

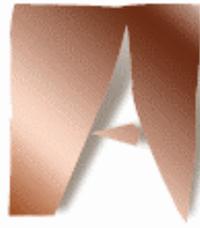


**Universitat  
Autònoma  
de Barcelona**

**Plegamiento y funcionalidad biológica del  
inhibidor de metalocarboxipeptidasas LCI  
(*Leech Carboxypeptidase Inhibitor*)**

**Sílvia Salamanca Seguí**

**Junio 2003**



**Universitat  
Autònoma  
de Barcelona**

Institut de Biotecnologia i Biomedicina  
Departament de Bioquímica i Biologia Molecular

**Plegamiento y funcionalidad biológica del inhibidor de  
metalocarboxipeptidasas LCI  
(*Leech Carboxypeptidase Inhibitor*)**

Memoria presentada por **Sílvia Salamanca Seguí**, licenciada en  
Bioquímica, para optar al grado de Doctora en Ciencias.

Trabajo realizado bajo la dirección del Dr. Josep Vendrell i Roca y la Dra.  
Julia Lorenzo Rivera

**Sílvia Salamanca Seguí**

**Dr. Josep Vendrell i Roca**

**Dra. Julia Lorenzo Rivera**

*“ La principal problemática de los tratamientos anticoagulantes la presenta la interacción frecuente con otros tratamientos que pueden potenciar o deprimir su efecto. Un hecho poco conocido de la historia de los anticoagulantes orales la tiene su relación con el inicio de las vicisitudes, que sufrió previas a su muerte, el antiguo jefe del estado español D.Francisco Franco. Este dictador padeció una tromboflebitis en 1975, siendo tratado por múltiples médicos, uno de ellos instauró un tratamiento con anticoagulantes orales, mientras que el otro, al ser requerido por el intenso dolor que sufría, desconociendo el tratamiento previo recetó fenilbutazona, un antiinflamatorio potente que desplaza al anticoagulante de la albúmina, proteína que actúa como transportadora, aumentando enormemente la concentración anticoagulante en estado libre y sufriendo una súbita intoxicación. El exceso de anticoagulación propició una hemorragia digestiva, con shock y posterior deterioro del estado general del enfermo que le condujo después de múltiples complicaciones a su muerte . ”*

<b>PRESENTACIÓN</b>	1
<b>ABREVIATURAS</b>	3
<b>I. INTRODUCCIÓN</b>	5
<b>I.A. EL PLEGAMIENTO PROTEICO</b>	6
I.A.1. LOS FUNDAMENTOS DEL PLEGAMIENTO PROTEICO DESDE EL POSTULADO DE ANFINSSEN HASTA LA “ <i>NEW VIEW</i> ” (NUEVA VISIÓN)	8
I.A.1.a. El postulado de Anfinsen y la paradoja de Levinthal	8
I.A.1.b. La existencia de intermediarios del plegamiento proteico	9
I.A.1.c. Modelos clásicos de plegamiento proteico	10
I.A.1.d. La nueva visión: el “ <i>energy landscape</i> ” (paisaje de plegamiento) y el “ <i>folding funnel</i> ” (embudo de plegamiento)	12
I.A.2. PLEGAMIENTO EN EL ENTORNO BIOLÓGICO	14
I.A.3. ESTUDIO DEL PLEGAMIENTO PROTEICO DE LAS PROTEÍNAS PEQUEÑAS RICAS EN PUENTES DISULFURO	15
I.A.3.a. El plegamiento del inhibidor pancreático de tripsina bovino (BPTI)	16
I.A.3.b. La diversidad en los caminos de plegamiento de las proteínas pequeñas ricas en puentes disulfuro.	18
I.A.3.c. Estudios de desplegamiento de las proteínas pequeñas de tres puentes disulfuro	20
I.A.3.d. Estudio del plegamiento y desplegamiento de las proteínas pequeñas de cuatro puentes disulfuro	22
<b>I.B. EL LCI Y LAS CARBOXIPEPTIDASAS</b>	24
I.B.1. EL LCI	24
I.B.2. LAS CARBOXIPEPTIDASAS	25
I.B.2.a. Descubrimiento de TAFI y nomenclatura	27
<b>I.C. HEMOSTASIA Y COAGULACIÓN</b>	29
I.C.1. REACCIONES COAGULANTES: LA COAGULACIÓN	31
I.C.1.a. Reacciones anticoagulantes	33
I.C. 2. REACCIONES FIBRINOLÍTICAS: LA FIBRINÓLISIS	36
I.C.2.a. Función de TAFI en la fibrinólisis.	39
I.C.3. TAFI.	40
I.C.3.a. Síntesis y caracterización de TAFI	40

I.C.3.b. Activación de TAFI	41
I.C.3.c. Inactivación de TAFI	42
I.C.3.d. El papel de TAFI en el modelo de la coagulación revisado	43
I.C.4. IMPORTANCIA DE TAFI A NIVEL PATOLÓGICO	45
I.C.4.a. Nuevo enfoque de las terapias coadyuvantes	46
<b>I.D. CONEXIÓN DEL SISTEMA FIBRINOLÍTICO CON EL LCI</b>	<b>47</b>
<b>II. OBJETIVOS</b>	<b>49</b>
<b>III. TRABAJO 1. Major Kinetic Traps for the Oxidative Folding of Leech Carboxypeptidase Inhibitor</b>	<b>51</b>
<b>III.A. INTRODUCTION</b>	<b>51</b>
<b>III.B. EXPERIMENTAL PROCEDURES</b>	<b>52</b>
III.B.1. Materials	52
III.B.2. Control folding experiments	53
III.B.3. Folding of LCI in the presence of redox agents	53
III.B.4. Analysis of folding intermediates of LCI	53
III.B.5. Characterization of the disulfide structure of the major kinetic traps	53
III.B.6. Amino Acid Sequencing and Mass Spectrometry	54
<b>III.C. RESULTS</b>	<b>55</b>
III.C.1. Oxidative folding of LCI in the absence of redox agents	55
III.C.2. Oxidative folding of LCI in the presence of redox agents	56
III.C.3. Structural analysis of the major 3-disulfide kinetic traps	57
<b>III.D. DISCUSSION</b>	<b>59</b>
III.D.1. Diversity of disulfide folding pathway	59
III.D.2. The underlying chemistry of the diversity of disulfide folding pathway	61
III.D.3. The pathway of oxidative folding of LCI	61
III.D.4. Molecular basis for the stability of 3-disulfide kinetic traps of LCI	64
<b>III.E. REFERENCES</b>	<b>65</b>
<b>IV. TRABAJO 2. The Unfolding Pathway of Leech Carboxypeptidase Inhibitor</b>	<b>67</b>
<b>IV.A. INTRODUCTION</b>	<b>67</b>



<b>V.C. RESULTS</b>	87
V.C.1. The importance of properly produced protein inhibitors	87
V.C.2. <i>In vitro</i> effect of LCI on plasma clots	87
V.C.3. LCI levels influence TAFI activity in plasma	90
V.C.4. Effect of LCI on clot structure	90
V.C.5. LCI does not appear to alter fibrin monomers assembly	91
<b>V.D. DISCUSSION</b>	91
<b>V.E. REFERENCES</b>	95
<b>VI. RESUMEN Y DISCUSIÓN GENERAL</b>	99
<b>VI.A. RESUMEN DE LOS TRABAJOS</b>	99
<b>VI.B. DISCUSIÓN GENERAL</b>	101
VI.B.1. El plegamiento del LCI	101
VI.B.2. El efecto fibrinolítico del LCI	106
<b>VII. CONCLUSIONES</b>	111
<b>VIII. BIBLIOGRAFÍA GENERAL</b>	113
<b>IX. APÉNDICE. Material y métodos generales</b>	123
<b>IX.A. Reactivos</b>	123
<b>IX.B. Expresión heteróloga del LCI</b>	123
<b>IX.C. Determinación de la masa molecular del LCI recombinante</b>	124
<b>IX.D. Experimentos de plegamiento y desplegamiento del LCI</b>	124
IX.D.1. Experimentos de replegamiento oxidativo en ausencia de agentes redox (control + y control -)	124
IX.D.2. Plegamiento del LCI en presencia de agentes redox	125
IX.D.3. Desnaturalización y desplegamiento del LCI en presencia de un iniciador de tioles	125
IX.D.4. Análisis de los intermediarios de replegamiento oxidativo del LCI	126
IX.D.5. Análisis estructural (identificación del apareamiento de cisteínas) de las trampas cinéticas III-A y III-B y de los isómeros scrambleds del LCI aislados del desplegamiento	126
IX.D.6. Secuenciación de aminoácidos y Espectrometría de Masas	127

IX.D.7. Espectrometría de fluorescencia del LCI tratado por urea y por hidrocloreuro de guanidinio	127
IX.D.8. Trazado de las curvas de desnaturalización y desplegamiento del LCI	127
<b>IX.E. Experimentos de funcionalidad biológica del LCI</b>	128
IX.E.1. Producción de proteínas recombinantes	128
IX.E.2. Preparación de las muestras de plasma	128
IX.E.3. Ensayo de fibrinólisis	128
IX.E.4. Determinación de la actividad de TAFI en plasma	129
IX.E.5. Microscopía electrónica de rastreo (SEM)	130
IX.E.6. Microscopía electrónica de transmisión (TEM)	131

En la estructura de la presente tesis los tres capítulos de los resultados siguen el formato de artículos de investigación; dos de ellos están publicados y uno se halla en vías de ser enviado a publicación. Por ello se ha mantenido la estructura y el lenguaje original de los artículos. No obstante, al final de la tesis se presenta un anexo de Material y Métodos que pretende recopilar y traducir aquella metodología que ha sido utilizada a lo largo de todo el trabajo.

La presente tesis ha sido realizada en el *Institut de Biotecnologia i Biomedicina* (IBB) de la Universidad Autónoma de Barcelona y su finalización ha sido posible gracias a la amable acogida y la orientación que me brindaron diversos grupos y profesores ajenos a la UAB.

Entre ellos me gustaría destacar al grupo de Hemostasia y Trombosis, en la unidad de Hematología del Hospital de St.Pau, donde la Dra. Montserrat Borrell me inició en la metodología clínica aplicada a los análisis de los factores de coagulación *in vitro*; al grupo de Estructura de Macromoléculas del Centro Nacional de Biotecnología, en la Universidad Autónoma de Madrid, donde el Dr. José M<sup>a</sup>Valpuesta y el Dr. José Carrascosa hicieron posible que obtuviera imágenes de microscopía electrónica de transmisión y, finalmente, al grupo de Química de Proteínas del *Institute of Molecular Medicine*, en Houston (USA). El profesor Rowen Chang ha guiado mi trabajo en el estudio del plegamiento proteico. Gracias a su dedicación los dos primeros trabajos de esta tesis han visto la luz.

**APC** - Proteína C activa

**BPTI** - Inhibidor de tripsina pancreático bovino

**CPA1, CPA2, CPB** - Carboxipeptidasa pancreática A1, A2, B

**CPE, CPM, CPN, CPU, CPR** - Carboxipeptidasa E, M, N, U, R

**pCPB** - Carboxipeptidasa B de plasma

**CTX-III** - Cardiotoxina III

**DIC** - Coagulación intravascular diseminada

**DTT** - Ditiotreitól

***E. Coli*** - *Esterichia Coli*

**EDTA** - Ácido 1-(4-Aminobenzil)etilendiamino-N,N,N',N'-tetraacético

**EGF** - Factor de crecimiento epidérmico

**FPA, FPB** - Fibrinopéptidos A y B

**FV, FVII, FVIII, FIX, FX, FXI, FXII** - Factor V, VII, VIII, IX, X, XI, XI

**FVa, FVIIa, FVIIIa, FIXa, FXa, FXIa, FXIIa** - Factor V, VII, VIII, IX, X, XI, XI activo

**GdnHCl** - Hidrocloruro de guanidinio

**GdnSCN** - Tiocianato de guanidinio

**GEMSA** - Ácido 2-Guanidinoetilmercaptosuccínico

**GSH** - Glutación reducido

**GSSG** - Glutación oxidado

**HEPES** - [Ácido 2-etansulfónico]-N-[2.hidroxi metilo]- piperazina

**HK** - Quininógeno de elevado peso molecular

**HPLC** - Cromatografía líquida de alta presión

**HRG** - Glicoproteína rica en histidina

**ILGF** - Factor de crecimiento parecido a la insulina

**IPTG** - Isopropil-β-D-galactopiranosido

**LBS** - Lugares de unión a lisina

**LCI** - Inhibidor de carboxipeptidasas de sanguijuela

**MALDI-TOF** - *Matrix assisted laser desorption/ionization-time-life-of-flight.*

**O/N** – *Overnight*, cultivo de noche

**PBS** - Suero (tampón) fisiológico salino

**PCI** - Inhibidor de carboxipeptidasas de patata

**PL** - Fosfolípidos

**PPACK** - D-Phe-Pro-Arg clorometil cetona

**SEM** - Microscopía electrónica de rastreo

## *Abreviaturas*

**TAFI** - Inhibidor de la fibrinólisis activable por trombina

**TAP** - Péptido anticoagulante de garrapata

**TEM** - Microscopía electrónica de transmisión

**TF** - Factor tisular

**TFPI** - Inhibidor de la vía del factor tisular

**TM** - Trombomodulina

**t-PA** - Activador tisular del plasminógeno

**TR** - Trombina

**u-PA** - Activador tipo urokinasa

# **I. INTRODUCCIÓN GENERAL**

El universo proteico es de gran complejidad. Muchas proteínas poseen una elevada masa molecular, forman complejos con otras biomoléculas, sufren cambios al llevar a cabo su actividad o son inestables. Esta situación contrasta con las características de un subgrupo de proteínas: aquellas que se hallan estructuradas en un solo dominio, son estables y su función depende de interacciones simples.

Este tipo de proteínas permite, a la vez, estudiar conceptos básicos referidos a la estructura y estabilidad proteica y llevar a cabo análisis sobre su funcionalidad.

El LCI (*leech carboxypeptidase inhibitor*) es una de estas moléculas. Con ella hemos intentado llevar a cabo estudios sobre su plegamiento, utilizándola como molécula modelo, mientras que, en otro aspecto bien diferenciado del trabajo, hemos analizado su capacidad como inhibidor de proteasas y su potencialidad como regulador fisiológico.

A partir de esta dualidad en los apartados de introducción y discusión generales se tratarán los dos temas independientemente, aunque manteniendo su relación a través de la molécula que protagoniza el estudio en ambos casos.

## I.A. EL PLEGAMIENTO PROTEICO

Uno de los retos más importantes a los que se enfrenta la bioquímica actual consiste en discernir las reglas que determinan de qué modo una cadena polipeptídica recién sintetizada adquiere su estructura tridimensional y funcional. De hecho, los rápidos progresos en la secuenciación del genoma, incluyendo al genoma humano, han generado una enorme avalancha de datos que hacen que la solución al problema urja. Al secuenciar un gen, sólo obtenemos una pequeña parte del mensaje a descifrar: es esencial conocer la naturaleza, la estructura y la función de la proteína que está codificada por dicho gen.

Es posible que existan cerca de 100.000 genes en el genoma humano y alrededor de  $10^{11}$  en todos los organismos. Las proteínas codificadas por estos genes, para llegar a ser biológicamente activas deben plegarse en una estructura tridimensional estable. De acuerdo con una clasificación reciente y a pesar de la gran diversidad del proteoma, el número de estructuras diferentes descritas hasta ahora es relativamente pequeño, menos de 700, compuestas por tan sólo 32 dominios proteicos (Thornton et al, 1999). La naturaleza ha creado la complejidad a través de la combinación de un número pequeño de elementos sencillos, del mismo modo que ocurre a otra escala, con los dos elementos más comunes de la estructura secundaria observados en las proteínas: las hélices- $\alpha$  y las láminas- $\beta$ .

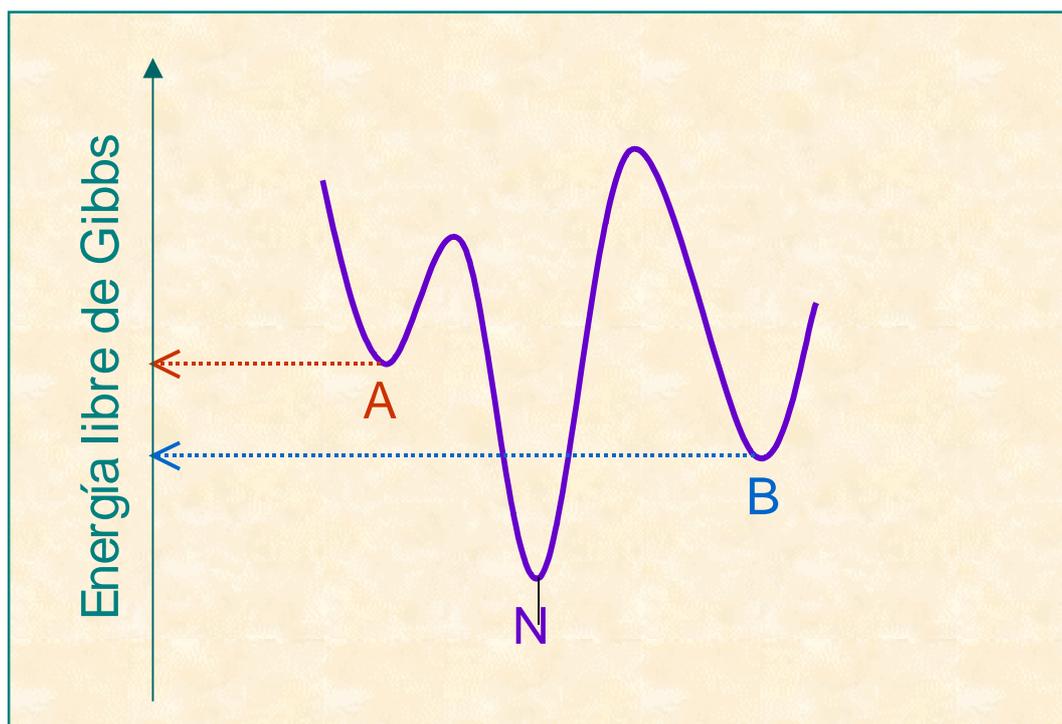
El plegamiento proteico, la segunda traducción del mensaje genético, completa la información transferida del DNA al producto final, la proteína activa. En otras palabras, para entender completamente este proceso, es necesario descifrar el código de plegamiento, la segunda parte del mensaje genético:

DNA $\rightarrow$ RNA $\rightarrow$ cadena polipeptídica $\rightarrow$ ¿proceso de plegamiento? $\rightarrow$ proteína activa.
--

La cuestión del mecanismo del plegamiento proteico ha intrigado a los científicos durante varias décadas. Ya en 1929, una época en la que no se conocía prácticamente nada de la estructura proteica, Wu analizó el reverso del proceso de plegamiento: la desnaturalización. Poco tiempo después, Northrop en 1932 y Anson y Mirsky en 1934-35 tuvieron éxito en la reversión de la desnaturalización de varias proteínas, como la hemoglobina, el quimotripsinógeno y el tripsinógeno.

Las cuestiones fundamentales parecen ser las siguientes: ¿cómo codifica la secuencia para el plegamiento dado que el esqueleto (la sucesión de enlaces peptídicos más los enlaces del C $^{\alpha}$  con N y el carbonilo) de todas las proteínas tiene la misma composición? En otras

palabras, ¿cómo dictan las cadenas laterales el plegamiento? ¿Cómo encuentra una determinada secuencia su estructura nativa entre el enorme número de conformaciones posibles que puede adoptar un polipéptido? ¿Cómo se inicia el proceso de plegamiento y cuál es (son) la(s) vía(s) de plegamiento? Y finalmente, ¿las reglas que podamos deducir de los estudios de plegamiento *in vitro* son válidas para los procesos de plegamiento *in vivo*?



**Figura 1. La controversia de la hipótesis termodinámica versus la hipótesis cinética.** La figura muestra tres estados energéticos donde N es la forma nativa, termodinámicamente estable, que representa el mínimo global de energía, A y B son formas parcialmente plegadas (intermediarios) que representan mínimos locales de la energía libre de Gibbs. El estado B se halla más cercano a N energéticamente (termodinámicamente, por tanto), mientras que el estado A se halla más cercano cinéticamente (la barrera energética a superar es menor) a N, de ahí la nombrada controversia.

Hasta hoy, el plegamiento proteico es un campo extremadamente activo de investigación que involucra los campos de la biología, la química, la bioquímica, la bioinformática y la física. Los principios fundamentales tienen aplicaciones prácticas en los recientes avances de la investigación sobre genomas, en la comprensión de diversas patologías y en el diseño de proteínas *de novo* con funciones especiales.

## **I.A.1. LOS FUNDAMENTOS DEL PLEGAMIENTO PROTEICO. DESDE EL POSTULADO DE ANFINSSEN HASTA LA “NEW VIEW” (NUEVA VISIÓN)**

### **I.A.1.a. El postulado de Anfinsen y la paradoja de Levinthal.**

El logro considerable de C. Anfinsen y su grupo al replegar la ribonucleasa reducida y desnaturalizada recobrando el enzima plenamente activo marcó el inicio de la era moderna en el estudio del plegamiento proteico. A partir de sus resultados el autor concluía que “toda la información necesaria para conseguir la conformación nativa de una proteína en un ambiente determinado se halla contenida en su secuencia aminoacídica” (Anfinsen, 1973). Es decir, que la estructura nativa de una proteína es termodinámicamente estable y representa un mínimo en la energía libre de Gibbs. A partir del postulado de Anfinsen surgió el corolario del control termodinámico en el plegamiento proteico. Este proceso fue discutido por Levinthal, en consideración al poco tiempo que requiere el proceso de plegamiento en la naturaleza. Según él, una búsqueda al azar (representada por la hipótesis del control termodinámico de Anfinsen) de todas las conformaciones proteicas posibles no puede darse en realidad, dado el enorme tiempo que requeriría. Tomemos por ejemplo una cadena polipeptídica de 100 aminoácidos; si asumimos tan sólo dos posibles conformaciones para cada residuo, existen  $10^{30}$  posibles conformaciones para la cadena completa. Si tan sólo se requieren  $10^{-11}$  segundos para convertir una conformación en otra, una búsqueda al azar de todas las conformaciones requeriría  $10^{11}$  años. Una cifra completamente irreal en un contexto biológico donde el tiempo de plegamiento es del orden de segundos o minutos. Además, es evidente que la evolución ha hallado una solución efectiva a este problema combinatorial. A este principio se le conoce como la paradoja de Levinthal (Levinthal, 1968). De ahí que el autor sugiera que las proteínas no realizan una búsqueda al azar, sino que siguen un camino concreto de plegamiento. Dicho camino vendría conducido por un control cinético del plegamiento proteico (en contraposición al control termodinámico, postulado por Anfinsen). La figura 1 ilustra las discrepancias entre la hipótesis termodinámica y la hipótesis cinética. A partir de este punto, se han dirigido numerosos esfuerzos para hallar los posibles caminos de plegamiento de las proteínas, de los que han surgido diversos modelos.

### **I.A.1.b. La existencia de intermediarios del plegamiento proteico.**

Durante mucho tiempo se consideró que la transición entre proteína plegada y desplegada bajo condiciones fisiológicas era un proceso en dos estados en que tan sólo se hallarían representados de manera significativa la especie nativa y la especie desnaturalizada. Sin embargo, gracias a los estudios cinéticos que se realizaron con el objetivo de explicar la paradoja de Levinthal se demostró la existencia de intermediarios de plegamiento.

Por consiguiente, podían darse dos casos al estudiar las cinéticas de plegamiento:

1. Que los datos cinéticos de plegamiento se ajustaran a una exponencial simple. En este caso, la proteína se plegaría siguiendo un modelo de dos estados: de la especie desnaturalizada a la especie nativa.
2. Que los datos cinéticos de plegamiento se ajustaran a una exponencial múltiple. En este otro caso la proteína se plegaría siguiendo un modelo multiestado: de la especie desnaturalizada al intermediario o intermediarios y de ahí a la proteína nativa.

Los intermediarios, según la hipótesis del control cinético, corresponderían a mínimos locales de la energía libre de Gibbs, de ahí su relativa estabilidad.

Una observación más profunda mostró que no todos los intermediarios poseían la misma estabilidad y concentración, y que, además, no todos ellos conducían a la proteína nativa. Por ello se separó a los intermediarios de plegamiento en dos tipos:

1. Intermediarios “*off-pathway*” (fuera del camino).
2. Intermediarios “*on pathway*” (dentro del camino).

Mientras que los primeros no darán como fruto la conformación nativa de la proteína, los segundos deberían hallarse en el camino “correcto” del plegamiento y nos marcarían, por tanto, la vía a través de la cual la proteína hallaría su conformación nativa. Como consecuencia, la identificación y caracterización estructural de los intermediarios de plegamiento se vio transformada en un tema de importancia crucial.

Por otro lado, si bien los intermediarios de “fuera del camino” no aportan información sobre la vía de plegamiento, nos proporcionan una explicación coherente acerca de los fenómenos de plegamiento incorrecto o malplegamiento (*misfolding*): constituyen trampas cinéticas donde la proteína puede permanecer y, en casos extremos, agregar.

### **I.A.1.c. Modelos clásicos de plegamiento proteico.**

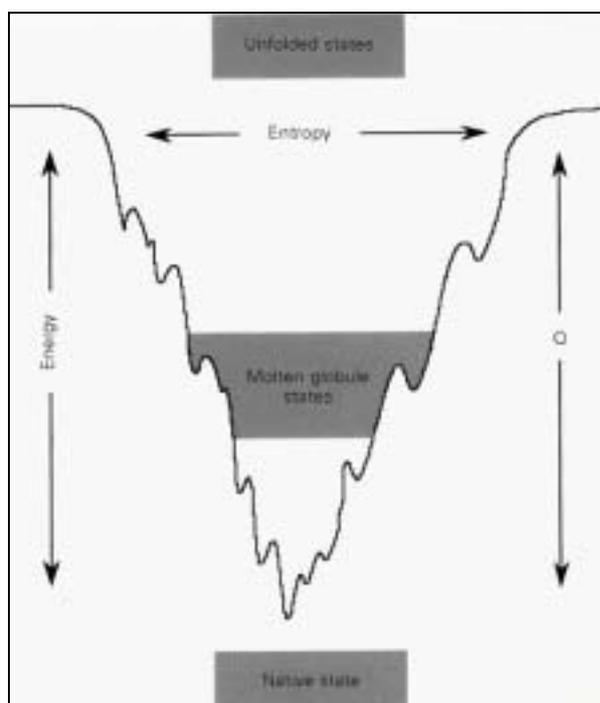
Para discernir cómo una cadena polipeptídica puede evitar verse atrapada en la paradoja de Levinthal, se han propuesto diferentes modelos de plegamiento tanto a partir de observaciones experimentales (Kim & Baldwin, 1990) como desde consideraciones teóricas (Karplus & Sali, 1995). Aquí citamos los más representativos:

1- El modelo de nucleación-propagación. En este modelo se forman uno o más núcleos estructurales, alrededor de los cuales crece (“se propaga”) la estructura. La etapa limitante del plegamiento es la formación de dichos núcleos. Fersht (Fersht, 1997) refinó dicho modelo a la luz de los estudios cinéticos sobre el plegamiento en ribonucleasas añadiendo la idea de la formación de un núcleo local débil que se estabiliza por interacciones de largo alcance.

2- El modelo de “*framework*”. Tras el primer suceso, la nucleación, se forman las estructuras secundarias que a su vez se asociarán para formar estructuras suprasecundarias, de ahí a los dominios y eventualmente al monómero activo. Una jerarquía semejante es la que se utiliza para describir la estructura proteica (Jaenicke, 1987, Kim & Baldwin, 1990).

3-El modelo de difusión-colisión. Se desarrolló en 1976 por Karplus y Weaver y se reconsideró en 1994 a la luz de datos experimentales más recientes (Karplus & Weaver, 1994). En un modelo semejante, la nucleación sucede simultáneamente en partes diferentes de la cadena polipeptídica generando microestructuras que tienen un tiempo de vida controlado por la difusión de un segmento. Según este modelo el plegamiento de una cadena polipeptídica que contenga de 100 a 200 aminoácidos puede ocurrir en un tiempo muy corto, menos de un segundo, y el plegamiento se da a través de varios pasos de difusión-colisión.

4-El modelo del colapso hidrofóbico. El primer evento del plegamiento proteico consiste en un colapso a través de interacciones hidrofóbicas de largo alcance y sucede previa a la formación de una estructura secundaria (Dill, 1985). La idea se originó a partir del trabajo de Kauzmann que considera el efecto hidrofóbico como la fuerza líder en el plegamiento y la estabilización proteica. Una de las aportaciones del modelo del colapso hidrofóbico fue el concepto del “*molten globule*” (glóbulo fundido). El glóbulo fundido es un intermediario temprano que aparece en los primeros estadios del plegamiento proteico y que se caracteriza por poseer un elevado grado de estructura secundaria. Tras su formación se llevaría a cabo una reordenación de un número pequeño de estados condensados. La figura 2 muestra el estado energético del glóbulo fundido.



**Figura 2. Representación esquemática bidimensional de la energía en el camino del plegamiento.** La anchura representa la entropía y la profundidad la energía libre.  $Q$  es la fracción de contactos nativos para cada estado energético. Nótese la posición de los intermediarios “*molten globule*”.

Recientemente ha sido identificado otro intermediario predecesor al glóbulo fundido por Ptitsyn (Ptitsyn, 1995) y Uversky y Ptitsyn (Uversky & Ptitsyn, 1996). Esta especie, menos compacta que un *molten globule* tiene un contenido significativo de estructura secundaria pero menor que la del *molten globule*, y muestra regiones hidrofóbicas accesibles a solvente. Jeng y Englander lo llamaron “*pre-molten globule*” (Jeng & Englander, 1991). La formación rápida de intermediarios de transición, tanto glóbulo fundido o pre-glóbulo fundido, o ambos, con un elevado contenido de estructura secundaria y un porcentaje pequeño de una estructura terciaria fluctuante, tiene el soporte de numerosas observaciones (Matouschek & Fersht, 1991, Fink, 1995). De todos modos, es posible que la formación de estas especies sea precedida por un evento que aún no podemos detectar.

Si condensáramos genéricamente los modelos clásicos concluiríamos que los sucesos de los primeros estadios del plegamiento de una proteína consisten en un colapso hidrofóbico acompañado, o no, de la formación de estructuras secundarias, dependiendo de la proteína. Los sucesos intermedios en escala de tiempo de plegamiento suceden después de la formación del *molten globule* y antes de la etapa limitante del proceso de plegamiento que genera la estructura nativa y funcional. En estas fases intermedias, se puede observar la aparición de sitios de unión a sustratos o a ligandos.

En el paso final del plegamiento, la proteína adquiere su conformación nativa con la emergencia de sus propiedades funcionales. Estos sucesos finales corresponden a la ordenación precisa de los elementos de estructura secundaria, el correcto empaquetamiento del núcleo hidrofóbico, el apareamiento correcto de los dominios en las proteínas multidominio, el reordenamiento de los puentes disulfuro, la isomerización cis-trans de la prolina y el ensamblaje de las subunidades en proteínas oligoméricas.

#### **I.A.1.d. La nueva visión: el “*energy landscape*” (paisaje de plegamiento) y el “*folding funnel*” (embudo de plegamiento).**

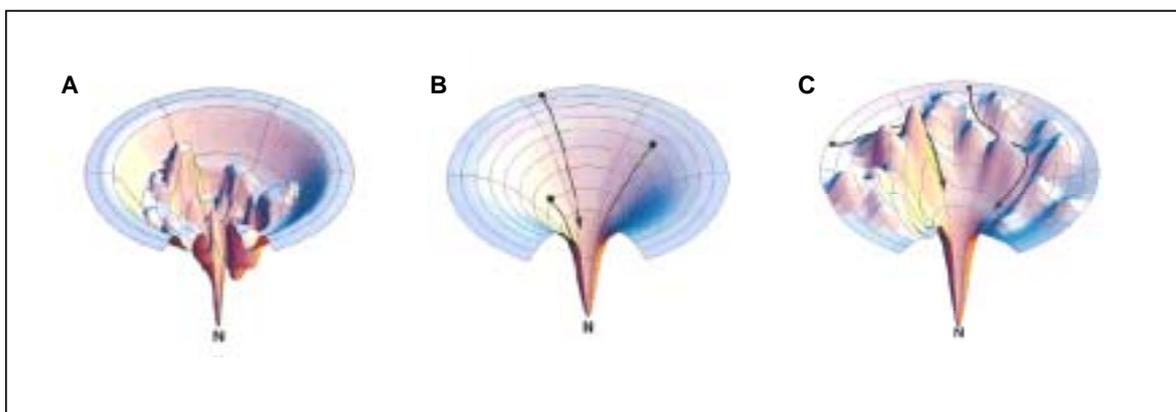
Como acabamos de ver, a lo largo de los años se han propuesto numerosos modelos fenomenológicos; sin embargo, recientemente ha emergido un modelo unificado de plegamiento proteico basado en la superficie energética efectiva. Dicho modelo ha sido llamado “*new view*” (nuevo enfoque o nueva visión) y surge de estudios teóricos. De hecho, han surgido varios modelos teóricos para sobrepasar o explicar la paradoja de Levinthal que, aunque tengan limitaciones debidas al tiempo de computación, pueden simular el plegamiento desde la proteína desplegada a la estructura nativa. Usando modelos computacionales, Karplus y colaboradores (Dobson et al, 1998) estudiaron un polipéptido de 27 aminoácidos y los resultados indicaron que de todas las conformaciones posibles que surgen al azar, sólo una minoría encuentran su estado nativo fácilmente, mientras la mayoría de estas conformaciones al azar jamás lo hallan, debido al gran salto energético que deben realizar para superar la barrera entre el mínimo energético local que ellas representan y el mínimo energético global que representa el estado nativo. Bajo estas condiciones, el plegamiento puede ser muy rápido. El problema del salto energético se supera al focalizar la proteína su plegamiento en fragmentos pequeños y actuar estos como unidades de plegamiento. De este modo, se puede superar la paradoja de Levinthal mediante la formación rápida y simultánea de microestructuras en diferentes regiones de la proteína, restringiendo el número de conformaciones posibles durante el ensamblaje subsiguiente de la estructura final.

La llamada *new view* ha evolucionado durante los últimos años gracias a datos tanto teóricos como experimentales a través del uso de modelos mecánicos simplificados y viene ilustrado por el modelo del “embudo de plegamiento” que presentaron Wolynes y colaboradores (Wolynes et al, 1995). El modelo se presenta en términos de energía de plegamiento y describe el comportamiento termodinámico y cinético de la transformación de un conjunto de moléculas desplegadas a un estado predominantemente nativo (figura 3). Del “paisaje energético” emerge una gran variedad de comportamientos de plegamiento

dependiendo de los parámetros y condiciones energéticas. Los autores sugieren que la tasa de plegamiento se ralentiza por “ondas” o “huecos” que corresponden a los mínimos locales de energía donde se encuentran los intermediarios estables. De acuerdo con este modelo, hay pequeñas vías paralelas y cada cadena polipeptídica seguiría su propia vía individual. Cercano al fondo del embudo de plegamiento, el número de conformaciones de la proteína disminuye como lo hace a su vez la entropía. Cuanto más inclinada sea la pendiente más rápido será el plegamiento. En la figura 3b se representa un embudo ideal y liso en tres dimensiones; en este caso, la proteína se puede plegar muy rápidamente exhibiendo un comportamiento en dos estados. En contraste, en la figura 3a, se representa un paisaje arisco con trampas cinéticas que representan barreras locales de energía. En dicho caso, el plegamiento será más lento y en multiestado (Dill & Chan, 1997). Cuando las barreras energéticas locales son lo suficientemente elevadas, las moléculas proteicas pueden quedarse atrapadas y, posiblemente, agregar.

La nueva visión ha reemplazado progresivamente la idea de las vías de plegamiento. La metáfora del “paisaje energético” nos provee de un marco conceptual para entender a ambas cinéticas de plegamiento, la de dos estados y la de multiestado, así como el mal plegamiento proteico y su posible agregación.

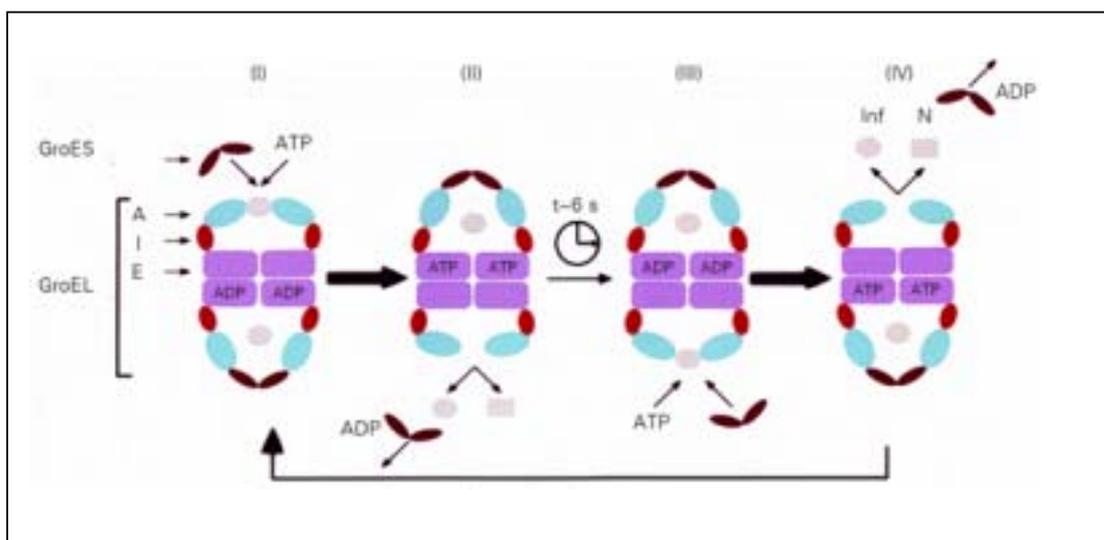
Diversos resultados experimentales apoyan esta teoría y demuestran de manera creciente que la cadena polipeptídica inicialmente extendida se pliega a través de una población heterogénea de intermediarios parcialmente plegados en un equilibrio fluctuante (Dobson et al, 1998).



**Figura 3. Diferentes paisajes energéticos representados en forma de embudos de plegamiento.** El eje vertical representa la energía libre interna. Cada conformación posible está representada por un punto en el paisaje, mientras que los dos ejes horizontales muestran los grados de libertad. A) Representación de un paisaje abrupto, con colinas y trampas, donde probablemente la cinética de plegamiento sea una exponencial múltiple. B) Representación de un paisaje ideal, y por tanto, poco probable. El paisaje es liso, el plegamiento es rápido. Los caminos de plegamiento y desplegamiento de la proteína serán los mismos. C) Representación de un embudo de plegamiento de una proteína de plegamiento rápido. El plegamiento se halla limitado por la pendiente, por tanto, de esta forma la proteína puede hallar vías de plegamiento simultáneas fácilmente, tal y como muestra la figura.

## I.A. 2. PLEGAMIENTO EN EL ENTORNO BIOLÓGICO

Las reglas básicas del plegamiento proteico se han establecido en estudios *in vitro*, y es aceptado comúnmente que el modelo de plegamiento que se ha dilucidado de ellos es válido para entender el mecanismo por el que una cadena polipeptídica naciente alcanza su estructura tridimensional en un entorno celular. Sin embargo, este punto de vista se vio modificado con el descubrimiento de las chaperonas moleculares en 1987, proteínas que ayudan al correcto plegamiento de otras proteínas. Desde dicha fecha, se ha aportado muchísima información al respecto. Los datos estructurales que poseemos actualmente junto a los estudios funcionales han arrojado mucha luz sobre el funcionamiento de las chaperonas en su asistencia al plegamiento proteico en la célula. Se ha podido establecer que no se requiere información adicional a la proteína para que una chaperona pliegue correctamente una proteína. Las chaperonas moleculares, por su asociación transitoria con péptidos nacientes desestabilizados por el estrés o con proteínas translocadas, juegan un papel importante en la prevención del malplegamiento y la subsiguiente agregación. Las chaperonas no interaccionan con proteínas ya plegadas, sino que se asocian transitoriamente mediante interacciones hidrofóbicas a un intermediario de plegamiento de los primeros estadios. No llevan información capaz de dirigir una proteína a asumir una estructura diferente que la que viene dictada por su secuencia aminoacídica. Aquí se ilustra brevemente el funcionamiento de una de las chaperonas mejor conocidas hasta hoy día: el sistema Gro-El-Gro-Es (figura 4).



**Figura 4. El ciclo GroES-GroEL.** Inf representa la proteína desplegada, N sería la proteína plegada. A es el dominio apical (en azul), el cual se une a la proteína desplegada y a GroES, I es el dominio intermedio (en rojo) y E es el dominio ecuatorial (en magenta) el cual une e hidroliza el ATP.

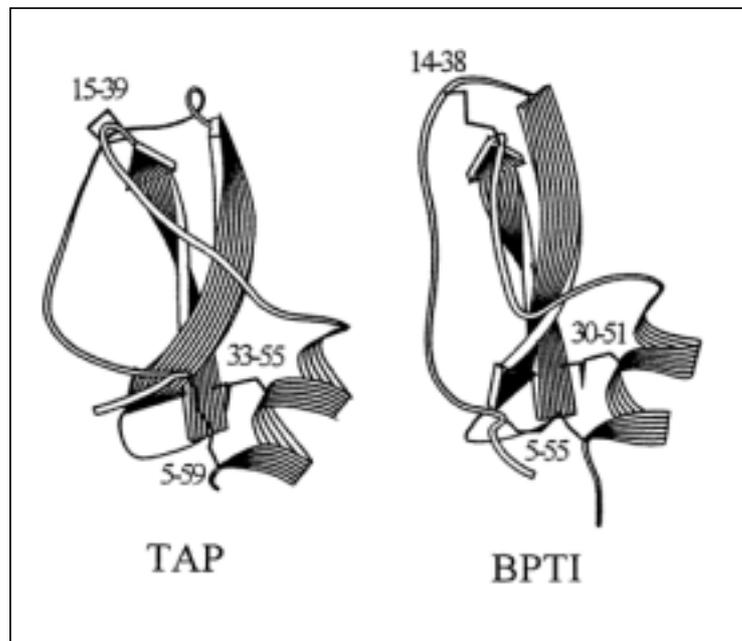
Como ya hemos comentado, los experimentos de plegamiento proteico *in vitro* han demostrado que el mal plegamiento resulta de una competición cinética entre el camino correcto de plegamiento y un proceso fuera del camino, resultante de la aparición de mínimos locales de energía en el embudo de plegamiento ya descrito. Cuando la formación de la estructura correcta está cinéticamente favorecida *in vivo*, no se requiere de chaperonas moleculares. Este hecho se ha demostrado no sólo para proteínas monoméricas (Neira et al, 1997), sino también para proteínas diméricas como el represor p22-Arc. *In vivo*, se calcula que menos de un 15% de las proteínas biosintetizadas necesitan de la asistencia de chaperonas moleculares para adquirir su estructura nativa (Ellis & Hartl, 1999). Por tanto, parece razonable la aproximación al estudio del plegamiento mediante experimentos *in vitro*. Un caso particularmente útil viene representado por las proteínas monodominio. Entre ellas, se destaca a las proteínas pequeñas ricas en puentes disulfuro por ser el objeto de estudio del presente trabajo.

### **I.A.3. ESTUDIO DEL PLEGAMIENTO PROTEICO DE LAS PROTEÍNAS PEQUEÑAS RICAS EN PUENTES DISULFURO**

El plegamiento proteico de las proteínas pequeñas ricas en puentes disulfuro viene guiado por la formación y reorganización de sus puentes disulfuro. Esto las distingue del resto de proteínas, en las que las interacciones no covalentes juegan el papel predominante, y hace que constituyan un grupo aparte para el que los modelos descritos en el apartado I.A.1.c. deben ser revisados y adaptados.

Las vías de plegamiento de las proteínas pequeñas ricas en puentes disulfuro se analizan estudiando el mecanismo de formación de los puentes disulfuro en los intermediarios de plegamiento que aparecen a lo largo del proceso. Por este motivo es de suma importancia la caracterización estructural de dichos intermediarios, lo que es posible utilizando métodos de seguimiento del plegamiento oxidativo (Creighton & Goldemberg, 1984, Wedemeyer et al, 2000, ...). Mediante esta técnica, las proteínas se reducen completamente y se desnaturalizan en presencia de un agente reductor y uno desnaturalizante. Después de la eliminación de ambos agentes caotrópicos se inicia el replegamiento de la proteína, completamente reducida y desnaturalizada, en presencia de un tampón redox. El camino de replegamiento se puede caracterizar finalmente y definir por la heterogeneidad y las estructuras de los intermediarios con puentes disulfuro que se acumulan a lo largo del proceso del replegamiento oxidativo. Es

posible “atrapar” o “capturar” los intermediarios de plegamiento pueden mediante diversas técnicas, siendo una de las más efectivas la captura en medio ácido. Mediante esta técnica, por la que se acidifica la muestra a pH 3-4, se puede parar el proceso de plegamiento de una proteína de manera sencilla (Weissmann & Kim, 1991, Creighton, 1990). Además, la captura de disulfuros es reversible (no como una técnica anterior que permitía atrapar intermediarios mediante la derivatización por alquilación) y permite continuar la monitorización del proceso de plegamiento (experimentos *stop/go*). Esta técnica permite, asimismo, estudiar la influencia de los agentes redox en el proceso de plegamiento y la eficiencia en la recuperación de la proteína nativa.



**Figura 5. Estructuras nativas del TAP y del BPTI.** Ambas proteínas poseen estructuras secundarias similares y adoptan conformaciones tridimensionales comparables. Las seis cisteínas que forman los tres puentes disulfuro se hallan numeradas.

### I.A.3.a. El plegamiento del inhibidor pancreático de tripsina bovino (BPTI).

El inhibidor pancreático de tripsina bovino (BPTI, *bovine pancreatic trypsin inhibitor*) es una proteína monodominio compacta estabilizada por tres puentes disulfuro (figura 5). El mecanismo del desplegamiento y replegamiento del BPTI, dilucidado por captura y caracterización estructural de los intermediarios que contenían puentes disulfuro, ha sido sujeto de numerosas investigaciones y representa uno de los modelos mejor y más extensamente caracterizados. El modelo original fue descrito por primera vez por Creighton e identificaba siete intermediarios predominantes, dos de los cuales contenían un puente disulfuro y cinco que contenían 2 puentes disulfuro, con un 75% de dichos puentes disulfuro

nativos (Creighton, 1974, 1975, 1977). En el modelo revisado propuesto por Weissmann y Kim, se describen cinco intermediarios mayoritarios, dos que contienen un disulfuro y 3 que contienen 2 disulfuros, conteniendo todos ellos exclusivamente puentes disulfuro nativos. De la prevalencia de los puentes disulfuro nativos a lo largo del camino del plegamiento se deduce un importante principio: que las interacciones covalentes no específicas que estabilizan el BPTI nativo y las estructuras locales del BPTI juegan un papel crucial en la guía del plegamiento en sus primeros estadios por lo que dictan la formación de un determinado y escaso número de intermediarios mayoritarios que admiten únicamente puentes disulfuro nativos.

Según Creighton, a partir de sus trabajos sobre el plegamiento del BPTI se podía deducir la existencia de un camino específico de plegamiento, en donde la proteína pasaría a través de una serie de intermediarios específicos en su camino hacia el estado nativo. Ese fue un punto de inicio de controversia, ya que Jonathan Weissman y Peter Kim obtuvieron datos posteriores que indicaban que los intermediarios más comunes diferían a los hallados por Creighton (Weissmann y Kim, 1991, 1992). Aunque Creighton estaba de acuerdo con los datos experimentales aportados por Weissman y Kim, no lo estaba, en cambio, con la interpretación de dichos datos ni en el mecanismo de plegamiento propuesto por dichos autores. La cuestión a discutir era cuál de los dos tipos de intermediarios (del tipo “nativo” o “no nativo”) dirigía el proceso de plegamiento, hecho ejemplificado en el rol cinético de uno de los intermediarios de dos puentes disulfuro, que contiene 2 de los disulfuros nativos del BPTI (Cys<sup>30</sup>-Cys<sup>51</sup>, Cys<sup>5</sup>-Cys<sup>55</sup>) (N<sub>SH</sub><sup>SH</sup>) y que es el precursor inmediato del BPTI nativo. La formación del tercer puente disulfuro, Cys<sup>14</sup>-Cys<sup>38</sup>, completa el plegamiento y es el paso final del proceso de plegamiento del BPTI. Aunque todos los autores están de acuerdo en que el paso limitante sería el de la conversión de este intermediario de 2 puentes disulfuro a la proteína nativa de tres, Creighton opina que son los intermediarios no-nativos los que tienen importancia. Según Creighton, el intermediario nativo y su conformación pueden causar la mayor dificultad a la hora de formar el tercer puente disulfuro. Creighton aduce que las moléculas no nativas son más flexibles y por tanto mejores para el plegamiento, haciendo posible que la molécula se pliegue rápidamente. Según Creighton, Kim y Weissman confundieron el que un intermediario sea cinéticamente estable con que tenga importancia termodinámica. Por otro lado, Kim y Weissman concluyeron que el énfasis en las especies no nativas como un factor determinante del plegamiento del BPTI no se hallaba justificado. Por tanto, los hechos esenciales de la reacción de plegamiento del BPTI, incluyendo el paso de reordenamiento, vienen determinados en gran parte por la estructura nativa que se observa en

la formación de sus intermediarios. Subsecuentemente, se demostró que el plegamiento nativo del BPTI podía obtenerse sin la presencia de especies no nativas de intermediarios en el medio, y que no había motivo para creer que la secuencia aminoacídica del BPTI codificara para interacciones no nativas que fueran importantes para el plegamiento proteico. Tanto estudios cinéticos como estructurales daban soporte a su conclusión de que las interacciones nativas juegan un papel predominante en el plegamiento del BPTI.

Durante mucho tiempo se ha tomado el modelo del plegamiento del BPTI (en cualquiera de sus versiones) como un patrón general que debían seguir todas las proteínas pequeñas ricas en puentes disulfuro. Estudios posteriores, como se explica a continuación, demostraron una mucho mayor variedad en las vías de plegamiento, y pusieron en evidencia el hecho de que el BPTI parece ser un caso único en la naturaleza.

### **I.A.3.b. La diversidad en los caminos de plegamiento de las proteínas pequeñas ricas en puentes disulfuro.**

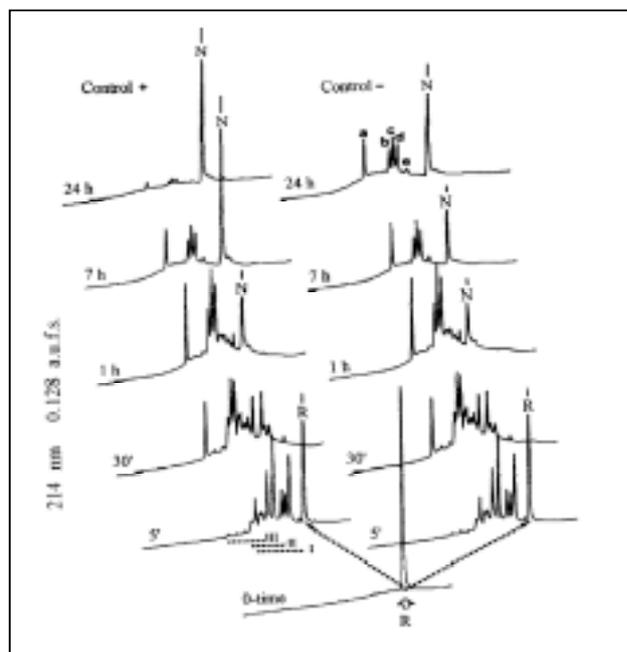
La aplicación de la técnica del plegamiento oxidativo ha permitido la elucidación del camino de plegamiento de varias proteínas de tres puentes disulfuro, entre ellas el BPTI (pionera, como acabamos de ver) (Creighton 1978,1984, 1986, 1990, Weissman & Kim, 1991), la hirudina (Chatreinet & Chang, 1993), el TAP (*tick anticoagulant peptide*) (Chang, 1996, Chang & Ballatore, 2000), el PCI (*potato carboxypeptidase inhibitor*) (Chang et al, 1994, Venhudova et al, 2001), el EGF (*epidermal growth factor*) (Wu et al, 1998, et al, 2001), el ILGF (*insuline like growth factor*) (Narhi et al,1993, Yang et al, 1999) entre otras. Los resultados obtenidos demuestran que incluso entre las proteínas pequeñas de tres puentes disulfuro, los mecanismos de plegamiento pueden variar substancialmente. Estas diferencias vienen ilustradas por: a) la tasa de heterogeneidad de los intermediarios de plegamientos, b) la predominancia de intermediarios que contienen puentes disulfuro nativos, c) la acumulación de intermediarios “*scrambled*” de 3 disulfuros como intermediarios. Los isómeros *scrambled* (“mezclado”) son especies completamente oxidadas que contienen, al menos, dos puentes disulfuro no nativos (en el caso de proteínas de tres puentes disulfuro).

Es decir, que todas estas proteínas difieren del camino de plegamiento del BPTI en que sus intermediarios de plegamiento, incluyendo especies de uno y dos disulfuros, son mucho más heterogéneos que los que se describen para el BPTI y, aparte del EGF, no existe ninguna evidencia de que haya, además, una acumulación importante de intermediarios predominantes que lideren el plegamiento en una vía concreta y única. La diferencia más remarcable, de todos modos, es la presencia de estos isómeros de 3 puentes disulfuro llamados

*scrambled*, como intermediarios de plegamiento, rasgo que no se observa en absoluto en el BPTI. Los intermediarios *scrambled* existen en concentraciones elevadas a lo largo de los caminos de plegamiento de la hirudina, el PCI, el TAP y también en el EGF.

Un caso extremo, en contraposición al modelo del BPTI, y bien estudiado en nuestro laboratorio, es del PCI, para el cual y en presencia de 0.5 mM cistina, más del 98% de la proteína se encuentra atrapada en formas *scrambled*, incapaces de reordenar sus puentes disulfuro y de alcanzar la estructura nativa a menos que se hallen tioles libres a su alrededor como catalizadores. El proceso oxidativo de estas proteínas no se halla guiado por la acumulación de intermediarios mayoritarios con estructuras de la proteína nativa. Las vías de plegamiento del PCI contiene un estadio inicial de empaquetamiento no específico, que conduce a la formación de las especies *scrambled* como intermediarios de plegamiento. A continuación la reorganización y la consolidación de las especies *scrambled* permite alcanzar su estructura nativa. En la figura 6 podemos observar el camino de plegamiento oxidativo del PCI.

Entonces, parece que pudiera haber dos grupos de proteínas: las de plegamiento tipo BPTI (o como en el caso del EGF, que a pesar de hallarse especies *scrambled*, se hallan intermediarios mayoritarios) o las de plegamiento tipo PCI (hirudina, TAP, ILGF).



**Figura 6. Camino del plegamiento oxidativo del PCI.** Análisis por HPLC de los intermediarios de plegamiento del PCI atrapados en medio ácido. Se dejó replegar al PCI reducido y desnaturalizado, en dos condiciones: control – y control +, que indican ausencia y presencia, respectivamente, de  $\beta$ -mercaptoetanol (promotor de la formación de puentes disulfuro). Los intermediarios de plegamiento del PCI son muy heterogéneos: las especies de uno, dos y tres disulfuros (*scrambled*) están superpuestas. Las especies *scrambled* se eluyen en las fracciones a-e.

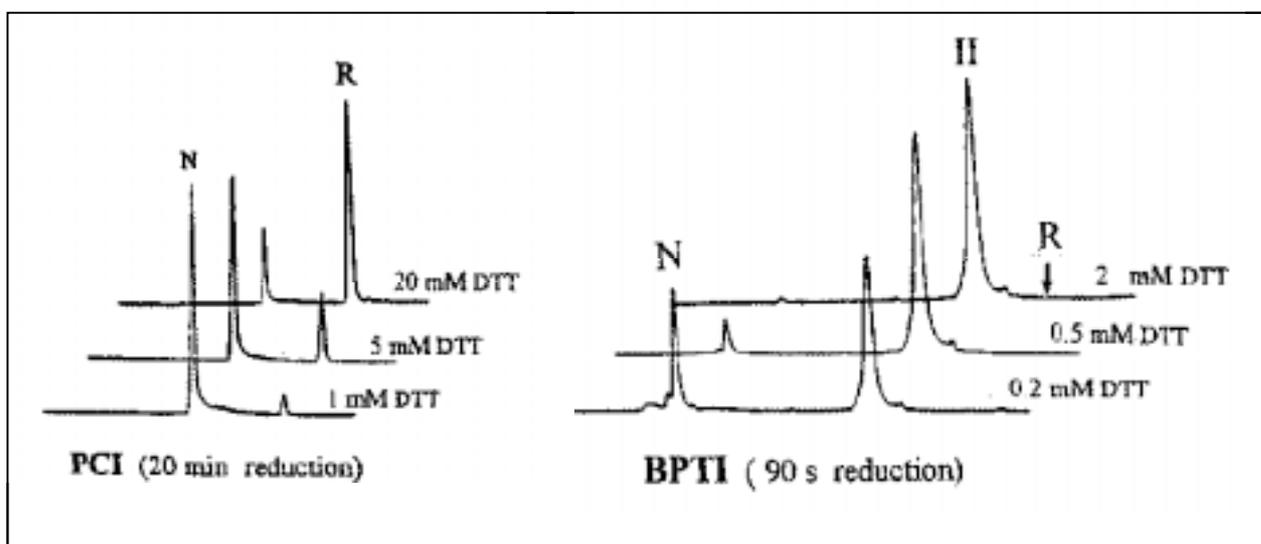
### **I.A.3.c. Estudios de desplegamiento de las proteínas pequeñas de tres puentes disulfuro.**

La causa subyacente de la diversidad de estos mecanismos de plegamiento ha sido asociada al modo en que los elementos estructurales locales de las proteínas con disulfuros son estabilizados y pueden ser predichos por el mecanismo de desplegamiento reductor de estas proteínas. (Chang et al, 2000). Por ello, también es de gran ayuda establecer las vías de desplegamiento proteico, ya que sus características se pueden relacionar directamente con las extraídas de los caminos de plegamiento. Los estudios realizados son de desplegamiento en un entorno reductor, de modo que la proteína al desplegarse va perdiendo puentes disulfuro y “abriéndose”.

Hasta hoy día, se ha observado que existen diferencias considerables en los mecanismos de desplegamiento del BPTI (Oas & Kim, 1988, van Mierlo et al, 1991) EGF (Chang, 1995), TAP (Chang, 1999), PCI (Chang et al, 2000) e hirudina (Bulychev & Chang, 1999). La correlación es la siguiente: 1) la hirudina, PCI, y TAP comparten el mismo mecanismo de todo o nada en su desplegamiento reductor, lo que sugiere que sus puentes disulfuros nativos se hallan estabilizados de modo concertado e interdependiente. Estas tres proteínas muestran un patrón común y característico en sus caminos de plegamiento que incluye la heterogeneidad elevada de sus intermediarios de uno y dos puentes disulfuro y la presencia de isómeros *scrambled* de tres puentes disulfuro como intermediarios del plegamiento. 2) El BPTI se despliega y repliega a través de un mecanismo único; a diferencia del tipo anterior, los tres puentes nativos del BPTI no se hallan estabilizados de manera concertada. Existen estructuras nativas estables que contienen uno y dos puentes disulfuro, tanto en el camino de plegamiento como en el de desplegamiento del BPTI. Estas propiedades gobiernan la formación de un número limitado de intermediarios predominantes que admiten principalmente puentes disulfuro nativos y excluyen la formación de especies *scrambled* de tres puentes disulfuro. 3) Los mecanismos de plegamiento y desplegamiento del EGF parecen descansar entre ambos modelos, el del BPTI y el de la hirudina, exhibiendo características de ambos. De manera parecida a cómo lo hace el BPTI, existe un intermediario estable de desplegamiento que contiene dos disulfuros nativos. La acumulación de este intermediario en el camino de plegamiento del EGF es notoria. Sin embargo, de modo contrario al BPTI, los intermediarios mayoritarios de un puente disulfuro contienen tanto puentes disulfuro nativos como no nativos. Además, en el camino de plegamiento del EGF se encuentran *scrambleds*, rasgo característico de los caminos de plegamiento de la hirudina y PCI.

En síntesis podríamos decir que aquellas proteínas con sus puentes disulfuro reducidos de manera colectiva y simultáneamente exhiben un elevado grado de heterogeneidad en sus intermediarios de plegamiento y la acumulación de isómeros *scrambled* en sus caminos de plegamiento. Una reducción secuencial de los puentes disulfuro nativos es indicativa de la existencia de dominios estructurales estables. En dichos casos, un tipo limitado de intermediarios que adoptan las estructuras parecidas a las nativas tiende a dominar el camino de plegamiento. En la figura 7 podemos observar el desplegamiento reductor del BPTI y el PCI.

Entre toda esta diversidad, la que se observa entre el TAP y el BPTI es la más intrigante. Ambas proteínas son inhibidores tipo Kunitz y comparten homología estructural en términos de patrón de disulfuros y estructura tridimensional (Lim-Wilby et al. 1995) (figura 5). A pesar de su parecido estructural, sus puentes disulfuro nativos parece que se hallen estabilizados de formas muy diferentes. Los intermediarios de plegamiento del TAP son mucho más heterogéneos que los del BPTI, no sólo los de uno y dos disulfuros, sino también los *scrambled*, aunque, de todos modos, entre los intermediarios heterogéneos del TAP se detectan algunas especies predominantes que contienen puentes disulfuro nativos.



**Figura 7. Mecanismos de desplegamiento reductor del PCI y del BPTI.** Ambas proteínas nativas fueron tratadas con diferentes concentraciones de DTT en los tiempos indicados. Los productos finales fueron atrapados en medio ácido previamente al análisis por HPLC. N, II y R indican las posiciones de elución de las especies nativas de tres disulfuros, especies de dos disulfuros y especies reducidas, respectivamente.

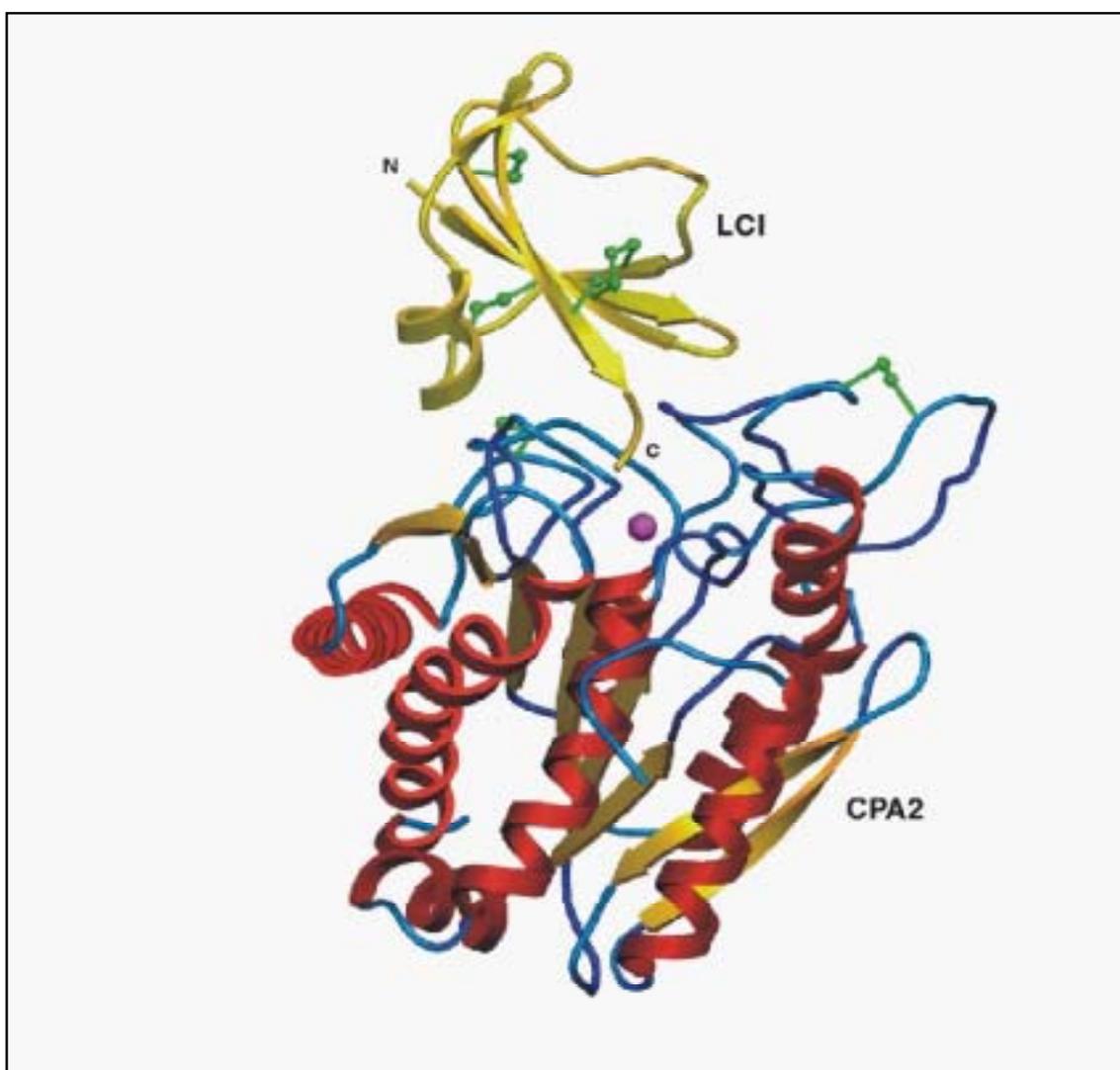
### **I.A.3.d. Estudio del plegamiento y desplegamiento de las proteínas pequeñas de cuatro puentes disulfuro.**

Como acabamos de explicar, las discrepancias que se tienen hasta hoy día sugieren que el camino de plegamiento de las proteínas pequeñas ricas en puentes disulfuro puede ser incluso más versátil que lo que se tenía por norma por el modelo del BPTI. Para entender los modelos del plegamiento proteico, se requiere de estudios posteriores de las vías de plegamiento de un mayor número de proteínas.

El análisis del camino de plegamiento de proteínas que contienen cuatro puentes disulfuro presenta otro nivel de reto técnico debido al incremento exponencial en el número de los isómeros disulfuro. Podemos encontrar 28, 210, 420 y 104 especies isoméricas posibles conteniendo, uno, dos tres y cuatro puentes disulfuro, respectivamente, y que pueden potencialmente servir como intermediarios. Incluso sin tan sólo aparece una fracción de esos isómeros, su separación e identificación conlleva aún un arduo trabajo. Uno de los modelos que recientemente se han investigado al respecto es la ribonucleasa A bovina. Sheraga y colaboradores mostraron que el camino de plegamiento de la ribonucleasa A comprende tanto intermediarios heterogéneos como especies de tres disulfuros que contienen puentes disulfuros nativos (Sheraga et al, 1984, 2001, Rothwarf, 1998). Otra proteína bien caracterizada al respecto es la  $\alpha$ -lactalbúmina, que también muestra un patrón de plegamiento complejo (Rao & Brew, 1989, Ewbanc & Creighton, 1993). Esta proteína presenta un dominio de unión a  $\text{Ca}^{2+}$ , que estructuralmente corresponde a una lámina- $\beta$ . Dependiendo de si el medio de plegamiento contiene o no calcio, la proteína presenta patrones de plegamiento distintos: en ausencia de calcio, la reducción de la  $\alpha$ -lactalbúmina procede a través de un mecanismo de todo o nada, y consecuentemente, se observa que su camino de plegamiento se parece a los del PCI y la hirudina, que vienen marcados por intermediarios de plegamiento heterogéneos, y la acumulación de isómeros de cuatro puentes disulfuro *scrambleds*, a lo largo de su camino de plegamiento. Sin embargo, la unión de calcio a la proteína sirve para generar un dominio de lámina- $\beta$  estable. Entonces, la proteína se despliega reduciéndose de manera secuencial, con una reducción selectiva de dos puentes disulfuro nativos que se hallan dentro del dominio de la  $\alpha$ -hélice. Consecuentemente, el camino de plegamiento de la  $\alpha$ -lactalbúmina en presencia de calcio consiste tan sólo en la aparición de dos intermediarios predominantes que adoptan estructuras parecidas a las nativas y se parece al observado en el caso del BPTI (Chang, 2002).

Todos estos datos, considerados en conjunto, muestran claramente que el mecanismo de plegamiento de las proteínas pequeñas ricas en puentes disulfuro es mucho más complejo de lo que se entendía previamente. Por tanto, se requieren estudios acerca del camino de

plegamiento de nuevas proteínas y, entre ellas, las proteínas pequeñas de cuatro puentes disulfuro plantean un reto interesante. Por ello, se planteó al inicio de este trabajo la realización de estudios de plegamiento y desplegamiento de una proteína de cuatro puentes disulfuro recientemente descubierta y caracterizada por nuestro grupo: el LCI (*leech carboxypeptidase inhibitor*). El LCI presenta un modelo idóneo para estudiar el plegamiento de las proteínas pequeñas ricas en puentes disulfuro, y este afán guió los trabajos primero y segundo de esta tesis, donde se han estudiado sus procesos de plegamiento y desplegamiento. A continuación se extiende en el siguiente apartado la introducción sobre esta molécula a sus enzimas diana, las carboxipeptidasas, por su relación con uno de los trabajos de la presente tesis.



**Figura 8. Estructura tridimensional del complejo LCI-CPA2 humana obtenida por difracción de rayos X.** Los elementos de estructura secundaria en la CPA2 se muestran en colores diferentes (las láminas- $\beta$  en rojo, hélices- $\alpha$  en naranja y el random coil en azul). El átomo de zinc se representa por una esfera lila y todos los puentes disulfuro se muestran en verde.

## I.B. EL LCI Y LAS CARBOXIPEPTIDASAS

### I.B.1. EL LCI

En nuestro grupo de investigación se ha aislado y caracterizado recientemente un nuevo inhibidor de metalocarboxipeptidasa, obtenido de la sanguijuela *Hirudo medicinalis* (Reverter et al, 1998). Este inhibidor, llamado LCI es un polipéptido de 66 aminoácidos rico en cisteínas (posee cuatro puentes disulfuro en su estructura) que se une fuertemente e inhibe a las carboxipeptidasas pancreáticas CPA1, CPA2, CPB y a la carboxipeptidasa de plasma pCPB. Esta proteína no muestra similitud con ninguna otra, excepto en su extremo C-terminal. En esta región, el inhibidor comparte su secuencia aminoacídica Thr-Cys-X-Pro-Tyr-Val-X con los inhibidores de carboxipeptidasa extraídos de *Solanacea*, sugiriendo un mecanismo de inhibición semejante, donde la cola C-terminal del inhibidor interactúa con el centro activo de las metalocarboxipeptidasas de modo similar al sustrato. En concreto, la homología que muestra con el PCI en su extremo C-terminal es:

LCI: (-C-I) -P63-Y-V-E

PCI: (-C-G)-P36-Y-V-G

En ambos casos el residuo C-terminal del inhibidor es hidrolizado después de la unión al enzima.

Recientemente se ha caracterizado la estructura tridimensional del LCI, tanto libre en disolución, como acomplejado con la carboxipeptidasa A2 humana. Se ha observado que el LCI define un motivo estructural que comprende 5 láminas- $\beta$  antiparalelas y una hélice- $\alpha$  corta (Reverter et al, 2000). La figura 8 muestra la estructura tridimensional del LCI unido a la CPA humana.

En nuestro grupo se ha expresado y purificado el inhibidor en un sistema de expresión heteróloga en *E. Coli*, obteniendo la forma plenamente activa del LCI, que denominamos rLCI. El hecho de tener a punto un sistema de producción recombinante que permite obtener elevadas cantidades de proteína hace fácil su disponibilidad para estudios de toda índole (en nuestro caso y como motivo inicial, los estudios de plegamiento).

Con la proteína recombinante así obtenida se ha demostrado que el LCI es un inhibidor competente de los diferentes tipos de carboxipeptidasas pancreáticas, con constantes de equilibrio de disociación del orden de  $(0,2-0,4) \cdot 10^{-9}$  M para los complejos con las carboxipeptidasas pancreáticas A1, A2 y B y la carboxipeptidasa B de plasma (Tabla I). Las

constantes de equilibrio de disociación para los complejos de LCI con diferentes metalocarboxipeptidasas fueron determinadas a pH 7.5. De aquí se puede concluir que el rLCI es un inhibidor potente y competitivo de las carboxipeptidasas pancreáticas y que posee una especificidad diferente para cada una de las distintas formas de carboxipeptidasa de mamífero, tal y como muestra la tabla I. Se debe destacar que presenta su máxima afinidad para la carboxipeptidasa B plasmática o TAFI, siendo el inhibidor más potente que se conoce hasta hoy día para este enzima. De la pCPB o TAFI (*thrombin activatable fibrinolysis inhibitor*) se hablará extensamente en el próximo apartado, aunque debemos señalar que recientemente se ha demostrado que retrasa la fibrinólisis. Asumiendo que el LCI es secretado por la sanguijuela durante su alimentación, la molécula probablemente participe en la eliminación de coágulos sanguíneos inhibiendo la CPB de plasma. El LCI podría entonces ayudar a mantener el estado líquido de sangre mientras la sanguijuela se alimenta y una vez la sangre se halla ya en el interior del parásito.

Antes de continuar con el hilo conductor de este trabajo, se introduce aquí brevemente las características de las carboxipeptidasas y más extensamente de TAFI.

**Tabla I.  $K_i$  del LCI para distintas carboxipeptidasas.**

<b>Tipo de carboxipeptidasa</b>	<b>Constante de inhibición (<math>K_i</math>) nM</b>
CPA1 bovina	0.25-0.48
CPA1 ( des-Glu <sup>66</sup> -LCI ) bovina	0.1-0.3
CPB porcina	0.27-0.51
CPA2 humana	0.17-0.78
TAFI (pCPB) humana	0.1-0.2

## **I.B.2. LAS CARBOXIPEPTIDASAS**

Las carboxipeptidasas son enzimas que hidrolizan los enlaces C-terminales de las proteínas. La familia de las carboxipeptidasas se divide en serín-carboxipeptidasa, cisteín-carboxipeptidasas y metalocarboxipeptidasas. Las serín y cisteín carboxipeptidasas poseen un residuo serina o cisteína, respectivamente en su centro activo, mientras que las metalocarboxipeptidasas conllevan un átomo de zinc que es importante en su acción catalítica. Las metalocarboxipeptidasas pancreáticas que juegan un papel en el aparato digestivo son los

miembros mejor caracterizados de este grupo. De acuerdo con su especificidad de sustrato, se las conoce como carboxipeptidasas del tipo A (CPA) que hidrolizan preferentemente residuos alifáticos, o carboxipeptidasas del tipo B (CPB), que hidrolizan preferentemente residuos básicos. Recientemente, se ha descrito un determinado número de carboxipeptidasas no digestivas, como la carboxipeptidasa A de mastocitos, la carboxipeptidasa E (CPE, también conocida como CPH) y la carboxipeptidasa M (CPM) unida a membrana.

Puesto que se conoce mucho sobre la estructura y los mecanismos catalíticos de la carboxipeptidasa B pancreática, y debido a su semejanza con la TAFI, a continuación se incorpora una breve descripción de aquella: CPB es un enzima digestivo cuyos niveles normales en suero sanguíneo son prácticamente indetectables. Sólo durante una pancreatitis pueden detectarse el enzima activo en el suero. Su vida media en sangre es del orden de minutos tal vez porque no está glicosilada (Erdős et al, 1963). La CPB pancreática tiene un peso molecular de 35kDa, se sintetiza en el páncreas en forma de zimógeno y posee dos dominios distinguibles: el dominio catalítico y el segmento de activación. El segmento de activación cubre el sitio activo, impidiendo la entrada de sustratos. La activación de la CPB por tripsina sucede mediante una rotura en la Arg-95, resultando en la liberación del segmento de activación y la exposición del centro activo de la enzima. La Arg C-terminal es liberada por la CPB activa y el segmento de activación es posteriormente hidrolizado en la Arg-81. La CPB activa es, en contraste con TAFIa, una proteasa estable (Burgos et al, 1991).

En el plasma, la familia de metalocarboxipeptidasas está representada por la carboxipeptidasa N (CPN, conocida también como lisin-carboxipeptidasa, kininasa I, o inactivador de anafilotoxina) y la TAFI. La CPN es constitutivamente activa, mientras que la actividad carboxipeptidasa de TAFI se genera durante la coagulación sanguínea. La CPN fue descubierta por Erdős y Sloane en 1962 como un enzima que inactivaba la bradikina mediante la hidrólisis de su Arg C-terminal (Erdős & Sloane, 1962). Se ha demostrado que la CPN puede hidrolizar un gran número de péptidos endógenos y de sustratos proteicos incluyendo kininas, anafilotoxinas (C3a, C4a y C5a) y los fibrinopéptidos A y B. La CPN se sintetiza en el hígado y se secreta a la circulación en su forma activa. Su concentración en plasma es de 30µg/ml. La CPN es una glicoproteína con un peso molecular de 280 kDa. Es un tetrámero con dos subunidades catalíticas de 83 kDa que están fuertemente glicosiladas. Las subunidades se mantienen juntas gracias a interacciones no covalentes (Skidgel, 1995). La CPN contiene un ión de zinc en su centro activo, y su pérdida comporta la total pérdida de actividad enzimática de la molécula. En contraste con TAFI, la CPN no modula la fibrinólisis y no resulta inhibida por el inhibidor de carboxipeptidasa de patata PCI.

### **I.B.2.a. Descubrimiento de TAFI y nomenclatura.**

La presencia de una actividad carboxipeptidasa lábil en el suero, que interfería con las medidas de la actividad de CPN fue descrita por primera vez en 1988 (Hendriks et al, 1989a y 1989b). Esta actividad carboxipeptidasa nueva no se hallaba presente en el plasma pero aparecía después de la coagulación. El enzima difería de la CPN en términos de estabilidad, pH óptimo y especificidad de sustrato. Además, la separación de suero humano fresco mediante cromatografía de intercambio iónico mostraba dos picos con actividad carboxipeptidasa. La actividad carboxipeptidasa del primer pico no se mantenía estable a 37°C, mientras que la actividad carboxipeptidasa del segundo era estable a 37°C. Debido a esta inestabilidad, el nuevo enzima fue llamado carboxipeptidasa U (CPU, *unstable*-inestable-*carboxypeptidase*) (Hendriks Et al, 1989b, Hendriks et al 1990). Campbell y Okada también describieron un incremento en la actividad arginin-carboxipeptidasa inducida durante la coagulación (Campbell & Okada, 1989). Estos autores encontraron el mismo incremento de actividad carboxipeptidasa en plasma carente de CPN y por tanto concluyeron, que esta actividad arginin- carboxipeptidasa no se hallaba relacionada con la CPN. El enzima responsable fue llamado CPR (R por Arg). Eaton y colaboradores descubrieron este mismo enzima como un contaminante durante la purificación de la  $\alpha$ 2-antiplasmina, utilizando una columna de afinidad plasminógeno-sefarosa (Eaton et al, 1991). El cDNA del enzima fue clonado a partir de hígado humano y se dedujo la secuencia aminoacídica. Eaton denominó al enzima CPB plasmática, por semejanza en especificidad con la CPB pancreática. Aunque inicialmente se describió que un peso molecular para la CPU muy elevado (Hendriks et al, 1990) ajustes posteriores lo cifraron en 53-60 kDa. Wang y colaboradores confirmaron entonces que la CPU y la CPB plasmática eran proteínas idénticas (Wang et al 1994). En la búsqueda de un mecanismo detrás del efecto profibrinolítico de la proteína C activa, Bajzar y colaboradores purificaron el mismo enzima y demostraron que podía ser activado por trombina y que bajo dicha activación, inhibía la fibrinólisis (Bajzar et al, 1995). Este grupo introdujo el término de TAFI para denominar al zimógeno del cual se muestra el resultado del modelado tridimensional llevado a cabo recientemente por nuestro grupo (figura 9).

Hasta el momento, los diversos nombres barajados, CPU, CPR, CPB de plasma y TAFI se están usando indistintamente. Es patente la necesidad de un solo nombre aceptado generalmente. El nombre que se utiliza con más frecuencia en la literatura consultada, es el de TAFI, y éste será el nombre usado en este trabajo por su relación con la función fisiológica y

sus patologías relacionadas, que serán, como se verá más adelante, el objeto de nuestro estudio y conformarán el tercer capítulo de la tesis. Así pues, una vez realizados los estudios sobre plegamiento del LCI, se planteó la investigación sobre el posible efecto de este inhibidor en la fibrinólisis: dado que su enzima diana en el plasma, TAFI, la inhibe, un suministro de LCI debería favorecerla. Siguiendo esta línea de pensamiento, en el tercer trabajo que conforma esta tesis se ha estudiado el efecto *in vitro* del LCI en la fibrinólisis.

Dada la complejidad del sistema hemostático parece apropiado centrar la tercera parte de esta introducción general en los procesos de hemostasia y coagulación, que se presentan a continuación.



**Figura 9. Representación de la estructura de TAFI dilucidada por modelado molecular.** Los dominios pro- y enzimático se muestran en azul y dorado, respectivamente, y las diferentes

intensidades del color muestran las estructuras secundarias de hélices- $\alpha$  o láminas- $\beta$ . Las posiciones y las cadenas laterales de los cuatro puntos de glicosilación del segmento pro se muestran en rojo.

## I. C. HEMOSTASIA Y COAGULACIÓN

La hemostasia es el proceso mediante el cual se mantiene la integridad del árbol vascular, evitando y corrigiendo las disrupciones u obstrucciones en alguno de sus segmentos. Al producirse una lesión en los vasos sanguíneos la coagulación es el proceso por el cual se evita la pérdida masiva de sangre. La coagulación y la fibrinólisis son los procesos que forman y disuelven la fibrina, respectivamente. Estos procesos están finamente regulados y protegen al organismo del sangrado excesivo o del depósito excesivo de fibrina. La regulación de ambas cascadas se consigue mediante una gran variedad de mecanismos que incluyen respuestas celulares, flujo e interacciones proteína-proteína. Dentro de la coagulación sanguínea, la hemostasia primaria o provisional es proporcionada por las plaquetas, células que desempeñan un papel decisivo en la detención de las hemorragias gracias a la formación del trombo plaquetario. Las plaquetas también intervienen en otras fases de la hemostasia, como la vasoconstricción y la coagulación plasmática, de tal modo que el agregado de plaquetas constituye el sustrato sobre el que se forma la fibrina. No se profundiza aquí en la función plaquetaria debido a que no es objeto de estudio de esta tesis, centrándose en cambio en la interacción del LCI con otras proteínas solubles en el plasma.

**Tabla II. Factores y reguladores de la coagulación.**

PROTEÍNA	Concentración plasmática ( nM )	Vida media ( días )
Protrombina ( Factor II )	1400	2.5
Factor V	20	0.5
Factor VII	10	0.25
Factor VIII	0.7	0.3-0.5
Factor IX	90	1
Factor X	170	1.25
Factor XI	30	2.5-3.3
Factor XII	375	2-3
Quininógeno de alto peso molecular	6000	5
Precalicerina	450	-
Factor XIIIa	70	9-10
Factor tisular	-	-
Proteína C	60	0.25
Proteína S	300	1.75
Trombomodulina	-	-
Fibrinógeno	8800	3-5
Antitrombina	2400	2.5-4
Cofactor II de la heparina	1200	2.5
TFPI ( inhibidor de la vía del factor tisular )	2.5	-
TAFI ( inhibidor de la fibrinólisis activado por trombina )	73	-

La coagulación plasmática es el conjunto de reacciones que conducen a la transformación del fibrinógeno, una proteína soluble que se halla en circulación, en fibrina, proteína insoluble que evita la pérdida de sangre cuando se lesiona un vaso. Dicho proceso tiene lugar mediante una reacción en cascada de los factores de coagulación en la que tiene lugar la activación de proenzimas y que queda limitado a la lesión vascular gracias a los inhibidores naturales de los factores activados.

Los factores de coagulación (con la excepción del calcio) son proteínas que se sintetizan en los hepatocitos y se hallan en la sangre como zimógenos inactivos. Algunos de ellos como la protrombina, factores VII, IX y X pueden ser activados y transformados en enzimas con actividad serín-proteasa mediante proteólisis limitada. Estas reacciones, que en fase fluida son muy poco eficientes, tienen lugar *in vivo* sobre superficies fosfolípídicas (membranas de células endoteliales o plaquetas). El resultado es una respuesta masiva, local (muchos factores se activan en base a un contacto entre proteínas que se hallan en el endotelio de los vasos) y fuertemente amplificada de todas las proteínas que participan en el proceso.

Los inhibidores naturales de la coagulación sanguínea se dividen en dos tipos: los inhibidores de las serín-proteasas y los inhibidores de los factores V y VIII activados. Entre los inhibidores de las serín-proteasas podemos encontrar la antitrombina III, que es el principal inhibidor de la trombina, con la que forma un complejo irreversible, y el TFPI (inhibidor de la vía del factor tisular) que inhibe la formación del FXa (factor X activo), al inhibir el complejo TF-FVIIa. También se debe destacar el papel de la  $\alpha$ 2-macroglobulina, que contribuye al 25% del total de actividad antitrombina del plasma, y que además de su efecto sobre los factores procoagulantes, es una antiplasmina. Los inhibidores de los factores V y VII activados forman parte de la vía de la proteína C, y son fundamentalmente la proteína C, la proteína S y la trombomodulina. En la tabla II pueden observarse los principales factores y reguladores de la coagulación.

La coagulación se halla asimismo controlada por una vía anticoagulante compuesta en síntesis por la trombina, la trombomodulina y la proteína C. La proteína C se convierte mediante el complejo TR-TM en proteína C activa (APC), la cual cataliza la inactivación mediante proteólisis de los cofactores esenciales que se requieren para la formación de trombina, los factores Va y VIIIa. Análogamente, se ha descubierto recientemente una vía antifibrinolítica. Esta vía proveería de una aparente simetría a los procesos de coagulación y fibrinólisis y también se compone, curiosamente, de trombina, trombomodulina y un zimógeno (TAFI) que se convierte a enzima activo. Este enzima activo, proteolíticamente inactivaría a un cofactor para atenuar la fibrinólisis. De todos modos, mientras que la APC es

una serín-proteasa, este enzima antifibrinolítico es una metaloproteasa que exhibe actividad parecida a la carboxipeptidasa B. Como ya hemos mencionado, en un período de unos pocos años y de forma totalmente independiente y separada, cinco grupos de investigación describieron una molécula que exhibía actividad antifibrinolítica. Como ya hemos comentado, nos referiremos a esta molécula como TAFI (*thrombin activatable fibrinolysis inhibitor*) un nombre basado en sus propiedades funcionales.

Las reacciones fibrinolíticas son las encargadas de eliminar la fibrina cuando ya no es necesaria para evitar la pérdida de sangre, reestableciendo el flujo sanguíneo. Aunque no forman parte de la cascada misma de coagulación, se hallan extensamente relacionadas con ella (Dobrovolsky & Titaeva, 2002).

### I.C.1. REACCIONES COAGULANTES: LA COAGULACIÓN

En el proceso de la coagulación sanguínea se distinguen dos fases:

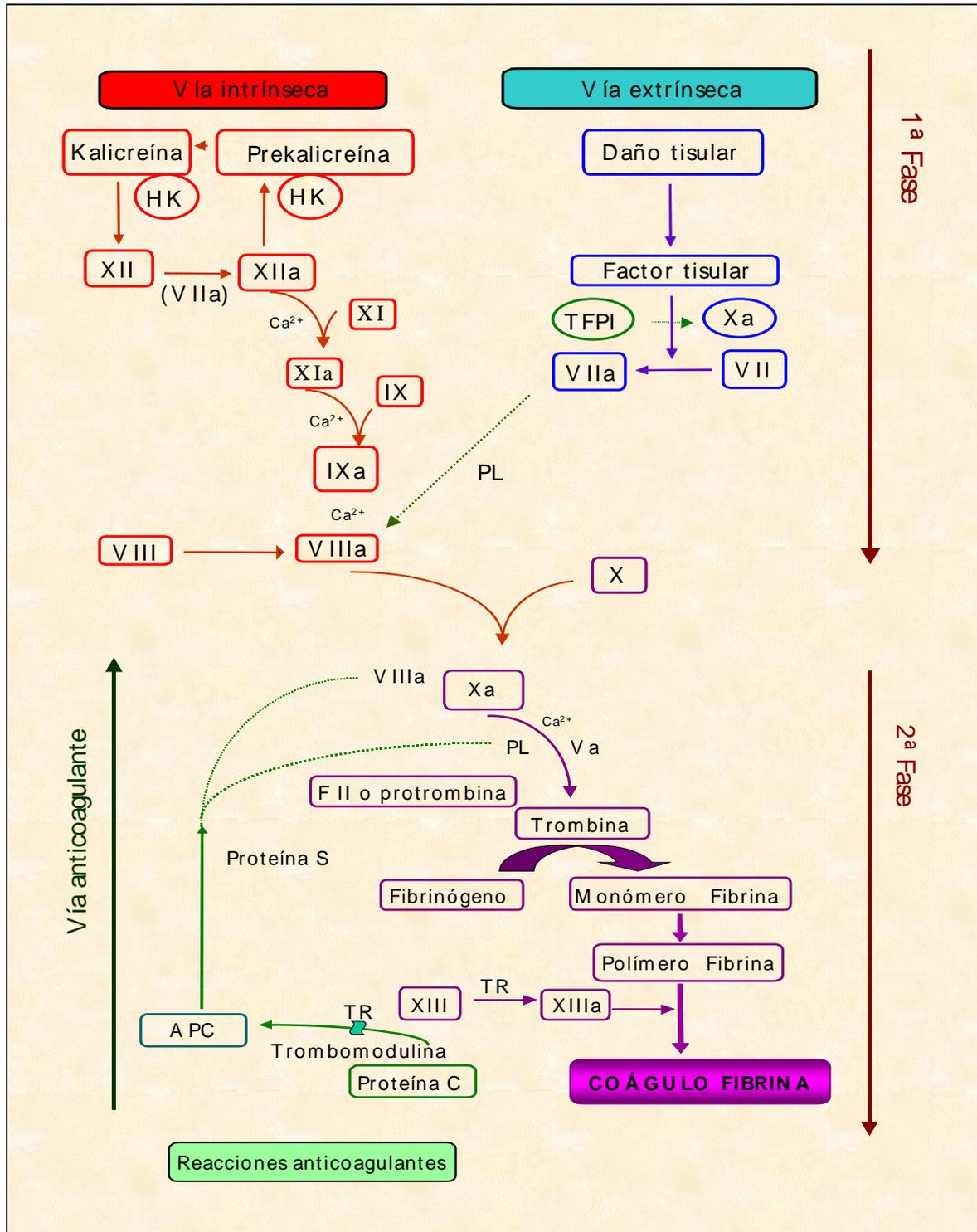
1ª Fase: cascada de reacciones que lleva a la conversión de la protrombina en trombina

2ª Fase: transformación del fibrinógeno en fibrina a merced de la trombina formada

**1ª Fase:** puede iniciarse por dos vías:

1-La vía extrínseca o vía del factor tisular (TF). Se desencadena cuando la sangre contacta con el factor hístico o factor tisular, que es un complejo lipoproteico procedente de las células endoteliales y los leucocitos. Por tanto, suele darse debido a una lesión en el vaso o a un proceso inflamatorio que libera el TF al torrente circulatorio. En este sistema interviene el factor VII, una serín-proteasa que se encuentra en mínimas cantidades en el plasma y se activa al unirse al TF, iniciando así la cascada.

2-La vía intrínseca. Se inicia con la activación del factor XII producida por el contacto con una superficie desprovista de endotelio. El factor XII, que es una serín-proteasa, en contacto con una superficie cargada negativamente sufre un cambio de conformación que determina su conversión a factor XII activado (XIIa) por acción proteolítica de la kaliceína. Ambas vías convergen en un punto central: la conversión de protrombina en trombina, una de las proteasas de mayor importancia en la cascada de coagulación, finalizando así la primera fase de la cascada de coagulación. La segunda fase es la que llevará a cabo la transformación del fibrinógeno en fibrina merced a la trombina formada.



**Figura 10. Esquema simplificado de las reacciones de la cascada de coagulación iniciada por las vías extrínseca e intrínseca.** Nótese la convergencia de las dos vías al final de la primera fase de la coagulación. En verde se hallan indicadas las reacciones anticoagulantes, mientras que la vía fibrinolítica no aparece, y se incluirá en una figura aparte. Las líneas continuas indican activación, mientras que las líneas discontinuas indican inhibición o degradación. Nótese también el papel dual de la trombina, y sus efectos antagónicos al actuar libre y/o acomplexada con la trombomodulina. Las abreviaturas utilizadas en la figura son las siguientes: V, VII, VIII, IX, X, XI, XI factores V, VII, VIII, IX, X, XI, XI, respectivamente; Va, VIIa, etc. factor V, VII, etc activos; HK, quinínogeno de elevado peso molecular; PL, fosfolípidos; TR, trombina; TF, factor tisular ; TFPI, inhibidor de la vía del factor tisular; TM, trombomodulina.

El proceso final de la primera fase de coagulación, llevado a cabo por el factor X activado (Xa), requiere fosfolípidos, calcio y factor V. El factor V, que es activado por la trombina, hace que la protrombina sea más accesible al efecto proteolítico del Xa.

En la figura 10 se puede observar el proceso de las reacciones de la cascada de coagulación, que se resumen en este esquema para no extender este capítulo excesivamente.

**2ª Fase:** transformación del fibrinógeno en fibrina y formación de la malla de fibrina.

El fibrinógeno es una glicoproteína de 340 kDa sintetizada en el hígado. La molécula está formada por dos mitades simétricas, cada una formada por tres tipos de cadena polipeptídicas ( $\alpha$ ,  $\beta$  y  $\gamma$ ). La transformación del fibrinógeno en fibrina se debe a la acción de la trombina, que libera de la molécula de fibrinógeno cuatro péptidos de bajo peso molecular llamados fibrinopéptidos A y B de las cadenas  $\alpha$  y  $\beta$ , respectivamente. En primer lugar, la trombina rompe el enlace Arg16-Gly17 de la cadena  $\alpha$  liberando el fibrinopéptido A. En segundo lugar la trombina hidroliza el enlace Arg14-Gly15 de la cadena  $\beta$  liberando el fibrinopéptido B y formando así el monómero de fibrina.

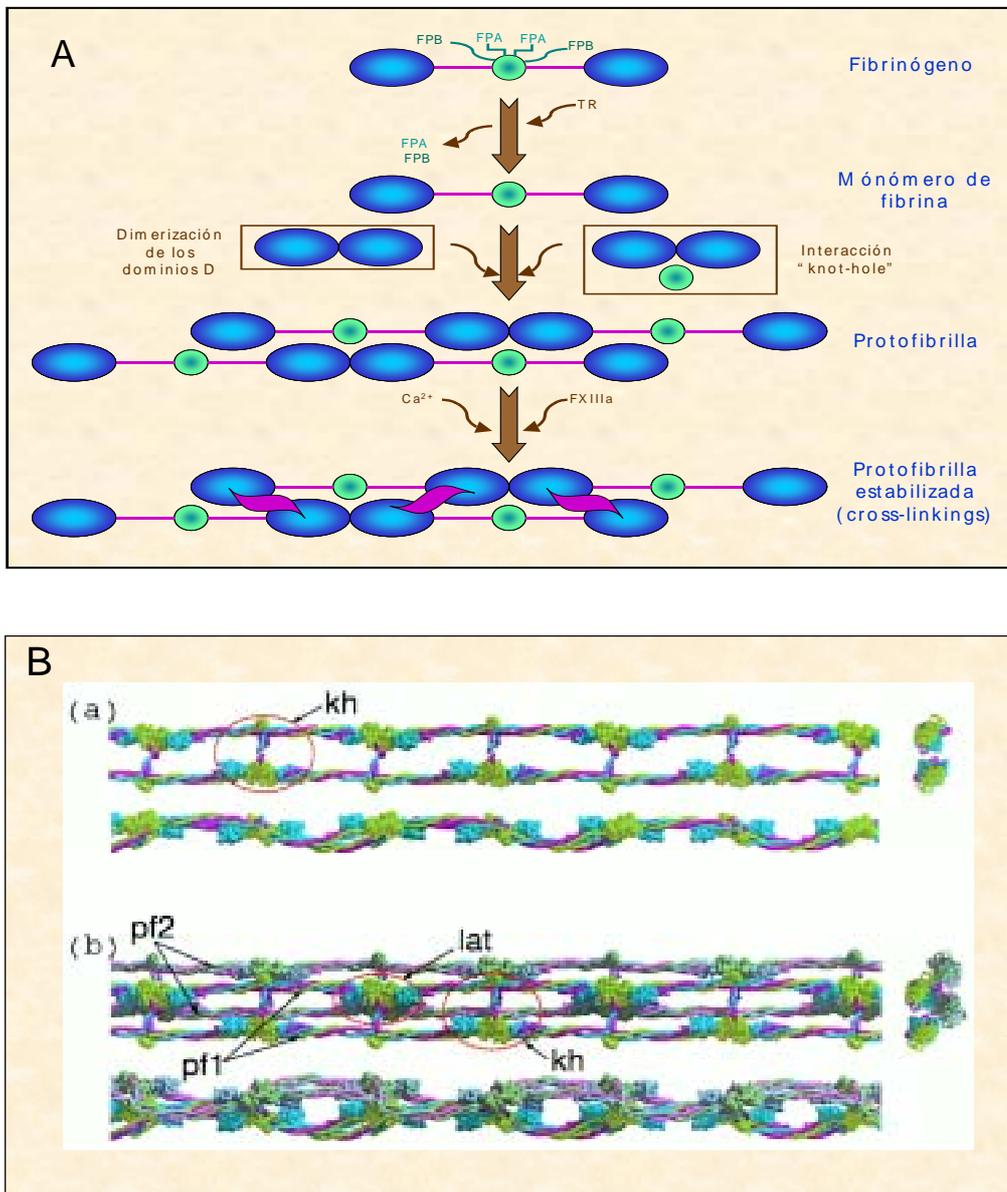
A continuación se produce una asociación lateral de las fibras aumentando su grosor. Finalmente, en presencia de Factor XIII activo, la fibrina que inicialmente estaba unida por interacciones no covalentes, sufre una serie de uniones intermoleculares mediante la formación de enlaces entre residuos lisina de una cadena  $\gamma$  y residuos glutamina de otra cadena  $\gamma$ . La incorporación de estos enlaces covalentes entre las fibras de fibrina aporta una mayor resistencia a la degradación por la plasmina, así como una mayor resistencia y elasticidad al coágulo (Sidelmann et al, 2000).

Dado que en el presente trabajo se hará hincapié en la estructura de la fibrina, a causa de estudios realizados sobre la posible influencia del LCI en dicha estructura, se incorpora en la figura 11 un esquema de la formación de dicha estructura.

### **I.C.1.a. Reacciones anticoagulantes.**

Las reacciones anticoagulantes (que pueden ser observadas en la figura 10) regulan la coagulación impidiendo su propagación indefinida. Estas reacciones son conducidas por dos proteínas: la trombomodulina y la proteína C. Ambas modulan la síntesis y la acción de la trombina que, como acabamos de ver, es la enzima que se genera al final de la cascada de activación de las propeptasas de la coagulación y que es capaz de transformar una proteína soluble, el fibrinógeno, en una red polimérica insoluble, la fibrina. La trombina tiene una gran afinidad por la trombomodulina, que se halla en el endotelio y es una proteína integral de membrana. Consta de una zona N-terminal semejante a la de la lectina (*lectin-like domain*)

seguida de 6 dominios semejantes al factor de crecimiento epidérmico (EGF), un dominio rico en aminoácidos serina y treonina, un dominio transmembrana y una cola citoplasmática.



**Figura 11. A) Esquema simplificado de la conversión del fibrinógeno a fibrina por trombina.**

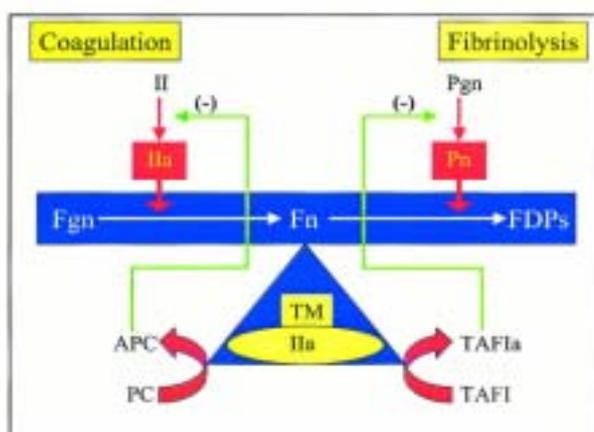
La activación de la coagulación conduce a la formación de trombina que hidroliza los fibrinopéptidos A y B (FPA y FPB) de la molécula de fibrinógeno, transformándolo en un monómero de fibrina que polimeriza en fibrina soluble. Simultáneamente, la trombina activa al factor de coagulación FXIII a FXIIIa, el cual estabiliza la fibrina soluble mediante la génesis de uniones entrecruzadas covalentemente que estabilizan el polímero.

**B) Estructura molecular de la fibrina y la formación de las fibrillas.**

Las cadenas  $\alpha$  se muestran en rosa, las cadenas  $\beta$  se muestran en azul y las cadenas  $\gamma$  se muestran en verde. Kh indica la interacción "knott-hob" (nudo-hueco) entre el dímero de dos cadenas  $\gamma$  y el dominio central de otra, que es el punto donde la fibrina se agrega lateralmente, formando así la protofibrilla. a) Protofibrilla estabilizada. b) Agregación lateral (lat) de dos protofibrillas (pf1 y pf2) por interacciones no covalentes: formación de las fibras que construirán el coagulo. Obsérvese que una vez formadas las interacciones en las protofibrillas las estructura se espiraliza, dándole rigidez y elasticidad.

En la vasculatura se localiza principalmente en los capilares y en menor medida en arterias y venas y sus niveles de expresión en la superficie endotelial son susceptibles a diversas condiciones patológicas. Cuando la trombina se une a la trombomodulina, sufre un cambio conformacional que permite el reconocimiento y activación de la proteína C, formando un complejo que pierde sus propiedades procoagulantes y adquiere otras anticoagulantes. La trombina se une con gran afinidad a la trombomodulina y, una vez formado el complejo, la trombina no transforma el fibrinógeno en fibrina, no activa al factor V, el VIII ni el XIII y tampoco activa a las plaquetas ni a las células endoteliales. Además, la trombomodulina ejerce una actividad semejante a la de la heparina acelerando la neutralización de la trombina por la antitrombina III. De hecho, se han dilucidado recientemente las bases estructurales para la actividad anticoagulante del complejo trombina-trombomodulina. Se ha demostrado que la trombomodulina se une al *anion-binding site* de la trombina, el lugar a través del que la trombina se une al fibrinógeno. De este modo, la trombomodulina inhibe la unión del fibrinógeno a la trombina, mientras se une a la proteína C de tal modo que coloca correctamente el péptido de activación de la proteína C como un sustrato para el sitio activo de la trombina. La trombomodulina utiliza uno de sus dominios EGF para bloquear estéricamente el acceso de sustratos de la trombina y, a su vez, utiliza otro dominio EGF para presentar un nuevo sustrato a la trombina como la proteína C, y, tal vez, la TAFI.

Un efecto recientemente identificado del complejo trombina-trombomodulina es su capacidad para inhibir la fibrinólisis mediante la activación de la TAFI, y sobre el que, de nuevo, se hablará extensamente más adelante. La figura 12 muestra la simetría de la vía anticoagulante y la posible vía antifibrinolítica.



**Figura 12. Modelo que ilustra la conexión entre las vías anticoagulante y antifibrinolítica.** En la figura la trombina viene indicada como FIIa (factor II activo). El complejo TR-TM (TM-IIa) cataliza la formación de proteína C activa (APC) y de TAFIa, lo cual inhibe la generación de IIa y plasmina, respectivamente. IIa cataliza la ruptura del fibrinógeno para la formación de fibrina, mientras que la plasmina degrada la fibrina, formándose los productos de degradación de la fibrina.

La proteína C, uno de los puntos clave del ataque del complejo trombina-trombomodulina, es una proteína vitamina K-dependiente sintetizada en el hígado que circula en plasma a una concentración de 50-80 nM. Una vez activada, la proteína C se une a los fosfolípidos de membrana o al receptor endotelial de la proteína C, y juntamente con la proteína S que actúa como cofactor, degrada los factores Va y VIIIa introduciendo por tanto un freno al proceso de la coagulación y conformando así la vía anticoagulante.

## **I.C. 2. REACCIONES FIBRINOLÍTICAS: LA FIBRINÓLISIS**

El proceso de la fibrinólisis destruye la fibrina depositada en el árbol vascular durante la coagulación, es decir, es un sistema reactivo a la activación de la coagulación y a la generación final de trombina. El enzima central de este sistema es la plasmina, una serín-proteasa de 85 kDa. En condiciones normales circula en el plasma humano en la forma de su proenzima, el plasminógeno. La transformación del plasminógeno en plasmina se lleva a cabo por los denominados activadores del plasminógeno: el activador tisular del plasminógeno (t-PA) y el activador tipo urokinasa (u-PA). La plasmina es un enzima extraordinariamente inespecífico, capaz de degradar la fibrina, el fibrinógeno, los factores de coagulación V y VIII y otros muchos sustratos. Para regular esta actividad proteolítica de la plasmina el plasma humano posee un inhibidor muy selectivo y potente, llamado  $\alpha$ 2-antiplasmina. Así mismo para regular los activadores del plasminógeno el organismo humano está dotado de los llamados inhibidores de la activación del plasminógeno tipo I, II y III.

En la figura 13 se puede observar el proceso de la fibrinólisis representado esquemáticamente, para, de nuevo, no extender demasiado este capítulo.

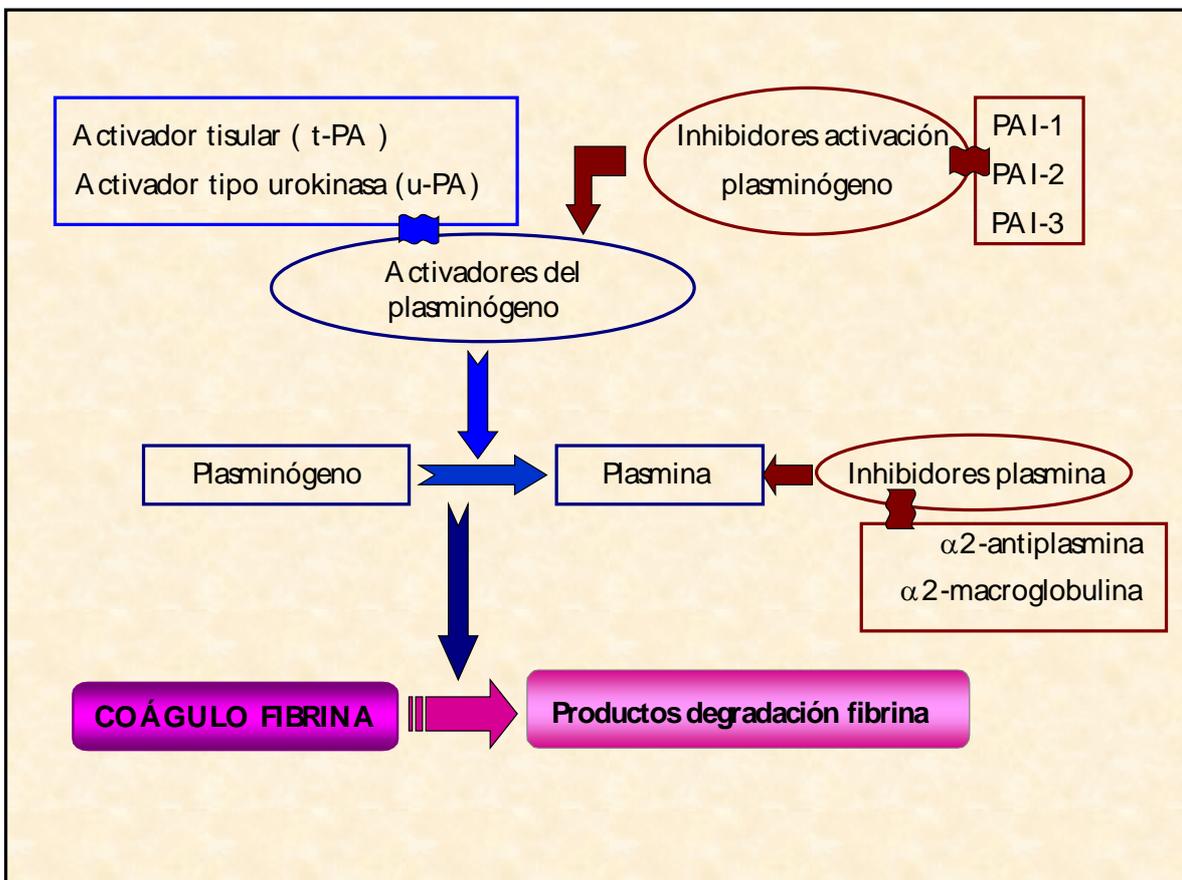
Se incluye aquí una breve descripción del plasminógeno y de su forma activa, la plasmina, dada la importancia de su acción en la fibrinólisis y su relación con la acción de la TAFI.

El plasminógeno humano es una glicoproteína de una sola cadena con un peso molecular cercano a los 90 kDa. La molécula de plasminógeno tiene 5 dominios homólogos llamados *kringles* que le confieren una gran afinidad por la fibrina, la  $\alpha$ 2-antiplasmina y la trombospondina. El plasminógeno nativo puede presentarse en dos formas: Glu-plasminógeno, forma que posee un ácido glutámico en su extremo amino-terminal o Lys-plasminógeno forma que posee una lisina en su extremo amino-terminal. El Glu-plasminógeno se convierte fácilmente en las formas que contienen en su extremo aminoterminal lisina, valina, metionina, etc., mediante proteólisis limitada.

Tanto el Glu-plasminógeno como el Lys-plasminógeno son proenzimas inactivos. La activación por proteólisis limitada del Lys-plasminógeno a plasmina se realiza por rotura de un

simple enlace peptídico entre los aminoácidos Arg560-Val561, llevado a cabo por t-PA o u-PA.

En condiciones normales tan solo un 20% del plasminógeno circula libre en plasma. El 50% del plasminógeno forma un complejo disociable con la HRG (glicoproteína rica en histidina) y alrededor de un 15% forma complejos disociables con la  $\alpha$ 2-antiplasmina. El resto, aproximadamente un 15% circula unido al fibrinógeno. Este equilibrio se mantiene a concentraciones plasmáticas del plasminógeno y del resto de componentes citados dentro del rango de la normalidad. Así pues, al formarse la fibrina dentro del árbol vascular, esta contiene ya plasminógeno, el cual se halla unido a la fibrina por una o más regiones llamadas *lysine-binding-sites* (lugares de unión a lisina, LBS), poseedoras de una elevada afinidad por los residuos de lisina.

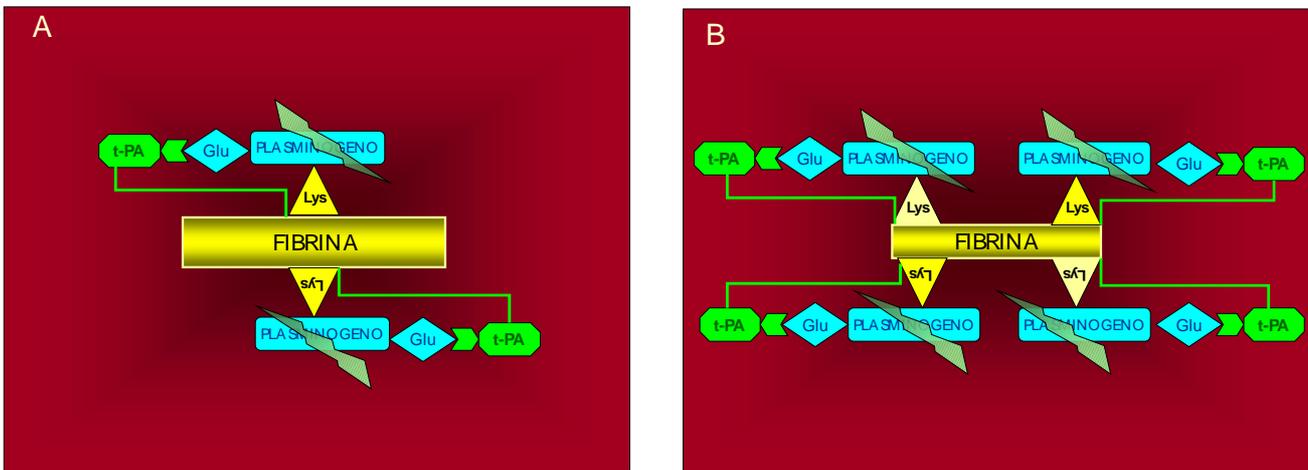


**Figura 13. Esquema de las reacciones de la fibrinólisis.** El inicio de la cascada corre a cargo de los activadores del plasminógeno, siendo el más común el t-PA. El plasminógeno se transforma en plasmina en la superficie de la fibrina, comenzando así la acción proteolítica sobre el coágulo.

Por estímulos aún mal definidos, las células endoteliales circundantes al coágulo o trombo liberan activadores del plasminógeno tipo t-PA o u-PA. Estos activadores, sobretudo el t-PA, tiene una gran afinidad por la fibrina (unas 100 veces superior a la que tienen por el fibrinógeno) y prácticamente no interactúan con el plasminógeno circulante sino tan solo con el que se halla unido a la fibrina mediante los LBS. En estas condiciones estos activadores transforman el plasminógeno en plasmina en la superficie del trombo. Aparte de esto, la plasmina unida a la fibrina no es inhibida por su inhibidor natural  $\alpha_2$ -antiplasmina, por poseer sus LBS ocupados por la fibrina.

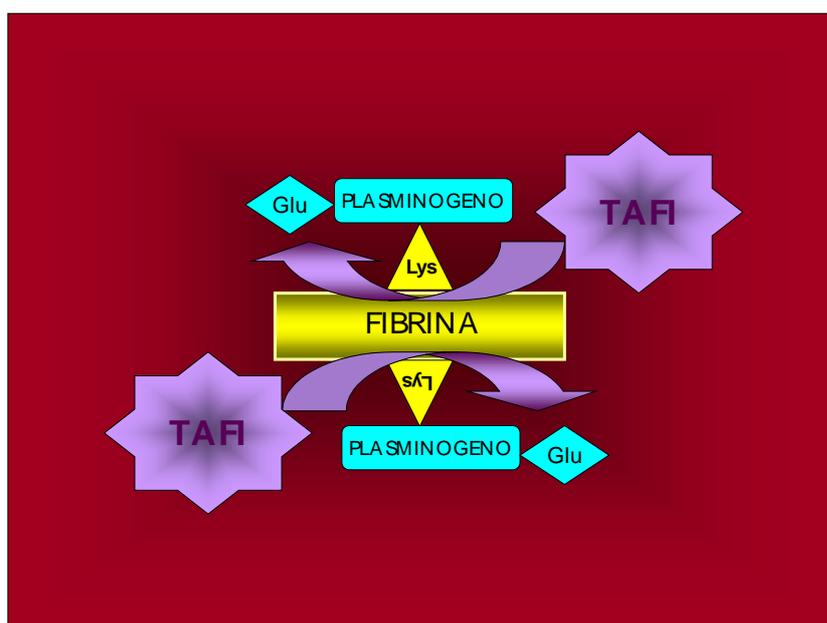
La proteólisis inicial de la fibrina genera tan sólo lisinas carboxiterminales que sirven de nuevos lugares de unión para el plasminógeno y el t-PA, acelerando así la fibrinólisis.

La molécula de plasmina está formada por dos cadenas: la cadena A o cadena pesada y la cadena B o ligera, que contiene la región C-terminal del polipéptido original y el centro activo de la molécula. La plasmina formada puede degradar de manera ordenada y secuencial tanto a la fibrina como al fibrinógeno, produciendo fragmentos conocidos como productos de degradación de la fibrina o el fibrinógeno (Gaffney, 2001).



**Figura 14. Complejo ternario formado en la superficie de la fibrina para iniciar la cascada fibrinolítica.** A) El plasminógeno se une a los residuos de lisina de la fibrina por sus dominios LBS (lysine binding sites). El t-PA se une a un residuo glutámico del plasminógeno y a la fibrina por los residuos lisina, activándose e hidrolizando el plasminógeno en plasmina. Las líneas truncadas indican la hidrólisis del plasminógeno por el t-PA. B) Nuevos residuos de lisina C-terminales aparecen en la fibrina parcialmente degradada, atrayendo a nuevos complejos de plasminógeno y t-PA. El resultado final es la concentración masiva y local de todos los elementos de la cascada fibrinolítica.

En determinadas situaciones, cuando se genera fibrina dentro del árbol vascular, la concentración de t-PA es extremadamente elevada en el punto donde se ha producido la trombosis y es en esta zona donde se produce la activación del plasminógeno en plasmina. En ausencia de fibrina, la eficiencia del t-PA para activar el Glu-plasminógeno circulante es muy baja. En cambio, en presencia de fibrina el t-PA tiene una afinidad 500 veces superior por el plasminógeno fijado a la fibrina. También se sabe que la forma parcialmente degradada del plasminógeno (Lys-plasminógeno) se transforma en plasmina por el t-PA 10 veces más rápido que la forma nativa (Glu-plasminógeno).



**Figura 15. Actuación de TAFI sobre los residuos lisina C-terminales de la superficie de la fibrina.** Al hidrolizar las lisinas de la superficie del coágulo TAFI impide que éste sea reconocido por el plasminógeno, y en consecuencia por el t-PA, con lo que se evita la concentración de los factores fibrinolíticos y el coágulo permanece estable por más tiempo.

### I.C.2.a. Función de TAFI en la fibrinólisis.

El sistema fibrinolítico se inicia después de la formación de fibrina cuando el plasminógeno y el t-PA se unen a la superficie de la fibrina. La unión se halla mediada por interacciones específicas de los residuos de lisina C-terminales de la fibrina parcialmente degradada y de los LBS en el plasminógeno y el t-PA. La unión del plasminógeno y el t-PA a la superficie de la fibrina da como resultado la formación de un complejo ternario y un incremento de la eficiencia catalítica en la generación de plasmina, tal y como observamos en las figuras 14 y 15. La fibrina unida al plasminógeno es un sustrato mejor para el t-PA que el plasminógeno en circulación. La plasmina degrada a la fibrina a la que se ha unido y entonces se generan nuevos residuos C-terminales de lisina que estimularán a su vez una mayor

formación de plasmina llegándose así finalmente a la lisis del coágulo. Se ha demostrado que TAFIa inhibe a la fibrinólisis mediante la lisis de estos residuos de lisina C-terminales de la fibrina, y, en consecuencia, limitando la formación de plasmina. La activación de TAFI en el plasma a cargo del complejo trombina-trombomodulina durante la coagulación da como resultado la lisis de las lisinas C-terminales de la fibrina parcialmente digerida, con un descenso concomitante de la unión del plasminógeno y el retardo de la ruptura del coágulo. TAFIa inhibe, por tanto, la activación del Glu-plasminógeno y su conversión a Lys-plasminógeno. El efecto de TAFIa en la fibrinólisis puede ser además potenciado por la activación de TAFI mediante plasmina y la inactivación de esta plasmina por TAFIa, aunque se ha demostrado que la plasmina puede también inactivar a TAFIa.

### **I.C.3. TAFI.**

#### **I.C.3.a. Síntesis y caracterización de TAFI.**

La TAFI en humanos se sintetiza en el hígado como una proteína de 423 aminoácidos. El péptido señal de 22 aminoácidos se hidroliza intracelularmente (Eaton et al, 1991). Recientemente, se ha identificado la presencia de TAFI en plaquetas, y se ha demostrado que la enzima puede ser secretada gracias a la estimulación de las mismas, mientras que por microscopía de inmunofluorescencia se observó un patrón consistente con la localización intracelular de la proteína en los gránulos- $\alpha$ .

La TAFI es una glicoproteína con un peso molecular de 55 kDa y un punto isoeléctrico de 5.0. Los carbohidratos unidos a la molécula constituyen aproximadamente un 20% de la masa total. TAFI circula en plasma a una concentración de 4-15  $\mu\text{g}/\text{ml}$ . Estudios de eliminación en ratones sugieren una vida media para TAFI de varias horas (Bajzar et al, 1996a, Mosnier et al, 1998)

Se ha sugerido que TAFI circula en plasma unida al plasminógeno (Tan & Eaton, 1995), aunque aún no se han hallado evidencias directas de ello. TAFI se une al plasminógeno, y lo hace con una afinidad 10 veces superior al Lys-plasminógeno que al Glu-plasminógeno. La conversión del Lys-plasminógeno a plasmina no es específica de fibrina y, por tanto, la formación de un complejo del Lys-plasminógeno con TAFI puede prevenir la formación de plasmina en fase fluída.

Ahora bien, dado que TAFIa posee una afinidad reducida para el plasminógeno y debido a que su masa molecular está por debajo del límite de filtración glomerular, probablemente el enzima activo no es retenido por la circulación, aunque se ha descubierto que dicha

eliminación puede venir modulada por la unión de TAFI a la  $\alpha 2$ -macroglobulina y a la *pregnancy zone protein* (Valnickova et al, 1996).

TAFIa muestra preferencia para los residuos C-terminales de arginina sobre los de lisina (Wang et al, 1994). Su pH óptimo de actuación es de 7,7 y se inhibe por EDTA, PCI, DTT,  $\beta$ -mercaptoetanol, GEMSA y LCI (Hendricks, 1990, Boffa, 1998, Reverter, 1998).

TAFI es un enzima conservado entre numerosas especies de mamíferos (se ha analizado su presencia en cerdo, rata, conejo, perro, ratón... (Marx et al, 2000, Schattermann et al, 1999, Bajzar et al, 1997); de todos modos, existen grandes diferencias en la actividad de TAFIa en las tres especies con referencia a la humana, encontrando desde un 20% de actividad en ratón a un 500% en cerdo. Al igual que en el hombre, TAFIa es inestable en todas las especies estudiadas, y se ha postulado que su función se ha conservado a lo largo de la evolución, dado que la inhibición de TAFI por PCI en todas ellas tenía como consecuencia la disminución del tiempo de fibrinólisis.

Finalmente, cabe decir que se han desarrollado con éxito medios de producción de TAFI recombinante, siendo el producto plenamente activo, y protocolos de purificación del enzima desde plasma, cada vez con mayor rendimiento (Zhao et al, 1998).

### **I.C.3.b. Activación de TAFI.**

La activación de TAFI puede ser producida por tripsina, plasmina, trombina o meizotrombina, y sucede por una sola hidrólisis en la Arg-92. Análogamente a la CPB pancreática, el centro de unión de sustrato de la unidad catalítica tal vez se halle cubierto por el péptido de activación como puede desprenderse de los modelos publicados. La liberación del péptido de activación descubre el sitio de unión, haciendo posible la interacción de la Arg-143 de la TAFIa con el grupo carboxilo C-terminal de un posible sustrato. Aunque TAFI posee 8 residuos de cisteína, no existen puentes disulfuro entre el péptido de activación y la unidad catalítica. La glicosilación del péptido de activación puede actuar de modo que estabilice y aumente la vida media de la TAFI circulante.

La activación de TAFI por trombina es un proceso ineficiente ( $K_M$  0.5-0.21  $\mu M$ ;  $k_{cat}$  0.0021  $s^{-1}$ ) y, en consecuencia, se requieren grandes cantidades de trombina (Bajzar et al, 1996b, Boffa et al, 1998, Kokame et al, 1998). El receptor de las células endoteliales, la trombomodulina, estimula la activación de TAFI por trombina unas 1250 veces, lo que se debe casi exclusivamente a un incremento en la eficiencia catalítica ( $K_{cat}$ , en presencia de trombomodulina es 0.4-1.2  $s^{-1}$ ). Comparado con la trombina, la  $K_M$  para la activación de TAFI mediada por plasmina es mucho más baja (55nM), mientras que la  $K_{cat}$  es 5 veces más baja

( $0.0004 \text{ s}^{-1}$ ) (Mao et al, 1999). La activación de TAFI por plasmina está estimulada por glicosaminoglicanos alrededor de unas 15 veces. Las demás enzimas de las cascadas de coagulación y fibrinólisis como la urokinasa, t-PA, APC, kalicreína y los factores VIIa, IXa y Xa no son capaces de activar TAFI (Schattmann, 2000).

La trombomodulina parece jugar un papel dual; por un lado, ralentiza la generación de trombina mediante el aumento de la activación de la proteína C por la propia trombina, mientras que por otro lado, hace que estas concentraciones bajas de trombina sean más efectivas en la activación de TAFI con una ralentización de la fibrinólisis como resultado. Aunque la activación de TAFI y de la proteína C pueden suceder simultáneamente, la concentración de trombomodulina es el factor determinante en el efecto final. La activación de TAFI se estimula a concentraciones bajas (5nM) de trombomodulina, mientras que la activación de TAFI disminuye a mayores concentraciones de trombomodulina (10nM). La reducción de la activación de TAFI a concentraciones más elevadas se ha demostrado que se debe a la inhibición de la generación de trombina por la proteína C activa. Esto sugiere que la trombomodulina es un agente antifibrinolítico a concentraciones bajas y profibrinolítico a concentraciones elevadas (Mosnier et al, 2001). Este fenómeno de la regulación diferencial de la fibrinólisis puede jugar un rol *in vivo* mientras la expresión de la trombomodulina en células endoteliales varía en diferentes tejidos. Además el tamaño del vaso sanguíneo puede ser un factor importante debido a que la concentración efectiva de trombomodulina se incrementa en el fluido sanguíneo desde la aorta hacia los capilares.

### **I.C.3.c. Inactivación de TAFI.**

La actividad enzimática de TAFIa es inestable, y se ha descrito su inactivación mediante proteólisis y por un proceso espontáneo dependiente de temperatura (Wang et al, 1994, Bajzar et al, 1996, Boffa et al, 1998). Después de la incubación con tripsina, plasmina o trombina se halló que su actividad incrementaba para luego decaer, puesto que después de la proteólisis inicial de activación, el enzima seguía sufriendo cortes proteolíticos que lo inactivaban (en la Arg-92 y Arg-330) (Eaton et al, 1991). La identificación de los lugares de proteólisis en TAFI que se generaban en su activación e inactivación por el complejo TR-TM indicaron que la forma activa de TAFI de 36 kDa era hidrolizada en la Arg-302 en dos polipéptidos de 25 y 11 kDa. Por mutagénesis dirigida se ha confirmado que Arg-302 es el principal lugar de hidrólisis por TR-TM, sugiriendo que la inactivación de TAFI viene causada por su inestabilidad conformacional tal y como había sido descrito.

TAFIa es altamente sensible a la temperatura: su vida media aumenta desde 10 min. a 37°C, y 45 min. a 30°C, hasta varias horas a 22°C. A 0°C, la enzima es estable (Boffa et al, 1998). Puesto que la pérdida de actividad de TAFIa está directamente relacionada con un decrecimiento de su fluorescencia intrínseca, se piensa que hay un cambio estructural espontáneo que provoca esta pérdida de actividad enzimática (Boffa et al, 2000). Probablemente, el cambio conformacional que provoca la inactivación hace que TAFI sea más susceptible a la proteólisis, y además, la escisión proteolítica hace de la inactivación de TAFI un proceso irreversible. De manera análoga, los residuos lisina de la fibrina pueden tener un efecto estabilizante en la actividad de TAFI.

### **I.C.3.d. El papel de TAFI en el modelo de la coagulación revisado.**

Se ha postulado un modelo para la coagulación sanguínea, basado en los resultados de los últimos años, donde no se hace distinción entre la vía intrínseca y la extrínseca. Según este modelo, la coagulación se inicia por la exposición del factor tisular (*tissular factor*, TF) (Bouma et al, 1998). El TF se une al factor VII(a) y este complejo activa a los factores IX y X. A concentraciones elevadas de TF, el factor X se activa principalmente por el complejo TF-VIIa, mientras que a bajas concentraciones de TF, la contribución del complejo TF-X en la activación del factor X se ve más pronunciada (Osterud & Rapaport, 1977, Marlar et al, 1982). Una vez que se han generado pequeñas cantidades de factor Xa, la vía extrínseca se inhibe rápidamente por el TFPI. El TFPI en primer lugar forma un complejo con el factor Xa, que al unirse al complejo TF-factor VIIa, resulta en un complejo cuaternario que no puede sostener por más tiempo la coagulación. Mientras, se generan cantidades suficientes de trombina para inducir la formación de fibrina. La formación de trombina continúa dentro del coágulo de fibrina y, de hecho, la mayor parte de la trombina se genera después de la formación del coágulo (Rand et al, 1996). La generación de trombina dentro del coágulo se produce principalmente a través de la vía intrínseca, mediante la activación del factor XI por la trombina unida a fibrina en el exterior del coágulo.

Aunque sólo se forman pequeñas cantidades del factor XI, se produce un segundo “*burst*” (explosión) de formación de trombina gracias a la vía intrínseca por la continua activación del factor XI por trombina y el poder de amplificación del complejo tenasa y protrombinasa. Estas concentraciones elevadas de trombina se necesitan para la activación de TAFI y la inactivación del u-PA, lo que resulta en un descenso de la fibrinólisis. En la figura 16 podemos observar un esquema del nuevo modelo de coagulación.



## I.C.4. IMPORTANCIA DE TAFI A NIVEL PATOLÓGICO

A partir de diversos estudios se ha determinado el nivel de TAFI en individuos sanos, hallándose en la población normal un amplio rango de niveles de TAFI (Chetaille et al, 2000). En individuos normales, los niveles antigénicos de TAFI se corresponden con la actividad de TAFI y con los ensayos de coagulación sanguínea (Mosnier, 1998). No se ha hallado ninguna diferencia en la concentración antigénica de TAFI entre hombres y mujeres. Los análisis del efecto asociados a la edad muestran que mientras no haya un incremento de TAFI relacionado con la edad en los hombres, éste sí se encuentra en mujeres (Van Tilburg et al, 2000). La terapia hormonal, incluyendo anticonceptivos orales, incrementa el nivel de TAFI, mientras que el embarazo no tiene efecto en la concentración (Meijers et al, 2000).

Recientemente se ha descrito una relación de los niveles de TAFI con trastornos tromboticos, por lo que niveles elevados de TAFI se asocian a un riesgo moderado de sufrir trombosis venosa. El incremento de riesgo de sufrir trombosis venosa puede explicarse por la inducción de un estado hipofibrinolítico (Silveira et al, 2000). También se han hallado niveles elevados de TAFI en pacientes de angina de pecho estable y enfermedades coronarias.

En humanos, las concentraciones de TAFI que exceden de 129nM se correlacionan con un riesgo 2 veces mayor de sufrir trombosis venosa profunda. Aunque TAFI tan solo sea un factor de riesgo moderado, la consistencia de las pruebas clínicas sugiere una implicación de TAFI en la regulación de la fibrinólisis (Van Tilburg et al, 2000).

También se ha sugerido un rol para TAFI en enfermedades inflamatorias. Los niveles de TAFI en plasma se correlacionan con marcadores de fase aguda (proteína C-reactiva y haptoglobina) en individuos sanos, por lo que es posible que existan mecanismos por los que el incremento en los niveles de TAFI durante la inflamación pueda contribuir al estado protrombótico y antifibrinolítico característico de la coagulación intravascular diseminada (DIC) (Takano et al, 1990). El rol de TAFI en la inflamación podría darse en ambos sentidos: por un lado, el aumento de los niveles de TAFI puede contribuir a la inactivación de los mediadores de la inflamación, como las anafilotoxinas y los péptidos inflamatorios (C3a y C5a), reduciendo la susceptibilidad a un choque séptico. Por otro lado, los niveles elevados de TAFI pueden contribuir a una inhibición elevada de la fibrinólisis y por tanto facilitar los efectos de deterioro de la DIC.

Otro estudio (Hori et al, 2002) ha demostrado que pacientes de *diabetes mellitus* tipo 2, con obesidad y resistencia a la insulina, presentan un nivel circulante elevado de TAFI, lo que puede ser la causa de su cuadro hipofibrinolítico.

En el otro extremo de las patologías se han detectado niveles bajos de TAFI en pacientes de leucemia promielocítica aguda, aunque, de manera interesante, la concentración antigénica de TAFI es normal, mientras que el nivel de actividad de TAFI desciende severamente un 60% (Meijers et al, 2000b). La deficiencia funcional de TAFI en la leucemia puede contribuir a la severidad de la diatesis hemorrágica debido a la capacidad dispareja del sistema de coagulación para proteger el coágulo de fibrina de la fibrinólisis.

Una activación defectiva de TAFI puede contribuir asimismo a la severidad de un desorden hemorrágico en las hemofilias A y B. Estos pacientes tienen una formación de trombina reducida por su vía extrínseca a bajas concentraciones de factor tisular y también tienen reducido la segunda explosión de generación de trombina. Debido a esto, la activación de TAFI, y en consecuencia el descenso de la regulación de la fibrinólisis de la vía intrínseca es defectivo, dato confirmado por la detección de bajos niveles circulantes de TAFI (Bouma et al, 1999).

### **I.C.4.a. Nuevo enfoque de las terapias coadyuvantes.**

El acrónimo TAFI implica que su papel fisiológico es la regulación de la fibrinólisis. Los datos obtenidos hasta hoy día tienden a soportar dicha hipótesis. TAFI se halla presente en el plasma de toda una variedad de mamíferos, y en todos los casos estudiados, puede ser activado por trombina y trombomodulina de modo que se produce un enzima térmicamente inestable, capaz de inhibir la fibrinólisis. Estos resultados sugieren que TAFI es parte de una vía conservada que involucra la regulación de la fibrinólisis. Redlitz et al proporcionaron la primera indicación de que TAFI estaba involucrada en la regulación de la fibrinólisis *in vivo*. Demostraron que la actividad de TAFIa, en muestras de plasma de perro con trombosis inducida eléctricamente, aumenta durante la trombosis y la terapia trombolítica (Redlitz et al, 1996). Además, la concentración de TAFIa se correlaciona con el tiempo de restauración del flujo sanguíneo. En consecuencia, y al detectar niveles elevados de TAFI correlacionados con desórdenes trombóticos, se planteó el uso de inhibidores específicos de TAFI para favorecer la fibrinólisis. Se han realizado numerosos ensayos *in vitro* demostrando el efecto del PCI en la aceleración de la fibrinólisis y un número ya significativo de ensayos *in vivo*, en los que se lo utiliza como un adyuvante trombolítico en conejos y ratones (Refino et al, 2000, Nagashima et al, 2000).

La coin fusión en conejos de un inhibidor de TAFIa (PCI) con t-PA (la molécula más utilizada en terapia trombolítica) aumenta de manera considerable la trombolisis, o bien reduce de manera también considerable la cantidad necesaria de t-PA para obtener la misma lisis

(Klement et al, 1999). La potenciación del agente trombolítico con un inhibidor de TAFIa puede incrementar la eficacia trombolítica sin que haya un descenso en la seguridad. Dado que TAFI se activa por trombina localizada en el trombo, la inhibición de TAFIa puede generar un incremento de la trombolisis con una especificidad adicional por fibrina. Un inhibidor de TAFIa (PCI, nuevamente) también reduce la mortalidad asociada al tromboembolismo inducido por trombina en ratones (Minemma et al, 1995). Por otro lado, los riesgos de sangrado de un tipo de terapia semejante serían bajos debido a que esta terapia no viene acompañada por los efectos anticoagulantes tan fuertes vistos en la heparina y en la hirudina, rasgo que confiere una innegable ventaja a esta aproximación terapéutica.

Aunque el rol patofisiológico de TAFI permanece todavía sin esclarecerse, es evidente que TAFI constituye un importante nexo de unión entre la coagulación y la fibrinólisis. Un mayor estudio de este mecanismo puede conducirnos a mejorar la eficiencia de la terapia trombolítica en el tratamiento de los desórdenes trombóticos, mientras que otros agentes que incrementen la eficiencia de TAFI pueden ser útiles en el tratamiento de los desórdenes hemorrágicos de las hemofilias A y B.

## **I.D. CONEXIÓN DEL SISTEMA FIBRINOLÍTICO CON EL LCI**

El hecho de que el LCI sea un potente inhibidor de un enzima que puede tener un papel fundamental en la regulación de la fibrinólisis sugiere que podría tener sentido pensar en él como modificador también de la cascada fibrinolítica. Es decir, establecida la hipótesis de que un exceso de actividad de TAFI puede provocar desórdenes de tipo trombótico (la velocidad normal de lisis de los coágulos se ve considerablemente retardada en esta condición), la inhibición de la actividad de TAFI podría reducir el tiempo de lisis del coágulo. Observados los resultados tan positivos obtenidos con el PCI, tanto en ensayos *in vitro* como en tratamiento *in vivo*, resulta atractivo estudiar el efecto del LCI en la fibrinólisis, en ensayos *in vitro* con plasma humano, de modo que pudiéramos observar si la capacidad fibrinolítica del LCI era comparable a la del PCI.

Por otro lado no debemos perder de vista que el LCI es un producto natural de la sanguijuelas y probablemente tenga una función biológica muy cercana a la nombrada, cosa que supondría una ventaja evolutiva del animal. Por tanto, es igualmente lógico pensar que el animal potencia sus capacidades anticoagulantes empleando un inhibidor que tenga como diana un enzima que ayuda a mantener la estabilidad del coágulo (en individuos sanos)

## *Introducción*

usándolo como coadyuvante de, por ejemplo, otras moléculas como la hirudina para mantener la sangre líquida el máximo tiempo posible y poder así alimentarse. Es decir, la hipótesis de que una molécula de semejantes características podía bien convertirse en un fármaco coadyuvante de las terapias trombolíticas debía de verse soportada por pruebas experimentales a un primer nivel *in vitro*. Éstas son las que serán aportadas en el tercer trabajo de esta tesis.

## **II. OBJETIVOS**

El objeto de estudio de la presente tesis es el inhibidor de carboxipeptidasas de sanguijuela (LCI) tanto desde un punto de vista estructural como desde un punto de vista funcional. Este estudio se inscribe dentro del marco general de la investigación sobre los inhibidores de metalocarboxipeptidasas, que se lleva a cabo en nuestro grupo.

Un primer bloque de objetivos comprende el estudio del plegamiento del LCI, de cara a una mayor caracterización estructural y de la estabilidad de la molécula.

**Los objetivos de esta primera parte han sido:**

1. Dilucidar la vía de plegamiento oxidativo del LCI mediante la identificación y aislamiento de sus intermediarios de plegamiento por la técnica de captura de intermediarios en medio ácido.
2. Caracterizar los posibles intermediarios de plegamiento del LCI. Definir el número de puentes disulfuro y el apareamiento de cisteínas de dichos intermediarios para observar la posible aparición de puentes disulfuro nativos y posibles formas *native-like*.
3. Comparar el camino de plegamiento oxidativo del LCI con otros caminos ya bien caracterizados de proteínas pequeñas ricas en puentes disulfuro.
4. Dilucidar la vía de desplegado del LCI mediante la técnica del intercambio de disulfuros (*scrambling*) y estudiar la influencia de diversos agentes desnaturizantes, incluidos los solventes orgánicos y las elevadas temperaturas.
5. Estudiar la estabilidad conformacional del LCI nativo mediante el trazado de sus curvas de desnaturización. Comparar la estabilidad conformacional del LCI con la de otras proteínas pequeñas ricas en puentes disulfuro, ya caracterizadas, mediante el cálculo de las concentraciones necesarias de diversos agentes desnaturizantes, requeridas para desplegar el 50% de la molécula.

El interés en el estudio de la funcionalidad biológica del LCI tiene su origen en su capacidad inhibidora de la pCPB o TAFI, y en el importante papel de este enzima en la regulación de la fibrinólisis.

**Los objetivos de esta segunda parte han sido:**

1. Estudiar el efecto del LCI en la fibrinólisis, en un sistema de coagulación *in vitro* (plasma humano) y comparar dicho efecto con el del PCI. Medir las variaciones de las velocidades de fibrinólisis de los coágulos en condiciones de elevada actividad TAFI, en ausencia o presencia de concentraciones crecientes de LCI y PCI, de cara

- a comparar las dosis necesarias de ambos inhibidores para obtener un mismo efecto.
2. Calcular la actividad enzimática de TAFI en plasma, reproduciendo las diversas condiciones del ensayo fibrinolítico, para hallar una correlación entre las variaciones de velocidades de fibrinólisis y las variaciones de actividad enzimática de TAFI debidas a la presencia de los inhibidores LCI o PCI.
  3. Investigar si el retraso fibrinolítico provocado por la actividad de TAFI tiene una base morfológica en los coágulos, mediante microscopía electrónica de rastreo y de transmisión. Detectar posibles anomalías en las redes de fibrina al reproducir las condiciones del ensayo fibrinolítico, tanto en ausencia como en presencia de inhibidores, y hallar una correlación entre la superestructura de las redes de fibrina y los resultados obtenidos en los estudios de fibrinólisis y de actividad enzimática.

### **III. TRABAJO 1.**

## **Major Kinetic Traps for the Oxidative Folding of Leech Carboxypeptidase Inhibitor**

### III.A. INTRODUCTION

A well established method for analyzing the mechanism of protein folding is **oxidative folding** of disulfide containing proteins (1-9). In this technique, proteins are initially fully reduced and denatured in the presence of a reducing agent and a denaturant. After removal of both reagents, the reduced and denatured protein is allowed to refold in the presence of a redox buffer (6). The disulfide folding pathway is then elucidated from the mechanism of formation of the native disulfide bonds, and characterized by the heterogeneity and the structures of disulfide isomers that accumulate along the process of oxidative folding. Folding intermediates can be trapped by several techniques. The most effective method is acid quenching of the protein solution, which allows the structural characterization of the trapped intermediates. This technique also permits to study the influence of redox agents on the folding pathway and the efficiency of recovery of the native protein (6,10).

A considerable number of 3-disulfide containing proteins have been studied by the technique of oxidative folding. Among them, the extensively reported BPTI (bovine pancreatic trypsin inhibitor) (1,2,7,8), as well as hirudin (10,11), potato carboxypeptidase inhibitor (12,13), tick anticoagulant peptide (14,15), epidermal growth factor (16,17), and insulin like growth factor (18,19). However, few models aside from ribonuclease A (20-23) and  $\alpha$ -lactalbumin (24-27) have been studied in detail among 4-disulfide containing proteins. For these proteins, the increasing number of possible isomers (a total of 762 including the 1-, 2-, 3- and 4-disulfide species) presents a formidable task of isolation and characterization of folding intermediates, even if only a fraction of these isomers exists along the pathway.

To date, no general consensus regarding the disulfide folding pathway has been reached from the proteins investigated. The accumulated results have shown that even amongst small 3-disulfide proteins, the folding mechanisms vary substantially. BPTI (1,2,8,9) and hirudin (10,11) exemplify two important cases of such divergent folding mechanisms. For BPTI, folding proceeds *via* a limited number of intermediates with native disulfide bonds and adopting native-like structures. Out of 75 possible disulfide isomers, only five to six 1- and 2-disulfide intermediates were shown to populate along the folding pathway of BPTI and all of them were found to adopt native disulfide bonds. In the case of hirudin, a highly heterogeneous population of 1- and 2-disulfide intermediates were identified along the folding pathway. Most importantly, fully oxidized 3-disulfide (scrambled) isomers, which have not been observed in the case of BPTI, were shown to serve as major folding intermediates of hirudin. For hirudin, folding proceeds through an initial stage of non-specific disulfide

formation (packing) to form the scrambled isomers, which is followed by disulfide reshuffling (consolidation) of the heterogeneous scrambled population to reach the native structure.

Therefore, further study of folding models from novel proteins is required in order to better understand the underlying causes that generate such a diversity of disulfide folding pathways. In the present work we report the pathway of oxidative folding of leech carboxypeptidase inhibitor (LCI), a 66-residue polypeptide rich in cysteines (28). The recently reported three-dimensional structure shows that LCI folds in a compact domain consisting of a five-stranded anti-parallel  $\beta$ -sheet and a short  $\alpha$ -helix (29). The molecule is stabilized by the presence of four disulfide bridges, all of them located within secondary structure elements. LCI displays no structural similarity to other protease inhibitors, except to potato carboxypeptidase inhibitor (30,31), with which it shares a C-terminal sequence homology and a substrate-like-manner inhibitory mechanism for pancreatic-like carboxypeptidases. LCI is also a potent inhibitor of plasma carboxypeptidase B, also called TAFI (Thrombin Activatable Fibrinolysis Inhibitor), an enzyme involved in the equilibrium of the blood clots. LCI could thus be a possible candidate for pharmacological application in thrombosis disease. Our laboratory has recently described the unfolding pathway of LCI (32), showing that the rate constant of the process is low when compared to other disulfide-containing proteins and that the protein is highly stable. All these features confirm the idoneity of LCI is a model for the study of oxidative folding pathways.

## **III.B. EXPERIMENTAL PROCEDURES**

### **III.B.1. Materials.**

Recombinant LCI was obtained by heterologous expression in *Escherichia coli* following a procedure previously described (28). The recombinant protein that contains a construction added glycine as the N-terminal residue was purified by ion-exchange chromatography on a TSK DEAE column (Amersham Biosciences) followed by reverse-phase HPLC. The protein was over 99% pure, as judged by HPLC, and its molecular mass was confirmed by MALDI-TOF. Thermolysin (P-1512), dithiothreitol, reduced glutathione (GSH), oxidized glutathione (GSSG),  $\beta$ -mercaptoethanol, were obtained from Sigma with purities greater than 99%.

### **III.B.2. Control folding experiments.**

Native LCI (0.5 mg) was initially reduced and denatured in Tris-HCl buffer (0.1M, pH 8.4) containing 6 M GdnCl and 30 mM dithiothreitol. The reaction was carried out at 22°C for 90 minutes. The sample was then passed through a NAP-5 column (Amersham Biosciences), previously equilibrated in Tris-HCl buffer (0.1 M, pH 8.4), to initiate the oxidative refolding. Desalting took approximately 1 minute. Reduced and unfolded LCI was recovered in 0.55 ml and immediately diluted to a final concentration of 0.5 mg/ml in the same Tris-HCl buffer, both in absence (control-) or presence (control +) of 0.25 mM  $\beta$ -mercaptoethanol. Folding intermediates were trapped in a time-course manner at selected times by mixing aliquots of the sample with an equal volume of 4% trifluoroacetic acid in water. Trapped folding intermediates were analyzed by HPLC-reverse phase chromatography. HPLC conditions are shown in the figure legends.

### **III.B.3. Folding of LCI in the presence of redox agents.**

Reduction, denaturation and desalting were carried out as described in the preceding paragraph. GSSG (0.5mM) or a mixture of GSSG/GSH (0.5mM/1mM) were added immediately after unfolded LCI was desalted by NAP-5 column. Folding intermediates were trapped with an equal volume of 4% trifluoroacetic acid in water and analyzed by HPLC-reverse phase chromatography.

### **III.B.4. Analysis of folding intermediates of LCI.**

Acid-trapped folding intermediates of LCI were purified from HPLC and freeze-dried. The samples were derivatized with 50  $\mu$ l of vinyl pyridine (0.1M) in the Tris-HCl buffer (0.1M pH 8.4) at 23°C for 35 minutes. The reaction was quenched with 4% trifluoroacetic acid. Vinyl-pyridine derivatized samples were then analyzed by mass spectrometry in order to characterize the number of disulfide bonds in the folding intermediates population.

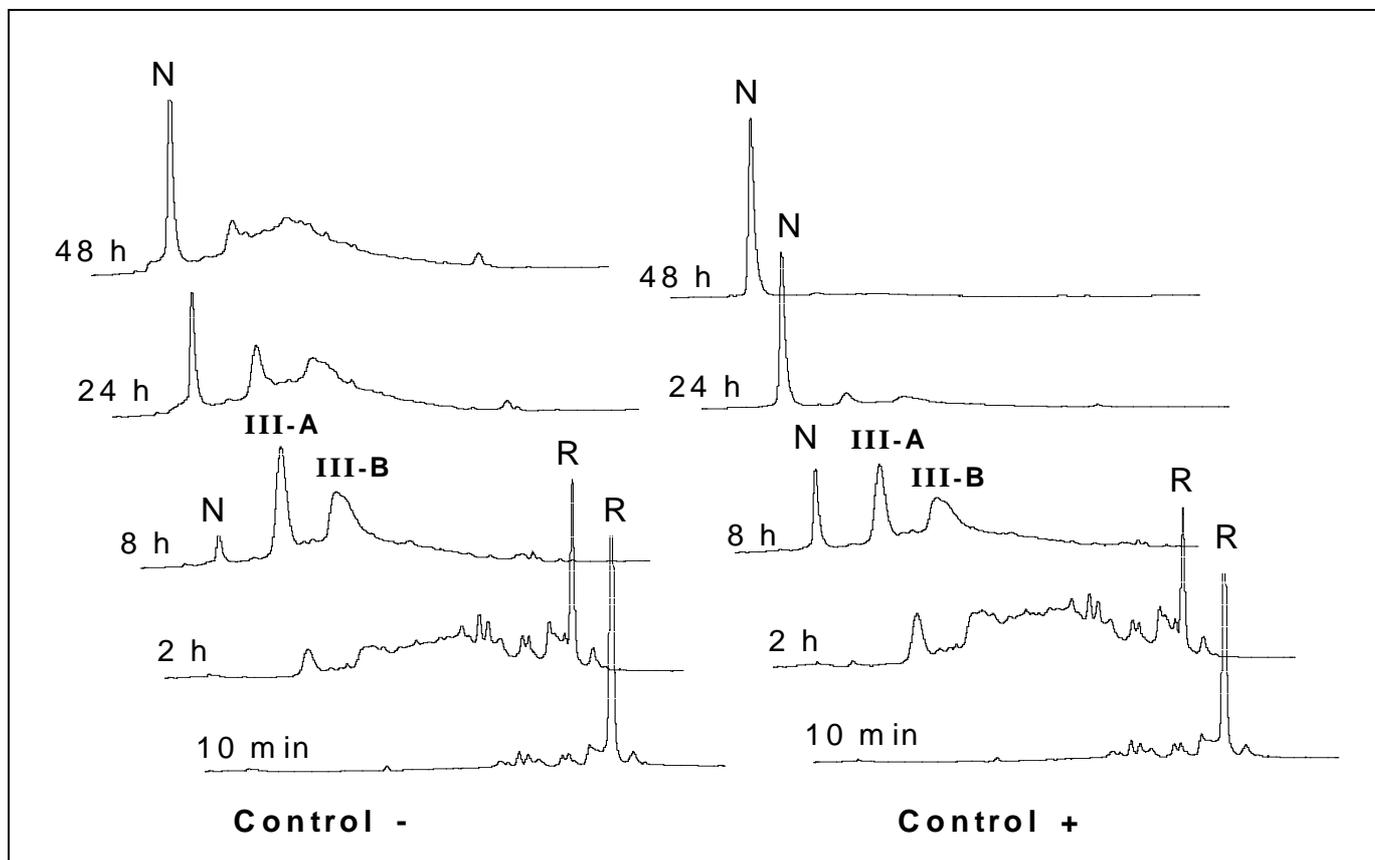
### **III.B.5. Characterization of the disulfide structure of the major kinetic traps.**

The major HPLC fractions containing non-fully reduced non-native LCI species detected in the control experiments were isolated, freeze dried and derivatized with vinyl pyridine as reported for the analysis of folding intermediates of LCI. Vinyl-pyridine derivatized samples were further purified by HPLC and freeze-dried. The isolated intermediates III-A1, III-A2, III-B (20  $\mu$ g each) were treated with 2  $\mu$ g of thermolysin (Sigma, P-1512) in 30  $\mu$ l of N-ethylmorpholine/acetate buffer (50 mM, pH 6.4). Digestions were

carried out for 16 h at 37 °C. Thermolytic products were then isolated by HPLC and analyzed by amino acid sequencing and mass spectrometry to identify the disulfide-containing peptides.

### III.B.6. Amino Acid Sequencing and Mass Spectrometry.

The amino acid sequence of disulfide-containing peptides was analyzed by automatic Edman degradation using a PerkinElmer Life Sciences Procise sequencer (model 494) equipped with an on-line phenylthiohydantoin-derivate analyzer. The molecular masses of disulfide-containing peptides were determined by a matrix-assisted laser desorption/ionization-time-of-flight mass spectrometer (Perkin Elmer Life Sciences Voyager-DE STR).



**Figure 1. HPLC analysis of the folding intermediates of LCI trapped by acidification.** Folding was carried out in the Tris-HCl buffer (0.1M, pH 8.4) in the presence (control +) or the absence (control -) of  $\beta$ -mercaptoethanol (0.25 mM). Folding intermediates were trapped by mixing with an equal volume of 4% aqueous trifluoroacetic acid and analyzed by HPLC using the following conditions: Solvent A was 0.1% trifluoroacetic acid in water; solvent B was acetonitrile/water (9:1, v/v) containing 0.085% trifluoroacetic acid. The sample was applied to a Zorbax 300SB C-18 column (4.6 mm  $\times$  5  $\mu$ m). A linear gradient was developed between 31 and 47% of solvent B in 45 min, with a flow rate of 0.5ml/min. Column temperature was 23°C. **N** and **R** indicate the elution position of the native and fully reduced species of LCI, respectively. III-A and III-B are two major fractions of 3-disulfide intermediates identified along the pathway of oxidative folding of LCI.

## III.C. RESULTS

### III.C.1. Oxidative folding of LCI in the absence of redox agents.

Folding experiments of reduced and denatured LCI were first performed using simple conditions in the Tris-HCl buffer alone (control -) and Tris-HCl buffer in the presence of  $\beta$ -mercaptoethanol (0.25 mM). HPLC profiles of folding intermediates trapped by acid quenching at selected time points are presented in Fig. 1. A high degree of heterogeneity of intermediates is observed during the early stages of folding. The chromatographic behavior of the folding intermediates clearly shows the accumulation of two fractions of major kinetic traps at the late stage of folding, designated here as III-A and III-B. Their patterns remain indistinguishable regardless of the presence of  $\beta$ -mercaptoethanol.

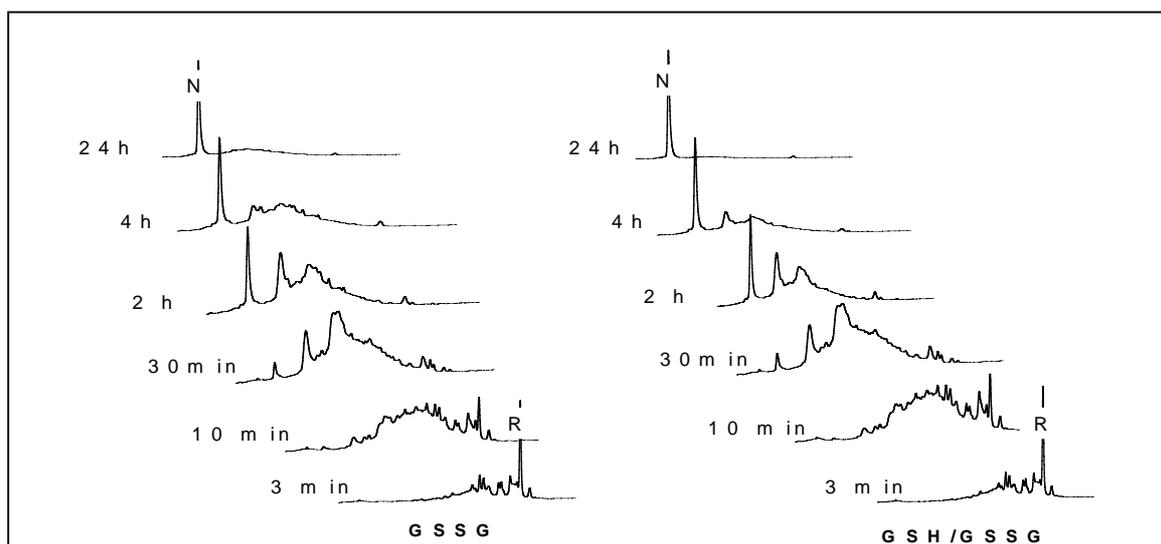
Purified intermediates from the HPLC analyses were reacted with vinyl pyridine and analyzed by MALDI mass spectrometry in order to evaluate the disulfide bond content of folding intermediates. The data (not shown) allow calculation of the relative concentrations of the fully reduced, 1-disulfide, 2-disulfide, 3-disulfide and 4-disulfide (“scrambled” isoforms plus the native) species for each time-course-trapped sample. The results are shown in Fig. 3 (parts A and B). The recovery of native LCI can be distinguished from that of 4-disulfide species by its unique HPLC elution position (Fig. 1) and is represented by bars in Fig. 3.

The folding of LCI could not reach completion in the absence of  $\beta$ -mercaptoethanol, since only about 40% of the protein was recovered as the native structure after 48h of folding, while the remaining 60% of LCI was trapped as folding intermediates. Among them, a 35% corresponds to 3-disulfide species represented by III-A and III-B and 25% to non-native 4-disulfide species made up of a heterogeneous population of scrambled isomers. In the presence of  $\beta$ -mercaptoethanol, the recovery of native LCI is significantly higher, with more than 90% of the protein recovered as the native structure after 48 hours of folding. In these conditions, the recovery rate is also more closely related to the generation of 4-disulfide LCI (Fig. 3B). These results confirm the co-existence of 3-disulfide and non-native 4-disulfide isomers as folding intermediates and major kinetic traps of LCI folding. They also demonstrate the role of  $\beta$ -mercaptoethanol as thiol catalyst for promoting the disulfide shuffling and the conversion of scrambled LCI to the native structure.

### III.C.2. Oxidative folding of LCI in the presence of redox agents.

Oxidative folding of LCI was subsequently performed in the presence of redox agents, including GSSG (0.5 mM) and a mixture of GSH/GSSG (1 mM/0.5 mM). Folding intermediates were similarly analyzed by HPLC. As shown in Fig. 2, HPLC patterns obtained in both conditions are almost indistinguishable, and they are also very similar to those obtained in the control experiments (Fig. 1). Fractions III-A and III-B were also shown to accumulate along the folding pathway under these conditions, confirming that isoforms eluted within these two fractions are major intermediates of the oxidative folding of LCI. However, the addition of GSSG accelerates the rate of LCI folding by approximately 10-12 fold. The increase in the rate of folding is observable both through the kinetics of disulfide oxidation that leads to the generation of native LCI and by the total recovery of the native species. Similar effects of GSSG on the efficiency of oxidative folding have been demonstrated in other proteins (10,12,14,17).

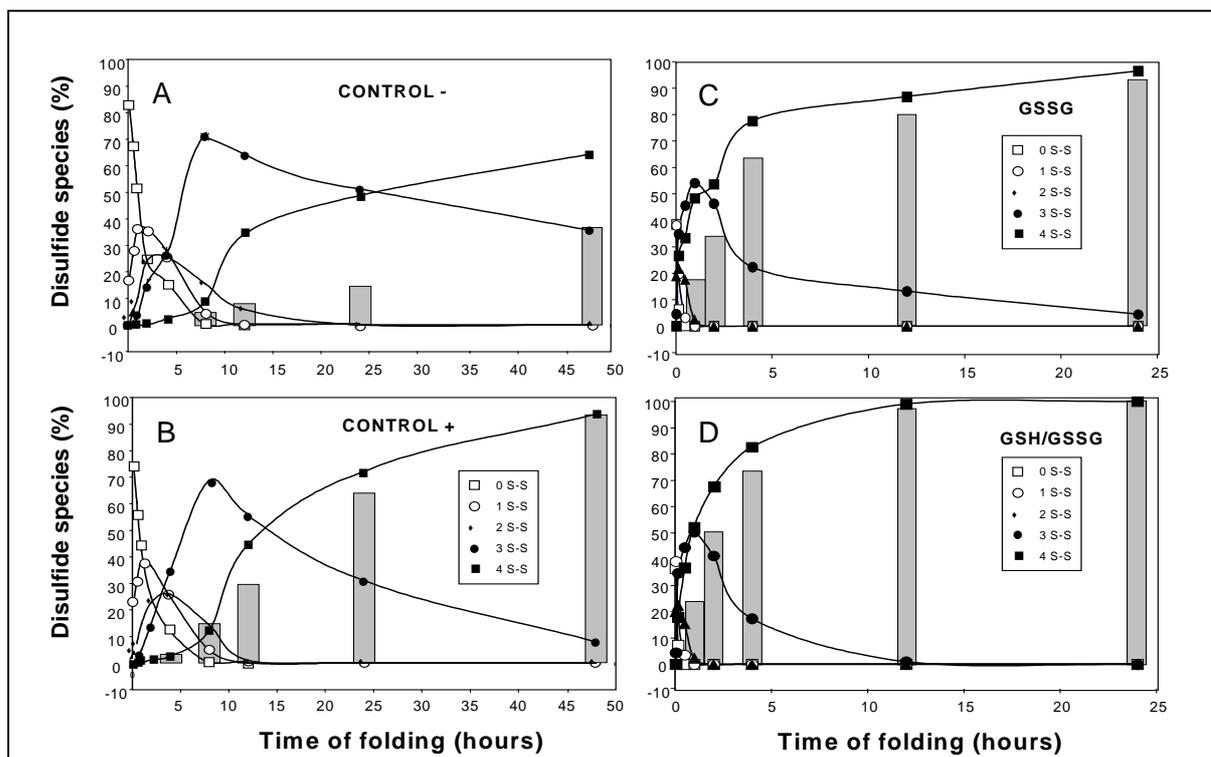
Analysis of the composition of disulfide species of time course-trapped intermediates (Fig. 3C and 3D) revealed the rapid accumulation of 3- and 4-disulfide LCI during the folding, along with the almost negligible recovery of native LCI beyond that of 4-disulfide species. The fact that native LCI was not accumulated in any significant amount above the observed levels of 4-disulfide scrambled isomers indicates that these species are essential folding intermediates of LCI.



**Figure 2. HPLC chromatograms of the intermediates of oxidative folding of LCI in the presence of redox agents.** Folding was performed in the Tris-HCl buffer (0.1M, pH 8.4) in the presence of GSSG (0.5 mM) or a mixture of GSH/GSSG (0.5 mM/1 mM). Folding intermediates were trapped by mixing with an equal volume of 4% aqueous trifluoroacetic acid and analyzed by HPLC using the same conditions described in the legend of Fig. 1. **N** and **R** indicate the elution position of the native and fully reduced species of LCI. III-A and III-B are two major fractions of 3-disulfide intermediates.

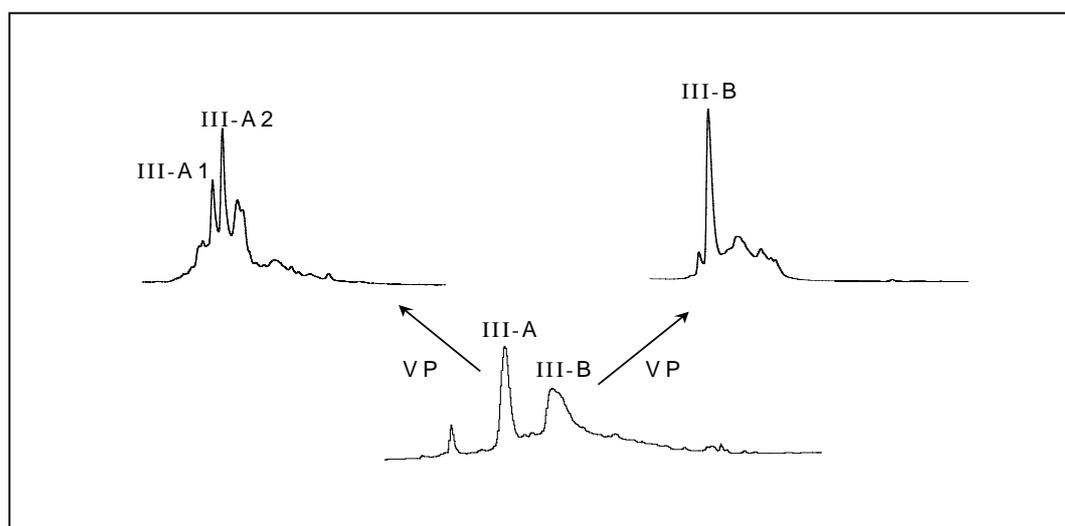
### III.C.3. Structural analysis of the major 3-disulfide kinetic traps.

Fractions III-A and III-B were isolated, freeze-dried, reacted with vinyl pyridine, and the derivatized samples were further purified by HPLC using the same conditions employed for the separation of folding intermediates. The chromatographic analysis shows that fraction III-B contains one predominant 3-disulfide species and III-A comprises two major 3-disulfide species (III-A1 and III-A2) (Fig. 4). III-A-1, III-A-2 and III-B were isolated, their molecular masses were determined by MALDI-TOF mass spectrometry, and aliquots were digested with thermolysin. The resultant thermolytic peptides were separated by HPLC (Fig. 5) and isolated for molecular mass analysis and Edman sequencing. Table I summarizes the  $M_r$  and sequence data of peptides containing disulfide bonds and free cysteines (modified with vinyl pyridine). Taken together, these data permit deducing the disulfide structures of the three major 3-disulfide intermediates (Fig. 6).



**Figure 3. Quantitative analysis of various disulfide species along the pathway of oxidative folding of LCI.** Folding was performed using the following conditions: (A) Tris-HCl buffer alone (control -). (B) Tris-HCl buffer in presence of 0.25 mM  $\beta$ -mercaptoethanol (control +). (C) Tris-HCl buffer containing GSSG (0.5mM). (D) Tris-HCl buffer containing a mixture of GSSG/GSH (0.5mM/1mM). 0 S-S, 1 S-S, 2 S-S, 3 S-S and 4 S-S represents the completely reduced, 1-disulfide, 2-disulfide, 3-disulfide and 4-disulfide species, respectively. The 4-disulfide species include the native and all scrambled isoforms. The recovery of native LCI is represented by bars shown at each graphic. Quantitative analysis of various disulfide species of LCI was based on the peak response of MALDI spectra (data not shown).

All three intermediates share a common thermolytic peptide (III-A1-3, III-A2-3, III-B-3) which was shown to have a molecular mass of 1653 and contain the native disulfide pairing Cys<sup>11</sup>-Cys<sup>34</sup>. Both III-A1 and III-A2 have free Cys<sup>22</sup> and Cys<sup>58</sup> derivatized with vinylpyridine, as unambiguously identified by Edman sequencing and mass spectrometry analyses. These coincidences, together with the initial difficulty of separating III-A1 and III-A2 intermediates, suggest that they have very similar isomeric structures. Analysis of the thermolytic peptides III-A1-5 and III-A2-5 further establishes that III-A1 and III-A2 adopt either the additional disulfide pairings of [Cys<sup>18</sup>-Cys<sup>43</sup>; Cys<sup>19</sup>-Cys<sup>62</sup>] and [Cys<sup>18</sup>-Cys<sup>62</sup>; Cys<sup>19</sup>-Cys<sup>43</sup>], or [Cys<sup>18</sup>-Cys<sup>62</sup>; Cys<sup>19</sup>-Cys<sup>43</sup>] and [Cys<sup>18</sup>-Cys<sup>43</sup>; Cys<sup>19</sup>-Cys<sup>62</sup>], respectively. Conventional methods of peptide mapping and Edman sequencing were unable to distinguish these two isomeric disulfide bond patterns in III-A1-5 or III-A2-5 because enzymatic cleavage between Cys<sup>18</sup>-Cys<sup>19</sup> was not possible. It is important to point out that [Cys<sup>18</sup>-Cys<sup>62</sup>; Cys<sup>19</sup>-Cys<sup>43</sup>] are native disulfide bonds whereas [Cys<sup>18</sup>-Cys<sup>43</sup>; Cys<sup>19</sup>-Cys<sup>62</sup>] are non-native disulfide bonds. Putative disulfide structures of III-A1 and III-A2 are presented in Fig. 6, in which one of them adopts three native disulfide bonds and the other contains one native and two non-native disulfide bonds. Since both species are detected along the folding pathway of LCI with comparable and significant concentration, this assignment should not affect the global understanding and interpretation of the data.



**Figure 4. HPLC isolation of 3-disulfide folding intermediates of LCI derivatized with vinyl pyridine.** Fractions III-A and III-B were isolated, freeze-dried and derivatized with vinyl pyridine. Derivatized III-A and III-B were further purified by HPLC using the same conditions described in the legend of Fig. 1. The results reveal that fraction III-A contains two major 3-disulfide intermediates (III-A1 and III-A2) and fraction III-B contains one predominant 3-disulfide intermediate (III-B).

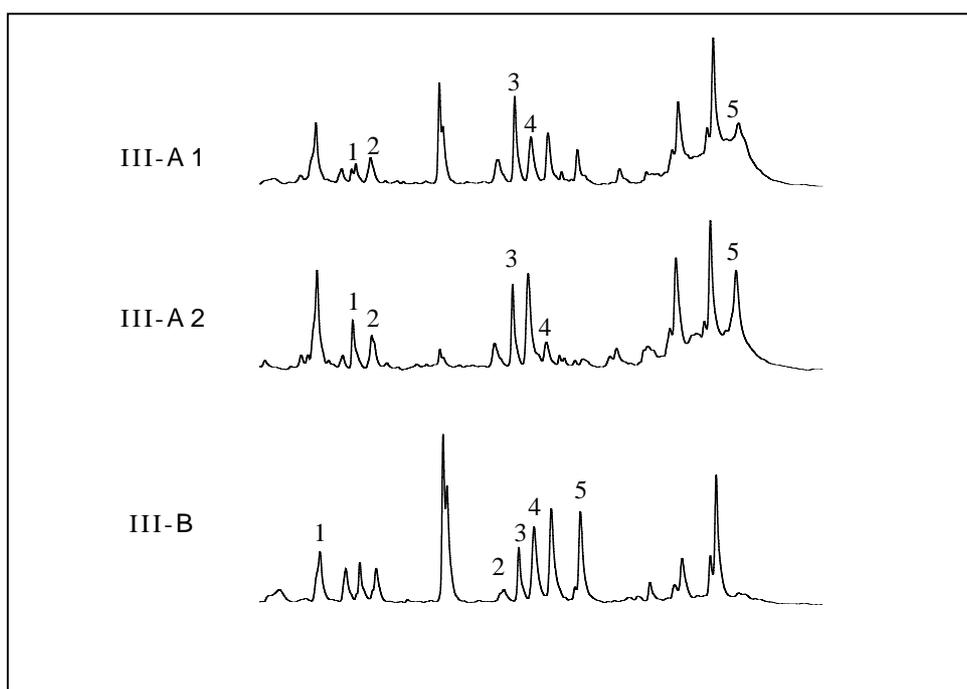
Assignment of disulfide pairings is unmistakable in the case of III-B. Aside from the native disulfide bond Cys<sup>11</sup>-Cys<sup>34</sup> verified by peptide III-B-3, Cys<sup>19</sup> and Cys<sup>43</sup> were shown to be free cysteines modified with vinyl pyridine. In addition, two native disulfide bonds Cys<sup>22</sup>-Cys<sup>58</sup> and Cys<sup>18</sup>-Cys<sup>62</sup> were identified by analysis of peptides III-B-1 and III-B-5 (Table 1). The disulfide structures III-B is also illustrated in Fig. 6, in which solid lines and dotted lines represent native and non-native disulfide pairings, respectively. All 3 isoforms share the native Cys<sup>11</sup>-Cys<sup>34</sup> native pairing. Among the nine disulfide bonds present in these three kinetic traps, seven are native and two are non-native.

## III.D. DISCUSSION

### III.D.1. Diversity of disulfide folding pathway.

The disulfide folding pathways of several model proteins have been elucidated by various laboratories (1,2,8,9,10-27). The results have shown that the mechanism of oxidative folding varies significantly from protein to protein. These differences are illustrated by: (a) the extent of the heterogeneity of folding intermediates; (b) the predominance of intermediates containing native disulfide bonds; and (c) the accumulation of fully oxidized scrambled isomers as intermediates. Among numerous proteins that have been investigated in detail, BPTI and hirudin represent two remarkable models with extreme folding characteristics, in spite of their comparable size and identical disulfide content (both contain 3 disulfide bonds). The original model of BPTI established by Creighton was based mainly on the analysis of folding intermediates trapped by iodoacetate and separated by ion-exchange chromatography (1,2,7). A subsequent study, using the method of acid trapping and HPLC analysis, revealed a somewhat different pattern of folding intermediates of BPTI (8). Specifically, intermediates with non-native disulfides are detected at much lower concentration than those found previously. Despite few discrepancies, the folding pathway of BPTI is characterized by the predominance of a limited number of folding intermediates that adopt native disulfide bonds and native like structures. Most importantly, 3-disulfide scrambled BPTI were absent and only 1- and 2-disulfide intermediates were observed during the folding of BPTI. By contrast, the folding mechanism of hirudin (10,11) differs from that of BPTI in three important aspects. (A) Folding intermediates of hirudin are far more heterogeneous. At least 30 fractions of 1- and 2-disulfide intermediates were identified along the folding pathways of hirudin. (B) Predominant folding intermediates adopting native disulfides are absent in this case. (C) Scrambled 3-disulfide isomers, not observed in the case of BPTI, were shown to serve as

folding intermediates of hirudin. Among the 14 possible scrambled isomers of hirudin, 11 have been identified as folding intermediates and their disulfide structures were determined (33,34). Moreover, accumulation of scrambled hirudin as folding intermediates can be greatly enhanced by the presence of oxidized glutathione or cystine (10). For instance, when folding of hirudin was performed in the presence of 2 mM of GSSG, a condition routinely applied for the oxidative folding of BPTI, the scrambled species accounted for more than 95% of the intermediates observed 5 min after the onset of the folding reaction. This accumulation took place even before any significant amount of native hirudin had appeared (10). For hirudin, folding apparently proceeds through an initial stage of non-specific disulfide formation (packing) followed by disulfide reshuffling (consolidation) of the heterogeneous scrambled intermediates to attain the native structure. The mechanism of oxidative folding of potato carboxypeptidase inhibitor (12,13) is nearly indistinguishable from that of hirudin.



**Figure 5. Peptide maps of thermolysin digested III-A1, III-A2 and III-B.** Thermolysin digestion products peptides were separated by HPLC using the conditions described in the legend of Fig. 1, except for the linear gradient used, which was 0-60% of solvent B in 60 minutes. Peptides were analyzed by Edman sequencing and MALDI mass spectrometry in order to identify the cysteine (vinyl pyridine derivatized) and disulfide containing peptides. The results are summarized in Table I. Data obtained were used to define the disulfide structure of the intermediates depicted in Fig. 6.

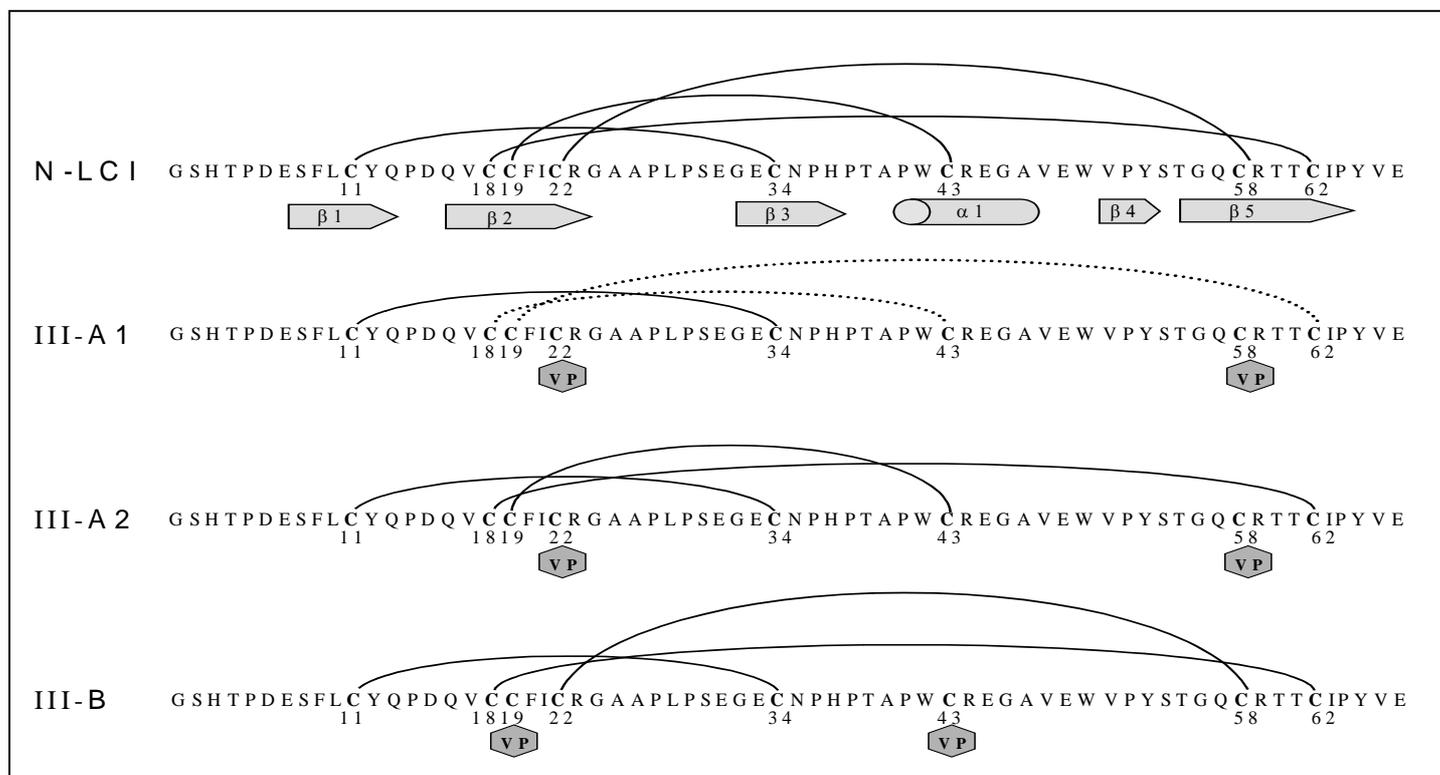
### III.D.2. The underlying chemistry of the diversity of disulfide folding pathway.

In order to search for the underlying causes of these diversities, we have previously conducted kinetic analysis of reductive unfolding of five different proteins, including BPTI, hirudin, potato carboxypeptidase inhibitor, tick anticoagulant peptide and epidermal growth factor (35). The experiment of reductive unfolding was intended to evaluate the relative stability and interdependency of disulfide bonds in the native protein. A striking correlation was observed between the mechanisms of reductive unfolding and oxidative folding. Those proteins with their native disulfide bonds reduced in a collective manner display both a high degree of heterogeneity of folding intermediates and the accumulation of scrambled isomers along the pathway of oxidative folding, as observed with hirudin. On the other hand, a sequential reduction of the native disulfide bonds is associated with the presence of predominant intermediates with native like structures, as is the case of BPTI. These studies lead to the conclusion that the pathway and mechanism of oxidative folding hinges critically upon the presence of localized stable domain structures. An outstanding example to illustrate this hypothesis is the mechanism of oxidative folding of  $\alpha$ -lactalbumin observed in the absence or the presence of calcium (27). The structure of  $\alpha$ -lactalbumin consists of an  $\alpha$ -helical domain that is primarily stabilized by hydrophobic force and a  $\beta$ -sheet domain that is stabilized upon binding to calcium (36). In the absence of calcium, the folding pathway of  $\alpha$ -lactalbumin resembles that of hirudin and potato carboxypeptidase inhibitor, which is depicted by heterogeneous 1-, 2- and 3-disulfide folding intermediates and the accumulation of 4-disulfide scrambled isomers along the folding pathway. Binding of calcium serves to generate a stable  $\beta$ -sheet domain. Consequently, the folding pathway of  $\alpha$ -lactalbumin in the presence of calcium involves only two predominant intermediates that adopt native-like structures of the  $\beta$ -sheet domain, thus bearing a resemblance to what is observed in the case of BPTI (7,8).

### III.D.3. The pathway of oxidative folding of LCI.

The pathway of oxidative folding of LCI elucidated here exhibits mechanisms observed in both BPTI and hirudin. The folding of LCI proceeds through a heterogeneous mixture of 1- and 2-disulfide intermediates, guiding the formation of two different populations of intermediates that act as kinetic traps. One population consists of a complex mixture of scrambled 4-disulfide species. The other consists of three predominant 3-disulfide isomers. Among them, two contain exclusively native disulfide bonds. Like in the case of hirudin, a sequential flow of heterogeneous 1- and 2-disulfide intermediates leads to the formation of scrambled (4-disulfide) species which are able to effectively convert to the native

structure only in the presence of thiol catalysts. On the other hand, and similarly to BPTI, intermediates adopting native disulfide bonds populate the late stages of folding and serve as rate limiting steps to attain the native structure. The pathway of oxidative folding of LCI thus further broadens the diversity spectrum of disulfide folding pathway.



**Figure 6. Disulfide structures of the major kinetic traps III-A1, III-A2 and III-B.** The disulfide pairing of native LCI (N-LCI) is shown at the top for comparison. Solid lines indicate native disulfide pairings. Dotted lines indicate non-native disulfide pairings. Free cysteines are labeled with vinyl pyridine (VP).

The effect of the GGS/GSH system in promoting the efficiency of oxidative folding has also been demonstrated in this study. Specifically, the kinetics of disulfide formation (i.e., the flow of intermediates) is accelerated by the addition of GSSG to the folding buffer and the kinetics of disulfide shuffling (scrambling) is promoted by the presence of GSH (or  $\beta$ -mercaptoethanol). In the case of LCI, addition of GSSG (0.5 mM) increases the rate of disulfide formation by about 12 fold. In the absence of any thiol reagent (“control – “experiment), scrambled 4-disulfide intermediates become trapped and unable to reshuffle their non-native disulfide bonds and convert to the native structure during the folding. However, the population and pattern of folding intermediates of LCI remain largely unchanged regardless of the presence of redox agents. This unique effect of GSSG/GSH redox system on the mechanism of oxidative folding has been likewise demonstrated with

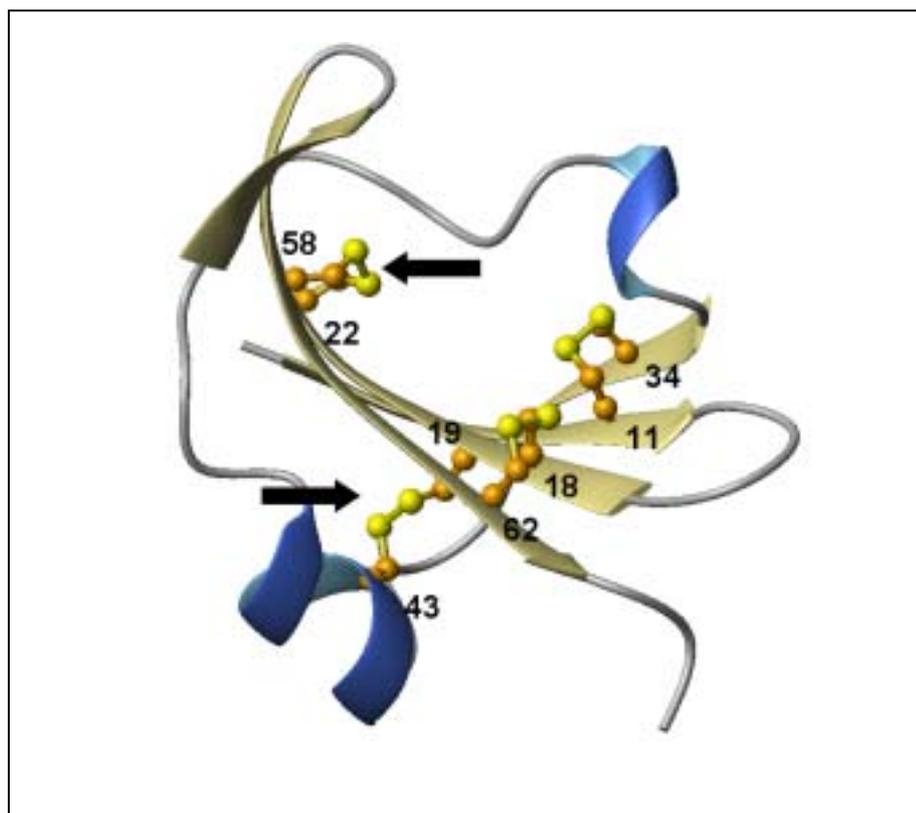
numerous disulfide proteins, including hirudin (10), potato carboxypeptidase inhibitor (12), tick anticoagulant peptide (14), and apo  $\alpha$ -lactalbumin (27).

**Table I. Amino acid sequence and molecular mass of cysteine and disulfide-containing peptides derived from thermolysin digestion of folding intermediates of LCI.** Peptides analyzed are those marked numerically in Fig.5. The underlined sequences are confirmed by Edman degradation. Vinyl pyridine (VP) modified peptides exhibit an additional molecular mass of 106.1.

Peptide	Sequence	Sequence position	Cys-Cys	M <sub>r</sub> found (exp.)
III-A1-1	<u>ICRG</u>	21-24	Cys <sup>22</sup> -VP	554.51 (552.65)
III-A1-2	<u>STGQCR</u>	54-59	Cys <sup>58</sup> -VP	756.16 (755.81)
III-A1-3	<u>LC</u> AAPLPSEGEENPHP	10-11 25-38	Cys <sup>11</sup> -Cys <sup>34</sup>	1652.62 (1650.85)
III-A1-4	<u>LCYQPDQ</u> AAPLPSEGEENPHP	10-16 25-38	Cys <sup>11</sup> -Cys <sup>34</sup>	2285.21 (2282.49)
III-A1-5	<u>YQPDQVCCFI</u> <u>TAPWC</u> <u>TTCIPYVE</u>	12-21 39-43 60-67	Cys <sup>18</sup> -Cys <sup>43</sup> Cys <sup>19</sup> -Cys <sup>62</sup>	2717.6 (2713.13)
III-A2-1	<u>ICRG</u>	21-24	Cys <sup>22</sup> -VP	554.51 (552.65)
III-A2-2	<u>YSTGQCR</u>	53-59	Cys <sup>58</sup> -VP	919.73 (918.98)
III-A2-3	<u>LC</u> AAPLPSEGEENPHP	10-11 25-38	Cys <sup>11</sup> -Cys <sup>34</sup>	1653.92 (1650.85)
III-A2-4	<u>GSHTPDESF</u>	1-9	No Cys	977.35 (975.97)
III-A2-5	<u>YQPDQVCCFI</u> <u>TAPWC</u> <u>TTCIPYVE</u>	12-21 39-43 60-67	Cys <sup>18</sup> -Cys <sup>62</sup> Cys <sup>19</sup> -Cys <sup>43</sup>	2717.6 (2713.13)
III-B-1	<u>ICRG</u> <u>STGQCR</u>	21-24 54-59	Cys <sup>22</sup> -Cys <sup>58</sup>	1097.9 (1096.26)
III-B-2	PHPTAPWCREGAV	36-48	Cys <sup>43</sup> -VP	1526.8 (1525.71)
III-B-3	<u>LC</u> AAPLPSEGEENPHP	10-11 25-38	Cys <sup>11</sup> -Cys <sup>34</sup>	1653.1 (1650.85)
III-B-4	<u>LCYQPDQ</u> AAPLPSEGEENPHP	10-16 25-38	Cys <sup>11</sup> -Cys <sup>34</sup>	2285.21 (2282.49)
III-B-5	<u>VCC</u> <u>TTCIP</u>	17-19 60-64	Cys <sup>19</sup> -VP Cys <sup>18</sup> -Cys <sup>62</sup>	962.5 (960.16)

### III.D.4. Molecular basis for the stability of 3-disulfide kinetic traps of LCI.

The structure of LCI consists of a five-stranded anti-parallel  $\beta$ -sheet and one short  $\alpha$ -helix (29), as seen in Fig. 7 and schematically displayed in Fig. 6. All four native disulfide bridges are located within regular secondary structure elements. The III-B intermediate has two free cysteines, Cys<sup>19</sup> and Cys<sup>43</sup>, which connect  $\alpha$ -1 helix and  $\beta$ -2 sheet in the native structure. The absence of this disulfide bridge suggests that the structure of III-B would display a disconnected  $\alpha$ -1 helical section and a main core of five anti-parallel  $\beta$ -sheets stabilized by the remaining three native disulfide bonds. The stability of the  $\beta$ -sheets core domain may account for the sluggish formation of the last native disulfide (Cys<sup>19</sup>-Cys<sup>43</sup>) and allow III-B to serve as one of the major kinetic traps. In the case of III-A1 and III-A2, the absence of disulfide Cys<sup>22</sup>-Cys<sup>58</sup> implies that the linkage between the C-terminal end of  $\beta$ -2 and the N-terminal end of  $\beta$ -5 and the subsequent stabilization of the complete  $\beta$ -sheet core domain represents a rate limiting step in reaching the native structure. All three kinetic traps, despite their native disulfide bonds and presumably native like structure, must somehow twist their conformation and overcome an energy barrier in order to form the native LCI.



**Figure 7. Schematic view of the native structure of leech carboxypeptidase inhibitor.** The four native disulfide bonds, Cys<sup>11</sup>-Cys<sup>34</sup>, Cys<sup>18</sup>-Cys<sup>62</sup>, Cys<sup>19</sup>-Cys<sup>43</sup> and Cys<sup>22</sup>-Cys<sup>58</sup> are numbered. The arrows indicate the two disulfide bonds, Cys<sup>22</sup>-Cys<sup>58</sup> and Cys<sup>19</sup>-Cys<sup>43</sup>, absent in intermediates III-A and III-B, respectively.

The structural properties of the intermediates III-A1 and III-A2 are reminiscent of those of two major isoforms described for insulin-like factor 1 (IGF-1) (37-40). Both in IGF and LCI, the two isoforms are products of isomerization *via* two adjacent cysteines (Figs. 6 and 7). Under physiological pH, the native IGF (Cys<sup>6</sup>-Cys<sup>48</sup>, Cys<sup>18</sup>-Cys<sup>61</sup>, Cys<sup>47</sup>-Cys<sup>52</sup>) exists in equilibrium with a scrambled species, designated as IGF-s (swap) (Cys<sup>6</sup>-Cys<sup>47</sup>, Cys<sup>18</sup>-Cys<sup>61</sup>, Cys<sup>48</sup>-Cys<sup>52</sup>) with a molar ratio of 60:40 (38,40). The same molar ratio is found between IIIA-2 and III-A1 of LCI (Fig. 4). If the structural property of IGF-1 constitutes a valid reference, we have indeed correctly assigned the disulfide structures of III-A2 and III-A1, in which the predominant species (III-A2) adopts all native disulfide bonds.

### III.E. REFERENCES

1. Creighton, T. E. (1978) *Prog Biophys. Mol. Biol.* 33, 231-297.
2. Creighton, T. E., and Goldenberg, D. P. (1984) *J. Mol. Biol.* 179, 497-526.
3. Wedemeyer, W.J., Welker, E., Narayan, M., and Scheraga, H.A. (2000) *Biochemistry.* 39, 4207-4216.
4. Frand, A.R., Cuozzo, J.W., Kaiser, C.A. (2000) *Trends. Cell. Biol.* 10, 203-210.
5. Woycechowsky, K.J., and Raines, R.T. (2000) *Curr. Opin. Chem. Biol.* 4, 533-539.
6. Creighton, T. E. (1986) *Methods Enzymol.* 131, 83-106.
7. Creighton, T. E. (1990) *Biochem. J.* 270, 1-16.
8. Weissman, J. S., & Kim, P. S. (1991) *Science.* 253, 1386-1393.
9. Goldengerg, D. P. (1992) *TIBS.* 17, 247-261.
10. Chang, J.-Y. (1994) *Biochem. J.* 300, 643-650.
11. Chatrenet, B., and Chang, J.-Y. (1993) *J. Biol. Chem.* 268, 20988-20996.
12. Chang, J.-Y., Cannals, F., Schindler, P., Querol, E., and Aviles, F. X. (1994) *J. Biol. Chem.* 269, 22087-22094.
13. Venhudova, G., Canals, F., Querol, E., and Aviles FX. (2001) *J. Biol. Chem.* 276, 11683-11690.
14. Chang, J.-Y. (1996) *Biochemistry.* 35, 11702-11709.
15. Chang, J.-Y., and Ballatore, A. (2000) *J. Prot. Chem.* 19, 299-310.
16. Wu, J., Yang, Y., and Watson, J. T. (1998) *Protein Sci.* 7, 1-17-1028.
17. Chang, J.-Y., Li, L. and Lai, P.-S. (2001) *J. Biol. Chem.* 276, 4845-4852.
18. Miller, J.A., Narhi, L.O., Hua, Q.X., Rosenfeld, R., Arakawa, T., Rohde, M., Prestrelski, S., Lauren, S., Stoney, K. S., Tsai, L., and Weiss, M.A. (1993) *Biochemistry.* 32, 5203-5213.

19. Yang, Y., Wu, J., Watson J.T. (1999) *J. Biol. Chem.* 274, 37598-37604.
20. Scheraga, H. A., Konishi, Y., and Ooi, T. (1984) *Adv. Biophys.* 18, 21-41.
21. Scheraga, H. A. , Wedemeyer, W.J., and Welker, E. (2001) *Methods Enzymol.* 341, 189-221.
22. Welker, E., Narayan, M., Volles, M.J., and Scheraga, H.A. (1999) *FEBS Lett.* 460, 477-479.
23. Rothwarf, D. M., Li, Y.-J., and Scheraga, H.A. (1998) *Biochemistry.* 37, 3760-3766.
24. Rao, K. R., and Brew, K. (1989) *Biochem. Biophys. Res. Commun.* 163, 1390-1396.
25. Ewbank, J. J., and Creighton, T. E. (1993) *Biochemistry.* 32, 3677-3693.
26. Ewbank, J. J., and Creighton, T. E. (1993) *Biochemistry.* 32, 3694-3707.
27. Chang, J.-Y., and Li, L. (2002) *Biochemistry.* 41, 8405-8413.
28. Reverter D., Vendrell J., Canals F., Hosrtmann J., Aviles F.X., Fritz H., Sommerhoff C. P. (1998) *J. Biol. Chem.* 273, 32927-32933.
29. Reverter D., Fernandez-Catalan C., Baumgartner R., Pfander R., Huber R., Bode W., Vendrell J., Holak T.A., Aviles F.X. (2000) *Nature Struct. Biol.* 7, 322-328.
30. Molina, M. A., Avilés, F. X., and Querol, E. (1992) *Gene* 116, 129-138.
31. Molina, M. A., Marino, C., Oliva, B., Avilés, F. X., and Querol, E. (1994) *J. Biol. Chem.* 269, 21467-21472.
32. Salamanca S., Villegas V., Vendrell J., Li L., Aviles F.X., Chang, J.-Y. (2002) *J. Biol. Chem.* 277, 17538-17543.
33. Chang, J.-Y., Schindler, P., Chatrenet, B. (1995) *J. Biol. Chem.* 270, 11992-11997.
34. Chang, J.-Y. (1995) *J. Biol. Chem.* 270, 25661-25666.
35. Chang, J.-Y., Li, L., and Bulychev, A. (2000) *J. Biol. Chem.* 275, 8287-8289.
36. Permyakov, E. A., and Berliner L.J. (2000) *FEBS Lett.* 473, 269-274.
37. Hober, S., Forsberg, G., Palm, G., Hartmanis, M., and Nilsson, B. (1992) *Biochemistry.* 31, 1749-1756.
38. Miller, J. A., Narhi, L.O., Hua, Q.X., Rosenfeld, R., Arakawa, T., Rohde, M., Prestrelski, S., Lauren, S., Stoney, K. S., Tsai, L., and Weiss, M.A. (1993) *Biochemistry.* 32, 5203-5213.
39. Narhi, L.O., Hua, Q.X., Arakawa, T., Fox, G.M., Tsai, L., Rosenfeld, R., Holst, P., Miller, J.A., and Weiss, M.A. (1993) *Biochemistry.* 32, 5214-5221.
40. Chang, J.-Y., Maerki, W., and Lai, P. S. (1999) *Protein Science.* 8, 1463-1468.