(2) J. E. Ash, and E. Hyde, Eds., "Chemical Information Systems", Ellis Horwood, Chichester, 1975.
(3) G. Vladutz, private communication, 1978.
(4) E. J. Schuegraf, "A Survey of Data Compression Methods for Non-numeric Records", *Can. J. Inf. Sci.*, **2** (1), 93–105 (1976).
(5) D. A. Huffman, "A Method for the Construction of Minimum-Redundancy Codes", *Proc. IRE*, **40**, 1098–1101 (1952).
(6) M. F. Lynch, "Variety Generation—a Reinterpretation of Shannon's Mathematical Theory of Communication, and Its Implications for Information Science", *J. Am. Soc. Inf. Sci.*, **28** (1), 19–25 (1977).
(7) I. J. Barton, M. F. Lynch, J. H. Petrie, and M. J. Snell, "Variable-Length Character String Analyses of Three Data Bases, and Their Application for File Compression", in "Informatics I: Proceedings of the 1st Informatics Conference, Durham, 1973", Aslib, London, 1974, pp 154–162.
(8) E. J. Schuegraf and H. S. Heaps, "A Comparison of Algorithms for Data Base Compression by Use of Fragments as Language Elements", *Inf. Storage Retr.*, **10** (9), 309–319 (1974).
(9) C. E. Shannon, "A Mathematical Theory of Communication", *Bell Syst. Tech. J.*, **27** (3), 379–423 (1948).
(10) K. Subramanian, S. Krishnan, and E. V. Krishnamurthy, "Huffman Binary Coding of WLN Symbols for File Compression", *J. Chem. Doc.*, **14** (3), 146–149 (1974).
(11) V. R. Walker, "Compaction of Names by x-Grams", in J. B. North, Ed., Proceedings of the 32nd Annual Meeting of the American Society for Information Science, 1969, pp 129–135.
(12) G. N. Metcalfe and G. B. Evans, "File Compression: a Report on BLCMP's R&D Programme", Birmingham Libraries Cooperative Mechanisation Project, Birmingham, 1977.
(13) A. R. Yeates, "Text Compression in the Brown Corpus Using Variety-Generated Keysets, with a Review of the Literature on Computers in Shakespearean Studies", M. A. Dissertation, University of Sheffield, 1977.
(14) M. A. Emly, "Compression of Continuous Text by *n*-Gram Encoding, with a Review of the Literature on Computer-Produced Concordances to Medieval Works", M.A. Dissertation, University of Sheffield, 1978.
(15) NOTE ADDED IN PROOF: This has now been done, and 55.64% compression was obtained, showing a small improvement over the performance without the modification.

# A Software Controlled Data Acquisition System for Chemical Relaxation Experiments

A. M. RAO,† P. SHAH,‡ R. HAIDLE, and G. CZERLINSKI*
Department of Biochemistry, Northwestern University, Chicago, Illinois 60611

Received March 28, 1978

A system is described for the acquisition of data with a Biomation 802 transient recorder. The transient recorder is connected to a Hazeltine 2000 CRT terminal with dual tape cassette unit through a PDP 8/e minicomputer. The minicomputer stores the controlling software which allows the user to vary the sensitivity of the vertical and horizontal scale of the Biomation 802 through the keyboard of the terminal. Various simple operations are available including the storage of the data matrix on dual tape cassettes. Systems information is added to the initial data set to allow a high degree of automation in later evaluation stages. The content of the cassettes is subsequently transferred onto the disk of a large computer, utilizing a disk stored program and telephone communications. The data is then further evaluated by a sequence of programs, proceeding through several stages of evaluation. Data evaluation thus becomes highly automated and may be done much more rigorously than by graphical means or by statistical evaluations on the calculator. Since the Biomation 802 transient recorder accepts data in the form of an analogue signal and a trigger, this system could easily be used to store and process signals coming from many other sources.

## INTRODUCTION

Interfacing to minicomputers was discussed previously[1-5] as well as the use of minicomputers in the handling of experimental data. Our application is somewhat different from previous ones insofar as we allow parallel transfer of the data from the Biomation 802 onto the PDP 8/e minicomputer, some preliminary editing, and subsequent transfer onto the Hazeltine tape cassette. It developed that the internal operation of this tape cassette system is such that we cannot effectively transfer more than about 1200 bits per second. The Hazeltine CRT 2000 terminal communicates with a large computer via a standard modem and phone lines at 300 band. We are also able to load horizontal and vertical scan into the Biomation 802 with the necessary information being typed in through the keyboard of the CRT terminal.

Before the use of the minicomputer, we had produced data on paper tape in rather large quantities. The rolls of paper tape were then transferred to a large computer via an overnight carrier. Unfortunately, the large volume of paper tape caused some operational problems at the main computer center. Furthermore, 10 min was required for the production of paper tape with about 1000 "points" (numbers between 1 and 256) per experiment from the Biomation 802. The conversion error of this instrument is 1 in 256 full scale (8 bit converter), adequate for most experiments. Three minutes is the optimal time between experiments for the following reasons. In our temperature-jump experiments, 3 min is required for (a) thermal equilibration after a temperature jump and (b) full attainment of a new equilibrium value after a concentration- or pH-jump. The Biomation 802 transient recorder is thus utilized on two types of chemical relaxation experiments.

The exponential decay curves may be inspected on a Tektronix 549 storage oscilloscope before conversion or on a Tektronix 602 oscilloscope after conversion. The digitized data is then transferred onto tape cassettes which are subsequently rerun under the control of a program and stored on disk at a CDC 6400 facility. Subsequently, a sequence of programs evaluates the data. Some information on the software and the details of the hardware configuration were reported before.[6] No changes were made on the Biomation 802, allowing us to use a teletypewriter and the previously employed interface (Dijiscan Model B203) as back-up for this computer system. However, we have not needed to use the back-up system since the minicomputer was first brought into operation in June of 1973. The equipment is used rather extensively in two research projects on the mechanism of action of enzymes.

## METHODS

In our biochemical experiments two temperature-jump instruments are utilized, one for the detection of transmission

† Program Analyst, Humiston-Keeling & Co., 233 E. Erie, Chicago, Ill. 60611.
‡ Equipment & Software Systems, Burroughs Corporation, Detroit, Mich. 48232.
* Author to whom inquiries should be addressed at Morton 4-609, 303 E. Chicago Ave., Chicago, Ill. 60611.

0095-2338/79/1619-0169$01.00/0 

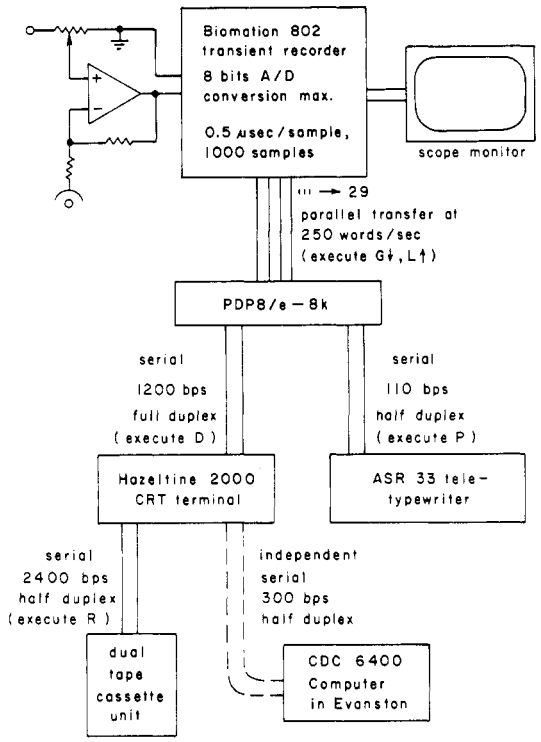


**Figure 1.** Overall diagram of the computerized data acquisition system. The amplifier shown in the upper left corner is inserted between photodiodes and the input of the Biomation 802 for concentration jump experiments only. For temperature-jump experiments a wide band and fast isolation amplifier is used, producing a ten times amplification for the adaptation to the input characteristics of the Biomation 802 transient recorder. A Tektronix 602 storage oscilloscope is connected to the Biomation 802 to show the stored information. The transient recorder is directly connected to a PDP 8/e minicomputer with control information normally appearing on the Hazeltine 2000 CRT screen. Some minor editing is available before the data are transferred from the core of the PDP 8/e onto the Hazeltine tape cassette unit. From there, data are transferred onto the disk of a large CDC 6400 via telephone lines. Command features are indicated in parentheses; they are available upon interactive transfer (and editing) of the original data (see Figure 2).

changes, the other one for the detection of fluorescence changes.[7] We also utilize three types of stopped-flow equipment for kinetic experimentation (including pH-jump measurements). The signal is generally fed through an amplifier and filtering network with a variety of rise and fall time constants available.[8]

The layout of the equipment around the PDP 8/e minicomputer is shown in Figure 1 (including a ten times wide band pre-amplifier, which amplifies the signal from temperature-jump experiments). The data acquisition is controlled through interaction with displays on a Hazeltine 2000 CRT terminal. The controlling messages and the full interactive communication for the various parts of the program are shown in Figure 2. Response by entering single letters increases speed of execution. The "Header Card" (=line) is manually added and provides vital information on the content of the set of data. The items, separated by a comma, mean: identification of experiment, vertical conversion factor (mV per full scale), horizontal conversion factor (ms per point), range multiplier for horizontal conversion factor (−1 gives factor $10^{-3}$, 0 gives factor 1.0, +1 gives factor $10^{+3}$), class descriptor for type of experiment (−1 describes blank, 0 describes indicating [= absorbing or fluorescing] species alone, +1 describes system with indicating species and one reactant, etc.); spaces thereafter (none submitted for blank) denote concentrations for indicator and any reactants. This information is submitted here to facilitate later stages of evaluation (see Discussion). The

```
SELECT MODE OF OPERATION (L,R,D,G,P,E)
R
PUT CASSETTE IN ONLINE-CONT-CHAR. MODE
(HIT RETURN TO CONTINUE)

TYPE IN THE HEADER CARD
82.75A,2000,45,0,-1,
OUTPUT COMPRESSION? (YES OR NO)
NO
WHICH CASSETTE DRIVE? (1 OR 2)
2
REWIND TAPE? (YES OR NO)
NO
STARTING LINE NUMBER=?
1
UPTO LINE NUMBER=?
4
IS EVERY THING READY?
(HIT RETURN TO CONTINUE)

82.75A,2000,45,0,-1,
208 208 209 208 208 207 205 204 204 206
207 207 208 208 209 212 212 213 213 211
212 211 210 208 207 208 206 204 206 207
208 207 206 207 206 206 208 208 210 210
SELECT MODE OF OPERATION (L,R,D,G,P,E)
```

**Figure 2.** Output of the control program on the Hazeltine 2000 CRT terminal, showing the user reponse by single letters. This output was obtained by transferring the contents of the video screen onto the Hazeltine thermal printer (temporarily on loan from the Hazeltine Corporation). The six letters stand for: R = Record data on tape cassette; D = Display data (in sections); G = Get data from the Biomation 802; P = Punch data onto paper tape; L = Load scales onto Biomation 802; E = End program. The "Mode of Operation" is executed as indicated in Figure 1 (parenthesized notation). Recording of noncompressed data is shown. Compressed data is used to conserve space (omission of blanks).

"Trailer Line" (at end of each set) contains information on buffer system, ionic strength, pH, temperature, etc. (needed in later evaluations). A teletypewriter may also be connected to the minicomputer and is essential for the initial loading of the binary programs from paper tapes. The electronic details and the instruction set of the interface have been reported previously.[6]

## RESULTS

Data acquisition was performed with the described equipment on three different instruments, namely, on the temperature-jump apparatus with detection of transmission changes (electron-transfer experiments on cytochrome *c*),[9] on the temperature-jump apparatus with fluorescence detection (binding of the fluorescing coenzyme NADH to liver alcohol dehydrogenase in the presence of inhibitors),[10] and on the concentration-jump apparatus with detection of transmission changes at 633 nm (detecting the various protonic forms of ferricytochrome *c*).[11] The routine for loading vertical and horizontal scan was implemented recently. Implementation of this routine reduces operator errors effectively. It is also possible (with the proper software for the computer) to enter scales for the Biomation 802.

After users became familiarized with the various parts and their integrated action, they rapidly acquired mastery of the equipment and could soon perform experiments faster than before. Previously, the output of the Biomation 802 was transferred through an interface into a teletypewriter paper tape punch. As the teletypewriter produces only ten characters per second, 10 min is required for the outputting of the data. With this new equipment, data are transferred out of the Biomation 802 and into the minicomputer within 4 s for a full set of 1024 "points". Subsequent transfer into the Hazeltine cassette unit takes only seconds. A new experiment can, in fact, be conducted every 3 min, adequate for thermal equilibration in the thermostatted temperature-jump cells and for almost all of the concentration-jump experiments.

Figure 3 depicts the flow diagram for data evaluation of a large number of recordings on several Hazeltine tape cassettes. They are combined successively until parameter values for the

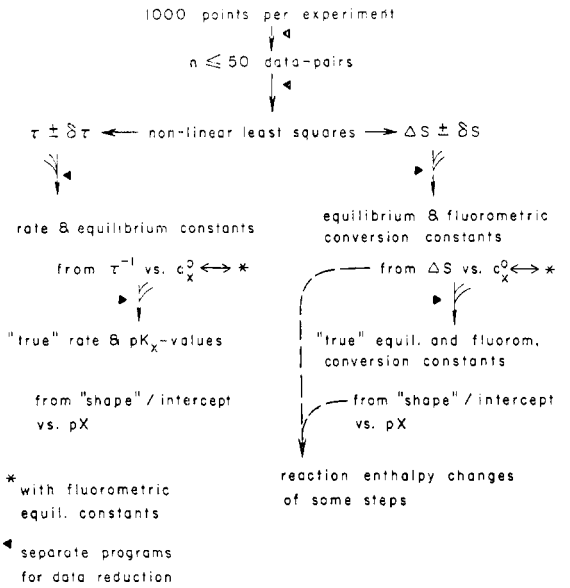


**Figure 3.** Flow of data evaluation, as employed for temperature-jump experiments with fluorescence detection. Although a floating point package was available, it was not used because of its size and as a subsequent program ("TRANSL") was already available on disk and operates on integer data input. The data, as produced by the Biomation 802, are stored in numbers from 0 to 255.

various physical constants in the system are obtained. Thus, only a small part of data processing is done on the PDP 8/e minicomputer. While Figure 3 refers to temperature-jump experiments with fluorescence detection, similar diagrams are available for temperature-jump experiments with observation of transmission changes and for concentration-jump experiments.

## DISCUSSION

The described equipment is adequately fast for ordinary needs in data acquisition and temporary initial storage (representing local backup storage). A current disadvantage is the slow transfer rate of the data into the large CDC 6400. An increase in the transmission rate to the CDC 6400 would be highly desirable (1200 bits per second is optimal). Nevertheless, the current configuration is of considerable advantage and should be of interest to others, producing transient data experimentally. Certainly, for many experiments it is not necessary to utilize a Biomation 802, which has a minimum sample conversion of 0.5 μs. Any other analog-to-digital conversion system with temporary core storage may be used in the place of the Biomation 802. However, the specifically described feature of loading for horizontal and vertical scales may not be available for comparatively simple systems. Furthermore, the temporary storage of data on cassettes provides a useful backup system for the CDC 6400.

Figure 3 presents only part of the whole evaluation process, needed in biochemical evaluations. One may describe the total evaluation process in a sequence of stages, which are briefly formulated in the following paragraphs. Stages 0, 1, and 2 are regularly used in all our evaluations. Stage 3 is used whenever sufficient information is available from the experiments. Stages 4 and 5 are currently not rigorously applied (the proper hardware and software facilities are currently not accessible).

**Stage 0.** This stage entails the conversion of analog signals into digital data, generally not directly usable. The data arrays consist of 250, 1000, or 2000 integer numbers which need to be converted to a smaller data matrix, containing the proper abscissa, the ordinate value, and the standard error of the ordinate value. These data matrices are called the "primary data" and actually stored as such (together with a "header line" and a "trailer line", which provide systems information; see Figure 2 and Methods section).

**Stage 1.** Primary data matrices (specifically of signal values with their standard errors vs. time) are evaluated for primary constants (chemical relaxation times and apparent rate constants). These primary constants are dependent upon various analytical concentrations and certain buffered equilibrium concentrations (mostly pH, more generally pX with X representing any sufficiently buffered concentration). The resulting secondary data matrices thus consist of columns of relaxation times, equilibrium signal changes, their errors, and the various analytical and buffered equilibrium concentrations.

**Stage 2.** Secondary data matrices are evaluated with reference to their concentration-independent parameters. These are pH-independent equilibrium and rate constants, as well as extinction coefficients (for detection of transmission changes, after proper conversion). If the chemical relaxation process is initiated by an external parameter (like temperature $T$), different values of the external parameter may be used, leading to tertiary data matrices.

**Stage 3.** Tertiary data matrices (like concentration-independent equilibrium and rate parameters vs. temperature $T$) are evaluated for the appropriate thermodynamic parameter (like enthalpy changes $\Delta H_j$ for the individual relaxation steps j [considering temperature-jump experiments]). One obtains then a quaternary data matrix of thermodynamic constants for a specific biochemical structure.

**Stage 4.** The various thermodynamic and other characteristic constants will be different from one biochemical configuration of a reactant to another and also from one enzyme configuration to another. A quaternary evaluation therefore compares the values of these various constants with the structural details of the biochemicals (possibly utilizing the three-dimensional coordinates from X-ray structure analyses). The various thermodynamic constants obtained are aligned in terms of crucial structural variables.

**Stage 5.** A set of thermodynamic parameters as a function of values of critical structural parameters would allow the extrapolation of thermodynamic characteristic values for compounds, which are not yet investigated, but can be structurally defined on the basis of the set of known data. This is the synthetic (or creative) stage of data evaluation. This is the only stage involving true (meaning predictive) advancement of knowledge.

## ACKNOWLEDGMENT

The authors appreciate funding of the PDP 8/e minicomputer by the National Science Foundation (GM-20823) and of all associated work by the National Institutes of Health (AA00282-02). They are also indebted to Drs. M. Wagner and R. Zabinski, as well as Ms. D. Kelley, for testing the usage of the equipment.

## REFERENCES AND NOTES

(1) Malmstadt, H. V.; Enke, C. G. "Digital Electronics for Scientists", W. A. Benjamin: New York, 1969.
(2) Enke, C. G.; Dessy, R. E. "Laboratory Computer Interfacing", prelim. ed., W. A. Benjamin, Menlo Park, Calif.
(3) Dessy, R. E.; Titus, J. A. *Anal. Chem.* **1973**, *45*, 124A.
(4) Holland, J. F.; Sweeley, C. C.; Thrust, R. E.; Teets, R. E.; Bieber, M. A. *Anal. Chem.* **1973**, *45*, 308.
(5) Holland, J. F.; Teets, R. E.; Timnick, A. *Anal. Chem.* **1973**, *45*, 145.
(6) Rao, P.; Haidle, R.; Czerlinski, G. Proceedings of the National Computer Conference, Chicago, 1974, pp 107–114.
(7) Czerlinski, G. *Rev. Sci. Instrum.* **1962**, *33*, 1184–1189.
(8) Czerlinski, G. *Rev. Sci. Instrum.* **1968**, *39*, 1730–1733.
(9) Zabinski, R. and Czerlinski, G. *J. Biol. Chem.* **1974**, *249*, 6125–6129.
(10) Czerlinski, G. H., Wagner, M., Erickson, J. O., and Theorell, H. *Acta Chem. Scand. B* **1975**, *29*, 797–810.
(11) Czerlinski, G.; Kelley, D., paper in preparation.