

Deriving the Isoelectric Point of Amino Acids from Acid-Base Equilibria

Gayan Senavirathne^{1,*}

¹Brooklyn, NY 11218

*Correspondence may also be addressed to: G. Senavirathne (senavirathne.83@gmail.com)

Abstract

This paper presents a refined method for calculating the isoelectric point (pI) of amino acids, focusing on the dominant zwitterionic form in solution. The approach improves pI calculations and extends to predicting pH values in amphoteric solutions, including titrations of polyprotic acids and bases.

Introduction

The isoelectric point (pI) of a molecule, such as an amino acid or protein, is the pH at which the molecule carries no net electrical charge (1-7). At this pH, the molecule exhibits no net movement in an electric field, a principle that underlies isoelectric focusing electrophoresis (1-3). For a simple amino acid like glycine, which contains two ionizable groups (α -COOH with pK_{a1} , and α -NH $_3^+$ with pK_{a2} , **Figure 1**), the predominant species at the pI is the amphoteric zwitterion (HA). This zwitterion is in equilibrium with small amounts of its conjugate acid (H $_2$ A $^+$) and conjugate base (A $^-$). At the pI, the H $_2$ A $^+$ and A $^-$ concentrations are equal, resulting in a zero net charge across all glycine species (4-6). Most pI estimates based on pK_a values are derived by averaging the corresponding pK_a values (e.g., for glycine, $pI = (pK_{a1} + pK_{a2})/2$), assuming the zwitterion is positioned between the two ionization equilibria (1-3, 7). However, previous derivations often overlook essential details regarding the underlying assumptions (4-6). The method presented here offers a more precise derivation of pI values using acid-base equilibria with additional considerations. This approach can be extended to more complex amino acids containing acidic or basic side chains and to determine the pH of solutions containing amphoteric species.

Derivation of the Isoelectric Point (pI) for Glycine

Before deriving the pI of glycine, consider the following reaction scheme: glycine undergoes two deprotonation steps as the solution pH transitions from acidic (low) to basic (high) conditions

(**Figure 1**). At very low pH, the α -NH₂ and α -COO⁻ groups are protonated, resulting in an overall +1 charge on the molecule. For simplicity, this protonated form is denoted as H₂A⁺. As the pH increases, the carboxyl group, which has the lower pK_a (higher acidity), begins to lose its proton, establishing the first equilibrium between the H₂A⁺ and HA forms. According to the Henderson-Hasselbalch equation, when the pH equals pK_{a1} the H₂A⁺ and HA concentrations are equal. As the pH continues to rise, the HA form becomes predominant. At the pI, virtually all glycine exists as the zwitterionic form (HA), where the molecule has no net charge.

As the pH moves further into the basic range, the α -NH₃⁺ group begins to lose its proton, forming the A⁻ species. The HA and A⁻ concentrations are equal at pH = pK_{a2}. Beyond this point, the A⁻ form predominates, and at a very high pH, the molecule exists almost entirely in the A⁻ form.

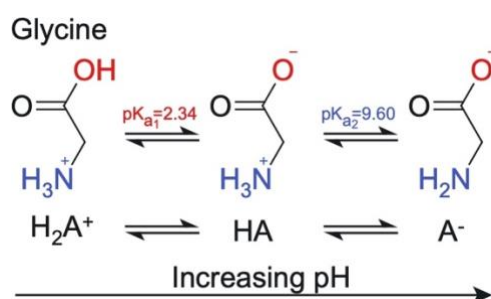
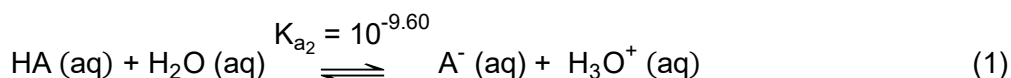


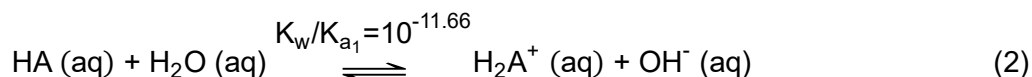
Figure 1. Sequential deprotonation of glycine as the solution pH transitions from acidic to basic. The acidic groups and their corresponding pK_a values are shown in the same color, with each pK_a value dictating the extent of the respective deprotonation step. The different forms of glycine—H₂A⁺, HA, and A⁻—are depicted below, with water (H₂O) and hydronium ions (H₃O⁺) omitted for clarity to simplify the derivation.

Consider the possible set of interconnected acid-base equilibria that may occur in solution at pI. The corresponding equilibrium constants (K_{eq}) for the forward reactions are also provided (3-5,7).

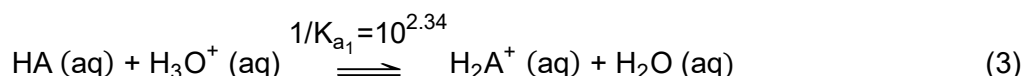
First equilibrium: HA can act as an acid, transferring a proton to water:



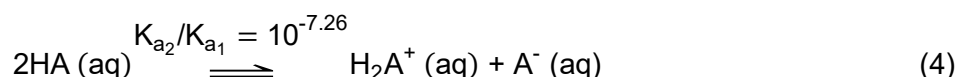
Second equilibrium: HA can also act as a base, hydrolyzing water. For the derivation of the equilibrium constant, see Appendix B:



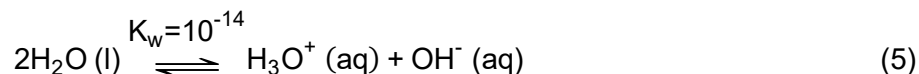
Third equilibrium: Instead of reacting with H_2O , HA can also react with the H_3O^+ produced by the first equilibrium:



Fourth equilibrium: HA can undergo autoprotolysis. For the derivation of the equilibrium constant, refer to Appendix C:



Fifth equilibrium: Autoprotolysis of water:



The question of which equilibria will occur in the solution to determine the pI depends on thermodynamic and kinetic feasibility. Based on these factors, the autoprotolysis of HA (the fourth equilibrium) is likely to dominate, followed by the acidic and basic reactions of HA (the first and second equilibria, respectively). The third equilibrium, which involves the interaction between HA and H_3O^+ , is the least likely to occur and can be excluded from further consideration (see Appendix D). The fifth equilibrium, governing the autoionization of water, operates in any aqueous solution by default, ensuring that the product of the concentrations of H_3O^+ and OH^- remains constant at a given temperature (e.g., 10^{-14} at 25°C) (3,5). However, it does not directly affect the pI of glycine.

The equilibrium constant for the most prevailing reaction from equation (4) can be written as:

$$\frac{K_{a_2}}{K_{a_1}} = \frac{[H_2A^+(aq)][A^-(aq)]}{[HA(aq)]^2} \quad (6)$$

At the pI, the H_2A^+ and A^- concentrations are equal (Appendix D). That is:

$$[H_2A^+(aq)] = [A^-(aq)] \quad (7)$$

Substituting this relationship into equation (6) results in:

$$\frac{K_{a_2}}{K_{a_1}} = \frac{[A^-(aq)]^2}{[HA(aq)]^2} = \frac{[H_2A^+(aq)]^2}{[HA(aq)]^2} \quad (8)$$

This simplifies to:

$$\frac{[A^-(aq)]}{[HA(aq)]} = \frac{[H_2A^+(aq)]}{[HA(aq)]} = \sqrt{\frac{K_{a_2}}{K_{a_1}}} \quad (9)$$

The fourth equilibrium does not directly consume or produce H_3O^+ or OH^- . However, its components—HA, H_2A^+ , and A^- —participate in the first and second equilibria, which adjust the concentrations of H_3O^+ and OH^- , respectively. At the pI, the relationship in equation (7) holds, ensuring that the H_2A^+ and A^- concentrations are equal, thereby influencing the overall acid-base balance. We can now use equation (9) to determine how the autoprotolysis of HA influences H_3O^+ in the first equilibrium or OH^- in the second equilibrium, and thus the pH of the solution at the isoelectric point.

From equation (1), the equilibrium constant for the second equilibrium can be written as:

$$K_{a_2} = \frac{[A^-(aq)][H_3O^+(aq)]}{[HA(aq)]} \quad (10)$$

Solving for the concentration of H_3O^+ the equation becomes:

$$[H_3O^+(aq)] = K_{a_2} \frac{[HA(aq)]}{[A^-(aq)]} \quad (11)$$

Equation (11) demonstrates that the ratio of concentrations $[HA(aq)]/[A^-(aq)]$ directly determines the H_3O^+ concentration in the solution. Equation (9) shows that this ratio is a direct result of the autoprotolysis of HA.

To determine the H_3O^+ concentration at equilibrium, equations (9) and (11) can be combined:

$$[H_3O^+(aq)] = K_{a2} \sqrt{\frac{K_{a1}}{K_{a2}}} \quad (12)$$

This simplifies to:

$$[H_3O^+(aq)] = \sqrt{K_{a1}K_{a2}} \quad (13)$$

Taking the negative logarithm (base 10) of both sides of equation (13):

$$-\log_{10}[H_3O^+(aq)] = -\log_{10}(\sqrt{K_{a1}K_{a2}}) = \frac{1}{2}[(-\log_{10}K_{a1}) + (-\log_{10}K_{a2})] \quad (14)$$

This becomes:

$$pI = \frac{pK_{a1} + pK_{a2}}{2} \quad (15)$$

Thus, the pH at the isoelectric point (pI) is the average of the pK_a values for the α -COOH and α - NH_3^+ groups.

Alternatively, an expression for the concentration of OH^- can be derived from equation (2) as follows:

$$[OH^-(aq)] = \frac{K_w}{K_{a1}} \frac{[HA(aq)]}{[H_2A^+(aq)]} \quad (16)$$

Equation (16) demonstrates that the ratio $[HA(aq)]/[H_2A^+(aq)]$ directly determines the OH^- concentration in the solution. Furthermore, equation (9) shows that this ratio is also a direct result of the autoprotolysis of HA.

Combining equations (9) and (16) to determine the OH^- concentration at equilibrium:

$$[\text{OH}^- (\text{aq})] = \frac{K_w}{K_{a1}} \sqrt{\frac{K_{a1}}{K_{a2}}} \quad (17)$$

$$[\text{OH}^- (\text{aq})] = K_w \sqrt{\frac{1}{K_{a1} K_{a2}}} \quad (18)$$

Taking the negative logarithm (base 10) of both sides of equation (18):

$$-\log_{10}[\text{OH}^- (\text{aq})] = -\log_{10} K_w (\sqrt{K_{a1} K_{a2}}) - \frac{1}{2} [(-\log_{10} K_{a1}) + (-\log_{10} K_{a2})] \quad (19)$$

This simplifies to:

$$\text{pOH} = \text{p}K_w - \left(\frac{\text{p}K_{a1} + \text{p}K_{a2}}{2} \right) \quad (20)$$

From the fifth equilibrium and by definition:

$$\text{p}K_w = \text{pH} + \text{pOH} \quad (21)$$

By combining equations (20) and (21), the same expression as equation (15) is obtained.

Applying equation (15) to the values for glycine from its $\text{p}K_a$ values (7):

$$\text{pI}_{\text{Gly}} = \frac{2.34 + 9.60}{2} = 5.97 \quad (22)$$

pI Calculation for Amino Acids with Acidic Side Chains

The pI calculation for amino acids with acidic side chains can be illustrated using aspartic acid, which has an acidic β -COOH side chain in addition to its α -NH₃⁺ and α -COOH groups (**Figure 2**). As the pH of the solution increases, Aspartic acid undergoes three deprotonation reactions, starting from the fully protonated form H₃A⁺. The α -COOH group, with the lowest pK_a, initially releases its proton, forming the zwitterionic species H₂A. As the pH rises, the β -COOH side chain deprotonates, forming the HA⁻ species. Finally, the α -NH₃⁺ group loses its proton, producing the A²⁻ form. These deprotonation steps are shown in **Figure 2**.

Like the derivation for glycine, the process can be applied to the amphoteric zwitterionic species H₂A, and it can be shown that the pI of aspartic acid is the average of the pK_a values for

the two ionizable groups involved in the protonation and deprotonation of H_2A . These groups are the α -COOH (pK_{a1}) and the β -COOH (pK_{a2}).

Therefore, the pI of Aspartic acid is calculated pK_a values (7) as:

$$pI_{Asp} = \frac{pK_{a1} + pK_{a2}}{2} = \frac{1.88 + 3.56}{2} = 2.72 \quad (23)$$

The β -COOH side chain shifts the pI of Aspartic acid to a more acidic region than glycine, which has a neutral side chain and a pI of 5.97.

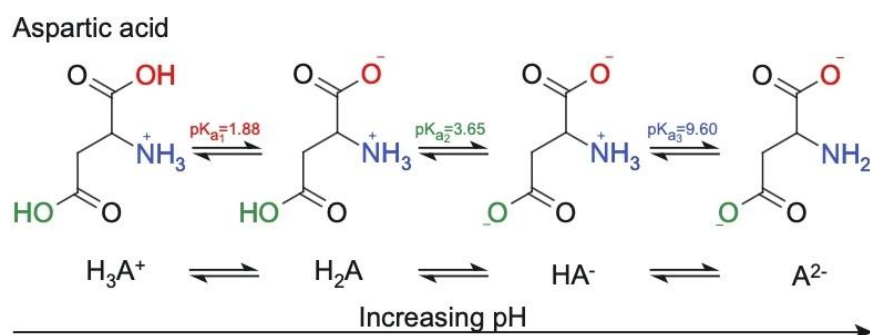


Figure 2. Sequential deprotonation of aspartic acid as the solution pH transitions from acidic to basic. The acidic groups and their corresponding pK_a values are shown in the same color, with each pK_a value dictating the extent of the respective deprotonation step. The different forms of aspartic acid— H_3A^+ , H_2A , HA^- and A^{2-} —are depicted below, with water (H_2O) and hydronium ions (H_3O^+) omitted for clarity to simplify the derivation.

pI Calculation for Amino Acids with Basic Side Chains

Consider the amino acid histidine, which contains a basic imidazole ring as its side chain, in addition to the α -NH₃⁺ and α -COOH groups (**Figure 3**). As the pH increases, histidine undergoes three deprotonation reactions, starting from the fully protonated form, H_3A^{2+} . Initially, the α -COOH group with the lowest pK_a releases its proton, forming the species H_2A^+ . As the pH rises, the imidazole side chain and α -NH₃⁺ groups sequentially lose their protons, forming the zwitterionic HA and A^- forms, respectively. These deprotonation steps are shown in **Figure 3**.

Using the derivation for the amphoteric zwitterionic species HA , it can be demonstrated that the isoelectric point (pI) of histidine is determined by averaging the pK_a values of the two

ionizable groups involved in protonation and deprotonation of HA: the imidazole (pK_{a_2}) and α - NH_3^+ (pK_{a_3}) groups (7).

$$\text{pI}_{\text{His}} = \frac{\text{pK}_{\text{a}_2} + \text{pK}_{\text{a}_3}}{2} = \frac{6.00 + 9.17}{2} = 7.59 \quad (24)$$

Having the extra basic side chain moves the isoelectric point of Histidine to a higher pH value compared to aspartic and glycine.

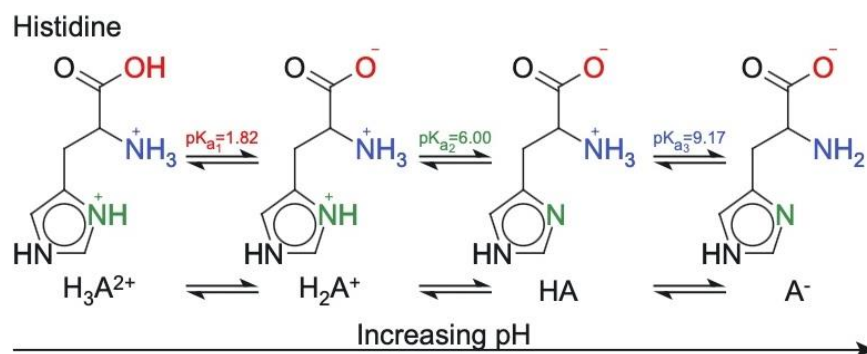


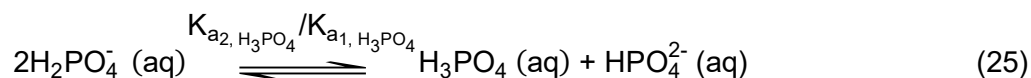
Figure 3. Sequential deprotonation of histidine as the solution pH transitions from acidic to basic. The acidic groups and their corresponding pK_{a} values are shown in the same color, with each pK_{a} value dictating the extent of the respective deprotonation step. The different forms of histidine— H_3A^{2+} , H_2A^+ , HA and A^- —are depicted below, with water (H_2O) and hydronium ions (H_3O^+) omitted for clarity to simplify the derivation.

In conclusion, the derivation presented here detailed a simple mathematical approach for determining the isoelectric point (pI) of any amino acid. The zwitterionic form of the amino acid, which is electrically neutral, behaves as an amphoteric species and autoprotolysis equilibrium that thermodynamically and kinetically dominates and determines the pI in solution, provided the corresponding equilibrium constant is higher than other possible acid-base equilibria. Additionally, this approach can be applied to theoretically calculate the pH values of amphoteric solutions, such as the equivalence points in the titration of polyprotic acids (e.g., H_3PO_4) or polyprotic bases (e.g., Na_2CO_3) (3-6).

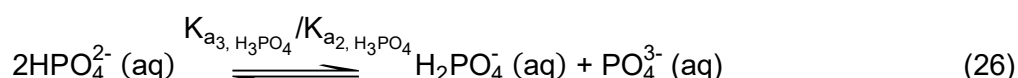
For example, consider the titration of H_3PO_4 with NaOH. At the first equivalence point, all the H_3PO_4 is converted to H_2PO_4^- , and at the second equivalence point, all H_2PO_4^- is converted

to HPO_4^{2-} . These two amphoteric species can then establish the following equilibria at their respective equivalence points, which determine the pH values:

At the first equivalence point:



At the second equivalence point:



By applying a derivation like that used for glycine (equations 1–22), the pH at the equivalence points, denoted as pH_{IP1} and pH_{IP2} , can be easily calculated using the following relationships:

$$\text{pH}_{\text{IP1}} = \frac{\text{p}K_{a1, \text{H}_3\text{PO}_4} + \text{p}K_{a2, \text{H}_3\text{PO}_4}}{2} \quad (27)$$

$$\text{pH}_{\text{IP2}} = \frac{\text{p}K_{a2, \text{H}_3\text{PO}_4} + \text{p}K_{a3, \text{H}_3\text{PO}_4}}{2} \quad (28)$$

References

1. Nelson, David L.; Cox, Michael M. *Lehninger Principles of Biochemistry* (8th ed.); W.H. Freeman and Company: New York, 2021.
2. Berg, Jeremy M.; Tymoczko, John L.; Gatto, Gregory J.; Stryer, Lubert. *Biochemistry* (9th ed.); W.H. Freeman and Company: New York, 2021.
3. Skoog, Douglas A. *Fundamentals of Analytical Chemistry*; Cengage - Brooks/Cole: Belmont, CA, 2012.
4. Segel, Irwin H. *Biochemical Calculations* (2nd ed.); Wiley: New York, 1993.
5. Harris, Daniel C.; Lucy, Charles A. *Quantitative Chemical Analysis*; Macmillan Learning: Austin, 2020.
6. Hawkes, S. J. (2000). Easy derivation of $\text{pH} \approx (\text{p}K_{a1} + \text{p}K_{a2})/2$ using autoprotolysis of HA^- : Doubtful value of the supposedly more rigorous equation. *Journal of Chemical Education*, 77(9), 1183. <https://doi.org/10.1021/ed077p1183>.
7. Lide, D. R. *Handbook of Chemistry and Physics*, 72nd ed.; CRC Press: 1991.
8. Senavirathne G, On the thermodynamic feasibility of acid-base equilibria. *ChemRxiv*. 2024; doi:10.26434/chemrxiv-2024-ms4fs-v2.

9. Atkins, Peter W.; De Paula, Julio; Keeler, James. *Atkins' Physical Chemistry*; Oxford University Press: New York, NY, 2023.

Appendix A: List of Terms

1. **Zwitterion** – A molecule or ion with both a positive and negative charge but is electrically neutral overall.
2. **Amphoteric species** – A substance that can act as an acid and base, depending on the environment.
3. **K_a** – The equilibrium acid dissociation constant.
4. **K_b** – The equilibrium base dissociation constant.
5. **K_w** – The ionization constant for water (1.0×10^{-14} at 25 °C)
6. **pK** – The negative logarithm of the equilibrium constant (K_{eq}) for a given reaction, expressed as $pK = -\log_{10} K_{eq}$

Appendix B: The derivations for the equilibrium constant for the second equilibrium

The equilibrium constant for this reaction ($K_{eq}(2)$) can be written as,

$$K_{eq}(2) = \frac{[H_2A(aq)][OH^-(aq)]}{[HA(aq)]} \quad (B1)$$

By multiplying the nominator and denominator by $[H_3O^+(aq)]$ and by rearranging the terms,

$$K_{eq}(2) = \frac{[H_2A(aq)][OH^-(aq)]}{[HA(aq)]} \times \frac{[H_3O^+(aq)]}{[H_3O^+(aq)]} \quad (B2)$$

$$K_{eq}(2) = [H_3O^+(aq)][OH^-(aq)] \times \frac{[H_2A(aq)]}{[HA(aq)][H_3O^+(aq)]} \quad (B3)$$

$$K_{eq}(2) = K_w/K_{a_1} \quad (B4)$$

Since $K_w = K_{a_1}K_{b_1}$, where K_{b_1} is the dissociation constant of amphoteric HA when it acts as a base,

$$K_{eq}(2) = K_{b_1} \quad (B5)$$

Appendix C: The derivations for the equilibrium constant for the fourth equilibrium

The equilibrium constant for this reaction ($K_{eq}(3)$) can be written as,

$$K_{eq}(3) = \frac{[H_2A^+(aq)][A^-(aq)]}{[HA(aq)]^2} \quad (C1)$$

By multiplying the nominator and denominator by $[H_3O^+(aq)]$ and by rearranging the terms,

$$K_{eq}(3) = \frac{[H_2A^+(aq)][A^-(aq)]}{[HA(aq)]^2} \times \frac{[H_3O^+(aq)]}{[H_3O^+(aq)]} \quad (C2)$$

$$K_{eq}(3) = \frac{[A^-(aq)][H_3O^+(aq)]}{[HA(aq)]} \times \frac{[H_2A^+(aq)][H_3O^+(aq)]}{[H_3O^+(aq)][HA(aq)]} \quad (C3)$$

$$K_{eq}(3) = K_{a2}/K_{a1} \quad (C4)$$

Appendix D: Thermodynamic and Kinetic Feasibility of the Reactions at the pI

The second, third, and fourth equilibria appear to be competing reactions (with H_2O , H_3O^+ , and HA acting as competing acids to protonate HA). The question, then, is which of these reactions is likely to occur in solution based on thermodynamic and kinetic feasibility (8,9). The third equilibrium is the most thermodynamically favorable (with the most prominent equilibrium constant), provided that sufficient H_3O^+ is available. However, at the isoelectric point (pI), the third reaction strictly depends on the H_3O^+ produced in the first equilibrium. The coupling between the two dampens the overall equilibrium constant from $1/K_{a1}$, reducing it to K_{a2}/K_{a1} . Moreover, the coupled first and third reactions resemble the autoprotolysis of HA (equation (4) from the fourth equilibrium), meaning that the third and fourth equilibria have the same thermodynamic feasibility. Based on thermodynamic feasibility, the third and fourth reactions are much more favorable than the second.

Although previous derivations of pI considered the coupled reactions rather than the HA autoprotolysis, it is easy to prove that both are equivalent using basic thermodynamics (8,9). First, consider the relationship between the equilibrium constants and the standard Gibbs free energy change (ΔG^0) of the first, second, and third reactions:

$$\Delta G^0(1) = -RT \ln K_{a2} \quad (D1)$$

$$\Delta G^0(3) = -RT \ln (1/K_{a1}) \quad (D2)$$

$$\Delta G^0(4) = -RT \ln (K_{a_2} / K_{a_1}) \quad (D3)$$

Since Gibbs free energy is a state function (9), the ΔG^0 for the coupled first and third equilibria can be found by adding equations (D1) and (D2) as follows:

$$\Delta G^0(1+3) = \Delta G^0(1) + \Delta G^0(3) = -RT \ln K_{a_2} + -RT \ln (1/K_{a_1}) \quad (D4)$$

$$\Delta G^0(1+3) = -RT \ln (K_{a_2} / K_{a_1}) \quad (D5)$$

Since equation (D3) is equivalent to equation (D5), regardless of their actual mechanisms, this shows that the net coupled reaction is equivalent to the autoprotolysis of HA in terms of the products and the extent. Both reactions are plausible for deriving pI, although previous derivations completely ignored the latter (4,5).

However, to understand which mechanism is more plausible, let us consider the kinetic aspect. For this, assume a solution of 0.1 M total glycine reaching the isoelectric point. Equation (10) can be used to estimate the concentration of H_3O^+ , assuming no other reactions occur. That is, the first equilibrium will produce equal amounts of H_3O^+ and A^- , which is set to x in the following:

From equation (10):

$$10^{-9.60} = \frac{x^2}{0.1 - x} \quad (D6)$$

Since $x \ll 0.1$, $0.1 - x$ can be approximated as 0.1. Therefore:

$$x = \sqrt{10^{-10.60}} \quad (D7)$$

$$x = \sqrt{10^{-10.60}} \approx 5 \times 10^{-6} \quad (D8)$$

For a 0.1 M glycine solution, the first equilibrium alone will produce approximately 5 μM H_3O^+ . Ignoring the autoprotolysis of HA (equation (4)), if the first equilibrium is coupled with the third equilibrium, both reactions will proceed until the pH settles at the pI. As a result, the maximum

concentration of H_3O^+ sustained at equilibrium through this coupling is approximately 1 μM (since the pI of glycine is ~ 6 , this corresponds to $\sim 1 \times 10^{-6}$ M H_3O^+ in the solution).

Once the HA autoprotolysis begins to dominate the total equilibrium, the higher concentrations of A^- and H_2A^+ (~ 74 μM , see equations (D9) and (D10)) produced will begin suppressing the forward reaction of the first three equilibria, according to Le Chatelier's principle. This explains how the isoelectric solution maintains its pH and equal concentrations of A^- and H_2A^+ , which, in turn, guarantees the average electrical neutrality of glycine (3-6).

A similar calculation using equation (6) can be used to estimate the equilibrium concentration of HA (set as $0.1 - x$; x is the H_2A^+ and A^- concentrations) when a 0.1 M glycine solution reaches isoelectric equilibrium.

From equation (6):

$$10^{-7.26} = \frac{x^2}{0.1 - x} \quad (\text{D9})$$

Solving for x without assuming $x \ll 0.1$ and $0.1 - x \approx 0.10$:

$$x = 7.41 \times 10^{-5} \quad (\text{D10})$$

Even without the assumption that x is negligible compared to 0.1 M, it can be shown that the concentration of HA at equilibrium is still close to 0.1 M. Additionally, the concentration of water in an aqueous solution is ~ 55.5 M. Now, with these numbers in hand and assuming the reactions occur by thermal diffusion limited collision between reactants (9), it can be argued that there is a much higher chance that ~ 0.1 M HA will react with ~ 55.5 M H_2O (as described in equations (1) and (2) from the first and second equilibria) than that it will react with $\sim 1\text{-}5$ μM H_3O^+ (as in equation (3) of the third equilibrium). Furthermore, there is a much higher likelihood that ~ 0.1 M HA will react with itself (autoprotolysis equilibrium, equation (4)). Thus, from a kinetic perspective, it can be argued that the first, second, and fourth equilibria are more plausible than the third. Therefore,

the third equilibrium was ignored from the derivation of pI , although one may use the coupled reactions to arrive at the same conclusion (4-6).