bring **L** vertical we have to increase ψ by ψ_0. This gives the components of **L** in X',Y',Z' as $(-\sin\psi_0, -\cos\psi_0, 0)$. Multiplication by **M** gives the expressions for L_x, L_y and L_z.

APPENDIX *B*

Seen along $-\mathbf{H}$, as in Fig. 5, the projection of **A** is perpendicular to **L**, as examination of Fig. 2(*b*) will show. The azimuth for **A** is therefore given by $\tan\psi_0 = -A_{y'}/A_{x'}$; the minus sign occurs because ψ_0 is by definition positive in this situation and $A_{y'}$ negative.

The components of **A** in X',Y',Z' are found from

$$\mathbf{A}_{X'Y'Z'} = \mathbf{M}^{-1}.\mathbf{A}_{XYZ},$$

where $\mathbf{M}^{-1} \equiv \mathbf{M}^T$, as **M** represents a pure rotation. Now the expression for $\tan\psi_0$ can be obtained.

References

ALEXANDER, L. E. & SMITH, G. S. (1962). *Acta Cryst.* **15**, 983–1004.

CLEGG, W. (1981). *Acta Cryst.* A**37**, 22–28.

GOLDSTEIN, H. (1959). *Classical Mechanics*, 6th ed., p. 143. Reading, MA & London: Addison-Wesley.

NELMES, R. J. (1975). *Acta Cryst.* A**31**, 273–279.

PAULUS, E. F. (1975). *Acta Cryst.* B**31**, 1191–1193.

RIECK, G. D. (1971). *High Temp. High Pressures.* **3**, 419–424.

TICHÝ, K. & BENEŠ, J. (1977). *Helv. Phys. Acta*, **50**, 459–466.

TICHÝ, K. & BENEŠ, J. (1979). *J. Appl. Cryst.* **12**, 10–14.

Acta Cryst. (1983). A**39**, 216–224

The Standard Crystallographic File Structure

BY I. D. BROWN

Institute for Materials Research, McMaster University, Hamilton, Ontario, Canada L8S 4M1

(*Received* 2 *August* 1982; *accepted* 27 *September* 1982)

Abstract

This paper describes a file structure that has been developed by a joint working party of the Data and Computing Commissions of the International Union of Crystallography. It is intended as a standard that can be used by those wishing to transfer any type of crystallographic data from one laboratory or program system to another.

Introduction

With the increasing use of computers in all branches of crystallography it has become necessary to define a standard file structure that will allow data files produced in one laboratory to be read directly into programs in a different laboratory.

At the Warsaw Congress of the International Union of Crystallography in 1978, the Data and Computing Commissions of the Union appointed a working party to propose a standard file structure for crystallographic data. The working party submitted its report at the 1981 Congress of the Union in Ottawa. This report, which is given below, was adopted by the Commissions with a recommendation to all authors of crystallographic programs that they write their programs so that they can read and write files with this structure.

Early in its deliberations the working group defined the following criteria to be met by the file structure. They are listed in decreasing order of importance.

1. *The file structure must be extendable to include all types of crystallographic data.*

2. *It must be compatible with current and future methods of data transmission.* Currently magnetic tapes are favoured with punch cards still sometimes used. Almost certainly there will be great changes in data transmission technology in the next few years.

3. *It should be easy to program for both reading and writing.* Files written in this structure are designed for machine-to-machine communication. Not all users will be experienced programmers or have access to large program systems. This implies the use of fixed formats. Users may well prefer to enter data in free format and use the computer to generate an exchange file in the standard form.

4. *The file should not require reread facilities* since these are not supported by all computers.

5. *A listing of a file written in this format should be easy to read visually* consistent with 3 above.

6. *The only records that must be included are those required for data management* (*e.g. END*). A standard crystallographic file will contain information of use to the writer and reader of the file. An author sending structural data to a journal will be interested in different

0567-7394/83/020216-09$01.50

data from workers exchanging powder patterns for phase identification.

7. *Provision should be made for the inclusion of derived data if required.* Some calculations, *e.g.* of structure factors [see Thomas (1978)], may be based on elaborate models using programs not available in other laboratories. It should be possible to transmit this information in a standard file.

8. *Provision should be made for comments* remembering that the interpretation of information on a comment card by a computer is difficult or impossible. This information is essentially for people only.

9. *The file structure is not intended for keyboarding data.* Although it can be used for this purpose, most crystallographers will prefer to use their own free-format routines for entering data by hand (see 3 above).

The following document is the report of the working party which consisted of I. D. Brown (Canada, Chairman), S. C. Abrahams (USA), R. Diamond (UK), S. R. Hall (Australia), A. C. Larson (USA), A. D. Mighell (USA), E. Parthé (Switzerland) and R. Shirley (UK). A number of others made valuable suggestions, most of which were incorporated in the final document. I would like to thank all of these people including the members of the working party for their contribution to the definition of this standard.

Standard crystallographic file structure

1. *Purpose*

The purpose of the standard crystallographic file structure is to assist in the exchange of crystallographic data between laboratories and to make it easy for the same data to be used as input to different programs.

2. *Transfer medium*

The file may be transferred by any acceptable medium. Unless otherwise specified by the user, the following conventions will be assumed (some common alternatives are included in parentheses).

Cards. Standard 80 column IBM cards punched using the 029 (026) punch convention.

Magnetic tape. 9 track, 800 bits per inch (9 track 1600 bpi or 7 track 800 bpi). Unlabelled. USASCII/7(EBCDIC): 80 characters/record; 45 record/block, zero filled if necessary (unblocked).

3. *The data structure of a standard crystallographic file*

(For simplicity, the file is described in terms of card images but without implying that it must physically exist in the form of cards.) A sample file is given in Appendix I.

3.1. *A file* consists of entries, each entry being logically independent of other entries. An entry normally will consist of data referring to one crystalline phase. Each entry begins with a TITLE card and ends with an END card.

3.2. *An entry* consists of a number of sections each including data of a particular type, *e.g.* atomic coordinates or structure factors. Each section begins with a header card and ends with an end-of-section (EOS) card (any card * in col. 1). The end-of-section card ensures that the program is ready to read the next card as a header.

3.3. *Each section* consists of formatted cards (or lines) containing 80 characters (including blanks). 5 characters at the end of each card are reserved for card sequence numbers (this is necessary as long as there is a danger of dropping a deck of cards).

3.4. *The character set* is restricted to the 46 characters 0–9, A–Z , . + − * / ( ) = *blank*. These characters are the only standard ones available on all machines.

3.5. *Cards* are of two types:

(*a*) *Header cards* are used to start a new section. The first 8 characters indicate the format of the following data cards and in some cases the information that is to be found on them. In addition, each header card may contain comments, such as alphanumeric column headings to help visual reading (see sample file).

(*b*) *Data cards* contain the data specified by the most recently read header card. An asterisk (*) in col. 1 indicates the last card in the section.

3.6. *Most data cards* begin with the following three fields:

(*a*) *EOS* (*end of section*) *col.* 1. This must be an asterisk (*) on the last card of each section, otherwise any other legal character (normally blank) may appear in this field. The end-of-section card may be a normal data card but in sections with a card identifier a special end-of-section card may be used containing *EOS in cols. 1–4 and cols. 5–75 blank.

(*b*) *CID* (*card identifier*) *cols.* 2–5. Within each section, the format of all the data cards is the same but in some sections the type of information stored on the card may be different. In these sections, each card contains a card identifier (CID) whose value determines the type of information carried (*e.g.* in the HKL section, the value of CID will determine whether the card contains intensities, calculated structure factors or powder data). CID uses an A4 field and is left justified. The last card of any section may have 'EOS' as a card identifier. No other data should appear on a card with CID = 'EOS'.

(*c*) *DSK* (*data set key*) *cols.* 6–9. Most cards also include a data set key (DSK) which allows the user to associate different data together. For example, the user

may wish to include data from several experiments in the same entry (*e.g.* data from X-ray and neutron diffraction experiments, or data for a native protein and several of its isomorphs). Cards from the different experiments would carry different values of DSK but could appear together within the same section (*e.g.* the structure factors of several isomorphs for a given reflection can be grouped together). DSK may contain any legal characters chosen by the user. Any card on which DSK is blank is assumed to contain data that applies to all the data sets included in the entry.

3.7. *Head cards* that cannot be interpreted are to be ignored. Some consequences are:

(i) Blank cards may be used to separate sections for visual effect.

(ii) The presence of an incorrect header may result in problems during reading of the file since cards will be ignored until an interpretable header is found.

(iii) Instruction or data cards for a user's program can be added to a file provided they do not mimic legal header cards. This can be ensured by using a character other than a letter or blank in columns 1–8.

(iv) Comments can be inserted between sections providing that columns 1–8 do not mimic header cards (*e.g.* if they are left blank). Unlike comments included in the REMARK section, these comments may not be read by a user's program and should be used with care.

3.8. *Data cards* that cannot be interpreted should be avoided. Since these will be read with a fixed-format read statement, they could cause a fatal read error.

3.9. Sections may follow each other in any order, and the same section may be included any number of times within an entry but where the file contains duplicate information (*e.g.* two CELL DIMension sections) the values appearing latest in sequence are the values to be used.

4. *Formats for the standard crystallographic file structure*-81

Each section starts with the header card shown. The first eight characters are reserved for an alphabetic section name. Otherwise the card may contain any other alphanumeric characters. All the other cards in the section are data cards and have the format shown. The last card in each section must have an asterisk in col. 1. Cols. 76–80 on all cards are reserved for a card or line-sequence number. The TITLE and END cards must appear in all entries. Other sections may be included as required by the user. Unless otherwise stated, all microscopic dimensions (a, b, c, λ) are in ångström units, all macroscopic dimensions are in mm, all angles in degrees and temperatures in Kelvin. No default values are assumed except where noted.

4.1. TITLE (A1, 66A1, 8A1)

This must be the first card of any entry.

Col.	Format		
1	A1	EOS	* This section may only contain 1 card and must therefore have an asterisk in col. 1.
2–67	66A1		Name of compound and other identification. Use only one card. Give other information in a REMARK section. This information may also be repeated in cols. 9–75 of the header card to allow visual identification of a card deck.
68–75	8A1		Entry number chosen by user to distinguish different entries in a multi-entry file. This field allows the user to identify and select the entry in which he is interested.

4.2. END

This mus be the last card of any entry. No data cards follow the END header.

4.3. REFERENCE (A1, 2A4, 1X, 65A1)

This section contains bibliographic information.

Col.	Format		
1	A1	EOS	* On the last card of the section
2–5	A4	CID	card identifier = JRNL, AUTH, TITL, RMRK or EOS (see below).
6–9	A4	DSK	Data set key (see *Introduction*).
10	1X		
11–75	65A1	DATA	Data whose contents depend on CID (see below).

The contents of DATA depend on CID as follows:

CID	DATA	
JRNL	Journal reference.	
11–12	2A1	Last 2 digits of date.
13–18	6A1	ASTM Journal Coden (given on cover of the journal or available from *Chem. Abstr. Source List*).
19–22	4A1	Volume number.
23–27	5A1	First page.
28–32	5A1	Last page.
33–75	43A1	Journal name.
AUTH	Authors' names, surname first, one name per card, use as many cards as necessary.	
TITL	Title of paper, use as many TITL cards as necessary.	
RMRK	Comments.	
EOS	End-of-section card. Contains no data.	

4.4. CELL DIMension (A1, 2A4, 1X, 6F10.4, F5.0)

Col.	Format		
1	A1	EOS	* On the last card of the section.
2–5	A4	CID	Card identifier, normally blank. ERRS: the data will be read as standard errors in the cell dimensions. EOS end-of-section card (contains no data).

6–9	A4	DSK	Data set key
10	1X		
11–20	F10.4	*A*	Unit-cell lengths in Å. All values must be given. This unit cell must correspond to the setting given in SPACE GROUP (§ 4.5).
21–30	F10.4	*B*	
31–40	F10.4	*C*	
41–50	F10.4	α	Unit-cell angles in degrees. All values must be given.
51–60	F10.4	β	
61–70	F10.4	γ	
71–75	F5.0	*Z*	The number of formula units (as given in FORMULA § 4.11) in the unit cell given on this card.

4.5. SPACE GRoup (3A1, 2X, A4, 2X, 11A1, 3A4)

Col.	Format		
1	A1	EOS	* Since this section should contain only 1 card there will always be an asterisk in column 1.
2	A1	LT	Lattice type P, A, B, C, F, I or H. Normally the first character of the Hermann–Mauguin space-group symbol but for rhombohedral space groups use .P for the rhombohedral setting, H for the hexagonal setting. Any program reading the file should generate the lattice translation operators from this symbol.
3	A1	CC	Center code, C = center of symmetry at the origin (A, N or other symbol = no center at origin). If C is specified, any program reading the file should automatically generate additional symmetry operators by inverting the operators given in SYMMETRY (§ 4.6) through the origin, *e.g.* if $x, -y, \frac{1}{2} + z$ is given, $-x, y, -\frac{1}{2} - z$ should be generated by the program.
4–5	2X		
6–9	A4	DSK	Data set key.
10–11	2X		
12–22	11A1	SG	Hermann–Mauguin space-group symbol for the setting actually used (see note below).
23–26	A4	XO	Origin shift in the form 1/8 – 3/8 1/4 *etc.* (4 characters per axis, right justified). This describes a vector from the origin given in the standard (*International Tables*) setting to the origin of the cell used in the description of the structure. The axis system in which the vector is given is that defined in the field SG above.
27–30	A4	YO	
31–34	A4	ZO	
35–75	41X		Undefined.

Note on the definition of the Hermann–Mauguin space-group symbols. If a SYMMETRY section (4.6) is given this symbol may be in any easily recognizable form. However, *if no SYMMETRY section is given the space-group symbol must adhere strictly to the following rules* to allow the symmetry operators to be unambiguously determined. If there is a discrepancy between the space-group symbol and the operators given in the SYMMETRY section, the symmetry section takes precedence. Give the symbol in the short form given in *International Tables for X-ray Crystallography* (1952) with the following conventions:

1. Left justify the space-group symbol.
2. Leave a space after the lattice type.
3. Leave a space between the symmetry symbols referring to the different directions.
4. Write $\bar{4}$ as –4 *etc.*
5. Write 2_1 as 21 *etc.*
6. Where alternative standard settings are given in *International Tables for X-ray Crystallography* the following conventions are assumed:

(*a*) For monoclinic space groups the *b*-axis setting is assumed unless otherwise indicated, *e.g.*:

P 21/N = P 1 21/N 1	(*b*-axis setting)
P 1 1 21/N	(*c*-axis setting)
P 21/N 1 1	(*a*-axis setting).

(*b*) In centrosymmetric space groups the setting with the origin at a centre of symmetry is assumed unless an origin shift is given in columns 23–34.

(*c*) The rhombohedral setting is assumed for rhombohedral space groups if LT = P, the hexagonal setting is assumed if LT = H.

7. If the Herman–Mauguin symbol is not the same as that given in *International Tables for X-ray Crystallography*, a cyclic permutation of axes is assumed in any case where there is an ambiguity, *e.g.* $P2_122$. Other permutations can be achieved through an origin shift.

8. Any other setting can be obtained by a suitable shift of the origin as given in cols. 23–34.

9. Some examples are given in Appendix II.

4.6. SYMMETRY (A1, 2A4, 1X, 3(3I2, F10.7, 4X))

Col.	Format		
1	A1	EOS	* On the last card of the section.
2–5	A4	CID	Normally blank. EOS if the last card of the section. An EOS card contains no data.
6–9	A4	DSK	Data set key.
10	1X		
11–12	I2	M11	
13–14	I2	M12	
15–16	I2	M13	
17–26	F10.7	T1	
27–30	4X		
31–32	I2	M21	
33–34	I2	M22	
35–36	I2	M23	
37–46	F10.7	T2	
47–50	4X		
51–52	I2	M31	
53–54	I2	M32	
55–56	I2	M33	
57–66	F10.7	T3	
67–75	9X		

The transformed coordinates (x′,y′,z′) are related to those given in the ATOM cards (x,y,z) by

$$x' = M11.x + M12.y + M13.z + T1$$
$$y' = M21.x + M22.y + M23.z + T2$$
$$z' = M31.x + M32.y + M33.z + T3.$$

In the event of a disagreement between the SPACE GRoup and SYMMETRY sections, the information in the SYMMETRY section will be assumed correct. The operators given here must correspond to the unit cell given in CELL DIMensions (§ 4.4) and the coordinates given in ATOM (§4.7).

4.7. ATOM (A1, 2A4, A2, A3, 6F8.5, I4, A1, A4, X4)

This section contains data on individual atoms.

Col.	Format		
1	A1	EOS	* On the last card of section.
2–5	A4	CID	Card identifier = ATCO, ATCE, UIJ, BETA, BIJ, UIJE, BETE, BIJE, or EOS (see below.)
6–9	A4	DSK	Data set key
10–11	A2	AN	Atom name, normally an element symbol left justified. (Iodine = I not J.)
12–14	A3	AI	Atom identifier, any legal characters.

These fields are used to associate data which appear on different cards but refer to the same atom.

15–22	F8.5	x_1	Parameters whose value depends on CID (see below).
23–30	F8.5	x_2	
31–38	F8.5	x_3	
39–46	F8.5	x_4	
47–54	F8.5	x_5	
55–62	F8.5	x_6	
63–66	I4	i	
67	A1	a_1	
68–71	A4	a_2	
72–75	4X		Undefined.

The cards may be grouped in any order, *e.g.* they may be ordered by CID or by AN and AI. The characters in cols. 10–14 must be the same on all cards referring to the same atom.

CID	Values of parameters	
ATCO	Atomic coordinates	
	$x_1 = x$	In fractions of unit cell
	$x_2 = y$	
	$x_3 = z$	
	$x_4 = U$	Isotropic temperature factor [$T = \exp(-8\pi^2 U \sin^2\theta/\lambda^2)$]
	$x_5 = \text{occ}$	Occupation number [≤ 1 (default = 1)]. For a fully occupied site $x_5 = 1$ even for special positions.
	x_6 = ionic charge or formal oxidation state.	
	i = Wyckoff position multiplicity (≥ 1).	
	a_1 = Wyckoff position letter.	
	a_2 = AT atom type. Any four characters used to define the form factor (see §4.8).	

ATCE	Errors in atomic coordinates.	
	$x_1 = \sigma(x)$	*N.B.* If these are written as integers (I8) they will be read as errors in the fifth decimal place.
	$x_2 = \sigma(y)$	
	$x_3 = \sigma(z)$	
	$x_4 = \sigma(U)$	
	$x_5 = \sigma(\text{occ})$	
UIJ	Anisotropic temeprature factors (U form)	
	$x_1 = U(11)$	As used in the expression $T = \exp\{-2\pi^2[\sum_i \sum_j U(ij)h_i h_j a_i^* a_j^*]\}$.
	$x_2 = U(22)$	
	$x_3 = U(33)$	
	$x_4 = U(12)$	
	$x_5 = U(13)$	
	$x_6 = U(23)$	
BETA	Anisotropic temperature factor (β form)	
	$x_1 = \beta(11)$	As used in the expression $T = \exp\{-[\sum_i \sum_j \beta(ij)h_i h_j]\}$
	$x_2 = \beta(22)$	
	$x_3 = \beta(33)$	
	$x_4 = \beta(12)$	
	$x_5 = \beta(13)$	
	$x_6 = \beta(23)$	
BIJ	Anisotropic temperature factor (B form)	
	$x_1 = B(11)$	As used in the expression $T = \exp\{-\frac{1}{4}[\sum_i \sum_j B(ij)h_i h_j a_i^* a_j^*]\}$.
	$x_2 = B(22)$	
	$x_3 = B(33)$	
	$x_4 = B(12)$	
	$x_5 = B(13)$	
	$x_6 = B(23)$	
UIJE		Standard errors in the values given on UIJ, BETA and BIJ cards respectively.
BETE		
BIJE		
EOS		Last card of section. Contains no data.

4.8. FORM FACtor (A1, 2A4, A2, A4, 6F10.6)

The form factor can be given in different ways according to the value of CID.

Col.	Format		
1	A1	EOS	* On the last card of the section.
2–5	A4	CID	Card identifier = blank, EXP1, EXP2 or EOS.
6–9	A4	DSK	Data set key.
10–11	A2	AN	Atom name (may be blank if AT is given).
12–15	A4	AT	Atom type (may be blank if AN is given, if both AN and AT are given they must both agree with the values given on the ATOM card).

16–25	F10.6	x_1	Parameters whose values depend on CID (see below).
26–35	F10.6	x_2	
36–45	F10.6	x_3	
46–55	F10.6	x_4	
56–65	F10.6	x_5	
66–75	F10.6	x_6	

CID	Values of parameters	
blank	$x_1 = \sin\theta/\lambda$	One card needed for each value of $\sin\theta/\lambda$ for each element or each atom type.
	$x_2 = f + f'$	
	$x_3 = f''$	

EXP1	$x_1 = a_1$	Two cards (EXP1 and EXP2) needed for each element or atom type. Parameters are used in the expression given below (see *International Tables for X-ray Crystallography*, 1974).
	$x_2 = b_1$	
	$x_3 = a_2$	
	$x_4 = b_2$	
	$x_5 = a_3$	
	$x_6 = b_3$	
EXP2	$x_1 = a_4$	
	$x_2 = b_4$	
	$x_3 = c$	
	$x_4 = f'$	
	$x_5 = f''$	

$$\text{form factor} = \sum_{j=1}^{4} a_j \exp[-b_j(\sin\theta/\lambda)^2] + c + f' + if''$$

EOS — Last card of section. Contains no data.

Form factors can be identified with atoms in one of two ways: either through the AN (normally an element symbol) or through AT (atom type). If both are given in this section both must agree with the values of AN and AT given in ATOM (§ 4.7). For neutron scattering lengths the EXP form can be used with $a_j = b_j = 0$ and c = scattering length.

4.9. HKL (A1, 2A4, 3I5, A1, 5F10.3)

This section contains information on Bragg reflections.

Col.	Format		
1	A1	EOS	* On the last card of the section.
2–5	A4	CID	Card identifier = blank, INT, CALC, PWDR, ISOM or EOS (see below).
6–9	A4	DSK	Data set key
10–14	I5	h	Miller indices of reflection
15–19	I5	k	
20–24	I5	l	
25	A1	a	Flag: 0 or 1 = normal reflection, 2 = unobserved reflection, 3 = unreliable measurement, 5 = space-group systematic absence.
26–35	F10.3	x_1	Parameters whose values depend on CID (see below).
36–45	F10.3	x_2	
46–55	F10.3	x_3	
56–65	F10.3	x_4	
66–75	F10.3	x_5	

CID	Values of parameters	
blank	x_1 |F(obs)|	Observed structure factor.
	x_2 σ[F(obs)]	Standard error in F(obs) – normally derived from counting statistics.
INT	To report intensity measurements	
	x_1 |F(obs)|	See above.
	x_2 σ[F(obs)]	
	x_3 I(net)	Observed intensity corrected for background but not absorption, extinction, *etc.*
	x_4 I (background)	I (observed) = I(net) + I(background).
	x_5 σ[I(net)]	Standard error in I(net) normally derived from counting statistics.
CALC	To report calculated structure factors	
	x_1 |F(obs)|	See above.
	x_2 σ[F(obs)]	
	x_3 |F(calc)|	Modulus of calculated structure factor.
	x_4 A	F(calc) = A + iB.
	x_5 B	
PWDR	To report powder patterns	
	x_1 I(obs)	Uncorrected intensity.
	x_2 σ(obs)	Standard error in I(obs).
	x_3 2θ	(Degrees).
	x_4 $\sigma(2\theta)$	Standard error in 2θ.
ISOM	To report data from isomorphs (*e.g.* of macromolecules).	
	x_1 |F(obs)|	See above.
	x_2 σ[F(obs)]	
	x_3 F_c	Amplitude and phase (degrees) due to isomorphous modification.
	x_4 α_c	

$$x_5\, g = \frac{\sigma^2[|F_o(hkl)| - |F_o(\bar{h}\bar{k}\bar{l})|]}{\sigma^2[|F_o(hkl)|] + \sigma^2[|F_o(\bar{h}\bar{k}\bar{l})|]}.$$

Structure factors for different isomorphs should use a different value for DSK but cards may be ordered within a section by *hkl*.

e.g.

CID	DSK	h	k	l	
ISOM	IS3	3	4	−5	+ data
CALC	NATV	3	4	−6	
ISOM	IS1	3	4	−6	
ISOM	IS2	3	4	−6	
ISOM	IS3	3	4	−6	
CALC	NATV	3	4	−7	

etc.

EOS — Last card of section. Contains no data.

4.10. CONDITIONs (A1, 2A4, A1, 6F10.0)

This section is used to give various data concerning the crystal and the conditions of measurement.

Col.	Format		
1	A1	EOS	* On the last card of the section.
2–5	A4	CID	Card identifier = blank, FACE, ORNT or EOS.
6–9	A4	DSK	Data set key.
10	A1	a	Parameters whose value depends on CID. Note that the x parameters may be written as integers and will be correctly read by F10.0 format. Other formats (*e.g.* F10.5) may be written and will also be correctly read if the decimal point is included.
11–20	F10.0	x_1	
21–30	F10.0	x_2	
31–40	F10.0	x_3	
41–50	F10.0	x_4	
51–60	F10.0	x_5	
61–70	F10.0	x_6	
71–75	5X		Undefined.

CID	Value of parameters	
blank	a	N = neutron diffraction, E = electron diffraction, X (or blank) = normal X-ray diffraction, S = synchrotron radiation.
	x_1 LAMBDA	Wavelength (Å).
	x_2 T	Temperature (K).
	x_3 SCALE	Scale factor. True F = SCALE* F given in HKL sections (§§ 4.9, 4.14 and 4.17) (default = 1·0).
	x_4 ABS	Linear absorption coefficient (mm^{-1}).
	x_5 RHO-M	Observed density.
	x_6 RHO-X	Calculated density.

FACE Used for defining the shape and size of the crystal.

x_1	D	Perpendicular distance from an arbitrary origin to a crystal face (mm).
x_2	h	Miller indices of the face.
x_3	k	
x_4	l	
x_5	χ	Polar coordinates of the normal to the face.
x_6	φ	

Note give (hkl) or $(\chi\varphi)$ but not both. One card is needed per crystal face.

If FACE cards are given, 3 ORNT cards must also be given. If χ and φ are given then ω is assumed to be the same as that on the first ORNT card and θ is assumed to be zero.

ORNT Used for defining the crystal orientation on the diffractometer.

a_1		Type of diffractometer, E = Enraf–Nonius, G = GE-XRD, H = Hilger Watts four circle, P = Picker, S = Syntex(Nicolet), W = Philips PW1100, X = other.
x_1	h	Miller indicates (not necessarily integral) of an arbitrary point in reciprocal space.
x_2	k	
x_3	l	
x_4	ω	Angular settings on diffractometer used corresponding to these Miller indices. Use the definition of ω, χ and φ normal for the diffractometer specified or define them in REMARK (§ 4.12).
x_5	χ	
x_6	φ	

At least 3 ORNT cards should be given to define the direction and sense of the axes.

EOS Last card of section. Contains no data.

4.11. FORMULA (A1, 2A4, 6(A2, F8.4))

Formula based on chemical analysis.

Col.	Format		
1	A1	EOS	* On the last card of the section.
2–5	A4	CID	Normally blank. EOS on the last card of the section if this card contains no data.
6–9	A4	DSK	Data set key.
10–11	A2		First element name.
12–19	F8.4		Relative number of atoms of first element present.
20–21	A2		Same for second element.
22–29	F8.4		
30–31	A2		Same for third element.
32–39	F8.4		
40–41	A2		Same for fourth element.
42–49	F8.4		
50–51	A2		Same for fifth element.
52–59	F8.4		
60–61	A2		Same for sixth element.
62–69	F8.4		

Use as many cards as necessary. The unit cell defined in § 4.4 should contain Z of these formula units.

4.12. REMARK (A1, 74A1)

Col.	Format		
1	A1	EOS	* On the last card of the section.
2–75	74A1		Any messages may be written in this section. Since it is difficult for computers to interpret this information, data should be given in other sections wherever possible.

4.13. ATOM COOrdinates (A1, A2, A3, A4, 3F8.5, 2F6.4, 3F6.5, 2F5.4)

This contains atomic coordinates in a compact form.

Col.	Format		
1	A1	EOS	* On the last card of the section.
2–3	A2	AN	Atom name (element symbol, left justified). This should normally be a recognized element symbol (use I not J for iodine).
4–6	A3	AI	Atom identifier (any legal characters).
7–10	A4	AT	Atom type (to identify form factors).
11–18	F8.5	x	Atomic coordinates in fractions of unit cell.
19–26	F8.5	y	
27–34	F8.5	z	
35–40	F6.4	U	Isotropic temperature factor $[T = \exp(-8\pi^2 U \sin^2\theta/\lambda^2)]$.
41–46	F6.4	OCC	Occupation number ($\leq$ 1·0), default = 1·0. For a fully occupied site this will be 1·0 even for special positions.
47–52	F6.5	$\sigma(x)$	Standard errors in x, y, z, U and OCC. *N.B.* If these are written as integers (I6 or I5), they will be read as errors in the fifth decimal place [fourth for $\sigma(U)$ and σ(OCC)].
53–58	F6.5	$\sigma(y)$	
59–64	F6.5	$\sigma(z)$	
65–69	F5.4	$\sigma(U)$	
70–74	F5.4	σ(OCC)	

4.14. HKL PACK (A1, 3I3, F8.2, F6.2, A1, 2(1X, 3I3, F8.2, F6.2, A1))

A compact form listing three structure factors per card.

Col.	Format		
1	A1	EOS	* On the last card of the section.
2–4	I3	h	Miller indices.
5–7	I3	k	
8–10	I3	l	
11–18	F8.2	$\lvert F(\text{obs})\rvert$	Observed structure factor for first reflection.
19–24	F6.2	$\sigma[F(\text{obs})]$	Error in F(obs).
25	A1	a	Flag [see HKL (§ 4.9) for definition].
26	1X		
27–29	I3	h	As above for second reflection.
30–32	I3	k	
33–35	I3	l	
36–43	F8.2	$\lvert F(\text{obs})\rvert$	
44–49	F6.2	$\sigma[F(\text{obs})]$	
50	A1	a	
51	1X		
52–54	I3	h	As above for third reflection.
55–57	I3	k	
58–60	I3	l	
61–68	F8.2	$\lvert F(\text{obs})\rvert$	
69–74	F6.2	$\sigma[F(\text{obs})]$	
75	A1	a	

4.15. ATOM MACromolecule (6A1, I5, 1X, A2, 3A1, A3, 1X, A1, I4, A1, 3X, 3F8.3, 2F6.2, 1X, I3, A4)

This section is designed for protein structures and is compatible with the format used in the Protein Data Bank. For further details see *Protein Data Bank File Record Formats* (1979).

Note that atomic coordinates are expressed in rectangular Cartesian coordinates in Å. The transformation matrix is given in the ORTHOGONal section. Since this format is compatible with the Protein Data Bank format the first six columns are used for the CID. The end of section is indicated using '*EOS' as a CID.

Col.	Format		Data
1–6	6A1	CID	ATOM, SIGATM, HETATM, TER or *EOS. ATOM records are used for atoms in the principle structural units. SIGATM records contain standard errors in coordinates. HETATM records are identical to ATOM records but are used for water and atoms in other HET groups. TER records use only columns 1–27 and occur after the terminal atom of each chain. * EOS is the end of section and contains no data.
7–11	I5		Atom serial number.
12	1X		
13–14	A2	AN	Atom name (element symbol, *right* justified).
15	A1		Remoteness indicator. } See Appendix *B* of *Protein Data Bank File Record Format.*
16	A1		Branch designator. } See Appendix *B* of *Protein Data Bank File Record Format.*
17	A1		Alternative location indicator.
18–20	A3		Residue name (see Appendix *C* of *Protein Data Bank File Record Format*).
			or
		AI	Atom identifier, any legal characters.
21	1X		
22	A1		Chain identifier.
23–26	I4		Residue sequence number.
27	A1		Code for insertion of residues.
28–30	3X		
31–38	F8.3	X	Atomic coordinates in Å [or
39–46	F8.3	Y	their standard errors (Å) if
47–54	F8.3	Z	CID = 'SIGATM'].
55–60	F6.2	OCC	Occupancy (standard error if CID = 'SIGATM'). Must be $\leq 1{\cdot}0$.
61–66	F6.2	B	Isotropic temperature factor (standard error if CID = 'SIGATM.)'
67	1X		
68–70	I3		Footnote number.
71–74	A4	AT	Atom type (to identify scattering factors). See ATOM (§ 4.7).

4.16. ORTHOGONal transformation (6A1, 1X, I3, 3F10.5, 5X, F10.5)

This section is compatible with the Protein Data Bank format and is used in conjunction with the ATOM MAC section.

Col.	Format		Card format (see below for explanation)
1–5	5A1	CID	SCALE, ORGIX, MTRIX, TVECT OR *EOS (*EOS = end of section card containing no data).
6	A1	LNUM	Line number of matrix (1, 2 or 3 below) (SCALE, ORGIX and MTRIX cards only).
7	1X		
8–10	I3	SNUM	Symmetry transformation number (MTRIX and TVECT cards only).
11–20	F10.5	M_{i1}	One row of matrix elements (i = LNUM).
21–30	F10.5	M_{i2}	
31–40	F10.5	M_{i3}	
41–45	5X		
46–55	F10.5	T_i	Translation element of matrix (i = LNUM).

This section contains the matrices useful in transforming the orthogonal coordinates given in the ATOM MAC section. If atomic coordinates are given in ATOM MAC at least three SCALE cards should be included to give the matrix that transforms orthogonal coordinates to crystal coordinates.

Each matrix requires three cards where the matrix is stored in columns 11–55 as follows:

$$\text{LNUM} = 1,\ X^1 = M_{11}X + M_{12}Y + M_{13}Z + T_1$$
$$\text{LNUM} = 2,\ Y^1 = M_{21}X + M_{22}Y + M_{23}Z + T_2$$
$$\text{LNUM} = 3,\ Z^1 = M_{31}X + M_{32}Y + M_{33}Z + T_3.$$

For SCALE cards $M = \underset{\sim}{S}$ and $T = \bar{U}$.
For ORGIX cards $M = Q$ and $T = \bar{T}$.
For MTRIX cards $M = \underset{\sim}{M}$ and $T = \bar{V}$.
For TVECT cards $M_{1i} = \bar{W}_i$.

The fractional (crystallographic) coordinates $\bar{x}$ are related to the orthogonal coordinates $\bar{X}$ given in the ATOM MAC section by $\bar{x} = \underset{\sim}{S}\bar{X} + \bar{U}$. The coordinates originally deposited in the Protein Data Bank ($\bar{X}_o$) are given by $\bar{X}_0 = Q\bar{X} + \bar{T}$.

When the asymmetric unit contains atoms related by non-crystallographic symmetry these can be generated by $\bar{X}' = \underset{\sim}{M}\bar{X} + \bar{V}$. In a polymeric structure the full structure can be generated by applying successively the vector given on the TVECT card to the coordinates in ATOM MAC $\bar{X}' = \bar{X} + \bar{W}$. Each MTRIX and TVECT card must contain a serial number in SNUM, each operator having a different value of SNUM.

The TVECT card should be blank in columns 46–55.

For further details of the use of these matrices see Appendix *A* of the *Protein Data Bank File Record Formats* (1979).

4.17. HKL PROTein (A1, 3I4, A4, 2(F6.0, F4.0), 3I3, 4F5.0, I6, I3)

This section is designed for protein structure factors.

Col.	Format		
1	A1	EOS	* On the last card of the section.
2–5	I4	*h*	Miller indices.
6–9	I4	*k*	
10–13	I4	*l*	
14–17	A4	DSK	
18–23	F6.0	$\|F(\text{obs})\|$	Observed structure factor.
24–27	F4.0	$\sigma\|F(\text{obs})\|$	Standard error in F(obs).
28–33	F6.0	Δ	$\|F(\text{obs})_{hkl}\| - \|F(\text{obs})_{hkl}\|$.
34–37	F4.0	$\sigma\Delta$	Standard error in Δ.
38–40	I3	αp	Most probable phase of native protein.
41–43	I3	αb	Best (*i.e.* centroid) phase of native protein.
44–46	I3	*m*	Figure of merit (×100).
47–51	F5.0	*A*	Coefficients in the expression log $P(\alpha) = A \cos\alpha + B \sin\alpha + C \cos 2\alpha + D \sin 2\alpha$ scaled by scale factor given in the CONDITIOns section.
52–56	F5.0	*B*	
57–61	F5.0	*C*	
62–66	F5.0	*D*	
67–72	I6	$\|F(\text{calc})\|$	Calculated structure factor and phase (in degrees) for native protein.
73–75	I3	$\alpha(\text{calc})$	

See comment below.

By a suitable choice of the scale given in the CONDITIOn section, these numbers can be given as integers if space is short. Fields given in F*n*.0 should normally contain only integers, would not include the decimal point and could be read as I*n*. Correctly scaled values are calculated by multiplying the integers given in these fields by the scale factor given in CONDITIOns (§ 4.10).

APPENDIX I
Sample of a standard crystallographic file

```
1234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890
TITLE      THALLIUM TRIACETATE
*THALLIUM(III) TRIACETATE                                                    001

REMARK
*  THIS IS AN EXAMPLE OF A STANDARD CRYSTALLOGRAPHIC FILE

REFERENCE
 JRNL      78ACRCAR  34 2845 2846   ACTA CRYST B
 AUTH      FAGGIANI R
 AUTH      BROWN I D
 TITL      THALLIUM(III) ACETATE
*EOS

CELL DIMENSION   A          B          C        ALPHA      BETA       GAMMA      Z
*ERRS          15.5400    9.6300     7.8480    90.       113.9200    90.        4
                  100         70         60                500
SPACE GROUP
*CC          C 2/C

SYMMETRY
*          1 0 0              0 1 0              0 0 1
          -1 0 0              0 1 0              0 0-1 0.500

ATOM          X          Y          Z                    VALENCE
 ATCD  TL     0.0      0.56330   0.25                      3.          4E
 ATCE  TL                  10
 ATCD  O 1    0.07070  0.79360   0.25380                   -2.         8F
 ATCE  O 1      90        160       180
 ATCD  O 2    0.05660  0.38460   0.10180                   -2.         8F
 ATCE  O 2      70        140       150
 ATCD  O 3    0.15010  0.45090   0.37770                   -2.         8F
 ATCE  O 3      90        180       200
 ATCD  C 1    0.0      0.86480   0.25                      4.          4E
 ATCE  C 1                270
 ATCD  C 2    0.0      0.04730   0.25                      4.          4E
 ATCE  C 2                420
 ATCD  C 3    0.21250  0.27620   0.21950                   4.          8F
 ATCE  C 3      130       260       290
 ATCD  C 4    0.13610  0.37270   0.23580                   4.          8F
 ATCE  C 4      120       200       260
 UIJ   TL     0.0321   0.04520   0.0460   0.0      0.0167
 UIJ   O 1    0.0590   0.0720    0.069   -0.0050   0.0310  -0.0050
 UIJ   O 2    0.0420   0.0690    0.0500   0.0180   0.0140   0.0120
 UIJ   O 3    0.0480   0.1000    0.0430   0.0170   0.0210  -0.0100
 UIJ   C 1      7500      4200      3100        0      2100         0
 UIJ   C 2     14400      6600     13500        0      5800         0
 UIJ   C 3      5200     11000      8100     4500      3300      1400
 UIJ   C 4      5700      4900      6100     1500      2900       300
*EOS

HKL        PARTIAL LIST
 CALC        3    1    -6      96.        3.       103.
 CALC        3    1   -72      38.        5.        26.
 CALC        3    1    -7      42.        6.        41.
 CALC        3    1    82      32.        6.        36.
 CALC        3    1    -8      44.        5.        49.
*EOS

CONDITIONS LAMBDA      TEMP(K)          ABS              D(CALC)
           X 0.71069   293.             16.7              2.57
*EOS

FORMULA
*          TL  1.0   C   6.0   O   6.0   H   9.0

END
1234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890
```

APPENDIX II
Examples of space-group descriptions

```
1234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890
 IA    0001 I 41/A M D     0 1/4-1/8                FIRST SETTING
 IC    0001 I 41/A M D                              SECOND SETTING
 RC    0001 R -3 C                          O AT 2B, RHOMBOHEDRAL SETTING
 RA    0001 R -3 C       1/4 1/4 1/4        O AT 2A, RHOMBOHEDRAL SETTING
 HC    0001 R -3 C                          O AT 2B, HEXAGONAL SETTING
 HA    0001 R -3 C         0   0 1/4        O AT 2A, HEXAGONAL SETTING
1234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890
```

References

International Tables for X-ray Crystallography (1952). Vol. I. Birmingham: Kynoch Press.

International Tables for X-ray Crystallography (1974). Vol. IV, p. 71. Birmingham: Kynoch Press.

Protein Data Bank File Record Formats (1979). Department of Chemistry, Brookhaven National Laboratory, Upton, NY 11973, USA.

THOMAS, J. O. (1978). *Acta Cryst.* A**34**, 819–823.