Brown, C. J. *Acta Cryst.* **21** (1966) 170.
6. Barker, D. L. and Marsh, R. E. *Acta Cryst.* **17** (1964) 1581.
7. Nyborg, J. *Acta Chem. Scand. To be published.*
8. Hanic, F. *Private communication.*
9. Hope, H. *Acta Cryst.* **18** (1965) 259.
10. Philippot, E. and Lindqvist, O. *Acta Cryst. To be published.*
11. *IUCr World List of Crystallographic Computer Programs*, 1966, NRC-4, No. 537.

Received February 11, 1970.

Enrichment of Gangliosides in Plasma Membranes of Hamster Kidney Fibroblasts

O. RENKONEN, C. G. GAHMBERG, K. SIMONS and L. KÄÄRIÄINEN

Department of Biochemistry, Department of Serology and Bacteriology, and Department of Virology, University of Helsinki, Helsinki, Finland

Earlier investigations have suggested that animal plasma membranes may contain gangliosides.[1-6] Plasma membranes of hamster kidney fibroblasts (BHK21 cells) have been isolated in this laboratory by using enzyme and antigen markers to follow the isolation steps.[7] These cells do contain gangliosides,[8] and the present report shows that they are enriched in the plasma membrane to a remarkable extent.

Material and methods. BHK21 cells, clone WI-2, were grown as described elsewhere.[9] Fragments of plasma membrane and endoplasmic reticulum were isolated by the method of Wallach and Kamat[10] as described previously.[7] Protein, marker enzymes, and plasma membrane antigens were assayed as described previously.[7] Lipids were extracted from lyophilized samples (2—30 mg protein) by two treatments at 20° with 3 ml of chloroform-methanol (2:1) for 4 h, followed by two similar extractions with chloroform-methanol (1:2). This procedure is believed to extract kidney gangliosides completely.[11] In some experiments the completeness of extraction was actually determined as described by Weinstein *et al.*;[6] only traces of lipids remained in the extracted protein.

To the combined chloroform-methanol extracts was added 6 ml chloroform and 4.5 ml water. The two liquid phases were equilibrated and separated. The lower phase was washed twice with fresh upper phase. The washed lower layer is called the *phospholipid extract.* The combined upper phases were dialyzed against distilled water and lyophilized. The dry residue was extracted 3 times with 3 ml of chloroform-methanol (2:1) and filtered. This solution is called the *ganglioside extract.*

The gangliosides were assayed by the method of Warren.[12] Semiquantitative analyses were carried out also by thin layer chromatography (TLC) on silica gel G plates with propanol-

APPENDIX III

Miscellaneous Programs*

PROGRAM SYSTEM DESCRIPTION

At present, a complete system of programs to generate E-maps or F-maps and to recycle information from these maps is available. Part of the thesis work this progress report period was devoted to becoming familiar with the programs, and to linking them up in an integrated system. The following is a description of the main features of the program system.

Two methods are available for generation of phases for an initial E-map: the program FAZE, and the set of three programs MULTAN. These programs together with other programs in the system are shown in Figure 1. FAZE must be supplied with an origin and starting phases, and tangent refinement, Equation (2), is performed only for a single set chosen by the operator. MULTAN is quite automatic, and chooses the origin and starting phases internally. MULTAN supplies several values to each of the unknown starting phases, and refines all of the multiple starting sets. These sets are listed and written on tape. A figure of merit, indicating internal consistency, is given for each set of phases as an indicator of which set is most likely to produce a correct E-map.

DDUMP and TDUMPER, shown in Figure 1, merely punch a card deck of phases from FAZE or MULTAN for input to NRC-2. NRC-2 is a listing program which stores information on tape for input to NRC-8, and other programs in the NRC series.

NRC-8 generates an E-map, F-map Patterson map, or difference map, depending on the control card settings and type of input data. The type of operation performed by NRC-8 during E-map generation is indicated by Equation (3).

*Reprinted from John A. Heitman's Progress Report No. 7, August 8, 1970.

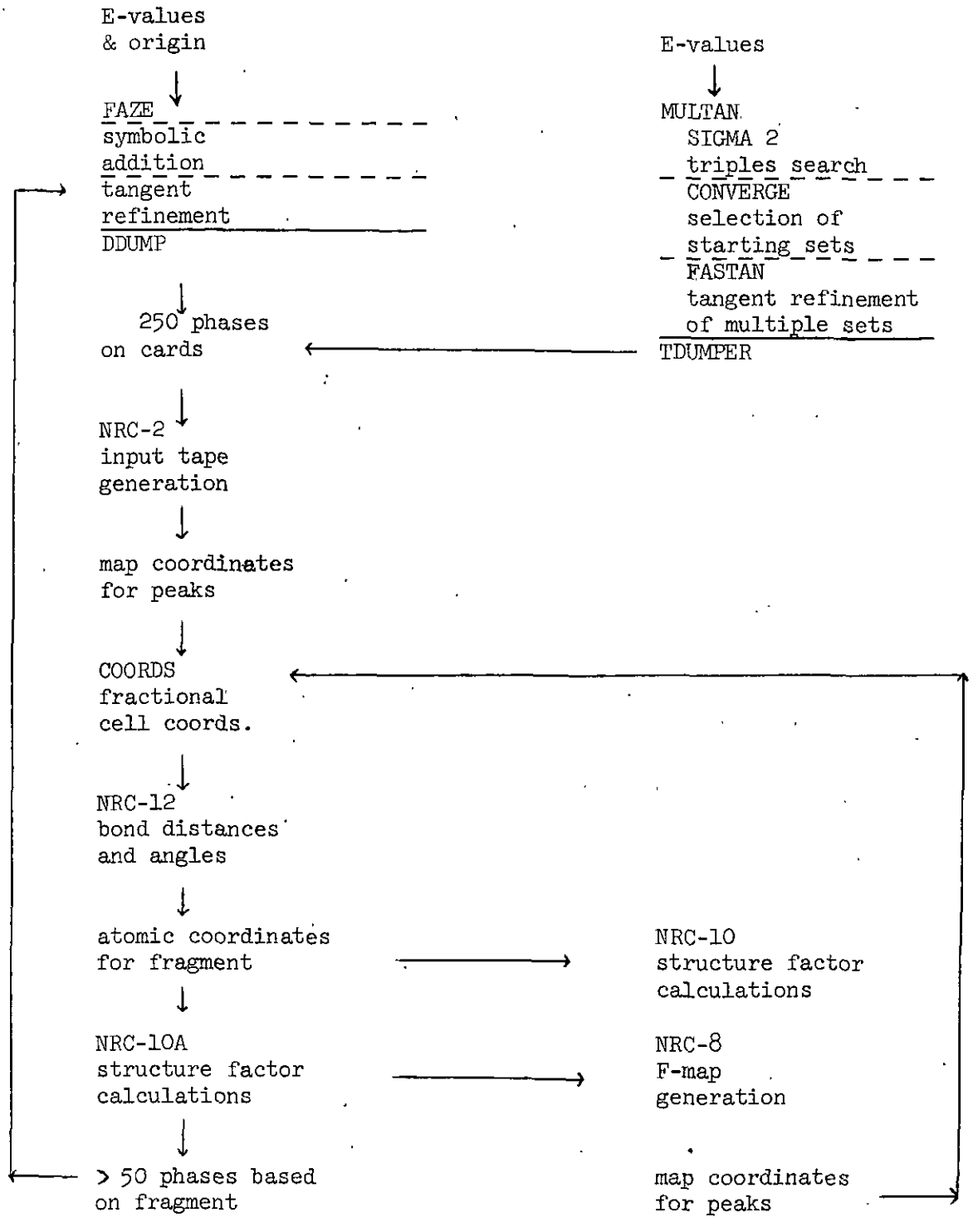


Figure 1. Miscellaneous Programs

Copy 4
File Proj 2612
JWS
7/16/73

An Attempt to Complex Certain
Metal Salts with Simple Carbohydrates
(Project 2612)

by

Gary A. Holcomb
January, 1971

An Attempt to Complex Certain Metal Salts with Simple Carbohydrates

by Gary Holcomb

It is well known that some metal salts will complex with simple carbohydrates. The purpose of this investigation was to attempt to discover different complexes than have been studied previously[1] and to obtain these complexes in crystalline form for x-ray crystallographic studies. Some complexes have already been made and studied in this manner.[1,2] This is a continuation of this program.

Experimental Procedure

The main procedure used was the one suggested and used by Weaver[1], Dale[3], and Isbell[4], where an excess of metal salt is added to insure complete complexation of the carbohydrate, rather than that of Rundleman[5], who used exact ratios of sugar to salt(1:1, 1:2).

The procedure for each possible complex attempted was as follows:1)dissolve the salt and carbohydrate in a minimum of water and add twice the volume of 95% ethanol(With the exception of the barium chloride compounds where they were simply dissolved in water, as barium chloride is insoluble in alchohol).

2) Evaporate the liquid in a light jet of air(to speed drying due to the shortness of time) until a sufficient amount of crystals ~~xxx~~ formed.

3) Filter and wash the crystals in a Buchner funnel.

4)Place the crystals in a dessiccator.

(Later samples were recrystallized in the same manner.)

Melting points were taken on all samples, mercury analyses using a Schöninger flask method with a Beckman "DU" Spectrophotometer were done on all mercury compounds, and optical rotations measured and x-ray diffraction powder patterns taken on all recrystallized samples. The optical rotations were measured with a Zeiss-Winkel Polarimeter, and the x-ray powder patterns were obtained through the use of a North American Phillips X-ray Diffraction Apparatus by Norelco.

Two samples were also analyzed for percent carbohydrates by the L. Borchardt Division of Natural Materials and Sciences with the use of a gas chromatograph.[6]

Salt: Mercury chloride

No.	Sugar	Ratio(Su/S)	Solvents	Results
1	D-xylose	1/Ex	w,Eth	Crystals
2	D-mannose	1/Ex	w,Eth	Crystals
3	D-xylose	1/Ex	w	Crystals
5	D-mannose	1/Ex	w	Crystals
7	D-mannose	1/Ex	w,Eth,MeCH	Crystals
8	D-arabinose	1/Ex	w,Eth	Crystals
9	Fructose	1/Ex	w,Eth	Crystals
11	D-arabinose	1/1	w	Crystals
15	(1)recryst.		Eth	Crystals
16	(2)recryst.		Eth	Crystals
17	D-xylose (recrystallized)	1/Ex	w,Eth	Crystals
21	Dextrose	1/Ex	w,Eth	Crystals
23	D-mannose (recrystallized)	1/Ex	w,Eth	Crystals
24	Dextrose (recrystallized)	1/Ex	w,Eth	Crystals
25	Fructose (recrystallized)	1/Ex	w,Eth	Crystals

Salt: Barium Chloride dihydrate

12	D-xylose	1/Ex	w	Crystals
13	D-mannose	1/Ex	w	Crystals
14	Fructose	1/Ex	w	Crystals
18	Galactose	1/Ex	w	Crystals
19	Cellobiose	1/Ex	w	Crystals
20	Dextrose	1/Ex	w	Crystals
26	D-xylose (recrystallized)	1/Ex	w	Crystals
27	D-mannose (recrystallized)	1/Ex	w	Crystals
28	Dextrose (recrystallized)	1/Ex	w	Crystals
29	Fructose (recrystallized)	1/Ex	w	Crystals

Salt: Cadmium chloride

4	D-xylose	1/Ex	w	Crystals
30	Fructose (recrystallized)	1/Ex	w wite solid	Wh Sol./ Crystals

Salt: Silver chloride

10	D-xylose	1/Ex	w,Eth	Rust dust
22	Cellobiose	1/Ex	w,Eth	Rust dust

Mercury Analyses

A/ "DU" on sensitivity 8
B/ "DU" on sensitivity 9
1/ Blank run correction: no paper burned.
2/ Blank run correction: paper burned.

Compound	%Hg	
D-arabinose-$HgCl_2$	A1/ 66.04%	A2/ 68.89%
Fructose-$HgCl_2$	A1/50.82%	A2/53.08%
	B1/ 56.81%	B2/ 59.07%
(5A)Mannose-$HgCl_2$	A1/ 68.33%	A2/ 71.20%
	B1/ 65.76%	B2/ 68.61%
D-xylose-$HgCl_2$	A1/ 65.24%	A2/ 63.66%
	B1/ 60.59%	B2/ 63.27%
(16)Mannose-$HgCl_2$	A1/ 63.33%	A2/ 55.38%
	B1/ 65.68%	B2/ 58.48%
Dextrose-$HgCl_2$	A1/ 57.57%	A2/ 60.59%
	B1/ 57.65%	B2/ 60.66%

Thoeretical % Hg (No. of assoc. waters)

Complex(1:1)	0	1	2	3	4	5	6
xy-$HgCl_2$*	47.6	45.6	43.9	42.2	40.7	39.3	37.9
man-$HgCl_2$**	44.4	42.8	41.2	39.7	38.4	37.1	35.9
Complex(1:2)	0	1	2	3	4	5	6
xy-$HgCl_2$*	57.9	56.5	55.1	53.7	52.5	51.3	50.1
man-$HgCl_2$**	55.5	54.2	52.9	51.7	50.5	49.4	48.3

*:Also for arabinose-$HgCl_2$
**: Also for Fructose- and Dextrose-$HgCl_2$

% Hg in $HgCl_2$: 73.9%
% Hg in $HgCl_2 \cdot 2H_2O$: 66.12%

Optical Rotation Measurements(Percent sugar)

D-xylose: Spec. rotation:Experimental:1)20.82°
2)20.46°
Actual: 19.1°
%error: 1)9.01%
2)7.12%

((1) denotes zero correction with water in a tube in the polarimeter; (2) denotes zero correction with no tube in the polarimeter)

D-xylose-$HgCl_2$:Rotations; 1).07° Percent sugar:1)13.81%
2).09° 2)17.76%

D-mannose-$HgCl_2$:~~Rotations~~:1) 21.87% 2)29.15%
Rotations:1).06° 2).08°

Dextrose-$HgCl_2$:Rotations:1).08° Percent sugar:1)4.0%
2).10° 2)5.0%

Fructose-$HgCl_2$:Rotations:1)-.96° Percent sugar:1)39.5%
2)-.98° 2)40.1%

D-xylose-$BaCl_2$:Rotations:1).08° Percent sugar:1)11.7%
2).10° 2)14.6%

D-mannose-$BaCl_2$:Rotations:1).09° Percent sugar:1)23.5%
2).11° 2)28.7%

Fructose-$BaCl_2$:Rotations:.01°, not above zero corrections significantly.

Theoretical Percent Sugar

	No. of associated waters						
Complex (1:1)	0	1	2	3	4	5	6
xylose-$HgCl_2$	35.5	34.1	32.8	31.5	30.4	29.3	28.3
Mannose-$HgCl_2$*	39.8	38.3	36.9	35.6	34.4	33.2	32.1
xylose-$BaCl_2$	41.9	39.9	38.1	36.4	34.9	33.5	32.2
Mannose-$BaCl_2$*	46.4	44.3	42.5	40.8	39.1	37.7	36.3
Complex (1:2)	0	1	2	3	4	5	6
xylose-$HgCl_2$	21.6	21.1	20.5	20.1	19.6	19.1	18.7
Mannose-$HgCl_2$ *	24.9	24.3	23.7	23.1	22.6	22.1	21.6
Xylose-$BaCl_2$	26.5	25.7	24.9	24.2	23.5	22.9	22.3
Mannose-$BaCl_2$ *	30.2	29.3	28.4	27.6	26.9	26.2	25.6

*:Also used for dextrose, fructose complexes.

X-ray Diffraction Graph Comparisons

These tables compare the powder patterns of the compounds with that of the pure metal salt, i.e. each catagory indicates something on the pattern of the compound that is different or the same as on the pattern of the pure salt.

Catagories

Peak Shifts: the no. of degrees difference between the angles at which the peaks appeared.

No. of Peaks High: A direct comparison with intensities of peaks of metal salt patterns

No. of Peaks Low: As above.

M.I.S.: Main intensity Sequence, i.e. the order of the four most intense peaks.

Main Ratios: Ratios between the four main peaks.

Small Ratios: Ratios between the numerous smaller peaks.

Extra peaks: Peaks appearing on the compound pattern and not on the pure salt pattern.

Missing Peaks: Peaks appearing on the pure salt pattern and not on the compound pattern.

$BaCl_2 \cdot 2H_2O$ Compounds

Compound	Peak Shifts	No. of Peaks High	No. of Peaks Low	M. I. S.	Main Ratios	Small Ratios	Extra Peaks	Miss. Peaks
Xy-$BaCl_2$	$-.1^o$ $-.2^o$	6	1	Bad	Bad	Good	5	1
Xy-$BaCl_2$(2)	$.1^o$	2	9	Bad	Bad	Good	0	1
Man-$BaCl_2$	$.1^o$ $-.2^o$	10	6	Bad	Bad	Bad	3	0
Dex-$BaCl_2$	$.0^o$ $-.2^o$	4	8	Ver Bad	Ver Bad	VERY Bad	2	0
Dex-$BaCl_2$(2)	$-.1^o$ $-.2^o$	1	9	Ver Good	Good	Fair	0	2
Fruc-$BaCl_2$	$-.1^o$ $-.2^o$	8	4	Fair	Ver Bad	Good	1	1

*Xylose-$BaCl_2$ to its own repeat: Only similarity is the angles at which the peaks occurred. M.I.S. close to each other, but two peaks were interchanged in the sequence. Everything else is very different in one than the other.

*Dextrose-$BaCl_2$ to its own repeat: Shifts are about the same. All intensities changed around (in repeat, much more like pure salt). Extra peaks in original disappeared in repeat.

X-ray Diffraction Graph Comparisons

$HgCl_2 \cdot 2H_2O$ Compounds

Compound	Peak Shifts	No. of Peaks High	No. of Peaks Low	M. I.	Main Ratios	Small Ratios	Extra Peaks	Miss. Peaks
Xylose-$HgCl_2$	.1°	1	3	Fair	Good	Excel.	0	1
Mannose-$HgCl_2$	.05° .2°	6	3	Excel	Excel	Poor	1	1
Dextrose-$HgCl_2$	.0°	7	2	Excel	Excel	Good	3	1
Fructose-$HgCl_2$	.0°	6	1	Bad	Terri	Terrible	3	1
Fructose-$HgCl_2$(2)	.2°	4	4	Fair	Fair	Poor	1	1

*Fructose-$HgCl_2$ to its own repeat: M.I.J. slightly switched around. All ratios are different, with little or no correlation between which peaks are high and which are low. The three extra peaks in the original disappeared in the repeat, and a newx one appeared.

Results and Conclusions

All the solution of barium chloride and mercury chloride formed crystals. My problem was to find out what the crystals were, being pure salt, pure carbohydrate, a mixture of carbohydrate and salt, or a complex. Results of the different analyses are as follows:

Melting Points: All of the mercury compounds charred while below 300°c. , showing the presence of carbohydrate. The barium compounds lost their water at temperatures other and above that of the water-loss temperature of Barium chloride dihydrate, and also charred when heated in a bunsen burner flame. This test gave only an indication of the presence of sugar in the sample melted, a sample which may or may not have been impure due to the smallness of the crystals chosen to be placed in the capillaries-they may have been just crystallized sugar. Therefore this test was inconclusive.

Mercury analysis: Two compounds, fructose--and dextrose-mercury chloride, came very close to the theoretical mercury percentages, and all others came closer to the percent mercury of mercuric chloride dihydrate. However, due to possible inaccuracies in the zero runs, main runs, and measurements, and impurities of the compounds due to incompletewashings(the crystals themselves dissolved too fast in the wash liquid), the percentage mercury values could have been all wrong, sugar impurities lowering the percentage, mercuric chloride impurities raising it.

Optical rotations: Mannose- and fructose-mercuric

chloride and mannose-barium chloride came close to or within the theoretical ranges of percent carbohydrate. The remaining compounds had percentages that were too low to come withing these ranges. The trial run, using xylose to see how accurate the test was, came withing 10% of the true value for the specific rotation of xylose. However, this solution was quite concentrated compared to the solutions of the compounds, and the error in the readings would increase as concentration decreased. If the solutions of the compounds could have been made more concentrated, this test may have been of greater reliability. The readings I obtained were so inconsistant that the chances of the average observed rotation being statistically significant are only fair, and, to compound this, even a difference of .01° in the average could have brought the final percentage off by 5-10%. Therefore the optical rotations were inconclusive also.

X-ray diffraction patterns :The traces of the mercury compounds as a group were very similar, as far as angles where peaking took place were concearned. The same can be said of the barium compounds. However, in several of the traces, extra peaks showed up that were not on the pure salt patterns. Repeats of another sample of the same compound did not show these peaks, however.

In addition, the intensities and intensity ratios of the peaks on the compound patterns were very different in most cases from the pure salt runs. It was thought that possibly the great size and reflecting ability of

the barium, mercury, and chloride ions was drowning out the reflections of the carbohydrates, and that the carbohydrates in turn were causing the shifts and different intensities and ratios to appear. This was at least partially borne out by comparison of a trace of a dry mixture of a carbohydrate(fructose) and salt($BaCl_2$) with that of the pure salt and of the pure carbohydrate. The mixture trace had most of the peaks from both pure compounds, but the carbohydrate responses were drastically reduced. Since the mixture was 50% each of salt and sugar by moles(approximately), in a complex of twenty to thirty percent carbohydrate, it is conceivable that the carbohydrate reflections might not show up at all, except under certain orientations. This would also help to explain the extra peaks which appeared and disappeared from the patterns.

Sugar analysis : Done on only two large samples, this analysis showed the compounds ran to contain less than .01% carbohydrate, completely resolving the question of whether or not complexing had taken place, at least for these two compounds. The compounds were "fructose-mercuric chloride" and "mannose-barium chloride," the two most hopeful complexes as seen from the results of the mercury analysis and optical rotation measurements. Since these came out containing no carbohydrates, it is reasonable to assume that none of the other compounds produced contains carbohydrates, from the information compiled.

The differing crystalline structures, the results of the first analyses, and the variation in the diffractograms all made the possibility of the compounds being complexes seem greater, it turned out, than they were. The only conclusions to be made are that mercuric chloride and barium chloride do not complex with the carbohydrates used and under the conditions present. This does not in any way imply that under different conditions, using other carbohydrates, other solvents, different procedures, etc., they will not complex.

REFERENCES

1) Weaver, John W., Unpublished work, 1970.

2) Richards, Gerald F., Unpublished work, 1970.

3) Dale, J.K., National Bureau of Standards Journal of Research, 3, 459, 1929.

4) Isbell, Horace S., Bureau of Standards Journal of Research, 5, 741, 1930.

5) Kendleman, J.A., Advances in Carbonydrate Chemistry, 21, 209-70, 1966.

6) Borchardt, L.G., Piper, C.V., TAPPI, 53, no. 2, 257-260, February, 1970.

Project Rpt No 5.
Project 2612
July 16, 1973

The Molecular Geometry of β-Pinene As Deduced From the Crystal and Molecular Structure of cis-Pinocarvyl-p-Nitrobenzoate

Gerald F. Richards, Robert A. Moran, and John A. Heitmann

The Institute of Paper Chemistry
Appleton, Wisconsin 54911

William E. Scott*

Department of Biological Sciences
Purdue University
West Lafayette, Indiana 47907

The crystal structure of cis-pinocarvyl-p-nitrobenzoate, (abbreviated name, cis-PNB) a derivative of β-pinene, has been determined from three-dimensional X-ray data obtained near the temperature of liquid nitrogen. The para-nitrobenzoate geometry is similar to that found in other para-nitrobenzene derivatives. The cyclobutane ring in the cis-pinocarvyl portion of the molecule is normal (internal dihedral angle 141°), leading to severe steric interactions between hydrogen atoms on C-3 and C-7. These interactions are minimized in cis-PNB through a decreased puckering in the (C-2, C-3, C-4) end of the molecule. The observed C-3—H.....C-7—H distance is 2.33Å. Since the same interactions exist in β-pinene, it is postulated that the cis-pinocarvyl grouping and β-pinene have very similar molecular geometries. cis-PNB exhibits pseudo mirror molecular symmetry, C-10 being the only nonhydrogen atom not related to another atom by the mirror plane.

β-Pinene (I) and related bicyclic terpenes undergo cyclobutane ring opening to yield monocyclic or acyclic molecules, and cyclobutane ring expansion to form bornane or fenchane compounds.[1,2,3] Such rearrangements dominate β-pinene chemistry, suggesting that the original bicyclic ring system is relatively unstable.

It has been postulated that the rigid cyclobutane ring leads to severe distortions in the rest of the bicyclic system. Support for this idea is found in the reported crystal structures of Cl-3-nopinone and Br-3-nopinone (II).[4] In these structures the cyclobutane ring is normal, but the dihedral angle between plane (C-2, C-3, C-4) and the best plane through C-1, C-2, C-4, and C-5 is considerably smaller than that found in cyclohexane rings. The structural details are not precisely known because the halogen atoms reduced the certainty with which the positions of the other atoms could be determined.

The crystal structure analysis of cis-pinocarvyl-p-nitrobenzoate (abbreviated name, cis-PNB) (III) was carried out in order to determine the precise geometry of a β-pinene derivative. cis-PNB was suitable for this purpose because the carbon, nitrogen, and oxygen atoms all have approximately equal electron densities. Consequently, it was possible to determine their positions with much higher precision than for the halogenated compounds.

EXPERIMENTAL SECTION

X-Ray Data - Crystals of cis-pinocarvyl-p-nitrobenzoate (mp 103.0-103.6°, $[\alpha]_D^{26°}$ -96.0°) suitable for X-ray diffraction analysis were kindly provided by Dr. James Farrard. The unit cell was found to be monoclinic with the following dimensions: at approximately -193°C, a = 10.583 ± 0.006Å, b = 6.740 ± 0.003Å, c = 10.443 ± 0.009Å, β = 90.46 ± 0.15°; at approximately

22°C, a = 10.904 ± 0.012Å, b = 6.778 ± 0.006Å, c = 10.906 ± 0.009Å, β = 91.77 ± 0.27°. The unit cell parameters were determined by the back-reflection Weissenberg technique. Systematic absences occurred for 0k0 when k was odd, indicating space group $P2_1$. Comparison of measured and calculated crystal densities indicated two molecules of cis-PNB per unit cell.

Three-dimensional X-ray data were obtained at approximately -193°C using the multiple film equiinclination Weissenberg method and a liquid nitrogen gas flow cryostat.[5] Two crystals were used to obtain layers (0kl) to (7kl) and (h0l) to (h4l). The reflection intensities were estimated visually by comparison with a calibrated intensity scale. Lorentz and polarization, but no absorption corrections were made. The data from the two crystal settings were scaled together by the method of Rollett and Sparks.[6] A total of 1361 unique, observed reflections were obtained. An additional 120 reflections were either unobserved or too weak to be measured with confidence. The relative intensities were converted to normalized structure factors, $|E_{hkl}|$, by the K-curve method of Karle and Hauptman.[7,8]

Structure Determination and Refinement - The unit cell positions of atoms O-11, C-12, O-13, C-14, C-15, C-16, C-17, C-18, C-19, and N-20 were determined by the Patterson search methods of Braun, Hornstra, and Leenhouts.[9,10] Phases calculated from this fragment were used to start the tangent formula recycle procedure.[12] Two cycles were required to determine the positions of the remainder of the nonhydrogen atoms in the molecule. The structure was refined by block-diagonal least-squares methods to an R index[13] of 0.10 using unit weights and anisotropic temperature factors. A three-dimensional difference electron density map was computed at this point, revealing the positions of all the hydrogen atoms. The least-squares refinement was

continued with the hydrogen atoms included. The final R index for all observed reflections was 0.085. Each hydrogen atom was assigned an isotropic temperature factor equal to the final isotropic temperature factor of the atom to which it was bonded, and all hydrogen parameters were held constant during the least-squares refinement.[14] Table I shows the final atomic coordinates and their estimated standard deviations. The anisotropic temperature factors were normal for a structure determined near liquid nitrogen temperature.[15]

Table I

Atomic Parameters

Atom	Fractional Coordinates X[a]	Y	Z
C-1	0.4411(6)	0.5260(9)	-0.2737(5)
C-2	0.4160(5)	0.3974(9)	-0.1591(5)
C-3	0.2922(5)	0.4530(9)	-0.0951(5)
C-4	0.1910(5)	0.5352(9)	-0.1901(5)
C-5	0.2543(6)	0.6389(9)	-0.3027(5)
C-6	0.3350(5)	0.4856(9)	-0.3791(5)
C-7	0.3801(6)	0.7338(10)	-0.2575(5)
C-8	0.2842(6)	0.2761(10)	-0.3945(5)
C-9	0.3723(6)	0.5654(11)	-0.5098(6)
C-10	0.4893(6)	0.2533(10)	-0.1204(6)
O-11	0.2392(4)	0.2817(6)	-0.0309(3)
C-12	0.2486(5)	0.2723(9)	0.0969(5)
O-13	0.2961(4)	0.3959(7)	0.1645(4)
C-14	0.1933(5)	0.0834(9)	0.1476(5)
C-15	0.1828(6)	0.0629(9)	0.2792(5)

contd.

Atom	X[a]	Y	Z
C-16	0.1327(5)	-0.1085(9)	0.3316(5)
C-17	0.0929(5)	-0.2566(9)	0.2487(5)
C-18	0.1001(5)	-0.2393(9)	0.1171(5)
C-19	0.1514(6)	-0.0675(10)	0.0655(5)
N-20	0.0360(4)	-0.4386(8)	0.3022(4)
O-21	0.0375(4)	-0.4579(7)	0.4189(4)
O-22	-0.0077(4)	-0.5599(7)	0.2275(4)
H-1(C-1)[b]	0.516	0.519	-0.322
H-1(C-3)	0.316	0.545	-0.030
H-1(C-4)	0.135	0.426	-0.219
H-2(C-4)	0.150	0.628	-0.136
H-1(C-5)	0.200	0.740	-0.340
H-1(C-7)	0.404	0.821	-0.332
H-2(C-7)	0.364	0.795	-0.176
H-1(C-8)	0.200	0.276	-0.445
H-2(C-8)	0.268	0.246	-0.309
H-3(C-8)	0.354	0.191	-0.434
H-1(C-9)	0.315	0.497	-0.550
H-2(C-9)	0.426	0.461	-0.530
H-3(C-9)	0.429	0.679	-0.505
H-1(C-10)	0.468	0.162	-0.057
H-2(C-10)	0.550	0.197	-0.170
H-1(C-15)	0.210	0.171	0.330
H-1(C-16)	0.118	-0.092	0.412
H-1(C-18)	0.076	-0.368	0.075
H-1(C-19)	0.176	-0.066	-0.020

[a] Estimated standard deviation times 10^4 in parentheses

[b] H-j(n-k) refers to the jth hydrogen atom bonded to the kth atom of type n.

RESULTS AND DISCUSSION

Figure 1 shows the bond-lengths and bond-angles in cis-PNB.

p-Nitrobenzoate Grouping - The geometry of the p-nitrobenzoate portion of cis-PNB is similar to that reported for a number of other p-nitrobenzene derivatives.[11,23,24,25,26] The average C-C bond-length within the benzene ring is 1.386Å and, within experimental error, all the benzene C-C bonds are equivalent. The average bond-angle within the benzene ring is 120.0°. Angle (C-15—C-16—C-17) is significantly smaller, and angle (C-16—C-17—C-18) is significantly larger than the average. The dihedral angle between the plane of the nitro group and the best plane through the six benzene carbons is 5°, and the dihedral angle between the carboxylate plane and the benzene plane is 7°.

β-Pinene Grouping - The bond-lengths and bond-angles in the four-membered ring composed of C-1, C-5, C-6, and C-7 agree closely with the values reported for other cyclobutane derivatives.[27,28,29,30] The dihedral angle between plane (C-1, C-7, C-5) and plane (C-1, C-6, C-7) is 141°. Values of 149°-180° have been reported for other cyclobutane rings. It is likely that the increased pucker aids in relieving severe steric interactions between hydrogen atoms on C-3 and C-7 in cis-PNB. Bonds C-1—C-2, C-2—C-3, C-3—C-4, and C-4—C-5 all have values close to the normal value of 1.530Å. The double bond between C-2 and C-10 is shorter than in ethylene (1.334Å), but close to the length of the exocyclic double bond in trans-2,8-dihydroxy-1(7)-p-menthene (1.314Å).[31,32] The (noncyclobutane) ring bond-angles at atoms C-1, C-2, C-3, C-4, and C-5 all have values within the range (109°-114°) observed in other cyclohexane structures.

Table II contains the values found for the conformation-angles within the cis-PNB bicyclic ring system.

Table II

Ring Conformation-Angles in the β-Pinene Portion of cis-PNB

Atoms i-1, i, j, j+1	τ_{ij}[a] (°)
C-6, C-1, C-2, C-3	-64.4
C-7, C-1, C-2, C-3	28.7
C-1, C-2, C-3, C-4	30.0
C-2, C-3, C-4, C-5	-27.9
C-3, C-4, C-5, C-6	61.9
C-3, C-4, C-5, C-7	-33.6
C-4, C-5, C-6, C-1	-82.9
C-7, C-5, C-6, C-1	27.3
C-4, C-5, C-7, C-1	83.4
C-6, C-5, C-7, C-1	-27.7
C-2, C-1, C-6, C-5	83.7
C-7, C-1, C-6, C-5	-26.9
C-2, C-1, C-7, C-5	-81.6
C-6, C-1, C-7, C-5	27.1

[a] Looking down bond {i-j}, τ_{ij} is defined as the clockwise angle between bond {(i-1)-i} and bond {j-(j+1)}.

Ideal cyclohexane boat and chair conformations have conformation-angles of 0° and ±60°. The values in Table II indicate that the cis-PNB molecule is flatter at the C-3 end and more puckered at the C-6 and C-7 ends than an ideal boat or chair cyclohexane conformation. The greater

deviation from ideality exhibited by cis-PNB and Cl-3-nopinone compared to some reported cyclohexane derivatives is shown in Table III and Figure 2.

Table III

Dihedral Angles $\delta(1)^a$, $\delta(2)$, and $\delta(3)$ in cis-PNB, Cl-3-Nopinone, and Various Cyclohexane Compounds.

Compound	$\delta(1)°$	$\delta(2)°$	$\delta(3)°$	Reference
Cyclohexane[b]	60.0	60.0	---	34
Cyclohexane[c]	54.6	54.6	---	34
Cyclohexylammonium chloride	50.0	49.0	---	35
Bicyclohexylidene	49.4	51.1	---	33
cis-PNB	28.2	71.5	69.8	--
Cl-3-Nopinone	30.0	75.0	70.0	4

[a] See Fig. 2

[b] Ideal model

[c] Electron diffraction

Increased ring distortion in β-pinene bicyclic structures appears to be necessary in order to minimize the steric interactions between hydrogen atoms bonded to C-3 and hydrogen atoms bonded to C-7. For example, in cis-PNB a small amount of relief is gained through puckering of the cyclobutane ring, as mentioned earlier. However, the major relief arises from the movement of C-3 toward the (C-1, C-2, C-4, C-5) plane and away from C-7, leading to small $\delta(1)$ values (i.e. a "flattened" ring). The shortest (C-3--H) --- (C-7—H) distance observed in cis-PNB is 2.33Å, which is close to the sum of the van der Waals radii (2.4Å).

The ring strain in the molecule could have been partially decreased if the bond-angle (C-1—C-2—C-3) were near 120°. However, the observed

value is 111.5°. The tendency for ring bond-angles at carbons involved in exocyclic double bonds to be close to 110° (rather than 120°) has been found previously.[32,33]

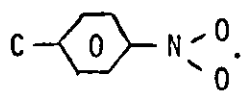
In summary, the β-pinene portion of cis-PNB lies between a classical bridged chair and a Y-shaped molecule. Previous NMR studies on iso-pinocampheol[36] and nopinone[37] have been interpreted in terms of the same kind of conformation. Since the ring distortion arises from the need to minimize repulsion between hydrogens bonded to C-3 and hydrogens bonded to C-7, it is very likely that β-pinene itself will closely resemble the β-pinene portion of cis-PNB.

Molecular Symmetry - The conformation of the benzoate linkage allows a pseudo-mirror plane to be constructed through C-3, C-6, C-7, C-8, C-9, and the nitrobenzoate group. Atoms C-1 and C-2 are approximately related to atoms C-5 and C-4, respectively, by this mirror plane. Of the nonhydrogen atoms, only C-10 is not related by pseudo mirror symmetry to another atom in the molecule.

Crystal Packing - *Figure 3 shows a projection of the unit cell down the* b-axis. The p-nitrobenzoate grouping is oriented perpendicular to the (001) plane. Most of the intermolecular distances are very much larger than normal van der Waals distances, only eleven distances involving C, N, or O being less than 3.4Å. The strongest intermolecular interactions occur along the b-axis, where the molecules are packed so that O-13, C-12, C-14, and C-15 of each molecule lie near O-22, N-20, C-17, and C-18 of another molecule one b-axis translation away. In the c-axis direction, the closest intermolecular distances occur between each nitro oxygen and either C-16 or C-19 of another cis-PNB molecule. In the a-axis direction the packing involves alternating C-10—O-13' and C-4—O-22", C-5—O-22" interactions.

1. J. L. Simonsen and L. N. Owen, "The Terpenes", Second Edition, Volume II, Cambridge University Press, Cambridge, England, 1952.
2. D. V. Banthorpe and D. Whittaker, Chem. Rev. 66, 643 (1966).
3. A. A. Newman, ed., "Chemistry of Terpenes and Terpenoids", Academic Press, New York, N.Y., 1972.
4. Y. Barrans, Comp. Rend. Acad. Sci. 259, 796 (1964).
5. For a description of the liquid nitrogen gas flow cryostat see G. Richards, Ph.D. Dissertation, University of Iowa, Iowa City, Iowa, February 1964.
6. J. S. Rollett and R. A. Sparks, Acta Crystallogr. 13, 273 (1960).
7. J. Karle and H. Hauptman, Acta Crystallogr. 9, 635 (1956).
8. The relative observed intensities, I_{hkl}, were converted to normalized structure factors by the relationship

$$|E_{hkl}| = [I_{hkl} \exp(4.09+0.57(\sin^2\theta/\lambda^2)^{1.5})]^{1/2}.$$

The following statistics were found: $\langle|E_{hkl}|\rangle = 0.809$, $\langle|E_{hkl}|^2\rangle = 0.994$, $\langle\,||E_{hkl}|^2-1|\,\rangle = 0.891$.
9. P. B. Braun, J. Hornstra, and J. I. Leenhouts, Philips Res. Reports 24, 85,427 (1969).
10. The molecular fragment used for the Patterson search procedures was  . The detailed geometry was taken from the published structure of p-nitrobenzoic acid.[11]
11. T. D. Sakore and L. M. Pant, Acta Crystallogr. 21, 715 (1966).
12. J. Karle, Acta Crystallogr. B24, 182 (1968).
13. The crystallographic R-index is simply defined as $R = \dfrac{\Sigma\,\big||F_o|-|F_c|\big|}{\Sigma|F_o|}$
14. The atomic scattering factors for carbon, nitrogen and oxygen were taken from Hanson *et al.*,[16] and the hydrogen scattering factors were those of Stewart *et al.*[17] The scaling and intensity data corrections were carried out with programs supplied by The Crystallography

Laboratory, University of Pittsburgh.[18,19] The three dimensional Fourier synthesis, tangent formula, and least squares refinement programs were provided by F. R. Ahmed.[20] Weights were assigned during the final refinement cycles according to the equation

$$w = [(c_1|F_0|^2+c_2|F_0|+c_3)(c_4\sin^2\theta+c_5\sin\theta+c_6)]^{-1}$$

where the least squares constants $c_1 = 1.38$, $c_2 = -0.20$, $c_3 = 0.014$, $c_4 = 2.30$, $c_5 = -5.76$, and $c_6 = 5.24$ were evaluated by the method suggested by Cruickshank.[21,22]

15. The tables and figures listed in this footnote will appear following these pages in the microfilm edition of this journal: tables containing anisotropic temperature factors, observed and calculated structure factors, and distances of atoms in cis-PNB from the least squares plane through the molecule.

16. H. P. Hanson, F. Herman, J. D. Lea, and S. Skillman, Acta Crystallogr. 17, 1040 (1964).

17. R. F. Stewart, E. R. Davidson, and W. T. Simpson, J. Chem. Phys. 42, 3175 (1965).

18. P. T. Beurskens, Technical Report No. 45, The Crystallography Laboratory, The University of Pittsburgh, Pittsburgh, Pa., 1963.

19. R. Shiono, D. Hall, and S. C. Chu, Technical Report No. 43, The Crystallography Laboratory, The University of Pittsburgh, Pittsburgh, Pa., 1963.

20. F. R. Ahmed, S. R. Hall, M. E. Pippy, and C. P. Huber, NRC Crystallographic Programs for the IBM 360 System, World List of Crystallographic, Computer Programs, Second Edition, Appendix, p. 52, 1966.

21. D. W. J. Cruickshank, In Crystallographic Computing, p. 187, edited by F. R. Ahmed, Copenhagen, Munksgaard, 1970.

22. D. W. J. Cruickshank, D. E. Pilling, A. Bujosa, F. M. Lovell, and M. R. Truter, In Computing Methods and the Phase Problem in X-Ray Crystal Analysis, edited by R. Pepinsky and J. M. Robertson, pp. 32-78, Pergamon Press, New York, 1961.

23. R. Guttormson and B. E. Robertson, Acta Crystallogr. B28, 2702 (1972).

24. K. N. Trueblood, E. Goldish, and J. Donohue, Acta Crystallogr. 14, 1009 (1961).

25. P. Coppens and G. M. J. Schmidt, Acta Crystallogr. 18, 62,654 (1965).

26. J. V. Barve and L. M. Pant, Acta Crystallogr. B27, 1158 (1971).

27. E. Benedetti, P. Corradini, and C. Pedone, Acta Crystallogr. B26, 493 (1970).

28. B. Greenburg and B. Post, Acta Crystallogr. B24, 918 (1968).

29. E. Adman and T. N. Margulis, J. Am. Chem. Soc. 90, 4517 (1968).

30. I. L. Karle, J. Karle, and K. Britts, J. Am. Chem. Soc. 88, 2918 (1966).

31. L. S. Bartell and R. A. Bonham, J. Chem. Phys. 27, 1414 (1957).

32. W. E. Scott and G. F. Richards, J. Org. Chem. 36, 63 (1971).

33. K. Sasvari and M. Low, Acta Crystallogr. 19, 840 (1965).

34. R. A. Wohl, Chimia 18, 219 (1964).

35. S. T. Rao and M. Sundaralingam, Acta Crystallogr. B25, 2509 (1969).

36. G. Zweifel and H. C. Brown, J. Am. Chem. Soc. 86, 393 (1964).

37. R. J. Abraham, F. H. Bottom, M. A. Cooper, J. R. Salmon, and D. Whittaker, Org. Mag. Res. I, 51 (1969).

Figure 1 Bond-lengths and bond-angles in cis-PNB. The estimated standard deviation for C-C and C-N bonds is 0.008Å; for C-O bonds, 0.007Å; and for the C-N bond, 0.006Å. The molecule is shown in orthogonal projection in the (010) plane.

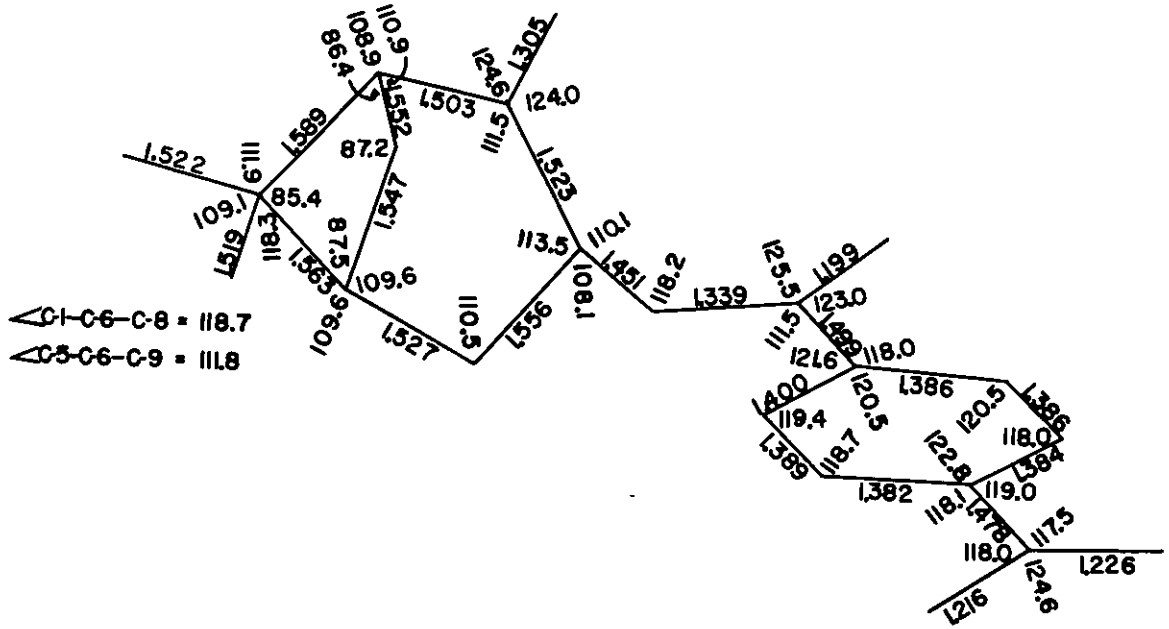

∠C-1-C-6-C-8 = 118.7
∠C-5-C-6-C-9 = 111.8
1.522
109.1
85.4
111.9
1.589
118.3
1.519
1.563
109.6
87.5
1.547
87.2
1.552
86.4
108.9
110.9
109.6
1.527
110.5
1.556
108.1
113.5
110.1
1.451
118.2
1.339
1.503
124.6
1.305
124.0
111.5
1.523
125.5
1.199
123.0
111.5
1.488
118.0
121.6
1.386
1.400
120.5
119.4
118.7
1.389
1.382
118.1
1.478
118.0
124.6
1.216
117.5
1.226
1.386
120.5
118.0
1.384
119.0
122.8

Figure 2 Dihedral angles $\delta(1)$, $\delta(2)$, and $\delta(3)$ formed by the boat and chair conformations of the pinane bicyclic ring system.

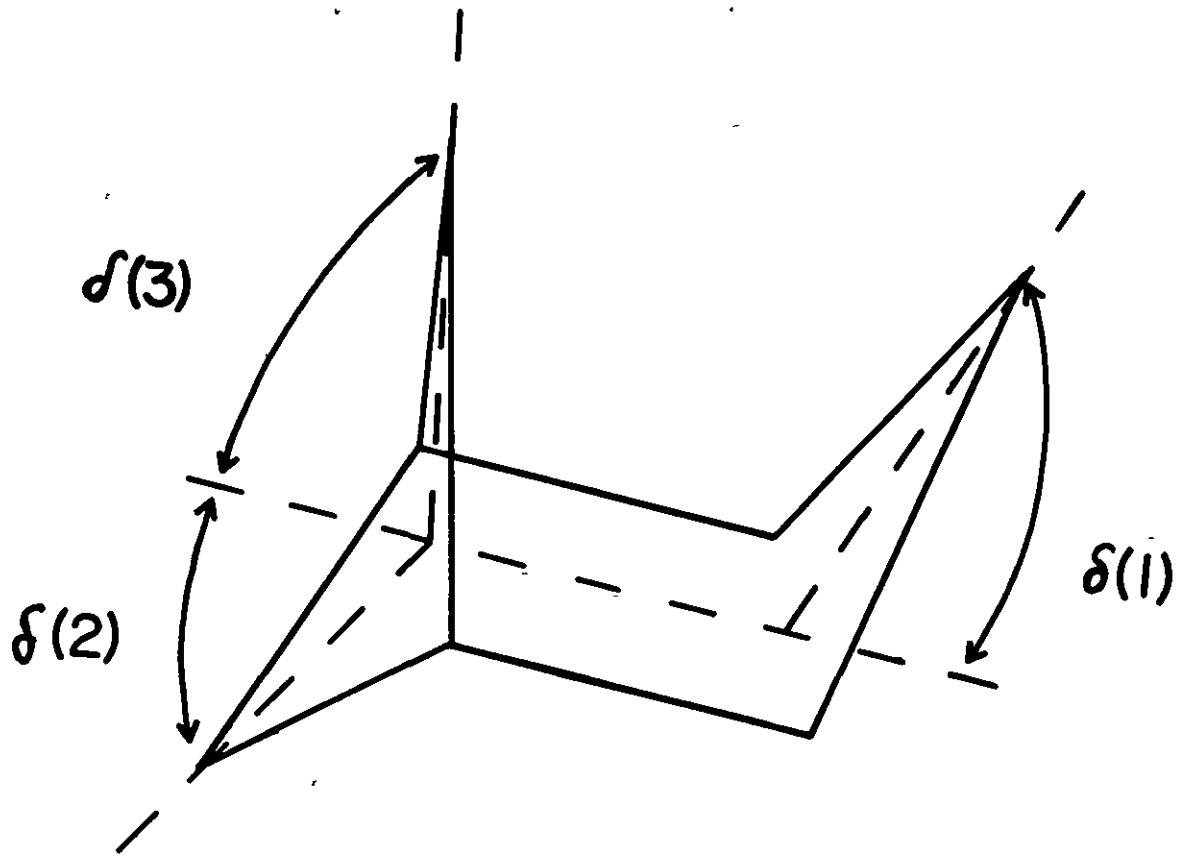
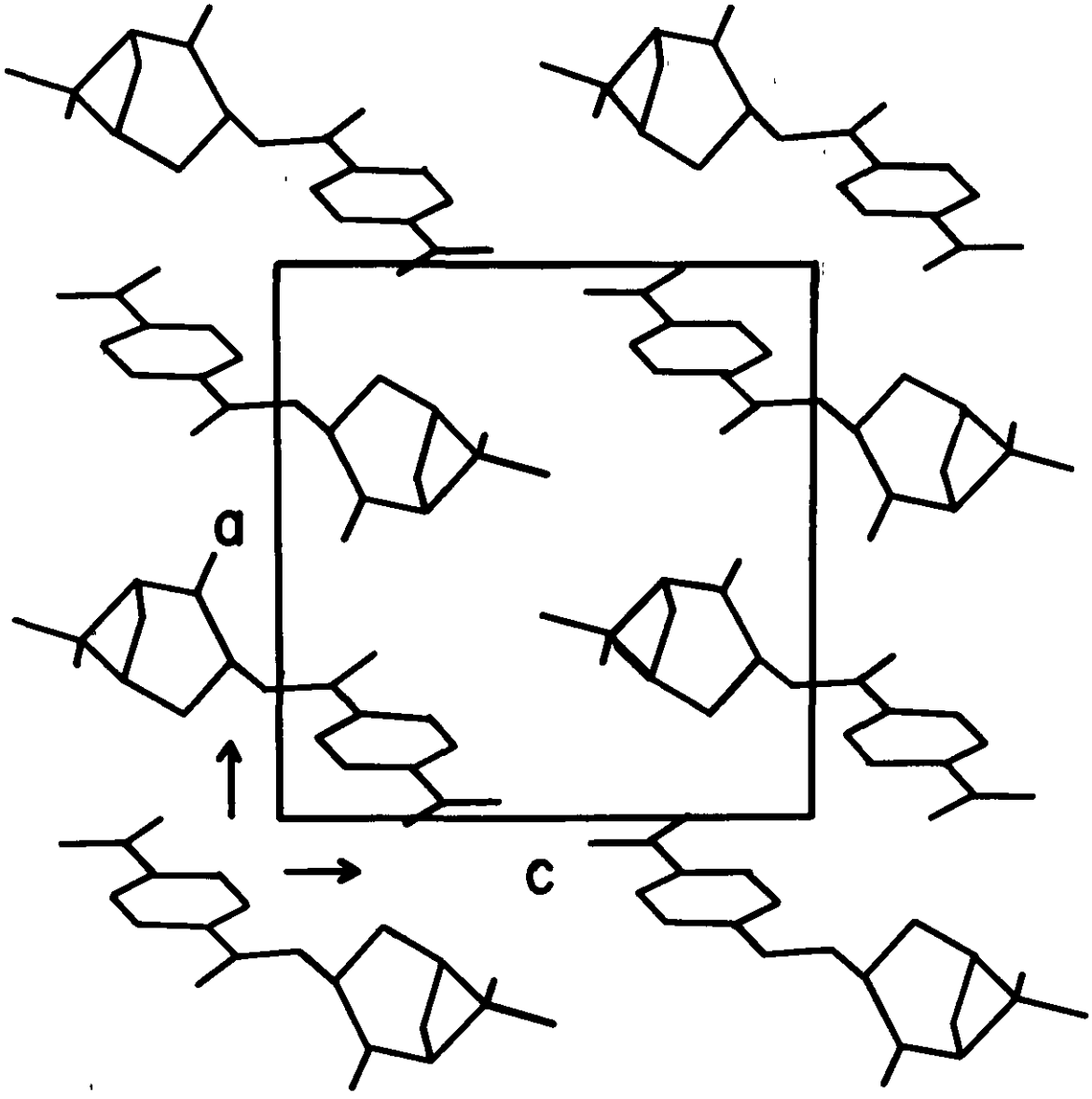
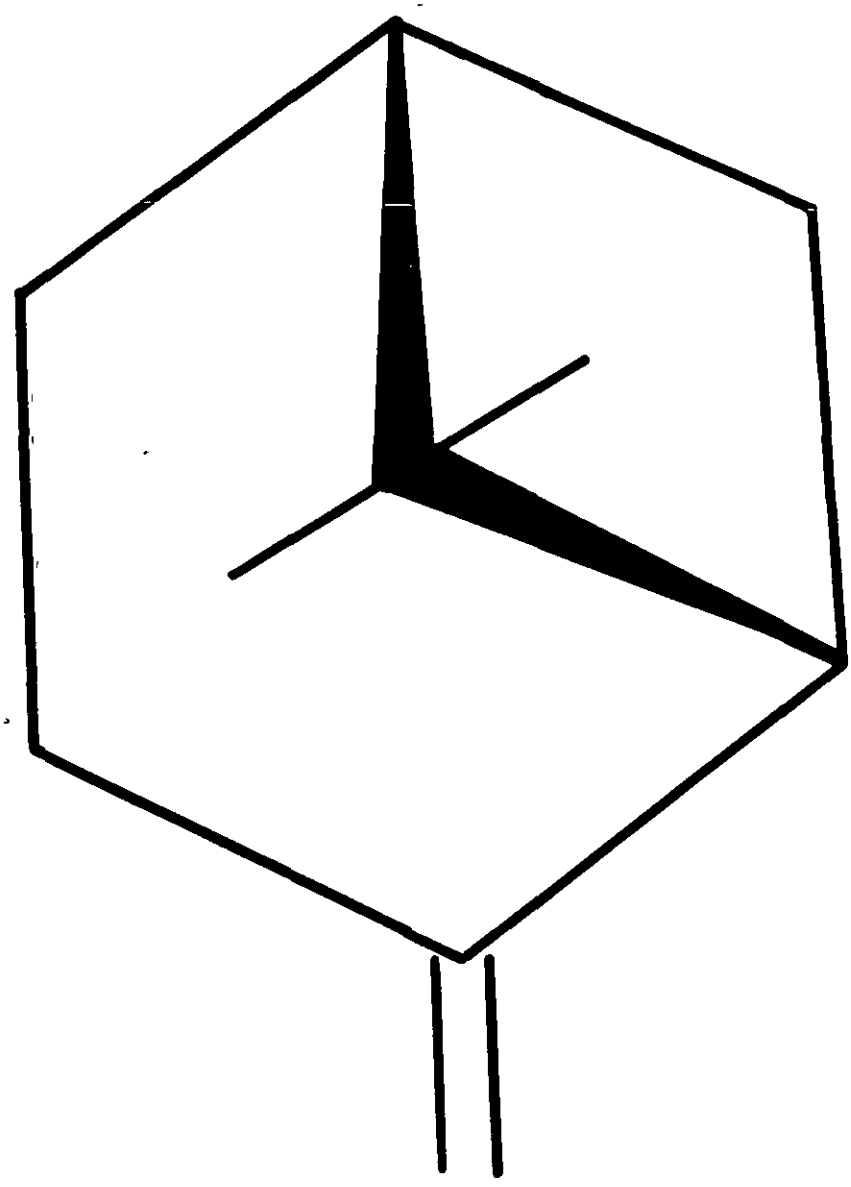
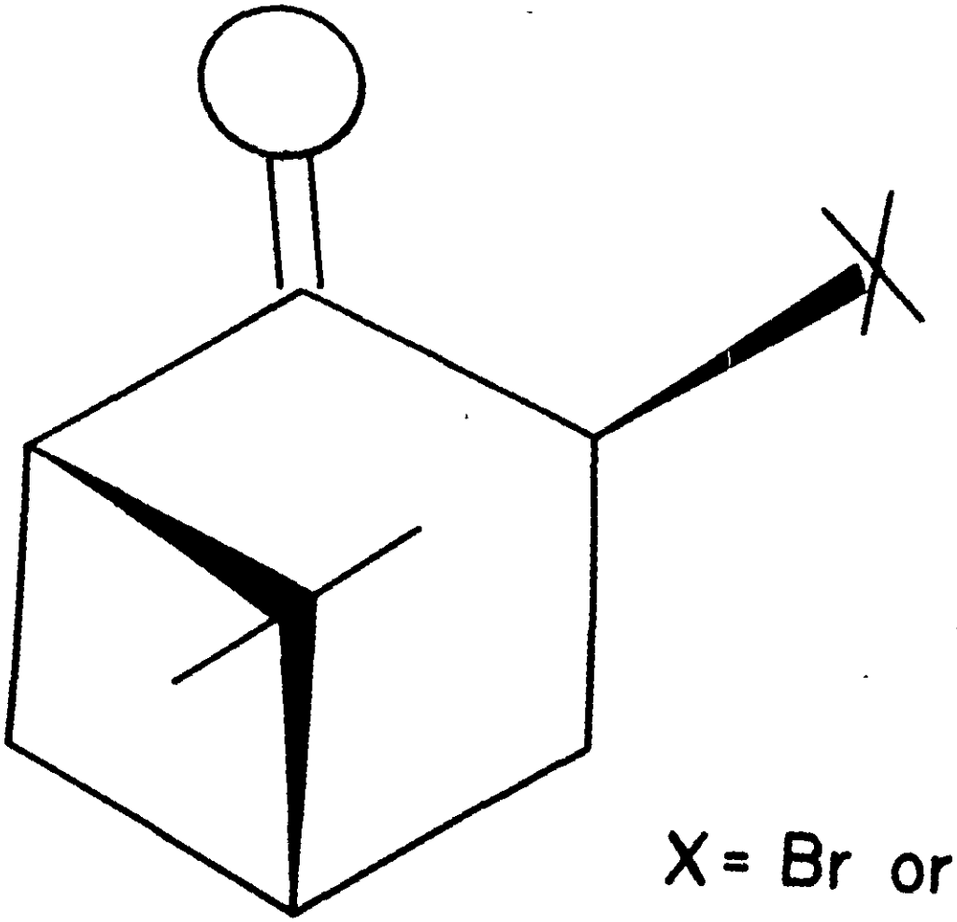

δ(3)
δ(2)
δ(1)

Figure 3 cis-PNB crystal packing shown in orthogonal projection in (010) plane.


a
c





X = Br or Cl

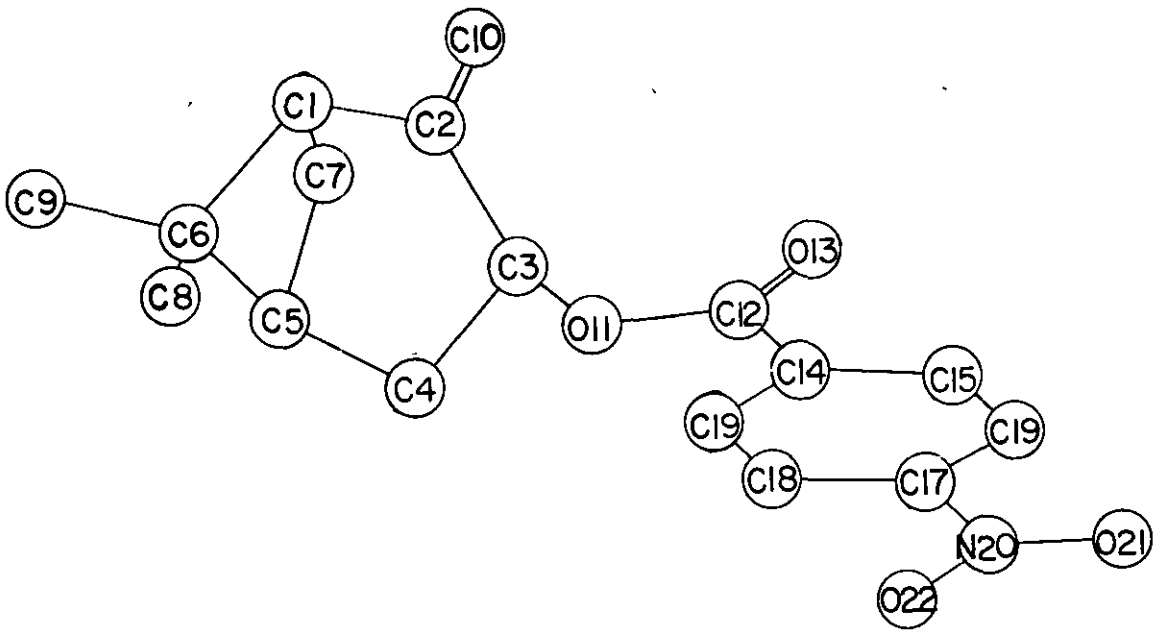

C10
C1
C2
C7
C9
C6
C8
C5
C4
C3
O11
O13
C12
C14
C15
C19
C19
C18
C17
N20
O21
O22

Table IV

Anisotropic Temperature Factors for Carbons and Oxygens in cis-PNB ($x10^5$)

Atom	B_{11}[a]	B_{22}	B_{33}	B_{23}	B_{13}	B_{12}
C-1	355	882	371	-126	234	6
C-2	305	689	389	-138	23	-114
C-3	289	482	331	44	158	108
C-4	317	776	344	46	206	6
C-5	443	422	290	154	151	149
C-6	329	643	382	287	137	190
C-7	406	662	416	302	280	-215
C-8	540	863	319	-169	-18	479
C-9	354	1521	419	-86	327	517
C-10	455	1172	405	-54	168	92
O-11	391	558	353	103	76	-273
C-12	280	479	422	215	140	314
O-13	589	795	411	82	44	-479
C-14	262	617	359	256	-18	97
C-15	458	583	380	-268	-26	75
C-16	365	812	250	104	115	290
C-17	241	498	407	380	216	368
C-18	287	433	423	1	68	183
C-19	457	653	281	176	183	44
N-20	270	743	405	74	136	3
O-21	615	1131	386	445	363	-351
O-22	485	920	570	212	54	-657

[a] The anisotropic temperature factor expression used in the refinement was of the form

$$\exp[-(B_{11}h^2+B_{22}k^2+B_{33}l^2+B_{23}kl+B_{13}hl+B_{12}hk)].$$

Table V

Distances of Atoms in cis-PNB From Least Squares Plane Through the Molecule[a]

Atom	Distance, Å
C-1	1.035
C-2	1.187
C-3	-0.141
C-4	-1.374
C-5	-1.088
C-6	0.094
C-7	-0.127
C-8	0.184
C-9	0.191
O-11	-0.154
C-12	0.001
O-13	0.134
C-14	0.011
C-15	0.007
C-16	0.018
C-17	0.022
C-18	0.003
C-19	0.003
N-20	-0.002
O-21	0.102
O-22	-0.106

[a] The equation of the least squares plane is $9.627X - 2.788Y + 0.304Z = 1.663$. C-10 was not included in the least squares plane calculation.

K	FO	FC	ALPHA
	D L	0	
2	217	229	150 64
4	283	288	21 59
6	301	310	177 90
8	154	138	5 12
	E L	0	
2	187	236	12 10
4	56	60	102 90
6	133	137	2 1 5
8	183	35	148 90
1	25	26	61 88
3	82	72	170 25
	Z L	0	
0	110	102	180 00
1	197	236	26 26
2	120	127	12 87
4	61	55	314 77
6	115	119	110 20
8	134	140	65 19
1	26	25	236 36
3	47	48	240 83
	I L	0	
0	105	9	0 U
1	216	626	355 96
2	216	149	335 77
3	294	215	2 3 78
4	82	77	2 4 46
5	304	313	51 30
6	51	55	329 46
7	40	47	230 82
8	11	16	242 72
	K L	0	
0	292	291	0 0
1	47	344	245 58
2	63	58	126 72
3	92	58	261 90
4	98	17	35 65
5	105	75	55 59
6	120	117	271 72
7	105	75	128 04
8	34	19	120 39
	L L	0	
0	40	63	0 0
1	451	459	217 92
2	319	309	38 55
3	196	178	24 72
4	159	136	113 93
5	93	44	151 72
6	73	87	97 86
7	162	144	41 58
8	127	114	185 72
	M L	0	
1	102	78	336 05
2	365	334	353 70
3	92	46	351 70
4	76	56	224 44
5	124	112	349 52
6	51	47	74 29
7	65	61	24 73
	N L	0	
0	111	37	0 0
1	114	85	225 35
2	326	276	305 65
3	155	137	291 33
	O L	0	
0	57	44	0 0
1	71	75	269 58
2	250	231	136 95
3	87	85	21 97
	P L	0	
0	40	34	180 00
2	49	58	299 03
3	279	293	3 30
	Q L	0	
0	51	51	0 0
1	37	39	262 33
2	54	59	236 06
3	63	81	351 55
	R L	0	
0	0	46	0 0
1	0*	39	349 39
2	66	71	97 93
3	57	31	204 36
	S L	0	
0	0	2	180 00
1	55	56	331 56
2	77	83	203 56
3	57	64	219 59
	T L	0	
0	45	50	180 00
1	0	21	90 00
2	64	82	229 03
	U L	0	
0	266	242	0 0
1	340	366	88 29
2	392	415	229 11
3	195	200	177 92
4	115	112	261 76
5	50	59	19 01
6	182	170	223 44
7	52	55	157 63
8	32	34	80 16
	V L	1	
1	257	256	252 71
2	100	99	118 39
3	159	153	200 19
4	168	168	329 47
5	55	65	298 91
6	183	133	311 82
7	49	57	76 74
8	19	24	321 60
	W L	1	
1	460	415	315 03
2	46	52	247 45
3	283	276	270 73
4	80	88	167 81
5	67	74	346 75
6	47	114	173 66
7	57	61	171 69
8	63	52	215 36
	Z L	1	

K	FO	FC	ALPHA
1	624	557	314 76
2	349	121	186 10
3	216	187	332 16
4	148	194	221 22
5	42	71	316 92
6	119	153	346 31
7	76	84	166 03
8	31	55	306 35
	A L	1	
0	36	49	180 00
1	0	20	170 42
2	0	15	246 39
3	243	259	181 95
	B L	1	
0	63	67	0 0
1	518	541	227 73
2	49	54	42 79
3	81	69	208 74
4	46	37	41 79
5	0	17	71 69
6	41	67	76 67
7	0	18	317 06
	C L	1	
0	59	61	0 0
1	67	63	40 25
2	32	40	280 79
3	105	112	183 17
	D L	1	
0	74	85	0 0
1	197	133	163 72
	E L	1	
0	274	269	180 00
1	249	254	247 76
2	140	134	189 25
3	227	227	337 77
4	101	102	227 48
5	43	54	52 87
6	58	55	106 95
7	220	216	57 26
8	29	28	151 40
	F L	1	
0	42	434	180 00
1	145	142	46 87
2	180	160	203 66
3	202	269	141 74
4	88	86	52 46
5	43	27	111 53
6	122	151	244 18
7	74	73	259 89
8	43	37	171 95
	G L	1	
0	42	435	180 00
1	437	430	156 55
2	116	123	347 40
3	142	148	51 67
4	22	16	121 30
5	51	102	171 04
6	137	145	47 74
7	58	51	94 84
8	72	73	132 57
	H L	1	
0	84	74	0 0
1	362	220	58 99
2	290	284	189 27
3	370	461	165 98
4	42	44	12 33
5	114	124	106 95
6	43	50	288 08
7	76	91	257 71
8	110	106	24 29
	I L	1	
0	114	116	180 00
1	152	257	29 13
2	294	267	51 19
3	173	167	185 96
4	108	165	155 97
5	34	44	2 50
6	104	102	49 49
7	68	71	191 90
8	51	45	25 05
	J L+	2	
0	35	41	180 00
1	88	69	325 17
2	33	30	321 49
3	42	60	294 58
	K L	2	
1	84	77	319 20
2	49	85	188 31
3	81	48	318 72
	L L	2	
0	22	23	0 0
1	37	47	253 52
	M L	2	
0	43	72	0 0
1	95	52	318 47
2	31	45	25 56
	N L	2	
0	45	82	180 00
1	406	404	6 19
2	119	115	126 86
3	144	156	358 76
4	41	37	43 13
5	194	222	355 82
6	103	109	327 07
7	93	109	315 08
8	0	13	70 22
	O L	2	
0	155	176	0 0
1	270	277	101 06
2	190	115	269 07
3	119	118	3 73
4	43	48	154 62
5	149	152	236 69
6	55	57	155 31
7	66	76	248 32
8	30	32	196 68
	P L	2	
0	198	206	0 0
1	169	161	212 04
2	26	10	9 10
3	46	49	230 36
4	130	135	183 92

K	FO	FC	ALPHA
5	126	137	45 18
6	114	124	127 75
7	116	136	342 51
8	56	61	77 21
	Q L	2	
0	224	224	180 00
1	176	159	41 56
2	314	283	165 17
3	109	102	322 88
4	78	87	172 65
5	0	21	100 26
6	41	51	266 90
	R L	2	
0	43	42	0 0
1	118	141	227 73
2	49	54	42 79
3	81	69	208 74
4	46	37	41 79
5	0	17	71 69
6	41	67	76 67
7	0	18	317 06
	S L	2	
0	195	201	180 00
1	92	89	107 58
2	75	58	153 42
3	125	148	174 46
4	183	164	243 19
5	0*	27	323 46
6	45	44	146 92
7	46	36	245 36
	T L	2	
0	200	188	180 00
1	238	219	188 09
2	45	57	236 35
3	102	108	14 50
	U L	2	
0	163	158	0 0
1	105	100	128 86
2	154	157	32 25
3	34	43	234 43
	V L	2	
1	117	97	258 80
2	52	66	166 06
3	36	39	43 13
	W L+	2	
0	0	28	0 0
1	116	112	245 74
2	90	95	150 56
3	101	97	234 18
	X L+	2	
0	90	91	180 00
1	109	108	325 97
2	47	58	169 77
3	60	72	252 80
	Y L+	2	
0	38	41	180 00
1	55	57	44 80
2	0*	28	67 57
3	145	138	164 28
	Z L+	2	
0	139	155	180 00
1	38	36	11 10
2	62	76	62 51
	A L	2	
0	0	18	180 00
1	69	58	270 62
2	65	67	6 72
3	0	16	278 96
	B L	2	
0	139	132	180 00
1	55	36	123 41
2	38	54	209 13
	C L	2	
0	35	41	180 00
1	88	69	325 17
2	33	30	321 49
3	42	60	294 58
	D L	2	
0	288	303	0 0
1	152	169	193 52
2	249	209	48 39
3	36	26	137 55
4	233	216	110 97
5	158	161	149 64
6	126	117	93 62
7	65	77	282 91
8	81	75	67 37
	E L	2	
0	274	240	180 00
1	264	254	165 93
2	264	259	66 87
3	56	41	73 35
4	184	189	43 80
5	313	313	236 45
6	92	100	21 31
7	86	101	101 73
8	92	47	268 36
	F L	2	
0	97	105	180 00
1	98	106	210 98
2	214	175	159 01
3	43	55	197 20
4	175	171	39 58
5	135	136	198 28
6	54	25	175 15
7	43	41	344 28
8	43	56	295 51
	G L	2	
0	186	98	0 0
1	117	112	100 42
2	489	485	208 26
3	213	215	88 84
4	124	121	18 28
5	92	92	330 41
6	0	40	135 04
7	56	60	49 60
8	186	160	27 27
	H L	2	
0	45	52	180 00
1	359	300	121 90
2	22	30	244 92
3	36	38	201 10
4	30	23	23 15
5	244	237	214 64

K	FO	FC	ALPHA
2	50	56	157 44
3	11	17	306 49
	I L	2	
0	150	158	180 00
1	127	136	148 84
	J L	2	
0	0*	45	180 00
1	0*	22	25 86
2	36	44	215 94
3	0*	33	304 04
	K L	2	
0	91	106	180 00
1	55	58	26 40
2	36	57	344 02
	L L	2	
0	0*	4	180 00
1	35	38	197 04
2	29	34	47 10
3	48	59	30 16
	M L	2	
0	44	46	106 33
1	54	42	47 27
2	37	53	16 49
	N L	2	
0	0*	3	180 00
1	48	59	30 16
	O L	2	
0	267	293	180 00
1	218	214	71 90
2	302	264	116 54
3	0	27	272 15
4	82	78	80 12
5	72	75	352 4
6	31	48	349 81
7	35	32	352 59
8	80	84	118 77
	P L	2	
0	364	363	0 0
1	100	95	91 16
2	194	146	73 65
3	234	232	75 13
4	267	243	29 60
5	44	49	154 10
	Q L	2	
0	143	144	0 0
1	107	100	243 24
2	124	115	101 59
3	118	104	330 51
4	103	108	16 80
5	0	34	276 35
6	59	59	275 67
7	87	78	324 77
	R L	2	
0	34	35	180 00
1	117	123	64 66
2	243	247	172 49
3	54	49	182 57
4	177	180	182 95
5	43	45	34 41
6	50	53	256 86
7	42	49	97 61
	S L	2	
0	90	74	180 00
1	203	192	279 63
2	284	270	209 11
3	169	159	353 59
4	67	66	289 10
5	77	67	353 57
6	84	85	195 62
7	91	78	56 17
	T L	2	
1	79	78	66 53
2	43	35	62 71
3	79	73	221 29
4	65	67	32 32
5	69	93	278 69
6	70	78	5 52
7	27	27	176 49
	U L	2	
0	135	150	180 00
1	19	16	93 94
2	104	103	321 11
3	124	111	124 14
4	84	82	201 25
5	132	119	34 42
6	52	51	74 80
7	47	37	78 10
	V L	2	
0	60	54	0 0
1	244	222	60 23
2	67	68	110 44
3	143	143	154 54
4	94	90	104 75
5	145	135	326 41
6	44	45	149 12
	W L	2	
0	66	55	0 0
1	106	45	23 83
2	140	115	126 86
3	64	66	327 45
	X L	2	
1	128	122	184 22
2	95	97	46 33
3	136	141	29 55
	Y L	2	
0	97	93	0 0
1	226	209	76 22
2	107	296	320 11
3	92	91	75 07
4	147	151	225 14
5	94	95	73 42
6	31	35	13 3
7	59	54	44 84
	Z L	2	
0	215	213	180 00
1	131	122	254 64
2	32	38	163 37
3	44	51	277 51
4	43	45	177 46
5	46	47	295 10
6	126	126	308 47
7	24	26	94 07
	A L	2	
0	413	449	0 0
1	94	93	242 15
2	113	111	259 43
3	31	21	236 51
4	103	93	324 52
5	32	26	35 65
6	38	34	87 40
7	70	60	292 71
	B L	2	

K	FC	FC	ALPHA
0	97	96	0 0
1	139	138	231 57
2	96	89	235 56
3	72	69	163 93
4	99	98	3 94
5	117	111	87 96
6	80	75	208 72
7	62	67	29 06
	C L+	4	
0	97	95	180 00
1	46	58	8 67
2	86	83	244 39
3	146	148	28 31
4	42	46	172 77
5	48	39	274 39
6	65	56	207 91
7	75	51	163 89
	D L+	4	
0	112	107	43 05
1	124	118	163 72
2	106	97	337 33
3	83	85	210 71
4	75	72	281 97
5	70	106	219 59
	E L+	4	
0	84	79	180 00
1	227	193	85 97
2	41	42	332 40
3	123	112	24 21
	F L+	4	
1	57	58	90 47
2	141	136	61 03
3	75	76	34 46
	G L+	4	
0	94	96	180 00
1	294	302	78 64
2	171	174	73 06
3	151	146	359 14
4	59	55	47 26
5	46	70	340 94
6	42	50	73 41
7	0*	21	231 03
8	72	64	205 08
	H L+	4	
0	97	100	0 0
1	274	282	18 10
2	125	130	257 46
3	290	282	22 37
4	129	139	245 43
5	125	125	346 44
6	94	102	266 46
7	31	30	181 29
8	44	45	296 03
	I L+	4	
0	45	48	48 68
1	41	48	48 68
2	106	125	288 79
3	37	37	286 31
	J L+	4	
0	67	73	180 00
1	48	50	196 07
2	0*	27	94 46
3	15	16	17 39
	K L+	4	
1	113	120	125 59
2	35	36	301 91
	L L+	4	
0	65	76	0 0
1	51	65	162 88
2	48	55	57 39
	M L+	4	
0	55	52	0 0
1	169	169	132 64
2	102	108	178 57
3	87	88	134 95
	N L+	4	
0	86	86	180 00
1	63	63	180 73
2	37	41	329 61
3	0*	26	169 01
	O L+	4	
0	195	194	0 0
1	135	138	248 76
2	147	143	87 71
3	57	57	297 29
4	32	38	101 46
5	43	44	98 91
6	61	74	2 73
7	28	22	265 94
8	19	16	78 76
	P L+	4	
0	45	47	0 0
1	59	50	94
2	209	197	280 57
3	32	32	268 20
4	93	87	223 56
5	70	78	222 61
6	88	88	152 97
7	39	27	242 05
	Q L+	4	
0	85	77	0 0
1	161	151	335 87
2	99	62	80 77
3	75	79	250 31
4	139	141	45 66
5	46	47	78 02
7	106	86	135 25
	R L+	4	
0	100	103	0 0
1	41	39	180 83
2	81	81	105 55
3	61	58	173 68
4	105	100	332 82
5	92	90	74 65
6	116	96	280 91
7	45	33	45 91
	S L+	4	
0	28	18	0 0
1	234	225	263 46
2	74	75	221 93
3	71	66	224 86
4	26	28	267 29
5	129	135	116 05
6	35	33	150 96
7	53	27	342 20
	T L+	4	
0	48	45	0 0
1	40	49	231 62
2	93	86	139 43
3	124	123	306 87
4	91	75	330 54
5	123	99	254 25
6	138	103	276 91
7	60	59	243 94
	U L+	4	
0	131	126	0 0
1	86	82	159 41
2	145	143	353 97
3	88	82	27 92
4	72	54	186 33
5	42	52	142 62
6	66	78	26 18
7	22	45	29 64
	V L+	4	
0	0*	16	180 00
1	37	43	316 45
2	36	42	179 88
3	0*	26	114 66
	W L+	4	
0	53	52	0 0
1	169	169	132 64
2	102	108	178 57
3	87	88	134 95
	X L+	4	
0	86	75	0 0
1	0*	25	58 84
2	0*	32	276 47
3	236	187	166 00
	Y L+	4	
0	16	33	0 0
1	175	99	288 88
2	32	109	77 59
	Z L+	4	
0	59	59	0 0
1	46	46	237 12
2	0	14	23 0
3	0*	17	133 36
	A L+	5	
0	23	53	33 90
3	66	52	333 99
	B L	5	
0	33	47	0 0
1	41	39	180 83
2	81	105	55
3	61	58	173 68
4	105	100	332 82
5	92	90	74 65
6	116	96	280 91
7	45	33	45 91
	C L	5	
0	28	16	0 0
1	234	225	263 46
2	74	75	221 93
3	71	66	224 86
4	26	28	267 29
5	129	135	116 05
6	35	33	150 96
7	53	27	342 20
	D L	5	
0	48	55	0 0
1	40	49	231 62
2	93	86	139 43
3	124	123	306 87
4	91	75	330 54
5	123	99	254 25
6	138	103	276 91
7	60	59	243 94
	E L	5	
0	131	126	0 0
1	86	82	159 41
2	145	143	353 97
3	88	82	27 92
4	72	54	186 33
5	42	52	142 62
6	66	78	26 18
7	22	45	29 64
	F L	5	
0	0*	16	180 00
1	37	43	316 45
2	36	42	179 88
3	0*	26	114 66
	G L	5	
0	53	52	0 0
1	169	169	132 64
2	102	108	178 57
3	87	88	134 95
	H L	5	
0	86	75	0 0
1	0*	25	58 84
2	0*	32	276 47
3	236	187	166 00
	I L	5	
0	16	33	0 0
1	175	99	288 88
2	32	109	77 59
	J L	5	
0	59	59	0 0
1	46	46	237 12
2	0	14	23 0
3	0*	17	133 36
	K L	5	
0	23	53	33 90
3	66	52	333 99
	L L	5	
0	33	47	0 0
1	41	39	180 83
2	81	105	55
3	61	58	173 68
4	105	100	332 82
5	92	90	74 65
6	116	96	280 91
7	45	33	45 91
	M L	5	
0	28	16	0 0
1	234	225	263 46
2	74	75	221 93
3	71	66	224 86
4	26	28	267 29
5	129	135	116 05
6	35	33	150 96
7	53	27	342 20
	N L	5	
0	48	55	0 0
1	40	49	231 62
2	93	86	139 43
3	124	123	306 87
4	91	75	330 54
5	123	99	254 25
6	138	103	276 91
7	60	59	243 94
	O L	5	
0	131	126	0 0
1	86	82	159 41
2	145	143	353 97
3	88	82	27 92
4	72	54	186 33
5	42	52	142 62
6	66	78	26 18
7	22	45	29 64
	P L	5	
0	0*	16	180 00
1	37	43	316 45
2	36	42	179 88
3	0*	26	114 66
	Q L	5	
0	53	52	0 0
1	169	169	132 64
2	102	108	178 57
3	87	88	134 95
	R L	5	
0	86	75	0 0
1	0*	25	58 84
2	0*	32	276 47
3	236	187	166 00
	S L	5	
0	16	33	0 0
1	175	99	288 88
2	32	109	77 59
	T L	5	
0	59	59	0 0
1	46	46	237 12
2	0	14	23 0
3	0*	17	133 36
	U L	5	
0	23	53	33 90
3	66	52	333 99
	V L	5	
0	33	47	0 0
1	41	39	180 83
2	81	105	55
3	61	58	173 68
4	105	100	332 82
5	92	90	74 65
6	116	96	280 91
7	45	33	45 91
	W L	5	
0	28	16	0 0
1	234	225	263 46
2	74	75	221 93
3	71	66	224 86
4	26	28	267 29
5	129	135	116 05
6	35	33	150 96
7	53	27	342 20
	X L	5	
0	48	55	0 0
1	40	49	231 62
2	93	86	139 43
3	124	123	306 87
4	91	75	330 54
5	123	99	254 25
6	138	103	276 91
7	60	59	243 94
	Y L	5	
0	131	126	0 0
1	86	82	159 41
2	145	143	353 97
3	88	82	27 92
4	72	54	186 33
5	42	52	142 62
6	66	78	26 18
7	22	45	29 64
	Z L	5	
0	0*	16	180 00
1	37	43	316 45
2	36	42	179 88
3	0*	26	114 66
	A L	6	
0	53	52	0 0
1	169	169	132 64
2	102	108	178 57
3	87	88	134 95
	B L	6	
0	86	75	0 0
1	0*	25	58 84
2	0*	32	276 47

K	FO	FC	ALPHA
H= -12, L= 5			
1	73	37	109.41
2	35	47	112.01
H= 0, L= 6			
0	73	66	180.00
1	56	67	147.88
2	44	56	135.35
3	38	42	129.73
4	221	251	130.51
5	76	87	121.06
6	55	65	126.18
7	39	27	281.83
H= 1, L= 6			
0	171	68	0.0
1	61	59	221.05
2	77	64	250.75
3	61	55	161.76
4	204	220	54.24
5	64	75	124.97
6	0+	13	213.83
7	0+	26	329.51
H= -1, L= 6			
0	171	166	180.00
1	54	67	21.48
2	55	67	150.88
3	46	48	177.76
4	114	105	240.79
5	85	95	228.25
6	78	96	158.56
7	96	99	171.01
H= 2, L= 6			
0	31	43	180.00
1	240	244	172.70
2	82	86	164.25
3	70	77	104.32
4	46	71	53.72
5	246	248	231.32
6	46	43	191.18
7	47	50	66.43
H= -2, L= 6			
0	285	285	180.00
1	241	245	201.52
2	101	110	235.48
3	49	53	121.27
4	84	94	264.80
5	100	120	296.10
6	72	98	8.91
H= 3, L= 6			
0	150	151	180.00
1	116	137	140.77
2	46	50	45.43
3	105	111	178.27
4	113	116	201.50
5	154	151	170.01
6	0+	10	25.21
7	61	43	226.46
H= -3, L= 6			
0	37	30	180.00
1	109	106	24.40
2	108	106	222.22
3	248	254	175.17
4	68	67	74.07
5	56	64	151.53
6	35	40	127.40
7	57	56	232.69
H= 4, L= 6			
0	0+	13	0.0
1	43	50	332.72
2	87	92	174.06
3	75	79	199.13
4	25	24	327.42
5	114	104	116.02
6	0+	21	275.04
7	49	86	257.23
H= -4, L= 6			
0	125	122	180.00
1	159	151	308.32
2	135	133	338.21
3	160	152	218.12
4	45	52	167.04
5	0+	2	40.00
7	66	68	209.88
H= 5, L= 6			
0	66	67	0.0
1	44	39	150.47
2	117	119	187.32
3	149	159	159.24
4	157	38	51.76
5	87	76	75.93
6	127	108	220.72
7	72	58	194.09
H= -5, L= 6			
0	64	59	0.0
1	117	109	235.16
2	121	121	60.29
3	93	97	279.68
4	66	49	41.35
5	19	50	138.45
6	46	46	141.33
7	47	39	194.32
H= 6, L= 6			
0	121	121	0.0
1	187	182	175.74
2	120	115	42.69
3	77	73	155.16
4	154	128	54.55
5	38	25	224.05
6	87	65	147.66
H= -6, L= 6			
0	146	144	180.00
1	97	69	175.46
2	131	153	195.20
3	61	77	40.16
4	46	65	5.78
H= 7, L= 6			
0	39	37	0.0
1	50	45	317.04
2	67	80	15.43
3	107	101	181.25
H= -7, L= 6			
0	0+	9	180.00
1	76	76	329.55
2	123	126	104.32
3	0+	28	262.28
4	42	46	96.00
5	165	175	54.24
6	46	51	86.54
H= -8, L= 6			
0	41	33	180.00
1	0+	39	327.53
2	40	20	55.43
3	141	104	174.34
H= -9, L= 6			
0	171	166	180.00
1	54	67	21.48
2	55	67	150.88
3	46	48	177.76
4	114	105	240.79
5	85	95	228.25
6	78	96	158.56
7	96	99	171.01
H= 2, L= 6			
0	31	43	180.00
1	240	244	172.70
2	82	86	164.25
3	70	77	104.32
4	46	71	53.72
5	246	248	231.32
6	46	43	191.18
7	47	50	66.43
H= -7, L= 6			
0	285	285	180.00
1	241	245	201.52
2	101	110	235.48
3	49	53	121.27
4	84	94	264.80
5	100	120	296.10
6	72	98	8.91
H= 3, L= 6			
0	150	151	180.00
1	116	137	140.77
2	46	50	45.43
3	105	111	178.27
4	113	116	201.50
5	154	151	170.01
6	0+	10	25.21
7	61	43	226.46
H= -3, L= 6			
0	37	30	180.00
1	109	106	24.40
2	108	106	222.22
3	248	254	175.17
4	68	67	74.07
5	56	64	151.53
6	35	40	127.40
7	57	56	232.69
H= 4, L= 6			
0	0+	13	0.0
1	43	50	332.72
2	87	92	174.06
3	75	79	199.13
4	25	24	327.42
5	114	104	116.02
6	0+	21	275.04
7	49	86	257.23
H= -4, L= 6			
0	125	122	180.00
1	159	151	308.32
2	135	133	338.21
3	160	152	218.12
4	45	52	167.04
5	0+	2	40.00
7	66	68	209.88
H= 5, L= 6			
0	66	67	0.0
1	44	39	150.47
2	117	119	187.32
3	149	159	159.24
4	157	38	51.76
5	87	76	75.93
6	127	108	220.72
7	72	58	194.09
H= -5, L= 6			
0	64	59	0.0
1	117	109	235.16
2	121	121	60.29
3	93	97	279.68
4	66	49	41.35
5	19	50	138.45
6	46	46	141.33
7	47	39	194.32
H= 6, L= 6			
0	121	121	0.0
1	187	182	175.74
2	120	115	42.69
3	77	73	155.16
4	154	128	54.55
5	38	25	224.05
6	87	65	147.66
H= -6, L= 6			
0	146	144	180.00
1	97	69	175.46
2	131	153	195.20
3	61	77	40.16
4	46	65	5.78
H= 7, L= 6			
0	39	37	0.0
1	50	45	317.04
2	67	80	15.43
3	107	101	181.25
H= -7, L= 6			
0	0+	9	180.00
1	76	76	329.55
2	123	126	104.32
3	0+	28	262.28
4	42	46	96.00
5	165	175	54.24
6	46	51	86.54
H= -8, L= 6			
0	41	33	180.00
1	0+	39	327.53
2	40	20	55.43
3	141	104	174.34
H= -9, L= 6			
0	171	166	180.00
1	54	67	21.48
2	55	67	150.88
3	46	48	177.76
4	114	105	240.79
5	85	95	228.25
6	78	96	158.56
7	96	99	171.01
H= 2, L= 6			
0	31	43	180.00
1	240	244	172.70
2	82	86	164.25
3	70	77	104.32
4	46	71	53.72
5	246	248	231.32
6	46	43	191.18
7	47	50	66.43
H= -2