Optimizing Experimental Parameters in Isothermal Titration Calorimetry

Joel Tellinghuisen*

Department of Chemistry, Vanderbilt University, Nashville, Tennessee 37235

Received: June 29, 2005; In Final Form: August 26, 2005

In isothermal titration calorimetry, the statistical precisions with which the equilibrium constant (K) and reaction enthalpy ($\Delta H°$) can be estimated from data for 1:1 binding depend on a number of quantities, key among them being the products $c \equiv K[M]_0$ and $h \equiv \Delta H°[M]_0$, the stoichiometry range (R_m, ratio of total titrant X to total titrate M after the last injection), and the number of injections of titrant. A study of the statistical errors as functions of these quantities leads to the following prescription for optimizing throughput and precision: (1) Make 10 injections of titrant. (2) Set the concentrations in accord with the empirical equation $R_m = 6.4/c^{0.2} + 13/c$ (but no smaller than 1.1). (3) Make the starting concentration $[M]_0$ as large as possible within the large-signal limits of the instrumentation but limited to $c < 10^3$ for estimating K. With this procedure, both K and $[M]_0$ are predicted to have relative standard errors <1% over large ranges of K. Systematic errors in the concentrations, $[X]_0$ and $[M]_0$, are fully compensated by the "site number" or stoichiometry parameter (n). On the other hand, altering and freezing any of the fit parameters leads to a deterioration of the fit quality and to predictable changes in the other parameters. Fit divergence at very small c is avoidable through a simple redefinition of the fit parameters; however, unless n can be fixed from other information, $\Delta H°$ may be statistically ill-defined in this region.

Introduction

Isothermal titration calorimetry (ITC) is fast becoming the method of choice for characterizing chemical binding interactions of widely ranging types in even more widely ranging applications. For example, a recent check of the Science Citation Index for papers citing this method in the title, abstract, or keywords turned up over 300 papers in more than 60 journals in the year 2004 alone. The breadth of interest in this technique can be appreciated from a sampling of these papers, which examine chemical solution processes spanning size scales from metal-ion chelation to polymer interactions, binding affinities from modest (carbohydrate−protein) to strong (drug−protein), and applications from food science to drug design.[1−23] Although titration calorimetry dates back over 40 years,[24] the current burst of interest in this method can be traced to the more recent development of instruments that permit relatively rapid generation of titration curves for small samples (∼1 mL volume, ∼1 mM concentration of titrate).[25−29]

A strong selling point for ITC is that in a single experiment one can estimate both $\Delta H°$ and the equilibrium constant (K) (hence, also $\Delta G°$ and $\Delta S°$) for the targeted binding process at a specified temperature. The experiment involves sequential injection of one reactant (titrant) into a cell containing the other (titrate), producing a titration curve of heat (q) versus extent of reaction. The shape of this curve is closely related to the K value, while its scale is proportional to $\Delta H°$. Under favorable circumstances, both primary quantities can be obtained with very good precision (1% relative standard errors). The analysis involves nonlinear least-squares (LS) fitting, with most workers using the software provided by the manufacturers of the instruments.

* Phone: (615) 322-4873. Fax: (615) 343−1234. E-mail: joel.tellinghuisen@vanderbilt.edu.

Some readers may raise eyebrows over the just-stated 1% precision, because that quality is seldom achieved in routine work. Often, the precision is limited by complications inherent in the chemical system under study, which prevent it from conforming to the standard fit model for complexation. The resulting systematic errors are generally unknown, but they can lead to large fit residuals and hence large apparent statistical errors, as well. However, there are also indications that large uncertainties may stem from poor choices of operating parameters by experimenters. For example, in a recent comparative study of a test reaction by 14 participating groups,[30] the experimental parameters, raw data quality, and final results all exhibited considerable spread, and the ensemble statistics of the results were an order of magnitude worse than what would be suggested by the results from the individual experiments.

Among the parameters set by the ITC experimentalist are the concentrations of titrate and titrant, the injection volume, and the number of injections. The optimal choice of these parameters for characterizing 1:1 complexation, $M + X \rightleftarrows MX$, is the topic of the present work. A useful step in this direction was made in a recent article by Turnbull and Daranas,[31] who noted that good results can be achieved for values of the key quantity called "c" ($=K[M]_0$) much smaller than generally thought, provided that one extend the titration to a sufficiently large excess of titrant X. However, these authors did not address one important quantity, the number of injections. Also, it is possible to go much further than they did in specifying the conditions needed to best characterize the binding process by ITC for various limitations of nature and instrumentation.

In the present work, I have employed an approach first described in a theoretical consideration of statistical error in ITC, which focused on the possible role of nonconstant error.[32,33] Nonconstant error, or heteroscedasticity, requires weighted LS for optimal parameter estimation, whereas the commercial

10.1021/jp053550y CCC: $30.25
Published on Web 09/29/2005

software most often used for this purpose[31] employs unweighted fitting by default. In very recent work, I have analyzed the experimental variance function and confirmed the heteroscedasticity, for data obtained using the popular model VP-ITC instrument from MicroCal.[34] Knowledge of the data error permits reliable prediction of the LS parameter errors from the variance−covariance matrix for exactly fitting data,[35] thus avoiding the inconvenience and uncertainty of the Monte Carlo approach that is often used for this purpose.[31,36]

There are two major outcomes of the present effort:

(i) Under most circumstances, optimal results are obtained with 10 or fewer injections.

(ii) The concentrations should be adjusted to achieve a final excess of titrant over titrate ($[X]_0/[M]_0$ after the last injection) given by the empirically derived equation

$$R_m = \frac{6.4}{c^{0.2}} + \frac{13}{c} \tag{1}$$

where R_m is the stated ratio of total concentrations in the cell after the last injection and c is as defined above. The first of these results was already strongly indicated in the initial study[32] and was bolstered by the recent analysis of the data variance function.[34] For heat-starved reactions, even fewer injections may by optimal, resulting in even greater savings of time over the traditionally used ~25 injections. Equation 1 summarizes the results discussed below and holds for optimal estimation of K over the range $0.1 < c < 1000$, independent of $\Delta H°$ for the reaction. Of course, it requires at least approximate knowledge of the K value. Also, both results are premised upon confident foreknowledge that the process is truly 1:1 complexation.

In subsequent sections, I first review the nonlinear LS variance−covariance matrix, its use in experiment design, and its specific application to the problem of 1:1 complexation by ITC. I then discuss results for predicted statistical errors in K and $\Delta H°$ computed as functions of the number of injections (m) and the range (R_m) of the titration, which lead to the recommendation of 10 injections and to eq 1. These results are used to investigate the following question: "What should my starting titrate concentration $[M]_0$ be?" The answer is provided in the form of relative error contour plots, obtained as functions of K and $[M]_0$, with realistic experimental limitations taken into account.

Experiment Design with Least Squares

Linear least squares has several important "guarantees" that facilitate its use in experiment design.[37−39] First, if the data are of finite variance, the least-squares solutions for the adjustable parameters will be minimum-variance estimates, provided the data are weighted as σ_{yi}^{-2}, where σ_{yi} is the standard error in the dependent variable y at the ith point. Of course, this means that this data error must be known in at least a relative sense; it is also assumed that there is only one variable subject to random error, here taken as y. If the data are normally distributed about the "true" values of y (i.e., the errors have the Gaussian form), then the parameter estimates will be normally distributed about their true values (hence unbiased), with variances that are exactly calculable from the variance−covariance matrix ($\mathbf{V}$). The latter depends only on the structure of the independent variables (e.g., the x_i values in the simplest two-variable cases) and the error structure in the data. Thus, one can reliably predict the errors in the adjustable parameters for various assumed structures of the data and their statistical errors. By way of illustration, the fit model, $y = a + bx + cx^2 + d/x$, is a linear LS model with independent variable x (error free), dependent variable y, and adjustable parameters $a-d$. On the other hand, $y = 1/A$ is a nonlinear model designed to estimate the constant A from the single dependent variable y.

The guarantees of linear LS do not carry over in general to nonlinear LS. For example, in the last case just noted, one can readily verify that, even though $a = 1/A$ is a proper linear estimator, A is not normal and does not even have finite variance.[35,40] However, this distinction becomes unimportant when a is sufficiently precise in a relative sense. To quantify this statement, I have suggested a "10% rule of thumb": If the relative standard error in a nonlinear parameter is less than 10%, its distribution function will be close enough to Gaussian to permit estimation of its confidence limits in the usual way, with 10% reliability in these limits. This rule has been found to hold in a number of common nonlinear fitting problems, including the analysis of ITC data for 1:1 binding.[32,35] Furthermore, just as linear problems may be characterized a priori, from the structure of the data and error, so may nonlinear problems. In this case, however, the $\mathbf{V}$ matrix may depend on the values of the dependent variable, so it is evaluated for exactly fitting data, yielding what I have called the "exact" nonlinear $\mathbf{V}$.

This last statement requires clarification, since most scientists estimate LS parameter errors using the scatter in the fitted y_i values to determine the scale of the errors, which would result in zero errors for exactly fitting data. To this end, and for proper use of some commercial data analysis software, it is useful to distinguish the a priori variance−covariance matrix ($\mathbf{V}_{\text{prior}}$) from the a posteriori one ($\mathbf{V}_{\text{post}}$). The two differ by a scale factor

$$\mathbf{V}_{\text{post}} = \frac{\chi}{\nu}\mathbf{V}_{\text{prior}} \tag{2}$$

pwhere χ is the sum of weighted squared residuals and ν is the degrees of freedom, equal to the number of data points minus the number of adjustable parameters. The parameter variances are the diagonal elements of $\mathbf{V}$, and the notation emphasizes that $\mathbf{V}_{\text{prior}}$ is the appropriate form when the data error is known a priori (as in a Monte Carlo calculation), while $\mathbf{V}_{\text{post}}$ is needed when the data error is being estimated from the fit itself. In the latter case, χ/ν is an estimate of the variance in y for data of unit weight, while, for data of known σ_{yi}, χ/ν is an estimate of the reduced χ^2 for the fit (for which the statistical expectation value is unity). Both forms require that the data be given weights proportional to σ_{yi}^{-2}. For $\mathbf{V}_{\text{prior}}$, the weighting proportionality factor is 1.0; for $\mathbf{V}_{\text{post}}$, it is arbitrary (since the scale of σ_{yi} is now taken as unknown).

Those commercial data analysis programs that support weighted least-squares fitting typically require the "weights" to be provided as σ_{yi} values in an additional column in the data sheet. For example, this is the case for the Origin program (from MicroCal) and for KaleidaGraph (Synergy). For the latter, when weighting is invoked, the program assumes the data errors are known absolutely and returns the $\mathbf{V}_{\text{prior}}$-based parameter errors. Thus, if the supplied σ_{yi} values are intended to be valid in only a relative sense, as from a data transformation, the user must scale the resulting parameter errors by $(\chi/\nu)^{1/2}$ (with χ called "Chisq" in the fit results box). When weighting is invoked in the Origin program, the default is $\mathbf{V}_{\text{prior}}$, but the user can check the box "Scale errors by square root of reduced χ^2" to switch to $\mathbf{V}_{\text{post}}$.

In practice, the main weighting abuse is the use of unweighted LS to fit heteroscedastic data. There are two consequences of this neglect: (1) The fit no longer returns minimum-variance estimates of the parameters, and (2) the resulting $\mathbf{V}$ matrices

are incorrect. The magnitude of these effects depends on the scale of the heteroscedasticity. It is not likely to be a major problem if the data error ranges over a factor of ~5 or less, which is often the case for ITC data.[34]

To summarize the main points of this section, one can straightforwardly predict the parameter standard errors for any proposed LS fit model, linear or nonlinear, and one can do this using commercial data analysis packages such as Origin and KaleidaGraph. It is necessary to know or assume a data error structure (as is true also for a Monte Carlo calculation). For commercial programs other than the two I have covered specifically, it is important to understand the difference between $\mathbf{V}_{\text{prior}}$ and $\mathbf{V}_{\text{post}}$ and to know how the program handles this distinction. Monte Carlo calculations can provide useful instruction and corroboration but are only truly needed if the relative parameter errors exceed 10%, or if details about the parameter distributions and biases are desired. The results described below were carried out using in-house-devised FORTRAN codes that have been described elsewhere,[32−35,39] though they could have been done with the commercial packages, albeit with less efficiency, since these programs are oriented toward the analysis of only single data sets at a time.

Fit Model

The ITC instruments that appear to be most widely used are of the perfusion type, in which injection of volume v_i from the syringe results in the expulsion of the same volume of material up the stem of the cell. The model I have used to analyze the data for 1:1 binding assumes that (1) prior to each injection the system is uniform and at equilibrium and (2) the expelled material is of this equilibrium composition and is not involved in either the mixing or the heat production in the active volume (V_0) of the cell. For most of the results described below, the injection volume (v_i) is taken to be constant (v) for all injections. After the ith injection, the total concentrations (free and complexed) of X and M are given by[32]

$$[\mathrm{X}]_{0,i} = [\mathrm{X}]_0(1 - d^i) \quad \text{and} \quad [\mathrm{M}]_{0,i} = [\mathrm{M}]_0 d^i \tag{3}$$

where $[\mathrm{X}]_0$ is the concentration of titrant in the syringe and $[\mathrm{M}]_0$ is the starting concentration of M in the reaction vessel. The dilution factor is $d = 1 - v/V_0$. At equilibrium, the concentrations of reactants and product satisfy the expression

$$\frac{[MX]_i}{([X]_{0,i} - [MX]_i)([M]_{0,i} - [MX]_i)} = K \equiv K^\circ \times (\mathrm{L\ mol^{-1}}) \tag{4}$$

The number of moles of complex produced by the ith injection is thus

$$\Delta n_i = V_0[\mathrm{MX}]_i - (V_0 - v)[\mathrm{MX}]_{i-1} = V_0\{[\mathrm{MX}]_i - d[\mathrm{MX}]_{i-1}\} \tag{5}$$

and the associated heat is

$$q_i = \Delta H^\circ \Delta n_i \tag{6}$$

As before, I ignore experimental complications such as the need to estimate heats of dilution for the titrant, and the related concentration dependence of q_i. I also use the dimensionless thermodynamic K° interchangeably with K, which is equivalent to assuming all activity coefficients are unity.

As was noted above, a key parameter is the extent to which the titration is driven with excess titrant. In terms of the fit model just described, the parameter R_m is defined as

$$R_m = \frac{[X]_{0,m}}{[M]_{0,m}} \tag{7}$$

where the concentrations are the totals (bound and unbound) in the cell after the final (mth) injection.

This fit model is the same as that used by El Harrous et al.[41] but differs slightly from that employed in the Origin software provided by MicroCal and used by Turnbull and Daranas.[31] In essence, the model used here assumes slow mixing after injection, while the Origin model assumes fast mixing. To some extent, the two models represent extremes that are accessible to the experimentalist (through the adjustable stirring speed). I am aware of no results that might permit critical comparison of the two. This might anyway be difficult, since my computational checks show maximum parameter differences <1%, in a case where the injection volume is 20 μL. Thus, we can be confident that the parameter standard errors will be insensitive to these model differences.

The fit model has two adjustable parameters, ΔH° and K, and as many data points as injections. The software in general use includes a third parameter, the "site number" (n). For well-understood 1:1 complexation, this is simply a concentration correction factor, needed to put the concentrations of X and M on a common footing. Its inclusion is often important for achieving a good fit of typical ITC data; here, it is (as usual) taken as a correction factor to $[\mathrm{M}]_0$.

The data error was taken to be the preferred function E from the recent analysis of 37 data sets for the complexation of Ba^{2+} with 18-crown-6 ether[34]

$$\sigma_i^2 = \sigma^2[1 + a^2 q_i^2 + b^2(q_i/v_i)^2] \tag{8}$$

with $\sigma = 0.8$ μcal, $a = 0.003$ $\mu\mathrm{cal}^{-1}$, and $b = 0.02$ μL/μcal. This means that the proportional error term from a exceeds that from the constant error for $q \approx 300$ μcal, which is relatively large for many realistic situations. By comparison, the data error assumed by Turnbull and Daranas[31] was $\sigma = 0.5$ μcal and $a = 0.01$ $\mu\mathrm{cal}^{-1}$, with a switchover from constant to proportional error at $q = 100$ μcal. Thus, the experimentally derived data error[34] actually makes heteroscedasticity less of a problem than implied in ref 31. Accordingly, while weighted LS is used at all times in the present work, we can still confidently predict that unweighted LS will suffice for heat-starved reactions.

The volume error term in b was not present in ref 31. This error equals that in a when $v \approx 7$ μL. The calculations are done assuming a fixed total titrant volume of mv, so for large m, this term will exceed the a term. However, as m increases, q_i/v becomes roughly constant, neutralizing the expected increased error for small v.

Results and Discussion

Figure 1 shows typical titration curves for K and c values differing by a factor of 100, for titration to the typical range $R_m = 2$, and for titration to $R_m = 20$ for the smaller K value, as directed by eq 1. For $R_m = 2$, the reaction is 97% complete for c ($=K[\mathrm{M}]_0$) = 100 but only 56% complete for $c = 1$. Increasing R_m 10-fold brings the low-c reaction to 94% completion and greatly increases the precision in the determination of K and ΔH°.

Figures 2 and 3 illustrate how the relative parameter errors can be examined as functions of the key operational parameters, for extremes in the range of c and h ($\equiv \Delta H^\circ[M]_0$) covered in

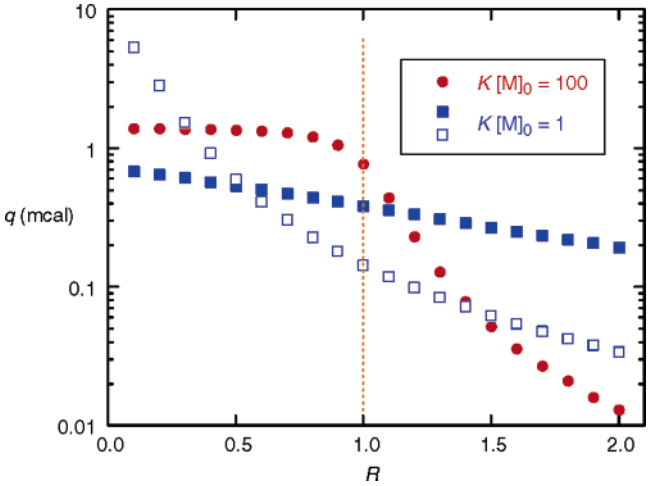

Figure 1. Synthetic ITC data for $[M]_0 = 1$ mM, $\Delta H° = 10$ kcal/mol, $V_0 = 1.4$ mL, and $K = 10^5$ M^{-1} (circles) and 10^3 M^{-1} (squares), displayed as a function of the stoichiometry ratio (R). The open points have their R axis compressed by a factor of 10. The dashed vertical line represents the nominal titration end point for the solid points. Identical displays result from other combinations of $[M]_0$, K, and $\Delta H°$, yielding c ($=K[M]_0$) $= 1$ and 100 and $\Delta H°[M]_0$ ($\equiv h$) $= 10$ cal/L.

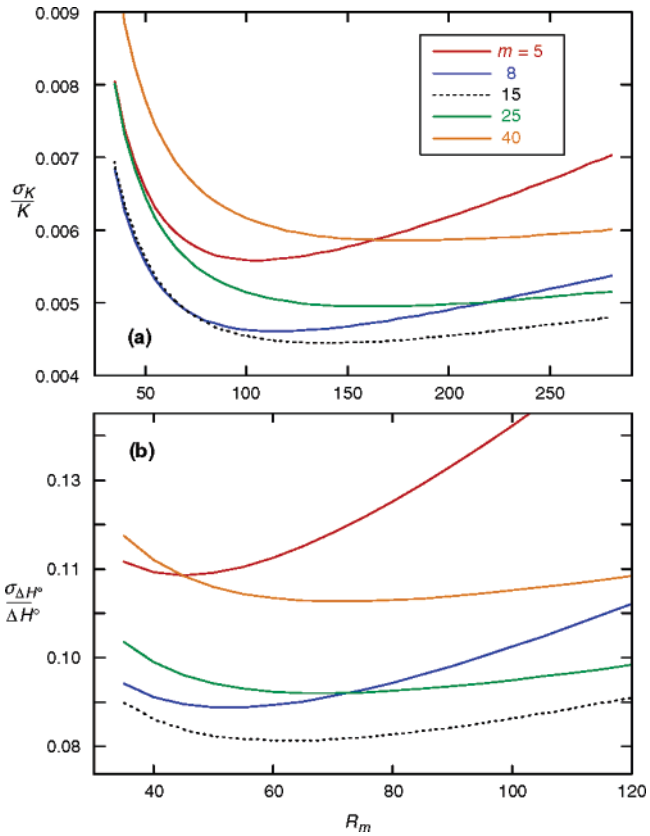

Figure 2. Relative errors in K (a) and $\Delta H°$ (b) as functions of the range of titration (R_m) and the number of injections, for $c = 0.1$ and $h = 10$ cal/L.

the present work. In most such calculations, the total titrant volume was fixed at 0.1 mL, but the relative errors are not very sensitive to this choice. The structure in Figure 3 is real, as the parameter errors are quite sensitive to the precise values of R_m and m for large c. The figures show one general result of the study, namely, that the optimum range for determining $\Delta H°$ is somewhat smaller than that for K. It is also noteworthy that, for m and R_m near the minima in such displays, there is little loss in precision from altering the optimal values by ~25%.

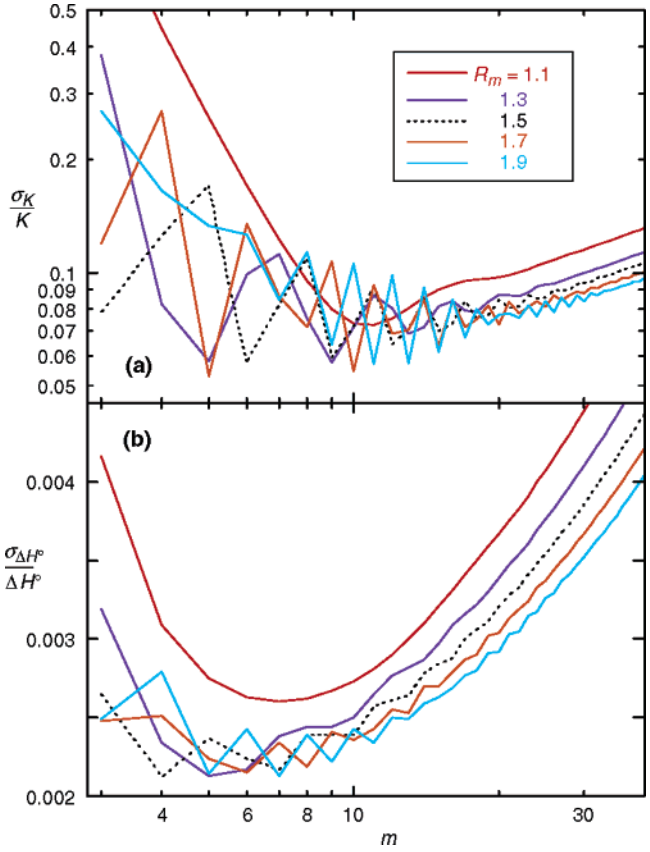

Figure 3. Relative errors in K (a) and $\Delta H°$ (b) as functions of the number of injections and the range of titration, for $c = 1000$ and $h = 1$ cal/L.

One other trend worthy of comment is the manner in which the relative errors first decline with increasing m and then increase again. This behavior is a consequence of the playoff between the constant error term and the "a" term in the proportional error.[34] Without the latter, the error almost invariably increases with increasing m. This somewhat surprising result follows from the manner in which a fixed total signal (the heat of reaction) is subdivided into progressively smaller pieces.[32,42] In this process, for constant σ_q, the relative error increases in proportion to m, more than off-setting the statistical averaging effect, which improves the precision only as $m^{1/2}$.

From results such as those in Figures 2 and 3, I have estimated the m and R_m values that yield the best precision, leading to Figure 4. Considering first part b, the choice of R_m shows no dependence on h for K and only a weak dependence for $\Delta H°$. The fitted curve for K is eq 1; it was obtained by fitting both sets of data together to a sum of powers, with subsequent rounding of the fitted parameters to the indicated values. From part a, both K and $\Delta H°$ prefer larger m for larger heat. This is consistent with the explanation above regarding the increased precision for small m in the weak signal limit. The logic in favor of $m = 10$ as the "default" is as follows: Even though larger m values are favored for large h, this is the regime where results are typically so precise that they are more likely to be limited by systematic error than by random error, for example, from limitations in the fit model, related to the way in which heats of dilution and activity coefficients are treated.[43] For heat-starved reactions, $m = 10$ is a reasonable compromise that gives some redundancy and will not lead to serious loss of precision as compared with the optimal value.

There is a practical reason that in some cases covered here it may be necessary to go to larger m. The calculations to this

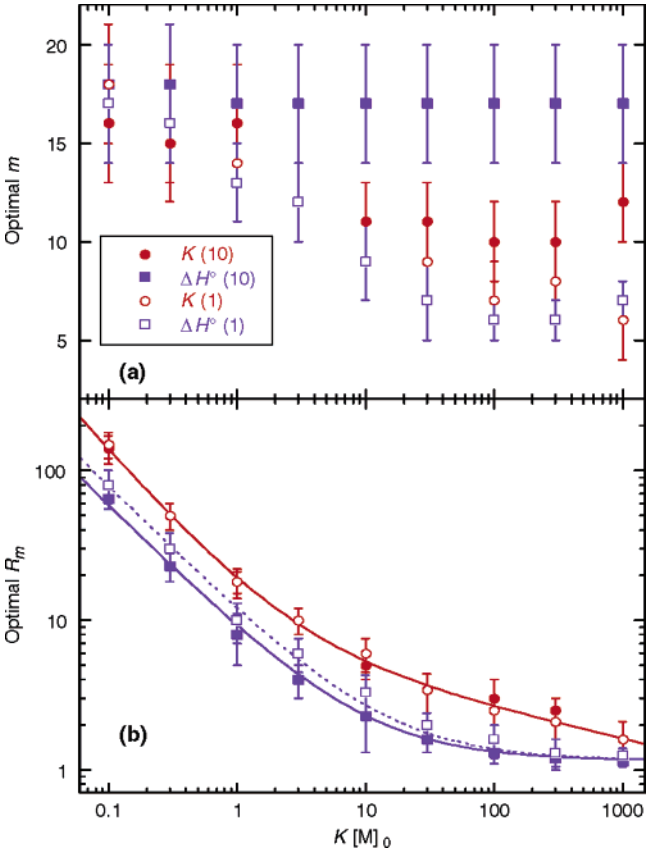

Figure 4. Values of m (a) and R_m (b) yielding the best precision in K and $\Delta H°$, as functions of c, for $h = 10$ cal/L (solid points) and $h = 1$ cal/L (open points). The error bars are intended as guides to the precision with which these estimates could be made from plots such as those of Figures 2 and 3. The fitted curve on the results for K in part b is eq 1; the fitted curves for $\Delta H°$ are of the form $R_m = 1.16 + g/c^{0.85}$, with $g = 8.2$ (solid) and 11 (dashed).

point have not considered experimental limitations on the large-signal side. Most binding processes are exothermic, and for power compensation instruments, too much heat can cause the power to "go negative", where the results become unusable. For such cases, it may be necessary to increase m or otherwise modify the injection procedure to reduce the peak heat production. The low-c cases are particularly prominent in this regard, because a constant-v injection program with 10 injections can lead to more than half the total reaction occurring in the first injection. One way to avoid this problem is by varying the injection volume so as to yield roughly constant q for each injection. This is done by solving the equilibrium eq 4 to yield constant incremental product formation. I have verified that this approach yields precision insignificantly different from that of the standard constant-v procedure, in the minimum-error region.

From Figure 4b, the best $\Delta H°$ estimates require smaller R_m than dictated by eq 1, so we might ask how important this difference is. The answer is "not very", as is illustrated in Figure 5, where eq 1 is used to define R_m for ITC data spanning the range of c of interest, with results compared with the best possible. There is ~50% loss in precision in $\Delta H°$ at large c, but in this region, the precision is already better than generally permitted by subtle systematic problems. The loss at small c is comparable but still not a great concern in most cases.

It is interesting that when the recommended procedure—$m = 10$ and R_m given by eq 1—is followed, the relative precision in K is roughly independent of c over much of the range of interest, while that in $\Delta H°$ varies by almost 3 orders of

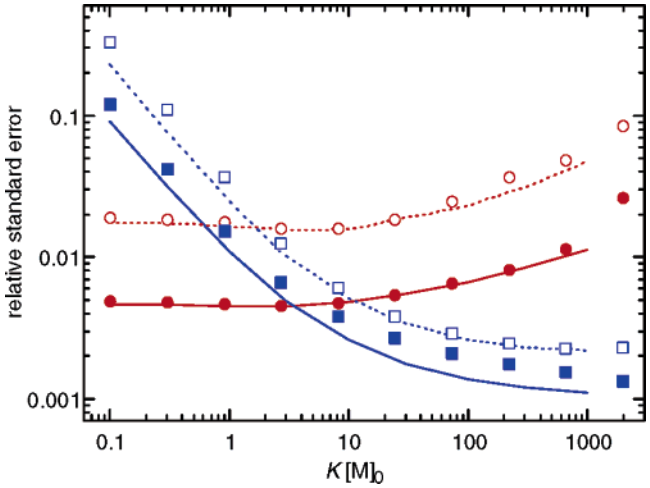

Figure 5. Relative standard errors in K (circles) and $\Delta H°$ (squares), as computed over the indicated range of c, using eq 1 to set R_m, with $m = 10$. The lines represent the optimal values obtained from results such as those in Figures 2 and 3. Solid points and lines are for $h = 10$ cal/L, and open points and dashed lines are for $h = 1$ cal/L.

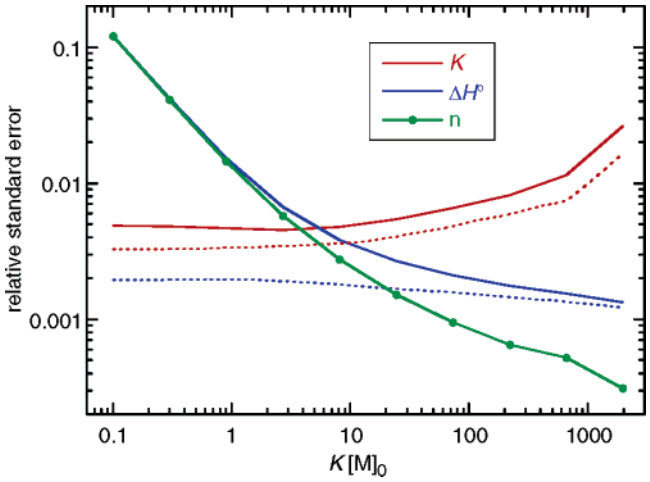

Figure 6. Relative errors in K, $\Delta H°$, and n, computed as for the solid points in Figure 5 and then recomputed after fixing $n = 1$ (dashed curves).

magnitude. At small c, the precision in $\Delta H°$ tracks that in n, as shown in Figure 6. Both become quite imprecise in the small-c limit. However, they are also strongly correlated, as is evident from the greatly improved precision for $\Delta H°$ when n is fixed and removed from the fit. Thus, if the stoichiometry can be confidently established independent of the ITC experiment, fixing n at its known value can greatly improve the determination of $\Delta H°$, as was noted by Turnbull and Daranas.[31] However, it is still wise to examine the fitted value of n to confirm that it is statistically consistent with its supposedly known value.

The results shown in Figures 4−6 provide useful guidance, but they do not quite address the following question: "What $[M]_0$ should I use"? The intuitive answer is "as large as possible", and that turns out to be close to the truth. To provide a more quantitative answer, I have computed the relative standard errors in K and $\Delta H°$ as funcions of K and $[M]_0$ and plotted precision contours. In doing so, I have also attempted to include realistic instrumental limitations, by (1) discarding any conditions that led to a total $q > 50$ mcal (which is likely to yield defective data for exothermic reactions) and (2) taking $[X]_0 = 1$ M as the maximum allowed titrant concentration. There is also a needed modification of eq 1, which was derived by

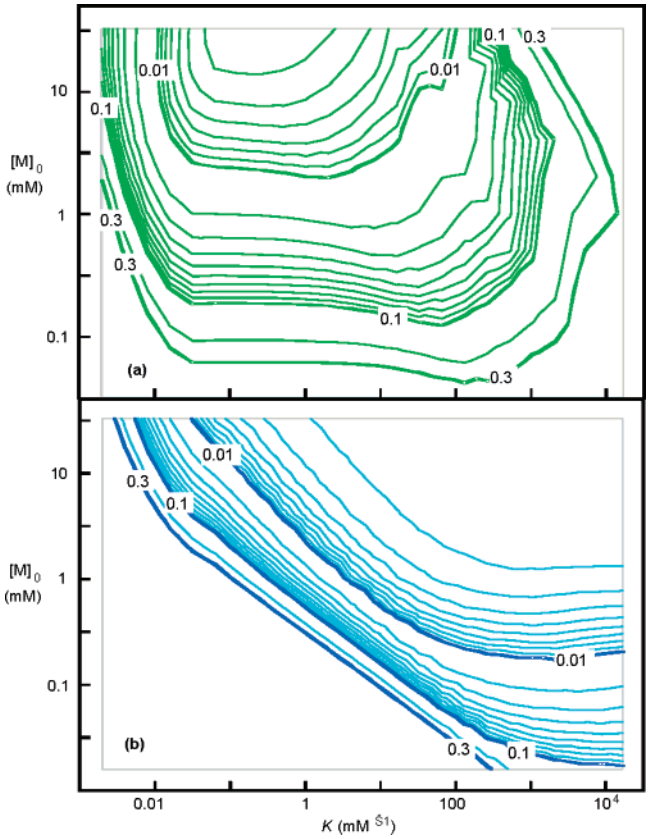

Figure 7. Precision contours (relative standard errors) for K (a) and $\Delta H°$ (b), for $\Delta H° = 1$ kcal/mol. For each K and $[M]_0$, R_m values were computed using eq 1 but were limited to $R_m = 1.1$ on the small end and were also effectively limited on the upper end by setting the maximum value of $[X]_0$ at 1 M. The number of injections was $m = 10$ at all times, with $v = 30$ μL. Conditions yielding a total reaction $q > 50$ mcal were deleted, effectively providing the upper cutoff on these plots.

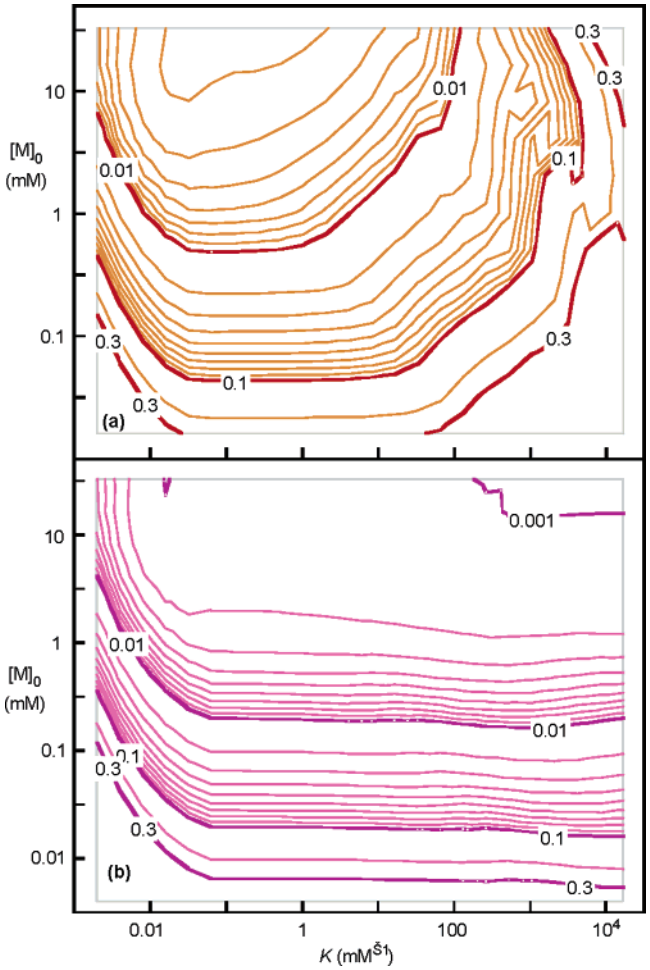

Figure 8. Precision contours (relative standard errors) for K (a) and $\Delta H°$ (b), when n is frozen at 1.00. Other conditions are the same as those in Figure 7.

fitting data over only the range $0.1 < c < 10^3$: At large c, this equation can yield values of $R_m < 1$, which was avoided by setting a minimum value of 1.1.

Figure 7 displays results of such calculations for $\Delta H° = 1$ kcal/mol. The preference for the upper-right corner in part b for $\Delta H°$ is consistent with expectations from Figures 5 and 6, and the creation of a "valley of precision" for K is also hinted at in these figures. It is interesting, however, that the nominal 1% region spans the approximate c range 0.1−3000, which is much broader than normally thought. The presently realized precision at low c is a direct consequence of the large R_m value of ∼100 dictated by eq 1 as compared with the much smaller values typically chosen in past work.

The results for $\Delta H° = 10$ kcal/mol were qualitatively similar to those in Figure 7 but with the $[M]_0$ axes shifted down about a factor of 10. This is because the upper large-q cutoff now occurs for a factor of 10 smaller $[M]_0$. When n was fixed at 1.0 and removed from the fit, the new contours were lowered appreciably, and those for $\Delta H°$ were rendered almost horizontal (Figure 8). Both effects are as anticipated from Figure 6. Note that when n can be fixed in this manner, the 1% region for both key quantities extends almost down to $c = 0.01$.

It is interesting that eq 1 leads to a value of $[X]_0$ in the syringe that is almost independent of $[X]_0$ at small c, as is illustrated in Figure 9. In this range, the second term in eq 1 dominates, making $[X]_{0,m} \approx 13/K$ and $[X]_0$ in the syringe larger by approximately the cell/syringe volume ratio.

The procedure of fixing n from other information about concentrations was mentioned earlier. Sometimes, it is also

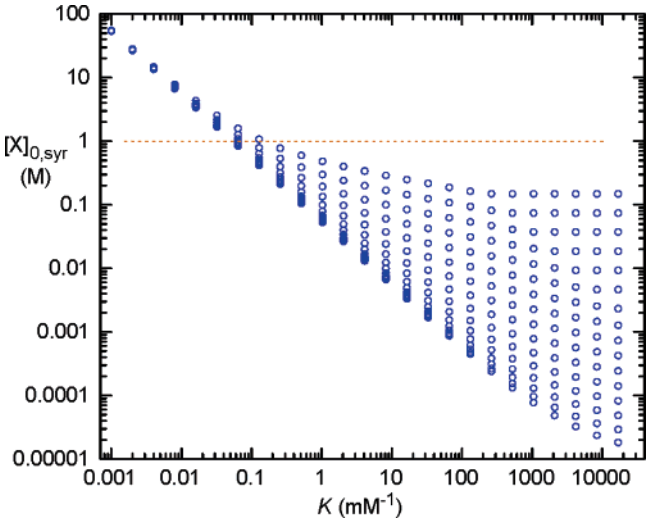

Figure 9. Computed titrant syringe concentrations for the results illustrated in Figures 7 and 8, showing low-c dependence on K alone. The total titrant volume was taken as 0.3 mL, and the cell volume, as 1.4 mL. The dashed horizontal line shows the adopted cutoff value of 1 M.

justifiable to fix K or $\Delta H°$ from other work, so it is useful to examine the effects of changes in the stated concentrations on the estimated values of the parameters, and of any one parameter on the others. Such effects are illustrated in Table 1, where it can be seen that some changes have no effect on $\int$ for the fit

TABLE 1: Effects of Changes in Concentrations and *n* on Other ITC Fit Quantities[a,b]

c	$[X]_0$	$[M]_0$	*n*	*K*	$\Delta H°$	χ^2 [c]
	×1	**×*f***	$\times f^{-1}$	unch	unch	unch
	×*f*	**×1**	$\times f^{-1}$	$\times f^{-1}$	$\times f^{-1}$	unch
	×*f*	**×*g***	$\times f/g$	$\times f^{-1}$	$\times f^{-1}$	unch
1	**×1.01**	**×1**	**×1**	×0.988	×1.0005	+0.63
1	**×1**	**×1.01**	**×1**	×1.0026	×0.9896	+0.65
100	**×1.01**	**×1**	**×1**	×0.931	×1.0043	+161
100	**×1**	**×1.01**	**×1**	×1.066	×0.9849	+176
1	**×1**	**×1**	$+f\sigma_n$	$+0.71f\sigma_K$	$-0.98f\sigma_{\Delta H°}$	$+f^2$
1	**×1**	**×1**	$+0.71f\sigma_n$	$+f\sigma_K$	$-0.63f\sigma_{\Delta H°}$	$+f^2$
1	**×1**	**×1**	$-0.98f\sigma_n$	$-0.63f\sigma_K$	$+f\sigma_{\Delta H°}$	$+0.97f^2$
100	**×1**	**×1**	$+f\sigma_n$	$+0.69f\sigma_K$	$-0.63f\sigma_{\Delta H°}$	$+f^2$
100	**×1**	**×1**	$+0.69f\sigma_n$	$+f\sigma_K$	$-0.27f\sigma_{\Delta H°}$	$+0.99f^2$
100	**×1**	**×1**	$-0.63f\sigma_n$	$-0.27f\sigma_K$	$+f\sigma_{\Delta H°}$	$+f^2$

[a] Quantities given in bold are controlled or held fixed in the indicated manner; the other quantities represent the resulting changes for 1% increments or for the general factors *f* and *g* (positive or negative). *n* is treated as a correction factor to the concentration $[M]_0$. [b] The results in the first three rows are general; the others were obtained for 10 injections, using eq 1 to set $[X]_0$, eq 8 for the data error, $v = 20$ μL, and $h = 10$ cal/L. In the last six rows, changes are in terms of the parameter standard errors from the original three-parameter fit and are strictly valid only for small f (<1). [c] $\chi^2 = \mathcal{J}/\nu$, where ν is the number of degrees of freedom, here, 7 (=10 points − 3 adjustable parameters); "unch" means unchanged. The entries in the last six rows represent the changes when the standard errors are from $\mathbf{V}_{prior}$. When $\mathbf{V}_{post}$ is used, χ^2 increases by the factor $(1 + f^2/\nu)$ for linear fits, or by approximately $(1 + xf^2/\nu)$ for the present nonlinear fits, where $x \neq 1$ in two cases.

and can be determined exactly for any data set. Thus, the first three rows in Table 1 show that changes in the stated concentrations can be completely absorbed by *n*, with no effect on the apparent quality of the fit, as shown by $\mathcal{J}$ (hence χ^2). On the other hand, when either concentration is altered with *n* held constant, there is an effect on the parameters and on χ^2, with the increase in χ^2 being quite drastic for $c = 100$ (cf. the χ^2 average value of 7). The results in rows 4−7 show that errors in concentration have a bigger relative effect on *K* than on $\Delta H°$. This seems at odds with the results illustrated by Turnbull and Daranas[31] (their Figure 4). However, they discussed the effect on $\Delta G°$ rather than on *K*, and because of the logarithmic relation ($\Delta G° = -RT \ln K°$), the indicated changes in *K* do represent very small absolute changes in $\Delta G°$. Also, these authors allowed *n* to float for their large-*c* results ($c = 50$), so those are exactly predictable from the first three rows in Table 1. Note that the effect on χ^2 of fixing $n = 1$ for erroneous concentrations would normally be considered acceptable at $c = 1$ but would never be tolerated at $c = 100$.

The last six rows of Table 1 show the effects of altering any one of the parameters from its fitted value. These results were obtained by changing the targeted parameter using $f = 1$ and refitting the others. However, the coefficients that appear in the parameter columns are also the correlation coefficients from the fit, obtainable from the **V** matrix using $C_{ij} = V_{ij}(V_{ii}V_{jj})^{-1/2}$. Strictly speaking, these are valid only in the small-*f* limit, though they would hold exactly for any *f* if the fit were linear (in which case, the last six entries in the last column would all be exactly f^2).[44] Also, the effects of parameter changes depend on *c* and *h*, as the degree of interparameter correlation changes with these quantities. Thus, the factor −0.98 in rows 4 and 6 illustrates the very strong correlation (negative) between *n* and $\Delta H°$ at small *c*. In contrast, the correlation between *K* and $\Delta H°$ is small at large *c*. The correlations also depend on R_m and *m*, so the

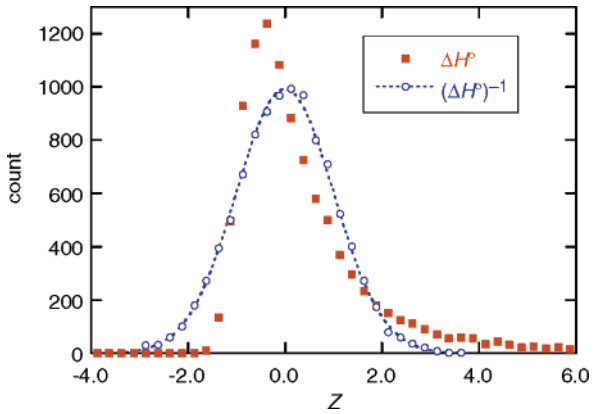

Figure 10. Histogrammed results for $\Delta H°$ and its reciprocal, from a Monte Carlo run of 10^4 data sets having $K = 100$ M^{-1}, $[M]_0 = 1$ mM, $\Delta H° = 10$ kcal/mol, $m = 10$, $R_m = 140$, $v = 20$ mL, and data error 3 times that given in eq 8. The argument *Z* is the displacement from the true value in units of the nominal standard errors, which are 33.9% of the value in both cases. The curve shown for the reciprocal is a fitted Gaussian having displacement and width insignificantly different from the true values of 0 and 1, respectively, and yielding $\chi^2 = 25.4$ (27 fitted points).

specific values of the coefficients in the last six rows of Table 1 will not hold for *m* and R_m different from the values used here.

Because of the very strong correlation between *n* and $\Delta H°$ at small *c*, there is a potential problem even carrying out the nonlinear fits needed to estimate all three parameters simultaneously.[31] To investigate this matter, I conducted Monte Carlo computations on simulated data sets having $c = 0.1$ and $h = 10$ cal/L. Equation 1 prescribes a titration range of 140 for this case. With this choice and the data error set by eq 8, I experienced only one divergence in 10^4 simulated data sets. However, as the data error was scaled up progressively, divergences did occur—3% of the time for an error scale factor of 2 rising to 19% for a factor of 5. At this point, the nominal relative statistical errors in both *n* and $\Delta H°$ exceed 50%, so fixing $n = 1$ would normally be statistically justified. Interestingly, dropping the titration range to $R_m = 40$ actually decreased the divergence rate by almost half. This can be understood by reference to Figure 2b, which shows better inherent precision for $\Delta H°$ at $R_m = 40$ than at 140 (where the precision in *K* is optimized).

There are computational tricks that can be used to tease convergence out of even the most troublesome data set at small *c*. Although *n* and $\Delta H°$ are very highly corelated and thus ill-defined, their product is very well-defined and not strongly correlated with either *n* or $\Delta H°$ individually, nor with K. Thus, defining the nonlinear adustable parameters as *K*, $\Delta H°$, and $(n \times \Delta H°)$ leads to rapid convergence of the nonlinear fit and no divergences, up to an error scale factor of at least 10. The MC statistics from such runs helps explain the divergences in the normal fit. Even with an error scale factor of only 3, the divergent runs from the normal fit yield such anomalously large $\Delta H°$ values as to make the statistics for this quantity diverge, for example, an average value 31 times too large and a statistical error 10^3 larger in one run of 10^4 data sets! A closer examination of the statistics of $\Delta H°$ and its reciprocal show that the latter is well-behaved and close to normal (Figure 10), confirming that the divergence of $\Delta H°$ is another example of "reciprocal behavior" identified earlier for other nonlinear fitting problems.[35,40] In further confirmation of this, the nonlinear fit parameters can be redefined as *K*, *n*, and $(1/\Delta H°)$, again producing perfect convergence behavior up to an error scale

factor of at least 20, where the relative errors in n and $\Delta H°$ are more than 220%. Of course, making the fit converge in no way changes the pathological statistical behavior of $\Delta H°$ itself under these circumstances.

Conclusion

The statistical errors in the key parameters K and $\Delta H°$ from ITC data are functions of several quantities dictated by nature and several under the control of the experimentalist. These dependences are examined as functions of the number of injections of titrant and the stoichiometry range of the titration, for the case of 1:1 binding and a range of values of $c = K[M]_0$ and $h = \Delta H°[M]_0$. The method involves computation of the exact a priori variance−covariance matrix for the nonlinear least-squares model, with the data error taken from an earlier comprehensive study of ITC experimental data. The results lead to the recommendation to use no more than 10 injections of titrant in most circumstances and yield a simple prescription for adjusting the titrant concentration from approximate knowledge of the value of K for the reaction. When this prescription is followed, both K and $\Delta H°$ are predicted to have relative error below 1% over broad ranges of K—from 10 to 10^5 M^{-1} for K itself and from $K = 30$ M^{-1} up for $\Delta H°$.

In practice, the 1% goal may remain elusive in much ITC work, because of materials limitations (insolubility, scarcity) and systematic errors in procedures and fit models. Among the identified systematic errors are ones associated with backlash in the injection of titrant[45] and active volume ambiguities.[46] Backlash problems are easily avoided through a simple procedural modification and should anyway seldom exceed 1%. The volume error may be as large as 3% but is again easily corrected. Limitations in fit models remain a greater unknown, but their presence is revealed in, for example, persistent statistical inconsistency between estimates of $\Delta H°$ obtained directly and those extracted from the T dependence of K (van't Hoff method).[43] Although there has been a tendency to blame such inconsistencies on K, inadequate treatment of dilution heat effects may also be compromising the calorimetric estimates of $\Delta H°$. This could be an even greater problem if large initial concentrations $[X]_0$ are used in the small-c case. In this regard, it is useful to note again the relatively broad nature of the curves in Figure 2. Thus, dropping R_m to 30 from the prescribed value of 140 for $c = 0.1$ increases the error in K by less than a factor of 2 and in $\Delta H°$ by only ~20%. For the many ITC applications that tolerate relative errors of 10% or more, these increases will be unimportant. In this regard, note that Turnbull and Daranas[31] set 0.4 and 1 kcal/mol as their precision benchmarks for $\Delta G°$ and $\Delta H°$, respectively. The first represents a relative standard error >65% in K, and the second is 10−100% of the typical 1−10 kcal/mol range of $\Delta H°$ covered in the present computations.

The highest $[M]_0$ values covered in Figures 7 and 8 may also exceed the high-signal limitations of the instrument for exothermic binding, unless special methods are used. These might include increasing the number of injections or spreading each injection out in time. Such methods will certainly be needed at very low c, where the usual constant-v program can yield half the total enthalpy in the first injection. In this case, a variable-v, constant-q scheme may be particularly useful.

It is relatively easy to reproduce specific results illustrated here with the widely used Origin software provided by MicroCal and other data analysis programs such as KaleidaGraph. In the former, one must access the weighting option by opening the Control box under the Options menu and selecting one of the weighting modes. Some weighting modes can be invoked by just checking a box, but to use error functions such as eq 8, it is necessary to compute the data errors in a separate column of the data sheet. One can then compute the parameter errors using both the a priori and a posteriori approaches, by simply repeating the computations before and after checking the box "Scale errors by square root of reduced χ^2". To incorporate convergence tricks such as redefining the fit parameters, it is necessary to write one's own fit definition functions; a KaleidaGraph example requiring the entry of eight short statements in the program's macro library is available on request.

Acknowledgment. I have benefited greatly from the facilities of the Center for Structural Biology at Vanderbilt, where the experiments leading to the estimation of the data error function were conducted. I also wish to thank Laura Mizoue for assistance with the experiments, Mike Brandts of MicroCal for generously sharing information about the VP-ITC instrument, and Ed Lewis, Mike Doyle, and Brad Chaires for discussions and suggestions that have led to many of the results displayed and discussed here.

References and Notes

(1) Gourishankar, A.; Shukla, S.; Ganesh, K. N.; Sastry, M. *J. Am. Chem. Soc.* **2004**, *126*, 13186.
(2) Thongngam M.; McClements, D. J. *J. Agric. Food Chem.* **2004**, *52*, 987.
(3) Arnaud A.; Bouteiller, L. *Langmuir* **2004**, *20*, 6858.
(4) Dai, S.; Tam, K. C.; Wyn-Jones, E.; Jenkins, R. D. *J. Phys. Chem. B* **2004**, *108*, 4979.
(5) Saro, D.; Klosi, E.; Paredes, A.; Spaller, M. R. *Org. Lett.* **2004**, *6*, 3429.
(6) Sinn, C. G.; Dimova, R.; Antonietti, M. *Macromolecules* **2004**, *37*, 3444.
(7) Peters, W. B.; Edmondson, S. P.; Shriver, J. W. *J. Mol. Biol.* **2004**, *343*, 339.
(8) Wu, Y. G.; Tabata, M. *J. Solution Chem.* **2004**, *33*, 777.
(9) Kolobe, D.; Sayed, Y.; Dirr, H. W. *Biochem. J.* **2004**, *382*, 703.
(10) Spencer, S. D.; Raffa, R. B. *Pharm. Res.* **2004**, *21*, 1642.
(11) Mikulecky, P. J.; Feig, A. L. *Nucleic Acids Res.* **2004**, *32*, 3967.
(12) Loregian, A.; Appleton, B. A.; Hogle, J. M.; Coen, D. M. *J. Virol.* **2004**, *78*, 9084.
(13) Talhout, R.; Stafforst, T.; Engberts, J. B. N. *J. Colloid Interface Sci.* **2004**, *276*, 212.
(14) Li, Z. G.; Mintzer, E.; Bittman, R. *Chem. Phys. Lipids* **2004**, *130*, 197.
(15) Kitagishi, H.; Satake, A.; Kobuke, Y. *Inorg. Chem.* **2004**, *43*, 3394.
(16) Heerklotz, H. *J. Phys.: Condens. Matter* **2004**, *16*, R441.
(17) Gerebtzoff, G.; Li-Blatter, X.; Fischer, H.; Frentzel, A.; Seelig, A. *ChemBioChem* **2004**, *5*, 676.
(18) Kim, H. S. *J. Ind. Eng. Chem.* **2004**, *10*, 273.
(19) Jacobs, S. A.; Fischle, W.; Khorasanizadeh, S. *Methods Enzymol.* **2004**, *376*, 131.
(20) Goncalves, E.; Debs, R. J.; Heath, T. D. *Biophys. J.* **2004**, *86*, 1554.
(21) Boland, J. S.; Davidson, P. M.; Bruce, B.; Weiss, J. *J. Food Prot.* **2004**, *67*, 285.
(22) Piervincenzi, R. T.; Chilkoti, A. *Biomol. Eng.* **2004**, *21*, 33.
(23) Seong, H. R.; Kim, D. S.; Kim, S. G.; Choi, H. J.; Ahn, K. H. *Tetrahedron Lett.* **2004**, *45*, 723.
(24) Christensen, J. J.; Izatt, R. M.; Hansen, L. D. *Rev. Sci. Instrum.* **1965**, *36*, 779.
(25) Beaudette, N. V.; Langerman, N. *Anal. Biochem.* **1978**, *90*, 693.
(26) Spokane, R. B.; Gill, S. J. *Rev. Sci. Instrum.* **1981**, *52*, 1728.
(27) Ramsay, G.; Prabhu, R.; Freire, E. *Biochemistry* **1986**, *25*, 2265.
(28) Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L.-N. *Anal. Biochem.* **1989**, *179*, 131.
(29) El Harrous, M.; Gill, S. J.; Parody-Morreale, A. *Meas. Sci. Technol.* **1994**, *5*, 1065.
(30) Myszka, D. G.; Abdiche, Y. N.; Arisaka, F.; Byron, O.; Eisenstein, E.; Hensley, P.; Thomson, J. A.; Lombardo, C. R.; Schwarz, F.; Stafford, W.; Doyle, M. L. *J. Biomol. Technol.* **2003**, *14*, 247.
(31) Turnbull, W. B.; Daranas, A. H. *J. Am. Chem. Soc.* **2003**, *125*, 14859.
(32) Tellinghuisen, J. *Anal. Biochem.* **2003**, *321*, 79.
(33) Tellinghuisen, J. *Methods Enzymol.* **2003**, *383*, 245.
(34) Tellinghuisen, J. *Anal. Biochem.* **2005**, *343*, 106.

(35) Tellinghuisen, J. *J. Phys. Chem. A* **2000**, *104*, 2834.
(36) Alper, J. S.; Gelb, R. I. *J. Phys. Chem.* **1990**, *94*, 4741.
(37) Mood, A. M.; Graybill, F. A. *Introduction to the Theory of Statistics*, 2nd ed.; McGraw-Hill: New York, 1963.
(38) Hamilton, W. C. *Statistics in Physical Science: Estimation, Hypothesis Testing, and Least Squares*; The Ronald Press Co.: New York, 1964.
(39) Tellinghuisen, J. *J. Chem. Educ.* **2005**, *82*, 157.
(40) Tellinghuisen, J. *J. Phys. Chem. A* **2000**, *104*, 11829.
(41) El Harrous, M.; Mayorga, O. L.; Parody-Morreale, A. *Meas. Sci. Technol.* **1994**, *5* 1071.
(42) Doyle, M. L.; Simmons, J. H.; Gill, S. J. *Biopolymers* **1990**, *29*, 1129.
(43) Mizoue, L. S.; Tellinghuisen, J. *Biophys. Chem.* **2004**, *110*, 15.
(44) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes*; Cambridge University Press: Cambridge, U.K., 1986.
(45) Mizoue, L. S.; Tellinghuisen, J. *Anal. Biochem.* **2004**, *326*, 125.
(46) Tellinghuisen, J. *Anal. Biochem.* **2004**, *333*, 405.