

Tesis Doctoral

DISEÑO, SÍNTESIS Y ESTRUCTURA DE DOMINIOS
HELICOIDALES
INFLUENCIA DE LA INTRODUCCIÓN DE AMINOÁCIDOS D

Carmen Giovana Granados Ramírez



Departamento de Química Orgánica
Facultad de Química
Universidad de Barcelona
Barcelona, 2007

INTRODUCCIÓN y OBJETIVOS

Introducción

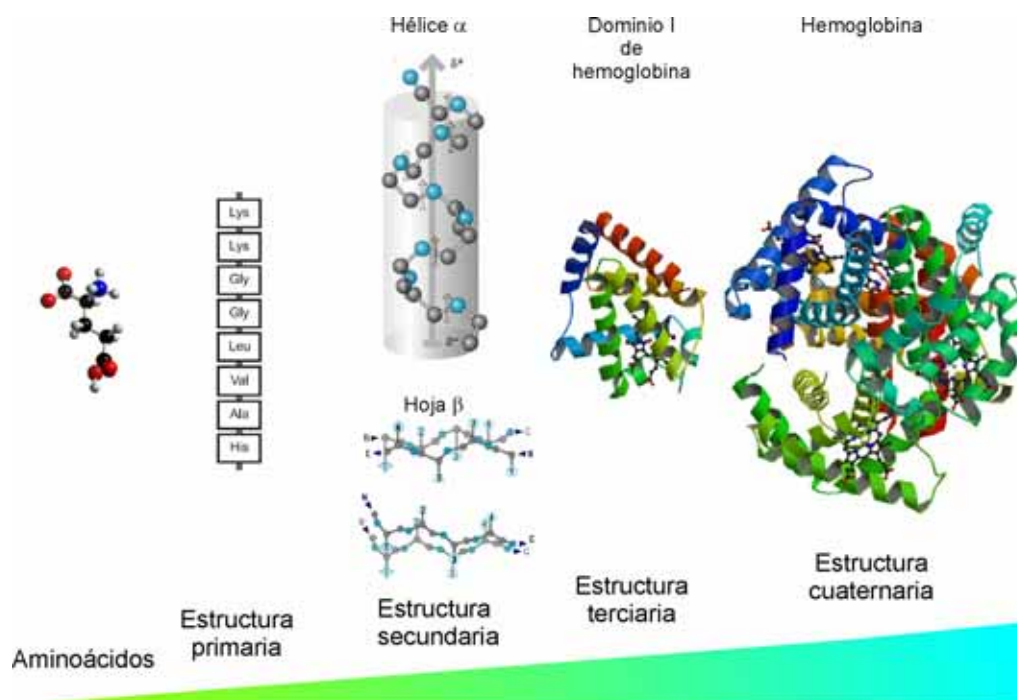
Las proteínas codificadas ribosomalmente están compuestas principalmente de 20 aminoácidos naturales. Con este pequeño número de piezas, las proteínas son capaces de alcanzar una diversidad sorprendente de estructuras, propiedades y funciones. Casi todos los aminoácidos presentes en proteínas son L estereoisómeros con excepción de la glicina que es aquiral. La forma como están organizados los aminoácidos en las proteínas se ha dividido en cuatro niveles estructurales¹. A la secuencia de los diferentes aminoácidos de una proteína se denomina estructura primaria. Las estructuras generadas por la interacción entre los grupos N-H y C=O de la cadena principal estructura definen la estructura secundaria. Entre los principales tipos de estructura secundarias presentes en proteínas se encuentran las hojas β , los giros β y las hélices α . En las hojas β el patrón de enlaces de hidrógeno involucra segmentos extendidos de cadena peptídicas. Los giros β , usualmente sólo se ven envueltos 3 o 4 residuos y consisten en un giro formado por la interacción entre residuos separados una o dos posiciones. En las hélices α los residuos se organizan de manera helicoidal por la interacción consecutiva por enlace de hidrógeno entre el grupo -NH de cada residuo i con el grupo C=O de cada residuo $i+4$. La organización espacial de los elementos de estructura secundaria de la cadena polipeptídicas da lugar a la estructura terciaria de la proteína. La estructura cuaternaria se refiere a la organización de varias cadenas polipeptídicas en un único complejo (Figura 1).

El diseño de proteínas se ha enfocado en obtener complejos pequeños y bien plegados, con el objeto de entender los mecanismos que producen la estabilidad de las proteínas. Posteriormente, este objetivo se ha ampliado hasta también conferir a las proteínas de diseño propiedades y actividades específicas². Los métodos de síntesis química de péptidos y la expresión de proteínas recombinantes han permitido manipular las secuencias peptídicas para obtener nuevas proteínas tanto *de novo*³, como nuevas modificaciones de las proteínas naturales⁴.

El diseño racional de proteínas se basa en el conocimiento detallado de la estructura y función de las proteínas para proponer nuevas variantes; no obstante, es necesario el conocimiento previo de la estructura tridimensional de la proteína para así manipular la secuencia y obtener mejores propiedades o mejor estabilidad. Por ejemplo, la estabilidad térmica de la lisozima T4 aumenta 23°C con respecto a la estructura natural por la introducción de tres enlaces disulfuro⁵. Con la aplicación de métodos de evolución dirigida se ha podido ampliar el espectro de posibles estructuras y propiedades. Con esta técnica se genera una quimioteca de genes que poseen mutaciones al azar de la proteína de interés,

que tras la selección de los individuos con las propiedades deseadas permite amplificar las que son de mayor interés⁶. Recientemente, por medio de esta metodología se ha descrito el uso de esta técnica para seleccionar y aislar enzimas capaces de catalizar la unión de dos moléculas de ARN⁷.

Figura 1. La organización estructural de una proteína viene definida por cuatro niveles estructurales: estructura primaria, estructura secundaria, estructura terciaria y estructura cuaternaria.



Por otro lado, el diseño *de novo*⁸ se basa en la construcción de proteínas de estructura tridimensional bien definida con secuencias no relacionadas con las de las proteínas naturales. Este término también incluye el diseño de secuencias que podrían adoptar un plegamiento determinado sin conocimiento previo de los detalles atómicos de la proteína. La ubicación de residuos hidrofóbicos en la secuencia es suficiente para promover un tipo determinado de plegamiento, patrones determinados que involucran residuos polares e hidrofóbicos, puede llegar a ser suficiente para diseñar quimiotecas de proteínas y obtener nuevas proteínas con estructura globular o fibrilar bien definidas⁹. El diseño *de novo* también se ocupa de la construcción de proteínas simplificadas, compuestas sólo de un pequeño conjunto de aminoácidos. Por ejemplo, la sustitución de gran parte de los aminoácidos del dominio SH3, el cual es una pequeña proteína con estructura de hoja β , por sólo aminoácidos como isoleucina, lisina, ácido glutámico, alanina y glicina, condujo a la

obtención de análogos de la proteína con una secuencia simplificada que sólo conservaba los residuos del núcleo hidrofóbico¹⁰.

A medida que ha aumentado el conocimiento de los mecanismos por los cuales las proteínas se organizan estructuralmente, la introducción de algoritmos computacionales en el diseño molecular se han ido convirtiendo en un instrumento extremadamente útil en el diseño de proteínas estables o con nuevas funciones¹¹. Inicialmente, el diseño totalmente automático de una proteína de 28 residuos, con estructura similar al motivo de dedo de zinc¹². El posterior diseño de haces de hélices estables, para formar controladamente dímeros, trímeros o tetrámeros a partir de una secuencia de 33 residuos diseñada *de novo*¹³. Y más recientemente, el reporte del diseño *de novo* de una proteína de 93 residuos cuya estructura fue diseñada partiendo una topología definida α/β que no ha sido reportada para proteínas naturales¹⁴.

En la naturaleza las proteínas tienen un gran espectro de funciones muchas veces relacionada con la forma en que interacciona con otros elementos. Uno de los desafíos del diseño de nuevas proteínas es la introducción de pequeñas moléculas o cofactores dentro de su estructura¹⁵. La actividad de muchas enzimas esta relacionada con la presencia en sus estructuras de moléculas pequeñas, orgánicas o inorgánicas. A este respecto, se han diseñado proteínas *de novo* capaces de unir metales¹⁶ o algunas moléculas orgánicas como flavinas¹⁷ o grupos hemo¹⁸.

Otro campo del diseño de proteínas que abre un nuevo panorama de estructuras, propiedades y actividades es la incorporación de residuos no comunes en sitios específicos de la cadena polipeptídica. Aunque algunos péptidos naturales incorporan en su secuencia residuos diferentes a los 20 aminoácidos comunes¹⁹, por procesos no ribosomales, estos sólo se encuentran en péptidos pequeños y en posiciones puntuales. Entre estas modificaciones no ribosomales están la incorporación de aminoácidos D, ácidos carboxílicos, residuos N-metilados, así como la adición de anillos heterocíclicos o ácidos grasos²⁰.

Recientemente, también se ha logrado incorporar amino ácidos no naturales dentro de proteínas producidas por organismos como *Escherichia coli*, levaduras y células eucariotas²¹⁻²³. Hasta ahora, alrededor de 30 aminoácidos no naturales se han podido introducir en proteínas producidas *in vivo*. No obstante, las modificaciones que se pueden realizar sólo se pueden realizar en posiciones puntuales de. La incorporación de aminoácidos D en el diseño de proteínas también se ha explorado principalmente cuando afecta posiciones puntuales.

Las modificaciones por aminoácidos D se han realizado usando métodos que involucran síntesis química^{24,25} o métodos de ligación química nativa que hacen uso de síntesis química y expresión de proteínas²⁶⁻²⁸. La incorporación de aminoácidos D en forma consecutiva en la cadena polipeptídica parcial o totalmente se ha explorado en el diseño de péptidos y sólo en algunas proteínas. Estas modificaciones se han realizado mediante síntesis química en el caso de péptidos²⁹ y la síntesis total del enantiómero todo-D de la proteasa de VIH-1³⁰.

El estudio de las modificaciones realizadas, tanto en proteínas naturales como en el diseño *de novo*, está motivado por la necesidad de entender cómo se organizan las proteínas en estructuras, que les confieren las propiedades y actividades diversas que presentan en los seres vivos. La introducción dentro de las cadenas polipeptídicas de otros elementos estructurales diferentes a los aminoácidos comunes hace posible ampliar el espacio conformacional asequible de las proteínas³¹⁻³³, y de esta forma llegar a obtener nuevas estructuras y propiedades.

Objetivos

A lo largo de los tres capítulos de esta tesis se mostrará el desarrollo de diferentes estrategias de diseño aplicadas a la modificación de proteínas o péptidos de origen natural. Desde el punto de vista del diseño racional de análogos no naturales que introducen D aminoácidos hasta la modificación de dominios naturales a partir de una aproximación de química combinatoria. Los objetivos de este trabajo que serán desarrollados en los tres capítulos siguientes:

1. Evaluación de la influencia del cambio de la quiralidad, tanto total como parcialmente, en la actividad y propiedades estructurales del péptido C34, inhibidor de la fusión derivado de gp41 de VIH-1. En primer lugar, se plantea evaluar las consecuencias sobre la actividad del cambio total de la quiralidad de los residuos de C34, tanto en el enantiómero como en el isómero con la secuencia invertida con aminoácidos D (retro-enantiómero). Y en segundo lugar, se busca establecer una estrategia para el diseño de péptidos heteroquirales derivados de C34, que conserven la capacidad de formar complejos con el fragmento N36 de gp41 y, por consiguiente, la propiedad de inhibir la fusión viral.
2. Diseñar una quimioteca de péptidos todo-D para identificar péptidos capaces de formar complejos heteroquirales no covalentes con una secuencia natural, en este caso se utilizó como modelo el fragmento correspondiente al fragmento N-terminal del dominio B de la

proteína A de *S. aureus*.

3. Analizar los efectos, sobre la estructura y la actividad, de mutaciones puntuales en la región de la hélice C-terminal (H₃) del haz de tres hélices presente en el dominio de B de la proteína A de *S. aureus*. Las mutaciones Ser44Phe Leu47Lys y Ser44Phe Leu47Val se realizaron de acuerdo con lo encontrado previamente tras el análisis de la interacción entre el péptido correspondiente a las dos primeras hélices del dominio natural (H₁₋₂) y péptidos de una quimioteca de análogos de H₃³⁴. El efecto de estas modificaciones se analizó desde dos aproximaciones diferentes. Por un lado, la síntesis y caracterización de los dominios análogos en forma covalente; y por otro, el estudio de los complejos no covalentes formados entre los fragmentos H₁₋₂ y H₃ o análogos.

Bibliografía

- (1) Branden, C.; Tooze, J. *Introduction to protein structure*; 2 ed.; Garland Publishing: New York, 1998; 426.
- (2) Robertson, M. P.; Scott, W. G. Biochemistry: Designer enzymes. *Nature* **2007**, *448*, 757-758.
- (3) Baltzer, L.; Nilsson, H.; Nilsson, J. De novo design of proteins-what are the rules? *Chem. Rev.* **2001**, *101*, 3153-3164.
- (4) Leisola, M.; Turunen, O. Protein engineering: opportunities and challenges. *Appl. Microbiol. Biotechnol.* **2007**, *75*, 1225-1232.
- (5) Matsumura, M.; Signor, G.; Matthews, B. W. Substantial increase of protein stability by multiple disulphide bonds. *Nature* **1989**, *342*, 291-293.
- (6) Joyce, G. F. Forty years of in vitro evolution. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 6420-6436.
- (7) Seelig, B.; Szostak, J. W. Selection and evolution of enzymes from a partially randomized non-catalytic scaffold. *Nature* **2007**, *448*, 828-831.
- (8) DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nistri, F.; Lombardi, A. De novo design and structural characterization of proteins and metalloproteins. *Annu. Rev. Biochem.* **1999**, *68*, 779-819.
- (9) Hecht, M. H.; Das, A.; Go, A.; Bradley, L. H.; Wei, Y. De novo proteins from designed combinatorial libraries. *Protein Sci.* **2004**, *13*, 1711-1723.
- (10) Riddle, D. S.; Santiago, J. V.; Bray-Hall, S. T.; Doshi, N.; Grantcharova, V. P. et al. Functional rapidly folding proteins from simplified amino acid sequences. *Nat. Struct. Biol.* **1997**, *4*, 805-809.
- (11) Jones, D. T. Structural biology. Learning to speak the language of proteins. *Science* **2003**, *302*, 1347-1348.
- (12) Dahiyat, B. I.; Mayo, S. L. De novo protein design: fully automated sequence selection. *Science* **1997**, *278*, 82-87.
- (13) Harbury, P. B.; Plecs, J. J.; Tidor, B.; Alber, T.; Kim, P. S. High-resolution protein design with backbone freedom. *Science* **1998**, *282*, 1462-1467.
- (14) Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L. et al. Design of a novel globular protein fold with atomic-level accuracy. *Science* **2003**, *302*, 1364-1368.
- (15) Koder, R. L.; Dutton, P. L. Intelligent design: the de novo engineering of proteins with specified functions. *Dalton Trans.* **2006**, 3045-3051.
- (16) Mulholland, S. E.; Gibney, B. R.; Rabanal, F.; Dutton, P. L. Determination of nonligand amino acids critical to [4Fe-4S]₂^{+/+} assembly in ferredoxin maquettes. *Biochemistry* **1999**, *38*, 10442-10448.

- (17) Sharp, R. E.; Moser, C. C.; Rabanal, F.; Dutton, P. L. Design, synthesis, and characterization of a photoactivatable flavocytochrome molecular maquette. *Proc. Natl. Acad. Sci. U S A* **1998**, *95*, 10465-10470.
- (18) Grosset, A. M.; Gibney, B. R.; Rabanal, F.; Moser, C. C.; Dutton, P. L. Proof of principle in a de novo designed protein maquette: an allosterically regulated, charge-activated conformational switch in a tetra-helix bundle. *Biochemistry* **2001**, *40*, 5474-5487.
- (19) Sieber, S. A.; Marahiel, M. A. Learning from nature's drug factories: nonribosomal synthesis of macrocyclic peptides. *J. Bacteriol.* **2003**, *185*, 7036-7043.
- (20) Marahiel, M. A.; Stachelhaus, T.; Mootz, H. D. Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chem Rev* **1997**, *97*, 2651-2674.
- (21) Jackson, J. C.; Duffy, S. P.; Hess, K. R.; Mehl, R. A. Improving nature's enzyme active site with genetically encoded unnatural amino acids. *J. Am. Chem. Soc.* **2006**, *128*, 11124-11127.
- (22) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. Expanding the genetic code of escherichia coli. *Science* **2001**, *292*, 498-500.
- (23) Wiltshi, B.; Budisa, N. Natural history and experimental evolution of the genetic code. *Appl. Microbiol. Biotechnol.* **2007**, *74*, 739-753.
- (24) Bang, D.; Makhatadze, G. I.; Tereshko, V.; Kossiakoff, A. A.; Kent, S. B. Total Chemical Synthesis and X-ray Crystal Structure of a Protein Diastereomer: [D-Gln 35]Ubiquitin. *Angew. Chem., Int. Ed.* **2005**, *44*, 3852-3856.
- (25) Bang, D.; Gribenko, A. V.; Tereshko, V.; Kossiakoff, A. A.; Kent, S. B. et al. Dissecting the energetics of protein α -helix C-cap termination through chemical protein synthesis. *Nat. Chem. Biol.* **2006**, *2*, 139-143.
- (26) Valiyaveetil, F. I.; Sekedat, M.; MacKinnon, R.; Muir, T. W. Glycine as a D-amino acid surrogate in the K⁺-selectivity filter. *Proc. Natl. Acad. Sci. U S A* **2004**, *101*, 17045-17049.
- (27) Anil, B.; Song, B.; Tang, Y.; Raleigh, D. P. Exploiting the right side of the Ramachandran plot: substitution of glycines by D-alanine can significantly increase protein stability. *J. Am. Chem. Soc.* **2004**, *126*, 13194-13195.
- (28) Anil, B.; Craig-Schapiro, R.; Raleigh, D. P. Design of a hyperstable protein by rational consideration of unfolded state interactions. *J. Am. Chem. Soc.* **2006**, *128*, 3144-3145.
- (29) Wehofsky, N.; Thust, S.; Burmeister, J.; Klusmann, S.; Bordusa, F. All-D-polypeptides: Novel targets for semisynthesis. *Angew. Chem. Int. Ed.* **2003**, *42*, 677-679.
- (30) Milton, R. C.; Milton, S. C.; Kent, S. B. Total chemical synthesis of a D-enzyme: the enantiomers of HIV-1 protease show reciprocal chiral substrate specificity. *Science* **1992**, *256*, 1445-1448.
- (31) Horne, W. S.; Stout, C. D.; Ghadiri, M. R. A heterocyclic peptide nanotube. *J. Am. Chem. Soc.* **2003**, *125*, 9372-9376.
- (32) Sharma, G. V. M.; Jadhav, V. B.; Ramakrishna, K. V. S.; Jayaprakash, P.; Narsimulu, K. et al. 12/10- and 11/13-mixed helices in α/γ and β/γ hybrid peptides containing C-linked carbo- γ -amino acids with alternating α - and β -amino acids. *J. Am. Chem. Soc.* **2006**, *128*, 14657-14668.
- (33) Petersson, E. J.; Craig, C. J.; Daniels, D. S.; Qiu, J. X.; Schepartz, A. Biophysical characterization of a β -peptide bundle: comparison to natural proteins. *J. Am. Chem. Soc.* **2007**, *129*, 5344-5345.
- (34) Pastor, J. J. Una nueva estrategia de química combinatoria para el diseño y modificación de dominios proteicos. In *Departamento de Química Orgánica*; Universidad de Barcelona: Barcelona, 2004; pp 230.