Acta Cryst. (1994). D**50**, 459–465

Two Distinct Approaches to Crystallization Results-Recording Databases

BY ANNE M. HASSELL

Bioanalytical and Structural Chemistry Department, Glaxo Research Institute, Research Triangle Park, NC 27709, USA

THOMAS J. HARROCKS AND ERIC H. DASHMAN

Research Computing Department, Glaxo Research Institute, Research Triangle Park, NC 27709, USA

AND ANIL MISTRY

Protein Biochemistry Department, Glaxo Group Research, Greenford, Middlesex, England

(*Received* 21 *October* 1993; *accepted* 20 *December* 1993)

Abstract

Two different systems for recording crystallization results are described. Both programs are written in Microsoft *Excel*. Results can be entered either as text or numerical scores. In each system, one can search or query on the results, macromolecule name, experimental parameters, *etc.* and correlate these to specific crystallization conditions or trends in the crystallizations. Such information is valuable when optimizing crystal growth conditions and determining new crystallization parameters to test. *XTAL RES* is a Macintosh program that is linked to *PCRYSTAL* (software used to enter details of the crystallization experiments) through an Oracle database on a VAX network. A second system, the *Electronic Notebook*, was designed as a 'stand-alone' *Excel* application running under Microsoft Windows.

Introduction

Crystallization of biological macromolecules is a time-consuming process and can often be the rate-limiting step in determining the three-dimensional structure of macromolecules (Giegé & Ducruix, 1992; McPherson, 1990). Because of the extensive nature of the screening required to determine and optimize conditions for crystal growth, vast quantities of data are often generated. Indeed, it was the accumulation of such large quantities of crystallization data that motivated our two laboratories to design results-recording databases. It quickly became apparent that we needed a system that would not only record the results accurately, but would also allow us to analyze these results for use in the design of future experiments. In this paper, we present two totally different approaches, *XTAL RES* and the *Electronic Notebook*, for electronically recording results of crystallization experiments. Although this software was designed to meet the specific needs of the two individual laboratories involved, it clearly reflects influences from systems used by other crystallography groups (Weber, 1990; Carter, 1990; Hannick, Perozzo, Schultz & Ward, 1992).

The *XTAL RES* system utilizes a client/server technology with Oracle* as the central database. The *Electronic Notebook* is a purely 'stand-alone' system and stores data locally. User interfaces for both systems are written in *Excel*.† Results may be entered as textual descriptions, color and pattern-coded text, or, in the case of the *Electronic Notebook*, as numerical scores, using a 4×6 Linbro tray format. Each system has search capabilities that enable the researcher to use previous crystallization results to plan future experiments. The user can search on any single parameter or combinations of parameters, such as, macromolecule name + temperature + precipitating agent + a result of 'crystals'. While these two databases share many common features, *e.g.* color and textual descriptors for each

* Oracle, Oracle Corporation, Redwood Shores, CA, USA.

† Microsoft *Excel*, Microsoft Corporation, Redwood, WA, USA.

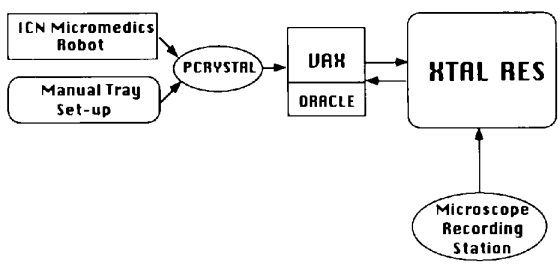


Fig. 1. *XTAL RES* organization. Experimental parameters are entered into the system using *PCRYSTAL* software. This information is stored in a server database (Oracle) on a VAX and is accessed by the results-recording database, *XTAL RES*. Trays (set up manually or with the ICN robot) are examined under a microscope. The results are entered into *XTAL RES* and stored in an Oracle database on a VAX.

ISSN 0907-4449

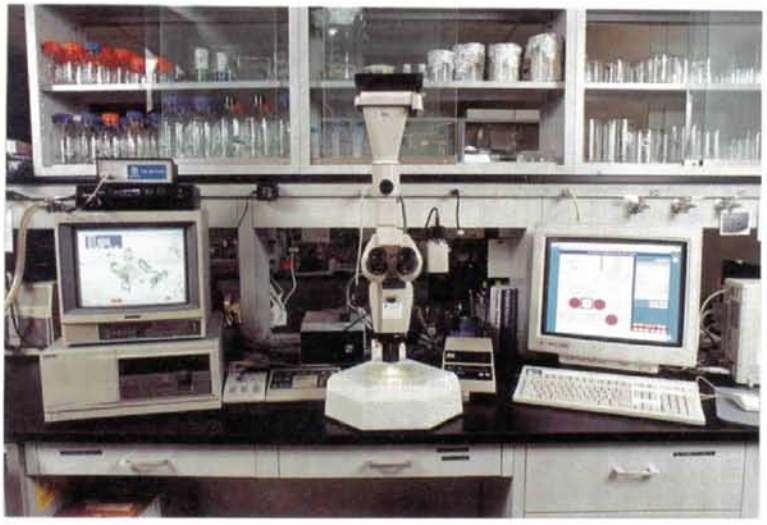
result, they are designed quite differently. Such differences are prominently reflected in the choices of a network *versus* a 'stand-alone' system, as well as in the hardware and software utilized. We hope that the following descriptions of these two results-recording systems will provide ideas for others in the development of even more comprehensive systems.

Materials and methods

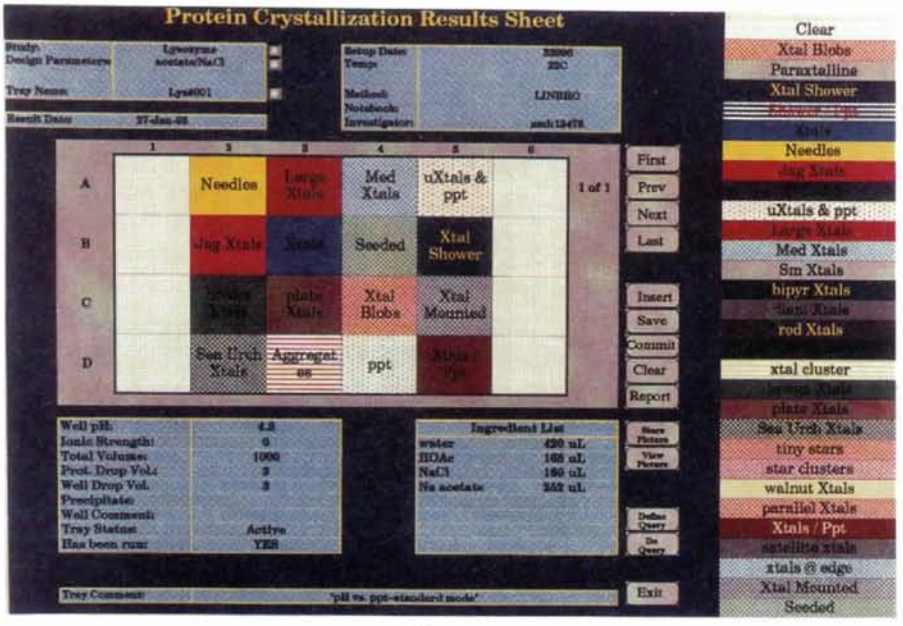
XTAL RES/network station

XTAL RES runs under *Excel* (Version 4.0) on a Macintosh, and the code was written primarily in *Excel* Macro language. It shares a common Oracle database with *PCRYSTAL* (Fig. 1), which is the software used to enter crystallization conditions. *PCRYSTAL* defines manual experiments and drives the ICN Micromedics Protein Crystallization Robot. *Second Wind* (Version 1.2)† and *SQL*NET* (Version 1.5)‡ provide the connectivity between the *XTAL RES* network database (Oracle) and the *Excel* spreadsheet. This network connection allows the researcher to view precipitation conditions for each

† *Second Wind*, Anju Technologies, Los Altos, CA, USA.
‡ Oracle *SQL*NET*, Oracle Corporation, Redwood Shores, CA, USA.



(*a*)



(*b*)

Fig. 2. (*a*) *XTAL RES* workstation. The *XTAL RES* workstation consists of a Nikon SMZ-U microscope connected to an imaging system (Sony Trinitron Color Monitor, Sony Color Video Printer UP-5000, Optronics CS-450 camera and Oracle RGB-9000 Trackball) which is linked to a Macintosh Quadra 950 equipped with a SuperMac monitor, an Alphatronics Inspire II optical disk drive, an Apple LaserwriterII printer and a Seiko SII/CH4104 color printer. (*b*) *XTAL RES* data-entry screen. *XTAL RES* allows you to choose from a list of approximately 50 results. Each result is color and pattern coded and has a specific text description. Results are entered by selecting the wells of interest and then clicking on the desired results button. Additional results can be added to this master list whenever necessary. The user can view details of the precipitation conditions for each well while entering results. The First, Prev, Next and Last buttons allow the user to look at all results entries for a particular tray. The Insert button allows one to enter new results, while the Commit command saves the data to an Oracle database.

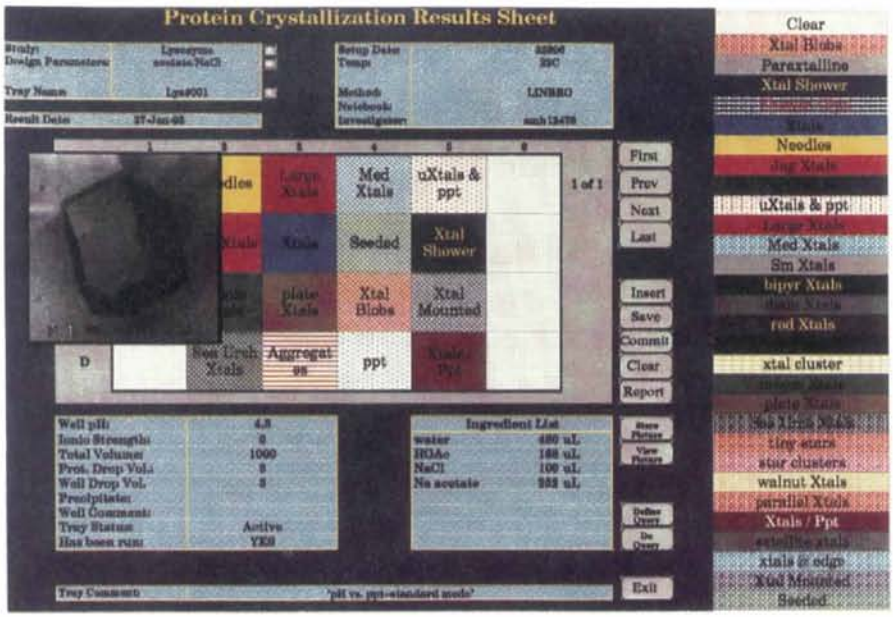


Fig. 3. Storage of image results in *XTAL RES*. Pictures of the crystals may be stored using *VideoSpigot* NuBus and C code for image display. This can be accomplished by first selecting the well and then using the Store Picture and View Picture buttons. More than one image may be stored for each well. The user may view the image for only one well at a time or for the entire tray. These pictures may be enlarged for better viewing, as shown above.

Study:	**Lysozyme**	**SetupDate:**	**27-Jan-93**
Design Parameters:	**acetate/NaCl**	**Temp:**	**22C**
Tray Name:	**Lys#001**	**Method:**	**LINBRO**
Investigator:	**amh12478**	**Notebook:**	
Tray comment:	**pH vs ppt.--standard mode**		

	Date	1	2	3	4	5	6
A	27-Jan-93		Needles	Large Xtals	Med Xtals	uXtals & ppt	
	08-Mar-93		Sm Xtals	rod Xtals	Med Xtals	uXTALS	
	09-Mar-93		Med Xtals	rod Xtals	plate Xtals	plate Xtals	
	14-Sep-93		Med Xtals	rod Xtals	plate Xtals	plate Xtals	
B	27-Jan-93		Jag Xtals	Xtals	Seeded	Xtal Shower	
	08-Mar-93		Jag Xtals	Xtals	Seeded	uXTALS	
	09-Mar-93		Shower / Ppt	Xtals	Slight ppt	plate Xtals	
	14-Sep-93		Shower / Ppt	walnut Xtals	walnut Xtals	walnut Xtals	
C	27-Jan-93		broom Xtals	plate Xtals	Xtal Blobs	Xtal	
	08-Mar-93		broom Xtals	walnut Xtals	walnut Xtals	walnut Xtals	
	09-Mar-93		broom Xtals	walnut Xtals	walnut Xtals	walnut Xtals	
	14-Sep-93		broom Xtals	walnut Xtals	walnut Xtals	walnut Xtals	
D	27-Jan-93		Sea Urch	Aggregates	ppt	Xtals / Ppt	
	08-Mar-93		Sea Urch	Aggregates	ppt	rod Xtals	
	09-Mar-93		Sea Urch	Slight ppt	Slight ppt	rod Xtals	
	14-Sep-93		Xtals	Xtals	Xtals	Xtals	

Fig. 4. *XTAL RES* report spreadsheet. Results from several observation dates may be summarized in a spreadsheet by using the Report button. These sheets are suitable for pasting into a notebook/file. Approximately 15 observations will fit onto one page, but this can be customized to fit any notebook being used.

```
SELECT STUDYNAME, EXPNAME, PLATENAME, WELLID, RES_DATE, RESULT_NAME FROM PCRYSTAL.PC_ST_EX_PL_WE_RE WHERE 1 = 1
AND STUDYNAME = 'Lysozyme'
AND EXPNAME = 'acetate/NaCl'
AND ( RESULT_NAME = 'Xtals' OR RESULT_NAME = 'Large Xtals' OR RESULT_NAME = 'Sea Urch Xtals' OR RESULT_NAME = 'Needles' )
ORDER BY STUDYNAME, EXPNAME, PLATENAME, RES_DATE, WELLID
```

STUDYNAME	EXPNAME	PLATENAME	WELLID	RES_DATE	RESULT_NAME
Lysozyme	acetate/NaCl	Lys#001	2	27-Jan-93	Needles
Lysozyme	acetate/NaCl	Lys#001	20	27-Jan-93	Sea Urch Xtals
Lysozyme	acetate/NaCl	Lys#001	3	27-Jan-93	Large Xtals
Lysozyme	acetate/NaCl	Lys#001	9	27-Jan-93	Xtals

Fig. 5. *XTAL RES* query results. Results of a search performed on a set of RESULTS ('Xtals, Large Xtals, Needles, or sea urchin crystals') for lysozyme using acetate/NaCl crystallization conditions are presented above. The query results are listed by tray number and well number.

well as the results are being entered. Databases other than Oracle (*RDB*, *MS Acess*, *Sybase*) could serve as a repository for the crystallization data. Oracle was our choice since it was already in use on our VAX network. The code for *XTAL RES* is written to allow the program to interface with local programs (*VideoSpigot*, graphic display code, external C routines) as well as the remote network database. This allows *XTAL RES* to display image results captured and stored locally using *VideoSpigot* (Version: *Screenplay* by Peter Barrett).* Custom C code to display the images was compiled into XCMD's (X commands). The XCMD's are called directly from the *Excel* Macro code.

Electronic notebook/'stand-alone' station

The *Electronic Notebook* is a PC-based data entry and analysis program. The station consists of a 486 PC running at 33 MHz with a minimum of 8 Mb RAM and a 200 Mb hard disk drive. It contains a MicroEye 2C† card and is equipped with a 17 inch Super VGA color monitor. The system requires Microsoft *Excel* (Version 4.0) and *Q + E* (Version 3.01),‡ both running under Windows 3.1.§ The PC is linked to a Sony CCD RGB high-resolution (720 × 512 pixels with capability to grab 24-bit color images) video camera which is attached to a Wild M5A stereomicroscope. Images are captured using MicroEye *Capture* (Version 1.0). A single 256-color image (in bitmap format) requires approximately 1 Mb of memory. *Visual Basic* (Version 3.0)* allows the images to be automatically archived and sent to Poem *ColorBox* III† for compression. Compression to 1/100th of the original file size is achieved quite effectively with no visible loss in resolution. These compressed files can be stored on either a floppy or optical disc.

The DLL's (dynamic link libraries) are written in Visual C++‡ for viewing the compressed images within the notebook. In addition, the user has direct access to *Word* (Version 2.0)§ from the *Electronic Notebook*. Experiments set up either manually or with a robot can be designed using the *PIPEX* utility (Eiselé, 1993). *PIPEX* is a utility routine which provides the preparation of crystallization buffers with a computer-controlled motorized pipette. The relevant information may be either typed directly into the *Electronic Notebook* or automatically imported from

* *VideoSpigot* NuBus, SuperMac Technology, Sunnyvale CA, USA.

† Digithurst Ltd, Royston, Hertfordshire, England.

‡ *Q + E*, Pioneer Software Systems, Microsoft Corporation, Redwood, WA, USA.

§ Microsoft Windows, Microsoft Corporation, Redwood, WA, USA.

* Microsoft *Visual Basic*, Microsoft Corporation, Redwood, WA, USA.

† Iterated Systems Inc., Norcross, GA, USA.

‡ Microsoft Visual C++, Microsoft Corporation, Redwood, WA, USA.

§ Microsoft *Word*, Microsoft Corporation, Redwood, WA, USA.

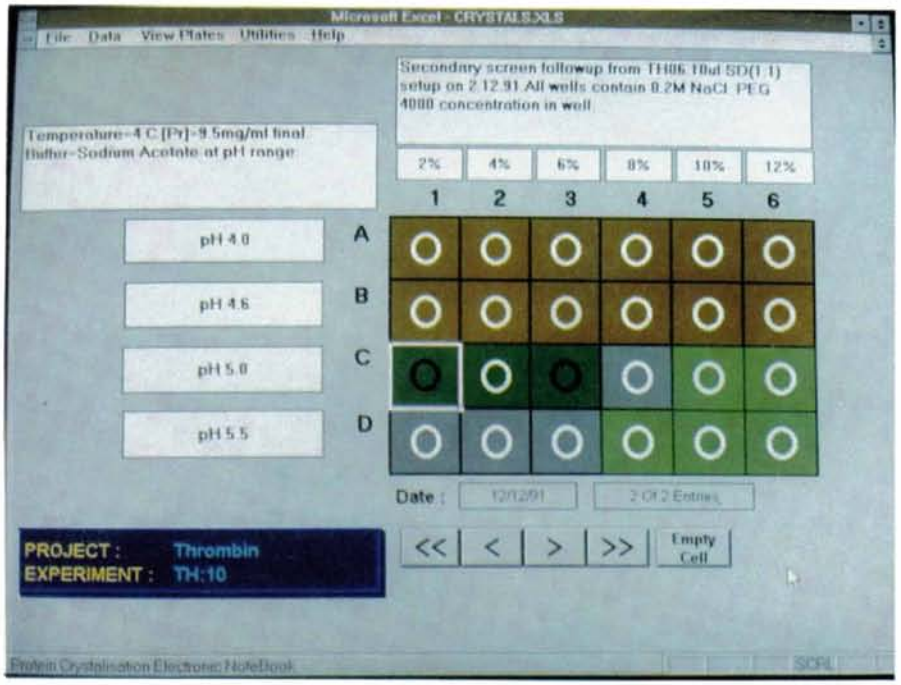


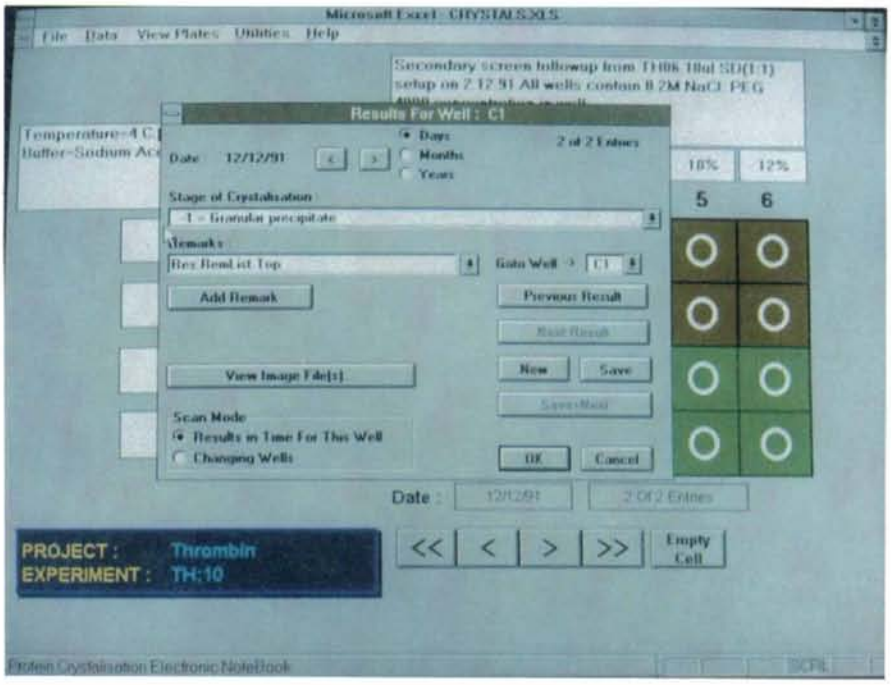
Fig. 6. *Electronic Notebook* start-up screen. The menu bar at the top provides the following functionalities. File allows selection/creation of projects and experiments. Data permits data input into the Header, Results and Conditions windows, View Plates allows viewing of all plates within a project for trends analysis. Utilities provides additional features, including the generation of three-dimensional graphs, the Search facility, experiment tracking function, access to Microsoft *Word*, and access to the color-coded *Q*-number system. The Help function is still in development.

the *PIPEX* utility to the program. The *Electronic Notebook* gathers and archives different types of information concerned with the crystallization experiments. It provides access to this data from within one application.

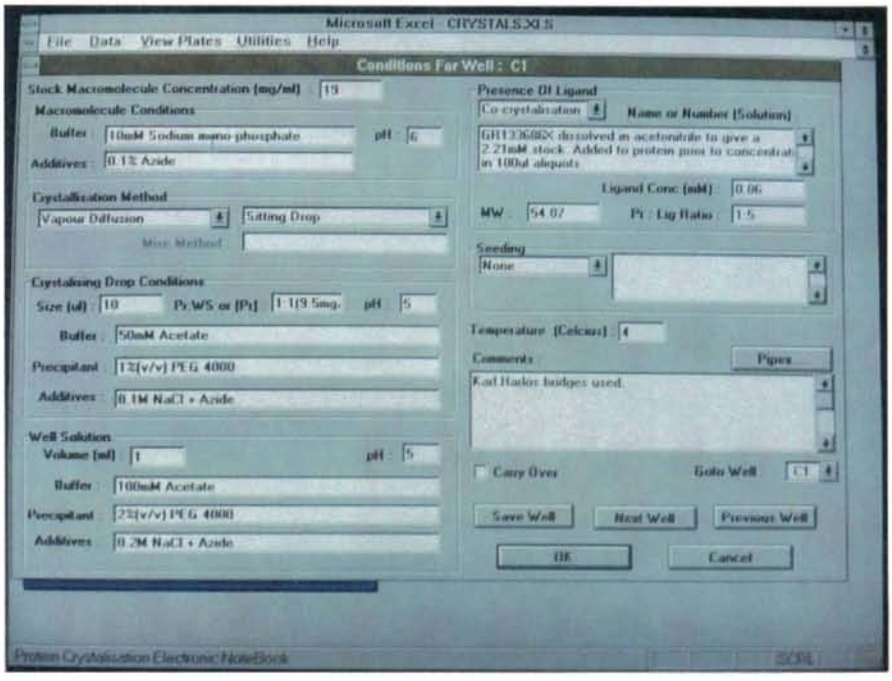
Results and discussion

XTAL RES/network station

The results-recording workstation is configured with a stereomicroscope linked to an imaging system



(*a*)



(*b*)

Fig. 7. (*a*). *Electronic Notebook* results sheet. The Results page allows data entry as a *Q* number which is linked to a specific color code and a descriptive results scheme. Additional textual/descriptive data may also be entered on this screen. Results images of the experiment concerned may be viewed from the Results page. (*b*) *Electronic Notebook* well conditions screen. The Conditions page allows detailed entry of the conditions in the well solution and in the protein droplet of each experiment. Solution volumes can be incorporated automatically using the Pipex button if the volumes were generated using the *PIPEX* utility.

connected to a Macintosh computer (Fig. 2*a*). *XTAL RES* utilizes color and text for each result to be recorded. The results list contains approximately 50 different types of results, including various types of precipitates, crystals, microcrystals and phase separation. Additional information, such as whether or not the crystal has been mounted for X-ray studies or used in seeding experiments, can also be stored. New results descriptors may be added at any time. The results data-entry screen with its color/pattern-coded results list and text descriptors is depicted in Fig. 2(*b*). The user may store the results simply as text or may include images of the crystals with each result (Fig. 3). This feature enables the researcher to moni-

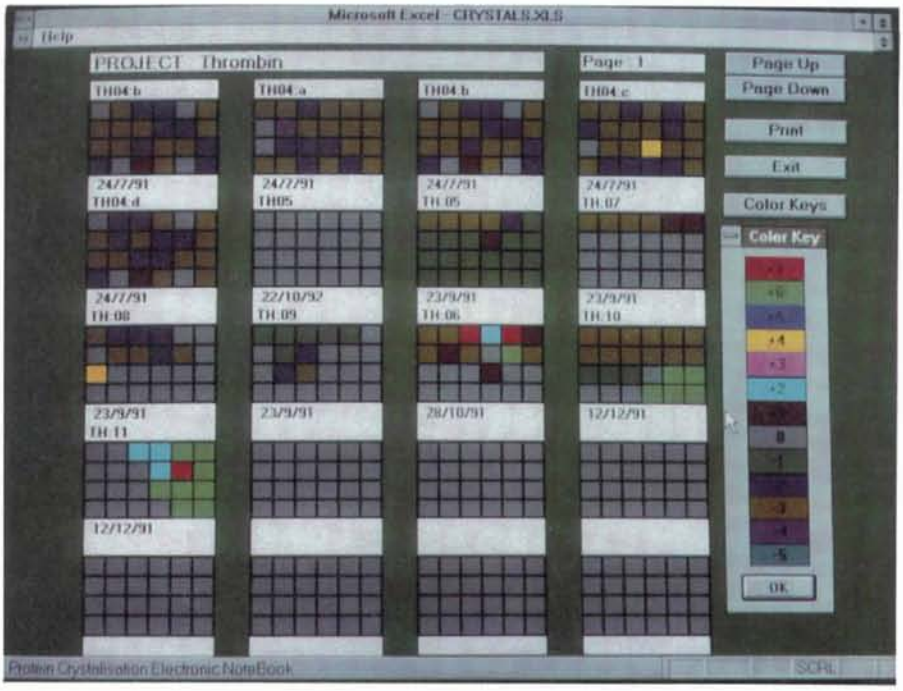


Fig. 8. *Electronic Notebook* data analysis. Trends in crystallization can be observed either in terms of the entire project or in terms of one plate through time. In both cases, all results sets are displayed as miniaturized 24-well plates with results shown as the *Q*-number color code for the last observed set of results for each plate.

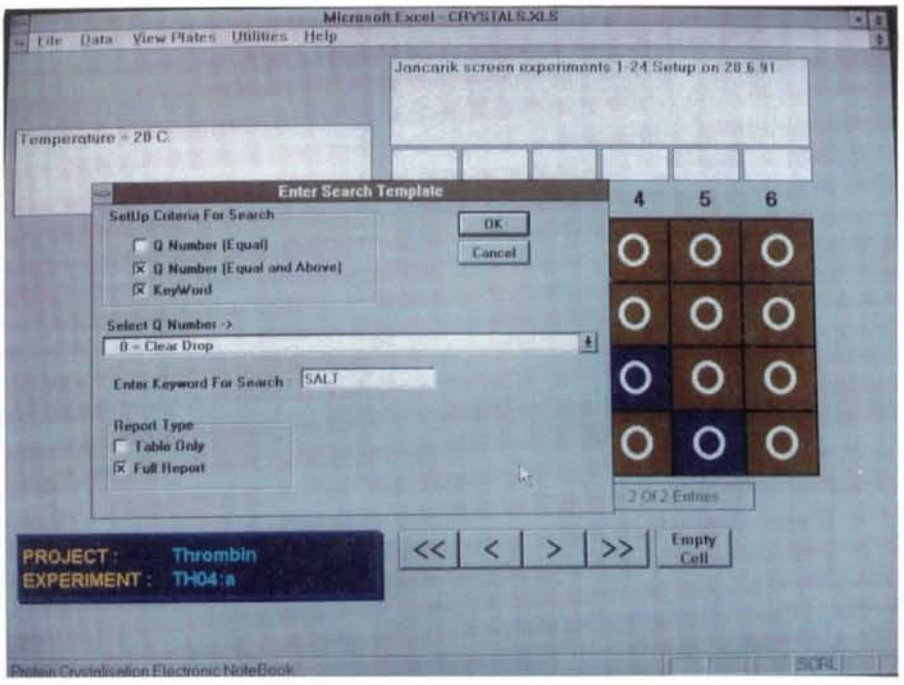


Fig. 9. Key results can be identified by the use of the Search facility. Search modes include keywords (SALT, Tris, ammonium sulfate, *etc.*), specific *Q* numbers, or all results within a *Q*-number unit (*e.g.* all results 5 or higher). A complete report or summary table can be generated from the search and exported directly to Microsoft *Word* for printing or editing.

tor changing crystal habits. Results for each tray may be viewed as a spreadsheet (Fig. 4) suitable for a notebook. Several observations of results may be printed on one sheet. Alternatively, one may print the actual data-entry screen; this may be desirable when the researcher wishes to view pictures of the crystals. Once the data are entered, one can search on any single parameter or combinations of parameters used in setting up the tray and the results obtained, such as, macromolecule name, date, temperature, crystallization conditions, protein lot number, *etc.* to determine patterns of crystallization for a particular macromolecule (Fig. 5). Integration of the results (*XTAL RES*) and set-up databases (*PCRYSTAL*) allows correlation of results to specific crystallization conditions.

Electronic notebook/'stand-alone' station

The *Electronic Notebook* startup screen is displayed in a typical 4 × 6 Linbro plate format with the results depicted as colors in Fig. 6. Each result is assigned a quality number or Q number (Carter & Carter, 1979) on a scale from −5 to +7. The descriptive terms for these Q parameters are listed in Table 1.

These Q numbers are an extension of Carter's original six-point scale (Carter & Carter, 1979). Each number represents a color which is displayed on the screen as a result. Multiple data sets (taken on different days) for the same experiment may be entered along with images of the results as viewed under a microscope. Descriptions of the experimental conditions are displayed in header windows around the plate, and a menu bar is displayed at the top of the screen (Fig. 6). In addition, textual information regarding experimental conditions can be entered into a header page that is seen when an experiment is viewed on the main screen. The results sheet describes the results over time for each well (Fig. 7*a*). Detailed information about the precipitation conditions for each well is entered into the *Electronic Notebook* on the 'Conditions' page (Fig. 7*b*). Data analyses in terms of 'trends' in crystallization are depicted in Fig. 8. The user can query on Q numbers or keywords (crystallization parameters) using the Search facility (Fig. 9). A summary report of the search is produced as a Microsoft *Word* document.

Although *XTAL RES* and the '*Electronic Notebook*' differ dramatically in their approach to managing large volumes of crystallization data, both programs provide easy methods for entering data and for performing searches on single parameters or combinations of parameters tested. Both systems are currently in use in our laboratories, but are clearly in a developmental stage. Despite this, they have provided improved efficiency in managing large quantities of data. Instead of spending days pouring over endless pages of data in search of conditions that produced crystals, we can now scan approximately 50 000 crystallization results in a matter of minutes. Although our experience with the data-analysis features is limited, these databases have already alleviated the need for paper-based results-recording systems. However, as more data are entered into *XTAL RES* and the *Electronic Notebook*, we will be able to utilize the results-analysis features fully to detect trends in crystallization. Such information could then be used in planning future experiments or in the design of new factorial screens. Another challenge is to devise software tools for extracting the 'hidden' information or variables that are important for successful crystallizations. These data represent another level of information, similar to that found in the NIST/CARB Database for the Crystallization of Biological Macromolecules (Gilliland & Bickhan, 1990), that can be used in developing new crystallization strategies.

Table 1. *Q-number descriptors*

Q number	Description
+7	Large, single crystals (> 100 μm thick)
+6	Prisms, cubes, rods, bipyramids (< 100 μm thick), '3-D' crystals
+5	Plates, small rods (< 50 μm thick), '2-D' crystals
+4	Needles, spherulites (< 10 μm thick), '1-D' crystals
+3	Rosettes, small aggregates, microcrystalline birefringent particles
+2	Particulate solution, birefringent
+1	Phase separation, microdroplets
0	Clear drop
−1	Granular precipitate, slowly developing
−2	Heavy precipitate
−3	Heavy precipitate and drop shrinkage
−4	Dry drop
−5	Protein denaturation

The authors wish to thank Dr Steven R. Jordan for critical reading of the manuscript.

References

CARTER, C. W. JR (1990). *Methods: a Companion to Methods in Enzymology*, Vol. 1, pp. 12–24. New York: Academic Press.

CARTER, C. W. JR & CARTER, C. W. (1979). *J. Biol. Chem.* **254**, 12219–12223.

EISELÉ, J.-L. (1993). *J. Appl. Cryst.* **26**, 92–96.

GIEGÉ, R. & DUCRUIX, A. (1992). *Crystallization of Nucleic Acids and Proteins: a Practical Approach*, edited by A. DUCRUIX & R. GIEGÉ, pp. 1–18. Oxford: IRL Press.

GILLILAND, G. L. & BICKHAM, D. M. (1990). *Methods: a Companion to Methods in Enzymology*, Vol. 1, pp. 6–11. New York: Academic Press.

HANNICK, L. I., PEROZZO, M. A., SCHULTZ, L. W. & WARD, K. B. (1992). *J. Cryst. Growth*, **22**, 303–305.

MCPHERSON, A. (1990). *Eur. J. Biochem.* **189**, 1–23.

WEBER, P. C. (1990). Personal communication.