Relations between Protonation Constants and Titration Curves in Polyprotic Acids: A Critical View

G. Matthias Ullmann[‡]

[‡] IWR - Biocomputing; University Heidelberg; Im Neuenheimer Feld 368; 69120 Heidelberg, Germany; email: Matthias.Ullmann@iwr.uni-heidelberg.de; Fax: +49-6221-54-8868

Supporting information

1. Inflection Points of Titration Curves

Monoprotic Acid. The titration curve of a monoprotic acid is given by eq 1 which is algebraically equivalent to the Henderson-Hasselbalch equation.

$$\langle x \rangle = \frac{10^{\mathrm{p}K_a-\mathrm{pH}}}{1+10^{\mathrm{p}K_a-\mathrm{pH}}} \tag{1}$$

The first and second derivative of the titration curve in eq 1 are given by eq 2 and 3, respectively.

$$\frac{\partial \langle x \rangle}{\partial \mathrm{pH}} = \frac{a10^{\mathrm{p}K_a-\mathrm{pH}}}{(1+10^{\mathrm{p}K_a-\mathrm{pH}})^2} \tag{2}$$

$$\frac{\partial^2 \langle x \rangle}{\partial \mathrm{pH}^2} = \frac{(\ln 10)^2 10^{\mathrm{p}K_a-\mathrm{pH}}(1-10^{\mathrm{p}K_a-\mathrm{pH}})}{(1+10^{\mathrm{p}K_a-\mathrm{pH}})^3} \tag{3}$$

At the inflection point, the second derivative of eq 1, i.e. eq 3, becomes zero, which is exactly the case when $\mathrm{p}K_a$ value equals the pH value. For a monoprotic acid, the $\mathrm{p}K_a$ value calculated from the partition function, from the protonation probabilities, and the inflection point, and the $\mathrm{p}K_{1/2}$ value all coincide with the $\mathrm{p}K_a$ value of the protonatable group.
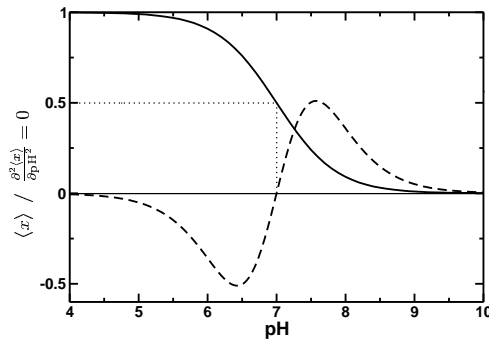

Figure 1: Titration curve of a monoprotic acid with $\mathrm{p}K_a$=7.0 (solid line) and the second derivative of this titration curve (dashed line). The $\mathrm{p}K_a$ value of the monoprotic acid which is 7.0 coincides with the inflection point of the titration curve and the pH at which the protonation probability is 0.5.
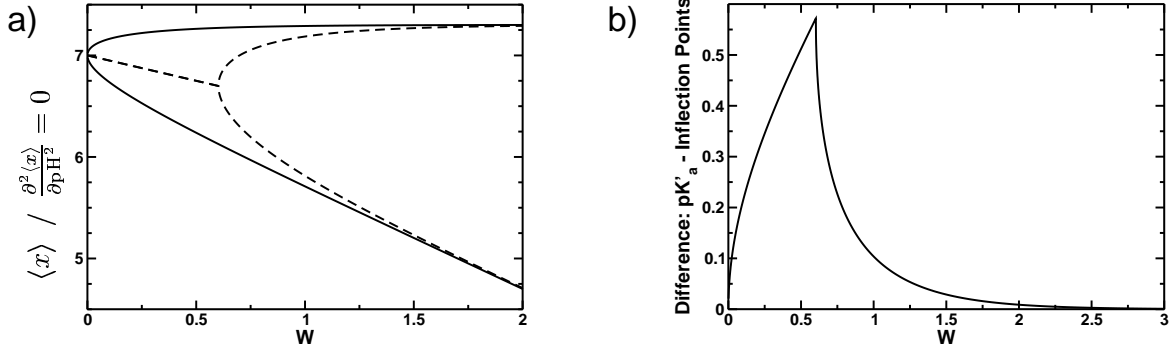
Figure 2: a) Comparison of the real part of the inflection points (dashed lines) of the titration curves and the quasi-site pK_j' values (solid lines) of a system of two interacting identical sites having the microscopic binding constants $pK_{00}^{10}=pK_{00}^{01}=7.0$ in dependence on the interaction W between the two sites. b) Differences between the inflection points of the titration curve and the quasi-site pK_j' values of this diprotic acid.

Diprotic Acid. The total protonation $\langle X \rangle$ of a molecule is given by the sum of the probability of each state multiplied by the number of protons bound to it: $\langle X \rangle = \langle (10) \rangle + \langle (01) \rangle + 2 \langle (11) \rangle$. For the sake of simplicity, we define $a = 10^{pK_{00}^{10}}$, $b = 10^{pK_{00}^{01}}$, $c = 10^{pK_{00}^{10}+pK_{00}^{01}-W}$, and $\lambda = 10^{-pH}$ and obtain eq 4.

$$\langle X \rangle = \frac{(a+b)\lambda + 2c\lambda^2}{1 + (a+b)\lambda + c\lambda^2} \tag{4}$$

This curve would be measured by techniques that look at the system as a whole such as potentiometry and not at individual chemical groups in the molecule.

The total titration curve of a molecule with two interacting sites has three inflection points, which are given by eq 5. Some of these inflection points may be complex even though the two protons repel each other.

$$Y_o = \frac{1}{2} \left(pK_{00}^{10} + pK_{00}^{01} - W \right) \tag{5}$$

$$Y_{1,2} = \lg \left(\frac{(a+b)^2 - 8c \pm \sqrt{(a+b)^4 - 20(a+b)^2c + 64c^2}}{2(a+b)c} \right)$$

The inflection points do not correspond to the macroscopic $p\overline{K}_k$ or the quasi-site pK_j' values of the system. Only for large W, the inflection points Y_1 and Y_2 and the macroscopic $p\overline{K}_k$ or the

3

quasi-site pK'_j values of the system coincide.

For large W or ΔpK, the inflection points $Y_{1,2}$ from eq 5 give good estimates for the quasi-site pK'_i values. $Y_{1,2}$ in eq 5 gets complex when the two branches in Figure 2a merge. The titration curve has then only a single real inflection point Y_o, which is identical with the real parts of $Y_{1,2}$. Figure 2 shows the deviation of the inflection points from the quasi-site pK'_a values for two identical sites in dependence on the interaction between the two sites. Since protons should repel each other, we consider only positive and not negative interaction, which would correspond to attraction. The deviation between the inflection points from the macroscopic pK'_a values is the biggest, when the interaction between the sites is small.

The equations for the inflection points of the individual titration curves are very complicated and would fill several pages. In general one can say that each of the individual titration curves has four inflection points of which some may be complex. The inflection points of the total titration curves and of the individual titration curves approach each other only for large interaction energies.

Polyprotic Acid. From the decoupled sites representation, it is known that the total titration curve of a molecule that can bind N protons can be written as a sum of N titration curves of non-interacting quasi-sites each having a pK'_j value which is related to the free energy of binding a proton to this this quasi-site by $\Delta G_j^{o'} = -\beta^{-1} \ln 10\, pK'_j$.

$$\langle X \rangle = \sum_{j=1}^{N} \frac{e^{-\beta(\Delta G_j^{o'} - \mu_{H+})}}{1 + e^{-\beta(\Delta G_j^{o'} - \mu_{H+})}} \tag{6}$$

The first and the second derivative of the total titration curve become therefore:

$$\frac{\partial \langle X \rangle}{\partial \mu_{H+}} = \sum_{j=1}^{N} \frac{\partial \langle y_j \rangle}{\partial \mu_{H+}} = \sum_{j=1}^{N} \frac{\beta e^{-\beta(\Delta G_j^{o'} - \mu_{H+})}}{(1 + e^{-\beta(\Delta G_j^{o'} - \mu_{H+})})^2} \tag{7}$$

$$\frac{\partial^2 \langle X \rangle}{\partial \mu_L^{2}} = \sum_{j=1}^{N} \frac{\partial^2 \langle y_j \rangle}{\partial \mu_{H+}^{2}} = \sum_{j=1}^{N} \frac{\beta^2 e^{-\beta(\Delta G_j^{o'} - \mu_{H+})}(1 - e^{-\beta(\Delta G_j^{o'} - \mu_{H+})})}{(1 + e^{-\beta(\Delta G_j^{o'} - \mu_{H+})})^3} \tag{8}$$

$$= \frac{\sum_{j=1}^{N} \left(\beta^2 e^{-\beta(\Delta G_j^{o'} - \mu_{H+})}(1 - e^{-\beta(\Delta G_j^{o'} - \mu_{H+})}) \prod_{s=1, s \neq j}^{N} (1 + e^{-\beta(\Delta G_s^{o'} - \mu_{H+})})^3 \right)}{\prod_{j=1}^{N} (1 + e^{-\beta(\Delta G_j^{o'} - \mu_{H+})})^3} \tag{9}$$

The titration curves of individual sites are a linear combination of the titration curves of the

quasi-sites.

$$\langle x_i \rangle = \sum_{j=1}^{N} a_{ij} \frac{e^{-\beta(\Delta G_j^{\circ\prime} - \mu_{\mathrm{H}+})}}{1 + e^{-\beta(\Delta G_j^{\circ\prime} - \mu_{\mathrm{H}+})}} \tag{10}$$

The first and the second derivative of the titration curves of the individual sites are:

$$\frac{\partial \langle x_i \rangle}{\partial \mu_{\mathrm{H}+}} = \sum_{j=1}^{N} a_{ij} \frac{\partial \langle y_j \rangle}{\partial \mu_{\mathrm{H}+}} \quad = \quad \sum_{j=1}^{N} a_{ij} \frac{\beta e^{-\beta(\Delta G_j^{\circ\prime} - \mu_{\mathrm{H}+})}}{(1 + e^{-\beta(\Delta G_j^{\circ\prime} - \mu_{\mathrm{H}+})})^2} \tag{11}$$

$$\frac{\partial^2 \langle x_i \rangle}{\partial \mu_{\mathrm{H}+}^{\ 2}} = \sum_{j=1}^{N} a_{ij} \frac{\partial^2 \langle y_j \rangle}{\partial \mu_{\mathrm{H}+}^{\ 2}} \quad = \quad \sum_{j=1}^{N} a_{ij} \frac{\beta^2 e^{-\beta(\Delta G_j^{\circ\prime} - \mu_{\mathrm{H}+})}(1 - e^{-\beta(\Delta G_j^{\circ\prime} - \mu_{\mathrm{H}+})})}{(1 + e^{-\beta(\Delta G_j^{\circ\prime} - \mu_{\mathrm{H}+})})^3} \tag{12}$$

From the equations and from the graphical representation of the equations in the main text one can see that the inflection points of total and individual titration curves can not be identified with pK_a values in the case of polyprotic acids.

How to Get Microscopic Constants from the Decoupled Sites Representation

The quasi-site pK'_j values can be obtained from the root of the partition function which can be written in terms of macroscopic constants. The quasi-site pK'_j values can then be used to obtain the coefficients a_{ij} of the linear combination by fitting the titration curves of individual sites. The constraints that the sum over all rows and over all columns must be one can be used as a constraint in the fit, for instance by using Lagrange multipliers. The obtained pK'_j values and a_{ij} coefficients can then be used to calculate the microscopic constants for di- and triprotic acids. In the case of polyprotic acids with more then three sites, either symmetry or special assumptions are required to obtain the microscopic constants from individual titration curves.

Diprotic Acid. From eqs 17 to 19 in the main text follows that

$$
\begin{aligned}
10^{pK_{00}^{10}} &= a_{11}10^{pK'_1} + a_{12}10^{pK'_2} \\
10^{pK_{00}^{01}} &= a_{21}10^{pK'_1} + a_{22}10^{pK'_2} \\
10^{pK_{00}^{10}+pK_{10}^{11}} &= a_{11}10^{pK'_1+pK'_2} + a_{12}10^{pK'_1+pK'_2} \\
10^{pK_{00}^{01}+pK_{01}^{11}} &= a_{21}10^{pK'_1+pK'_2} + a_{22}10^{pK'_1+pK'_2}
\end{aligned}
\tag{13}
$$

From this equation, one can derive eq 21 in the main text that give the microscopic constants in terms of quasi-site constants and the linear coefficients a_{ij} of the DSR. One should note that the there is only one free linear coefficients, the other coefficient are all defined because of constraints that the sum over all row and over all columns must be 1.

Triprotic Acid. From eqs 17 to 19 in the main text and by setting the free energy of the totally deprotonated state to zero follows that

$$
\begin{aligned}
e^{-\beta G_{000}^{O}} &= 1 \tag{14}\\
e^{-\beta G_{100}^{O}} &= a_{11}10^{pK'_1} + a_{12}10^{pK'_2} + a_{13}10^{pK'_3} \tag{15}\\
e^{-\beta G_{010}^{O}} &= a_{21}10^{pK'_1} + a_{22}10^{pK'_2} + a_{23}10^{pK'_3} \tag{16}\\
e^{-\beta G_{001}^{O}} &= a_{31}10^{pK'_1} + a_{32}10^{pK'_2} + a_{33}10^{pK'_3} \tag{17}\\
e^{-\beta G_{110}^{O}} + e^{-\beta G_{101}^{O}} &= a_{11}(10^{pK'_1+pK'_2}+10^{pK'_1+pK'_3}) + a_{12}(10^{pK'_1+pK'_2}+10^{pK'_2+pK'_3}) + a_{13}(10^{pK'_2+pK'_3}+10^{pK'_1+pK'_3}) \tag{18}\\
e^{-\beta G_{110}^{O}} + e^{-\beta G_{011}^{O}} &= a_{21}(10^{pK'_1+pK'_2}+10^{pK'_1+pK'_3}) + a_{22}(10^{pK'_1+pK'_2}+10^{pK'_2+pK'_3}) + a_{23}(10^{pK'_2+pK'_3}+10^{pK'_1+pK'_3}) \tag{19}\\
e^{-\beta G_{101}^{O}} + e^{-\beta G_{011}^{O}} &= a_{31}(10^{pK'_1+pK'_2}+10^{pK'_1+pK'_3}) + a_{32}(10^{pK'_1+pK'_2}+10^{pK'_2+pK'_3}) + a_{33}(10^{pK'_2+pK'_3}+10^{pK'_1+pK'_3}) \tag{20}\\
e^{-\beta G_{111}^{O}} &= a_{31}10^{pK'_1+pK'_2+pK'_3} + a_{32}10^{pK'_1+pK'_2+pK'_3} + a_{33}10^{pK'_1+pK'_2+pK'_3} \tag{21}
\end{aligned}
$$

6

These eight equations are sufficient to calculate the energy of all the eight possible protonation states and thus of the difference between the energies of the protonation states. The energy of the totally deprotonated state is set to zero (eq 14). The energy of the single protonated and the totally protonated states can be directly obtained by taking the negative of the logarithm of eqs 15, 16, 17, and 21 and dividing this number by β. To obtain the energies of the doubly-protonated states, we define eqs 22-24.

$$A = a_{11}(10^{\mathrm{p}K'_1+\mathrm{p}K'_2} + 10^{\mathrm{p}K'_1+\mathrm{p}K'_3}) + a_{12}(10^{\mathrm{p}K'_1+\mathrm{p}K'_2} + 10^{\mathrm{p}K'_2+\mathrm{p}K'_3}) + a_{13}(10^{\mathrm{p}K'_2+\mathrm{p}K'_3} + 10^{\mathrm{p}K'_1+\mathrm{p}K'_3}) \tag{22}$$

$$B = a_{21}(10^{\mathrm{p}K'_1+\mathrm{p}K'_2} + 10^{\mathrm{p}K'_1+\mathrm{p}K'_3}) + a_{22}(10^{\mathrm{p}K'_1+\mathrm{p}K'_2} + 10^{\mathrm{p}K'_2+\mathrm{p}K'_3}) + a_{23}(10^{\mathrm{p}K'_2+\mathrm{p}K'_3} + 10^{\mathrm{p}K'_1+\mathrm{p}K'_3}) \tag{23}$$

$$C = a_{31}(10^{\mathrm{p}K'_1+\mathrm{p}K'_2} + 10^{\mathrm{p}K'_1+\mathrm{p}K'_3}) + a_{32}(10^{\mathrm{p}K'_1+\mathrm{p}K'_2} + 10^{\mathrm{p}K'_2+\mathrm{p}K'_3}) + a_{33}(10^{\mathrm{p}K'_2+\mathrm{p}K'_3} + 10^{\mathrm{p}K'_1+\mathrm{p}K'_3}) \tag{24}$$

From these equations, we finally obtain the energies of the doubly-protonated states by taking the negative of the logarithm of eqs 25, 26, and 27.

$$e^{-\beta G^{\mathrm{o}}_{110}} = \frac{1}{2}(A + B - C) \tag{25}$$

$$e^{-\beta G^{\mathrm{o}}_{101}} = \frac{1}{2}(A - B + C) \tag{26}$$

$$e^{-\beta G^{\mathrm{o}}_{011}} = \frac{1}{2}(-A + B + C) \tag{27}$$

The microscopic $\mathrm{p}K^{\mathbf{P}}_{\mathbf{r}}$ are given by

$$\mathrm{p}K^{\mathbf{P}}_{\mathbf{r}} = -\beta \ln 10 \, (G^{\mathrm{o}}_{\mathbf{p}} - G^{\mathrm{o}}_{\mathbf{r}}) \tag{28}$$

where the superscripts **p** marks the product state and the subscript **r** marks the reactant state.