Acta Cryst. (1988). A44, 401-410

Crystallographic Publishing in Retrospect and Prospect

BY S. C. ABRAHAMS AND R. A. MATULA

AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA

(*Received 26 January 1988; accepted 18 March 1988*)

Abstract

Long-term trends in the number of crystallographic papers published annually in all journals during the last 20 years differ substantially from those in the International Union of Crystallography (IUCr) journals during the same period. The IUCr journals published about 7·5% of all crystallographic papers in 1968, a proportion that increased to about 11% in 1987. The total number of papers in all fields abstracted by Chemical Abstracts Service increased annually by about 6·5% before 1979, but only by about 1·5% thereafter. The breadth of crystallographic coverage in the IUCr journals is comparable to and generally more consistent than that of other leading journals, but many important areas are severely under-represented in comparison with the coverage at the last IUCr Congress of Crystallography. Subscription costs for 15 journals that publish the largest number of crystallographic papers varied in 1986 from 2·3 to 11·8 US cents per 1000 characters, excluding journals with page charges. *Acta Crystallographica* cost 3·5 US cents per 1000 characters in 1986. The mean publication time for the only monthly IUCr journal is generally under 4 months, comparing well with the leading competitive journals publishing crystallographic papers. Relationships with the crystallographic databases continue to evolve constructively. Innovative publishing methods of relevance to all journals include the use of compuscripts, automated optical character recognition, machine-readable binary-coded print information, electronic publication and other non-print article distribution. The state of the IUCr journals is expected to remain robust as the total number of high-quality papers published continues to increase and the extent of crystallographic coverage broadens further.

Introduction

Acta Crystallographica's fortieth anniversary provides an opportune juncture for examining the extent to which one of the four Objects adopted at the First General Assembly of the International Union of Crystallography (1948), 'to promote international publication of crystallographic research and of crystallographic works', has been realized. This objective was amplified by the first Editor and principal leader in founding the journal in his opening Editorial Preface (Ewald, 1948). Following a broad description of crystallography in terms of ten examples of major development at that time, Ewald noted that *Acta* was 'intended to offer a central place for publication and discussion of all research in this vast and ever-expanding field.' It is noteworthy that the previous 'central' journal, the *Zeitschrift für Kristallographie*, ceased publication early in 1945 and did not resume publication again until late in 1954. Ewald had been an editor of *Z. Kristallogr.* until the outbreak of World War II.

The International Union of Crystallography (IUCr) initiated several 'crystallographic works' in 1948 in addition to *Acta*. The most important of these were *Structure Reports*, in succession to the pre-war *Strukturbericht* series, and *International Tables for X-ray Crystallography*, in succession to the 1935 *International Tables for the Determination of Crystal Structures*. Emphasis in this paper hereafter is entirely on journal publication and is based on a talk presented at an Open Meeting of the IUCr Commission on Journals in Perth (Abrahams, 1987).

The number of papers published in successive volumes of *Acta* is considered below both in terms of total world output of crystallographic papers and of all papers abstracted by Chemical Abstracts Service. The breadth of subject coverage is also considered. Attention is given to some factors that influence an author's choice of journals, and a brief analysis of comparative journal costs is presented. The results of closer interaction between the IUCr's journals and the crystallographic databases, and the possible consequences of introducing currently available technological advances, are considered.

Bibliometric analysis

The number of papers published in a given field and year was obtained by interrogating the CA SEARCH (CAS) database, produced by Chemical Abstracts Service and maintained on-line by the DIALOG® Information Retrieval Service. CAS was selected as the most comprehensive of the available databases. Most references in CAS to crystallographic papers published in the general literature are placed in

0108-7673/88/040401-10$03.00

Table 1. *Identification of Subsections used in CAS Section* 75

(Subsections applicable from 1982 to present with 1, 9 and 11 omitted; see also *Bibliometric analysis*)

0	Reviews
2	Crystal morphology, orientation, crystallinity, non-epitaxial film deposition.
3	Crystal defects, color centers, twinning, grain boundaries.
4	Mechanical properties of crystals.
5	Lattice dynamics and energetics.
6	Other crystal properties.
7	Polytypism, polymorphism, crystal phase transitions, ordering.
8	Crystal structures.
10	Crystallographic methods and apparatus for structure determination.
12	Other.

Section 75, entitled 'Crystallography and Liquid Crystals'. However, many equally relevant papers are placed in other CAS Sections, and the totals obtained for a given year are hence a function of the detailed search procedure.

The method used in the present study, in order to achieve a balance between the unwanted inclusion of irrelevant papers in the total count and the undesirable exclusion of normal crystallographic papers assigned an unexpected CAS indexing term, was to take the entire content of Section 75 from 1982 to the present less those papers in Subsection 1 (crystallization, nucleation, growth, epitaxy, recrystallization), Subsection 9 (experimental determination of structures of amorphous and vitreous substances by crystallographic methods) and Subsection 11 (liquid crystals). Only Subsection 1 was omitted for the period 1972 to 1981; the entire Section was taken for the period 1967 to 1971, since subsections were not then identified. Section 70 was assigned to 'crystallization and crystal structures' prior to 1975. In addition, all papers for all periods in other CAS Sections that contain 'crystal structure' or 'crystal structure types' or 'crystal structure determination' or 'crystal structure-property relationship' or 'crystallog?' as controlled vocabulary terms; or 'crystal structure(s)' as an identifier; or 'X-ray', 'electron', or 'neutron' with 'diffraction'; also 'X-ray' or 'neutron' with 'scattering' as free text terms; or 'X-ray diffractometer(s)' or 'neutron diffractometer(s)' as additional free-text terms; or 'charge density' in the title were included. Finally, all references to books, conference proceedings, dissertations, patents or technical reports were excluded. The Subsections in Section 75 used in this study are listed in Table 1. The distribution of crystallographic papers among the various journals depends upon the use of Codens; since these often change with time due to changes in volume titles, volume divisions *etc.*, particular care is required in their use. Codens are unique six-letter code representations of journal titles; for examples see *Notes for Authors* (1983). The first year of CA SEARCH was 1967; hence, no results taken from this source can originate earlier.

Variations of the search procedure outlined above would necessarily result in different totals but are not expected to change any conclusions. A bibliometric and citation analysis of the 1972–1976 crystallographic literature was reported by Hawkins (1980).

Crystallographic papers published

The total number of refereed papers published annually in the IUCr journals (*i.e.* excluding the categories IUCr Announcements, Crystallographers, New Commercial Products and Book Reviews) since their inception, as given in the reports of the IUCr Commission on Journals published each year in *Acta* (*e.g.* International Union of Crystallography, 1986), is shown in Fig. 1. It may be noted that these totals are not proportional to the number of pages published, since the average length of a paper has decreased from 4·5 pages in 1948 to 3·5 pages in 1986 and most unnecessary blank space on the pages of *Acta* has been eliminated in the last decade for economic reasons. It is interesting to note that the entire content of *Acta* in its first year was only 76 papers, which rose in number to a plateau of about 230 papers published annually between 1952 and 1960. As crystallographic computations became increasingly performed on digital computers, and the world population of crystallographers (defined here as scientists listed in the current *World Directory of Crystallographers*; the number of entries may not be completely reliable, owing to variations in national inclusion policies) grew (see Table 2), a prolonged climb in the rate of publication also began. A further increase may be associated with the division of *Acta* into two separate sections (A and B) and the launching of *J. Appl. Cryst.* in 1968. By contrast, the later division of *Acta* B into *Acta* B and *Acta* C in 1983 (Abrahams, 1982) had little immediate effect on the manuscript flow.

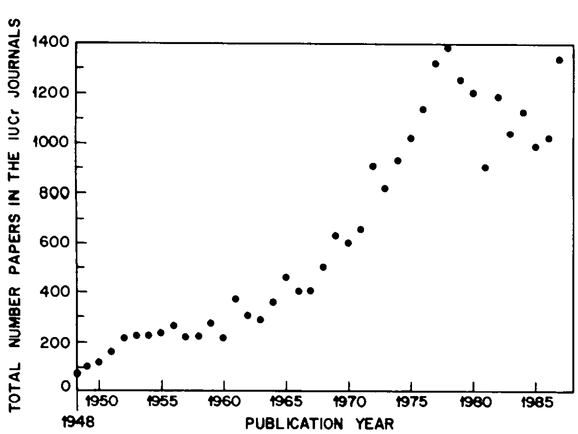

Fig. 1. Total number of refereed papers published annually in *Acta* and *J. Appl. Cryst.* by publication year. See text concerning the totals for 1981 and 1986.

Table 2. *World population of crystallographers*

WDC edition*	Year	Number	Countries
1	1957	2208	52
2	1960	3557	54
3	1965	5037	51
4	1971	6982	57
5	1977	7638	71
6	1981	8174	68
7	1986	8968	69

* WDC: *World Directory of Crystallographers.* Dordrecht: Reidel.

The total number of papers published in the IUCr journals reached a yearly maximum of about 1350 in 1977 and 1978, declining thereafter until 1987, for which the total is 1328. The low point of 906 papers in 1981 was a direct result of the page restriction that had to be imposed upon *Acta* in response to the substantial deficit incurred in 1980; see International Union of Crystallography (1983). The number published in 1986 was adversely affected by the resignation of the Technical Editor and the death in the same year of the Deputy Technical Editor, which led to slower processing than usual of accepted manuscripts.

The overall trends in Fig. 1 may be compared with the annual variation in number of papers published in *Acta* A and *J. Appl. Cryst.*, as given in Fig. 2. Both *Acta* A and *J. Appl. Cryst.* also reached maxima in 1976–1980 of about 192 papers for the former and about 168 for the latter in 1978–1979. Both declined thereafter to about 100 papers each in 1986, following essentially the same trend as in Fig. 1.

Consideration of the total number of papers abstracted by CAS yearly since 1967, as presented in Fig. 3 (open circles), places the trends shown in Figs. 1 and 2 in sharper perspective. Attempts to ascertain the actual number of journals abstracted annually, and also the annual variation, if any, in the effectiveness with which the abstracting service covered the literature, and hence the stability of the source-material base, were unsuccessful. The rate of growth in the total numbers shown in Fig. 3 is not constant over the period 1967–1986, but appears to change in slope around 1977–1979. Before the break, the annual growth was about 6·5%, thereafter becoming about 1·5%. Contrastingly, the IUCr rate of growth (Fig. 1) in the 1967–1977 period was about 22% annually, declining thereafter over the period 1978–1987 by about 0·6% annually (or by a 3·3% annual decline over the period 1978–1986).

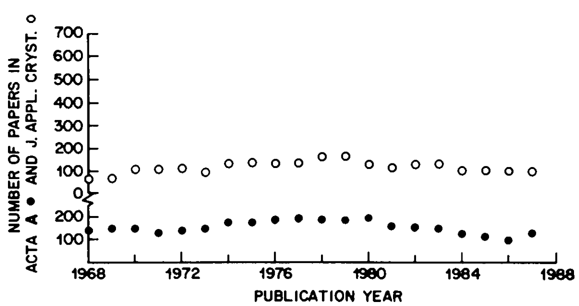

Fig. 2. Number of refereed papers published annually in *Acta* A (closed circles) and, separately, in *J. Appl. Cryst.* (open circles), by publication year.

A further comparison may be made by considering the annual world output of crystallographic papers, taken as the entire content of Section 75 of CAS each year but modified as noted above in *Bibliometric analysis*, also presented in Fig. 3 (closed circles). The initial increase of about 11% annually from 1968–1972 is only half that experienced by the IUCr journals, but approaches twice the annual increase in all papers abstracted by CAS in this period. However, between 1972 and 1974 the output decreased, changing thereafter to a steady increase of about 3% annually through 1986; this increase includes crystallographic papers published in biological journals.

The proportion of the annual world output of crystallographic papers published in the IUCr's journals has risen from about 7·5% in 1968 to about 11% in 1987.

Principal journals publishing crystallographic papers

Hawkins (1980) compiled a list of 15 journals that published the largest number of crystallographic papers between 1972 and 1976. A re-examination of

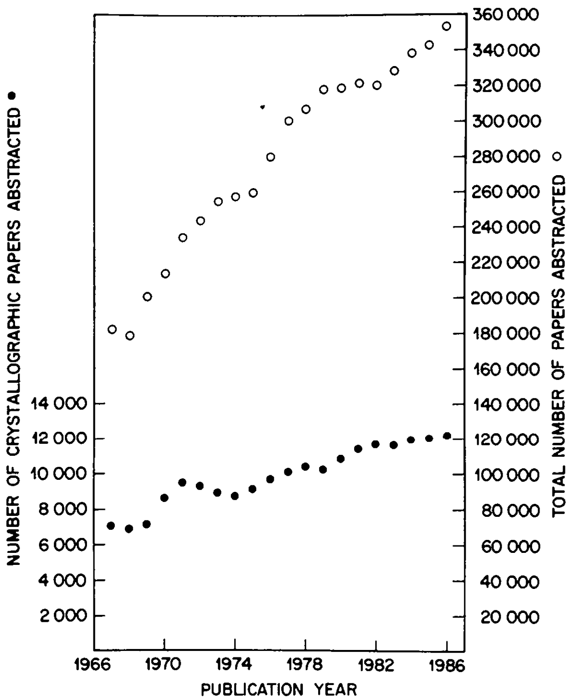

Fig. 3. Number of crystallographic papers given in CAS Section 75 with modifications noted in text (closed circles), and total number of papers in all fields abstracted by CAS (open circles) from all journals examined, by journal publication year.

Table 3. *Journals with most crystallographic papers in 1986*

Journal	Number of papers	Journal	Number of papers
Acta Cryst. C	726	*J. Org. Chem.*	104
Inorg. Chem.	379	*Izv. Akad. Nauk SSSR, Neorg. Mater.*	99
J. Am. Chem. Soc.	352	*J. Phys. (Paris) Colloq.*	95
J. Organomet. Chem.	291	*Solid State Commun.*	94
Phys. Rev. B	261	*Phys. Status Solidi B*	91
J. Chem. Soc. Chem. Commun.	234	*Zh. Strukt. Khim.*	90
J. Chem. Soc. Dalton Trans.	230	*Dokl. Akad. Nauk SSSR*	88
Physica B+C *(Amsterdam)*	202	*J. Cryst. Growth*	88
Inorg. Chim. Acta	198	*Z. Kristallogr.*	85
J. Less-Common Met.	158	*Zh. Neorg. Khim.*	85
J. Solid State Chem.	144	*J. Appl. Phys.*	84
Kristallografiya	143	*Acta Cryst.* A	83
Z. Anorg. Allg. Chem.	142	*J. Appl. Cryst.*	81
Angew. Chem.	139	*J. Crystallogr. Spectrosc. Res.*	80
Phys. Status Solidi A	125	*Acta Cryst.* B	79
Z. Naturforsch. Teil B	125	*Cryst. Res. Technol.*	79
Fiz. Tverd. Tela (Leningrad)	110		
J. Phys. C	104		

See *Bibliometric analysis* section for retrieval method.

the crystallographic literature published in 1986 based on retrieval from Section 75 of CAS, modified as in *Bibliometric analysis* above, leads to the ordered list presented in Table 3.

Acta Cryst. C is clearly the journal that publishes most crystallographic papers. It should be noted that the number of papers presented in Table 3 by the retrieval method used is six fewer than actually appeared in 1986. Similar underestimates are found for *Acta* A, *Acta* B and *J. Appl. Cryst.* of 15, 11 and 23 papers respectively. An examination of the unretrieved papers revealed three causes for the underestimation. One was the difficulty in selecting search terms for papers that lie on the border between crystallography and another science. A second was lack of adequate indexing by the abstracting services for some papers that should clearly be described as crystallographic. The latter suggests that closer cooperation between the abstracting services and the crystallographic community would improve the indexing of papers published either in the core journals or elsewhere. A third was the possibility that indexing for 1986 was not complete for all journals in the penultimate CAS update for 1987.

The two other Sections of *Acta* rank 30th and 33rd, with *J. Appl Cryst.* in 31st position; see Table 3. It may be noted that, of the two other 'traditional' crystallographic journals, *Kristallografiya* occupied 12th rank in 1986 whereas *Z. Kristallogr.* was in 27th place; furthermore, both *J. Crystallogr. Spectrosc. Res.* and *Cryst. Res. Technol.* were closely comparable to the three smaller IUCr journals in number of papers published in 1986. The 34 journals listed in Table 3 collectively contained 5468 abstracted crystallographic papers in 1986, about 45% of the total retrieved. Numerous other journals published fewer than 80 crystallographic papers last year. The biological journal with most crystallographic papers was *J. Mol. Biol.* with 61 papers. Hawkins (1980) reported that the 34 highest-ranking journals (in number of crystallographic papers published) for the period 1972–1976 collectively contained about 65% of the total.

Other journals publishing a high percentage of crystallographic papers include *J. Mol. Struct.*, with 61 such papers as defined above in 1986 and *Mol. Cryst. Liq. Cryst.* with 36 such papers. Very recently two additional journals of crystallographic interest have been launched. *Powder Diffraction*, in 1986, covers the practical techniques of its speciality and contained 37 crystallographic papers in its first year. *Crystallography Reviews* was inaugurated in August 1987.

Breadth of published crystallographic research

Consideration of the extent to which the IUCr journals have succeeded in becoming 'the central place' for all crystallographic publications (see *Introduction*) requires an investigation not only into the numerical distribution of such papers among the major journals (see Table 3) but also one into the degree of representation by all major categories of crystallography in these journals. Rather than attempt a determination of such breadth as a widely ranging function of both publication year and many major journals, this study has been restricted to the last three publication years and to the IUCr's and five other journals, partly because of changing crystallographic categories with time and partly due to space limitations. A pragmatic determination of the major categories in crystallography of high current interest has been taken for this study from the principal groupings of papers presented at the Fourteenth International Congress of Crystallography in Perth, 12–20 August 1987, as given in Table 4, based on the generally accepted premise that such Congresses best represent the entire field. See the *Communicated Abstracts* (1987) for all papers therein presented.

The following results regard nearly all categories in Table 4 as forming separate fields; categories 2 and 3, however, were combined since they could not be distinguished bibliometrically, as also were categories 4 and 5 and categories 17 and 18. The categories listed in Table 4 were augmented by additional characteristic descriptive terms common to papers within each field, as found in the *Communicated Abstracts* (1987). The on-line record for each of the 969 crystallographic papers retrieved from *Acta* and *J. Appl. Cryst.* in 1986 (see Table 3), as provided by the abstracting service, could then be searched for a match with these descriptive terms. Each paper presented at the Congress was assigned by the IUCr Program Committee to a single category; hence a similar assignment procedure was initially adopted in the search and match process.

Table 4. *Major categories of current crystallographic research with percentage distribution of each at the* 14*th IUCr Congress and in the IUCr journals for* 1984, 1985 *and* 1986

Category	14th IUCr Congress*	IUCr journals 1984	1985	1986
1 Determination of Macromolecular Structures	2·4	1·2	1·4	1·2
2 Structural Molecular Biology	10·1	5·6	5·5	5·3
3 Crystallography in Biochemistry and Pharmacology	6·1			
4 Atomic Scale Mechanisms and Chemical Properties	4·5	4·1	4·5	4·6
5 Physical Properties and Structure	7·3			
6 Charge, Spin and Momentum Densities	3·3	0·9	0·8	1·1
7 Materials Science	9·6	1·6	1·3	1·0
8 Inorganic and Mineralogical Crystallography	6·1	26·0	26·0	26·8
9 Structures of Organic, Organometallic and Coordination Compounds	11·6	41·4	38·9	42·3
10 Polymer Crystallography	0·5	1·0	0·8	0·5
11 Real and Ideal Crystals	8·1	1·9	2·3	2·2
12 Advances in Powder Diffraction	4·4	1·1	1·4	0·9
13 Neutron Diffraction	2·0	2·4	2·6	1·8
14 Electron Diffraction and Electron Microscopy	4·2	1·2	1·1	1·2
15 Utilization of Synchrotron Radiation	2·3	0·5	0·5	0·6
16 Apparatus and Techniques	3·3	9·3	10·3	8·7
17 Computational Methods and Error Analysis	5·1	0·9	1·2	0·9
18 Computing	1·8			
19 Data Retrieval and Crystallographic Teaching	2·0	0·0	0·0	0·2
20 Symmetry and its Generalization	3·8	0·8	1·0	0·5
21 Structural Results from Methods other than Diffraction	0·2	—	—	—
22 Gases, Liquids and Amorphous Compounds	1·2	0·1	0·4	0·2
Residual†	—	55	38	25
Total‡	—	1081	952	969

* Percentage of papers in each category of the total number presented, including Main Lectures and post-deadline abstracts.
† Number of papers not assigned to one of the 22 categories.
‡ Total number of papers abstracted by CAS as of late 1987.

Starting with category 1, 18 papers were found in that field among the original 969, corresponding to 1·9% of the total. Eliminating these 18 papers, categories 2 and 3 were thereupon searched and the matching papers again identified and eliminated. Each subsequent category was similarly searched, following the elimination of all previously matched papers. This approach, however, gives a quantitative distribution that is strongly dependent upon the order in which the categories are searched, since many papers may correctly be placed in more than one category.

The final approach adopted in assessing the breadth of crystallographic coverage counts each of the n matches that a given paper makes with the categories in Table 4, assigning to each such category the fraction $1/n$ for that paper. The sum of these fractions for all 969 papers is presented in Table 4 in the form of a percentage. It is notable that the two approaches, which result in a difference as large as 12·2% (for category 9), give a rather similar order of category population. The latter approach was hence considered to be reliable and was also applied to the 952 papers retrieved in 1985 and to the 1081 papers retrieved in 1984, with a percentage distribution by category for each year as given in Table 4.

Comparison of the distribution by category in the IUCr journals with that presented at the 14th International Congress shows that papers in many fields are not well represented in the IUCr journals (see Table 4). Categories 8, 9, 16 and, to a lesser extent, categories 10 and 13 are over-represented and all others are under-represented. It may be noted that the distribution in each of the three years sampled is remarkable constant. The distribution range for papers published in 1984 to 1986 in the next three leading journals, as shown in Table 3, and also the distribution range in the other two journals that traditionally seek to cover the whole field of crystallography, were determined similarly and are presented for comparison in Table 5.

The three chemistry journals in Table 5 exhibit, not unexpectedly, great strength in categories 8 and 9 with categories 4 and 5 also moderately strong. Although numerically weak, category 10 is relatively almost as strong or stronger in all three journals as at the International Congress. The remaining field of comparable relative strength is category 22, for *Inorg. Chem.* and *J. Am. Chem. Soc.* These chemistry journals publish few crystallographic papers in the remaining fields defined in Table 4.

Kristallografiya experiences moderately wide variations from year to year in its distribution of fields. Category 8 is most heavily represented in this journal on a consistent basis, with categories 4 and 5, 16, 9 and 11 also rather strong, followed by category 7.

Z. Kristallogr. emphasizes inorganic and mineralogical crystallography, with categories 9, 4 and 5, 16 and 7 also often rather strong. It is notable that this journal has shown a steady decline in total number of papers recently, with 118 retrieved in 1984, 104 in 1985 and 85 in 1986.

It is apparent from Table 5 that no journal attracts papers from all crystallographic fields as effectively as the International Congresses. Fields from which

Table 5. *Distribution of category percentage range in* 1984–1986 *for leading journals publishing crystallographic papers*

Category	IUCr	*Inorg. Chem.*	*J. Am. Chem. Soc.*	*J. Organomet. Chem.*	*Kristallografiya*	*Z. Kristallogr.*
1	1·2–1·4	0·0–0·1	0·5–0·6	0·0	0·0–0·5	0·0
2 and 3	5·3–5·6	0·9–1·1	3·4–5·6	0·0–0·2	0·8–1·4	0·7–1·6
4 and 5	4·1–4·6	14·9–16·2	9·8–12·1	12·1–12·2	9·4–20·8	4·6–16·3
6	0·8–1·1	0·9–1·5	1·6–2·3	0·3–0·5	0·0–1·2	0·0–2·1
7	1·0–1·6	0·7–1·3	0·7–1·3	0·0–0·2	5·4–7·4	4·2–6·2
8	26·0–26·8	39·8–40·8	29·2–30·4	30·3–32·6	26·6–38·0	43·4–47·2
9	38·9–42·3	34·6–36·9	43·3–44·5	51·4–53·3	8·6–15·0	14·7–25·6
10	0·5–1·0	0·3–0·9	0·6–1·5	0·4–1·0	0·0–0·4	0·0–4·7
11	1·9–2·3	0·2–0·7	1·3–1·9	1·3–1·5	8·9–11·3	1·5–2·3
12	0·9–1·4	0·1–0·5	0·0	0·0–0·2	0·4–0·5	0·8–2·3
13	1·8–2·6	0·6–1·2	0·3–1·0	0·0–0·4	0·5–1·7	0·8–2·3
14	1·1–1·2	0·5–0·8	0·5–0·9	0·2–0·6	0·5–3·3	0·8–1·4
15	0·5–0·6	0·0	0·0	0·0	0·0–0·4	0·0
16	8·7–10·3	1·1–1·2	1·7–2·5	0·5–1·3	9·8–12·1	6·2–7·0
17 and 18	0·9–1·2	0·0	0·0–0·2	0·0	1·6–4·2	0·0–1·5
19	0·0–0·2	0·0	0·0	0·0	0·0	0·0–0·8
20	0·5–1·0	0·1–0·3	0·1–0·5	0·0	1·2–3·3	0·0–1·4
22	0·1–0·4	1·2–1·3	1·3	0·2–0·6	1·9–4·1	0·0–1·4
Residual*	118	4	9	54	35	32
Total†	3002	1081	974	797	435	307

* Number of crystallographic papers published in 1984–1986 not assigned to one of the 22 categories in Table 4.
† Total number of crystallographic papers published in 1984–1986 and abstracted by CAS as of late 1987.

Table 6. *Comparative subscription costs in* 1986 *of* 15 *journals publishing crystallographic papers*

Journal	Number of pages (1986)	Characters/page	1986 cost (US $)	US cents/ 1000 characters
Acta (A+B+C)	3120	5200	571	3·5
Angew. Chem.	1134	5800	299	4·5
Inorg. Chem.	4910	7300	399	1·1
Inorg. Chim. Acta	3121	5100	980	6·2
J. Appl. Cryst.	500	4900	136	5·6
J. Am. Chem. Soc.	8315	7500	299	0·5
J. Chem. Soc. Dalton Trans.	2694	6900	635	3·4
J. Less-Common Met.	4188	2900	1429	11·8
J. Phys. C	8180	3400	840	3·0
J. Solid State Chem.	2069	3300	530	7·8
Physica B+C *(Amsterdam)*	2205	3900	563	6·5
Phys. Rev. B	17904	6300	850	0·8
Phys. Status Solidi A	5607	3500	695	3·5
Z. Anorg. Allg. Chem.	2648	2900	175	2·3
Z. Naturforsch. Teil B	1594	4500	215	3·0

the IUCr journals should clearly seek greater representation are those indicated in Table 4 by categories 2 and 3, 4 and 5, 7 and 11, and also categories 6, 12, 14, 15, 17 and 18, 19, 20 and 22. The breadth of crystallographic coverage in the IUCr journals is comparable to that in the other journals presented in Table 5, with smaller variations over the three-year period examined.

Comparative journal costs

Most libraries, faced with a continued proliferation of new journals and rising subscription costs, have found it necessary to eliminate multiple subscriptions; many have cancelled complete subscriptions. In consequence, most journals are confronted with an eroding subscription base at a time that production costs continue to rise. It is instructive, in these circumstances, to examine the cost of subscribing to the IUCr journals in comparison with journals that Table 3 shows to be direct competitors.

In comparing subscription costs for different journals, it is important to choose a unit that allows for the variety of page and point size in current use. A convenient unit is the subscription cost of 1000 printed characters in US cents, taken as the basis for Table 6. This unit often varies from year to year for each journal since neither the number of pages published nor the subscription cost generally remain constant. It is noteworthy that the subscription cost to the IUCr journals has been virtually constant, in Danish kroner, for the last five years and has remained at the same level again in 1988. It is also noteworthy that *Acta* contains much less blank space than most journals, since nearly all papers are run on whereas many other journals start each paper on a new page, in addition to making less intensive use of each page. It is estimated that the unit cost for *Acta* given in

Table 6 should be reduced a further 5% in comparison with such journals.

The three lowest-unit-cost journals in Table 6 request page charges, thereby achieving unusually low subscription costs. The two highest-cost journals in the table are published by commercial companies; subscriptions to commercially published journals are generally higher than those to journals published by scholarly societies, on a unit-cost basis. Excluding the journals with page charges, the remaining twelve have an average cost of 5·1 US cents per 1000 characters. A recent study of physics journals by Barschall (1986) presented a 1985 cost range from 3·4 to 31 US cents per 1000 characters, excluding journals with page charges. If, in addition, the commercial journal costing 31 US cents per 1000 characters is also excluded as aberrant, then the average cost for the remaining ten physics journals in 1985 was 6·7 US cents per 1000 characters. A similar study was made by the American Mathematical Society (1983) of the 1983 cost of primary typeset American mathematical research journals. These also covered a wide range, from 0·9 to 35 US cents per 1000 characters, with the highest unit costs again associated with commercial publishing companies and the lowest costs with academic societies that request page charges.

The unit cost for subscribing to *Acta* is hence at the lower end of the general price range, with that for *J. Appl. Cryst.* close to average.

Mean publication time

Among the major services provided to authors and readers by publishers of crystallographic papers is the reduction of mean publication time [defined by the International Union of Crystallography (1986) as the averaged elapsed time between the final acceptance date of a given manuscript and its nominal publication date] to the least possible, consistent with high technical-editing, typesetting, printing and production quality. For the only IUCr monthly publication, *Acta* C, the mean publication time was 3·9, 3·8 and 4·8 months respectively in 1984, 1985 and 1986. In 1987, it was 4·7 months. The increased time in 1986 and 1987 was caused by unanticipated and major changes in personnel in the Technical Editor's office (see *Crystallographic papers published*, above), but the mean time is expected to become less than 4 months again by early 1988. In addition, a new monthly Fast Communications section may soon be inaugurated with mean publication time reduced to about 2 months. Few comparable international typeset monthly journals publish more rapidly.

The remaining IUCr journals, issued bimonthly, necessarily take about a month longer in publication than *Acta* C. As facsimile transmission of proofs between the Technical Editor's office and the author increasingly replaces the use of postal services, it is expected that all mean publication times will decrease by nearly a month. Such transmissions are now used regularly between the Technical Editor and the typesetters for the IUCr.

IUCr journals and the crystallographic databases

The traditional relationship between primary journals such as those of the IUCr and the numerical databases, as described in *Crystallographic Databases* (1987), is one in which information flows entirely from the former to the latter. During the last decade, as the IUCr journal Co-editors have assumed the burden of checking much of the crystallographic data presented in manuscripts for error and internal consistency, their work load has increased sharply. At the same time, the databases have separately developed sophisticated checking routines as part of their evaluation process before entering the published numerical data into their database. It soon became apparent that much duplication of effort could be avoided by a closer relationship between journals and databases, thus relieving the Co-editors and allowing evaluated data to enter the databases more rapidly. Following successful preliminary trial tests, larger-scale cooperation with the databases is being undertaken in which the primary crystallographic data received by the IUCr journals are passed directly to the databases in machine-readable form, where they are carefully checked numerically; the possibility that a previous investigator might already have reported on the same or a closely related material without the knowledge of the present author will similarly be examined by interrogating on-line bibliographic databases. It is expected that this centralized checking process, once operating on a routine basis, will be undertaken as part of the regular manuscript review process.

Compuscripts

Crystallographers have historically been innovative in their application of computers to diffraction problems [*cf.* Pepinsky (1952) and Pepinsky, Robertson & Speakman (1961)]. The recent widespread use of word processors has led many authors to consider the possibility of transferring the keystrokes made in producing their final typescript to their editor, in machine-readable form, for subsequent use in producing their typeset paper. Valuable experience with such compuscripts has already accumulated in some scientific societies. The editors of *Phys. Rev.*, for example, have provided detailed instructions for authors wishing to submit compuscripts able to drive their standard typesetting system, thereby resulting in normal journal page output; see Judd (1983). The number of usable compuscripts received subsequently has remained rather small, due probably to the onerous requirements placed on the keyboarder. Furthermore, each such compuscript requires an effort from the technical editors that is comparable with that needed for handling typescripts; the total cost saving is estimated to be less than about 15%. Other journals publishing a high percentage of mathematical and tabular material report similar experience.

An alternative approach is for authors to transmit, either in hard-medium machine-readable form or over a communications network, the direct keystroke output from their word processor without inclusion of typesetting commands. Translation programs already exist for such intercomputer communication. This approach is being used on an experimental basis by the authors and editors of the IUCr *International Tables for Crystallography*, Volumes B and C. It is still too early to know if this approach will be of general applicability. Among its advantages are the elimination (in principle) both of subsequent proofreading by the author after he has transmitted a fully corrected text and the postal delays often associated with normal manuscript transmission; however, under normal production conditions, publishers generally request a 'hard-copy' back-up together with the compuscript before they will process the latter. A major disadvantage is that many if not all specific typesetting codes required must be manually inserted in the compuscript in addition to the normal technical-editing corrections. This process may take as long as re-keyboarding the original typescript by an experienced operator. It also opens the possibility of introducing inadvertent changes in the textual or numerical content of the compuscript.

A separate development that is likely to be combined with computer-operated typesetting is the rapid approach of automatic full-page-make-up technology, initially leaving appropriate spaces for the figures but later handling the entire page content, resulting in a complete compuscript. Present scientific journal publishing methods continue to cut and paste both figures and galleys on the camera-ready page. Scanning methods that convert line and half-tone figures to digital form will eliminate the need for such stripping-in and will allow entire papers to be transmitted directly to full-text databases as complete documents. It is noted that a number of chemistry journals are presently available, without graphics, in full-text databases.

Wide availability of comparable methods would allow authors to transmit their compuscript directly to their editor of choice who, following normal review, could send it on to the appropriate referees electronically. Once accepted, the final version of the full compuscript (including figures) could be transmitted directly to the Technical Editor and, thereafter, to the typesetting device. Such papers could be displayed on readers' screens as soon as they were finally accepted in addition to being entered simultaneously as full documents in full-text databases. Before such a system is instituted, as is now nearly possible in principle, it will be necessary to solve the controversial problem of cost recovery by the publisher.

Automated optical-character recognition

An alternative approach to producing a compuscript that also eliminates re-keyboarding a previously fully corrected typed or keyboarded paper is the use of a device that accurately recognizes each character on the page. A highly successful system that indeed transforms a typescript into machine-readable form both accurately and rapidly is the Kurzweil Data Entry Machine (KDEM), now produced and marketed by Xerox Corporation. The KDEM reads any typewriter or typeset typeface, and does not require the author to use a special typeface. A number of publishers presently use such machines, including the American Institute of Physics. Provided the typescript is free from erasures, mechanical corrections and other blemishes such as originate in ill-adjusted photocopiers, these machines efficiently generate computer-readable material even from rather short normally prepared typescripts more accurately than normally competent keyboarders. It has been judged that the KDEM is in fact almost always as fast as and is usually faster than a keyboarder of average experience.

The major advantage of such machines is that they do not disturb the traditional responsibility of the author for submitting a well prepared typescript of his paper for publication. Informal polls taken at crystallographic meetings in recent years have revealed that a wide majority of authors prefer this traditional approach to that of producing a compuscript (but see also below under *Electronic publication and non-print distribution*).

Machine-readable binary-coded print information

High-density information has long been encoded in binary form on magnetic media. Very recently, the encoding of similar information in traditional print media has been introduced by Cauzin Systems, Inc. of Waterbury, CT, USA, under the name Softstrip®. In the new system, a strip of data encoded in a series of black and white spaces of overall dimension 1×9 inches contains 5·5 kbytes; a full journal page hence contains the equivalent of 20–40 kbytes, depending on the print quality. The encoding system provides for high-quality error detection. An inexpensive reader scans and reads the information presented directly into computer memory, taking about 25 s to read each strip. The new system is already used to provide an alternative Table of Contents in some commercial journals; the Softstrip® version allows the reader to produce an easily stored and computer-searchable database of authors and titles. The system is also used in some biological journals to present long sequences of nucleic acids and proteins. An application of current interest in crystallographic journals is the direct distribution of immediately usable computer programs on the journal page. It is, however, notable that digital data of all kinds may be distributed by this system within the limitations of available space.

Electronic publication and non-print distribution

The combination of compuscripts (or automated optical character recognition) with digitized figure capability opens the possibilities of full-scale electronic publication, as noted above. Authors presently wishing to produce a compuscript face a major problem in that each publisher accepting them has his own rigid set of format requirements; each set must be learned by the keyboarder, who is often the author. This lack of uniformity has been recognized and met by the introduction of an international Standard Generalized Mark-up Language (SGML) in 1986. The Association of American Publishers (AAP) has now issued an Electronic Manuscript Standard based on the SGML that offers all authors and publishers a consistent procedure for preparing and encoding compuscripts that may easily be exchanged and processed. As software developers produce packages that automatically insert the AAP Standard codes in a variety of widely available word-processing programs, it is anticipated that many authors will find production of compuscripts no more difficult than typescripts.

Electronic publication in the future is likely to be accompanied by journal distribution in media other than the traditional printed paper page. Although the print medium is unlikely to be displaced in the foreseeable future, it is expected that either digitally recorded optical disc or cassette will become increasingly popular in libraries, fully displacing the less-convenient and less-legible microfilm and microfiche, particularly for older volumes, as a means of conserving space.

CD-ROMs (compact disc read-only memory) are already available for textual and tabular data, but many publishers still regard them as too expensive. However, the cost of forming the master followed by disc reproduction is rapidly dropping and this medium may well become important for distributing crystallographic publications. Access to many CD-ROMs without manual intervention is already feasible in 'juke-box'-type (multiple disc) players.

An alternative medium with about five times more storage than a standard CD-ROM but with slower (*i.e.* serial) access time is the digital video cassette. At least one manufacturer currently offers a tape drive that stores 2·2 Gbytes on a regular T-120 video cassette. This medium may become a strong competitor to the compact disc.

Prospects for crystallographic publishing

Crystallography continues to be a field in which common underlying theory and experimental technique are often applied to the solution of interesting problems in other branches of science (see Table 4). Many excellent crystallographic papers are published in non-crystallographic journals for a variety of reasons, including the high prestige of some major journals, institutional and departmental preferences, ease of manuscript acceptance and, perhaps regarded as most important, the exposure to specialists in the branch of science specifically served by that journal. A major reason for dividing *Acta* B in 1983 into the new *Acta* B and *Acta* C was to provide a new and attractive journal specifically for publishing a high percentage of this kind of paper presenting original work in structural science.

Papers that make fundamental contributions to the common crystallographic core not only remain most welcome in *Acta* A, but are likely to enter general use more rapidly than if published in a non-crystallographic journal. The number of structure determinations undertaken as an end in themselves and presented in *Acta* C has started to increase, at least in part because of the popularity of the new Short Format category. Compared with only eight Short Format papers in 1985 when the category opened, this number rose to 75 in 1986 and 174 in 1987. This rising trend is welcomed and may be hastened by the forthcoming introduction of a machine-readable version of Short Format compuscript. Steps have also been taken to attract more papers in the field of applied crystallography to *J. Appl. Cryst.*, including publication of the refereed contributions to a recent International Conference on the Applications and Techniques of Small-Angle Scattering.

In addition to the expected normal increase in number of papers published in the IUCr journals, mergers with other crystallographic journals similar to that made in 1983 with *Crystal Structure Communications* may be mutually advantageous (see Table 5). Further efforts to attract crystallographic papers in the fields of biological macromolecules, property-structure relationships, materials science, and real and ideal crystal studies in particular, as well as in the other under-represented fields noted in *Breadth of published crystallographic research* above, would improve the present balance between written and oral or poster communication within the organs of the IUCr. The early introduction of publishing innovations such as those discussed above or others that are presently less well developed, where appropriate, in combination with an increased coverage of the total field can only enhance the level of service now provided to the crystallographic community. It is likely that fresh financial and legal problems will emerge as new distributional media are utilized, but continued vigilance by the oversight mechanism set in place since the Twelfth General Assembly may be expected to maintain the present robust state of IUCr crystallographic publishing and financial solvency.

It is a pleasure to thank C. E. Bugg, M. H. Dacombe, Th. Hahn, J. H. Robertson, J. R. Rumble (National Bureau of Standards), J. T. Scott (American Institute of Physics), T. Selover (American Institute of Chemical Engineers), and D. W. Weisgerber (Chemical Abstracts Service) for discerning comments on a draft version of this paper.

References

ABRAHAMS, S. C. (1982). *Acta Cryst.* **A38**, 1-2.
ABRAHAMS, S. C. (1987). *Acta Cryst.* **A43**, C300.
AMERICAN MATHEMATICAL SOCIETY (1983). *Not. Am. Math. Soc.* **30**, 715-724.
BARSCHALL, H. H. (1986). *Physics Today*, December, pp. 34-36.
Communicated Abstracts of the Fourteenth International Congress of Crystallography (1987). *Acta Cryst.* **A43**, C1-C351.
Crystallographic Databases (1987). Bonn, Cambridge, Chester: IUCr Data Commission.
EWALD, P. P. (1948). *Acta Cryst.* **1**, 1-2.
HAWKINS, D. T. (1980). *Acta Cryst.* **A36**, 475-482.
INTERNATIONAL UNION OF CRYSTALLOGRAPHY (1948). Statutes and By-Laws. *Acta Cryst.* **1**, 275-276.
INTERNATIONAL UNION OF CRYSTALLOGRAPHY (1983). Appendices to the Agenda of the Twelfth General Assembly. *Acta Cryst.* **A39**, 444-465.
INTERNATIONAL UNION OF CRYSTALLOGRAPHY (1986). Report of the Executive Committee for 1984. *Acta Cryst.* **A42**, 58-79.
JUDD, P. (1983). *Bull. Am. Phys. Soc.* **28**, 957 (B1-B57).
Notes for Authors (1983). *Acta Cryst.* **A39**, 174-186.
PEPINSKY, R. (1952). Editor. *Computing Methods and the Phase Problem in X-ray Crystal Analysis.* Pennsylvania State College: X-ray Crystal Analysis Laboratory.
PEPINSKY, R., ROBERTSON, J. M. & SPEAKMAN, J. C. (1961). Editors. *Computing Methods and the Phase Problem in X-ray Crystal Analysis.* Pergamon: Oxford.

Acta Cryst. (1988). **A44**, 410-413

On the Application of Phase Relationships to Complex Structures. XXVII. Phase Extension for Small Proteins

BY M. M. WOOLFSON

Department of Physics, University of York, York YO1 5DD, England

AND YAO JIA-XING

Institute of Physics, Beijing, China

(*Received* 8 *August* 1987; *accepted* 4 *February* 1988)

Abstract

An example is given of phase extension for a small protein, avian pancreatic polypeptide, by the use of the Sayre-equation tangent formula (SETF). Initial data were the phase estimates of 129 reflexions with large values of $|E|$ within the 3 Å resolution sphere. The mean error of these phases, estimated by a combination of isomorphous replacement and anomalous scattering, was 26·5°. Random values were then given to 1371 other phases out to 1 Å resolution and refinement was carried out with SETF. In 20 trials, 11 gave mean phase errors less than 34° for all 1500 reflexions with the best set having a mean phase error of 31·9°. Maps computed with these phases showed the general form of the molecule.

Introduction

Much thought is being given to the application of direct methods to protein-structure problems. The most successful ideas so far have involved some combination of information from physical methods, such as the heavy-atom method, isomorphous replacement or anomalous scattering, with direct-methods theory (*e.g.* Hauptman, 1982; Karle, 1984, 1986; Fan Hai-fu, 1983). However, there is a tendency to illustrate ideas with ideal calculated data for known structures and, until recently, the only unknown protein structures to have been solved using direct methods have involved starting from heavy-atom positions found by using *MULTAN* (Wilson, 1978). However, there is an interesting example of the application of

0108-7673/88/040410-04$03.00