

mental conditions will show different degrees of shifting effectiveness. These results therefore, provide strong support for the model (Figure 1) previously proposed (Glassman *et al.*, 1971).

#### Acknowledgment

We thank M. A. Kelling for her assistance in the prelimi-

nary stages of this work.

#### References

- Glassman, T. A., Cooper, C., Harrison, L. W., and Swift, T. J. (1971), *Biochemistry* 10, 843.  
 Swift, T. J., and Connick, R. E. (1962), *J. Chem. Phys.* 37, 307.

## Mathematical Models for Interacting Groups in Nuclear Magnetic Resonance Titration Curves

Richard I. Shrager, Jack S. Cohen, Stephen R. Heller, David H. Sachs, and Alan N. Schechter

**ABSTRACT:** Mathematical models were developed for the interpretation of nuclear magnetic resonance titration curves of chemical shift *vs.* pH with emphasis on the case of two interacting titrating groups. The equations were applied by means of computer curve fitting to data obtained for the imidazole C-2 and C-4 proton resonances of L-histidine, *N*-acetyl-L-histidine, L-histidine methyl ester, and several L-histidine dipeptides. Inflections are observed in these titration curves due to the effects of the neighboring titrating amino and car-

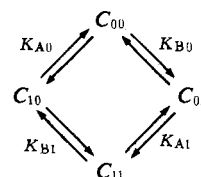
boxyl groups. The curve-fitting procedure provided accurate values for the ionization constants of the imidazole and the interacting titrating groups, even when the pK values were sufficiently close that the inflections had merged to produce an asymmetric curve. Application of these procedures to the titration curves of L-histidyl-L-histidine indicated that the asymmetry noted in the high pH region of one of the curves is due to the effect of an adjacent amino group and not to an interaction between the two imidazole rings.

Proton magnetic resonance spectroscopy has been used to determine the pK values of the imidazole rings of individual histidine residues in peptides and proteins (Bradbury and Scheraga, 1966; Roberts *et al.*, 1969; Cohen, 1969, 1971). The values obtained give information about the local environments of these residues, and may be used to study the effects of ligands and other perturbants (Meadows *et al.*, 1969; Ruterjans and Witzel, 1969). In several cases the titration curves of imidazole proton chemical shift as a function of pH have deviated from the theoretical curve describing a simple proton association equilibrium (Ruterjans and Witzel, 1969; Cohen *et al.*, 1970a; King and Bradbury, 1971; Ruterjans and Pongs, 1971). These deviations represent interactions between the imidazole residues and other charged groups within the molecule.

We present here a general mathematical treatment of such nuclear magnetic resonance (nmr) titration curves, using computer curve fitting for calculating pK values for interacting groups. These methods facilitate identification of interacting groups and thus may prove useful as a "probe" of the environment around histidine residues in proteins.

**Theoretical Section. A. GENERAL FORMULA FOR INTERACTING SITES.** We assume that two sites, called A and B, can inter-

act in two ways. First, the state of each site, *i.e.*, occupied or unoccupied by some ligand, may influence the equilibrium constant of the other site for that ligand. Second, the state of each site may influence the chemical shift of the other site. The overall state of the two-site systems may be described by four concentrations:  $C_{00}$  = concentration with both sites unoccupied;  $C_{10}$  = concentration with only site A occupied;  $C_{01}$  = concentration with only site B occupied;  $C_{11}$  = concentration with both sites occupied. This may be illustrated as follows (Edsall *et al.*, 1958),



Assuming hydrogen to be the sole occupant of the sites, we define the equilibrium constants as follows, where  $H$  is the concentration of hydrogen ions

$$\begin{aligned} K_{A0} &= C_{10}/C_{00}H, & K_{B0} &= C_{01}/C_{00}H, \\ K_{A1} &= C_{11}/C_{01}H, & K_{B1} &= C_{11}/C_{10}H \end{aligned} \quad (1)$$

from which it follows that

$$K_{A0}K_{B1} = K_{A1}K_{B0}$$

The equilibrium constants are identified by two subscripts:

\* From the Physical Sciences and Heuristics Laboratories, Division of Computer Research and Technology, and the Laboratory of Chemical Biology, National Institute of Arthritis and Metabolic Diseases, National Institutes of Health, Bethesda, Maryland 20014. Received October 7, 1971. A preliminary account of this work was given in Cohen *et al.* (1971). This is the second of a series of papers on the Nuclear Magnetic Resonance Titration Curves of Histidine Ring Protons; the paper by Sachs *et al.* (1971) is the first of this series.

TABLE I: Definition of Parameters in General Equation.

| Site A   | Site B   |
|--|--|
| $\delta_A = \frac{a_1 + a_2H + a_3H^2}{1 + a_4H + a_5H^2}$ | $\delta_B = \frac{a_6 + a_7H + a_8H^2}{1 + a_4H + a_5H^2}$ |
| $a_1 = \delta_{A00}$                                       | $a_6 = \delta_{B00}$                                       |
| $a_2 = K_{A0}\delta_{A10} + K_{B0}\delta_{A01}$            | $a_7 = K_{B0}\delta_{B01} + K_{A0}\delta_{B10}$            |
| $a_3 = K_{A1}K_{B0}\delta_{A11}$                           | $a_8 = K_{A1}K_{B0}\delta_{B11}$                           |
| $a_4 = K_{A0} + K_{B0}$                                    |  |
| $a_5 = K_{A1}K_{B0}$                                       |  |
| Directly resolvable:                                       |  |
| $\delta_{A00} = a_1$ $\delta_{A11} = a_3/a_5$              | $\delta_{B00} = a_6$ $\delta_{B11} = a_8/a_5$              |
| Remaining parameters                                       |  |
| $K_{A0}, K_{A1}, K_{B0}, \delta_{A01}, \delta_{A10}$       | $\delta_{B01}, \delta_{B10}$                               |
| Must be resolved from                                      |  |
| $a_2, a_4, a_5$  | $a_7$  |

the first referring to the site in question (A or B) and the second to the state of the other site (0 or 1). The chemical shift parameters ( $\delta$ 's) require three subscripts; the first referring to the site in question, the second to the state of site A, and the third to the state of site B. Thus  $\delta_{B01}$  is the  $\delta$  of site B when site A is unoccupied and site B is occupied. We assume that the observed  $\delta$  is a weighted average of chemical shift parameters, *i.e.*

$$\delta_A = (C_{00}\delta_{A00} + C_{01}\delta_{A01} + C_{10}\delta_{A10} + C_{11}\delta_{A11})/C \quad (2)$$

$$\delta_B = (C_{00}\delta_{B00} + C_{01}\delta_{B01} + C_{10}\delta_{B10} + C_{11}\delta_{B11})/C$$

where,  $C = C_{00} + C_{01} + C_{10} + C_{11}$ . This assumption is generally used, and appears valid,<sup>1</sup> for exchange processes which are fast compared to the nuclear magnetic resonance time scale, such as the proton association-dissociation equilibrium (Emsley *et al.*, 1965). Relation 1 may be reexpressed as

$$\begin{aligned} C_{10} &= K_{A0}C_{00}H & C_{01} &= K_{B0}C_{00}H \\ C_{11} &= K_{A1}C_{01}H = K_{A1}K_{B0}C_{00}H^2 \end{aligned} \quad (3)$$

Substituting 3 in 2 to eliminate  $C_{10}$ ,  $C_{01}$ , and  $C_{11}$ , the  $C_{00}$  factor cancels yielding the expressions in Table I.

In a formula of the general form

$$\frac{b_1 + b_2H + b_3H^2}{b_4 + b_5H + b_6H^2}$$

only five parameters are resolvable. Since the numerator and denominator can be multiplied by any factor without altering the value, we can arbitrarily factor  $b_4$ , yielding the unambiguous form

$$\frac{a_1 + a_2H + a_3H^2}{1 + a_4H + a_5H^2}$$

<sup>1</sup> The good fit of the titration data of imidazole to a simple proton association equilibrium (Sachs *et al.*, 1971) is an experimental verification of this assumption.

as in Table I. If the five  $a$ 's are, in turn, functions of many other parameters (more than five), then the same set of  $a$ 's can be generated by an infinite number of sets of those other parameters (functional dependence).

As shown in Table I,  $\delta_A$  and  $\delta_B$  fairly bristle with parameters. If  $\delta_A$  is considered alone, there are seven parameters of interest (four  $\delta$ 's and three  $K$ 's) in a model with only five resolvable coefficients ( $a_1$  through  $a_5$ ). Therefore, we must discover two additional relations between the parameters of interest, or impose two arbitrary relations to arrive at a unique set of numbers. Similarly, if  $\delta_A$  and  $\delta_B$  are considered together, there are eleven parameters of interest with only eight resolvable coefficients, obliging us to produce three additional relations. Mathematically, the only requirement is that these relations enable one to solve the system in Table I. When resolving the chemical shift parameters, one is always obliged to include either  $\delta_{A01}$  or  $\delta_{A10}$  in a relation since they both appear in the coefficient  $a_2$  and only in that coefficient. Thus, for example, one can constrain these intermediate  $\delta$  values by relating them to the maximum or minimum  $\delta$  values, *e. g.*,  $\delta_{A01} = 0.1(\delta_{A11} - \delta_{A00}) + \delta_{A00}$  providing the required additional relations. The physical justification for such constraints will be dealt with later.

It should be noted that the same general expression is obtained assuming a simple two-state equilibrium, but with a pH-dependent ionization constant.

B. SPECIAL CASES OF THE GENERAL EQUATION. When the equilibrium constant of one site is not significantly affected by the state of the other site two parameters may be eliminated as follows;  $K_{A0} = K_{A1}$ , which may be denoted  $K_A$ , and  $K_{B0} = K_{B1}$  ( $= K_B$ ). We define the nondimensional quantities  $\rho$  to represent the fractional distance between  $\delta_{A00}$  and  $\delta_{A11}$ ,

$$\rho_A = (\delta_A - \delta_{A00})/(\delta_{A11} - \delta_{A00})$$

$$\rho_{A10} = (\delta_{A10} - \delta_{A00})/(\delta_{A11} - \delta_{A00})$$

$$\rho_{A01} = (\delta_{A01} - \delta_{A00})/(\delta_{A11} - \delta_{A00})$$

which can also be expressed as a percentage.

From Table I, one gets

$$\begin{aligned} \rho_A &= \frac{(K_A\rho_{10} + K_B\rho_{01})H + K_AK_BH^2}{(1 + K_AH)(1 + K_BH)} \\ &= \frac{C_AK_AH}{1 + K_AH} + \frac{C_BK_BH}{1 + K_BH} \end{aligned}$$

where  $C_A = (K_B - K_A\rho_{10} - K_B\rho_{01})/(K_B - K_A)$ ,  $C_B = (K_A - K_A\rho_{10} - K_B\rho_{01})/(K_A - K_B)$ ,  $C_A + C_B = 1$ . This is the sum of two Henderson-Hasselbalch expressions (Cohen *et al.*, 1970a).

If, in addition to the restrictions  $K_{A0} = K_{A1}$  and  $K_{B0} = K_{B1}$ , the intermediate chemical shift values are equal to the maximum ( $\rho_{A10} = 100\%$ ) and minimum ( $\rho_{A01} = 0\%$ ) chemical shift values respectively, *i.e.*,  $\delta_{A01} = \delta_{A00}$  and  $\delta_{A10} = \delta_{A11}$ , then site A is no longer interacting with site B, and  $\delta_A$  will be a single Henderson-Hasselbalch expression (Cohen *et al.*, 1970b).

C. SUMMARY OF MODELS. For convenience the various forms of the equations described in sections A and B will be treated as different models (Table II). Examples of curves generated from models 1 and 2 with various choices of parameters are shown in Figures 1 and 2. These illustrate the effects on the shapes of the curves which result from changes in the inter-

TABLE II: Definition of Models Used for Curve Fitting.

| Model | Site    | Constraints of Parameters  | Description   |
|-------|---------|--|---|
| 1a    | A       | $\delta_{A01} = \delta_{A00}, \delta_{A10} = \delta_{A11}$<br>$K_{A0} = K_{A1} = K_A, K_{B0} = K_{B1} = K_B$   | Single simple equilibrium. $K_A$ and $K_B$ are macroscopic constants  |
| 1b    | A       | $K_{A0} = K_{A1} = K_A, K_{B0} = K_{B1} = K_B$   | Sum of simple equilibria. $K_A$ and $K_B$ are macroscopic constants   |
| 2     | A       | $\delta_{A01} = \delta_{A00}, \delta_{A10} = \delta_{A11}$   | Intermediate $\delta$ values ignored. $K$ 's are microscopic values   |
| 3     | A       | $\delta_{A01} = \delta_{A00} + \rho_{A01}(\delta_{A11} - \delta_{A00})$<br>$\delta_{A10} = \delta_{A00} + \rho_{A10}(\delta_{A11} - \delta_{A00})$   | $\rho_{A01}$ and $\rho_{A10}$ are fractional differences between $\delta_{A11}$ and $\delta_{A00}$ , which serve to define the intermediate $\delta$ values |
| 4     | A and B | $\delta_{A01} = \delta_{A00}, \delta_{A10} = \delta_{A11}$<br>$\delta_{B01} = \delta_{B00}, \delta_{B10} = \delta_{B11}$   | As in model 2, but assuming mutual interaction between two sites for both of which $\delta$ vs. pH curves are fitted simultaneously                         |
| 5     | A and B | $\delta_{A01} = \delta_{A00} + \rho_{A01}(\delta_{A11} - \delta_{A00})$<br>$\delta_{A10} = \delta_{A00} + \rho_{A10}(\delta_{A11} - \delta_{A00})$<br>$\delta_{B01} = \delta_{B00} + \rho_{B01}(\delta_{B11} - \delta_{B00})$<br>$\delta_{B10} = \delta_{B00} + \rho_{B10}(\delta_{B11} - \delta_{B00})$ | As in model 4 but with intermediate $\delta$ values defined as in model 3   |

mediate chemical shift values (Figure 1) or microscopic pK values (Figure 2).

In the mathematical procedures employed the equilibrium constants are defined as ionization constants and evaluated as pK values directly. When one pK value is quoted for the application of model 1a only one curve for a simple proton association equilibrium has been fitted to all the data; when two or more pK values are quoted for model 1b the sum of two or more such equilibria have been fitted.

D. RELATIONSHIPS OF MICROSCOPIC AND MACROSCOPIC IONIZATION CONSTANTS. The pK values derived in model 1 may be defined as apparent individual macroscopic ionization constants, and may be understood as values for discrete groups. The pK values derived in models 2 and 3 are defined as apparent microscopic ionization constants for the four-state mechanism defined in section A. These two kinds of ionization con-

stants are related as shown by Edsall *et al.* (1958), and can be expressed in terms of our mechanism as follows:  $K_A^{-1} = K_{A0}^{-1} + K_{B0}^{-1}$  and  $K_B = K_{A1} + K_{B1}$ .

#### Experimental Section

Experimental details and materials used have been described previously for data recorded both at 60 MHz (Sachs *et al.*, 1971) and at 220 MHz (Cohen *et al.*, 1970b). The experimental data upon which the computer fits of Figures 3–5 are based have been published (Sachs *et al.*, 1971). Chemical shift values are quoted downfield from external tetramethylsilane, and pH values are direct meter readings. L-Histidyl-L-histidine was obtained from Fox and DL-histidyl-DL-histidine from K & K Chemical Co. The MODELAIDE curve-fitting

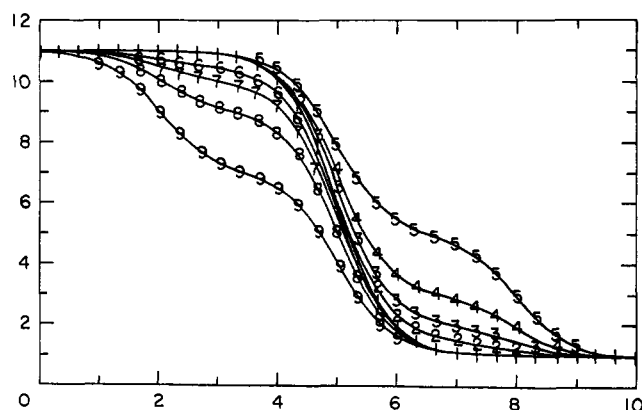


FIGURE 1: Curves generated by models 1a (curve 1) and 1b (curves 2–9) with increasing values of the intermediate chemical shift parameters  $\delta_{A01}$  (curves 2–5,  $\rho_{A01} = 5, 10, 20, 40\%$ , respectively) and  $\delta_{A10}$  (curves 6–9,  $100 - \rho_{A10} = 5, 10, 20, 40\%$ , respectively). The value of pK<sub>A</sub> is 5 throughout, while pK<sub>B</sub> is 8 in curves 2–5 and 2 in curves 6–9. The ordinate represents any physical or chemical parameter. The abscissa represents pH.

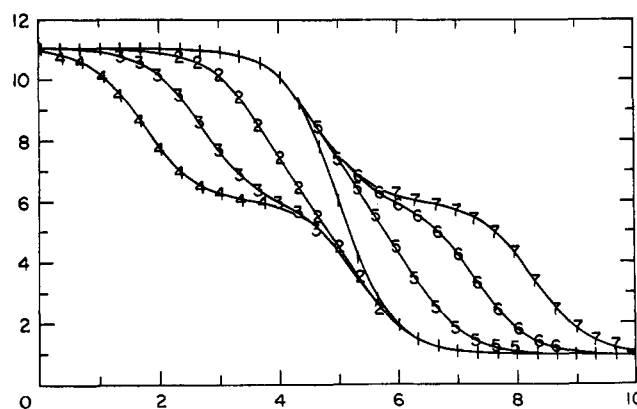


FIGURE 2: Curves generated by model 2 with the following microscopic pK values: curve 1, pK<sub>A00</sub> 5, pK<sub>A01</sub> 5, pK<sub>A10</sub> 5, pK<sub>A11</sub> 5; curve 2, pK<sub>A00</sub> 5, pK<sub>A01</sub> 4, pK<sub>A10</sub> 5, pK<sub>A11</sub> 4; curve 3, pK<sub>A00</sub> 5, pK<sub>A01</sub> 3, pK<sub>A10</sub> 5, pK<sub>A11</sub> 3; curve 4, pK<sub>A00</sub> 5, pK<sub>A01</sub> 2, pK<sub>A10</sub> 5, pK<sub>A11</sub> 2; curve 5, pK<sub>A00</sub> 6, pK<sub>A01</sub> 5, pK<sub>A10</sub> 6, pK<sub>A11</sub> 5; curve 6, pK<sub>A00</sub> 7, pK<sub>A01</sub> 5, pK<sub>A10</sub> 7, pK<sub>A11</sub> 5; curve 7, pK<sub>A00</sub> 8, pK<sub>A01</sub> 5, pK<sub>A10</sub> 8, pK<sub>A11</sub> 5. The chemical shift parameters  $\delta_{A01} = \delta_{A00} = 1$  and  $\delta_{A10} = \delta_{A11} = 11$  throughout. Ordinate and abscissa are as in Figure 1.

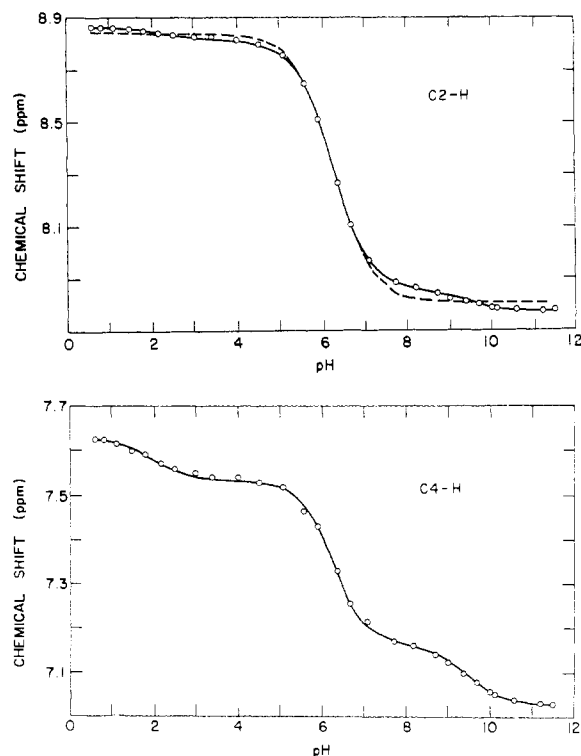


FIGURE 3: Chemical shift data as a function of pH (○) for the C-2 (upper half) and C-4 (lower half) proton magnetic resonances at 60 MHz of L-histidine. The dotted line in the upper figure is the best fit obtained assuming a single simple equilibrium (model 1a). The solid line in both cases is the fit for the sum of three such equilibria (model 1b with three terms).

program has also been described elsewhere (Shrager, 1970). Standard deviations of pK values derived by curve fitting are generally  $\pm 0.05$  pH unit.

## Results

The application of the above mathematical models to the analysis of imidazole proton nmr titration curves will be illustrated with the following examples.

**L-Histidine.** Computer fits of several models to the data for the titration of the imidazole C-2 and C-4 protons of L-histidine are shown in Figure 3. It is clear that a single simple equilibrium (model 1a) does not adequately represent either of these titrations. In fact the sum of three such transitions (model 1b) are required, giving a best fit with the following pK values: C2 - H; 1.98, 6.21, 9.24; C4 - H; 1.87, 6.24, 9.43. In view of the relatively small proportion of the carboxyl ( $100 - \rho_{A10} \approx 5\%$ ) and amino ( $\rho_{A01} \approx 5\%$ ) transitions in the total curve for the C-2 proton, the pK's derived from the C-4 proton curve (proportion  $\sim 15\%$ ) are probably the more reliable values.

**N-Acetyl-L-histidine.** Two inflections are seen in the C-2 and C-4 proton titration curves of N-acetyl-L-histidine (Figure 4) and model 1b gives an adequate representation of the data. In addition models 2 and 3 were also used. Several values of the chemical shift parameters  $\rho_{A01}$  and  $100 - \rho_{A10}$  (for the intermediates  $C_{01}$  and  $C_{10}$ ) were selected and held constant during a curve fit, to determine if the choice was critical in determining pK's. Ranges up to 10% for the C-2 H and 20% for the C-4 H were chosen. As will be seen from the values in Table III the sum of squares for the best fit remained

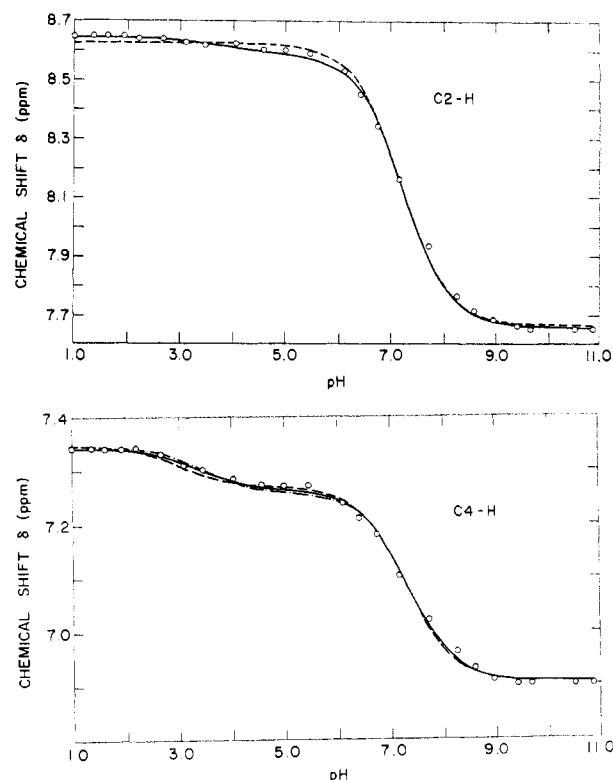


FIGURE 4: Chemical shift data as a function of pH (○) for the C-2 (upper) and C-4 (lower) proton magnetic resonances at 60 MHz of N-acetyl-L-histidine. The dashed line in the upper figure is the best fit obtained assuming a single simple equilibria. The solid line in both cases is the fit for the sum of two such equilibria (model 1b with two terms). The dashed and the dot-dashed line in the lower figure represent the fits obtained with model 3 ( $100 - \rho_{A10} = 10$  and  $20\%$ , respectively).

essentially constant while the pK values obtained varied widely. This illustrates that in the application of models 2 and 3 when the  $\Delta pK$  between the groups is greater than two, the microscopic pK values are very sensitive to the selected value of the intermediate chemical shift parameters. Nevertheless, the macroscopic pK values calculated from all applications of models 2 and 3 are very similar to those obtained directly from model 1b.

**L-Histidine Methyl Ester.** The titration data for histidine methyl ester is not satisfactorily represented by a single simple equilibrium (Figure 5). When the simplified equation (model 1a) is fitted only to the data below pH 4 and above pH 9 the asymmetry in the data becomes more obvious. This asymmetry does not appear as a clearly separate inflection. In order to obtain the pK's for the interacting groups models 2 and 3 were necessary (Table IV). Again, as for N-acetyl-L-histidine, several values of  $\rho_{A01}$  and  $100 - \rho_{A10}$  were imposed to test their effect on the resulting pK's. In this case changes over a range of 0-20% in the values of  $\rho_{A01}$  and  $100 - \rho_{A10}$  did not affect the quality of the fit and did not appreciably alter the values of the microscopic constants. Only at the highest values chosen did significant variations occur, but even these did not affect the values of the macroscopic pK values calculated therefrom. This would indicate that in such a case, where the two inflections are merged, reasonable values of the intermediate chemical shift parameters may be chosen in order to determine pK values. Values of 5% for the C-2 and 20% for the C-4 proton for  $100 - \rho_{A10}$ , which are reasonable choices as indicated by data for L-histidine (Figure 3), give internally

TABLE III: Ionization Constants of *N*-Acetyl-L-histidine.

|      | Model | $\rho_{A01}\%$ | 100 —<br>$\rho_{A10}\%$ | $pK_{A0}$ | $pK_{A1}$ | $pK_{B0}$ | $pK_{B1}$ | $pK_A$ | $pK_B$ | Sum of<br>Squares |
|------|-------|----------------|-------------------------|-----------|-----------|-----------|-----------|--------|--------|-------------------|
| C2-H | 1a    |                |                         |           |           |           |           | 7.18   |        | 0.010             |
|      | 1b    |                |                         |           |           |           |           | 7.24   | 3.79   | 0.0044            |
|      | 2     | 0              | 0                       | 7.22      | 5.05      | 6.00      | 3.83      | 7.25   | 3.81   | 0.0040            |
|      | 3     | 0              | 5                       | 7.23      | 6.49      | 4.47      | 3.72      | 7.23   | 3.72   | 0.0040            |
|      | 3     | 0              | 10                      | 7.31      | 10.73     | 1.91      | 5.33      | 7.31   | 3.73   | 0.0041            |
|      | 3     | 5              | 5                       | 7.23      | 5.86      | 5.07      | 3.69      | 7.23   | 3.69   | 0.0040            |
| C4-H | 1a    |                |                         |           |           |           |           | 7.08   |        | 0.014             |
|      | 1b    |                |                         |           |           |           |           | 7.29   | 3.39   | 0.0018            |
|      | 2     | 0              | 0                       | 7.21      | 4.14      | 6.55      | 3.47      | 7.30   | 3.39   | 0.0014            |
|      | 3     | 0              | 10                      | 7.26      | 4.45      | 6.23      | 3.43      | 7.30   | 3.39   | 0.0014            |
|      | 3     | 0              | 15                      | 7.28      | 4.88      | 5.81      | 3.40      | 7.29   | 3.39   | 0.0014            |
|      | 3     | 0              | 20                      | 7.32      | 7.96      | 2.79      | 3.41      | 7.32   | 3.41   | 0.0017            |

consistent microscopic  $pK$  values. Very similar results were obtained for L-histidylglycine (Table IV).

*L*-Histidyl-L-histidine. Two imidazole C-2 and C-4 proton resonances were observed for L-His-L-His.<sup>2</sup> The titration data for one of these (curve 2) are fitted adequately (Figure 6A) by the single simple equilibrium of model 1a (to exclude the carboxyl transition these data were not extended below pH 2.5). By contrast curve 1 is not fitted well by model 1a because of the inflection, or asymmetry, in the high pH region. Model 4 was applied to both curves simultaneously with the two imidazoles considered as sites A and B (Table II). A good fit to both sets of data was obtained (Figure 6B). However, if these two curves do represent mutually interacting imidazoles, then by consistency, fitting each separately with model 2 should produce microscopic  $pK$  values which are very similar to those obtained with model 4. The values obtained with each model are shown in Table V and the graphical representations of the fits are shown in Figure 6. In Figure 6C,D the maximum and minimum  $\delta$  values for the curve *not* being fitted were fixed, and the remainder of its shape was then determined by the  $pK$  values derived from fitting the second curve. The resultant discrepancies between the data and the theoretical curves show that the results are not well explained by mutual imidazole-imidazole interaction, according to this mechanism.

The  $pK$  (7.65) of the group (B) interacting with the imidazole giving rise to the asymmetry in curve 1 is clearly in the range expected for an amino group, as is the smaller effect seen in curve 2.

## Discussion

In several proteins the histidine nmr titration curves follow a simple proton association equilibrium, and may be described by a corresponding equation relating chemical shift to pH (Cohen *et al.*, 1970b).

The observation of asymmetric titration curves for two of

the histidine residues of ribonuclease (Ruterjans and Witzel 1969; Cohen *et al.*, 1970a; King and Bradbury, 1971), which clearly do not fit such a relationship, prompted us to reexamine histidine model systems. Bradbury and Scheraga (1966) previously reported data on histidine model compounds, but they analysed their results assuming a histidine-histidine interaction, and without regard for the effects of the titrations of the amino and carboxyl groups. However, we found in-

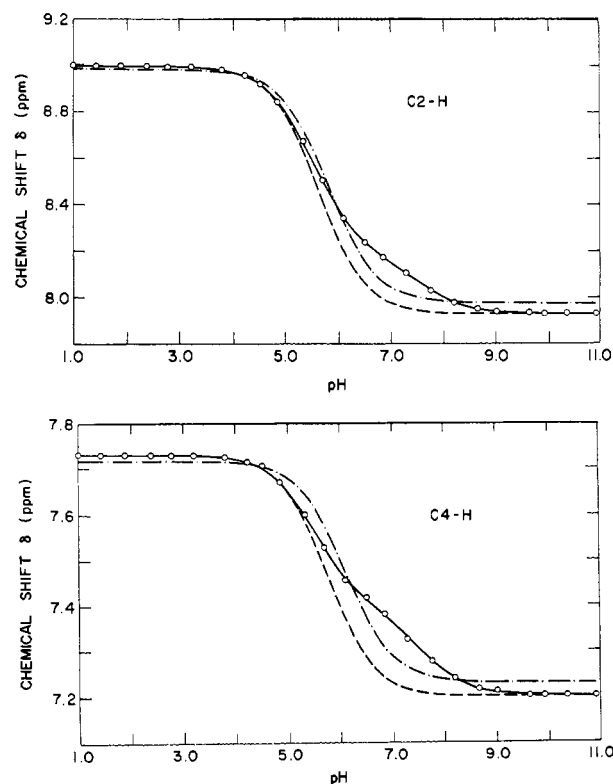


FIGURE 5: Chemical shift data as a function of pH (O) for the C-2 (upper) and C-4 (lower) proton magnetic resonances at 60 MHz of L-histidine methyl ester. The best fit obtained using a single simple equilibrium (model 1a) ignoring the data between pH 3 and 9 is represented by a dashed line, and that including all the data is represented by a dot-dashed line. The solid line is the best fit using models 2 or 3.

<sup>2</sup> Bradbury and Scheraga (1966) used the racemic mixture of DL-histidyl-DL-histidine and found several extra resonances. We observed the same phenomenon with a racemic mixture in which four clearly titrating resonances were resolved at 220 MHz. These could be separated into two pairs by their relative intensities. We were able to identify the smaller of our two sets as corresponding to L-histidyl-L-histidine by the superposition of these curves with the data from optically pure material.

TABLE IV: Ionization Constants of L-Histidine Methyl Ester.

|                 | Model | $\rho_{A01}$ % | 100 —<br>$\rho_{A10}$ % | $pK_{A0}$ | $pK_{A1}$ | $pK_{B0}$ | $pK_{B1}$ | $pK_A$ | $pK_B$ | Sum of Squares |
|-----------------|-------|----------------|-------------------------|-----------|-----------|-----------|-----------|--------|--------|----------------|
| C2-H            | 1a    |                |                         |           |           |           |           | 5.82   |        | 0.050          |
|                 | 2     | 0              | 0                       | 6.93      | 5.63      | 7.43      | 6.14      | 5.51   | 7.55   | 0.0002         |
|                 | 3     | 5              | 0                       | 6.84      | 5.61      | 7.45      | 6.22      | 5.51   | 7.54   | 0.0002         |
|                 | 3     | 10             | 0                       | 6.73      | 5.59      | 7.48      | 6.33      | 5.52   | 7.55   | 0.0002         |
|                 | 3     | 10             | 10                      | 6.78      | 5.60      | 7.47      | 6.28      | 5.52   | 7.55   | 0.0002         |
|                 | 3     | 5              | 5                       | 6.87      | 5.62      | 7.45      | 6.20      | 5.52   | 7.55   | 0.0002         |
|                 | 3     | 20             | 20                      | 6.34      | 5.54      | 7.52      | 6.72      | 5.51   | 7.55   | 0.0002         |
| C4-H            | 1a    |                |                         |           |           |           |           | 6.10   |        | 0.024          |
|                 | 2     | 0              | 0                       | 7.10      | 5.72      | 7.33      | 5.95      | 5.52   | 7.53   | 0.0001         |
|                 | 3     | 10             | 0                       | 7.01      | 5.68      | 7.37      | 6.06      | 5.52   | 7.53   | 0.0001         |
|                 | 3     | 20             | 0                       | 6.86      | 5.62      | 7.42      | 6.19      | 5.52   | 7.53   | 0.0001         |
|                 | 3     | 30             | 0                       | 6.53      | 5.57      | 7.48      | 6.52      | 5.52   | 7.53   | 0.0001         |
|                 | 3     | 20             | 20                      | 6.98      | 5.66      | 7.38      | 6.07      | 5.52   | 7.53   | 0.0001         |
| Histidylglycine |       |                |                         |           |           |           |           |        |        |                |
| C2-H            | 1a    |                |                         |           |           |           |           | 6.27   |        | 0.062          |
|                 | 2     | 0              | 0                       | 7.35      | 6.12      | 7.84      | 6.61      | 5.60   | 7.96   | 0.007          |
|                 | 3     | 10             | 0                       | 7.17      | 6.07      | 7.89      | 6.79      | 5.99   | 7.97   | 0.007          |
| C4-H            | 1a    |                |                         |           |           |           |           | 6.54   |        |                |
|                 | 3     | 10             | 0                       | 7.37      | 6.15      | 7.23      | 6.51      | 5.99   | 7.61   | 0.001          |

TABLE V: Ionization Constants of L-Histidyl-L-histidine.

| Curve | Model | $pK_{A0}$ | $pK_{A1}$ | $pK_{B0}$ | $pK_{B1}$ | $pK_A$ | $pK_B$ | Sum of Squares |
|-------|-------|-----------|-----------|-----------|-----------|--------|--------|----------------|
| 1     | 1a    |           |           |           |           | 6.25   |        | 0.095          |
| 2     | 1a    |           |           |           |           | 6.96   |        | 0.020          |
| 1     | 2     | 6.01      | 7.15      | 6.35      | 7.49      | 5.85   | 7.65   | 0.003          |
| 2     | 2     | 6.78      | 7.23      | 6.89      | 7.34      | 6.53   | 7.59   | 0.007          |
| 1 + 2 | 4     | 6.07      | 6.84      | 6.41      | 7.18      | 5.91   | 7.34   | 0.030          |

flections in nmr titration curves to be due to titration of the neighboring amino and carboxyl groups in L-histidine and N-acetyl-L-histidine, and we observed an asymmetric curve in the case of L-histidine methyl ester (Sachs *et al.*, 1971).

We give here the derivation of a general equation (Table I) for two interacting<sup>3</sup> sites, based upon the mechanism used by Edsall *et al.* (1958) for the spectrophotometric titration of interacting chromophores. Bradbury and Scheraga (1966) applied a form of this mechanism to the nmr titration data of histidylhistidine. A different analysis of interacting groups has been presented by Laskowski and Scheraga (1954).

When the  $pK$  values of the interacting groups are sufficiently separated ( $>2$  pH units), the general equation reduces to a special case corresponding to a sum of simple proton association equilibria (Cohen *et al.*, 1970a). This form of the equation (model 1b in Table II) requires a titration curve with

a clear-cut plateau to which the intermediate chemical shift value may be fitted (Figures 3 and 4). On the other hand, the use of models 2 and 3 for such a curve yields  $pK$  values which are very dependent on the chosen value of the intermediate chemical shift parameter (Table III). This is readily understood by the fact that a value for this parameter is implied by the plateau observable in the data, and if the chosen value of  $\delta_{A10}$  is different from this the values of the  $pK$ 's are altered by compensation in order to obtain a good fit.<sup>4</sup>

When the  $pK$  values of the interacting groups are closer ( $\leq 2$  pH units) such as for L-histidine methyl ester (Figure 5) an asymmetric titration curve results and there is no obvious intermediate chemical shift plateau. However, application of the complete equation (model 3) gives microscopic  $pK$  values which are not appreciably dependent on the chosen value of

<sup>3</sup> Interaction in the sense used here means competition between groups for the same protons as a result of proximity. This implies the direct influence of the electronic state of one group upon that of the other. This interaction could occur *via* hydrogen bonding, but the possibility is considered to be unlikely in the cases examined in this paper.

<sup>4</sup> It should be noted that the value of the sum of squares of differences between the observed and calculated points is not the only criterion for the nature of the fit. It is quite common to see long runs of positive errors followed by long runs of negative errors, even though the error level (sum of squares) seems reasonable. Thus inspection of the absolute errors, or more readily the graphical presentation, may indicate how the fit varies when the total sum of squares is a doubtful guide.

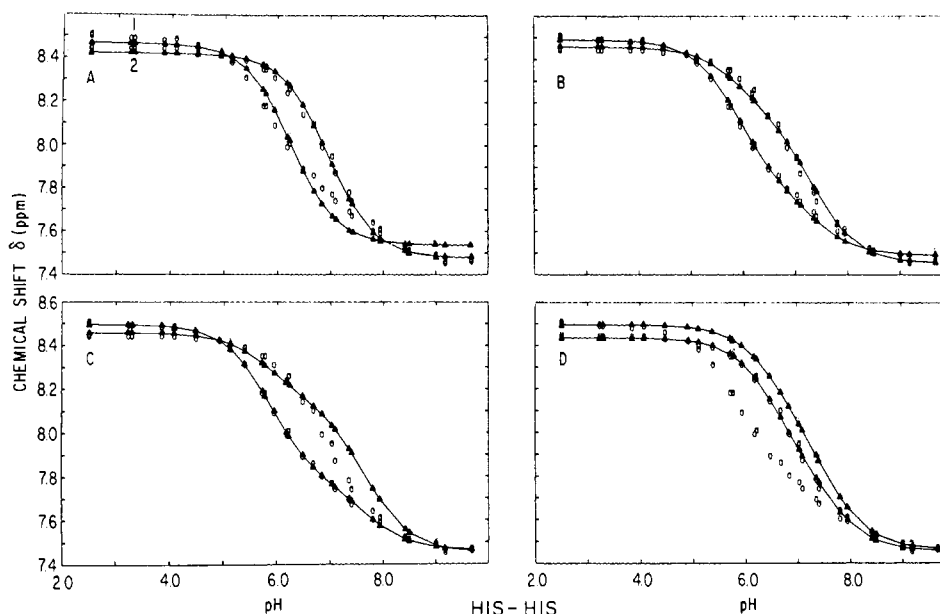


FIGURE 6: Chemical shift data as a function of pH (O) and the corresponding points calculated from models (▲) for the two C2 proton magnetic resonances of L-histidyl-L-histidine. A shows the best fits to each curve, taken separately, with a single simple equilibrium (model 1a). B shows the fit obtained from the two curves fitted simultaneously, assuming interaction between the two imidazole groups (model 4). C and D show the results derived from fitting curves 1 and 2, respectively, with model 2. In both C and D, the second theoretical curve shown is generated from the microscopic  $pK$  values derived from the fit of the first curve and from the maximum and minimum chemical shift values from the fits in B.

the intermediate chemical shift parameter. Furthermore, the  $pK$  value for the interacting group thus obtained is in good agreement with the accepted value for the  $pK$  of the amino group (Sachs *et al.*, 1971). This result is important with respect to the applicability of this method to the interpretation of asymmetries in protein nmr titration curves. It shows that an accurate solution is possible for an equation consisting of two quadratic expressions containing seven parameters, by the choice of a range of values for the two dependent parameters  $\delta_{A01}$  and  $\delta_{A10}$ .

It has been suggested that two imidazole groups in close juxtaposition to each other in ribonuclease may form a common hydrogen bond (Ruterjans and Witzel, 1969; King and Bradbury, 1971). Our analysis of the titration data for L-histidyl-L-histidine indicates no mutual imidazole-imidazole interaction in this case (Figure 6). Rather the  $pK$  value (7.7) of the interacting group giving rise to the asymmetry in this case is that expected for an amino group. It is therefore concluded from the similarities between L-histidine methyl ester, L-histidylglycine, and curve 1 of L-histidyl-L-histidine that this curve can be assigned to the amino-terminal residue in L-histidyl-L-histidine, and that an asymmetric nmr titration curve is characteristic of amino-terminal histidine residues.

In a further paper we will extend this method of analysis to the nmr titration curves of the histidine residues of ribonuclease where the imidazole group acts as a nondisturbing probe of the local environment. We are also applying these curve-fitting methods to data from nmr studies of other synthetic histidine-containing peptides.

#### Acknowledgments

We thank Professor J. Edsall for helpful discussions.

#### References

- Bradbury, J. H., and Scheraga, H. A. (1966), *J. Amer. Chem. Soc.* 88, 4240.
- Cohen, J. S. (1969), *Nature (London)* 223, 43.
- Cohen, J. S. (1971), *Biochim. Biophys. Acta* 229, 603.
- Cohen, J. S., Schechter, A. N., Shrager, R., Sachs, D. H., and Heller, S. R. (1971), *Fed. Proc., Fed. Amer. Soc. Exp. Biol.* 30, 1293.
- Cohen, J. S., Shrager, R. I., McNeel, M., and Schechter, A. N. (1970a), *Biochem. Biophys. Res. Commun.* 40, 144.
- Cohen, J. S., Shrager, R. I., McNeel, M., and Schechter, A. N. (1970b), *Nature (London)* 228, 642.
- Edsall, J. T., Martin, R. B., and Hollingworth, B. R. (1958), *Proc. Nat. Acad. Sci. U. S. A.* 44, 505.
- Emsley, J. W., Feeney, J., and Sutcliffe, L. H. (1965), *High Resolution Nuclear Magnetic Resonance*, Vol. I, New York, N. Y., Pergamon.
- King, N. L. R., and Bradbury, J. H. (1971), *Nature (London)* 229, 404.
- Laskowski, M., and Scheraga, H. A. (1954), *J. Amer. Chem. Soc.* 76, 6305.
- Meadows, D. H., Roberts, G. C. K., and Jardetzky, O. (1969), *J. Mol. Biol.* 45, 491.
- Roberts, G. C. K., Meadows, D. H., and Jardetzky, O. (1969), *Biochemistry* 8, 2053.
- Ruterjans, H., and Pongs, O. (1971), *Eur. J. Biochem.* 18, 313.
- Ruterjans, H., and Witzel, H. (1969), *Eur. J. Biochem.* 9, 118.
- Sachs, D. H., Schechter, A. N., and Cohen, J. S. (1971), *J. Biol. Chem.* 246, 6576.
- Shrager, R. (1970), *J. Ass. Comp. Mach.* 17, 446.