DOCUMENT RESUME

ED 065 162 LI 003 784

TITLE	Crystallographic Data Centre Services and Publications.
INSTITUTION	Cambridge Univ. (England). Chemical Lab.
SPONS AGENCY	Department of Education and Science, London (England).
PUB DATE	[72]
NOTE	16p.; (3 References)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Bibliographic Citations; Chemical Analysis; Chemistry; Computer Programs; *Data Bases; Foreign Countries; *Information Centers; *Information Services; Magnetic Tapes; Nuclear Physics; Publications; *Research; *Sciences
IDENTIFIERS	England; Molecular Structures; *Scientific and Technical Information

ABSTRACT

The Cambridge Crystallographic Data Centre is concerned with the retrieval, evaluation, synthesis, and dissemination of structural data based on diffraction methods. The source of input is almost entirely primary journals. Bibliographic information and numeric data on crystal and molecular structures are on magnetic tapes. The bibliographic file is ordered in 86 chemical classes, and by molecular formula within each class. The numeric file contains data corresponding to each entry in the bibliographic file published since January 1960. This paper discusses the Centre's publications and services including: (1) bibliographic magnetic tape service, (2) hard-copy alerting service, (3) retrospective searches, (4) manual searches, (5) numeric data magnetic tape service, and (6) computer programs. A list of the 86 chemical classes used for file ordering, information on the Protein Data Bank, and a breakdown of service charges are appended. (Author/SJ)

CRYSTALLOGRAPHIC DATA CENTRE

SERVICES AND PUBLICATIONS

University Chemical Laboratory
Lensfield Road
Cambridge
England

ENQUIRIES

Enquiries should be addressed to Mrs.O.Kennard or Dr.D.G.Watson (Tel. 0223 56491)

SUBJECT FIELD

Crystal structures of organic and organometallic compounds analysed by X-ray or neutron diffraction methods.

SPONSOR

Office for Scientific and Technical Information, Department of Education and Science, Elizabeth House, 39 York Road, London SE1, England.

DATA BASE

The source of input is almost exclusively primary journals. Bibliographic information and numeric data are held on magnetic tapes and these files are regularly updated on a batch basis - about 150 structures are processed every six weeks. Numeric data are subject to critical evaluation before incorporation in the data base. Four files are regularly maintained:-

1. Bibliographic (1935-) Computer-readable
2. Numeric data (1960-) Computer-readable
3. Chemical diagrams (1935-)
4. Reprints (1960-)

OUTPUT

Output from the data base is of two types - publications and mag.tape or hard-copy computer services.

BIBLIOGRAPHIC FILE

The bibliographic file is ordered in 86 chemical classes (see Appendix 1) and by molecular formula within each class. The total number of entries in the file at May 1972 was ca. 7000.

Each entry refers to a single structural investigation of an organic or organometallic compound by X-ray or neutron diffraction.

Entries have the following elements:-

- Compound name
- Compound synonym (as needed)
- Indication of polymorph type, experimental temperature, etc.
- Molecular formula
- Authors' names
- Journal reference
- Cross-reference to "Interatomic Distances" (for pre-1960 entries)
- Cross-reference to "Structure Reports" (for pre-1960 entries)
- Chemical classification code
- Accession date
- Identification code
- Ring Index notation has been introduced during 1972.

Bibliographic entries are illustrated in the booklet "Specimen File Entries".

NUMERIC DATA FILE

The numeric data file contains data corresponding to each entry in the bibliographic file published since Jan.1960.

Each entry has the following possible elements:-

- Unit cell dimensions, with standard deviations
- Space group symbol
- Symmetry operators for special space group settings
- Number of formula units per unit cell
- Calculated and measured densities
- Standard deviation of measured density
- Temperature of density measurement
- Temperature of measurement of intensities
- Melting point
- Atomic coordinates
- Bond lengths
- Connectivity specifications
- Critical evaluation flags
- Remarks
- Identification code

During 1972 the following data elements have been introduced:-

- Reliability factor (R)
- Bond lengths corrected for thermal vibration
- Flag for mode of intensity measurement

Numeric data entries are illustrated in the booklet "Specimen File Entries".

Data are released from the file only after computer processing and critical evaluation for internal consistency. Transcription errors will have been corrected and entries flagged with respect to their check status.

Atomic coordinates refer to covalently bonded residues - molecular or ionic - and in many cases have been transformed from the values listed in the original publications. In the provision of services facilities are available for the output of coordinates with respect to orthogonal cell axes, etc.

Some examples of the uses of the numeric data file are:-

(a) calculation of molecular geometry
(b) calculation of intermolecular packing geometry
(c) generation of molecular or packing diagrams - mono or stereo
(d) use of atomic coordinates in theoretical chemistry calculations

Standard crystallographic computer programs are generally available for (a)-(c)

A special numeric data file is being set up for protein structures. Details are given in Appendix 2.

A computerised data base covering inorganic crystal structures is being established under the direction of Professor D.Rogers,Department of Chemistry,Imperial College of Science and Technology, London, England.

PUBLICATIONS

A new series of standard reference volumes, "Molecular Structures and Dimensions", is published by the Centre in conjunction with the International Union of Crystallography. The series is a continuation and extension of the "Tables of Interatomic Distances and Configuration in Molecules and Ions" (Chemical Society Special Publication), which covered the literature until the end of 1959.

The first three volumes of the series contain classified bibliographic information for some 6500 compounds:-

Vol. 1 (1935-69): Classes 1-59 - General Organic Crystal Structures

Vol. 2 (1935-69): Classes 60-86 -Complexes and Organometallic Crystal Structures

Vol. 3 (1969-71): Classes 1-86 - Organic and Organometallic Crystal Structures

Cumulative indexes are provided in Vol.3 and will be included in future annual supplements.

The distributors for the series are:-

A.Oosthoek, Domstraat 11-13, Utrecht, Netherlands

Polycrystal Book Service, P.O.Box 11567, Pittsburgh, Pennsylvania 15238, USA

Copies can also be ordered through any bookseller.

The first of the volumes dealing with numeric data is planned for production in late 1972. X-ray and neutron studies of organic and organometallic compounds for the years 1960-65 will be presented. This compilation will include bibliographic information, bond lengths, bond angles, torsion angles, chemical diagrams, stereoscopic crystallographic diagrams and summary tables of bond lengths.

SERVICES

Details of the charges for services are given in Appendix 3.

1. BIBLIOGRAPHIC MAG.TAPE SERVICE

The bibliographic file may be leased for periods of one year commencing January 1st of any year. The base fee for the lease during the first year depends on the calendar year in which the service is first provided to the user.

The file can be leased for periods longer than one year by payment of an annual subscription fee and the Centre will provide a replacement copy of the file as it exists on January 1st of each year.

For a separate fee an update service can be provided whereby the Centre will supply a mag.tape of new entries whenever the bibliographic file is updated. Normally the file is updated every six weeks and each new batch contains ca. 150 entries. The update service would guarantee a minimum of 5 batches per annum.

There is no fee for the provision of services by the user to his own employees. Provision of external services by the user will require negotiation of a special contract. Ownership of the file remains vested in the Crystallographic Data Centre. Copies of the Basic Agreement and the Supplementary Licence Agreement required for the operation of a publicly available service may be obtained on request.

A specimen mag.tape is available from the Centre and its cost can be credited against the first full payment if the user subsequently subscribes to the leasing service.

2. BIBLIOGRAPHIC HARD-COPY ALERTING SERVICE

The Centre will provide hard-copy lists of new entries whenever the bibliographic file is updated. Normally the file is updated every six weeks and each new batch contains ca. 150 entries. The alerting service would guarantee a minimum of 5 batches per annum.

Two forms of output are available viz. line-printer listings and perforated cards.
Specimen output, free of charge, is available on application to the Centre.

As a back-up to this service it is possible to provide a hard-copy list of the complete bibliographic file (1935-72).

3. RETROSPECTIVE BIBLIOGRAPHIC SEARCH SERVICE

A study of the specimen bibliographic entries in Appendix 2 indicates the types of retrospective searches which can be undertaken.
Thus it is possible to search on the basis of chemical class, year of publication, authors' names, journals and for specified elemental compositions in the molecular formula. In addition one can search the qualifying phrases for neutron studies, low temp.studies etc. However, since we have no chemical connexion table at present it is not possible to undertake substructural searches. The Centre has plans for including chemical connexion tables in the bibliographic file at a later date.

For each enquiry the Centre will provide a cost estimate which will include an assessment of any charges for special programming.

4. MANUAL BIBLIOGRAPHIC SEARCH SERVICE

Since the bibliographic volumes of the series "Molecular Structures and Dimensions" are published only once each year the Centre is prepared to answer, free of charge, simple queries of the type "Has the structure of X been determined?"

The enquirer must submit the compound name, structural diagram and molecular formula.

5. NUMERIC DATA MAG. TAPE SERVICE

It is hoped that, in the future, it will be possible to maintain complete copies of the numeric data file (and bibliographic file) in various national centres. In the meantime users can obtain subsets of the file directly from the Crystallographic Data Centre.

It is anticipated that most searches will be based on our chemical classification scheme but, in any case, the provision of numeric data will normally be a 2-stage process.

Stage 1 consists of the provision of a listing of the appropriate bibliographic entries and a data index. The latter summarises the content of the data entries and their check status.

Stage 2 consists of the provision of mag.tapes containing numeric data entries with the corresponding bibliographic entries. This set may be the complete set indicated by stage 1 or a subset chosen by the user.

For stages 1 and 2 cost estimates will be provided.

A specimen mag.tape is available from the Centre and its cost can be credited against the first full payment to Stage 2 if the user subsequently subscribes to the service.

6. COMPUTER PROGRAMS FOR BIBLIOGRAPHIC FILE

The following programs can be purchased from the Centre:-

(a) List Programs

(i) Formatted List programs (upper-case only or upper/lower-case)

(ii) Listing of entries for specified classes

(iii) Listing of entries for specified years

(b) Index Programs

(i) Molecular formula

(ii) Author

(iii) Identification code

(c) Update Program

This program is used for merging new entries in the main bibliographic file and handles the problem of superseded entries.

Minimum Computer Configuration

(a) All need 96K bytes core. (i) and (iii) require 1 mag. tape deck.
(ii) requires 2 mag.tape decks.

(b) All need 256K bytes core and 1 mag.tape deck.
(i) requires 2 scratch files and (ii) requires 3 scratch files.

(c) Needs 256K bytes core, 3 mag.tape decks, 2 small scratch files, 1 card punch.

Above programs all require 1 card reader and 1 line-printer.

APPENDIX 1

LIST OF CHEMICAL CLASSES

Class No.	Class Name	No. of Entries
1	Aliphatic Carboxylic Acid Derivatives	195
2	Aliphatic Carboxylic Acid Salts (Ammonium, IA,IIA Metals)	118
3	Aliphatic Amines	132
4	Aliphatic (Nand S) Compounds	21
5	Aliphatic Miscellaneous	50
6	Enolates (Aliphatic and Aromatic)	9
7	Nitriles (Aliphatic and Aromatic)	38
8	Urea Compounds (Aliphatic and Aromatic)	72
9	Nitrogen-Nitrogen Compounds (Aliphatic and Aromatic)	69
10	Nitrogen-Oxygen Compounds (Aliphatic and Aromatic)	42
11	Sulphur and Selenium Compounds	104
12	Carbonium Ions,Carbanions,Radicals	34
13	Benzoic Acid Derivatives	79
14	Benzoic Acid Salts (Ammonium,IA,IIA Metals)	14
15	Benzene Nitro Compounds	40
16	Anilines	79
17	Phenols and Ethers	63
18	Benzoquinones	30
19	Benzene Miscellaneous	78
20	Monocyclic Hydrocarbons (3,4,5-Membered Rings)	51
21	Monocyclic Hydrocarbons (6-Membered Rings)	57
22	Monocyclic Hydrocarbons (7,8-Membered Rings)	23
23	Monocyclic Hydrocarbons (9- and Higher-Membered Rings)	26
24	Naphthalene Compounds	41
25	Naphthoquinones	26
26	Anthracene Compounds	54
27	Hydrocarbons (2 Fused Rings)	29
28	Hydrocarbons (3 Fused Rings)	44

Class No.	Class Name	No.of Entries
29	Hydrocarbons (4 Fused Rings)	34
30	Hydrocarbons (5 or More Fused Rings)	43
31	Bridged Ring Hydrocarbons	103
32	Hetero-Nitrogen (3,4,5-Membered Monocyclic)	84
33	Hetero-Nitrogen (6-Membered Monocyclic)	130
34	Hetero-Nitrogen (7-and Higher-Membered Monocyclic)	17
35	Hetero-Nitrogen (2 Fused Rings)	73
36	Hetero-Nitrogen (More than 2 Fused Rings)	73
37	Hetero-Nitrogen (Bridged Ring Systems)	44
38	Hetero-Oxygen	83
39	Hetero-Sulphur and Hetero-Selenium	144
40	Hetero-(Nitrogen and Oxygen)	53
41	Hetero-(Nitrogen and Sulphur)	78
42	Hetero-Mixed Miscellaneous	26
43	Barbiturates	45
44	Pyrimidines and Purines	123
45	Carbohydrates	124
46	Phosphates	33
47	Nucleosides and Nucleotides	65
48	Amino-Acids and Peptides	153
49	Porphyrins and Corrins	55
50	Antibiotics	76
51	Steroids	98
52	Monoterpenes	26
53	Sesquiterpenes	75
54	Diterpenes	47
55	Sesterterpenes	4
56	Triterpenes	34
57	Tetrapenes	9
58	Alkaloids	189
59	Miscellaneous Natural Products	97
60	Molecular Complexes	225
61	Clathrates	31
62	Boron Compounds	72
63	Silicon Compounds	52
64	Phosphorus Compounds	158
65	Arsenic Compounds	63
66	Antimony and Bismuth Compounds	29

Class No.	Class Name	No.of Entries
67	Groups IA and IIA Compounds	52
68	Group III Compounds	56
69	Germanium, Tin, Lead Compounds	94
70	Tellurium Compounds	30
71	Transition Metal-C Compounds	143
72	Metal π -Complexes (Open-Chain)	138
73	Metal π -Complexes (Cyclopentadiene)	170
74	Metal π -Complexes (Arene)	45
75	Metal π -Complexes (Miscellaneous Ring System)	129
76	Metal Complexes (Ethylenediamine)	175
77	Metal Complexes (Acetylacetone)	104
78	Metal Complexes (Salicylic Derivatives)	94
79	Metal Complexes (Thiourea)	62
80	Metal Complexes (Thiocarbamate or Xanthate)	61
81	Metal Complexes (Carboxylic Acid)	162
82	Metal Complexes (Amino-Acid)	89
83	Metal Complexes (Nitrogen Ligand)	384
84	Metal Complexes (Oxygen Ligand)	95
85	Metal Complexes (Sulphur or Selenium Ligand)	155
86	Metal Complexes (P,As,Sb Ligand)	225

Note:- The number of entries for each class corresponds to the file content at May 1972.

APPENDIX 2

PROTEIN DATA BANK

From 1972 a repository system for protein crystallographic data is operated jointly by the Centre and the Brookhaven National Laboratory, U.S.A. Atomic coordinates are stored and provision will be made for the storage of structure factors and electron density maps.

Requests for these data should be made to any of the following:-

(i) Cambridge Crystallographic Data Centre

(ii) Dr.W.C.Hamilton,Brookhaven National Laboratory, Upton, New York 11973, U.S.A.

(iii) Dr.E.F.Meyer, Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas 77843, U.S.A.

Cambridge and Brookhaven maintain identical files and distribution is on magnetic tape wherever possible. At present there is no charge for the service other than handling costs. The total holding will be announced annually in the bibliographic volumes of the series "Molecular Structures and Dimensions".

APPENDIX 3

SERVICE CHARGES

The charges set out below apply to users in academic and other non-profit research organisations. Should anyone from these institutions have difficulty in meeting the charges for services they should get in touch with the Centre. Users in industry and commerce should apply to the Centre for service charges.

1. BIBLIOGRAPHIC MAG. TAPE SERVICE

Base fee for 1972	£100
Subscription fee	£20 per annum
Update fee	£50 per annum
Specimen tape	£20 (can be credited against future service)
Postage	Extra
Cost of mag. tapes	Extra
Documentation	No charge

2. BIBLIOGRAPHIC HARD-COPY ALERTING SERVICE

Line-printer paper output	£15 per annum
Line-printer perforated card output	£25 per annum
Postage	Extra
Specimen output	No charge

Line-printer perforated card output from the complete bibliographic file £1.50 per 100 entries.

3. RETROSPECTIVE BIBLIOGRAPHIC SEARCH SERVICE

Minimum charge	£10
Postage	Extra

A cost estimate will be provided for each enquiry.

4. MANUAL BIBLIOGRAPHIC SEARCH SERVICE

No charge.

5. NUMERIC DATA MAG. TAPE SERVICE

Specimen numeric data mag. tape	£20 (can be credited against future Stage 2 service)

Stage 1: A cost estimate will be provided for each enquiry.

Minimum charge	£15
Postage	Extra

On the basis of the results from stage 1 an estimate will be provided for stage 2.

Stage 2:

Cost of data	£1 per data entry.
Cost of mag. tapes	Extra
Postage	Extra
Documentation	No charge

N.B. A data entry is charged only if it satisfies the following criteria:-

(a) atomic coordinates are present

(b) it has been evaluated and all clerical transcription errors rectified.

6. COMPUTER PROGRAMS FOR BIBLIOGRAPHIC FILE

List programs	£200
Index programs	£400
Update programs	£400
Complete package	£800

The above program charges include postage, tapes and documentation.