Biophysical Chemistry, 38 (1990) 159–165
Elsevier

BIOCHE 01514

Evaluation of Adair binding constants from experimental data obtained with an Imai cell

Statistical weighting functions

Michael L. Johnson and Ann E. Lassiter

Department of Pharmacology, Interdisciplinary Biophysics Program, Diabetes Endocrinology Research Center, and Biodynamics Institute, Box 448, University of Virginia Health Sciences Center, Charlottesville, VA 22908, U.S.A.

Received 2 May 1990
Accepted 22 May 1990

Hemoglobin; Oxygen binding; Adair constant; Self-assembly; Linkage; Cooperativity

One group of laboratories uses an 'even-weighted', or unweighted, nonlinear least-squares method for the analysis of experimental oxygen binding data obtained with an Imai oxygenation apparatus. Another group uses an 'end-weighted' nonlinear least-squares analysis. With end weighting each observation is assigned a statistical weight which is proportional to $Y(1-Y)$, where Y is the fractional saturation. In this work we discuss statistical weighting functions as applied to the Imai oxygenation apparatus and then determined what are the best weighting factors for an actual series of published experiments. Based on these calculations, it is concluded that the 'best' weighting for the Imai oxygenation apparatus is a very small amount of end weighting. Furthermore, the amount of end weighting is so small that even weighting, or no weighting at all, is also appropriate.

1. Introduction

For two decades the Imai oxygenation apparatus [1] has proven to be one of the premier methods which is used for the measurement of oxygen binding isotherms. The basic design of the Imai oxygenation apparatus is a spectrophotometer cell which contains an oxygen electrode, and a stirring mechanism that allows for the exchange of the gas phase which is in equilibrium with the solution being studied. The fractional saturation of a solution, typically hemoglobin in a dilute buffer, is monitored by an oxygenation-sensitive spectral change. The oxygen concentration in the solution is measured with the oxygen electrode. The oxygen concentration is titrated by flowing gas mixtures with different oxygen concentrations over the stirred solution. An entire oxygen titration can usually be obtained by flowing air over a deoxygenated hemoglobin solution and simultaneously monitoring the spectral change and oxygen electrode. Several laboratories have used the Imai oxygenation apparatus for extensive studies of hemoglobin cooperativity [2–17].

While the basic operation of the Imai oxygenation apparatus is generally the same in all of the laboratories, the methods of data analysis are distinctly different. One group of laboratories [1–5,17] performs their experiments at a single hemoglobin concentration while the other group [6–16] utilizes a series of different hemoglobin concentrations. The difference has been discussed in detail elsewhere [19,20] and will not be dealt with here.

Correspondence address: M.L. Johnson, Box 448, Department of Pharmacology, University of Virginia, Charlottesville, VA 22908, U.S.A.

A second significant difference between the laboratories is the amount of statistical weight to be assigned to each observation. For one group of laboratories the weighted nonlinear least-squares analysis uses even weighting [6–16] and for the other group the weighted nonlinear least-squares analysis uses end weighting [1–5,17]. Even weighting is where each observation is assigned the same statistical weight, i.e., the nonlinear least-squares parameter estimation is actually unweighted. With end weighting each observation is assigned a statistical weight which is proportional to $Y(1-Y)$, where Y is the fractional saturation [1,17]. It should be noted that the relative weights of the individual data points will vary by orders of magnitude with the end weighting scheme. The purpose of this work is to discuss the choice of statistical weighting functions, as applied to the Imai oxygenation apparatus, and then to determine what are the best weighting factors for an actual series of published experiments.

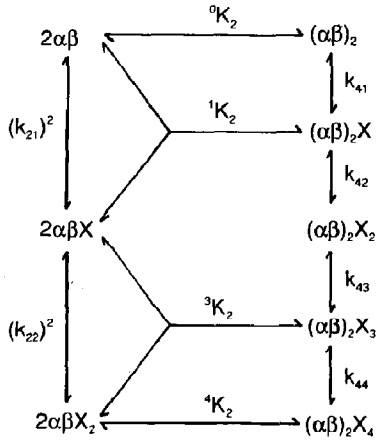

Fig. 1. The seven-parameter oxygenation-linked dimer, $\alpha\beta$, to tetramer, $(\alpha\beta)_2$, reaction scheme as defined by Ackers and Halvorson [13]. The $^iK_2$s are the subunit assembly constants to form a tetramer with i ligands bound. The k_{2i}s are the stepwise Adair constants for binding oxygen to hemoglobin dimers. The k_{4i}s are the stepwise Adair constants for binding oxygen to hemoglobin tetramers.

2. Methods

The experimental data was analyzed by a nonlinear least-squares method utilizing a series of different weighting functions. This procedure determines a set of equilibrium constants from a set of actual experimental data by an iterative comparison of the data with a calculated function. The calculated function, eqs 1–5, is given by the oxygenation-linked dimer-to-tetramer association scheme (fig. 1) [21]. The nonlinear least-squares method has also been published elsewhere [18].

Hemoglobin A_0 undergoes a reversible dimer-to-tetramer association reaction which is linked to the degree of oxygenation of both the dimeric and tetrameric species [21]. The stepwise Adair constants for dimer and tetramer and the dimer to tetramer are defined schematically in fig. 1.

The seven-parameter binding formulation is:

$$\bar{Y}_{2,4} = \frac{Z_2' + \left[Z_4' \left\{ \left(Z_2^2 + 4^0K_2Z_4[P_t] \right)^{1/2} - Z_2 \right\} / (4Z_4) \right]}{Z_2 + \left(Z_2^2 + 4^0K_2Z_4[P_t] \right)^{1/2}} \quad (1)$$

where

$$Z_2 = 1 + K_{21}[X] + K_{22}[X]^2 \quad (2)$$

$$Z_2' = K_{21}[X] + 2K_{22}[X]^2 \quad (3)$$

$$Z_4 = 1 + K_{41}[X] + K_{42}[X]^2 + K_{43}[X]^3 + K_{44}[X]^4 \quad (4)$$

$$Z_4' = K_{41}[X] + 2K_{42}[X]^2 + 3K_{43}[X]^3 + 4K_{44}[X]^4 \quad (5)$$

where K_{21} and K_{22} are the product Adair constants for oxygen binding to the $\alpha\beta$ dimer, K_{41}, K_{42}, K_{43} and K_{44} the tetrameric product Adair constants, $^0K_2$ the dimer-tetramer association constant ($\alpha\beta \leftrightarrow (\alpha\beta)_2$) at zero oxygen concentration, [X] the oxygen concentration, and $[P_t]$ the heme concentration. The product Adair constants, K_{ij}, are defined in terms of the stepwise Adair constants, k_{ij}, as

$$K_{21} = k_{21} \quad (6)$$

$$K_{22} = k_{21} {}^* k_{22} \quad (7)$$

$$K_{41} = k_{41} \quad (8)$$

$$K_{42} = k_{41}{}^{*}k_{42} \tag{9}$$

$$K_{43} = k_{41}{}^{*}k_{42}{}^{*}k_{43} \tag{10}$$

$$K_{44} = k_{41}{}^{*}k_{42}{}^{*}k_{43}{}^{*}k_{44} \tag{11}$$

Due to the cyclic nature of the oxygenation-linked dimer-tetramer association scheme, we were able to choose a particular set of seven parameters; ${}^{0}K_2$, ${}^{4}K_2$, K_{44}, ${}^{0}K_2/{}^{1}K_2$, ${}^{3}K_2/{}^{4}K_2$, K_{coop2}, and the square root of k_{43}. The definitions and the reasons for choosing these parameters have been discussed extensively elsewhere [6,16,20]. Once a nonlinear least-squares parameter estimation has been performed on this parameter set then the other parameters can be evaluated as published elsewhere [18].

3. Data set to be analyzed

We have chosen the hemoglobin A_0 pH 8.5 data set published by Chu et al. [10] because it contains 14 separate oxygen titrations at protein concentrations ranging from 2 to 85 μM heme and a total of 884 oxygen binding data points. In addition to these oxygen titrations, the data set includes a gel permeation column chromatography determination of the fully liganded dimer-to-tetramer association constant, ${}^{4}K_2$, and a kinetic determination of the forward and reverse rate constants for the unliganded dimer-to-tetramer association, and thus ${}^{0}K_2$. The experimental conditions for this data set were 0.1 M Tris, pH 8.5, 0.1 M NaCl, and 1.0 mM Na_2EDTA.

4. Results and discussion

The method of analysis used by both groups of laboratories for oxygen binding data is weighted nonlinear least-squares. In nonlinear least-squares analysis the parameter values (Adair constants or other equilibrium constants) with the highest probability of being correct are evaluated and secondly the statistical confidence (standard error) of these parameters is evaluated [18].

In order for a weighted nonlinear least-squares analysis to yield parameter values with the maximum likelihood (highest probability) of being correct, a number of assumption must be met. First, it must be assumed that all of the statistical uncertainties are in the dependent variable, i.e., fractional saturation, $\bar{Y}_{2,4}$ in eqs 1–5. This assumption means that the independent variable, i.e., the oxygen concentrations, X in eqs 1–5, is known to infinite precision. Second, it must be assumed that the experimental uncertainties in the dependent variable, i.e., fractional saturation, are random and follow a Gaussian probability distribution. Third, it must be assumed that no systematic uncertainties exist in the dependent variable or the independent variable. For a more complete discussion of these and other assumptions, readers are referred to the article of Johnson and Frasier [18].

Within these assumptions it can be shown that the parameter values with the maximum probability of being correct are obtained by any procedure which minimizes the weighted least-squares norm of the data [18]. The weighted least-squares norm, X^2, is

$$X^2 = \sum \left[\{ Y_i - Y_{i,\mathrm{cal}} \} / \sigma_{Y_i} \right]^2 \tag{12}$$

where the summation is performed over each of the 'i' data points, Y_i is the fractional saturation of data point i, $Y_{i,\mathrm{cal}}$ the 'calculated' value of the fractional saturation based on the current parameter values and dependent variable (oxygen concentrations) of data point i, and σ_{Y_i} the standard error of the fractional saturation (statistical weight) of data point i. When the least-squares norm, X^2, is minimized the parameter values are those with the maximum likelihood of being correct. There is a possibility that the nonlinear least-squares process can yield multiple sets of parameter values which have nearly equivalent minima, X^2 [18].

It is clear that the choice of values of σ_{Y_i} can and will significantly alter the value of X^2 and consequently may alter the estimated parameter values. It is impossible to predict a priori the magnitude of this alteration. However, it is clearly that the choice of σ_{Y_i} is not arbitrary. There are two possible types of experimental uncertainties in the Imai apparatus which will be discussed here; those arising from the spectrophotometer and those from the oxygen electrode. It should be

noted that the important item here is the relative contributions from the different sources of experimental uncertainties and the relative uncertainties of the individual data points, not the exact values.

The estimation of the spectrophotometric uncertainties is straightforward. The uncertainty of most spectrophotometers is a constant of about 0.1% of the full-scale setting of the spectrophotometer. However, it should be borne in mind that in most cases the spectrophotometer readings are transformed into fractional saturation values before the nonlinear least-squares analysis is performed. This transformation is given by eq. 13;

$$Y_i = \{A_i - A_0\}/\{A_\infty - A_0\} \tag{13}$$

where A_i is the absorbance at data point i, A_∞ and A_0 being those which correspond to fractional saturations of unity and zero. The value of A_0 is usually determined by a quadratic extrapolation of the absorbance values, A_i, to zero oxygen concentration. The value of A_∞ is usually determined by a quadratic extrapolation of the absorbance values, A_i, vs the inverse of the oxygen concentration to zero. Due to the inverse the extrapolation is, in effect, actually to infinite oxygen concentration. In most cases these extrapolations involve between five and ten data points where the fractional saturation is within 5% of the endpoints. In a quadratic extrapolation of this type the uncertainty in A_∞ and A_0 will be equal to the uncertainty in A_i divided by the square root of the number of data points used for the extrapolations. In other words, σ_{A_∞} and σ_{A_0} are both approximately equal to $\sigma_{A_i}/3$. Consequently, the estimate of the experimental uncertainties which originate with the spectrophotometer uncertainties includes contributions from the determination of A_∞ and A_0 in addition to the spectrophotometer uncertainties, A_i. The resulting spectrophotometer uncertainties, σ_{A_i}, are the root mean square of the uncertainties due to A_i, A_∞, and A_0. Combining all of these, the uncertainty in Y_i due to spectrophotometer errors is given by eq. 14;

$$\sigma_{Y_i}^2 = \left[\sigma_{A_i}^2 + \left\{Y_i\sigma_{A_i}/3\right\}^2 + \left\{(1 - Y_i)\sigma_{A_i}/3\right\}^2\right] /\{A_\infty - A_0\}^2 \tag{14}$$

where the Y_i and $\{A_\infty - A_0\}$ terms originate in the derivatives of Y (eq. 13) with respect to A_∞, A_0, and A_i.

If spectrophotometric errors were the only source of experimental uncertainties then the best possible values of σ_{Y_i} would be given by eq. 14. The best weighting scheme for spectrophotometric uncertainties, as predicted by eq. 14, is one where the center of the fractional saturation has a few percent less uncertainty than the ends and thus is actually a 'center-weighted' scheme. However, we have not as yet considered the consequences of uncertainties in the oxygen electrode measurements.

Including uncertainties in the oxygen electrode measurements is a direct contradiction of one of the assumptions which was required to derive the least-squares norm, X^2. It was assumed that all of the statistical uncertainties are in the dependent variable, i.e., fractional saturation. There are two possible solutions, either the use of a maximum likelihood method which correctly includes uncertainties in the independent variables [22–24] or the making of an approximation which allows the use of nonlinear least-squares. We will examine the approximate solution, since it is the most commonly used method.

The approximation is to project, or to reflect, the uncertainties of the independent variables, in this case oxygen concentration, onto the dependent variable, in this case fractional saturation. This is accomplished by assuming that the fitting function, in this case eqs 1–5, is linear in the region of the particular data point, X_iY_i. The value of the uncertainty in the fractional saturation, σ_{Y_i}, is then approximated as the root mean square of the uncertainty given by eq. 14 and the reflected uncertainty in the oxygen concentrations. The reflected uncertainty is calculated as the product of the uncertainty of the dependent variable, i.e., oxygen concentration, and the derivative of the dependent variable (i.e., fractional saturation) with respect to the independent variable (oxygen concentration), $\partial Y_i/\partial X_i$.

This is a standard approximation which is widely used in the statistical literature, usually without consideration of its validity. This approximation is probably reasonable if the frac-

tional uncertainty in the data is small. In the case of hemoglobin oxygen binding data as measured in an Imai apparatus, the variance of fit is on the order of $2-5 \times 10^{-5}$. Consequently, the approximation is probably reasonable.

This approximation is not without complications. At both ends of the oxygen titration the derivative of the saturation function with respect to the oxygen concentration is relatively small and as a consequence, the reflection of the uncertainties in the oxygen concentration onto the fractional saturation axis will be small. However, near the center of the titration the hemoglobin can be very cooperative and thus the derivative can be very large. In the present data set the maximal $\partial Y/\partial X$ is approx. 3.3×10^5 per mol oxygen. A consequence of this is very large values of the reflected uncertainties. This can lead to a situation where data points will have weighting functions which vary by orders of magnitude even though the actual uncertainties in the absorbance and oxygen concentration are not strong functions of the fractional saturation. The net result of this large variation in the weighting factors is to use so little statistical weight in the center of the fractional saturation range that the data points here are, in effect, discarded. Consequently, this procedure is suspect when the majority of the weighting factors originate from a reflection of the uncertainties of the independent variables.

In recapitulation, we have identified four sources of uncertainties (σ_{A_∞}, σ_{A_0}, σ_{A_i} and σ_{X_i}) which must be summed (as a root mean square) into the values of σ_{Y_i} to be used in the nonlinear least-squares parameter estimation. Furthermore, σ_{A_∞} and σ_{A_0} have been shown to have comparatively little effect. Consequently, there are two significant sources of random uncertainties; σ_{A_i} whose contribution is approximately constant and σ_{X_i} whose contribution is proportional to the derivative of the fractional saturation with respect to the ligand concentration. Therefore, if a least-squares parameter estimation is performed with the 'correct' weighting factors then the residuals, the weighted differences between the fitted function and the original data points, will not be a function of the derivative of the fractional saturation with respect to the ligand concentration.

Table 1

Analysis with equal weighting factors

Values presented below are the intrinsic (corrected for statistical factors) free energy changes for the binding of successive oxygens to tetrameric hemoglobin A_0. These were determined using the oxygenation-linked dimer-to-tetramer formulation of the binding equations, eqs. 1–5 in the text. The experimental data set used for this nonlinear least-squares analysis was the pH 8.5 data from Chu et al. [10]. σ^2 is the variance in fractional saturation units, $\Delta G'_{4i}$ is the free energy to bind an oxygen to a hemoglobin tetramer which already has $i-1$ oxygens bound (kcal/mol oxygen) and Q.E. is the 'quaternary enhancement' free energy change which has been reported elsewhere [7,10,16]. +C.I. and −C.I. denote the plus and minus one standard deviation confidence intervals. The experimentally determined values for the unliganded dimer-to-tetramer association constant, ${}^0K_2 = 6.949 \times 10^9$, and the liganded dimer-to-tetramer association constant, ${}^4K_2 = 7.252 \times 10^6$, were used for these calculations.

	Most probable	−C.I.	+C.I.
$\Delta G'_{41}$	−6.035	−6.078	−6.000
$\Delta G'_{42}$	−6.829	−6.986	−6.674
$\Delta G'_{43}$	−7.503	−7.720	−7.272
$\Delta G'_{44}$	−9.476	−9.587	−9.372
Q.E.	−1.010	−1.121	−0.906
σ^2	5.747×10^{-5}		

Table 1 lists the free energy changes for oxygen binding to hemoglobin A_0 tetramers, obtained by using a nonlinear least-squares parameter estimation on the pH 8.5 data set of Chu et al. [10] and an equal weighting of all the data points. Fig. 2A is a plot of the residuals, the differences between the fitted function and the original data points, as a function of the derivative of the fractional saturation with respect to the ligand concentration. It is clear from fig. 2A that there is a small, but significant, increase in the magnitude of the residuals as a function of the derivative of the fractional saturation with respect to the ligand concentration increases.

Consequently, a series of other weighting functions were tested. These weighting functions, σ_{Y_i}, were generated by eq. 15;

$$\sigma_{Y_i}^2 = 1 + [A\ \partial Y_i/\partial X_i]^2 \qquad (15)$$

where A is a relative scaling factor which is expressed as a percentage of the maximum oxygen concentration (2.78×10^{-4} M) in the data set. The

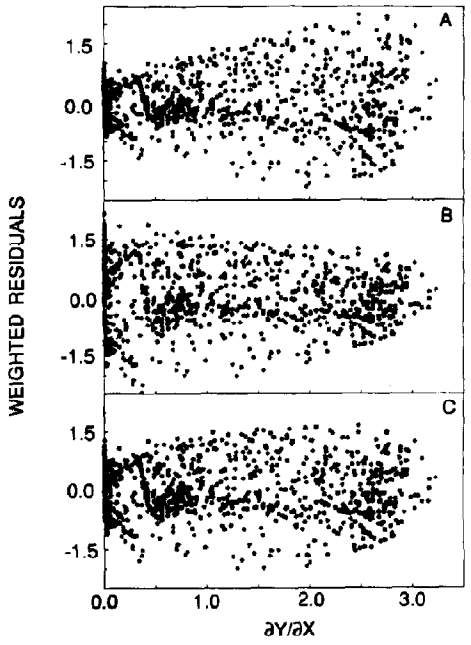

Fig. 2. A plot of the residuals, the differences between the fitted function and the original data points, as a function of the derivative of the fractional saturation with respect to the ligand concentration. These residuals are from a nonlinear least-squares parameter estimation on the pH 8.5 data set of Chu et al. [10]. Generated with (A) equal weighting of all the data points, (B) 5% weighting scheme using eq. 15, (C) 2% weighting scheme.

first term in eq. 15 originates from the approximately constant uncertainties due to σ_{A_i}. Thus, A is actually the contribution of σ_{X_i} relative to the contribution of σ_{A_i}. In each case, the sum of all of the weighting factors was normalized to be equal to the number of the data points, 884, so that the variances of fit, σ^2, can be directly compared.

Table 2 presents a calculation analogous to table 1 where the weighting functions were generated by eq. 15 with A being 5% of the maximal oxygen concentration. Fig. 2B presents the corresponding graph of residuals against $\partial Y/\partial X$. A comparison of fig. 2A and B indicates that a 5% weighting of the oxygen concentration uncertainties as compared to the fractional saturation uncertainties is more than enough to reverse the systematic trend in the residuals vs $\partial Y/\partial X$. A comparison of tables 1 and 2 indicates that including this small amount of end-point weighting due to the uncertainties in the oxygen electrode measurements has little or no effect on the resulting thermodynamic quantities.

Table 2

Analysis with 5% weighting

Values listed below are the intrinsic (corrected for statistical factors) free energy changes for the binding of successive oxygens to tetrameric hemoglobin A_0. These were determined as in table 1 except that a 5% weighting scheme was utilized.

	Most probable	−C.I.	+C.I.
$\Delta G'_{41}$	−6.015	−6.054	−5.982
$\Delta G'_{42}$	−6.864	−7.002	−6.728
$\Delta G'_{43}$	−7.539	−7.711	−7.356
$\Delta G'_{44}$	−9.426	−9.504	−9.352
Q.E.	−0.960	−1.038	−0.886
σ^2	6.633×10^{-5}		
Range of σ_{Y_i}	4.59		

Table 3 and fig. 2C show the corresponding information for a 2% weighting scheme. The 2% weighting yields a set of residuals with no dependence on the derivative of the fractional saturation with respect to the ligand concentration and less variation in the thermodynamic quantities. A 2% weighting scheme also yields the lowest variance.

It should be noted that the 2 and 5% weighting schemes generate a set of weights, σ_{Y_i}, which was end weighted with ratios of the maximum weight to the minimum weight to be 2.05 and 4.59, respectively. The $Y(1-Y)$ end weighting recom-

Table 3

Analysis with 2% weighting

Values shown are the intrinsic (corrected for statistical factors) free energy changes for the binding of successive oxygens to tetrameric hemoglobin A_0. These were determined as in table 1 except that a 2% weighting scheme was utilized.

	Most probable	−C.I.	+C.I.
$\Delta G'_{41}$	−6.035	−6.073	−6.002
$\Delta G'_{42}$	−6.814	−6.960	−6.667
$\Delta G'_{43}$	−7.553	−7.747	−7.349
$\Delta G'_{44}$	−9.443	−9.533	−9.356
Q.E.	−0.977	−1.067	−0.890
σ^2	5.139×10^{-5}		
Range of σ_{Y_i}	2.05		

mended by Imai and co-workers [1–5,17] would generate a ratio of 611.6 for this data set.

It should be noted that the above calculations do not mean that $Y(1 - Y)$ end weighting is wrong for all Imai oxygenation instruments. They only indicate that it is not the optimal weighting scheme for the Imai apparatus used to collect this set of saturation data. The best weighting scheme is so dependent upon the particular oxygen electrode, spectrophotometer, amplifier electronics, and experimental protocol that no general weighting can be applied to all Imai oxygenation apparatus implementations. However, it can be concluded that the best weighting for the Imai oxygenation apparatus in the laboratory of Ackers [6–16] is a very small amount of end weighting. Furthermore, the amount of end weighting is so small that even weighting, or no weighting at all, works equally well.

It is highly recommended that any, and every, laboratory which is conducting this type of experimentation should evaluate the best weighting function for their instrumentation by following a procedure like that outlined here.

Acknowledgements

This work was supported by the Biodynamics Institute of the University of Virginia, National Institutes of Health grants GM-28928 and DK-38942 and Office of Naval Research Grant N00014-88-K-0326.

References

1 K. Imai, H. Morimoto, M. Kotani, H. Watori, H. Waka and M. Kuroda, Biochim. Biophys. Acta 200 (1970) 189.
2 K. Imai, Biochemistry 12 (1973) 798.
3 K. Imai, J. Biol. Chem. (1974) 7607.
4 K. Imai and T. Yonetani, J. Biol. Chem. 250 (1975) 2227.
5 K. Imai and T. Yonetani, J. Biol. Chem. 250 (1975) 7903.
6 F.C. Mills, M.L. Johnson and G.K. Ackers, Biochemistry 15 (1976) 1093.
7 F.C. Mills and G.K. Ackers, Proc. Natl. Acad. Sci. U.S.A. 76 (1979) 273.
8 F.C. Mills and G.K. Ackers, J. Biol. Chem. 254 (1979) 2881.
9 D.H. Atha, M.L. Johnson and G.K. Ackers, J. Biol. Chem. 254 (1979) 12390.
10 A.H. Chu, B.W. Turner and G.K. Ackers, J. Biol. Chem. 256 (1981) 1199.
11 G.K. Ackers and M.L. Johnson, J. Mol. Biol. 147 (1981) 559.
12 M.L. Johnson and G.K. Ackers, Biochemistry 21 (1982) 201.
13 M.L. Johnson, B.W. Turner and G.K. Ackers, Proc. Natl. Acad. Sci. U.S.A. 81 (1984) 1093.
14 M.L. Johnson, Biochemistry 27 (1988) 833.
15 M. Straume and M.L. Johnson, Biophys. J. 56 (1989) 15.
16 M.L. Johnson and G.K. Ackers, Biophys. Chem. 37 (1990) 265.
17 D. Myers, K. Imai and T. Yonetani, Biophys. Chem. 37 (1990) 323.
18 M.L. Johnson and S.G. Frasier, Method Enzymol. 117 (1985) 301–342.
19 M.L. Johnson and G.K. Ackers, Biophys. Chem. 7 (1977) 77.
20 M.L. Johnson and A.E. Lassiter, Biophys. Chem. 37 (1990) 231.
21 G.K. Ackers and H.R. Halvorson, Proc. Natl. Acad. Sci. U.S.A. 71 (1974) 4312.
22 M.L. Johnson, Anal. Biochem. 148 (1985) 471.
23 F.S. Acton, Analysis of straight line data (Wiley, New York, 1959) p. 129.
24 Y. Bard, Nonlinear parameter estimation (Academic Press, New York, 1974) p. 67.