

Simulación estocástica

Parte I

En esta parte del curso, se utilizará un programa para simular el progreso de una reacción mediada por una enzima. El objetivo es que se obtengan datos simulados, en función del tiempo, de las concentraciones de sustratos y productos de una reacción, así como de la concentración de los intermediarios enzimáticos formados, y elaborar los datos obtenidos a fin de comprobar leyes de la cinética de enzimas.

1- El significado probabilístico de la constante de velocidad en las reacciones de 1er orden.

La velocidad (instantánea) con la cual disminuye la concentración de A para la reacción de primer orden $A \xrightarrow{k} B$ es:

$$\frac{d[A]}{dt} = -k[A] \quad 1$$

o bien:

$$-\frac{d[A]}{dt} = k[A] \quad 2$$

Se puede interpretar esta ecuación diciendo que la velocidad de desaparición (nótese el signo negativo) de $[A]$ es proporcional a su concentración y la constante de proporcionalidad es k .

Un significado importante de k se obtiene de su despeje:

$$k = -\frac{d[A]}{[A]dt} \quad 3$$

Cuando $dt = 1$, es decir, cuando ha transcurrido una unidad de tiempo, la constante k es el cociente entre lo que ha disminuido $[A]$ durante ese tiempo, cambiado de signo, $-d[A]$, y el valor de $[A]$.

En términos estadísticos, el valor absoluto de $d[A]/[A]$ representa la probabilidad de que, durante un tiempo unitario, desaparezca una cantidad $-d[A]$ del reactivo. La constante de proporcionalidad k es una medida de dicha probabilidad.

2- Ejercitación: cálculo de $[A]$ en función del tiempo.

En este punto conviene volcar en una planilla de Excel estos conceptos.

a- Cálculo determinista.

Intentaremos reproducir la curva que sigue $[A] = f(t)$ aproximando la ecuación 1 con:

$$\frac{[A]_{t+\Delta t} - [A]_t}{\Delta t} = -k[A]_t \quad 4$$

Entonces:

$$[A]_{t+\Delta t} = [A]_t - k[A]_t \Delta t = [A]_t (1 - k \Delta t) \quad 5$$

Digamos que, cuando $t = 0$, comenzamos con 1000 moléculas de A en un volumen unitario. Tomaremos $\Delta t = 1$ y construiremos una columna de tiempos (de 0 a 100, o más, según convenga) y de $[A]$ en una tabla de Excel de la siguiente forma:

Tabla 1

$k =$	0,05
<i>tiempo</i>	$[A]$
0	1000
1	950
2	902,5
3	857,375
...	...
100	5,92052922
...	...

Los valores de $[A]$ para $t > 0$ se obtienen con la ecuación 5, es decir que $[A]_{t+1} = [A]_t (1 - k)$, con $k = 0,05$.

Haga un gráfico con los datos de la tabla 1. ¿Qué ecuación representa estos valores?

Como ya se vio, la integración de la ecuación 1 entre $t = 0$ y un valor cualquiera de t es:

$$[A] = [A]_0 e^{-kt} \quad 6$$

Haga un gráfico de esta ecuación utilizando valores de $[A]_0 = 1000$ y $k = 0,05$, y superpóngalo al de los datos de la tabla. ¿Coinciden los gráficos? ¿Se puede mejorar el grado de coincidencia?

Agregue al gráfico una línea de tendencia exponencial para los datos, y elija la opción “Presentar la ecuación en el gráfico”. ¿Cuál es el valor de k en la ecuación?

b- Cálculo estocástico.

Utilizaremos aquí el concepto probabilístico de k para generar la curva.

Como en el caso determinista, consideraremos un volumen unitario y un número inicial de 1000 moléculas de A . La variación del tiempo también será unitaria.

Para calcular el número de moléculas remanentes a $t = 1$, tendremos que hacer que cada una de las 1000 moléculas decida si reaccionará o permanecerá sin cambio. Al final del proceso, algunas de las 1000 moléculas habrán reaccionado y su número se restará de las 1000 iniciales. Para $t = 2$ el ciclo recomienza pero ahora sobre la población remanente de moléculas, y así para los valores de tiempo siguientes.

¿Cómo decide la molécula si reaccionará o no?

Se utilizan números aleatorios con distribución de probabilidad uniforme estandarizada. Al ingresar en una celda de Excel: “=ALEATORIO()” (o bien “=RND()” para la versión en inglés), aparece un número real entre 0 y 1 con esta distribución (intervalo “[0, 1)” porque incluye el 0 pero no el 1).

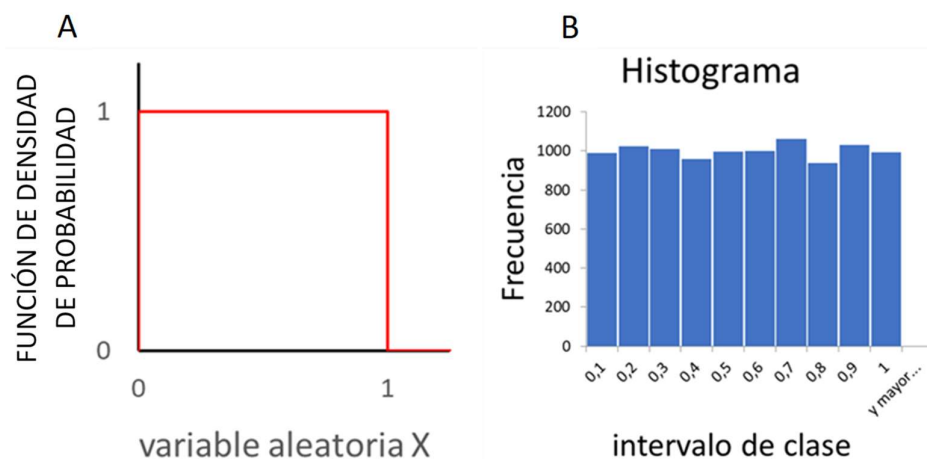


Figura 1. A: Distribución uniforme estandarizada. Fuera del segmento $[0, 1)$, la densidad de probabilidad de la distribución vale 0, y dentro del segmento vale 1. El área total bajo la curva vale 1. B: Histograma de 10.000 números aleatorios con distribución uniforme estandarizada.

Que la distribución sea uniforme significa que la probabilidad de que el número aleatorio obtenido se encuentre dentro de algún intervalo de longitud finita, incluido en el intervalo $[0, 1)$, es independiente de la ubicación del intervalo. Es decir, la probabilidad de que el número caiga entre 0 y 0,1 será la misma a que caiga entre 0,1 y 0,2, o entre 0,85 y 0,95 (todos estos son intervalos de longitud 0,1). Sin embargo, el número aleatorio tendrá el doble de probabilidad de caer en un intervalo del doble de longitud, por ejemplo, entre 0,55 y 0,75. En otras palabras, la probabilidad de que un número aleatorio, con distribución uniforme estandarizada, caiga en el intervalo $[0, 1)$ es 1, y la probabilidad de que caiga en un intervalo, incluido en el $[0, 1)$, será proporcional a la longitud del mismo.

Esto nos da la clave de cómo programar que una molécula tome la decisión de si reaccionará o no: se genera un número aleatorio con distribución uniforme; si el número cae en un intervalo de cierta longitud, determinada por el valor de k , se la hace reaccionar, si no, permanecerá sin cambio. Si el tiempo transcurrido es unitario, la longitud del intervalo coincidirá con el valor de k .

Hagamos la prueba en una hoja de Excel.

Tabla 2

	J	K	L	M	N	...	ALX
5	$k =$	0,05					
6			Total moléculas	Molécula N° →			
7	<i>tiempo</i>	[A]	que reaccionaron	1	2	...	1000
8	0	1000	=suma(M8:ALX8)	=expresión lógica
9	1	=K8-L8	=suma(M9:ALX9)
10	2	=K9-L9
...
108	100	=K99-L99
...

Esta hoja es algo más compleja que la anterior. Nótese que este es sólo un ejemplo, y los valores de celdas a los que nos referiremos en lo que sigue no tienen por qué ser los mismos que los de la hoja que prepare cada uno.

El punto central es la “expresión lógica” que figura en la celda M8 y que se deberá copiar en el resto de las celdas resaltadas en amarillo. Dicha expresión es:

=SI((ALEATORIO())<=\$K\$5)*(M\$7<=\$K\$8);1;0)

7

y se lee así:

si el número ALEATORIO() es menor o igual que un cierto valor (k), ubicado en la celda fija \$K\$5, y al mismo tiempo se cumple otra condición (sobre la cual volveremos en seguida) entonces aparecerá el valor “1”, de lo contrario aparecerá un “0”. Esta es la esencia del procedimiento por el cual, a cada una de las moléculas de A se le asigna la decisión de reaccionar o no. En la celda L8 (y siguientes hacia abajo) se calculará la suma de todas las celdas que están a la derecha en esa fila (desde la molécula N° 1 hasta la N° 1000). Se debe escribir “=suma(‘rango de celdas ubicadas a la derecha en esa

fila’)). El número obtenido será el de todas la moléculas que reaccionaron para ese valor de tiempo. Por ejemplo, en la celda L8 se escribirá “=suma(M8:ALX8)” y se copia hacia abajo; en la L9 quedará “=suma(M9:ALX9)”, y así, hasta completar la columna hasta el tiempo máximo considerado. Lo mismo se debe hacer con la columna que contiene el número remanente de moléculas a un tiempo dado. Por ejemplo, el valor en la celda K9 se calculará como “=K8-L8”.

Se entiende aquí que la longitud del intervalo representada por k es directamente el valor que aparece en la celda \$K\$5, ya que el límite inferior del intervalo es el 0. Pero podríamos igualmente haber escrito:

$$=SI((ALEATORIO()>=[\text{celda fija 1}])*(ALEATORIO()<=[\text{celda fija 2}])*...) \quad 8$$

donde [celda fija 1] y [celda fija 2] contienen los límites inferior y superior del intervalo, respectivamente. Obsérvese que el signo “*”, el de multiplicar en Excel, actúa en la expresión lógica como el operador “Y”, es decir, si se cumplen simultáneamente una condición A “Y” otra condición B.

Siguiendo con la expresión 7, la segunda condición, “M\$7<=\$K8”, está puesta para que, en la celda que se está evaluando (M8), solo aparezca un “1” cuando el número de la molécula (en M7) sea igual o menor al de las moléculas remanentes en ese instante (1000, en K8). De lo contrario, a medida que transcurre el tiempo, corremos el peligro de estar evaluando moléculas ya inexistentes. Nótese que, para la siguiente fila (tiempo = 1), el número de moléculas en K9 será menor, ya que se le habrán restado las que reaccionaron a tiempo = 0 (en la celda L8). Al escribir el signo \$ en “[columna]\$7” y “\$K[filas]” nos aseguramos que, sea cual fuere la celda “[columna][filas]” en el campo amarillo, siempre se referirá a un valor que esté en la columna K y a un valor que esté en la fila 7.

Una vez completada la tabla en la hoja de Excel, haga un gráfico de [A] en función del tiempo y agregue una línea de tendencia con su ecuación. Verá que, al actualizarse la hoja de Excel, se recalculan todos los valores aleatorios y se obtienen nuevos datos de [A] y, por lo tanto, el valor de k en la ecuación también

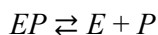
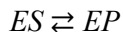
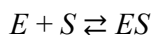
varía aleatoriamente. Calcule un promedio de los valores de k y compárelo con el valor en la celda \$K\$5, 0,05.

Nota: Para hacer este ejercicio, es posible que su versión de Excel no soporte un número de columnas tan alto como para computar 1000 moléculas. En ese caso, se deberá trabajar con filas en lugar de columnas y viceversa. Esto, naturalmente, cambia la sintaxis de algunas celdas. Por ejemplo, la condición “M\$7<=\$K8” deberá cambiar a “\$M7<=K\$8”. Pero esto dependerá del caso en particular y lo arreglaremos caso por caso. Para ilustrar el ejercicio, el mismo viene acompañado por un archivo de Excel, con una hoja para cada versión.

Parte II

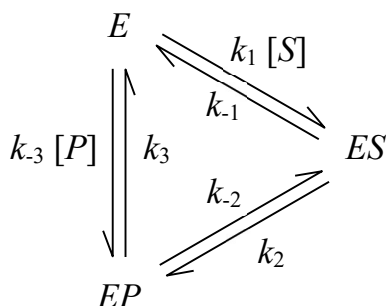
Simulación de una reacción enzimática con un solo sustrato y un solo producto.

Consideremos la reacción enzimática, vista anteriormente:



Esquema 1

Una manera de escribir esta serie de reacciones es:



Esquema 2

En las esquinas de este esquema 2 figuran los intermediarios de la enzima. Sobre las flechas de reacción hemos incluido las constantes de primer orden, k_{-1} , k_2 , k_{-2} y k_3 , y las de segundo orden, k_1 y k_{-3} , éstas últimas multiplicadas por las concentraciones del sustrato S y del producto P , respectivamente.

Nótese que cualquier intermediario de la enzima debe poder reaccionar, por lo menos, por dos caminos diferentes. Por ejemplo, E puede unir S para convertirse en ES o unir P para convertirse en EP . Esto es lo que hace reversible el ciclo enzimático. Desde un punto de vista cualitativo, el ciclo de catálisis de una enzima es siempre reversible. La irreversibilidad es una cuestión cuantitativa, ya sea porque una reacción del ciclo (o más de una) se encuentra fuertemente desplazada hacia uno de los intermediarios o por la ausencia de un ligando, por ejemplo, de S o de P . Veremos más adelante que el desplazamiento de una reacción en el ciclo depende finalmente de cuán desplazada se encuentra la reacción $S \rightleftharpoons P$.

En nuestra simulación consideraremos que:

- 1- el volumen de reacción y el paso de tiempo, dt , son unitarios. Por una razón estadística, convendrá utilizar un número de moléculas de enzima del orden de miles, y de moléculas de sustrato y/o de producto del orden de millones, pero esto no es imprescindible.
- 2- las constantes de velocidad de primer orden en el ciclo son probabilidades (ecuación 3), por lo cual deberán tomar valores en el intervalo $[0, 1)$.
- 3- las constantes de segundo orden, k_1 y k_{-3} , podrán actuar como probabilidades luego de multiplicarlas por $[S]$ y por $[P]$, respectivamente. Se aplican aquí las consideraciones para reducir una ecuación de velocidad segundo orden a una de pseudo primer orden. Sin embargo, estas probabilidades no serán generalmente constantes, ya que variarán con las concentraciones de $[S]$ y de $[P]$ (en función del tiempo, por ejemplo). Por lo tanto, deberemos asegurar que $k_1[S]$ y $k_{-3}[P]$ tomen siempre valores en el intervalo $[0, 1)$, para cualquier valor de $[S]$ y de $[P]$ y cualquier valor de tiempo.
- 4- Como cada molécula de un determinado intermediario puede reaccionar por dos caminos excluyentes (si E se combina con S no podrá hacerlo con P , por ejemplo), la elección de si reaccionará o no, y cómo reaccionará en caso de hacerlo, es un poco más compleja. En este caso, se inspeccionará si el número aleatorio que se genera queda incluido en uno de dos intervalos, excluyentes entre sí (no debe haber intersección o superposición entre estos dos intervalos) que corresponderán a cada vía de reacción. Si no queda incluido en ninguno de los dos intervalos, la molécula permanecerá sin reaccionar.

Enzyma.exe: Código del programa (puede saltarse esta parte, hasta “Instalación del programa”, si lo desea)

En la tabla 3 presentamos la parte central del código en Visual Basic® del programa que utilizaremos (columna izquierda) y una escueta explicación del mismo (columna derecha). En este tipo de lenguaje de computación, el signo “=” no es el que se entiende generalmente en matemáticas, sino que el valor sobre la derecha queda asignado al de la izquierda. Nótese que “EL” representa el número de moléculas de enzima libre, E.

Tabla 3. Parte central del código del programa ENZYMA.EXE

Private Sub MainTimer_Timer()	COMIENZO
NEL = EL	Se graban los valores de E, ES, EP, S y P, que pueden variar en lo que sigue, en las memorias NEL, NES, NEP, NS y NP, que permanecerán fijas hasta que cambie el valor de TIEMPO.
NES = ES	
NEP = EP	
NS = S	
NP = P	
For rep = 1 To NEL	Comienzo del análisis de una molécula de E
rndNumber = Rnd	Se genera un número aleatorio y se guarda en una memoria fija, para utilizarlo en lo que sigue.
If rndNumber <= k_1 *NS Then	Si se cumple condición para $E + S \rightarrow ES$
EL = EL - 1	El número de moléculas de E disminuye en 1
ES = ES + 1	El número de moléculas de ES aumenta en 1
S = S - 1	El número de moléculas de S disminuye en 1
ElseIf rndNumber <= k_1 *NS + k_3 *NP Then	Si no, si se cumple la condición para $E + P \rightarrow EP$
EL = EL - 1	El número de moléculas de E disminuye en 1
EP = EP + 1	El número de moléculas de EP aumenta en 1
P = P - 1	El número de moléculas de P disminuye en 1
End If	Si no, la molécula de E no reacciona
Next	Próxima molécula de E, hasta completar NEL
For rep = 1 To NES	Comienzo del análisis de una molécula de ES
rndNumber = Rnd	(Análogo a lo visto para el análisis de las moléculas de E)
If rndNumber <= k_{-1} Then	
EL = EL + 1	
ES = ES - 1	
S = S + 1	
ElseIf rndNumber <= $k_{-1} + k_2$ Then	
EP = EP + 1	
ES = ES - 1	
End If	
Next	Próxima molécula de ES, hasta completar NES
For rep = 1 To NEP	Comienzo del análisis de una molécula de EP
rndNumber = Rnd	(Análogo a lo visto para el análisis de las moléculas de E)
If rndNumber <= k_2 Then	
ES = ES + 1	
EP = EP - 1	
ElseIf rndNumber <= $k_2 + k_3$ Then	
EL = EL + 1	
EP = EP - 1	
P = P + 1	
End If	
Next	Próxima molécula de EP, hasta completar NEP
TIEMPO = TIEMPO + 1	Habiendo analizado todas las moléculas de E, ES y EP
End Sub	El tiempo aumenta en 1
	Se vuelve al COMIENZO

Para la evaluar la decisión de cada molécula se utiliza una técnica (“rulo For-Next”) por la cual un contador (nombre de fantasía “rep”, por “repetición”), avanza desde “1” hasta el número

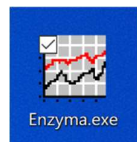
total de moléculas. Las órdenes se ejecutan secuencialmente de arriba abajo y al llegar a la orden “Next” se vuelve al “For” inmediatamente anterior, sumándole 1 al contador “rep” y comparando el nuevo valor con el del número total de moléculas. Si el valor de “rep” supera al de “NES”, por ejemplo, se salta a la posición inmediatamente posterior al “Next” y se sigue ejecutando el código hacia abajo. Al terminar de evaluar todas las moléculas existentes a un tiempo dado, el programa le suma una unidad al tiempo y vuelve a ejecutarse desde el comienzo.

Instalación del programa

En principio no se requiere instalación. Basta con extraerlo del archivo comprimido en un directorio del disco rígido que se elija y hacerlo correr, pero es posible que se requiera agregar un archivo complementario (MSCOMCTL.OCX) y/o activar este archivo en el registro de Windows. El archivo MSCOMCTL.OCX se provee junto con el programa. Como ENZYMA.EXE es un archivo ejecutable, puede que se requiera una autorización de su programa antivirus. Se le brindará ayuda para completar la instalación.

Breve guía del programa

Al copiar el programa, aparecerá un botón con la imagen:



Abierto el programa aparecerá la siguiente pantalla:

Ingresa aquí el número inicial de moléculas de S y P, así como el de EL, ES y EP.

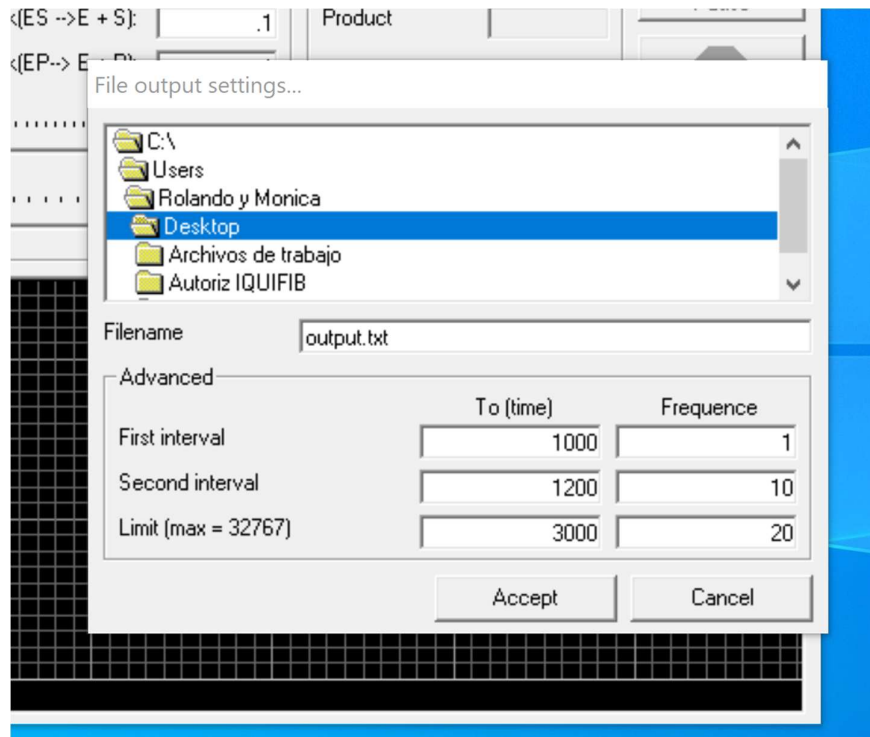
Ingresa aquí el valor de las constantes de reacción. Deben cumplir los requisitos para ser considerados como probabilidades (ver texto).

Aquí aparecerán los resultados actuales y el tiempo durante el curso de la reacción.

Pantalla gráfica y controles de escala. Aquí aparecerán graficados, en escalas arbitrarias, los valores de S, P y de los intermediarios enzimáticos, de acuerdo con el código de colores de arriba.

IMPORTANTE: Al ingresar los valores de las constantes, debe cuidarse cómo el sistema en su PC establece la separación decimal, si con una coma o con un punto, ya que esto alterará los resultados obtenidos. Luego de cerrar el programa, los valores de números de moléculas y de constantes de velocidad ingresados por el usuario no se mantienen (lamentablemente). Al volverlo a abrir, reaparecerán los valores por defecto.

En la pantalla se grafican los datos en función del tiempo. La capacidad gráfica es limitada y se utiliza solo en forma orientativa. Para aprovechar mejor los datos, conviene graficarlos en una hoja de Excel a partir del archivo de texto que se puede generar apretando el botón “Save”. El botón de guardar datos (“Save”) deberá oprimirse antes de comenzar un registro. Por ello, conviene ya tener fijados los valores del archivo en el cual se quieren guardar los datos. Cuando se oprime el botón “File output settings”, aparece la siguiente pantalla:



En la parte superior se elige el directorio donde se grabarán los datos y el nombre del archivo (de texto) que se generará, que por defecto se graba como “output.txt” en el mismo directorio en el cual se encuentra el programa. En la parte inferior se establecen las frecuencias de datos que se van a grabar. Por defecto, se graban cada unidad de tiempo hasta tiempo = 1000, cada 10 unidades hasta 1200 y cada 20 unidades hasta 3000. Luego de tiempo = 3000 ya no se grabarán más datos, a menos que el número en la casilla se incremente.

Los botones de la derecha (“Start”, “Pause” y “Stop”) no requieren mayor explicación. Una vez activado el botón “Pause”, los demás comandos quedan inhabilitados y sólo se podrá reanudar y pausar sucesivamente el registro.

EJERCICIOS

Enzyma.exe

Este programa simula en forma estocástica la evolución de las especies químicas involucradas en el esquema de reacción: $E + S \rightleftharpoons ES \rightleftharpoons EP \rightleftharpoons E + P$. En el programa, la E (enzima libre) se denota como EL.

En este ejercicio se propone simular curvas de estas especies en función del tiempo y obtener valores de velocidad (v) para distintas concentraciones iniciales de S, de manera de calcular los parámetros V_{max} y K_m .

Ingresa los valores de concentración inicial de EL (se sugiere 10 000), la concentración inicial de sustrato (ver valores de S sugeridos más abajo) y los de las constantes de velocidad de las reacciones del esquema que se transcriben al final. Mantenga los demás valores iniciales (las concentraciones de P, ES y EP) en 0.

En “File output settings” elija el nombre del archivo de texto donde se guardarán los datos (por defecto es “output.txt”, y el directorio en el que aparezca será el mismo en el cual ha copiado el programa) y los valores de intervalos de tiempo y cada cuantas unidades de tiempo se guardarán datos. Sugerencia:

	Time	Frecuence
First interval	200	1
Second interval	1000	10
Limit	5000	20

Con estos números, hasta tiempo = 200 guardará un valor de concentración de EL, ES, EP, S y P por cada unidad de tiempo; entre los tiempos 200 y 1000 los guardará cada 10 unidades y entre los tiempos 1000 y 5000 los guardará cada 20 unidades. En total grabará un máximo de 481 datos de cada especie química.

Para que los datos se guarden oprima el botón “Save” antes de comenzar el “experimento”. Para comenzarlo oprima “Start”. El programa continúa indefinidamente hasta que se oprima “Stop”. Si se hace esto antes de haber transcurrido el tiempo máximo (digamos 5000) se guardarán los datos hasta el tiempo de la detención. Tenga en cuenta que si cierra el programa, los valores de EL, S y de las constantes que Ud. ingresó se perderán y serán reemplazados con los valores por defecto.

Abra el archivo de texto, seleccione todos los datos y cópielos en un archivo de Excel (sugerencias: Control E para seleccionar todos los datos, Control C para copiar y Control V para pegar en la planilla de cálculo).

Grafique los datos de $P = f(\text{tiempo})$. Elija el intervalo de valores que cumplan con una línea recta (representativos de la generación de P en estado estacionario en las condiciones iniciales) y agrégueles una “línea de tendencia” (haga click derecho con el cursor del mouse sobre los datos del gráfico y una vez en el menú elija las opciones “Lineal”, “Presentar ecuación” y “Presentar R cuadrado”). El valor de la pendiente representa el de la velocidad (v) a la concentración de sustrato utilizada. Se sugiere realizar un experimento para cada una de las siguientes concentraciones iniciales de sustrato: 500 000, 1 000 000, 2 500 000, 5 000 000, 10 000 000, 25 000 000, 50 000 000, 100 000 000 (en millones: 0.5, 1, 2.5, 5, 10, 25, 50, 100).

Haga un gráfico de $v = f([S])$ y algún gráfico basado en la transformación de la ecuación de Michaelis-Menten a una forma lineal (por ejemplo $1/v = f(1/[S])$) para calcular los valores de K_m y de V_{max} .

Repita una curva en función de las mismas concentraciones de S para una concentración inicial de $P = 75\,000\,000$ (75 millones). En este caso, incluya también un experimento para $[S] = 0$.

Elabore un breve informe con los resultados y las conclusiones en un archivo de Excel.

Utilice los siguientes valores de constantes de velocidad:

$$k(E+S \rightarrow ES) \quad 7.5E-09$$

$$k(E+P \rightarrow EP) \quad 1.5E-09$$

$$k(ES \rightarrow EP) \quad 0.16$$

$$k(EP \rightarrow ES) \quad 0.25$$

$$k(ES \rightarrow E+S) \quad 0.018$$

$$k(EP \rightarrow E+P) \quad 0.072$$