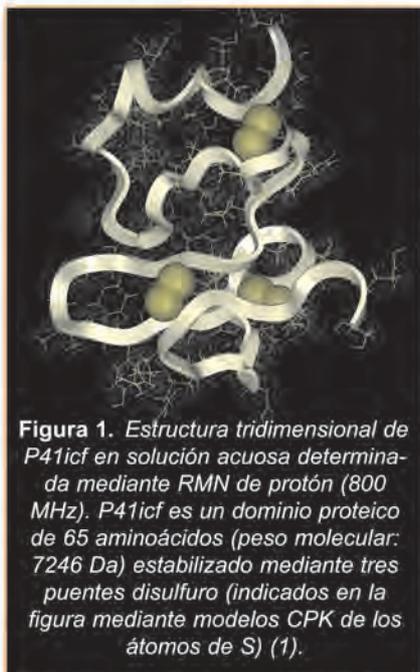


# Obtención de estructuras tridimensionales en disolución utilizando Resonancia Magnética Nuclear.

Hace pocas semanas, nuestro laboratorio publicó un artículo en el que se describe por primera vez la síntesis y la estructura del compuesto P41icf, que es un potente inhibidor de la enzima cathepsina L humana. (1) Dicha enzima está relacionada con una amplia variedad de procesos patológicos asociados a distintas enfermedades como osteoporosis, cáncer ó artritis reumatoide, por lo que se considera una importante diana terapéutica. P41icf es un inhibidor enormemente potente y muy selectivo de cathepsina L humana, por lo que la elucidación de su estructura tridimensional (**fig. 1**) abre las puertas al diseño y síntesis de versiones simplificadas de P41icf optimizadas de cara a su potencial uso terapéutico. La determinación de la estructura



de P41icf en disolución se ha llevado a cabo utilizando RMN de alto campo. El proceso para establecer la estructura tridimensional de una



Ernest Giralt

Departamento de Química Orgánica.  
Universidad de Barcelona e IRBB -  
Parque Científico de Barcelona. Josep  
Samitier 08028 Barcelona  
[egiralt@pcb.ub.es](mailto:egiralt@pcb.ub.es)

molécula como P41icf consta de tres fases:

- Asignación espectral.
- Medida de restricciones experimentales.
- Cálculo de la estructura.

## ASIGNACIÓN ESPECTRAL

Esta primera fase del proceso es frecuentemente la más laboriosa. Para facilitar el proceso, en muchas ocasiones se recurre al marcaje isotópico de la molécula con  $^{15}\text{N}$ ,  $^{13}\text{C}$  e, incluso,  $^2\text{H}$ . La idea es producir la proteína mediante expresión en un microorganismo como la bacteria *E.coli*. Resulta relativamente fácil "adaptar" la bacteria a crecer en un medio de cultivo en el que la única fuente de nitrógeno sea ión amonio. A partir de aquí, y utilizando, por ejemplo, cloruro amónico- $^{15}\text{N}$ , se obtiene una proteína en la que todos los átomos de hidrógeno están marcados. Repitiendo la misma operación, y utilizando como única fuente de carbono glucosa- $^{13}\text{C}$ , se obtiene una proteína marcada con  $^{13}\text{C}$ . Para proteínas de gran tamaño, resulta conveniente sustituir parcial ó totalmente los átomos de hidrógeno por átomos de deuterio. Aunque con mayor dificultad, las bacterias aca-

ban "adaptándose" también a crecer en agua pesada, lo que permite, a su vez, obtener la proteína de interés en forma parcial ó totalmente deuterada. La preparación de este tipo de muestras marcadas permite la utilización de una amplia variedad de experimentos, denominados de triple re-sonancia. Se llaman así porque, mediante la utilización de espectrómetros de RMN y sondas adecuados, se pueden irradiar y observar de manera combinada en el mismo experimento átomos de  $^1\text{H}$ ,  $^{15}\text{N}$  y  $^{13}\text{C}$ , a sus respectivas frecuencias de resonancia. Para moléculas de peso molecular muy elevado se debe utilizar además deuteración y las secuencias de pulsos del tipo TROSY, desarrolladas por Pervushin y Wüthrich (2). Este tipo de estrategia es el que nos ha permitido recientemente completar la asignación de la proteína Fuc-1 (3), un inhibidor de proteasas de gran importancia biotecnológica (4) que resulta especialmente difícil de estudiar mediante RMN, no sólo por su elevado tamaño (195 aminoácidos; peso molecular 23006 Da), sino también por su tendencia a formar en solución microagregados. Este tipo de estrategia basada en el marcaje isotópico no se ha podido utilizar, sin embargo, en el caso de P41icf. Alternativamente, P41icf se ha preparado mediante síntesis en fase sólida utilizando metodología desarrollada conjuntamente por nuestro laboratorio en Barcelona y el grupo de Shumpei Sakakibara en Japón (5, 6). No obstante, el uso de instrumentos de muy alto campo, en este caso concreto 800 MHz, ha permitido la asignación de todos los protones de la molécula utilizando exclusivamente experimentos multidimensionales homonucleares.

## MEDIDA DE RESTRICCIONES EXPERIMENTALES

Sin lugar a dudas, el experimento más importante para la determinación de estructuras tridimensionales mediante RMN es el experimento NOESY. Se trata de un experimento bidimensional que conduce a espectros en los que las señales que aparecen en la diagonal corresponden a cada uno de los átomos de hidrógeno de la molécula, mientras que las señales fuera de la diagonal conectan parejas de átomos de hidrógeno que se encuentran cercanos en el espacio. Cuando se estudian moléculas marcadas con  $^{15}\text{N}$  ó  $^{13}\text{C}$ , el experimento 2D NOESY bidimensional que acabo de describir puede sustituirse por experimentos tridimensionales del tipo 3D  $^1\text{H}$ - $^{15}\text{N}$  NOESY-HSQC y 3D  $^1\text{H}$ - $^{13}\text{C}$  NOESY-HSQC, e incluso experimentos tetradimensionales del tipo 4D  $^{13}\text{C}/^{15}\text{N}$  HMQC-NOESY-HMQC (7). En estos casos, se utiliza la dispersión de desplazamientos químicos de los átomos de  $^{13}\text{C}$  y/ó  $^{15}\text{N}$  para "separar" (en inglés, "edit") el conjunto de señales NOESY en distintos planos.

En una situación ideal, en la que sean válidas una serie de aproximaciones, la intensidad (ó volumen) de cualquier señal de un espectro NOESY que conecta a una pareja de átomos de hidrógeno es proporcional a la inversa de la sexta potencia de la distancia entre los átomos ( $1/r^6$ ). Si utilizamos esta sencilla expresión, y haciendo servir cualquier pareja de átomos de la molécula con distancia intraatómica conocida como calibración, es posible obtener distancias interatómicas. En el caso de P14icf, por ejemplo, trabajando a 800 MHz, fuimos capaces de determinar experimentalmente más de 900 distancias interatómicas, lo que, en principio, supone una garantía de cara a obtener una estructura tridimensional con buena resolución.

Curiosamente, en el campo de la determinación de estructuras tridimensionales mediante RMN se da frecuentemente la paradoja de que el estudio de moléculas de bajo

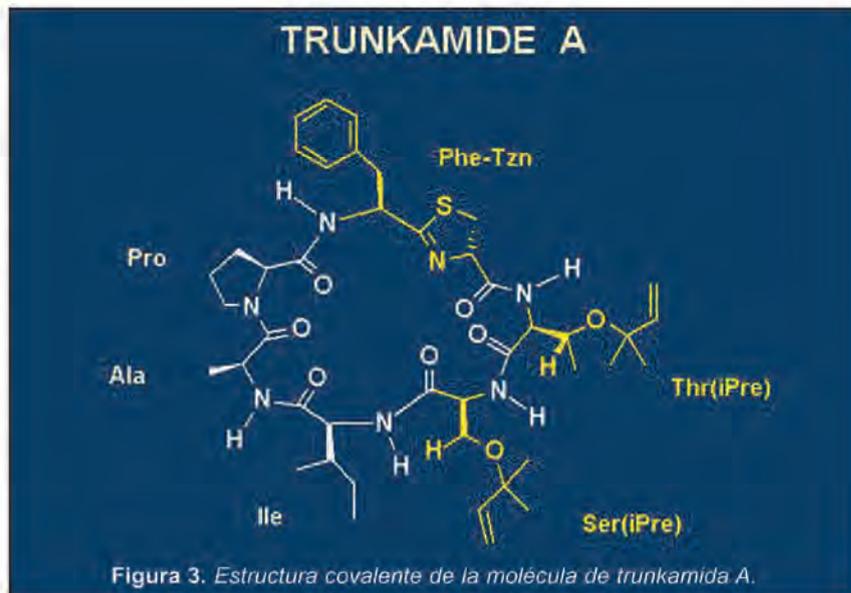
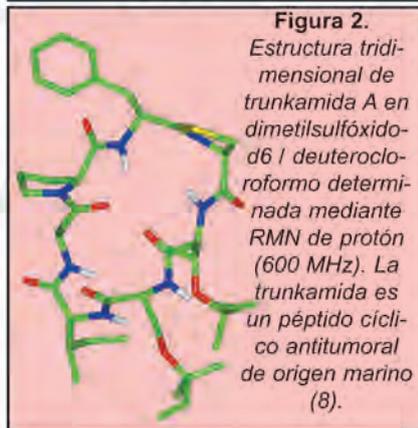


Figura 3. Estructura covalente de la molécula de trunkamida A.



peso molecular resulta en ocasiones mucho más difícil que el estudio de moléculas más complejas, de peso molecular más elevado. Esta circunstancia es el resultado de distintos factores. Por una parte, podríamos decir que, en términos relativos, una molécula pequeña tiene más superficie que una molécula grande. Hay que tener en cuenta que, en primera aproximación, por ejemplo, para un cuerpo esférico de radio  $r$ , la superficie aumenta con  $r^2$ , mientras que la masa aumenta con  $r^3$ . De la misma manera no es extraño que, para compensar la pérdida de calor a través de la piel, un animal pequeño como el ratón tenga unas necesidades alimenticias, proporcionalmente, mucho mayores que un animal de gran tamaño como el elefante.

En segundo lugar, algunas de las aproximaciones a las que nos hemos referido al hablar de la dependencia entre señales NOESY

y distancia, dejan de ser válidas con más frecuencia en moléculas pequeñas que en moléculas grandes bien plegadas (7). Pese a ello, en casos favorables y adoptando precauciones adecuadas, es posible obtener también para moléculas medianas y pequeñas estructuras tridimensionales en solución bien definidas. Este es el caso, por ejemplo, de la estructura que se muestra en la fig. 2. Se trata de la estructura en DMSO- $d_6$ /CDCl $_3$  30:70 del ciclopeptido antitumoral de origen marino trunkamida A, determinada recientemente en nuestro laboratorio mediante RMN a 600 MHz (8). En la fig. 3 se muestra la estructura covalente de la molécula que, como puede apreciarse, contiene dos sustituyentes tipo dimetilalilo que bloquean los grupos hidroxilo de los residuos de serina y treonina y un anillo de tiazolina que, probablemente, proviene biogénicamente del ataque nucleófilo del grupo tiol de un residuo de cisteína sobre el grupo carbonilo de un residuo adyacente de D-fenilalanina.

Las distancias interprotónicas determinadas mediante experimentos NOESY no son las únicas restricciones experimentales que pueden obtenerse mediante experimentos de RMN. Las constantes de acoplamiento escalar, especialmente aunque no únicamente, del tipo protón-protón pueden determinarse mediante una amplia variedad de experimentos y pueden

relacionarse fácilmente con ángulos de torsión utilizando ecuaciones tipo Karplus (7).

Los coeficientes de variación con la temperatura de desplazamientos químicos de protones NH nos permiten, aunque sea de manera indirecta, deducir la presencia de enlaces de hidrógeno con grupos carbonilo (**fig. 2**). En moléculas marcadas, experimentos tridimensionales del tipo HNCO han permitido recientemente la determinación de pequeñas constantes de acoplamiento a varios enlaces, incluyendo el enlace de hidrógeno, que permiten localizar sin ningún tipo de ambigüedad este tipo de restricciones.

Finalmente, el trabajo con muestras parcialmente orientadas ha abierto las puertas a la determinación de constantes de acoplamiento dipolar residuales que contienen valiosa información estructural sobre la orientación relativa de distintos enlaces de la molécula respecto al campo magnético externo (9). Este tipo de información puede ser utilizado para el refinamiento de estructuras generadas a partir de experimentos NOE, e incluso, en casos favorables, para generar estructuras de baja ó moderada resolución sin necesidad de medir efectos NOE ni de asignar de manera exhaustiva todas las resonancias de la molécula.

## CÁLCULO DE LA ESTRUCTURA

El tratamiento informático de todas las restricciones experimentales determinadas mediante experimentos de RMN efectuados sobre la sustancia problema permite, finalmente, obtener la deseada estructura tridimensional. El método más utilizado en la actualidad es el denominado dinámica molecular restringida ("restrained molecular dynamics, rMD"). Se basa en la utilización de los ya clásicos cálculos de dinámica molecular para el estudio de las conformaciones moleculares más estables, pero utilizando campos de fuerza "modificados". La modificación consiste en añadir a la expresión del campo de fuerzas un

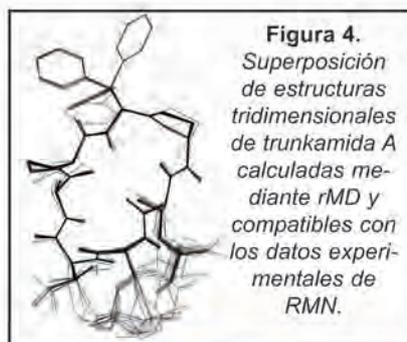
conjunto de potenciales (tantos como restricciones experimentales disponibles) que "penalicen" aquellas estructuras que no satisfagan la restricción experimental. De esta manera se consigue que estos potenciales conduzcan a la estructura hacia aquella conformación que sea compatible con todos los datos experimentales disponibles y que, al mismo tiempo, sea razonable desde un punto de vista energético.

El resultado en este tipo de cálculos, en general, un conjunto de conformaciones, todas ellas compatibles con los datos experimentales. La superposición de dichas estructuras conduce a figuras del tipo de la **fig. 4**, en las que junto a zonas de la molécula con estructura bien definida se observan zonas con estructura variable. Estas zonas de estructura variable son generalmente el resultado de una elevada movilidad local rápida a la escala de tiempo de la RMN. Sin embargo, un diagnóstico preciso de este comportamiento necesita, generalmente, que se realicen estudios de relajación. Sólo así pueden distinguirse las imprecisiones causadas por la movilidad de las imprecisiones causadas, simplemente, por la falta de datos experimentales (7).

De hecho, el estudio de estructuras tridimensionales de moléculas flexibles mediante RMN representa hoy en día un problema, sólo parcialmente resuelto pese a los intensos esfuerzos metodológicos por abordarlo de manera satisfactoria (10, 11).

## ELUCIDACIÓN ESTRUCTURAL MEDIANTE RMN: PASADO, PRESENTE Y FUTURO

La disponibilidad de espectrómetros de RMN supuso para la química, a partir de los años sesenta, la posibilidad de acceder de manera muy directa a la constitución de los compuestos objeto de estudio. La información de tipo tridimensional vino de la mano de la medida de constantes de acoplamiento (combinada, como ya se ha comentado, con ecuaciones de tipo Karplus) y, más tarde, de la medida puntual de efec-



**Figura 4.**  
Superposición de estructuras tridimensionales de trunkamida A calculadas mediante rMD y compatibles con los datos experimentales de RMN.



**Figura 5.** Imán superconductor del instrumento de 800 MHz situado en el Parque Científico de Barcelona. Mientras que los imanes superconductores de instrumentos de campo más bajo trabajan a la temperatura de ebullición del helio líquido (4 K), en este caso la temperatura de trabajo es 2 K. Esta temperatura tan baja se consigue enfriando el helio líquido en el interior del imán por efecto Joule-Thomson.

tos NOE. Sin embargo, esta utilización, podríamos decir "local", de datos espectroscópicos permitía diagnosticar únicamente situaciones relativamente sencillas: distinción entre isómeros *cis* y *trans*, posición de equilibrios silla-silla en compuestos ciclohexánicos, etc...

El presente está marcado, tal y como he ido comentando a lo largo de este artículo, por una serie de avances instrumentales, por el desarrollo de nuevos experimentos y por nuevas estrategias para la preparación de muestras. Entre los avances instrumentales se ha citado ya la carrera tecnológica por el desarrollo de imanes cada vez más potentes (**fig. 5**). Los imanes

potentes permiten ganar al mismo tiempo dispersión espectral y sensibilidad. Por otra parte, la ventaja de utilizar experimentos tipo TROSY en el estudio de moléculas de peso molecular elevados se pone de manifiesto únicamente a campos altos. La falta de sensibilidad, intrínseca en cierto modo a la técnica RMN, puede aliviarse (al menos para disoluciones de baja fuerza iónica) con el uso de las modernas criosondas, en las que el nivel del ruido se reduce espectacularmente mediante enfriamiento de los componentes electrónicos de la sonda a 20-30 K. La medida de constantes residuales dipolares en muestras acuosas parcialmente orientadas es una de las novedades de mayor impacto de cara a la determinación estructural de moléculas de tamaño

elevado. Por lo que refiere a la preparación de muestras, las técnicas de marcaje isotópico "global" a las que me he referido anteriormente pueden complementarse con una amplia variedad de métodos para el marcaje específico de determinados residuos ó grupos funcionales que permiten simplificar de manera muy importante los espectros de moléculas complejas (12).

Finalmente, si miramos hacia el futuro podemos anticipar a corto plazo que la carrera hacia campos más altos (1GHz!) continuará. Será fantástico disponer de métodos para orientar parcialmente (y, por tanto, medir constantes de acoplamiento dipolar residuales) compuestos apolares solubles únicamente en disolventes no acuosos. Gracias a los

avances en hardware informático, asistiremos a un impulso en el uso de experimentos de RMN homonucleares de elevada dimensionalidad (3D, 4D, 5D?, ...), experimentos que hoy por hoy presentan serias dificultades asociadas al manejo de grandes cantidades de datos. Por último, y relacionado directamente con el objeto del artículo que nos ocupa, el acceso a un mayor número de restricciones experimentales combinado con una utilización totalmente integrada de dichas restricciones, permitirá extender el campo de aplicación de la RMN a problemas de elucidación configuracional de una forma totalmente distinta a la manera cómo se abordan hoy día este tipo de problemas (8).

## BIBLIOGRAFIA

1. Synthesis and NMR Structure of P411c, a Potent Inhibitor of Human Cathepsin L. C. Chiva, P. Barthe, A. Codina, M. Gairi, F. Molina, C. Granier, M. Pugniere, T. Inui, H. Nishi, Y. Nishiuchi, T. Kimura, S. Sakakibara, F. Albericio and E. Giralt, *J. Am. Chem. Soc.*, **2003**, *125*, 1508-1517.
2. TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins, M. Sazman, K. Pervushin, G. Wider, H. Senn, K. Wüthrich, *Proc. Natl. Acad. Sci.*, **1998**, *95*, 13585-13590.
3.  $^1\text{H}$ (N),  $^{15}\text{N}$ ,  $^{13}\text{CO}$ ,  $^{13}\text{Ca}$ ,  $^{13}\beta$  assignment and secondary structure of a 20 kDa  $\alpha$ -L-fucosidase from pea using TROSY, A. Codina, M. Gairi, T. Tarragó, A.R. Viguera, M. Feliz, D. Ludevid and E. Giralt, *J. Biomol. NMR*, **2002**, *22*, 295-296.
4. The fuc1 gene product (20 kDa FUCI) of *Pisum sativum* has no  $\alpha$ -L-fucosidase activity, T. Tarragó, I. Martínez, M. Torrent, A. Codina, E. Giralt, P. Puigdomenech, D. Ludevid, *Plant Mol. Biol.*, **2003**, *51*,

877-884.

5. Synthesis and applications of a new base-labile fluorene derived linker for solid-phase peptide synthesis, F. Rabanal, E. Giralt and F. Albericio, *Tetrahedron*, **1995**, *51*, 1449-1458.

6. Chemical synthesis of the precursor molecule of the *Aequorea* green fluorescent protein, subsequent folding, and development of fluorescence, Y. Nishiuchi, T. Inui, H. Nishio, J. Bódi, T. Kimura, F. I. Tsuji, S. Sakakibara, *Proc. Natl. Acad. Sci. USA*, **1998**, *95*, 13549-13554.

7. Protein NMR Spectroscopy, J. Cavanagh, W. J. Fairbrother, A. G. Palmer III, N. J. Skelton, *Academic Press*, **1996**.

8. Solution Structure of the Antitumor Candidate Trunkamide A by 2D NMR and Restrained Simulated Annealing Methods, X. Salvatella, J. M. Caba, F. Albericio and E. Giralt, *J. Org. Chem.*, **2003**, *68*, 211-215.

9. Use of dipolar  $^1\text{H}$ - $^{15}\text{N}$  and  $^1\text{H}$ - $^{13}\text{C}$  couplings in the structure deter-

mination of magnetically oriented macromolecules in solution, N. Tjandra, J. G. Omichinski, A. M. Gronenborn, G. M. Clore, A. Bax, *Nature Struct. Biol.*, **1997**, *4*, 732-738.

10. Conformational Dynamics Detected by Nuclear Magnetic Resonance NOE Values and J Coupling Constants, H. Kessler, Ch. Griesinger, J. Lautz, A. Müller, W. F. van Gunsteren and H. J. C. Berendsen, *J. Am. Chem. Soc.*, **1998**, *110*, 3393-3396.

11. The Effect of Motional Averaging on the Calculation of NMR-Derived Structural Properties, X. Daura, I. Antes, W. F. van Gunsteren, W. Thiel and A. E. Mark, *PROTEINS: Structure, Function, and Genetics*, **1999**, *36*, 542-555.

12. Use of Selective Trp side Chain Labeling To Characterize Protein-Protein and Protein-Ligand Interactions by NMR Spectroscopy, R. Rodríguez-Mías, M. Pellecchia, *J. Am. Chem. Soc.*, **2003**, *125*, 2892-2893.