

We are indebted to H. Miyagishima and N. Hoshino for their assistance in the experimental work.

LITERATURE CITED

- (1) Smith, E. G., "The Wiswesser Line-Formula Chemical Notation," McGraw-Hill, New York, N.Y., 1968.
- (2) Granito, C. E., Becker, G. T., Roberts, S., Wiswesser, W. J., and Windlinx, K. J., "Computer-Generated Substructure Codes (Bit Screens)," *J. Chem. Doc.* 11(2), 106-10 (1971).
- (3) Ringdoc Instruction Bull. No. 7, "Peptides," 1964 (for subscribers only).
- (4) Revesz, G. S., "One-Letter Notation for Calculating Molecular Formulas and Searching Long-Chain Peptides in the *Index Chemicus* Registry System," *J. Chem. Doc.* 10(3), 212-15 (1970).
- (5) Hoffmann, E., "Use of a Modified Wiswesser Notation for the Encoding of Proteins," *Ibid.*, 9(3), 137-40 (1969).
- (6) Wiswesser, W. J., "A Protein-Spelling Alphabet," *Chem. Eng. News* 42, 4 (Sept. 14, 1964).

Cambridge Crystallographic Data Centre. IV. Preparation of "Interatomic Distances 1960-65"

F. H. ALLEN, N. W. ISAACS, OLGA KENNARD,* W. D. S. MOTHERWELL, R. C. PETTERSEN, W. G. TOWN, and D. G. WATSON**

Crystallographic Data Centre, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, England

Received August 20, 1973

The Cambridge Crystallographic Data Centre is concerned with the retrieval, evaluation, synthesis, and dissemination of structural data obtained by diffraction methods. Earlier papers in this series have described the organization of computer-based files of both bibliographic information and of evaluated numeric structural data. The present paper describes the use of a subset of the system to produce a compendium of interatomic distances including stereo diagrams of structures, numeric data, editorial comments, and a variety of indexes.

Earlier papers in this series have described the organization and information content of the bibliographic file¹ and the structural data file² of the Cambridge Crystallographic Data Centre. The process of evaluation of numeric data applied to the latter file has also been documented.³ The bibliographic information is disseminated via the Bibliographic Volumes of the series "Molecular Structures and Dimensions"⁴ (hereinafter called MSD). The computer-typesetting and other techniques used in the preparation of these volumes from our computer files will be described in Part V of this series.⁵

The present paper describes the use of both the bibliographic and structural data files in the preparation of Volume A1 of MSD: "Interatomic Distances 1960-65" (ID).⁶ This volume is a continuation of "Tables of Interatomic Distances and Configurations in Molecules and Ions,"⁷ which covered the literature until the end of 1959. It gave information on interatomic distances and, where relevant, bond angles, obtained by spectroscopy, electron, neutron, and X-ray diffraction for organic and inorganic molecules and ions, and for metals.

In our compilation the criteria for inclusion in the files^{1,2} necessarily restricts coverage to organic and organometallic compounds investigated by X-ray and neutron diffraction, while the spectacular increase in the number of studies since the early 1960's has forced us to consider the period 1960-65 in the first instance. Computer methods have been used both for the evaluation^{2,3} and presentation of results. In most cases, this has enabled us to extend the information given in the original publication by preparing stereoscopic illustrations of individual molecular structures and, where appropriate, by including torsion angles computed from the original data.

DATA BASE

The data base used in the preparation of ID was a merged subset of both bibliographic and structural data files for the chosen period. The file was in the form of card images on magnetic tape. Entries were grouped in 86 chemical classes¹ and ordered by C,H content within each class, as in the bibliographic volumes of MSD.^{4,5} Entries were numbered sequentially in each class⁴ and a card record ENTNO was added to each data entry in the form *m.n*, where *m* is the class number and *n* the serial number. In contrast to the bibliographic volumes where cross-references were provided to supplementary classes, each compound appears once only in ID, in its basic class. Finally a card type IDREF was added to the file, containing bibliographic references to studies of a given compound published outside the period 1960-65. This was generated automatically from the main bibliographic file current at that time.

ORGANIZATION

Responsibility for the data base was divided among six editors who ensured that all entries were complete, undertook any residual scientific checking, and added any additional textual material on the card types REMARK, DISORD, and ERROR² at their discretion. At a later stage, the editors chose a suitable view of the molecule for the crystallographic stereo pair, and edited the tables of numeric data. The responsibilities were allotted on the basis of chemical class, so that each editor had a relatively narrow range of compounds with which to become familiar, e.g., benzenoid compounds, natural products, etc. Frequent discussions were held to ensure editorial consistency.

* External Staff, Medical Research Council.

** To whom all enquiries should be addressed.

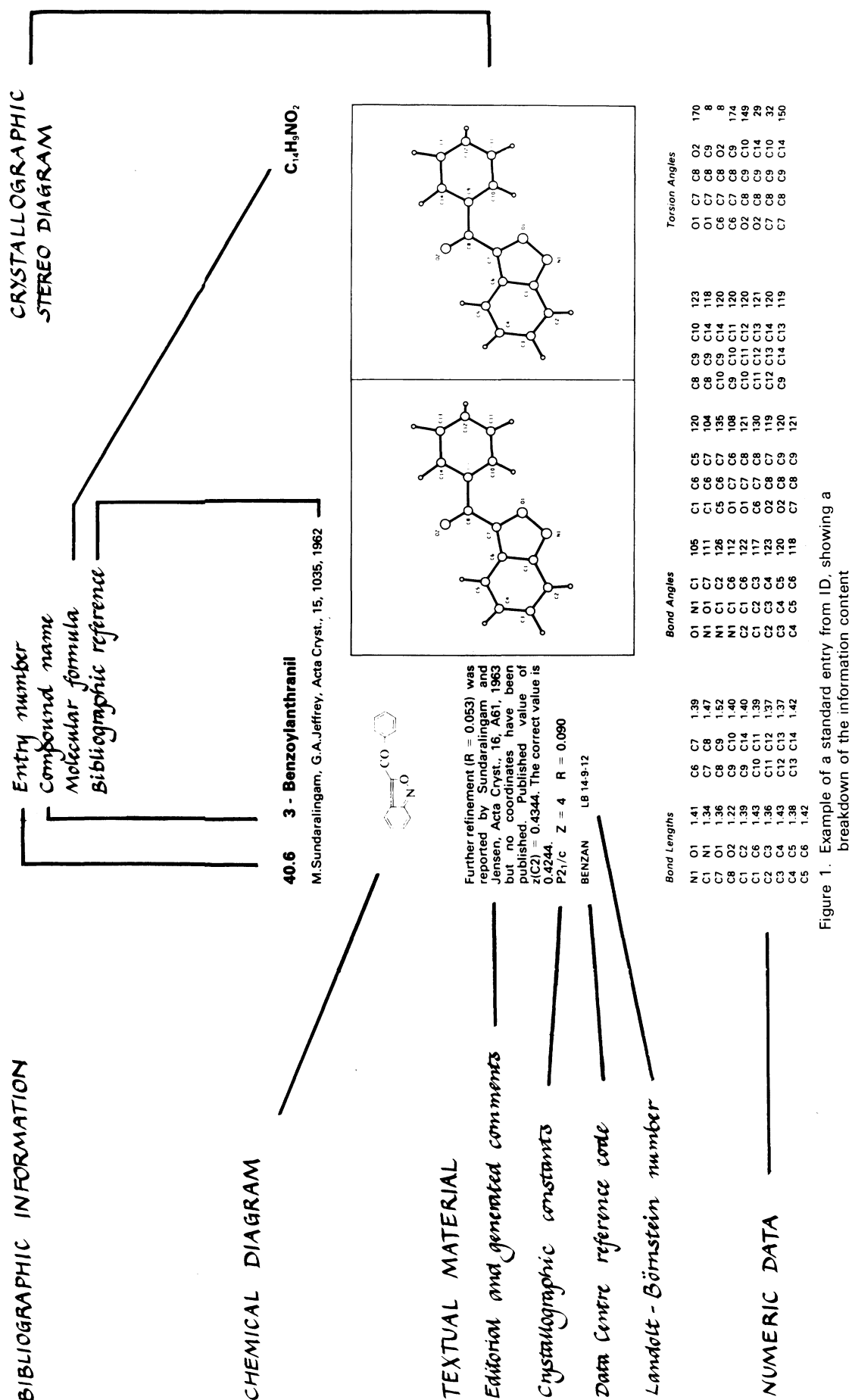


Figure 1. Example of a standard entry from ID, showing a breakdown of the information content

INFORMATION CONTENT

Figure 1 is a reproduction of a typical *standard entry* in the book. The possible categories of information are indicated; they may be broadly grouped as:

- (i) Bibliographic information
- (ii) Chemical diagram
- (iii) Textual material
- (iv) Crystallographic stereo diagram
- (v) Numeric data

Each of these will be discussed in detail below. In certain cases the standard form of an entry may be modified.

Partial entries contain information for groups (i) to (iii) only and occur in the following situations:

- (a) Studies in which atomic coordinates were not reported in the publication
- (b) Two-dimensional studies where only two coordinates were determined for each nonhydrogen atom
- (c) Structures where all, or a large proportion of, atoms were disordered
- (d) Clathrate structures
- (e) Structures where our evaluation procedures indicated incomplete or wrong connectivity, or other serious errors which could not be resolved. Such entries always contain textual comment indicating the nature of the error

Multiple entries for a given compound can occur for a variety of reasons:

- (a) Independent studies by different authors
- (b) Studies of different crystal forms
- (c) Studies at different temperatures
- (d) Studies with both X-rays and neutrons
- (e) Refinements using different data sets or different space-group models.

If these multiple entries have not been amalgamated (see below) then each entry has the form of a standard or partial entry.

Amalgamated Entries. For certain sets of multiple entries, the information has been amalgamated during page make-up into a composite, which allows for easy comparison of results. One entry is taken as the basic entry, information from additional entries being merged with it. The composite has all the information types (i) to (v), while additional entries are simply represented by bibliographic details, as shown in Figure 2.

The main section of the book contains 1305 entries. This is followed by summary tables of bond lengths arranged by atom pairs, author and formula indexes, and a permuted KWIC-type compound-name index.

BIBLIOGRAPHIC INFORMATION

This information is exactly as presented in the bibliographic volumes of MSD⁴ and was computer typeset⁵ via a Linotron 505. The information consists of:

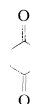
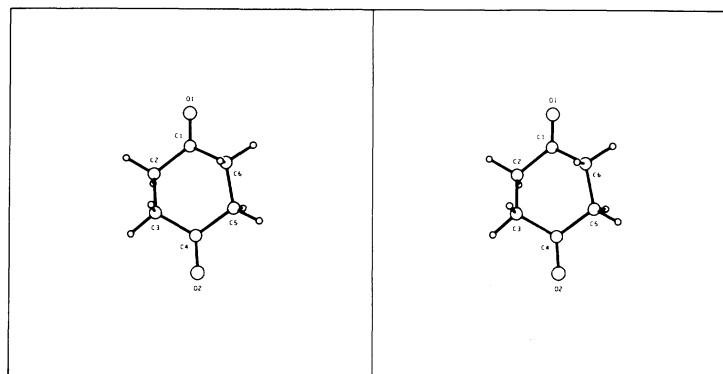
- (a) *Compound name*, usually that assigned by the authors
- (b) *Compound synonym*, where appropriate
- (c) *Qualifying phrase(s)*, introduced in parentheses following the compound name if the entry does *not* refer to a straight-forward X-ray study at room temperature
- (d) *Molecular formula*, expressed in terms of residues—i.e., discrete covalently-bonded groupings
- (e) *Bibliographic reference*, authors, names and journal reference to the paper from which the numeric data have been derived

21.1 Cyclohexane - 1,4 - dione



(i) A.Mossel, C.Romers, Acta Cryst., 17, 1217, 1964 (at -140°C)

(ii) P.Groth, O.Hassel, Acta Chem. Scand., 18, 923, 1964



(i)
P2₁ Z = 2 R = 0.081

CYHEXO LB 6-8-55

(ii)
P2₁ Z = 2 R = 0.069

CYHEXO01 LB 6-8-55

Bond Lengths				Bond Angles				Torsion Angles			
(i)				(i)				(i)			
C1	O1	1.21	1.22	O1	C1	C2	121	O1	C1	C2	C3
C4	O2	1.21	1.22	O1	C1	C6	123	C6	C1	C2	C3
C1	C2	1.53	1.51	C2	C1	C6	116	O1	C1	C6	C5
C1	C6	1.51	1.51	C1	C2	C3	113	C2	C1	C6	C5
C2	C3	1.54	1.55	C2	C3	C4	112	C1	C2	C3	C4
C3	C4	1.50	1.50	O2	C4	C3	122	C2	C3	C4	O2
C4	C5	1.52	1.51	O2	C4	C5	122	C2	C3	C4	C5
C5	C6	1.53	1.56	C3	C4	C5	116	O2	C4	C5	C6
				C4	C5	C6	113	C3	C4	C5	C6
				C1	C6	C5	111	C4	C5	C6	C1
							110				

21.2 Cyclohexane - 1,4 - dione



For complete entry see 21.1

Figure 2. Example of an amalgamated entry from ID

CHEMICAL DIAGRAM

This gives a conventional two-dimensional or linear formulation for each compound. We were grateful for permission to take many diagrams from two recent volumes in the Landolt-Börnstein series⁴; in certain cases, they were amended by a technical artist, who also drew additional material to the same scale and style.

TEXTUAL MATERIAL

Information in this section falls into three broad categories.

Editorial comments were added to the data base on card types² REMARK, DISORD, ERROR for any entry. These were keyboarded and included the special symbols required in the subsequent computer typesetting operation.⁵ Additional remarks were automatically generated from individual flags or combinations of flags on the summary card UNISUM,² as indicated in Table I.

Crystallographic constants were restricted to the space-group symbol (Hermann-Mauguin notation), Z-value (no. of formula units per cell), and R-factor(s). The first two, together with the crystallographic diagram, enable the user to deduce the molecular symmetry. The latter provides an indication of the "reliability" of the determination and of its associated data.

External cross-references were provided to the literature, for studies on a given compound published outside the period 1960-65, and also to the Landolt-Börnstein volumes⁸ via the LB entry number (Figures 1 and 2). Our own reference code¹ was also included to allow outside users to communicate with the Centre concerning a given compound.

CRYSTALLOGRAPHIC STEREO DIAGRAM

The stereo diagrams were prepared using the crystallographic connectivity established by the program UNIMOL.³ Initial attempts were made to transfer this project to a video-display system with joystick-control facilities, but either the distance from Cambridge was too great for reasonable editorial control or the required access time for tests and production runs was not available. All diagrams were therefore produced on the IBM 360/44 using a Calcomp 1627 10-inch pen plotter; diagrams for each compound used 9 inches of plot paper.

A preliminary run was made through the whole file using program PLOD. For each compound with coordinates and connectivity records, two "stick" drawings were produced. One of these gave a view down the shortest crystallographic axis and the other a projection onto the mean plane through all the atoms forming the molecule but omitting single atom residues, water, and NH_4^+ . Hydrogen atoms were included where possible and all atoms except C,H were labelled. These plots were used by the editors to check the connectivity of the system, and to "edit" the plot to obtain the best view for the final production run. Plots were usually prepared in batches of 50 compounds, which approached the plot-buffer limit, each one taking 5 to 10 seconds c.p.u. time and 1 minute plotter time.

Editing of plots was performed by program EDPLLOT. The instructions available to editors are shown in Table II and fall into three categories: definition of view, content of plot, and rotations of the chosen view. The rotations were defined by three mutually orthogonal plot axes, x_p parallel to the plot direction, y_p perpendicular to the plot direction, both in the plane of the plot, and z_p perpendicular to the plot itself and preserving a right-handed axial

Table I. Comments Automatically Generated from the Summary Card UNISUM

"Atomic coordinates were not reported"	
"No comparison of bond lengths is possible since they are not reported in the paper"	
"The structure is totally disordered"	
"The structure was determined in two dimensions only"	
"Z was not reported"	} and in any combination
"R was not reported"	
"The space group was not reported"	

set. A system for automatic positioning of atom labels was incorporated, as far as possible; this prevented the overlap of a given label with atoms, bonds, and other labels. In this pre-production run, only one diagram was plotted per compound, and all atoms of the asymmetric unit, except hydrogen, were labelled. Where molecular symmetry was coincident with a crystallographic symmetry element (i.e., SATOM records were present for atoms of type CIB³), the symmetry-generated atoms actually bonded to the asymmetric unit were also labelled. Every effort was made to align the plots with the chemical diagrams so that the reading of individual data elements from the numeric tables was made easier. For some compounds two or three passes were required to obtain an effective view compatible with clarity.

The final scaled, oriented coordinates for the atoms, together with coordinates for the atom labels, were compacted and written to magnetic tape. The final production of stereo pairs was performed on a CDC 7600 computer at the Los Alamos scientific laboratory of the University of California, using the program PLOP written in Cambridge. This produced the final diagrams (Figures 1 and 2) directly on 35-mm. film via an FR 80 microfilm plotter, which consists of a scope display with a 35-mm. camera

Table II. Instructions Used to 'Edit' the Crystallographic Diagrams

(a) Definition of basic view	
UNI	Short axis view
MAT	Projection onto mean plane (matrix) view
NONE	No plot required
(b) Definition of orientation	
$X \pm n$	Rotate by $\pm n^\circ$ about X_p axis
$Y \pm n$	Rotate by $\pm n^\circ$ about Y_p axis
$Z \pm n$	Rotate by $\pm n^\circ$ about Z_p axis
Any number of rotations may be specified. The final rotation is the cumulative result of working from left to right along the card.	
(c) Definition of content	
Default options: omit all single atom residues keep all hydrogen atoms	
OMIT	} Followed by
KEEP	
	(i) An atom name—e.g., C27
	(ii) A bond defined by two atom names—e.g., BOND Fe1 C35
	(iii) A residue defined by a constituent atom—e.g., R(Br1)
	(iv) SATOM, to omit all symmetry generated atoms
	(v) Any combination of (i), (ii), (iii)
CONN	followed by a pair of atom names creates a connection between those atoms
COMP	followed by two or more residues (defined as above) will destroy the crystallographic relationship between the residues and 'compose' the diagram so that it fills the available space to best advantage.
SATOM	Deal with <i>all</i> symmetry-generated atoms in labelling operation

on the face of the tube and uses a positive film process to get negative film strips.

A few of the most difficult plots, usually involving polymeric structures, were dealt with individually in Cambridge by direct photoreproduction of the EDPLOT output. Of the 1305 compounds in the book, plots were produced for 994. The negatives were printed in Cambridge and formed part of the page make-up input stream.

NUMERIC DATA

The numeric data synthesized from the file are *bond lengths* in Å to two decimal places, and *bond angles* and *torsion angles* to the nearest degree. This was done by program NUMBED, a molecular geometry program designed to use the crystallographic connectivity records. Certain special features were added for the production of ID.

To facilitate a consistent output order for numeric data, the input atoms were sorted internally by element type—i.e., transition metals (M); elements (X)—where X ≠ M, C, H, N, O, D; nitrogen, oxygen, carbon, hydrogen, and deuterium. Within the first two categories elements were arranged alphabetically. Further ordering in each category was by the numeric portion of the atom label. Symmetry-generated atoms appeared following those of the basic asymmetric unit for any category.

Algorithms were developed to ensure that the numeric data presented were nonredundant for structures with molecular symmetry.

Routines were included to read, decode, and act upon free-format edit cards; these enabled the editors to screen out unwanted information.

Initially a complete run was made through the file and all numeric data listed, residue by residue, for each entry. This list was then edited. The default options built into the program together with the available instructions are indicated in Table III. Instructions referring to an individual entry could be punched in any order, with the entry number in the first six columns. A system of continuation cards was developed for extensive editing of large lists, and checks introduced to ensure that instructions were not contradictory. The EXCLUDE cards referring to a complete class were assembled at the head of the deck. Attempts were made to keep the editing as consistent as possible. The final over-all content is summarized below.

In general, complete lists of *bond lengths* were presented for all entries, but were omitted for a few structures of low precision where a more precise study was available. Bond lengths involving hydrogen or deuterium were included for all neutron studies and for some X-ray studies at the editors' discretion. In certain cases, data for small residues—e.g., inorganic ions, solvent molecules—were omitted completely. For structures containing known errors which resulted in incorrect connectivity, no numeric data were presented. Where the error was more trivial, the authors' reported bond length, if present on a BOND record, was given in parentheses following our calculated value if the two differed by more than 0.05Å. This was done by comparing all published bond lengths present on BOND records with the calculated value. The bond lengths were grouped by residue and within each residue this list appeared as follows (where M and X are as defined above):

M-M	X-X	N-O	M-H	N-D
M-X	X-N	C-N	M-D	O-H
M-N	X-O	O-O	X-H	O-D
M-O	X-C	C-O	X-D	C-H
M-C	N-N	C-C	N-H	C-D

Table III. Instructions Used in Editing of Numeric Data by Program NUMBED

(a) Default options	
(i) Hydrogen and/or deuterium atoms will be omitted automatically unless a neutron study is indicated in the bibliographic qualifier.	
(ii) Bond lengths are always produced	
(iii) Bond angles are always produced	
(iv) Torsion angles are always produced	
(b) Edit instructions	
KEEP	Includes hydrogen and/or deuterium in bond length and bond angle calculations
OMIT R(C2)	Residue containing atom C2 excluded from all calculations
NOANG	Omit all bond angles
NOANG C1 C2 C3	Omit all angles at the vertices specified
ANG C1 C2 C3	Include only the angles at the vertices specified
EXCLUDE <i>n</i>	Do not calculate torsion angles for CLASS <i>n</i>
NOTOR	Omit all torsion angles
NOTOR C1 C2 N3 C7	Omit all torsion angles about the bonds specified, C1-C2, N3-C7, etc.
TOR	Include all torsion angles (only used for entries in a class where EXCLUDE is specified)
TOR C1 C2 N3 C7	Include only torsion angles about the bonds specified, C1-C2, N3-C7, etc.
OMIT C17 O5. . . .	Omit specified atoms from torsion angle calculation
BREAK C1 C2 N3 C7	'Breaks' the bonds C1-C2, N3-C7, etc., so that any atomic grouping attached to (here) C2 and C7 will not be included in the torsion angle calculation. This instruction used to screen out derivative side chains etc. (e.g. <i>p</i> -bromobenzoate)
CALC C1 C2 C3 C4	Include only the torsion angle specified, may be used as many times as required

Complete lists of *bond angles* were presented for chemical classes 1-70, but with the same exceptions applied to H, D and small residues as in the bond length tables. For classes 71-86, some angles were omitted—e.g., C-C-C angles in triphenylphosphine groups. Within each residue, angles were listed by apex atom in the order M, X, N, O, C.

The *torsion angles* presented were those which, in the editors' view, were of greatest conformational interest—i.e., they defined ring conformations or represented fairly accurately the dihedral angle between two planar groupings. The definition used throughout was that of Klyne and Prelog.⁹

The final deck of edit instructions was run against the data base to produce the *numeric galley* tape. This contained the card-image records listed in Table IV and included not only the numeric data but also all the textual information in a form ready for the typesetting process.

TYPESETTING AND PAGE ASSEMBLY

Material for the page make-up artist was provided in four streams:

- Bibliographic information was processed in Cambridge by program BIBTYP and converted to a steering

Ag-P		1.90	86.6	As-Pd		1.81	62.14	1.58	62.13	1.92*	51.1			
		1.91	65.7			1.82*	62.1 (2)		1.59		62.8	1.92	53.5	
	2.47	85.19	1.91*		65.9		1.82		62.10		1.62	62.15	1.93	35.3
			1.91*		65.11		1.83*		62.1				1.93*	36.2
			1.91*		65.15		1.83		62.5				1.93	51.2
	1.91*	84.20 (2)		1.83	62.13			1.93	58.8					
	1.91	86.6		1.84	62.10			1.93	60.1					
	1.91*	86.19 (2)		1.84	62.13			1.93	60.27					
Al-Br		1.92*	65.11	As-Pt		1.84	62.14	B-S		1.94	45.19			
		1.92	65.19			1.85*	62.1				1.94	53.6		
		1.92*	84.19			1.85	62.5				1.94	56.5		
		1.92*	84.20			1.85	62.10		1.92	62.14	1.94	62.10		
	1.93	60.32	1.92*		86.5	2.38*	86.7 (2)		1.86	62.14 (2)	1.93	62.14	1.95	1.22
2.36	60.32	1.92*	86.5			1.87*	62.1			1.95	16.18			
2.38	60.32	1.93	65.3 (2)			1.87	62.5 (2)			1.95	52.3			
2.50	60.32	1.93*	86.5			1.88	62.14			1.95	56.6			
		1.94	65.2			1.89	62.14			1.95	64.20			
		1.94*	86.7			1.90	62.13			1.96*	21.9 (2)			
		1.94*	86.19 (3)			1.92	62.13 (2)	Br-Br		1.96	22.2			
		1.95*	65.5	As-Ru		1.92	62.14				1.96*	28.1		
		1.95*	84.19				1.96*		62.6	2.25	60.8	1.96	45.20	
		1.95*	86.19 (2)				1.99*	62.6	2.32	60.8	1.96*	52.2		
		1.96*	65.5				1.99	62.13	2.54	9.6	1.96	60.11		
		1.96	65.12								1.97*	45.16		
		1.96	65.16							1.97	53.7			
		1.96*	86.19							1.97	62.10			
		1.97*	65.5							1.98*	19.18			
		1.97	65.8							1.98	26.4			
		1.97	65.13 (2)							1.98*	31.3			
		1.97*	65.15	As-S						1.98	38.2			
		1.97	65.17								1.98*	45.24		
		1.97	65.19								1.98	52.3		
		1.97	86.6								1.98	53.3		
		1.97*	86.19 (2)								1.98	53.4		
		1.98*	65.15							1.98	53.13			
		1.98	65.18							1.98	58.11			
		1.98*	84.19							1.99*	28.1			
		1.99*	65.15							1.99	53.11			
		1.99	65.17							1.99*	59.3			
		1.99	65.18							1.99	59.14			
		1.99*	86.7							2.00	58.11			
		1.99*	86.19	B-B						2.01	26.1			
		2.00	65.6								2.01*	51.4		
		2.00	65.19								2.02	1.26		
		2.01	65.3								2.03	26.4		
		2.01*	65.5								2.04*	51.4		
		2.01	65.18							2.06	38.2			
		2.01*	86.5							2.06	54.3			
		2.03	65.12											
		2.03	86.6											
		2.05*	86.19											
		2.06*	65.15 (2)											
		2.07	65.1											

Figure 3. Part of the bond length summary tables

tape for the Linotron 505 filmsetter at INSPEC (Information Service in Physics, Electrotechnology, Computers and Control). These procedures will be fully described in Part V of this series⁵; for the present volume, changes were made to the INSPEC programs to alter the typeface, point size, and layout of the information.

(ii) The file of *chemical diagrams* was assembled by the technical artist in Cambridge.

(iii) *Crystallographic stereo diagrams* were printed.

(iv) *Textual material and numeric data* were processed in Cambridge by program NUMTYP. The program assembled text from card types 3–8 of Table IV, generated text (see Table I) from the flags on UNISUM, and inserted the typesetting signals in card types 9–11. The output from NUMTYP was converted to INSPEC file structure by program IDTAPES, which produced magnetic tapes compatible with ICL 1900 series computers. The output from IDTAPES was used to generate the LINOTRON 505 steering tape. This yielded paragraphed text of set line width, and a line-numbered galley, listing tables of numeric data with column widths automatically adjusted to accommodate the largest atom label for a given entry.

Camera-ready copy for the 571 pages of the main book was assembled in batches of *ca.* 50 pages. Each batch was carefully proof-read and small corrections were stripped in. This process took about 11 weeks and was achieved without major difficulty since a great proportion of the work was done locally.

BOND LENGTH SUMMARY TABLES

In these tables, interatomic distances presented in the main book were brought together for each bonded element pair. Part of the table is shown in Figure 3. Table V gives a breakdown of the 15,368 distances in the main text by element type ($X \neq \text{C, N, O}$) and shows that 80% of this total is accounted for by C–C, C–N and C–O distances. Since bond lengths in these categories are affected by bond type and molecular environment, it was decided to omit them from the summary list and present instead a table assembled by one of us (O.K.) and previously published in "International Tables for X-Ray Crystallography"¹⁰ where bond type and environment were taken into account. Of the remaining 3,021 distances, 19 were rejected because their recalculated values differed by more than 0.05 Å from the published values. The final table contains 3002 entries spread over 170 element pairs.

Program SUMBED extracted the required bond lengths from the data base and wrote a compacted magnetic tape in entry number order. This was processed by program SUMTAB which produced a magnetic tape ready for typesetting and a listing of the table in its final form. It performed the following functions (see also Figure 3):

(i) Distances were arranged in element-pair order, with the element-pair sorted alphabetically. For each pair, the distances were sorted into increasing order and rounded to two decimal places (in Å).

(ii) Within each distance interval of 0.01 Å, the associated entry numbers were sorted in increasing order to facilitate reference to the main text.

(iii) If, for a given element pair, there were two or more equal distances in an individual main entry, then the number of such distances was indicated in parentheses following the entry number. This avoided needless repetition of the distance and entry number.

(iv) The distance was marked with an asterisk if it was obtained from a three-dimensional study with an *R*-factor ≤ 0.120 .

The table provides an element-pair index to the book and, in conjunction with the main text, we would expect it to be useful to the specialist crystallographer looking for comparison data and to the nonspecialist—e.g., a theoretical chemist attempting to obtain model dimensions.

Table IV. Information Content of Card Records on Numeric Galley Tapes

1. ENTNO	Entry number
2. UNISUM	Summary card from structural data file
3. REMARK	Editorial remarks
4. DISORD	Details of partial disorder
5. ERROR	Details of known errors, if these have been obtained directly from authors this is indicated
6. IDREF	Bibliographic reference(s) to studies outside the period 1960–65
7. SCOPE	<i>R</i> -factors and Landolt-Börnstein entry number
8. IDCRYS	Space-group symbol and <i>Z</i> value
9. BOND	Atom labels and bond length, one per record
10. ANGLE	Atom labels and bond angle, one per record
11. TOR	Atom labels and torsion angle, one per record
12. END	End of entry

Table V. Breakdown of Bond Lengths Presented in Main Text by Element Type (X Represents an Element other than C, N, O)

Element Pair	No. in Main Text	No. in Summary Tables
X–X	1,379	1,379
C–X	1,364	1,364
N–N	68	68
N–O	201	201
C–N	1,688	—
O–O	9	9
C–O	2,025	—
C–C	8,634	—
Totals	15,368	3,021

INDEXES

Three indexes are provided to the main text. Those based on *molecular formula* and *authors' names* are self-explanatory and were produced by similar procedures as for the bibliographic volumes.^{4,5} The permuted *compound name* index is, however, a relatively novel departure in books of this type.

A portion of this index is shown as Figure 4 and has a KWIC-style layout with the indexing points arranged alphabetically down the page center. The residual portion of the compound name wraps around to the left and right. The program, NAMINDEX, begins analysis of a given compound name at the indexing point indicated by an asterisk (*) in our bibliographic file.¹ The character string following the asterisk is compared with an input "stop list"—i.e., a list of common fragments of compound names which are not to be used as index points, for example aquo, acid, hydrate, di, tri, mono, etc. If a given string is found in the stop list, a pointer is moved to the end of the string and the comparison repeated with a new string. When a string is found which has no analogue in the stop list, a new indexing point has been found, and the program does the formatting and, if necessary, makes provision for the wrap-arounds. The compound name is scanned for further hyphens, spaces, or right parentheses, and the stop list comparison is taken up again to generate further indexing points for the name. This process stops at the end of the name (multiple spaces or end of last COMPND card) or at the start of a qualifier¹ (recognized by the combination: space, left parenthesis).

The process is repeated for all compound names, and finally the list of indexing points is sorted and printed with proper interpretation of typesetting signals. Multiple entries relating to the same compound are treated only once in this procedure, second and subsequent entries being omitted from the generation procedure.

The schedule for book production did not allow time for the development of typesetting procedures for this index,

	1,4-Di Thiane - disulfoxide	39. 8
	1,4-Di Thiane - iodine complex	39. 6
	1,4-Di Thiane-iodoform	60.21
	DL-1,2-Di Thiane-3,6-dicarboxylic acid	39.13
	Benzotritiroxan-13,14-di Thiatricyclo(8,2,1,1(4,7))tetradeca-4,6,10,12-tetraene	60.30
	(2-(1-Methylguanine))- (2-(3-methylbenzo- Thiazole)) methyl-monothethine cyanine iodide	41.15
	2-Benzoylimino-3-methyl Thiazolid-5-one	41.12
	2-Thio-3-ethyl-5- (2'- (3'-methylthiazolidinyldiene))- Thiazolidine-2,4-dione	41.10
	2-Thio-3-allyl-5- (2'- (3'-methylthiazolidinyldiene))- Thiazolidine-2,4-dione	41.11
	Phenyl Thiazolidinedione sodium salt	41. 8
	2-Thio-3-allyl-5- (2'- (3'-methyl Thiazolidinyldiene))-thiazolidine-2,4-dione	41.11
yl-5-thiomethyl-2,4-dihydro-1,3,4-thiadiazol-2- (1'-ethyl-4'-	Thiazolidinyldiene-2'-thione-5'-one)	[3-Meth 41. 9
	2-Thio-3-ethyl-5- (2'- (3'-methyl Thiazolidinyldiene))-thiazolidine-2,4-dione	41.10
	2- (2'-Hydroxyethylamino)-2- Thiazoline	41. 5
	Thiete sulfone	39. 1
	Tetracyano-1,4-di Thiin	39.17
	Ethyl Thio iron tricarbonyl dimer	85.18
	3-Ethyl-5- (4-methyl Thio)-cdelta(2)-1,3,4-(thiadiazolin-5-ylidene) rhodanine	41. 9
	Di Thio-oxamide	4. 1
-dione]	2- Thio-3-allyl-5- (2'- (3'-methylthiazolidinyldiene))-thiazolidine-2,4	41.11
,4-dione]	2- Thio-3-ethyl-5- (2'- (3'-methylthiazolidinyldiene))-thiazolidine-2	41.10
	Thioacetamide	4. 2
	tetrakis(Thioacetamido) chloro copper(i)	85.15
	Dimethylarsino dimethyl-di Thioarsinate	65. 5
	4,4'-Dihydroxy Thiobenzophenone monohydrate	17.13
	2-Methyl Thiobenzothiazole	41. 7
	Copper(ii) diethyl-di Thiocarbamate	80. 5
	Nickel(ii) diethyl-di Thiocarbamate	80. 6
	Silver(i) diethyl-di Thiocarbamate hexamer	80.11
	Diethyl-di Thiocarbamate copper(i) tetramer	80. 9
	Nitroso-(dimethyl-di Thiocarbamate) cobalt	80. 3
	bis(Diethyl-di Thiocarbamate) zinc(ii) dimer	80.10
	Tri Thiocarbonic acid	11. 2
	Aniline p- Thiocyanate	16.10
	bis(Ethylenediamine) copper(ii) Thiocyanate	76. 9
	Mercury(ii) tetrathiourea cobalt(ii) tetra Thiocyanate	79. 5
	bis-Ethylenethiourea cadmium Thiocyanate	79. 7
	Palladium(ii) bis(tri-n-propyl phosphine) Thiocyanate	86. 8
	Silver Thiocyanate - tri-n-propyl phosphine complex	85.19
	Cobalt(ii) di Thiocyanate dipyridine	83.24
	trans-bis(Ethylenediamine) bis(iso Thiocyanato) nickel(ii)	76.13
	trans-N-Methyl-N-benzyl- Thioformamide	4. 6
	2,3-Dimethyl-di Thiofurophthene	39.15
	Molybdenum tris(di Thioglyoxal)	85. 8
	DL-Homocysteine Thiolactone hydrochloride	48. 7
	Di(tetra-n-butylammonium) cobalt(ii) bis(maleonitrile di Thiolate)	85. 9
	tetra-n-butylammonium copper(ii) bis(maleonitrile di Thiolate)	85.10
	bis-(cis-1,2-bis(trifluoromethyl)ethylene-1,2-di Thiolato) cobalt(ii) dimer	85.20
	1,5-Diamino-1,2-di Thiolum iodide	39. 2
	3-Phenyl-1,2-di Thiolum iodide	39.20
	3-ethyl-5- Thiomethyl-2,4-dihydro-1,3,4-thiadiazol-2- (1'-ethyl-4'-thiazolid	41. 9
inyldiene-2'-thione-5'-one)]	3-Bromo-2,4-dinitro Thionaphthen-5-diazo-4-oxide	39.18
	trans-Dichloro-bis(ethylenediamine) cobalt(iii) hexa Thionate monohydrate	76. 6

Figure 4. Part of the KWIC-style compound name index

and, for this volume, it was produced by photo-offset directly from the computer output.

SUMMARY

The main volume of the "Tables of Interatomic Distances 1960-1965" contains entries for 1305 compounds, of which 277 had no numeric data for one of the reasons given above. The remaining 1028 entries contained numeric data for 1174 residues and comprised 15,378 bond lengths, 20,878 bond angles, and 10,736 torsion angles.

The book was produced from a computer-readable data base, and computer techniques were extensively used for the checking of numeric data, and for the production of diagrams, tables, text, and indexes; only the page make-up was undertaken by traditional methods. Although the computer programs were developed for the specific task in hand, the general organization of the work is applicable to the production of other types of scientific reference volumes.

ACKNOWLEDGMENT

We thank the Office for Scientific and Technical Information for financial support and the Institute of Theoretical Astronomy for computing facilities on the IBM 360/44. We are indebted to the printers, Messrs Unwin Bros., especially their representatives P. Crew and F. Lovett, for their advice and assistance. Other individuals who have contributed to the preparation of this volume are: A. C. Larson, programmer at Los Alamos scientific laboratory; P. Simmons, Linotron 505 filmsetter; D. Halls, technical artist; and R. S. Hammans, printer.

LITERATURE CITED

- (1) Kennard, O., Watson, D. G., and Town, W. G., *J. Chem. Doc.* 12, 14 (1972).
- (2) Allen, F. H., Kennard, O., Motherwell, W. D. S., Town, W. G., and Watson, D. G., *Ibid.*, 13, 119, (1973).
- (3) Allen, F. H., Kennard, O., Motherwell, W. D. S., Town, W. G., Watson, D. G., Scott, T., and Larson, A. C., *J. Appl. Cryst.*, in press.
- (4) Kennard, O., and Watson, D. G., "Molecular Structures and Dimensions," Vols. 1, 2, 3, and 4 (with Town, W. G.). Published for the Crystallographic Data Centre and the International Union of Crystallography by N. V. A. Oosthoek's Uitgevers Mij, Utrecht, 1970, 1972, 1973.
- (5) Kennard, O., Town, W. G., Simmons, P., and Watson, D., G., in preparation, Cambridge Crystallographic Data Centre, 1973.
- (6) Kennard, O., Watson, D. G., Allen, F. H., Isaacs, N. W., Motherwell, W. D. S., Pettersen, R. C., and Town, W. G., "Molecular Structures and Dimensions, Vol. A1, Interatomic Distances 1960-1965." Published for the Crystallographic Data Centre and the International Union of Crystallography by N. V. A. Oosthoek's Uitgevers Mij, Utrecht, 1973.
- (7) Sutton, L. E., "Tables of Interatomic Distances and Configurations in Molecules and Ions," Chemical Society Special Publications No. 11, London, 1958, and No. 18, London, 1965.
- (8) Hellwege, K.-H., and Hellwege, A. M., "Numerical Data and Functional Relationships in Science and Technology" III, Vols. 5a, 5b. Berlin, Springer, 1971.
- (9) Klyne, W., and Prelog, V., *Experientia* 16, 521 (1960).
- (10) "International Tables for X-Ray Crystallography," Volume III, Birmingham, The Kynoch Press, p. 275, 1962.