Modelling Movement of Ions During Isoelectric Focusing

CHOOSE-FOCUS-ANALYZE (CFA) EXERCISE

BT 3011: TRANSPORT PHENOMENA IN BIOLOGICAL SYSTEMS

DEBAYAN CHAUDHURY
BE14B035
November 4,2016

Contents

1 Objective 3

2 Introduction 3
- 2.1 Importance of pH in biological systems . . . 3
- 2.2 Isoelectric Focusing . . . 4

3 Approach 5
- 3.1 Brief Description of System . . . 5
- 3.2 Modelling the Ionization of Proteins . . . 6
- 3.3 Shell Balance on a differential strip in the Gel . . . 6
- 3.4 Charge as a function of distance travelled in the gel . . . 7
- 3.5 Force Balances . . . 9
- 3.6 Drag Force and Velocity Function . . . 9
- 3.7 Time Taken to Complete the Procedure . . . 10

4 Discussion 12
- 4.1 Qualitative Analysis . . . 12
  - 4.1.1 Charge . . . 12
  - 4.1.2 Acceleration . . . 12
  - 4.1.3 Velocity . . . 12
  - 4.1.4 Time . . . 13
- 4.2 Quantitative Analysis . . . 13

5 Limitations and Further Work 14

6 Conclusion 15

7 Acknowledgment 16

8 References 16

1 Objective

Isoelectric focusing is a common biochemical technique used in the separation of protein mixtures on the basis of their isoelectric points. As a part of my Choose-Focus-Analyze (CFA) exercise, I attempt to develop mathematical models based on shell and force balances to analyze the dependence of charge, acceleration and velocity of protein samples moving under the influence of an electric field in a uniform pH gradient on the distance travelled by the sample in the gel during isoelectric focusing.

2 Introduction

2.1 Importance of pH in biological systems

pH is an essential parameter when it comes to dealing with biological systems. Most biological systems involve aqueous solutions and thereby are highly susceptible to the effects of pH.[1] This is particularly true for proteins, which are biological marcro-molecules made from constituent amino acids. Due to the highly ionic carboxylic acid (-COOH) and amine (-NH_2) groups present in these amino acids, proteins are generally charged in nature. Some amino acids (such as lysine and aspartic acid) also contain ionic side chains which contribute to the charge of the protein. This charge is often key to the proper functioning of the protein and is greatly affected by changes in pH . Therefore, protein function is highly dependent on the pH of the protein solution.[1],[2],[3]

At low pH, all the ionic groups in the protein are protonated due an abundance of H^+ ions in the solution. Therefore, there is a net positive charge on the protein at low pH as the amine groups pick up protons, while the carboxylic acid groups remain neutral.As the pH increases, both the amine and carboxylic acid groups begin to lose their protons, gaining a net neutral and negative charge respectively at conditions of extremely high pH .In between, there exists a pH at which the net positive charge of the amine groups is exactly cancelled by the net negative charge on the carboxylic acid groups, which results in the protein being electrically neutral . This pH is called the *isoelectric point (pI)* of the protein, and is often used to characterize the functioning of the protein.[4] The ion formed by at the isoelectric point (pI), which is overall neutral, is called a *zwitterion.*[5] Figure 2.1 shows the ionization of the simplest amino acid, glycine.[6]

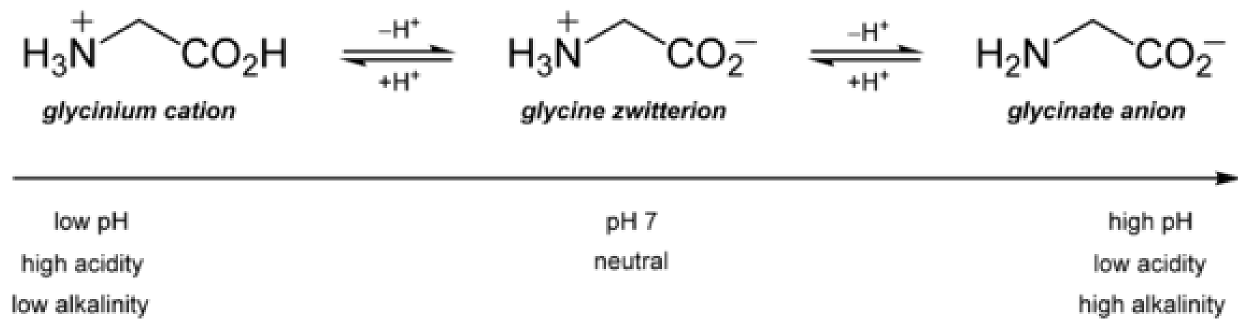


Figure 2.1: Ionization of Glycine-Source:Wikipedia

2.2 Isoelectric Focusing

Isoelectric focusing is a powerful technique used in the separation of protein mixtures which takes advantage of the diverse isoelectric points of different proteins. In isoelectric focusing, samples are loaded on one end of a gel which has a uniform pH gradient, i.e. the pH increases linearly down the length of the gel. This pH gradient is introduced by the use of ampholytes,which are low molecular weight amphoteric molecules used to prepare the gel. A uniform electric field is applied to the gel and the proteins which are initially charged start moving inside the gel in the direction of increase in pH . This is because the proteins are positively charged at low pH and therefore move away from the electric field in the direction of increasing pH .[7]

As the samples move along the gel, the pH increases and the positive charge on the protein decreases. As a consequence, the acceleration of the proteins, which is directly proportional to the charge of the protein, also decreases.In addition to the force due the electric field, the proteins also experience a drag force exerted by the gel. When these samples reach the region where the pH is equal to their pI, the charge on these samples is reduced to zero. Consequently, the acceleration at this point is also zero. On further travelling forward in the gel due to any residual velocity (which is already reduced due to presence of drag), they experience an acceleration in the opposite direction (as they now acquire a small negative charge) due to the electric field in addition to the already present drag. Therefore, they come to rest near the point where the pH is equal to the pI. Since different proteins have different isoelectric points, they travel different distances in the gel during isoelectric focusing. The protein of interest can thereby be identified and separated from a mixture of proteins by running the mixture against the pure protein of interest (control). The band

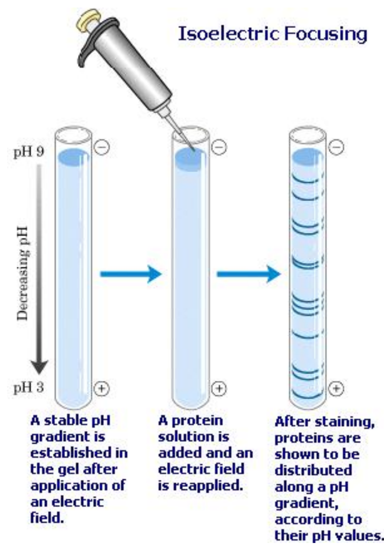


Figure 2.2: Principles of Isoelectric Focusing-Source:Lehninger Biochemistry

from the mixture which is at the same distance as the pure protein in the contains the protein of interest and is cut out from the gel and purified for further use.

A similar situation arises in the case of a decreasing pH gradient. The proteins are initially negatively charged at high pH and move towards the electric field, losing their negative charge as the pH decreases with distance travelled in the gel. Figure 2.2 illustrates the basic principles of isoelectric focusing.[7]

3 Approach

3.1 Brief Description of System

The gel is considered to be similar to the gels used in standard electrophoretic procedures, i.e. a cuboid having length L, breadth B and width W. The width of the gel is considered to be very small so that the motion in the vertical direction due to gravity can be neglected. A linearly increasing pH gradient is introduced in the gel by means of ampholytes. The pH is a linear function of distance x along the length of the gel and is given by:

$$pH = \frac{14}{L}x \tag{3.1}$$

Protein samples are loaded at $x = 0$ and a uniform electric field of magnitude E is applied on the gel.

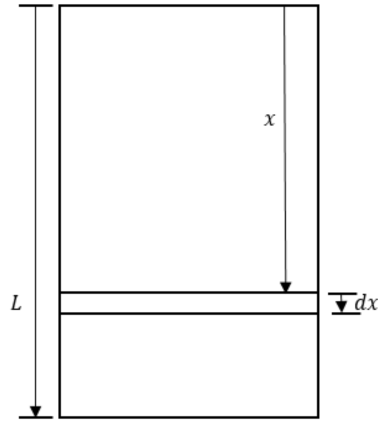
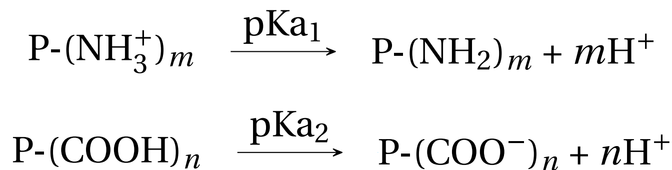


Figure 3.1: Shell Balance on a Differential Strip

3.2 Modelling the Ionization of Proteins

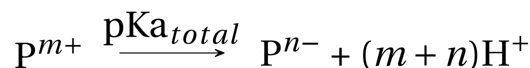
Ionization for a protein having M amine groups and N carboxylic acid groups is modelled by the following chemical equations:



$$\text{P-}(\text{NH}_3^+)_m \xrightarrow{\text{pKa}_1} \text{P-}(\text{NH}_2)_m + m\text{H}^+$$

$$\text{P-}(\text{COOH})_n \xrightarrow{\text{pKa}_2} \text{P-}(\text{COO}^-)_n + n\text{H}^+$$

It is assumed that both reactions take place simultaneously and at comparable rates. This implies that the concentration of the uncharged (Note: uncharged is different from neutral) product,which was initially very low due to the highly acidic starting pH, does not vary significantly with time. Therefore, the above two equations are added to obtain the net equation for protein ionization as:



$$\text{P}^{m+} \xrightarrow{\text{pKa}_{total}} \text{P}^{n-} + (m+n)\text{H}^+$$

$$\text{pKa}_{total}\text{=pKa}_1\text{+pKa}_2$$

3.3 Shell Balance on a differential strip in the Gel

Figure 3.1 illustrates the differential strip of length dx under consideration. The initial concentration of the protein loaded onto the gel is taken to be c_0. It is assumed that at distance x, a fraction y of the initial cationic form has dis-

sociated to give the anionic form and H^+. Therefore, the concentration of the cationic form at a distance x is $c_0(1-y)$, and the concentration of the anionic form is $c_0 y$. The conservation equation for charge Z is given by:

$$Input - Output + Generation - Cousumption = Accumulation \quad (3.2)$$

The expressions for the relevant terms at steady state are listed below:

$$Input = Z|_x$$

$$Output = Z|_{x+dx}$$

$$Generation = -nc_0 VF(y|_{x+dx} - y|_x)$$

$$Consumption = mc_0 VF((1-y)|_x - (1-y)|_{x+dx})$$

$$Accumulation = 0$$

Here,V is the total volume of the sample loaded and F is Faraday's constant.Replacing the terms in the equation and dividing throughout by dx, the relation between the charge Z and fraction of anion y is obtained as:

$$\frac{dZ}{dx} = -(m+n)FVc_0\frac{dy}{dx} \quad (3.3)$$

Integrating, the charge Z is obtained as a function of the anionic fraction y

$$Z = -(m+n)c_0 VFy + k$$

Here, k is the constant of integration which can be obtained by imposition of a boundary condition.The boundary condition is that when the fraction of anion present (y) is zero, net charge on the protein (Z) is $mc_0 VF$ as the entire protein is present in the cationic state. Therefore, the final expression for Z reduces to:

$$Z = c_0 VF(m - (m+n)y) \quad (3.4)$$

$$y = \frac{mc_0 VF - Z}{(m+n)c_0 VF} \quad (3.5)$$

3.4 Charge as a function of distance travelled in the gel

The fraction of the anionic component y can be related to the pH by the Henderson-Hasselbach Equation.[8] For the model of ionization of proteins that was con-

sidered, this relation is given by:

$$pH = pKa_{total} + log\frac{[P^{n-}]}{[P^{m+}]}$$

The anionic component concentration $[P^{n-}]$is equal to $c_0 y$. Substitution of this value in the equation yields:

$$pH = pKa_{total} + log\frac{y}{1-y} \tag{3.6}$$

Differentiating equation 3.6 with respect to x, the following relationship is established:

$$\frac{dpH}{dx} = \frac{1}{y(1-y)}\frac{dy}{dx} \tag{3.7}$$

Now, the pH gradient $\frac{dpH}{dx} = \frac{14}{L}$. Therefore, equation 3.7 reduces to:

$$\frac{dy}{y(1-y)} = \frac{14}{L}$$

Integration by partial fractions gives:

$$ln\frac{y}{(1-y)} = \frac{14}{L}x + k \tag{3.8}$$

Now, the boundary condition is that at $pH = pI$, $Z = 0$.The x corresponding to when the pH becomes pI is given by:

$$x = \frac{L}{14}pI$$

At this x, $Z = 0$. The corresponding value of y which is obtained by the use of equation 3.5 is:

$$y = \frac{m}{(m+n)}$$

Hence, the boundary conditions change to at $x = \frac{L}{14}pI$, $y = \frac{m}{(m+n)}$.

Substitution in equation 3.8 gives the value of the integration constant k as:

$$k = ln\frac{m}{n} - pI$$

Therefore, the final expression for the anionic component fraction y in terms of the distance travelled x is:

$$y = \frac{e^{(ln\frac{m}{n} - pI + \frac{14}{L}x)}}{1 + e^{(ln\frac{m}{n} - pI + \frac{14}{L}x)}} \tag{3.9}$$

The final expression for the charge Z as a function of the distance travelled x is obtained by substituting the expression for y in equation 3.4:

$$Z(x) = c_0 VF(m - (m+n)\frac{e^{(ln\frac{m}{n} - pI + \frac{14}{L}x)}}{1 + e^{(ln\frac{m}{n} - pI + \frac{14}{L}x)}}) \tag{3.10}$$

3.5 Force Balances

In this section, an attempt is made to analyze the forces acting on the protein samples as they move through the gel.The major forces acting on the samples are the force due to the electric field and the drag force. Therefore, net acceleration of the sample is given by:

$$a = (Z(x)|E| - DragForce)/M \tag{3.11}$$

Here, $Z(x)$ is the charge on the sample, $|E|$ is the magnitude of the constant electric field and M is the mass of the sample.

3.6 Drag Force and Velocity Function

The movement of protein samples through the gel can be considered analogous to the movement of small spherical particles through a highly viscous liquid. Low values of the density, velocity and diameter of the samples coupled with the high value of viscosity for the gel implies that the Reynolds Number (N_{Re}) is extremely small.Therefore, Stokes Law[9] can be used to model the viscous force acting on the samples. According to Stokes Law, the viscous drag force is given by:

$$F_d = 6\pi\mu r v \tag{3.12}$$

Here, μ is the viscosity of the gel, r is the radius of the sample and v is velocity of the sample. Using this expression, Equation 3.11 is reduced to:

$$a = \frac{Z(x)|E| - 6\pi\mu r v}{M}$$

Replacing a with $v\frac{dv}{dx}$ and rearranging, the following relationship is established:

$$Mv\frac{dv}{dx} + 6\pi\mu r v = Z(x)|E| \tag{3.13}$$

This equation gives velocity v as a function of the distance travelled x but is quite hard to solve analytically. However, if an assumption is made that the net force on the sample is nearly zero i.e. the acceleration term is negligible, the velocity function can be easily obtained. Therefore, for a nearly equal to zero, the velocity is given by:

$$v(x) = \frac{Z(x)|E|}{6\pi\mu\, r} \tag{3.14}$$

Therefore it is seen that the velocity of the sample has the same form of dependence on the distance travelled x as the charge of the sample. One interesting observation is that at the x where $pH = pI$, $Z = 0$, and therefore the velocity must also be equal to zero. This means that the samples must come to rest at the x where the pH equals their pI, which is what is observed in practice. Therefore, the assumptions made are valid and the developed model is fairly accurate.

3.7 Time Taken to Complete the Procedure

The time taken to complete the procedure is computed by substituting the velocity as the rate of change in distance. For this computation a new variable $u = \frac{14}{L}x - pI + ln\frac{m}{n}$ is defined. Therefore:

$$du = \frac{14}{L}dx$$

$$dx = \frac{L}{14}du \tag{3.15}$$

Now, replacement of v in Equation 3.14 gives:

$$\frac{dx}{dt} = \frac{c_0 VF|E|}{6\pi\mu r}(m - \frac{(m+n)e^u}{1+e^u})$$

$$\frac{dx}{dt} = \frac{c_0 VF|E|}{6\pi\mu r}\frac{(m-ne^u)}{(1+e^u)}$$

$$\frac{(1+e^u)dx}{(m-ne^u)} = \frac{c_0 VF|E|}{6\pi\mu r}dt$$

$$\frac{(m-ne^u)dx}{(m-ne^u)} + \frac{(m+n)e^u dx}{m-ne^u} = \frac{mc_0 VF|E|}{6\pi\mu r}dt$$

Integration on both sides gives the final expression as:

$$x - (m+n)\frac{L}{14n}ln(m-ne^u) = \frac{mc_0 VF|E|}{6\pi\mu r}t + constant \tag{3.16}$$

Now at $t = 0,\ x = 0$. Therefore value of the integration constant comes out to be:

$$Constant = -(m+n)\frac{L}{14n}(m - ne^{ln\frac{m}{n} - pI})$$

Therefore, the final equation is given by:

$$x - (m+n)\frac{L}{14n}ln\frac{(m-ne^u)}{m - ne^{(ln\frac{m}{n} - pI)}} = \frac{mc_0 VF|E|}{6\pi\mu r}t \tag{3.17}$$

Time taken to complete the procedure is given by the time at $x = \frac{L}{14}pI$. However, as x approaches this value, the $(m-ne^u)$ term in the numerator approaches zero, due to which the ln term which returns a large negative value. This is converted to a positive value by the - sign which increases the t value by a large amount. This implies that the samples will technically never reach the x value at which the pH is equal to their pI value in finite time. This is quite consistent with the actual protocol which requires the gel to be run for large time periods (around 16-24 hours)[10] in order to obtain results.

For the numerator in the logarithmic term to be greater than zero, the following condition has to satisfied:

$$m - ne^u > 0$$
$$u < log\frac{m}{n}$$
$$\frac{14}{L}x - pI < 0$$
$$x < \frac{L}{14}pI \tag{3.18}$$

Which is in accordance with the conclusion that was reached qualitatively.

4 Discussion

4.1 Qualitative Analysis

4.1.1 Charge

It is seen that the charge on the protein decreases exponentially with the distance travelled in the gel as given by Equation 3.10. The sample starts at $x = 0$ (pH approximately equal to 0) with a net positive charge and becomes electrically neutral at $x = \frac{L}{14}pI$ where the $pH = pI$.It is also seen that the distance at which the sample stops moving is directly proportional to the pI, which implies that proteins with different pI move different distances before coming to rest.This is the basic operational principle of isoelectric focusing.

4.1.2 Acceleration

The acceleration can be found as a function of distance by solving Equation 3.13. In this analysis, the net force on the sample and thereby the acceleration has been considered to be nearly zero in order to simplify the model.

4.1.3 Velocity

Under the assumption that the net acceleration of the sample is nearly zero, the velocity is given as a function of distance by Equation 3.14. It is seen that the velocity is directly proportional to the charge and has the same dependence on the distance travelled x as the charge Z. Additionally, at the distance where the charge reduces to 0 (where $pH = pI$), the net velocity also reduces to zero

indicating that the particles come to rest at that distance.

It is also seen from Equation 3.14 that when the drag force is modelled by Stokes Law, the velocity is inversely proportional to the size (radius) of the sample and viscosity of the gel.Additionally, the velocity is found to be independent of the mass of loaded sample. The results so obtained are in accordance with what is observed in the actual procedure, and thereby justify that the assumptions made while modelling are reasonable.

4.1.4 Time

The time taken by the samples to move a distance x in the gel is given by Equation 3.17. It is observed that the samples are theoretically not expected to reach the distance where $pH = pI$ in finite time.Therefore, the real life protocol involves a large running time for the gel in order to ensure that the samples reach as close to the expected distance as possible. Additionally,it is seen that the equation is not valid for x at which $pH > pI$, as the value of the numerator in the log turns out to be negative and the logarithmic function is not defined for negative numbers. This implies that the samples cannot ever travel a distance greater than the distance at which the $pH = pI$.

4.2 Quantitative Analysis

In this section, the relationship between the variables of interest and the distance travelled in the gel is analyzed by inputting feasible sample values. The results of analysis are illustrated in Figure 4.1 and Figure 4.2.

A protein having 1 amine group and 1 carboxylic acid group is considered. Therefore, m=n=1. The pI of the protein is taken to be 7 and the initial concentration is taken to be 0.1M. The volume loaded onto the gel is assumed to be 10uL (10^{-5}L) which is approximately the volume used in the actual procedure. Faraday's constant is 96500 $C/mole$. The length of the gel is taken to be 28 cm, which is a fairly typical value for gels. The graph between charge Z vs the distance travelled in the gel x is given in Figure 4.1. It is seen that the charge Z begins by considering a fully positively charged sample with $charge = mc_0VF = +0.0965C$. The charge then exponentially decreases with x, hitting 0 at $x = 14$ as shown in Figure 4.1. It then reaches a fully negative value at $x = 28$, where the pH is 14. The charge at this point should be $-nc_0VF = -0.0965C$, which corresponds to the value at that point in the plot.

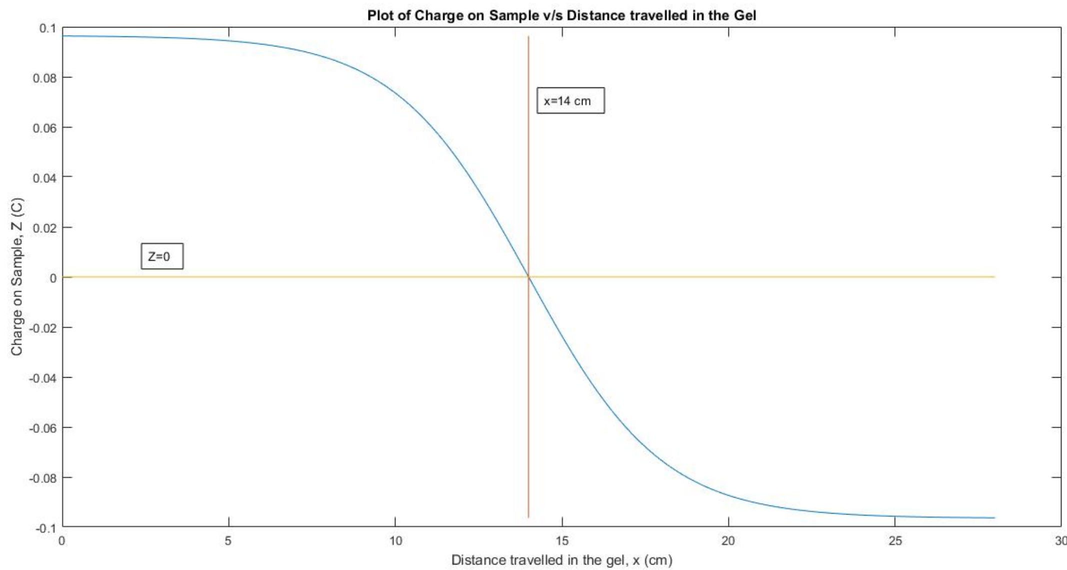


Figure 4.1: Variation of Z with x for defined parameters

The graph between the velocity v and the distance travelled x has the same shape as the Z vs x curve as v is linearly dependent on Z. The only difference is that the values on the y-axis get scaled by the constant factor k as per Equation 3.14.

The exact graph between distance travelled x and time taken t is difficult to obtain as the value of the coefficient of t given in equation 3.17 depends on the parameters of the actual experimental set up. It is difficult to obtain a reliable average value for this coefficient. However, by arbitrarily setting the coefficient to 1, an idea of how x varies with t can be obtained. This is shown in Figure 4.2 for the same protein that we considered for Z vs x. From Figure 4.2, it is seen that the distance x increases with t till x approaches 14.Here,the curve becomes asymptotic to the $x = 14$ line indicating that the proteins do not reach the distance at which $pH = pI$ in finite time.

5 Limitations and Further Work

One major limitation with the derived model is that in modelling of the ionization of proteins, the concentration of the uncharged protein is considered to be negligible and at steady state. This assumption does not hold in practice as the concentration of uncharged protein reaches significant values when the

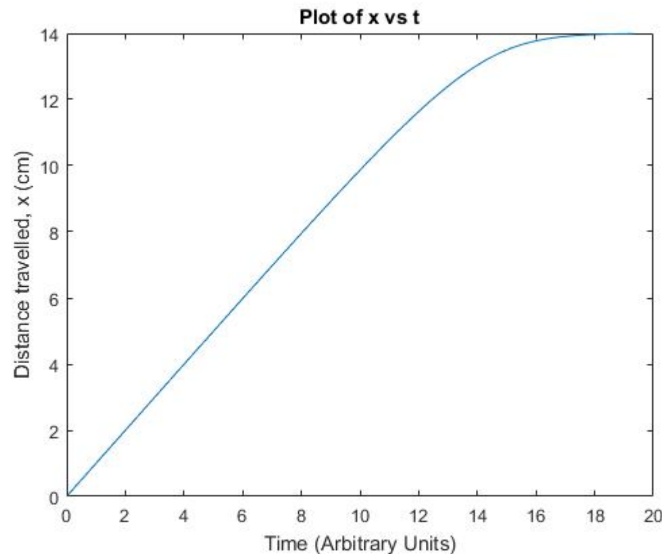


Figure 4.2: Variation of x with t for arbitrary parameters

sample reaches the region where the pH tends towards pI. Because of this, the sample appears scattered over a range of x values near the point where the $pH = pI$ instead of a sharp band at one particular x as indicated by the developed model.

The gel has been considered as a viscous fluid and the drag force experienced by the protein samples in the gel has been modelled by Stokes Law. Better models of drag force may give more accurate velocity and acceleration functions.

In this analysis, diffusion effects have been ignored. However,since the gel has to be run for a long time in order to obtain results, diffusive fluxes might contribute significantly to the movement of the samples in the gel. Incorporation of motion of samples due to diffusion might further improve the model.

6 Conclusion

Therefore in this CFA exercise the movement of charged protein samples through a gel having a linear pH gradient under the influence of an electric field is investigated. Expressions are derived for the charge, acceleration and velocity of the moving samples as a function of the distance travelled along the length of the gel which reasonably model the real-life protocol. The time for which the gel must be run in order to obtain results is analyzed and some interesting insights are derived from this expression.

7 Acknowledgment

I would like to thank Dr.GK Suraishkumar for providing me an opportunity to carry out this interesting exercise. I would also like to thank my friend Onkar for organizing a LaTeX bootcamp for my benefit, without which this report would have looked a lot shabbier.
All images in this report are original, unless otherwise credited.

8 References

[1]-https://groups.chem.ubc.ca/courseware/pH/section19/index.html
[2]-http://biology.stackexchange.com/questions/388/what-effect-has-changing-ph-and-salt-concentration-on-protein-complexes
[3]-Talley,K. and Alexov,E.: On the pH optimum activity and stability of proteins. Proteins,2010 Sep: 78(12): 2699-2706
[4]-Wikipedia Article:Isoelectronic Point
[5]-Wikipedia Article:Zwitterion
[6]-Wikipedia Article:Glycine
[7]-Lehninger Principles of Biochemistry, 4th Edition
[8]-Wikipedia Article: Henderson-Hasselbalch Equation
[9]-Wikipedia Article: Stokes Law
[10]-https://www.nationaldiagnostics.com/electrophoresis/article/isoelectric-focusing