The calculation of kinetic constants of enzyme-catalyzed reactions using digital computers

A program has been written for the DATATRON 220 digital computer that will calculate, by means of a previously described orthogonal polynomial procedure[1,2], the so-called initial velocities of enzyme-catalyzed reactions, use these velocities, in conjunction with other known experimental parameters, to calculate kinetic constants, test the data, using the "t" test[3] to detect possible invalid points, and recalculate the constants after excluding such points. While the program, currently in use in these laboratories, is specifically designed for evaluation of the constants K_S and k_3 for systems exhibiting simple Michaelis-Menten kinetics, *e.g.*, those described by the equation $[S]_0[E]_0/v_0 = K_S/k_3 + [S]_0/k_3$, it can be readily adapted to other kinetic forms.

The program requires the introduction of nine chart coordinates of a trace of extent of reaction *vs.* time, taken at equal time intervals, the initial enzyme and substrate concentrations, the conversion factors for the chart scales, and the constants for the enzyme and substrate "blank" reactions. On introduction of a suitable command, the coefficients of the "least squares" curves of orders one through five are determined. The lowest, *i.e.*, the first-order coefficient, is then tested for significance according to the 90 % "t" test. If the coefficient is found to be significantly non-zero, the next higher coefficient, C_1, is tested by the same criterion and the process continued until the order of curvature, M, is found such that C_n, $n \leqslant M$ are significant and C_{M+1} is insignificant. The initial velocity is determined as before[1] from the initial slope of the curve of degree M, is corrected for the "blank" reactions and the quantities $y = [S]_0[E]_0/v_0$ and $x = [S]_0$ calculated and stored in appropriate memory locations. A tally is kept of the number of initial velocities evaluated and the x and y values for consecutive runs are introduced into consecutive memory locations.

When from 4 to 40 such calculations have been completed, and upon introduction of a suitable command, a linear least-squares fit of y *vs.* x is performed and values of K_S and k_3 computed from the apparent intercept and slope. The intercept, a, the slope, b, k_3 and K_S, and their respective uncertainties, x_i, y_i, $y'_i = a + bx_i$ and $d_i = y'_i - y_i$, for $i = 1$ to n are printed out.

The "worst point", *i.e.*, the one with the largest value of Z_i, where

$$Z_i = | d_i / \{ [(n+1)/n] \sum_{j=1}^{n} (x_j - \bar{x})^2 + (x_i - \bar{x})^2 \}^{\frac{1}{2}} |$$

and $\bar{x}$ is the average value of x, is chosen and the x and y values of this point removed from the active data and stored in a special memory location. A new least-squares fit is performed and the 98 %-confidence region about the new relationship is determined from the equation

$$A = t_{(0.98,\, n-2)} [\sum_{j=1}^{n} d_j^2/(n-2)]^{\frac{1}{2}} \{ [(n+1)/n] \sum_{j=1}^{n} (x_j - \bar{x})^2 + (x_i - \bar{x})^2 \}^{\frac{1}{2}},$$

where $t_{(0.98,\, n-2)}$ is the 98 % "t" coefficient for n--2 degrees of freedom. A is the interval along the y axis such that the probability that $(a + bx_i - A) < y_i < (a + bx_i + A)$ is 98 %, provided (x_i, y_i) is a member of the family of normally distributed points used to calculate a and b. The y coordinate of the "worst point" is then examined and

if it lies within the required interval, the computer is placed in a ready state for the introduction of new data. If the point lies outside the interval, the results based upon the new relationship are printed out and a new "worst point" is chosen and tested as above. This process is continued until the "worst point" passes the significance test or the data is reduced to three points.

Access can be made to each section of the program without involving any other. Velocities can be calculated without determination of kinetic constants and the constants can be evaluated from predetermined velocities. The program can also be keyboard modified to compute velocities according to any desired order of curvature, from one to five, rather than relying upon a significance test. On high-speed-tape output, each velocity calculation requires approx. 15 sec. Most of the time is used by the extensive output and this can be simplified if time of operation is a serious consideration. The y *vs.* x least-squares fit requires 15 sec for each iteration.

The preceding program was developed from an earlier one written for the DATATRON 205 computer, a model of lesser capability than the 220. Aside from differences arising from the different characteristics of the two computers the two programs were similar up to the point of obtaining initial velocities. However, in contrast to the 220 program, which employs a point by point significance test in the evaluation of K_S and k_3, the 205 program determined the confidence region about the initial least-squares fit of y *vs.* x and discarded all points lying outside this region before proceeding to a second and final least-squares fit. The 205 program also was unable to accept initial-velocity data obtained at more than one enzyme concentration and lacked self-clearing procedures which in the 220 program allows a virtually unlimited number of experiments to be evaluated without recycling the program tape.

One feature of the 205 program, which was not incorporated in the 220 program, involving a stepping procedure which first calculated a value of the initial velocity using n equally spaced time coordinates and then repeated the calculation for $n-1$ coordinates dropping the initial or $t = 0$ point. While this procedure is useful for detecting an aberrant initial point, generally arising from operational difficulties, it was not included in the 220 program because the same information can be obtained in the majority of cases by direct examination of the trace of extent of reaction *vs.* time. In those cases where it cannot, it is a simple matter, with the 220 program, to introduce a second set of data excluding the initial point.

This investigation was supported in part by a grant from the National Institutes of Health, U.S. Public Health Service.

*Gates and Crellin Laboratories of Chemistry**,
California Institute of Technology, Pasadena, Calif. (U.S.A.)

HENRY I. ABRASH
ABRAHAM N. KURTZ
CARL NIEMANN

[1] K. A. BOOMAN AND C. NIEMANN, *J. Am. Chem. Soc.*, 78 (1956) 3642.
[2] W. E. MILNE, *Numerical Calculus*, Princeton Univ. Press, Princeton, N.J., 1949.
[3] A. M. MOOD, *Introduction to the Theory of Statistics*, McGraw-Hill, New York, N.Y., 1950.

Received August 22nd, 1960

* Contribution No. 2615.