CARBONIC ANHYDRASE: KINETICS OF REMOVAL OF Zn(II) By 2,6-PYRIDINECARBOXYLATE

Paul Nickerson*

Department of Chemistry, University of Florida, Gainesville, FL, US

E-mail: pvnick@ufl.edu

Abstract

In the past decade, supervised activity recognition methods have been studied by many researchers, however these methods still face many challenges in real world settings. Supervised activity recognition methods assume that we are provided with labeled training examples from a set of predefined activities. Annotating and hand labeling data is a very time consuming and laborious task. Also, the assumption of consistent pre-defined activities might not hold in reality. More importantly, these algorithms do not take into account the streaming nature of data, or the possibility that the patterns might change over time. In this chapter, we will provide an overview of the state of the art *unsupervised* methods for activity recognition. In particular, we will describe a scalable activity discovery and recognition method for complex large real world datasets, based on sequential data mining and stream data mining methods.

*To whom correspondence should be addressed

Introduction

Carbonic anhydrase (CA) catalyzes the interconversion of carbon dioxide and carbonic acid/bicarbonate as follows:

$$CO_2 + H_2O \overset{CA}{\rightleftharpoons} H_2CO_3 \tag{1}$$

In the active form, CA is bound to a Zn^{2+} cofactor (denoted as CA·Zn), which it relies upon for its catalytic activity. The zinc ion can be stripped from the enzyme using a Lewis base ligand, which donates electrons to the ion to form a covalent bond. The ligand being studied in this experiment is 2,6-pyridinecarboxylate, commonly called dipicolinate (or dipic). Figure 1 shows the structure of dipic. In this experiment, the rate of zinc removal by dipic will be measured.

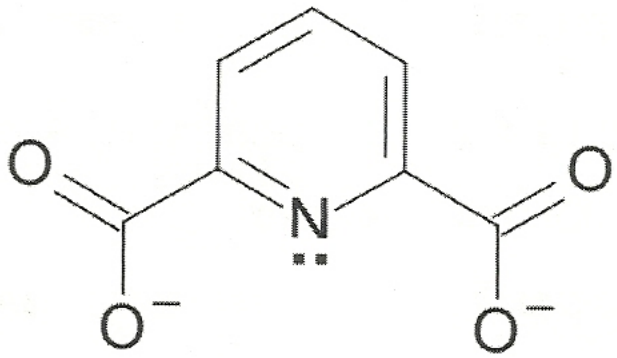

Figure 1: Structure of 2,6-pyridinecarboxylate (dipic)[1]

Mechanism

When $[dipic] >> [CA]$, that is, when $\frac{[dipic]}{[CA]} \geq 25$, then the removal of zinc is pseudo-first-order with respect to CA·Zn because the concentration of dipic, denoted as L, does not change appreciably. Thus the formation of the inactive enzyme, apoCA, can be modeled

using the following rate equation:

$$\frac{d[\text{apoCA}]}{dt} = k_{obs}[\text{CA·Zn}] \tag{2}$$

The pseudo-first-order rate constant, k_{obs}, increases as [L] increases, but levels off at sufficiently high concentrations of L. Biochemists will recognize behavior similar to Michaelis-Menten enzyme kinetics in which the enzyme, CA·Zn, and the substrate, L, reversibly form a CA·Zn·L complex with association constant K_{EML} (EML stands for Enzyme-Metal-Ligand):

$$\text{CA·Zn} + \text{L} \overset{\text{K}_{\text{EML}}}{\rightleftharpoons} \text{CA·Zn·L} \tag{3}$$

This can be modeled as follows:

$$K_{EML} = \frac{[\text{CA·Zn·L}]}{[\text{CA·Zn}][\text{L}]} \tag{4}$$

CA·Zn·L can either revert back to the original species or irreversibly convert to the inactive form of the enzyme, apoCA, and the covalently bound zinc-dipic molecule, ZnL:

$$\text{CA·Zn·L} \xrightarrow{\text{k}_\text{d}} \text{apoCA} + \text{ZnL} \tag{5}$$

This yields the following differential rate law:

$$\frac{d[\text{apoCA}]}{dt} = k_d[\text{CA·Zn·L}] \tag{6}$$

Note the difference in assumptions between this reaction and the Michaelis-Menten model; M-M kinetics assumes the enzyme is reformed and that only the substrate is modified, while (5) shows that both the substrate and the enzyme are permanently modified. Recall that [L] >> [CA·Zn], so [L] can be assumed to stay constant at $[\text{L}]_0$, which is substituted into a

rearranged form of equation (4),

$$[\text{CA}\cdot\text{Zn}\cdot\text{L}] = K_{EML}[\text{CA}\cdot\text{Zn}][\text{L}]_0 \tag{7}$$

Carbonic anhydrase can exist in one of three forms: the metalloenzyme CA·Zn, the enzyme-metal-ligand complex CA·Zn·L, or the inactivated enzyme apoCA. Initially, all CA is tied up in the metalloenzyme, and none exists as CA·Zn·L or apoCA. As the activated form of the enzyme gets bound to L and then inactivated,

$$[\text{CA}\cdot\text{Zn}] = [\text{CA}\cdot\text{Zn}]_0 - [\text{apoCA}] - \text{CA}\cdot\text{Zn}\cdot\text{L}, \tag{8}$$

which can be combined with equation (7) to yield

$$[\text{CA}\cdot\text{Zn}] = [\text{CA}\cdot\text{Zn}]_0 - [\text{apoCA}] - K_{EML}[\text{CA}\cdot\text{Zn}][\text{L}]_0 \tag{9}$$

and can be rearranged as follows:

$$[\text{CA}\cdot\text{Zn}] = \frac{[\text{CA}\cdot\text{Zn}]_0 - [\text{apoCA}]}{1 + K_{EML}[\text{L}]_0} \tag{10}$$

Equations (6) and (7) can be combined to give

$$\frac{d[\text{apoCA}]}{dt} = k_d K_{EML}[\text{CA}\cdot\text{Zn}][\text{L}]_0 \tag{11}$$

Therefore, the rate of apoCA formation is first-order with respect to CA·Zn. Combining equations (11) and (10) yields

$$\frac{d[\text{apoCA}]}{dt} = k_d K_{EML}[\text{L}]_0 \frac{[\text{CA}\cdot\text{Zn}]_0 - [\text{apoCA}]}{1 + K_{EML}[\text{L}]_0} \tag{12}$$

Rearranging and integrating,

$$\begin{aligned}\int_{[apoCA]_0}^{[apoCA]_t} \frac{d[\text{apoCA}]}{[\text{CA·Zn}]_0 - [\text{apoCA}]_t} &= \int_0^t \frac{k_d K_{EML}[\text{L}]_0}{1 + K_{EML}[\text{L}]_0} dt \\ &= \frac{k_d K_{EML}[\text{L}]_0}{1 + K_{EML}[\text{L}]_0} t\end{aligned} \tag{13}$$

The left side must be integrated using u-substitution:

$$u = [\text{CA·Zn}]_0 - [\text{apoCA}]_t$$

$$du = -d[\text{apoCA}]$$

To change the integral boundaries,

$u(t = 0) = [\text{CA·Zn}]_0$, since no inactivated enzyme has been formed

$$u(t = t) = [\text{CA·Zn}]_0 - [\text{apoCA}]_t$$

Therefore,

$$\begin{aligned}\int_{[apoCA]_0}^{[apoCA]_t} \frac{d[\text{apoCA}]}{[\text{CA·Zn}]_0 - [\text{apoCA}]_t} &= -\int_{[\text{CA·Zn}]_0}^{[\text{CA·Zn}]_0 - [\text{apoCA}]_t} \frac{du}{u} \\ &= -ln(u)\Big|_{[\text{CA·Zn}]_0}^{[\text{CA·Zn}]_0 - [\text{apoCA}]_t} \\ &= -ln\left(\frac{[\text{CA·Zn}]_0 - [\text{apoCA}]}{[\text{CA·Zn}]_0}\right)\end{aligned} \tag{14}$$

Combining the evaluated integrals from equations (14) and (13) yields

$$ln\left(\frac{[\text{CA·Zn}]_0 - [\text{apoCA}]}{[\text{CA·Zn}]_0}\right) = -\frac{k_d K_{EML}[\text{L}]_0}{1 + K_{EML}[\text{L}]_0} t \tag{15}$$

Assuming steady-state conditions for CA·Zn·L, that is, [CA·Zn·L] is always negligible, then $\frac{[\text{CA·Zn}]_0 - [\text{apoCA}]}{[\text{CA·Zn}]_0}$ is simply the fraction of CA·Zn remaining after reaction time t and can be

referred to as $F_{\mathrm{CA{\cdot}Zn}}$:

$$ln\left(F_{\mathrm{CA{\cdot}Zn}}\right) = -\frac{k_d K_{EML}[\mathrm{L}]_0}{1 + K_{EML}[\mathrm{L}]_0} t \tag{16}$$

Since $-\frac{k_d K_{EML}[\mathrm{L}]_0}{1+K_{EML}[\mathrm{L}]_0}$ is a constant for a given concentration of dipic, equation (16) exhibits linear behavior while the solution contains active enzyme. Therefore, a linear least squares regression procedure can be performed for measurements of $ln\left(F_{\mathrm{CA{\cdot}Zn}}\right)$ over time (until those measurements level off, which indicates that all the enzyme is used up), and the slope, denoted $-k_{obs}$, will be

$$-k_{obs} = -\frac{k_d K_{EML}[\mathrm{L}]_0}{1 + K_{EML}[\mathrm{L}]_0} \tag{17}$$

Taking the reciprocal of equation (17) yields

$$\begin{aligned} \frac{1}{k_{obs}} &= \frac{1 + K_{EML}[\mathrm{L}]_0}{k_d K_{EML}[\mathrm{L}]_0} \\ &= \frac{1}{k_d K_{EML}[\mathrm{L}]_0} + \frac{K_{EML}[\mathrm{L}]_0}{k_d K_{EML}[\mathrm{L}]_0} \\ &= \frac{1}{k_d K_{EML}} \times \frac{1}{[\mathrm{L}]_0} + \frac{1}{k_d} \end{aligned} \tag{18}$$

Again, a linear pattern is observed; measuring k_{obs} at several different dipic concentrations enables one to perform a least squares regression procedure on $\frac{1}{k_{obs}}$ versus $\frac{1}{[\mathrm{L}]_0}$ to determine the slope, $m = \frac{1}{k_d K_{EML}}$, and intercept, $b = \frac{1}{k_d}$. k_d and K_{EML} are thus calculated as follows:

$$k_d = \frac{1}{b} \tag{19}$$

$$K_{EML} = \frac{b}{m} \tag{20}$$

Rate Measurements

While equations (16) and (17) seem to imply that the determination of k_{obs} requires knowing [CA·Zn] or [apoCA] at any point in time, in reality a plot of $f(t)$ vs t, where $f(t)$ is proportional to $F_{\mathrm{CA{\cdot}Zn}}$, will still exhibit the same slope $-k_{obs}$. The Michaelis-Menten kinetics

model states that, for sufficiently large substrate concentrations, the enzyme is fully saturated and the reaction velocity asymptotically approaches a number, $V_{max} = k_{cat}[\text{CA}\cdot\text{Zn}]$, where k_{cat} (the "turnover number") is amount of substrate that can be converted to product by a single saturated enzyme molecule in a given unit of time.[2] Thus for sufficiently large concentrations of substrate, the enzyme catalytic activity fulfills the requirement of proportionality to $F_{\text{CA}\cdot\text{Zn}}$. Ideally, the determination of the enzymatic activity from V_{max}, that is, an enzyme assay, can be done using a substrate which is cheap, readily available, and easily measurable. As Figure 2 shows, carbonic anhydrase happens to hydrolyze para-nitrophenyl acetate (pNPA) to form para-nitrophenol and acetic acid:

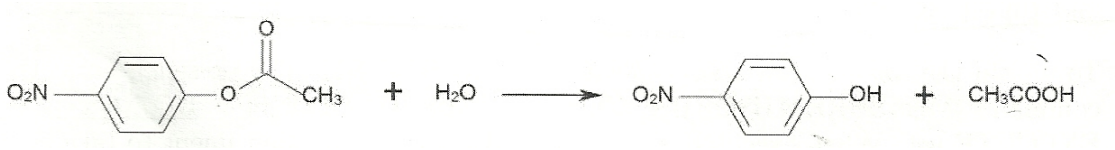

Figure 2: Carbonic anhydrase-catalyzed hydrolysis of para-nitrophenyl acetate (pNPA) to para-nitrophenol and acetic acid[1]

Para-nitrophenol absorbs strongly at 348nm. Recall that Beer's law states

$$A = \epsilon bc \tag{21}$$

where A is the measured absorbance at a particular wavelength, ϵ is the molar absorptivity, b is the measurement path length, and c is the concentration.[3] Therefore, $\frac{dA}{dt}$ is proportional to the change in para-nitrophenol (that is, the reaction velocity). Because the reaction occurs to a nontrivial extent without enzyme catalysis, a baseline correction is needed, $(\frac{dA}{dt})_{uncat}$, which can be obtained either from a solution containing only pNPA or a mixture of pNPA and apoCA (that is, after sufficient time has passed to inactivated all carbonic anhydrase before taking the assay). Thus, equation (16) can be implemented using $ln\left(\frac{\frac{dA}{dt}}{\left(\frac{dA}{dt}\right)_{uncat}}\right)$.

Procedure

This experiment was performed using a Spectral Instruments model SI 440 spectrophotometer, which includes a CCD detector and fiber optic probe. The spectrophotometer was set to detect absorbance at wavelength 348 nm every five seconds over the course of two minutes. Device precison was set to "low," and a blank was obtained and locked. A sample of deionized water was measured to ensure absorbance stayed constant, preventing systematic errors from faulty instrumentation.

A 0.003 M aqueous pNPA solution was prepared by dissolving 25 mg solid pNPA in 1.5 mL acetone within a 250 mL Erlenmeyer flask, then slowly adding 50 mL deionized water while stirring vigorously to prevent precipitation. Assays were constructed by combining 1.7 mL deionized water, 0.3 mL HEPES buffer (0.25 M, pH 8.0), and 1.0 mL pNPA solution.

To measure the uncatalyzed assay velocity, 40 μL deionized water was added to an assay, and the absorbance was measured for two minutes. $\left(\frac{dA}{dt}\right)_{uncat}$ was found to be $3.18 \times 10^{-5}\ min^{-1}$.

Five dipic concentrations were tested: 0.20 M, 0.10 M, 0.05 M, 0.032 M, and 0.016 M. For each [dipic], a 500 μL solution was prepared consisting of 250 μL aqueous bovine carbonic anhydrase (10^{-4} M in 0.125 M phosphate buffer, pH 7.5) and the required volumes of phosphate buffer (0.125 M, pH 7.4) and dipic (0.4 M in 0.125 M phosphate buffer, pH 7.4) needed to dilute to the specified [dipic]. In each run, the carbonic anhydrase and buffer were mixed and placed in a 25°C water bath to equilibrate for several minutes. Dipic was mixed just prior to starting measurements. After one minute, and then again at regular intervals over the course of one hour (or until $\frac{dA}{dt}$ measurements stopped decreasing, signifying full holoenzyme depletion), a 40 µL aliquot of CA/dipic solution was transferred to an assay cuvette, and a two-minute $\frac{dA}{dt}$ measurement was recorded.

Data and Results

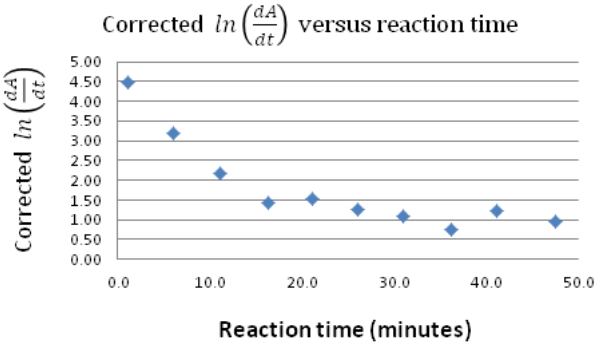

Figure 3: laksjdf askdfj asdkljfhaslkdjfhaklsdjfh asldkfjhasdlkjfhalksdjfhalksdjfh asldkfjhasdlkfjashdlfkajsdhflkajshdf asdfa sdiufhaisdu fas dfasd fasef asefasef asefas efasefkahsdkjfhaskldjfha sdfasldkfjhasdlfkjasdh falksdjfhalskdjfhasdlkfjah sdflkjashdlkfjah sdlkfjh aslkdjfh aslkdjfhalksdjfhalksjdh flkajsdhf alsdkjfash dflaskdjfah .

In the past decade, supervised activity recognition methods have been studied by many researchers, however these methods still face many challenges in real world settings. Supervised activity recognition methods assume that we are provided with labeled training examples from a set of predefined activities. Annotating and hand labeling data is a very time consuming and laborious task. Also, the assumption of consistent pre-defined activities might not hold in reality. More importantly, these algorithms do not take into account the streaming nature of data, or the possibility that the patterns might change over time. In this chapter, we will provide an overview of the state of the art *unsupervised* methods for activity recognition. In particular, we will describe a scalable activity discovery and recognition method for complex large real world datasets, based on sequential data mining and stream data mining methods.

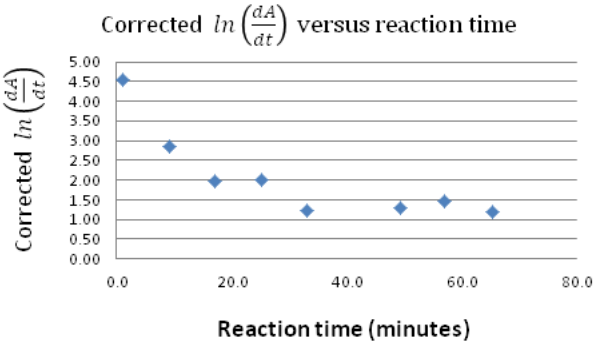

Figure 4: blahblah.

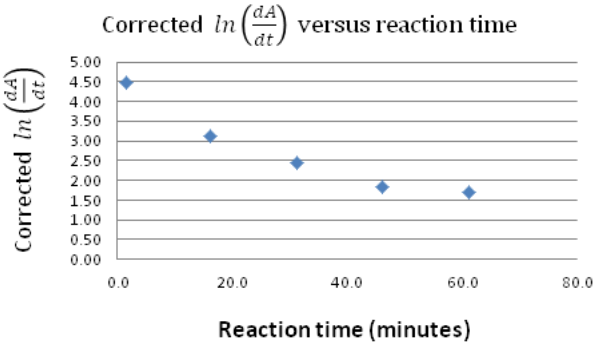

Figure 5: blahblah.

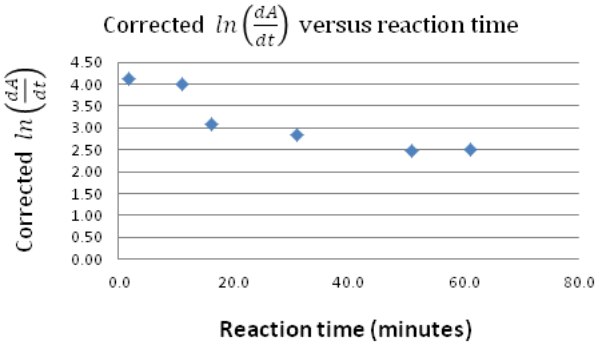

Figure 6: blahblah.

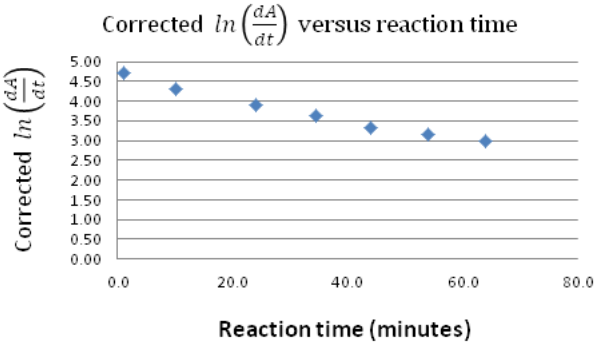

Figure 7: blahblah.

References

(1) Killian, B. J. *Experiments for Physical Chemistry Laboratory*, Summer 2014, Target Copy: Gainesville, **2014**. 45 - 50.

(2) Nelson, D. L.; Cox, M. M. *Lehninger Principles of Biochemistry*, 5th ed.; W. H. Freeman: New York, 2008. 195 - 200.

(3) Harris, D. C. *Quantitative Chemical Analysis*, 7th Ed.; W. H. Freeman: New York, 2006. 381.