Michaelis and Menten at 100: Still Going Strong

Binny J Cherayil

In this article, I provide an elementary introduction to enzyme kinetics at the single-molecule level that highlights the continuing relevance of the contributions made a hundred years ago by Leonor Michaelis and Maud Menten to the study of the mechanisms of enzyme action.

Binny J Cherayil is a Professor at the Department of Inorganic and Physical Chemistry, IISc. His research interests are soft matter, biophysics and statistical mechanics.

1. Introduction

In the annals of science, the year 1913 may not be quite as memorable as that *annus mirabilis* of 1905, the year Einstein published four path-breaking papers in physics. But 1913 was not without its share of interesting scientific developments; it was then, for instance, that Niels Bohr presented his quantum model of the atom, that Robert Millikan determined the charge on an electron, and that Johannes Stark showed the splitting of spectral lines in the presence of an electric field. 1913 also marks the publication of a minor classic in biochemistry: Michaelis and Menten's paper on a mechanism of enzyme action [1]. The paper is important not only for its science, but for what it represented: a research contribution by a woman (Maud Menten) at a time when women in Menten's native Canada were not even allowed to do research. Social conditions for women were, of course, not much better anywhere else in the world, including the US and England, where they had still to be given the right to vote, and could be sentenced to penal servitude (as Emmeline Pankhurst was in England) for demanding that right. So at a period in history when most women were still struggling to earn a measure of respect and fair treatment, Maud Menten, working with the German biochemist Leonor Michaelis, managed to pursue a career in research that was notable for the

Keywords
Michaelis–Menten mechanism, enzyme kinetics, single molecules.

The speed with which enzymes operate in bulk solution tends to have a hyperbolic dependence on the concentration of added substrate.

survival to this day of an equation that now includes her name: the Michaelis–Menten equation.

So what is this equation, and why are we talking about it a hundred years later? The equation deals with the kinetics of enzymes, with the study, in other words, of rates of enzyme activity. Enzymes, as we all know, are large biological catalysts that speed up the conversion of one molecule (the substrate S) to another (the product P), without themselves being changed at the end of the reaction. Because they're critical to a cell's survival, they're studied extensively for insights into their structure, function and mode of action. To characterize an enzyme's activity, we typically make a solution of the enzyme in a buffer, add known amounts of the substrate to it, and then measure how fast we get products in a fixed interval of time. The rate of product formation gives us an idea of how fast the enzyme operates, and if we were to plot this speed against the amount of added substrate, we would end up with a curve that looks something like the one below (*Figure* 1).

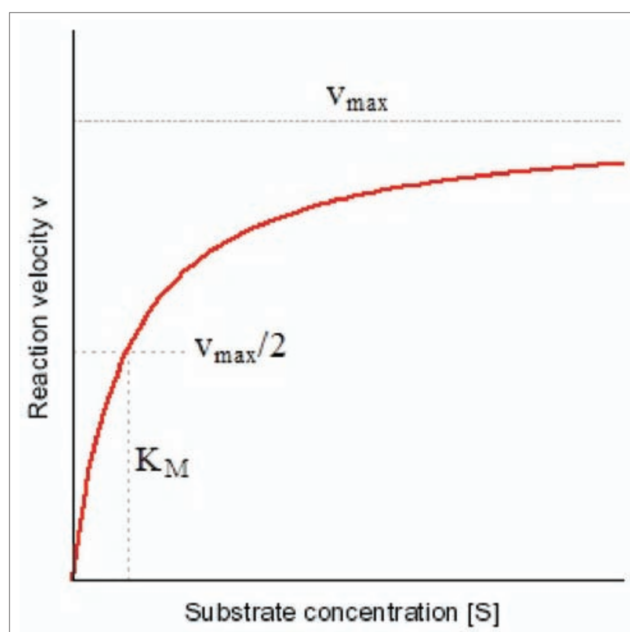
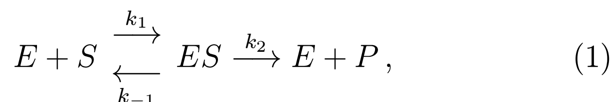


Figure 1. The speed of an enzyme versus the substrate concentration.
Source: http://depts.washington.edu/wmatkins/kinetics/mm1.gif

The curve says the following: the enzyme's speed v starts out slow (because at low substrate concentrations, there's essentially nothing for the enzyme to act on, so very little product is formed), it then increases (because as more substrate is added, more of the previously idle enzymes get to work, and more product is generated in the same interval of time), and finally levels off (because eventually so much substrate has been added that virtually every enzyme molecule is in action, and the addition of more substrate at this stage doesn't lead to more product per time). Michaelis and Menten saw how this behaviour could arise if the mechanism of enzyme action followed these steps:

The Michaelis–Menten mechanism postulates the existence of an intermediate, *ES*, in the pathway between substrate and product.



$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_2} E + P\,, \qquad (1)$$

Here E stands for the enzyme, and ES for an intermediate – the enzyme-substrate complex – that they assumed had to precede the formation of P. Each of the steps in this sequence proceeds at a definite rate, and given these rates it's easy to write down a set of equations for how the concentrations of E, ES and P change with time. These equations are the following:

$$\frac{\mathrm{d}[E]}{\mathrm{d}t} = -k_1[E][S] + (k_{-1} + k_2)[ES] \qquad (2a)$$

$$\frac{\mathrm{d}[ES]}{\mathrm{d}t} = k_1[E][S] - (k_{-1} + k_2)[ES] \qquad (2b)$$

$$\frac{\mathrm{d}[P]}{\mathrm{d}t} = k_2[ES] \qquad (2c)$$

and their meanings should be fairly clear. The first equation says, for instance, that the changes in concentration of E are the net result of the loss of E through the formation of ES, and the gain of E through the decomposition of ES, each process taking place with a characteristic rate constant. The other two equations say similar things about the concentration changes of ES and P.

The Michaelis–Menten equation relates the speed of an enzyme's activity to the amount of added substrate, and is quite successful in reproducing experimentally observed enzyme behaviour.

In general, these equations can't be solved exactly, but if we were to assume – not unreasonably – that the intermediate ES quickly achieves a *steady state*, where $\mathrm{d}[ES]/\mathrm{d}t = 0$ and ES decomposes about as fast as it's formed, then we *can* solve them. The solution leads to an expression for the velocity that is given by [2]

$$v = \frac{v_{\max}[S]}{[S] + K_{\mathrm{M}}} \tag{3}$$

and that has now come to be known as the Michaelis–Menten equation. Here, $v_{\max}$, the maximum velocity, is the product of the catalytic constant k_2 and the *total* enzyme concentration, while K_{M}, the so-called Michaelis constant, is defined by the relation $K_{\mathrm{M}} = (k_{-1} + k_2)/k_1$, and is the substrate concentration at which v is half $v_{\max}$. When plotted, the Michaelis–Menten equation reproduces the shape of the curve in *Figure* 1, and is more or less how a large number of enzymes actually behave in real life. So the scheme proposed by Michaelis and Menten seems to work, and indeed it's been shown to do so quantitatively in numerous examples of enzyme action.

Everything we've said so far about enzymes and enzyme action refers to the behaviour of large *ensembles* of these molecules, typically of the order of Avogadro's number. Should we expect anything different if we were to look at the behaviour of a *single* enzyme molecule?

2. Watching Single Molecules at Work

Till not so long ago, this question would have been purely academic, because it would have been essentially impossible to make any meaningful measurements on just one molecule (if that molecule happened to be in a liquid or other condensed phase), but in the last few years, since about the late 1980s, a slew of new and improved spectroscopic and detection methods have made such measurements almost routine. How was this achieved?

In most problems of chemical and physical characterization, spectroscopy is the tool of choice, and spectroscopy on single molecules basically involves the same two steps that we use to record the spectrum of bulk samples, which means doing two things: (1) shining light on a sample to induce a transition between distinct quantum states of the system, and (2) detecting this transition in some fashion. The only difference is that in single-molecule spectroscopy it's necessary to ensure that just one molecule occupies the region of space illuminated by the light source, and that the signal of interest can be distinguished from the signals of everything else that happens to be in the vicinity.

Single-molecule spectroscopy involves the same two steps used to record the spectrum of bulk samples: (1) shining light on the sample to induce a quantum transition, and (2) detecting the transition in some fashion. But the sample volume must be small enough that just a single molecule is found in it.

To excite just a single molecule with light, we typically have to focus the light onto a very small volume, of the order of 1–100 μm^3, and to use very low concentrations of the solution that contains the molecule of interest, of the order of 10^{-10} M. It's easy to show that at such concentrations, if the probe volume is, say, $10 \mu m^3$, then there will be essentially just a single molecule in the volume.

Even when it is possible to excite a single molecule with light, how would we know that we had actually excited it? Well, if the molecule happened to be fluorescent, i.e., if it emitted light of a different wavelength after excitation by the incident beam, then one way to follow what happens to it subsequently is to monitor this fluorescence emission. Of course, this is generally not too easy, because not only is the amount of light emitted small, it can and often does suffer interference both from the light emitted by other molecules in the probe volume and from other processes (like Rayleigh and Raman scattering) that occur there. Nevertheless, these problems *can* be overcome. What this means, in the final analysis, is that single molecules in condensed phases can be excited and detected.

Xie *et al* devised an assay that allowed a single molecule of the enzyme β-galactosidase to be observed in real time as it catalyzed the conversion of RPG to resorufin.

Let's now turn to an experiment [3] that actually followed the activity of a single enzyme (the enzyme β-galactosidase, to be specific), to get some idea about how all this works in practice.

3. Single-Molecule Activity of β-Galactosidase

β-galactosidase is a tetramer that acts on the sugar lactose and converts it to glucose and galactose. To follow the reaction of just one molecule of the enzyme with a substrate, a group in Cambridge, Massachusetts, led by X S Xie, did the following. They attached the enzyme to a polystyrene bead and then covalently linked the bead to a glass cover slide, effectively immobilizing the enzyme. They then took galactose and modified it by attaching a polyaromatic species to it, producing a molecule that's best referred to by its initials, RPG, since its full form is something of a mouthful. RPG is also acted on by β-galactosidase, one of the products of the reaction being resorufin, a fluorescent molecule. Xie and co-workers illuminated the immobilized enzyme with light in the presence of known concentrations of RPG. As the enzyme acted on one molecule of RPG, it generated one molecule of resorufin, which emitted a burst of fluorescence that was captured on film, before meandering out of the probe volume by simple diffusion. This whole process was followed over a long period of time. The experimental set-up they used is shown in *Figure* 2.

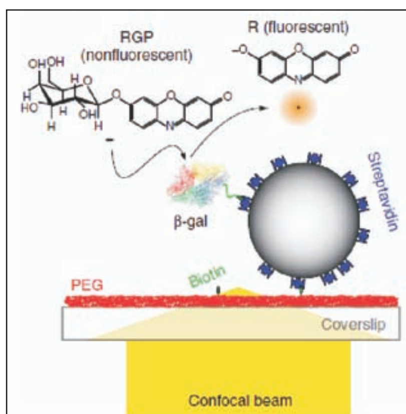


Figure 2. A schematic drawing of the experimental set-up used to observe single enzymes in action.
Reproduced by permission from [3], Nature Publishing Group.

What they finally measured was a series of spikes of fluorescence intensity, representing the instants of time when resorufin was released into the probe volume following catalytic action by β-galactosidase (see *Figure* 3).

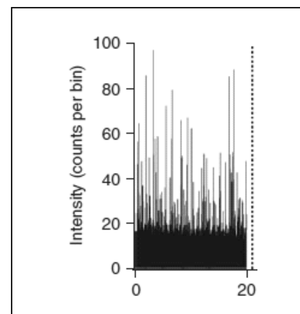


Figure 3. Random spikes of fluorescence.
Reprinted by permission from [3], Nature Publishing Group.

So first of all, in this single molecule spectroscopic measurement, what was recorded was not a spectrum (as that term is conventionally understood – a plot of intensity vs. wavelength, for instance), but a *stochastic time trace of light intensity* originating from a product molecule. What does this time trace tell us about single molecule activity?

4. Taking the Bus: Randomness Rules

To answer this question, it helps to think about a somewhat more familiar situation, one that all of us have almost certainly dealt with at some point in our lives. So forget about enzymes for the moment, and consider the experience of waiting for and catching a bus in one of our Indian cities, say Bangalore.

Imagine, then, that you're on your way back home from somewhere. The bus you need to take is scheduled to arrive at your particular bus-stop at, say, 5:15 pm, so you reach there at 5:10. On this particular day, let's suppose that you have to wait 9 minutes for the bus to arrive. The following day, you head back to the bus-stop at the same time, but now your wait is 7 minutes, and the day after that, it's some other interval of time that you faithfully record in your diary. The reason you're making these observations is to determine whether the Bangalore Metropolitan Transport Corporation (BMTC) is doing a good job of providing on-time bus service. After a month of noting down the times you've had to wait for the bus, you plot your data in the form of a graph, which might look like *Figure* 4.

Single-molecule measurements of enzyme activity often record the time trajectories of random flashes of light.

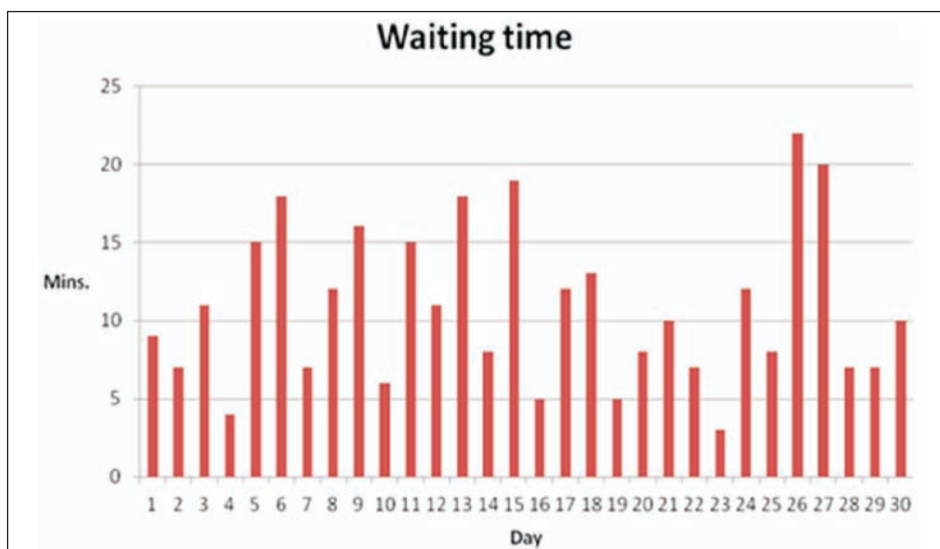


Figure 4. Taking the bus: a lot like waiting for Godot.

This is one of the simplest and most direct ways of depicting the data, but often, rather than showing the exact chronology of your waiting times at the bus-stop, it can be more informative to show instead the number of times you've had to wait 5 minutes or less, the number of time you've had to wait between 5 and 10 minutes, the number of times you've had to wait between 10 and 15 minutes, and so on for all the other relevant intervals of time. Now your graph looks like *Figure* 5.

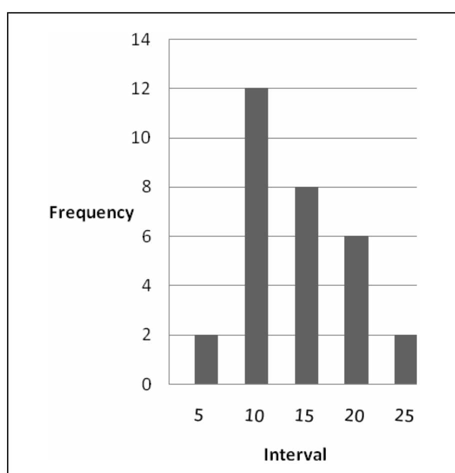


Figure 5. Bus waiting times represented in histogram form.

This graph is a frequency histogram. Recording your bus-stop experiences in the form of a frequency histogram can tell you quite a few things. It tells you at a glance that most of the time you've had to wait about 10–12 minutes to get the bus. And it tells you that you never had to wait more than 25 minutes for the bus. All this could be useful information if you needed to decide how long you could afford to stay back at whichever place you happened to be before heading out to the bus-stop to catch the bus.

A waiting time distribution is a useful way of summarizing data on the performance and efficiency of bus services.

A very minor modification of this graph provides yet another perspective on the data. The modification is to divide each of the y-axis values by a constant amount, this amount being the total number of days for which you've made observations of the waiting time. By doing this, the frequency histogram is normalized, and this means that the area under the normalized graph is unity. The normalized frequency histogram looks like *Figure* 5, but the y-axis now extends only up to 1, and it represents a fractional occurrence. We refer to this histogram as a *waiting time distribution*, $f(t)$. The waiting time distribution is useful because it tells us the *fraction* of the times that the wait for the bus has been between 0 and 5 minutes during the course of a 30-day observation period, the *fraction* of the times that it's been between 5 and 10 minutes, and so on, and because these are fractions that add up to unity, they effectively represent the *probability* of having had to wait these given intervals of time. Now, looking at this graph, you could make *probabilistic* statements about the BMTC bus service. In particular, you could ask how long you could *expect* to wait for a bus in the future. You could ask, in other words, what the *expectation* of your waiting time is. The expectation of the waiting time is just another way of saying the *average* or *mean* of your waiting time, a very familiar notion. The mean waiting time is easily

In some circumstances, the waiting time distribution for bus arrivals that stick to no particular schedule can be calculated fairly easily.

calculated from the following formula:

$$\bar{t} = \int_0^\infty \mathrm{d}t\, t\, f(t) \tag{4}$$

and from the waiting time distribution, you can decide that the mean or expected waiting time is about 11 minutes.

If your mean waiting time really were of the order of 10–11 minutes, you'd have a very high opinion of the BMTC. But in fact you probably feel that waiting for one of its buses is a lot like waiting for Godot, and that the best that can be said about the coming and going of these buses is that some fixed number reach the bus-stop during some interval of time. In other words, if the schedule of bus timings distributed by the BMTC announced one bus at, say, 5:15 pm, and another at 5:45 pm, you wouldn't be too confident of actually seeing any buses arriving at those times. What you could hope for is to see two buses *some* time between 5:00 pm and 6:00 pm. The waiting time distribution for this random (but more realistic) schedule turns out to be not too difficult to construct [4]. Let's see how.

Suppose that the number of buses per hour is fixed, and is denoted by λ. λ could be a quantity like 2 per hour, or 1 every 15 minutes, or 5 per day, or 10 per week, or anything else. All we require is that this rate be fixed. We don't require that the buses show up at precise and pre-ordained times. In these circumstances, how *probable* is it that there is a bus between two given instants of time, say, t and $t + \mathrm{d}t$ if we regard $\mathrm{d}t$ as being very very small, and if we start our clocks at some arbitrary initial time $t = 0$? To answer this question, let's divide the time interval t into a large number of sub-intervals n of equal duration. Each sub-interval clearly lasts t/n units of time. The first time we see a bus between t and $t + \mathrm{d}t$ is if we *don't* see a bus in the first

n sub-intervals and *then* we see a bus in the next interval lasting dt units of time. To estimate the probability that a bus does *not* arrive during the very first sub-interval, lasting t/n units, we argue as follows. Suppose n is large, so that t/n is small; then to a reasonable approximation the probability that it *does* arrive is simply the average rate λ, multiplied by the length of the given interval, t/n. Once we have an estimate of this probability, we can immediately say that the probability that there is *no* bus during this time is simply $1 - \lambda t/n$. And from this, we can also immediately say that the probability that there is no bus in n successive such intervals is $(1 - \lambda t/n)^n$. In other words, the probability that there is no bus between 0 and t is $(1 - \lambda t/n)^n$. Suppose we now ask for the probability that there is a bus in the very next interval of time, which we assume is dt. This probability is λdt. And therefore the probability that the first appearance of the bus takes place between t and $t + \mathrm{d}t$ is

$$P(\lambda, t)\mathrm{d}t = (1 - \lambda t/n)^n \lambda \mathrm{d}t. \tag{5}$$

When n is very large as we've assumed, the term in parentheses on the right-hand side of (5) tends towards an exponential. That is,

$$P(\lambda, t)\mathrm{d}t = \lambda \exp(-\lambda t)\mathrm{d}t\,. \tag{6}$$

This is called the Poisson interval distribution [4], and processes that lead to such distributions are called Poisson processes. The distinguishing characteristic of such processes is that some event occurs at a fixed rate and that in some chosen interval of time the probability that the event happens is small. A number of different kinds of random situations can be described as a Poisson process or with an exponential distribution, including the time it takes a bank customer waiting in a queue to reach the teller's window.

The Poisson interval distribution is probably a reasonably good description of the distribution of bus arrival times in Bangalore. It is characterized by an exponential dependence on the waiting time.

The Poisson interval distribution is actually a waiting time distribution; t in this formula is the time you have

Stochastic time traces of fluorescence intensity in measurements of single enzyme activity show similarities with bus arrival times in Bangalore.

to wait to see a bus for the first time when the average rate of arrivals is λ, and $P(\lambda, t)\mathrm{d}t$ – a probability – can be interpreted as the fraction of the times you've had to wait t units to get a bus. And being a waiting time distribution, you can easily verify by integration that the distribution is indeed normalized to unity, as it's supposed to be.

5. Interpretation of Single-Molecule Data

It will probably have become apparent by this time that the fluorescence bursts in *Figure* 3 are a lot like Bangalore's unreliable buses – they stick to no particular schedule, but show up more or less when they like. And since we found it useful to analyze the bus data in terms of frequency histograms and waiting time distributions, let's do the same for the enzyme data. When we construct the corresponding waiting time distribution, we get *Figure* 6.

The symbols are the actual data points, and the lines are guides to the eye. Each set of data corresponds to a given substrate concentration $[S]$, the lowest curve corresponding to the highest $[S]$ and the highest curve

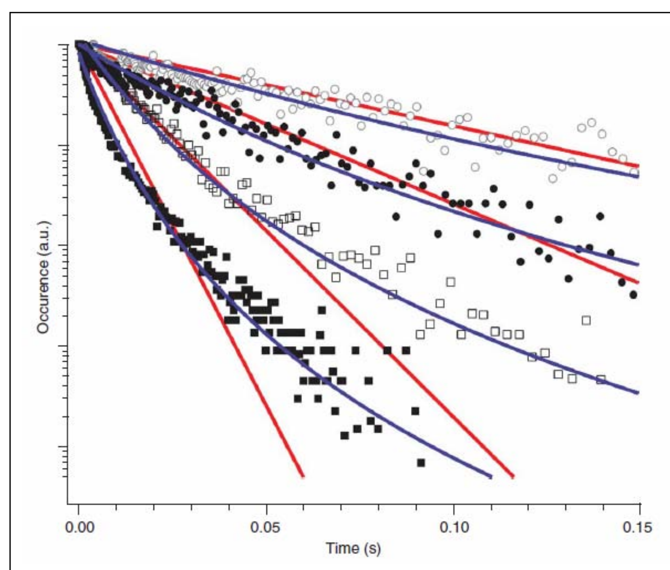


Figure 6. Waiting time distributions for enzyme activity at four substrate concentrations. Reproduced by permission from [3], Nature Publishing Group.

to the lowest $[S]$. The graphs have been drawn on a log-linear scale, so one thing is obvious: at low concentrations, the waiting time distribution is exponential (the curve is essentially a straight line), and at higher and higher concentrations, the curves become increasingly non-exponential. What are we to make of these results?

At low substrate concentrations, enzymatic reaction rates tend to be diffusion-controlled.

(i) ***Low Substrate Concentration***

Let's look at each of these limiting behaviours in turn. We know from our study of bus arrival times that an exponential waiting time distribution is the signature of a Poisson process. Remember a stochastic process is Poissonian when it is characterized by a constant rate of occurrence of some event, each occurrence of the event in a given interval of time having a low probability. If we apply these considerations to the enzyme, we would have to conclude that when there aren't too many substrate molecules around, they tend to reach the site of reaction one at a time, at irregular intervals. So the reaction rate is effectively controlled by the time it takes individual substrate molecules to execute a random walk from wherever they are to the enzyme's active site. That time, *on average*, is a constant, and if it happens to be much longer than the time it actually takes to make a molecule of P from S, then the time between the generation of one product molecule and the next is also random. But because the average rate of product generation is fixed, these times tend to follow an exponential distribution.

We can actually put some mathematical flesh on the bones of this explanation. Let's return to the Michaelis–Menten mechanism. The equations we introduced there cannot be applied directly to the single molecule situation because when we deal with just a single molecule, it no longer makes sense to talk about concentrations. However, we can speak meaningfully about the *probability* that the enzyme is in one or other of the states

Rate equations analogous to the ensemble equations (2a)–(2c) can be written down for enzyme activity at the single-molecule level.

corresponding to E and ES. But when we do this we also have to recognize something that was only implicit in the mechanism of (1), that the regeneration of E from ES is not necessarily instantaneous, but tends to proceed (if only fleetingly) through an intermediate E^0. A complete catalytic cycle should really be thought of as proceeding through two half reactions, one corresponding to $E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_2} E^0 + P$, and the other to $E^0 \xrightarrow{k_3} E$, which resets the enzyme to its original condition. Equations (2a)–(2c) really describe the first half-reaction. At the ensemble level, the existence of E^0 can be ignored because in general $k_3 >> 1$, so E^0 is indistinguishable from E, and the two half-reactions effectively reduce to (1). But at the single-molecule level, where the interest is in reaction times and their distribution, the inclusion of E^0 in the catalytic pathway makes it possible – as we'll see – to decide *when* a molecule of S is likely to have been converted to P. Furthermore, because there's just one molecule of the enzyme, not much of the substrate ever gets consumed, even when $[S]$ is small, and its concentration therefore stays effectively the same for the entire period that the enzyme acts. So for all intents and purposes the factor of $k_1[S]$ in (2a)–(2c) is itself a constant, and we can denote it k_1^0 and think of it as a 'pseudo-first order' rate constant. Now, by analogy with the ensemble equations describing the Michaelis–Menten mechanism, we can write down a corresponding set of single-molecule equations for the first half-reaction; they are [5]

$$\frac{\mathrm{d}P_E(t)}{\mathrm{d}t} = -k_1^0 P_E(t) + k_{-1} P_{ES}(t) \ , \tag{7a}$$

$$\frac{\mathrm{d}P_{ES}(t)}{\mathrm{d}t} = k_1^0 P_E(t) - (k_{-1} + k_2) P_{ES}(t) \ , \tag{7b}$$

$$\frac{\mathrm{d}P_{E^0}(t)}{\mathrm{d}t} = k_2 P_{ES}(t) \ , \tag{7c}$$

where $k_1^0 = k_1[S]$, as stated.

That's not all; because we're now talking about probabilities, we need to include one other equation, and that is

$$P_E(t) + P_{ES}(t) + P_{E^0}(t) = 1, \qquad (7d)$$

which simply says that at any given time the enzyme *must* be in one of its three possible states. We also have to specify a set of initial conditions, and these are $P_E(0) = 1$, $P_{ES}(0) = 0$ and $P_{E^0}(0) = 0$, which in turn say that at the time we start making measurements, before there's been enough time for the substrate to react, the enzyme exists wholly in the state E. It turns out that (7a)–(7d), unlike the ensemble equations, *can* be solved exactly, and this means we can obtain the complete expressions for each of the above probabilities. But the probability we really need is the one for the enzyme to be in the state E^0, because once the enzyme is in this state, the intermediate ES must have decomposed and formed the product P. And once we have $P_{E^0}(t)$, we can actually calculate the waiting time distribution. How do we do that?

The single-molecule Michaelis–Menten equations of (7a)–(7d) can be solved for the distribution of the time intervals between successive catalytic events.

Well, suppose the enzyme is in E^0 at time t, and is still in E^0 at time $t + \delta t$, with δt infinitesimal. That means a product molecule has been formed between t and $t + \delta t$. In other words the probability of seeing a product molecule between these two times is $[P_{E^0}(t + \delta t) - P_{E^0}(t)]/\delta t$, which means that [5]

$$f(t) = \frac{\mathrm{d}P_{E^0}(t)}{\mathrm{d}t} = k_2 P_{ES}(t) \ . \qquad (8)$$

We can calculate this $f(t)$, and it turns out that it is exponential at *all* concentrations of the substrate (more precisely, it is the difference of an exponential rise and an exponential decay that reduces to a single exponential in the limit $k_2 << k_{-1}$ that describes the steady state [6].) So at low $[S]$, we can treat the single molecule Michaelis–Menten equations as we have done above and get the

exponential form we anticipated on physical grounds. But this approach clearly doesn't explain why $f(t)$ is *non*-exponential at high substrate concentration.

(ii) *Single-molecule Reaction Speeds*

Before we get to that question, there's one other issue that's worth addressing at this stage. We've seen that at the ensemble level the expression for the enzyme velocity is given by the Michaelis–Menten equation. At the single-molecule level, should we expect to see any other expression for the velocity? The reason for asking this question is the following: An ensemble experiment essentially records the average of the properties of a huge number of molecules in different conformational states at *one* instant of time. A single-molecule measurement, on the other hand, monitors the properties of one molecule over a long period of time. A moment's reflection should make it apparent that every possible conformation that's represented in the ensemble is a conformation that this single molecule can assume if it's given a long enough time. And so if we averaged single molecule properties over time, we should get the same results as averaging over an ensemble of properties at any one instant of time. So the single-molecule measurements should, if carried out long enough, lead eventually to the Michaelis–Menten equation. But how can a waiting time distribution – which is what we record in a single molecule experiment – lead to an equation for the enzymatic velocity as a function of the substrate concentration?

The equivalence of ensemble and time averages allows an equation for the speed of a single enzyme to be derived from the expression for its waiting time distribution.

The answer lies in calculating the *mean* or *expectation* of the waiting time distribution, just as we calculated the mean or expected waiting time in the example of the bus arrivals. Once we've determined this quantity, $\bar{t}$, the reciprocal is clearly a rate, and if we identify – at least provisionally – this rate with the enzymatic velocity, then in principle we have a way of reconstructing

the Michaelis–Menten equation theoretically and experimentally. Theoretically, it's easy to calculate $\bar{t}$ from the expression for $f(t)$ we derived earlier. When we do this and take the reciprocal, we find indeed that we recover the Michaelis–Menten equation [6]. A similar calculation using the experimental $f(t)$ also leads to a result consistent with this equation. A nice comparison of the theoretical and experimental velocity curves is shown in [3].

The Michaelis–Menten equation can be recovered from the expression for the waiting time distribution that characterizes single enzyme activity.

So we've found that the theoretical single-molecule scheme above in which we use probabilities in place of concentrations in the ensemble reaction mechanism is consistent with the ensemble Michaelis–Menten equation, but fails to account for the non-exponentiality of the waiting time distribution at high substrate concentrations. So let's get back to this other piece of the puzzle.

(iii) *High Substrate Concentration*

What we need to do is come up with a scenario that might explain why at these concentrations the single-molecule enzymatic reaction is no longer Poissonian. As before, it helps to consider this question in terms of the bus arrival problem. For that problem, it seems clear that the only time we would expect the waiting time distribution to deviate from Poisson behavior is if the buses *didn't* arrive at a constant average rate, because that's what we said was the defining characteristic of a Poisson process. That would happen if the BMTC decided to make things even more difficult for its long-suffering public by having different average rates of arrival during different intervals of time, say, λ_1 per hour during the first 15 minutes after the hour, λ_2 per hour during the second 15 minute interval, and so on for the other 15 minute intervals. It's then possible to show that the waiting time distribution does in fact become non-exponential.

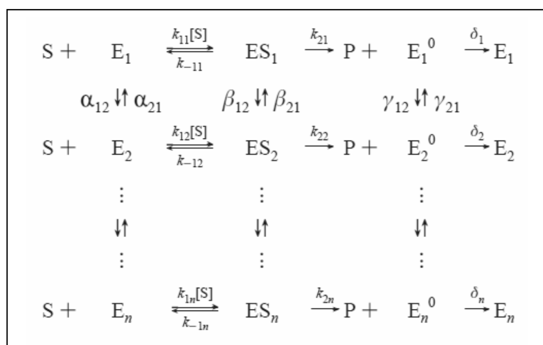


Figure 7. A multi-state generalization of the Michaelis–Menten mechanism.
Reprinted by permission from [6], American Chemical Society.

To translate this idea into the language of enzymes, we need to assume, first of all, that when there are a large number of substrate molecules in the solution, the enzyme doesn't have to wait for one or other of them to wander into its active site – there will almost always be a molecule of S already there. So the rate of reaction is now determined by how fast the enzyme chews up S and spits out P; the time it takes to do that has to be random; if it weren't, we'd see molecules of P coming off the enyzme in assembly-line fashion, at regular intervals, which we don't. So the reaction rate keeps changing. Kinetic processes where rate constants are not really constant at all but change with time are said to be governed by *dynamic disorder*. For an enzyme, dynamic disorder might plausibly occur if the enzyme existed in a number of distinct conformational states, in each of which its catalytic rate was different. A schematic picture of what this means is shown in *Figure* 7.

Non-exponentialities in the waiting time distribution of β-galactosidase could plausibly occur if the enzyme were assumed to be governed by dynamic disorder, i.e., if it fluctuated between states with different reactivities.

It's actually possible to write down a set of single-molecule equations for this scheme, along the lines of (7a)–(7d). It's even possible to solve them too. The equations are complicated, however, and so is the solution, but when it's assumed that the interconversion of the enzyme between its conformational substates is *slow*, they start to look a lot like the original single-molecule Michaelis–Menten equations; the only difference is that now they involve an average over the *distribution* of the different k_2 values. It turns out that if this distribution

is chosen in a special way, a waiting time distribution can be calculated that reproduces the experimental trends almost quantitatively. The mean waiting time can be calculated as well, and if its reciprocal is identified with a velocity as before, the resulting equation is of exactly the same form as the Michaelis–Menten equation (3), but again the equation involves an average over k_2 values, so the meaning of the Michaelis constant is a little different from its ensemble definition.

This brings up another question. Why do the reaction rates fluctuate the way they do, and is it possible to provide a more microscopic molecular picture of dynamic disorder?

6. Chain Fluctuations

We already have some idea of how reaction rates could change with time. They could do so if the conformation of the enzyme as a whole fluctuated with time, passing through a series of apparently locally stable sub-states that had their own characteristic reaction rates. But do they?

This question too was answered by single-molecule spectroscopy, and as it happens by the same group that looked at single enzyme catalysis [7]. They did this essentially by monitoring the rate at which an electron located on one part of a protein (a complex between fluorescein and anti-fluorescein, to be precise) migrated to another part of the protein under illumination by light of a definite colour. This rate actually says something about the *distance* x between the electron-donating and electron-accepting regions of the molecule because the rate is a sensitive function of the distance; in fact, the dependence of the rate on distance is known quite precisely. So a measurement of one is effectively a measurement of the other. And based on this equivalence, Xie *et al.* showed that in solution at room temperature a single enzyme is a fluctuating protean (pun intended) object,

Reaction rate fluctuations could arise from fluctuations in an enzyme's conformation.

Experimental measurements of electron transfer rates between donor and acceptor sites on a large protein support the idea that such molecules fluctuate randomly betwen distinct conformations.

constantly changing its shape as it flips from one conformation to another across a very wide range of timescales that extend at least as far down as milliseconds.

Of course, we already knew that (from NMR, for instance), but considerably more quantitative information can be teased out of the distance fluctuation data. What the Cambridge group did was to take the raw data on $x(t)$ and construct a time correlation function from it. That is, they measured the distance at some initial time $x(0)$, then multiplied it by the distance measured at some later time $x(t)$, to form $x(0)x(t)$, and then averaged this quantity over many such measurements, to yield $C(t) \equiv \langle x(0)x(t) \rangle$. The reason for recording the data in this fashion is that the time correlation function, as its name suggests, says something about correlation effects – how much the value of some measured quantity *remembers* the value it had earlier. That information can be useful in providing insights into the dynamics that underlie the correlations.

The experimental time correlation function is shown in *Figure* 8 [7]. Like the waiting time distribution it's highly non-exponential, and we're now confronted with another challenge: how to explain these data.

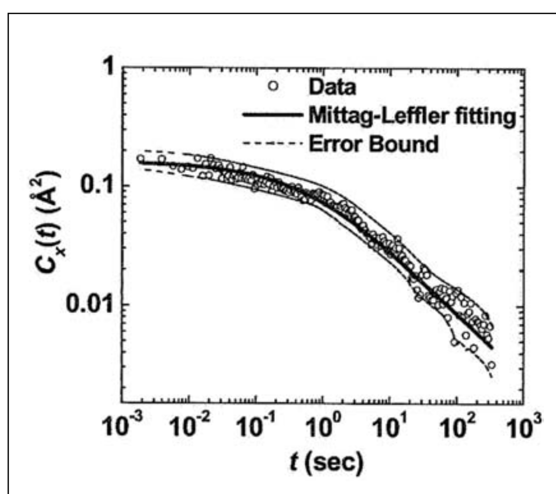


Figure 8. The time correlation function of distance fluctuations.
Reprinted by permission from [7], American Physical Society.

(i) *Modelling Distance Fluctuations*

As a first pass at the problem, we could try to explain the data by first writing down an equation for the way we think the distance between two groups on the protein changes with time. Since the protein is a large mechanical object, and since in solution its structure doesn't change all that much from moment to moment (unlike large organic polymers, which change their structures all the time), it's not too much of a stretch to think of its separate parts as being held together by springs that can only move so far apart. With that image in mind, the equation for fluctuations in the distance of one such spring is easy to write down; it is given by

$$m\frac{\mathrm{d}^2x(t)}{\mathrm{d}t^2} = F \qquad (9)$$

The motion of a particle that is acted on by a random force in a harmonic well provides a paradigm for describing distance fluctuations in proteins.

which of course will be recognized as Newton's law – force is equal to mass times acceleration. The simplicity of the equation is deceptive – it's actually extremely complicated because the expression for the force (taking into account all the little nitty-gritty details that its realistic description requires) is extremely complicated. Without lots of computing power, it's generally impractical to try to relate F to these details, but we can adopt a strategy that will at least allow us to get off the mark, to use a cricketing metaphor; in the equation for x, we replace the force F by a force that is much simpler to work with but that nevertheless captures what we hope is the right physics. The 'right physics' has to do at least two things: ensure that F has a random part (because $x(t)$ is random, and the two are linearly related to each other), and that it has a non-random part that provides a measure of the resistance a particle feels when it tries to move past the other particles in the medium All these considerations suggest that the equation should have the form

$$m\frac{\mathrm{d}^2x(t)}{\mathrm{d}t^2} = -\zeta\frac{\mathrm{d}x(t)}{\mathrm{d}t} - \frac{\mathrm{d}U(x)}{\mathrm{d}x} + \theta(t)\ . \qquad (10)$$

In complex fluid environments, a particle's motion is often better described by a generalized Langevin equation than by a Langevin equation.

Here U is an *effective* potential, i.e., a potential which is not necessarily realistic, but does what a realistic potential does without any of the fuss and bother. A certain amount of optimism is needed to believe that such potentials exist, but the choice $U \propto x^2$ seems to do the trick (basically because it keeps pairs of protein residues reasonably close to each other, as is required of a structure that on average is fairly well-defined). ζ is a friction coefficient, and the term in which it appears models the effects of resistance (which are often sufficiently large that it's safe to throw away the inertial term on the left-hand side.) $\theta(t)$ is the random part of F, and we'll say more about it later.

Equation (10) is a Langevin equation, and with the ma term discarded, it's about as simple as they come, perhaps a little *too* simple. It actually ignores an important effect: time variations in the frictional resistance. Once that effect is brought into the picture (and statistical mechanics tells us that there's a special way to do this, which we won't discuss here), the equation for $x$that we have to deal with is the following:

$$x(t) \propto \int_0^t \mathrm{d}t' K(t-t') \frac{\mathrm{d}x(t')}{\mathrm{d}t'} + \theta(t) \qquad (11)$$

which is a *generalized* Langevin equation, to get technical. Here $K(t-t')$, called a memory function, is a time-dependent generalization of ζ, and for reasons that are beyond the scope of this article, its form is actually determined by the form of the random term $\theta(t)$. So the choice of $\theta(t)$ is quite crucial to whatever information we can extract from the above equation. Unfortunately, to go into all the considerations that determine this choice sensibly would also take us too far afield. Suffice it to say that $\theta(t)$ seems to correspond to a stochastic process called *fractional Gaussian noise* (fGn, for short.)

Now 'noise' is a familiar concept, but it's not so easy

to define quantitatively. When is a set of sounds noise, for instance, and not music? The principal difference is of course that music is characterized by regularities, noise is not. The way we express these irregularities is to say something about the sound made at one instant of time and the sound made a short time afterward. If we were listening to music, we would expect sounds that succeeded each other quickly to be correlated. When we hear noise, however, what we're struck by is the lack of such correlation. And the simplest form of noise is one where there is compete lack of correlation between sounds. This kind of noise is called *white* noise. There are no characteristic frequencies that can be picked out in white noise, the way they can be in a Mozart concerto, for instance, or in a number by Led Zeppelin. Fractional Gaussian noise is another kind of noise in which there is some degree of correlation between sounds but not as much as when the sounds are part of a musical composition. Now noise doesn't have to refer exclusively to sounds, but it can refer to other kinds of excitations as well, including the kinds of mechanical excitations that lead a molecule like an enzyme to change its conformations at random.

A generalized Langevin equation in which the random force corresponds to fractional Gaussian noise leads to an expression for the time correlation function of the distance fluctuations in a protein that fits experimental data very well.

If $\theta(t)$ is fGn, then for technical reasons we require that

$$K(t-t') \propto |t-t'|^{\alpha} \ , \tag{12}$$

where α is a number less than 1. With these simplifications to Newton's laws, we can calculate $C(t)$. We find that $C(t)$ is given by a generalization of the exponential function called the Mittag–Leffler function. The actual expression is the following:

$$C(t) \ \propto \ E_a(-t/t_0) \ , \tag{13}$$

where $E_a(-x)$ is the Mittag–Leffler function, and is defined as $E_a(-x) = \sum_{n=0}^{\infty} (-x)^n/\Gamma(an+1)$, $\Gamma(\cdots)$ being the gamma function, while t_0 is some parameter with

Kramers' barrier crossing model is a good paradigm of chemical reactions in condensed phases.

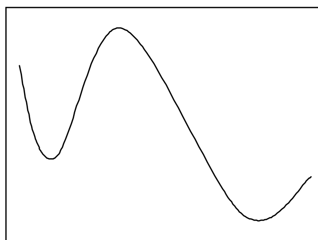
the units of time. This expression for $C(t)$ provides a very good fit to the experimental curve [7].

(ii) *Fluctuations and Reactivity*

If it is the case that the fluctuations of the enzyme are driven by fractional Gaussian noise, what effect does that have on the reactivity of the enzyme, in particular the waiting time distribution? To answer this question, we need to be able to formulate a model of reaction rates in which we can introduce random time varying processes, such as fGn. There is a way to do this that was developed by Kramers [8] a long time ago, in the 1940s, and that method is still applicable today. Kramers suggested that reaction rates can be modelled by the motion of a particle that travels between two wells over a barrier under the action of random forces from a solvent environment, as in *Figure* 9.

If the particle were initially located in the left well, and if the barrier between the two wells was fairly large, the particle would tend to stay there, all the while being jostled by its solvent neighbours. But every now and then, it would suffer an especially energetic encounter, and that would then give it enough of a boost to get to the other side. Such encounters of course occur at random, much as they do when we're walking down a crowded street, and someone brushes past us out of the blue with sufficient force to send us over a protective railing and into a ditch. So barrier crossing is a stochastic process, and Kramers argued that it was exactly the kind of process that could stand as a paradigm of chemical reactions in solution. After all, what happens in a chemical reaction; two reactants move around at random amongst a crowd of other molecules, usually doing nothing when they bump up against each other for want of the requisite energy, but every once in a while forming a new molecule when they do get that energy from somewhere. In general then, there is a waiting time

Figure 9. Kramers' barrier crossing model.



between putting reactants together, and seeing a product. Kramers showed from his double well model that if the particle is subject to white noise (and its motion is described by a Langevin equation) this waiting time distribution could be calculated. The distribution he found from the calculation turned out to be – any guesses? – an exponential. So the barrier crossing model is effectively a model of a Poisson process.

But what if the noise driving the barrier crossing were not white but fractional Gaussian, as we think it would be if the chemical reaction took place at the active site of an enzyme, where the effects of the fluctuations are expected to last rather longer than if they came from the random motions of simple solvent molecules. It turns out that the Kramers model for a reaction driven by fGn can still be solved for the waiting time distribution (but only after considerable effort [9].) This new distribution is *not* an exponential; it also involves the Mittag–Leffler function, and it has exactly the right shape to reproduce the high $[S]$ region of *Figure* 6, as shown in *Figure* 10.

We're almost at the end, but one last issue remains unaddressed. Although we've made something of a case for fGn as the kind of noise that governs enzyme fluctuations, we haven't really suggested why that should

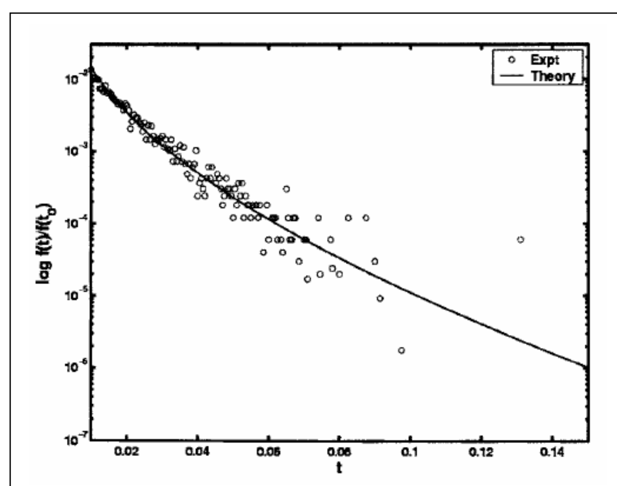


Figure 10. Comparing theory with experiment.
Reproduced by permission from [9], American Institute of Physics.

be the case. There are, after all, other kinds of noise that could conceivably be responsible for the effects we see, but apparently they are not. We think the reason lies in the nature of the protein itself, which in solution can't really fling itself around with complete abandon, like a dervish, but must move in a more controlled fashion, via the kinds of normal modes that dictate the way a molecule of, say, CO_2, vibrates. Protein normal modes are, of course, considerably more complicated, but put them all together, and the fluctuations they lead to between distinct sites just seem to acquire the character of fGn. (And therein lies a tale [10].)

Looking Back

We can now summarize everything we've learned so far about the changes we must make in our thinking when we pass from a description of enzyme kinetics based on ensembles of molecules to a description based on just one molecule.

1. The key quantity in single molecule measurements of enzyme kinetics is not a spectrum but a waiting time distribution $f(t)$.

2. $f(t)$ is generally non-exponential.

3. The non-exponentiality appears to originate in the phenomenon of dynamic disorder, the time variation of the rate constant k_2.

4. Dynamic disorder in turn appears to be a manifestation of conformational fluctuations.

5. The fluctuations have the same statistical characteristics as fractional Gaussian noise (at least in the case of β-galactosidase – they need not, of course, in general.)

6. These fluctuations ultimately originate in the slow highly correlated dynamics of the backbone of the protein.

7. Despite the occurrence of dynamic disorder, single molecule measurements are consistent with the ensemble Michaelis–Menten equation when the measurements are carried out to long enough times.

So as we write finis to this account of Michaelis and Menten's eponymous mechanism, it's nice to know that a hundred years later, their names can continue to be associated with enzyme kinetics at even the single-molecule level.

Acknowledgements

The author is grateful to Rati Sharma, IPC Department, IISc, for a critical reading of the manuscript, and for help with the figures.

Suggested Reading

[1] L Michaelis and M L Menten, *Biochem. Z.*, Vol.49, p.333, 1913.
[2] For a derivation, see, for instance, I H Segel, *Biochemical Calculations*, Wiley, New York, 1968.
[3] B P English, W Min, A M van Oijen, K T Lee, G Luo, H Sun, B J Cherayil, S C Kou and X S Xie, *Nat. Chem. Bio.*, Vol.2, p.87, 2006.
[4] Adapted from a discussion in *Random Walks in Biology*, by H C Berg Princeton University Press, New Jersey, 1983.
[5] S Xie, *Single Mol.*, Vol.2, p.229, 2001.
[6] S C Kou, B J Cherayil, W Min, B P English and X S Xie, *J. Phys. Chem. B,* Vol.109, p.19068, 2005.
[7] W Min, G Luo, B J Cherayil, S C Kou and X S Xie, *Phys. Rev. Lett.*, Vol.94, p.198302, 2005.
[8] H A Kramers, *Physica*, Vol.7, p.284, 1940.
[9] S Chaudhury and B J Cherayil, *J. Chem. Phys.*, Vol.125, p.024904, 2006.
[10] P Debnath, W Min, X S Xie and B J Cherayil, *J. Chem. Phys.*, Vol.123, p.204903, 2005.

Address for Correspondence
Binny J Cherayil
Department of Inorganic and Physical Chemistry
Indian Institute of Science
Bangalore 560 012, India.
Email:
cherayil@ipc.iisc.ernet.in