

# Proteine Bolognese

**R. Cubarsi**

UNIVERSITAT POLITÈCNICA DE CATALUNYA  
[rcubarsi@ma4.upc.edu](mailto:rcubarsi@ma4.upc.edu)

**E. Vázquez, A. Villaverde**

UNIVERSITAT AUTÒNOMA DE BARCELONA  
[Esther.vazquez@uab.cat](mailto:Esther.vazquez@uab.cat), [Antoni.villaverde@uab.cat](mailto:Antoni.villaverde@uab.cat)

---

## Abstract

*Se modeliza la penetración de una versión modificada de la proteína verde fluorescente llamada GFP en el citoplasma y el núcleo de una célula de mamífero. El modelo matemático proporciona una primera aproximación al experimento, con una disminución progresiva de la tasa de fluorescencia a medida que la luminosidad total va aumentando hasta saturarse. Sin embargo, se detecta una desviación localizada del modelo propuesto, debida a un proceso que no se realiza de forma totalmente diferenciable. El fenómeno es explicable a partir de la irrupción brusca de GFP en el citoplasma y el núcleo por la debilitación o permeabilización de sus membranas, y por la obstrucción del flujo de GFP por la membrana nuclear. El modelo matemático proporciona pues información relevante sobre la desviación del patrón que se ha adoptado.*

*Con la finalidad de unificar del Espacio Europeo de Educación Superior (EEES), y de acuerdo con las directrices del Proceso de Bolonia, las asignaturas de los nuevos planes de estudios han sufrido una revisión a fondo de sus contenidos. Para que la nueva metodología dé sus frutos, es necesario diseñar un temario muy optimizado, en el que cada tema entronque con temas de otras asignaturas.*

*El caso presentado evidencia como la Modelización Matemática es útil para aplicar y revisar los conocimientos adquiridos en diversas asignaturas, de matemáticas y biología en este caso, además de ser una forma eficaz de estimular el aprendizaje, ya que muestra la utilidad de lo que se estudia.*

*The penetration of a modified version of a green fluorescent protein called GFP in the cytoplasm and nucleus of a mammalian cell is modelled. The mathematical model provides a first approach to the experiment, with a progressive decreasing rate of fluorescence as the total luminosity increases until saturation. However, a local deviation from the proposed model is detected, due to a process that takes place in a non-differentiable way. The phenomenon is explained from the sudden entrance of GFP into the cytoplasm and nucleus by the weakening or permeabilization of their membranes, and from the obstruction of GFP flux by the nuclear membrane. Thus, the mathematical model provides relevant information on the deviation from the pattern that has been adopted. To unify the European Higher Education Area (EHEA), and in accordance with the guidelines of the Bologna Process, the subjects of new curricula have undergone a thorough review of their contents. So the new methodology be fruitful, a highly optimized agenda must be designed by linking every topic with topics from other subjects. The case presented evidences how Mathematical Modelling is useful to implement and review lessons learned in various subjects, from maths and biology in the current case, besides being an effective way to stimulate learning, since it is showing the usefulness of what is studied.*

## 1 Introducción

Para reforzar la asimilación de conceptos matemáticos y estimular el aprendizaje, sería conveniente introducir en los planes de estudios de carreras científico-técnicas una asignatura de Modelización Matemática, con ejemplos ligados a otras disciplinas (biología, economía, física etc.), donde se manejaran conceptos específicos de otras asignaturas de matemáticas, así como conocimientos transversales de más asignaturas. Ilustraremos este punto de vista con la modelización de un experimento de biología.

En todo experimento se selecciona un número finito de propiedades a observar y se obtienen datos para estas variables en diversas circunstancias, por ejemplo, para distintos tiempos. Construir un modelo matemático significa establecer la interrelación entre estas variables, determinar cuáles van a ser las constantes que caractericen al experimento y, a partir de ahí, establecer su dinámica. En el caso que se explica, el modelo es determinista y diferenciable, por lo que constituye un conocimiento específico de una asignatura de Ecuaciones Diferenciales. El mismo modelo es también aplicable a una gran variedad de problemas de distintos campos: ecología, química, economía, etc. El problema es sencillo, por lo que bastará utilizar herramientas de cálculo en una variable.

Para el caso estudiado, se puede dar una solución analítica, aunque es más cómodo usar software matemático para las representaciones gráficas y para los métodos de estimación de parámetros y errores. Si el modelo inicial, a posteriori, se quisiera ajustar más a los datos, sería también ineludible el uso de otras técnicas numéricas y, quizás, de simulación. Por tanto, en la Modelización Matemática converge una diversidad de técnicas.

No olvidemos que los métodos numéricos para la estimación de los parámetros del modelo –estimación por mínimos cuadrados, en nuestro ejemplo– y los métodos de estimación y propagación de errores son omnipresentes en todas las áreas de la ciencia. Los primeros se estudian en asignaturas de Cálculo Numérico (si no ha desaparecido todavía del plan de estudios), mientras que los segundos son propios de asignaturas de Probabilidades y Estadística. Los alumnos utilizan estas técnicas desde los primeros cursos, en las prácticas de laboratorio de física, química, biología, etc., aún sin haberlos todavía estudiado con rigor. Por lo tanto, la Modelización Matemática es una buena oportunidad para profundizar en todos estos conceptos y métodos.

## 2 Experimento

Se estudia cómo entra en una célula de mamífero, que posee un núcleo, una proteína verde fluorescente llamada GFP, que al llevar fusionado un péptido altamente catiónico, forma pequeñas partículas de 20 nm de diámetro. Dicha proteína se ha ingeniarizado [1] para penetrar células de mamífero y transportar genes u otros fármacos a su núcleo, con finalidades terapéuticas.

Las células de mamífero son más o menos esféricas, y cuando están en cultivo son un poco más planas, de forma que, para simplificar, supondremos que adoptan forma cilíndrica, con el núcleo centrado en el eje de simetría. Cuando se sumerge una célula en un medio que contiene esta proteína, se detecta la presencia de fluorescencia, primero en el citoplasma solamente, y después en el citoplasma y en el núcleo. La detección se hace a partir de secciones ópticas de la célula, de grosor constante y suficientemente pequeño, con un microscopio láser llamado confocal, pudiendo escoger el radio de la sección a fin de incluir sólo el núcleo, o bien citoplasma y núcleo conjuntamente. De las imágenes digitalizadas se obtienen distintas mediciones para

cada sección y, sumando, para todo el volumen considerado. Por ejemplo, número de píxeles, luminosidad de cada pixel, área o volumen del total de píxeles, valores máximo, mínimo y medio de la luminosidad, densidad lumínica por unidad de superficie o de volumen, etc. Con los datos obtenidos para diferentes tiempos, se desea obtener un modelo dinámico para la penetración de la GPF en el citoplasma y en el núcleo.

La variable escogida para estudiar la evolución temporal de la GPF ha sido la fluorescencia absoluta  $\phi(t)$ , medida en unidades de fluorescencia, bien sea de toda la célula, del núcleo, o por diferencia, del citoplasma. En la tabla 12.1 se muestran los datos para una de las células.

$t$	$\phi_T$	$\phi_N$	$\phi_C$
90	203109	84084	119025
105	217723	92602	125121
120	234113	102097	132016
135	254372	112749	141623
150	269190	122292	146898
165	289247	134582	154665
180	304716	143141	161575
195	320808	152154	168654
210	333268	159449	173819
225	346376	166432	179944
240	359224	173695	185529
255	370714	180322	190392
270	384155	187412	196743
285	398401	194708	203693
300	411066	201264	209802
315	427259	209381	217878
330	440822	215982	224840
345	452704	221297	231407
360	461427	226326	235101
375	465128	229619	235509

Table 12.1: Fluorescencia absoluta (en unidades de fluorescencia) de toda la célula  $\phi_T$ , del núcleo  $\phi_N$  y del citoplasma  $\phi_C$ , para cada tiempo  $t$  (min.).

### 3 Modelo

Consideremos, en primer lugar, la célula completa. La fluorescencia total  $\phi(t)$  (se omite en este caso el subíndice  $T$ ) en función del tiempo es casi nula al principio del experimento, y luego es una cantidad siempre positiva y continuamente creciente, que tiende a estabilizarse. En principio, por similitud a la dinámica de poblaciones, parece lógico tomar el modelo de la ley logística, dado por la ecuación diferencial:

$$\frac{1}{\phi} \frac{d\phi}{dt} = p\phi + q; \quad q > 0, \quad p < 0 \tag{12.1}$$

con  $p$  y  $q$  constantes. Definiendo  $F(t) = \ln \phi(t)$ , la tasa de variación o aumento relativo de la fluorescencia viene dada por  $F'(t) = \frac{d \ln \phi}{dt}$ . Al inicio, cuando  $\phi \approx 0$ ,  $F'$  vale  $q$ , que es la máxima tasa de fluorescencia. Después, la tasa  $F'$  baja progresivamente hasta anularse, cuando la fluorescencia se estabiliza. La máxima fluorescencia es  $L = -\frac{q}{p}$ . Estos son los dos únicos puntos de equilibrio del sistema:  $\phi = 0$ , inestable, y  $\phi = L$ , estable. A diferencia del modelo poblacional, no tiene sentido un valor  $\phi > L$ . Así pues, los parámetros del modelo son

Componente	$p$	$q$	$L$
Total	$-1.51 \cdot 10^{-8} \pm 1.23 \cdot 10^{-9}$	$0.82 \cdot 10^{-2} \pm 0.44 \cdot 10^{-3}$	$5.43 \cdot 10^5 \pm 5.31 \cdot 10^4$
Citoplasma	$-2.11 \cdot 10^{-8} \pm 2.89 \cdot 10^{-9}$	$0.63 \cdot 10^{-2} \pm 0.53 \cdot 10^{-3}$	$2.97 \cdot 10^5 \pm 4.80 \cdot 10^4$
Núcleo	$-4.05 \cdot 10^{-8} \pm 2.69 \cdot 10^{-9}$	$1.02 \cdot 10^{-2} \pm 0.46 \cdot 10^{-3}$	$2.53 \cdot 10^5 \pm 2.03 \cdot 10^4$

Table 12.2: Parámetros  $p, q$  y  $L$  para la célula completa, citoplasma y núcleo, con su desviación estandar.

$p$  y  $q$ , distintos para cada célula. La solución de la ecuación 12.1 es inmediata y su discusión se encuentra en muchos textos básicos de ecuaciones diferenciales [2].

Para determinar los parámetros del modelo a partir de un conjunto de  $m$  valores  $(t, \phi(t)); t = t_1, \dots, t_m$ , se usará la expresión discretizada

$$F'(t_i) = p\phi(t_i) + q \quad (12.2)$$

Para ello, es necesario hacer la estimación numérica de la derivada temporal  $F'(t)$ . Aunque en este caso las medidas temporales se han tomado a intervalos regulares, en otros casos no es así. Por lo tanto, para conseguir una mejor aproximación, hacemos una estimación de la derivada de la función  $F(t)$  en tres puntos. Así, en un tiempo intermedio  $t_i = t_2, \dots, t_{m-1}$  obtenemos [3]

$$F'(t_i) = \frac{(t_{i-1} - t_i)^2 F(t_{i+1}) - ((t_{i-1} - t_i)^2 - (t_{i+1} - t_i)^2) F(t_i) - (t_{i+1} - t_i)^2 F(t_{i-1})}{(t_{i+1} - t_i)(t_{i-1} - t_i)(t_{i-1} - t_{i+1})} \quad (12.3)$$

Una vez sustituida la ecuación 12.3 en la ecuación 12.2, el siguiente paso es resolver el sistema lineal sobredeterminado por el método de mínimos cuadrados, en los  $n = m - 2$  puntos donde se tiene calculada  $F'(t)$ . Escribimos dicho sistema en la forma matricial  $\mathbf{y} = \mathbf{G}\mathbf{x}$ , donde  $\mathbf{x} = (p, q)^T$ . Puesto que los datos iniciales son mediciones únicas y no llevan asociada una medida de error, asignaremos a cada una de ellas un error experimental  $\sigma_y$ , tal que el error cuadrático medio del ajuste de regresión sea  $\sigma_y^2 = \frac{1}{n}|\mathbf{y} - \mathbf{G}\mathbf{x}|^2$ . Finalmente, la estimación de los errores de los parámetros se calcula, como es usual [4], a partir de la matriz de covarianzas de los errores experimentales,  $\mathbf{V}_x = (\mathbf{G}^T \mathbf{G})^{-1} \sigma_y^2$ . En la tabla 12.2 se muestran los estimadores óptimos de los parámetros  $p, q$  para la célula completa y el núcleo, así como sus errores (desviación estandar).

El modelo propuesto sería básicamente aplicable a toda la célula y al núcleo, por encontrarse en situaciones parecidas: la proteína entra y la fluorescencia aumenta hasta llegar a la saturación. En cambio, el citoplasma podría no mostrar necesariamente una luminosidad creciente, pues su fluorescencia podría descender a medida que aumenta la del núcleo. Sin embargo, observamos que en muchos casos sí que es efectivamente posible usar el mismo modelo para aproximar los parámetros del citoplasma. Por tanto, también se listan los parámetros del citoplasma en la tabla 12.2. En la figura 1 se muestran las rectas de regresión, de acuerdo con la ecuación 12.2.

Para el parámetro  $q > 0$ , ordenados de menor a mayor, se cumple  $q_C \leq q_T \leq q_N$ . Así, la tasa inicial de toda la célula es un promedio de las tasas del citoplasma y del núcleo. Por otra parte,  $p < 0$  es el valor que hace que la fluorescencia se sature. Para el límite de saturación se cumple  $L_T = L_C + L_N$ , es decir  $-\frac{q_T}{p_T} = -\frac{q_C}{p_C} - \frac{q_N}{p_N}$ . Por tanto,  $-\frac{q_T}{p_T} \geq -\frac{q_C}{p_C}$  y  $-\frac{q_T}{p_T} \geq -\frac{q_N}{p_N}$ . Considerando la relación de los respectivos parámetros  $q$ , también se cumple  $-p_T \geq -p_C$  y  $-p_T \geq -p_N$ .

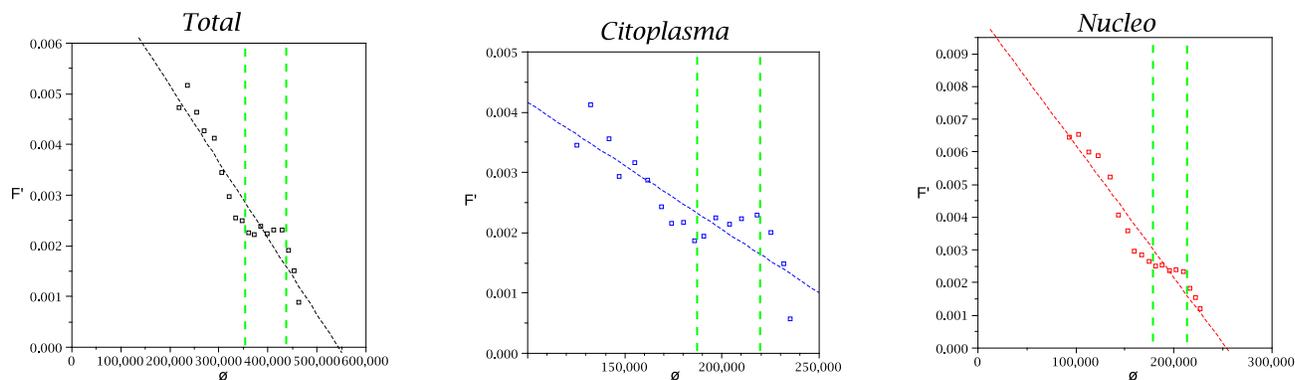


Figure 1: Tasa de fluorescencia  $F'$  en función de la fluorescencia  $\phi$ , según la ecuación 12.1, para toda la célula (izquierda), citoplasma (centro) y núcleo (derecha). Los cuadrados corresponden a datos experimentales y la línea continua al ajuste de regresión. El modelo no explica la zona entre líneas verticales discontinuas. Modificada de [6].

## 4 Discusión

Aunque el modelo proporciona una primera aproximación al proceso estudiado, con una disminución progresiva de la tasa de fluorescencia a ritmo casi constante a medida que la luminosidad total va aumentando hasta saturarse, observamos, sin embargo, que los gráficos de la tasa de fluorescencia tienen una tendencia característica intercalada en la línea de regresión: en el recorrido descendente de la tasa de fluorescencia, hay un período de estabilidad del ritmo de fluorescencia, incluso con un ligero aumento de dicha tasa (figura 1, entre las líneas verdes discontinuas). Ello podría ser debido a un proceso que no se realiza de forma totalmente diferenciable. Apuntamos las siguientes posibles causas:

- (a) Irrupción brusca de la GFP en el citoplasma y en el núcleo, por debilitación o ruptura de sus membranas, lo cual produciría una tendencia a la estabilización, o aumento de la tasa de fluorescencia, en la parte posterior a la membrana.
- (b) Obstrucción del paso de GFP por la membrana nuclear, lo cual produciría un aumento de la tasa de fluorescencia del citoplasma.

En la figura 2 se muestra la evolución temporal de la fluorescencia. Las curvas, para toda la célula (en negro) y para el núcleo (en rojo), son de tipo sigmoide, como corresponde a la solución de la ecuación 12.1. Para el citoplasma (en azul), obtenido como diferencia entre las anteriores, también se asemeja a una función sigmoide, aunque podría no ser así si los límites de fluorescencia entre el total y el núcleo fueran similares.

A partir de las anteriores figuras podemos verificar si esta “anomalía” del modelo se produce, o no, simultáneamente en el citoplasma y en el núcleo. En la figura 1 marcamos las fluorescencias inicial y final de esta zona con la línea discontinua verde. Luego, en las curvas temporales de la figura 1, leemos los tiempos correspondientes a estas fluorescencias. Los límites de la región seleccionada son comunes para el núcleo, el citoplasma y la célula completa. La simultaneidad es muy evidente en el extremo de la derecha de la zona señalada, para todas las partes de la célula (figura 1). El inicio de la anomalía, en cambio, es el mismo para la célula completa y para el citoplasma, con un cambio brusco de comportamiento de la tasa de fluorescencia, pero

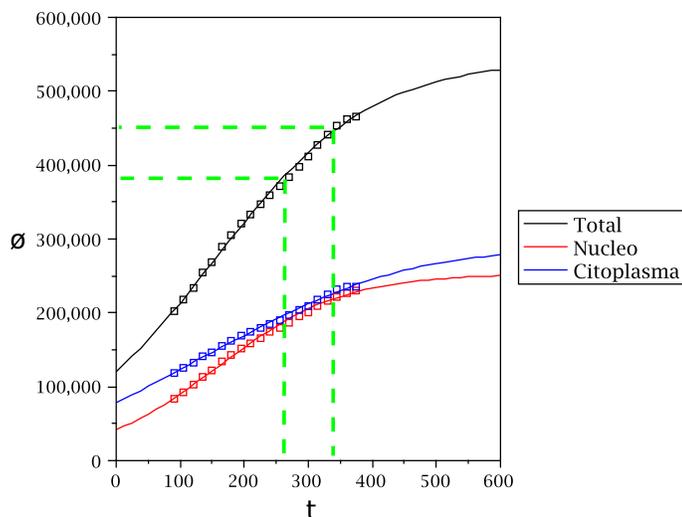


Figure 2: Evolución temporal de la fluorescencia, en negro para la célula completa, en rojo para el núcleo y en azul para el citoplasma. En todos los casos, los cuadrados corresponden a los datos experimentales, y las curvas, al modelo matemático.

se produce de forma mucho más suave en el núcleo.

En consecuencia, parece que la membrana nuclear retiene GFP en el citoplasma y, posteriormente, su ruptura reproduce una sobrefluorescencia similar en el núcleo. No obstante, mientras esto sucede en el núcleo, todavía hay un aumento de la tasa de fluorescencia en el citoplasma, en lugar de un descenso brusco, como sería lógico. Esto es posible explicarlo por una entrada súbita de GFP en el citoplasma, mayor que el flujo que sale hacia el núcleo. Es decir, por una ruptura o penetración súbita de la membrana del citoplasma.

La confirmación de esta hipótesis la proporciona el gráfico del comportamiento global de la célula (figura 1, Total), donde también se produce un sensible aumento de la tasa de fluorescencia durante este intervalo de tiempo. En [6] se halla una discusión detallada del experimento.

Concluimos pues, que no sólo es útil la correcta modelización de un experimento, sino que, de un modelo que no se ajusta totalmente a los datos, también se pueden deducir resultados interesantes, precisamente porque nos proporciona información relevante sobre la desviación del patrón que se ha adoptado.

## 5 Conclusión

Los nuevos planes de estudios, de acuerdo con las directrices del Proceso de Bolonia [5], iniciado en 1999 con la finalidad de unificar del Espacio Europeo de Educación Superior (EEES), han consolidado la estructura cuatrimestral de la mayoría de las asignaturas, forzando una revisión a fondo de sus contenidos. Se contempla un incremento de horas lectivas no presenciales –trabajo personal del alumno– para compensar la reducción de horas lectivas presenciales que supone la nueva reestructuración. En consecuencia, las asignaturas de matemáticas, impartidas en grados de titulaciones técnicas y experimentales, pueden únicamente contener los conocimientos estrictamente necesarios para poder realizar las otras asignaturas de la carrera. No vamos a discutir si esta nueva forma de enfocar las materias es mejor o peor, o si producirá alumnos mejor preparados, con una visión más completa de las matemáticas y con mayor creatividad.

Para que la nueva metodología dé sus frutos, hay que plantear un temario muy optimizado, en el que cada tema entronque con temas de otras asignaturas. Así, se pueden aprender unos contenidos específicos de cada asignatura, y practicar o reforzar los contenidos específicos de otras asignaturas, los cuales suelen llamarse conocimientos transversales. En nuestro caso, el problema descrito puede ser estudiado desde distintas asignaturas, a saber, biología celular, microbiología, genética, biología molecular, bioquímica, etc. y para ser modelizado matemáticamente requiere conocimientos de cálculo, ecuaciones diferenciales, métodos numéricos, álgebra, estadística, etc.

El tiempo nunca sobra y los temarios, heredados de planes de estudios anteriores, tienden a ser demasiado extensos. Especialmente los seres humanos aprendemos a base de imitar y repetir un mismo procedimiento, hasta conseguir hacerlo propio. Aquí podemos recordar aquello de “Ein Mal ist kein Mal” (literalmente, una vez es ninguna vez). Para asimilar los contenidos, es necesario repetir del derecho y del revés los mismos procesos de aprendizaje. Y si no ¿cómo se aprenden las tablas de multiplicar, los modos y tiempos de verbos, o se adquiere destreza con un instrumento musical?

La Modelización Matemática es un buen pretexto para aplicar y revisar la mayor parte de los conocimientos adquiridos en otras asignaturas, en particular de matemáticas y, además, una forma eficaz de estimular el aprendizaje, dado que muestra la utilidad de lo que se estudia.

**Agradecimientos.** Agradecemos el apoyo financiero recibido de MICINN (ACI2009-0919), FISS (PS0900165), AGAUR (2009SGR-108) y el CIBER de Bioingeniería, Biomateriales y Nanomedicina. AV ha sido distinguido con un premio ICREA ACADEMIA.



# References

- [1] E. Vázquez *et al.*, *Nanomedicine (Lond)*, 5(2):259-68, 2010.
- [2] M.W. Hirsch, S. Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*. Academic Press, New York, 1974.
- [3] D. Kincaid, W. Cheney, *Numerical Analysis*. Brooks/Cole Publishing, Belmont, CA, 1991.
- [4] R. Barlow, *Statistics, a guide to the use of statistical methods in the physical sciences*. John Wiley & Sons, Chichester, 1989.
- [5] [http://europa.eu/legislation\\_summaries/education\\_training\\_youth/lifelong\\_learning/c11088\\_es.htm](http://europa.eu/legislation_summaries/education_training_youth/lifelong_learning/c11088_es.htm)
- [6] E. Vázquez *et al.*, *Biomaterials*, 31, 9333-9339, 2010.

