



UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MEXICO

INSTITUTO DE INVESTIGACIONES BIOMEDICAS

ENSAMBLE DEL COMPLEJO DE ATAQUE A  
MEMBRANA DEL COMPLEMENTO:  
UN ACERCAMIENTO EVOLUTIVO.

T E S I S  
QUE PARA OBTENER EL TITULO DE:  
LICENCIADO EN INVESTIGACION  
BIOMEDICA BASICA  
P R E S E N T A :  
MARIANA MONDRAGON PALOMINO

DIRECTOR DE TESIS: DR. JUAN PEDRO LACLETTE SAN ROMAN.

CIUDAD UNIVERSITARIA.

OCTUBRE 1998.

TESIS CON  
FALLA DE ORIGEN

266415



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

**INSTITUTO DE INVESTIGACIONES BIOMÉDICAS**

APARTADO POSTAL 70228  
CIUDAD UNIVERSITARIA  
04510 MÉXICO, D.F.

Octubre 1ro., 1998.

DR. ALEJANDRO CRAVIOTO QUINTANA  
DIRECTOR  
FACULTAD DE MEDICINA  
PRESENTE.

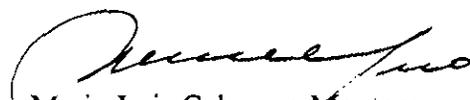
Por medio de la presente le informo que la tesis "Ensamble del complejo de ataque a membrana del complemento: Un acercamiento evolutivo", que presenta la alumna MARIANA MONDRAGON PALOMINO con número de cta. 9452983-4, reúne ampliamente los requisitos de una tesis de la Licenciatura en Investigación Biomédica Básica. Por lo tanto, los suscritos miembros del Jurado damos nuestro voto aprobatorio para que la tesis sea presentada en examen oral.

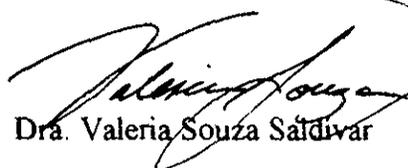
Atentamente,

  
Dra. Ma. del Carmen Gómez Eichelmann

  
Dr. Luis Enrique Eguiarte Fruns

  
Dr. Juan Pedro Laclette San Román

  
Dr. Mario Luis Calcagno Montans

  
Dra. Valeria Souza Saldivar



## Dedicatoria

Para Marta, Eduardo y Octavio, alas, raíces y sangre a la vez. Por la devota atención y cariño de toda la vida, por despertarme y hacerme descansar, por ponerme en el camino diario y ser un ejemplo de trabajo tenaz. Para ustedes que son en el sentido mas amplio y generoso el calor, la luz y el alimento.

## Agradecimientos

Esta tesis es mas que el resultado de mi último año de licenciatura, pues en ella se conjuntan muchas cosas que aprendí a lo largo de estos últimos cinco años, por eso quiero manifestar mi agradecimiento con todas las personas que han formado parte de mi estancia en Biomédicas.

Gracias Carmen, Enrique y Juan Pedro por pacientemente haberse ocupado de enseñarme a pensar y a usar las manos. De la manera mas noble y generosa han compartido su tiempo, conocimientos, recursos, entusiasmo e imaginación. Para mi resulta muy significativo que me brinden su confianza, me hayan permitido formar parte de sus equipos, conocer de sus vidas y trabajo. Gracias Daniel, Luis y Valeria por haberme adentrado en la Evolución Molecular y haberme estimulado a quedarme en ella. Gracias Tila y Anne por haber prodigado sustrato conceptual y enzima computacional para el proyecto del Complemento, su ayuda ha sido imprescindible. Gracias Jaime por haberme ayudado a perderle el miedo a las computadoras y a darles uso en las ciencias biológicas. Gracias Dr. Calcagno por su cuidadosa revisión de este manuscrito. Ha sido un placer y un privilegio trabajar y aprender de todos ustedes. No solo me han enseñado de ciencia, pues al observarlos absortos y diligentes en sus trabajos, a veces tristes o indignados y frecuentemente de buen humor, me han enseñado sin saberlo, sobre las motivaciones y características de la gente honesta y dedicada. Las cualidades de este manuscrito se deben en gran parte a lo que ustedes han contribuido directa e indirectamente como mis maestros y sinodales. Dificilmente puedo imaginar mejores guías. Mil gracias a Julieta Rubio, Saúl Cano -¿Carro?-, Auster Valderrama, Carlos Pompeyo, Dra. Yankelevich, Dr. León Cázares, Catherine, Eugenia y Claudia que nos guiaron durante el primera año, hicieron todo lo posible para que adquiriésemos nuevos y mejores hábitos de aprendizaje y procuraron proveernos con las herramientas fundamentales.

Gracias Luis Mendoza por haber sido un brillante y genuino compañero, por todo lo que aprendimos, por ser mi habito mas entrañable y contribuir a que la amistad quede intacta. Gracias Clelia, Alfonso -diméricos y puros- y Dayanica por todo su apoyo e inteligente amistad, que el tiempo los conserve igual de locos y tenaces. Gracias al ejército de dedicados y generosos compañeros de laboratorio, sin su ayuda y enseñanzas el camino habría sido muy áspero...allí van los nombres que no se olvidan: Jesús, Don Mike, Rafael, Tere, Judith, Guayabito, Juan e Irma, que entre mercapto y bacterias caprichosas me enseñaron sobre carteles, seminarios, pasteles esponjosos, a torear albures y cabulear al prójimo. Claudia Garay, Martha, Claudia Hallal, Irene, Rocío, Claudine, Manuel; domadoras del Western, los receptores y las células inmortales, precursoras del desayuno-seminario y el vegetarianismo a ultranza. Martín, Paty, Pepe, Laura Márquez, Carlos, Laura Vargas, Pavel, Dairo, Prima y Pedro. Asociación de Diversidades Locas con quienes he compartido desde puerquitos experimentales hasta teléfonos hiperactivos y una mudanza penosa. Agradezco que hayan tenido la prodigiosa paciencia de rescatar mis pertenencias, poner atención en mis seminarios y sobrellevar estoicamente el monopolio de las computadoras. Gracias también a Tata, Yazmin, Sebas, Alejandro, Daniel y Etelvina por haberme condimentado la licenciatura de maneras diversas e inesperadas.

También quiero expresar mi agradecimiento a todos ustedes que hicieron mi vida mas fácil en Biomédicas, suavizaron las complejidades administrativas y como era de esperarse cumplieron con sus trabajos de manera constante y capaz. Gracias Dr. Lamoyi por las becas, el congreso en Xalapa y por estar al tanto de nuestro apendizaje. Gracias Blanca Huerta por su ayuda y paciencia para con mis frecuentes visitas. Agradezco también al licenciado Brito, Alfonso, Guille y Magda que agilizaron las búsquedas bibliográficas e ignoraron mis frecuentes retardos. Gracias Don José por todos los impecables trabajos fotográficos. No se olvidan los proyectos que pudieron realizarse y las muchas veces que salí de dificultades gracias a la ayuda de los licenciados Francisco Puente, Moisés Paniagua, los laboriosos equipos de intendencia, vigilancia y las secretarias de los distintos departamentos donde estuve. Finalmente deseo hacer patente mi reconocimiento a la Fundación UNAM que brindó el apoyo económico para llevar a cabo muchos de mis planes.

# ÍNDICE

RESUMEN	i
<b>I. INTRODUCCIÓN</b>	
1. Funciones del sistema del complemento	3
2. Vías de activación del complemento	
2a. Vía alterna	5
2b. Vía clásica	9
2c. Vía dependiente de C1	11
2d. Vía de las lectinas	12
3. Etapa terminal de la cascada	
3a. Elementos de la fase terminal	14
3b. Características de las proteínas C6-C9	18
3c. Formación del complejo terminal	25
3d. Características del TCC	26
3e. Regulación de la vía terminal	27
3f. Efectos del TCC	29
<b>II. ANTECEDENTES DEL PROYECTO</b>	
1. Evolución molecular de la vía terminal	35
<b>III. OBJETIVO</b>	42
<b>IV. MATERIAL Y MÉTODOS</b>	
1. Secuencias	43
2. Alineación	43
3. Proceso de inferencia filogenética	45
4. Prueba de la tasa relativa	46
5. Análisis de Hillis-Huelsénbeck	46

## V. RESULTADOS

1. Características de la alineación generada	47
2. Análisis filogenético por parsimonia	47
3. Prueba de la tasa relativa	50
4. Análisis por método de distancia	56
5. Análisis de Hillis-Huelsenbeck	56

## VI. DISCUSIÓN

60

## VII ANEXOS

### Anexo 1 . Otras funciones del sistema del complemento.

- Complemento y solubilización de complejos inmunes	72
- Papel del complemento en el desarrollo de la memoria inmune	72
- Sistema del complemento y coagulación	73

### Anexo 2 . Evolución de las vías de activación del complemento

76

### Anexo 3 . Conceptos básicos sobre Evolución Molecular

- ¿Qué es Evolución Molecular?	83
- Supuestos fundamentales de la filogenética molecular.	84
- Alineación de secuencias	85

#### 1. Métodos de Inferencia filogenética

- Distancia y Parsimonia	87
- Máxima Parsimonia	87
- Métodos de Distancia.	90
- Prueba de "Bootstrap"	92
- Enraizado y grupos externos	93

#### 2. Relojes moleculares

94

Anexo 4. Alineación de secuencias empleadas en este estudio	97
Anexo 5. Filogenia de mamíferos.	102
VIII. BIBLIOGRAFÍA	103

## RESUMEN

Los resultados de los análisis de inferencia filogenética empleando las secuencias de aminoácidos de las perforinas y C6-C9 de diversas especies, sugieren que C6 y C7 fueron las primeras proteínas en aparecer durante la evolución de la vía terminal del complemento. Posteriormente, por duplicación del ancestro común y pérdida de módulos surgieron C8 y C9, moléculas de menor complejidad modular. También se ha llegado a esta conclusión mediante el análisis de las fases de intrones/exones de C6-C9 humanos, sin embargo el acercamiento empleado en este proyecto de tesis ofrece las ventajas de que mediante un método computacional estándar de inferencia filogenética se analizó toda la información disponible para estas proteínas y el orden de aparición se basó en las secuencias de las perforinas.

Los análisis de inferencia filogenética aquí presentados sugieren que las proteínas del TCC tuvieron un origen monofilético, lo que confirma que son proteínas cuyas características compartidas se deben a un origen evolutivo común. En este contexto las características funcionales compartidas por C9 y perforina son el resultado de un fenómeno de paralelismo, mientras que las similitudes a nivel de secuencias están dadas por que ambas proteínas tienen un ancestro común.

El orden de aparición de las moléculas del TCC correlaciona con las funciones de señalamiento que tiene este complejo y sus formas intermedias C5b-C7, C5b-C8, lo cual subraya el hecho de que la vía terminal del complemento tiene un papel relevante en la respuesta inmune por su participación en la traducción de señales más que por su capacidad para lisis ciertos tipos de células. Posiblemente las condiciones que favorecieron la aparición de moléculas formadoras de canales con funciones de señalamiento más que de lisis, están relacionadas con el desarrollo de mecanismos de evasión de las propiedades citolíticas del complemento por organismos patógenos.

La prueba de la tasa relativa además de haber proporcionado información sobre la variabilidad de las tasas de cambio de las proteínas C6-C9 con respecto a las perforinas, ha subrayado tendencias de cambio que podrían tener relación con los contactos intermoleculares dentro del TCC o de la interacción de sus componentes con la membrana. En base a los resultados obtenidos en este proyecto es posible descartar el orden de aparición de estos genes fundamentado en los módulos compartidos por C6-C9.

## I. INTRODUCCIÓN

La cascada del complemento es un sistema de defensa presente en el suero de los mamíferos, que tiene una vía de activación asociada a la inmunidad humoral y otras vías independientes. El complemento está formado por más de 30 proteínas séricas y membranales, de las cuales doce participan directamente en las vías que constituyen la cascada, mientras las proteínas restantes tienen funciones de regulación. Las características fisicoquímicas y funcionales de las principales proteínas del sistema del complemento se enlistan en la Tabla I. Más del 90% de los componentes del complemento se sintetizan en el hígado y circulan en la sangre constituyendo entre 5-10 % de la proteína sérica total. La fuente más importante de complemento extrahepático son los monocitos y los macrófagos. Ambas células producen las proteínas tempranas de la cascada del complemento, tanto de las vías clásica como alterna hasta C5, así como muchas proteínas que regulan el sistema (Liszewski, 1993; Hughes-Jones, 1986).

La activación del complemento inicia una cascada de reacciones bioquímicas donde cada componente activa al que le precede. Este mecanismo permite que una pequeña señal inicial se amplifique progresivamente y estimule la formación de cantidades mayores de productos activos.

El sistema del complemento es activado por cuatro vías, la alterna, la clásica, la de lectinas y la dependiente de C1. Estas cuatro vías convergen en una etapa de marcado mediada por C3, molécula que amplifica la cascada y se une covalentemente a la superficies activadoras haciendo que estas sean reconocidas por receptores específicos ubicados en la superficie de células fagocíticas. Esta etapa de marcado finaliza con la formación de una enzima capaz de cortar C5, el producto de esta reacción, C5b es el primer componente de una vía terminal común donde se forma el complejo de ataque a membrana, también conocido como complejo terminal del complemento (TCC). El TCC es un poro formado por las moléculas C5b, C6, C7, C8 y varios monómeros de C9. Este complejo se inserta en membranas y tiene funciones en la traducción de señales para las células del individuo, además de ocasionar la lisis de algunas partículas ajenas como bacterias, eritrocitos y virus.

## 1. Funciones del sistema del complemento

El sistema del complemento está implicado en diversas etapas de la respuesta inmune. En cualquier sitio del cuerpo donde se ejerza daño tisular, ya sea de origen traumático, químico o infeccioso, se desarrolla una reacción inflamatoria. Esta se caracteriza por vasodilatación localizada y permeabilidad capilar aumentada, que resulta en la entrada de líquido y migración de células fagocíticas al área. En el sitio inflamatorio se presenta tumefacción debida a la entrada de fluido a los tejidos, enrojecimiento por el flujo sanguíneo capilar exacerbado y aumento de la temperatura local ocasionado por factores pirógenos liberados en la zona. La activación de complemento en la zona inflamada ocasionada por un agente infeccioso es uno de los mediadores de los complejos eventos de inflamación y fagocitosis. Dicha activación ocurre en la superficie del microorganismo invasor y en el tejido dañado, promoviendo la liberación de fragmentos y complejos producidos durante la activación del complemento que movilizan células fagocíticas y las dirigen al sitio de daño tisular. En este sitio otros productos del complemento ayudan a la neutralización de organismos invasores, ya sea directamente al causar lisis o indirectamente al estimular la liberación de metabolitos tóxicos por fagocitos (Gallin, 1993).

Se puede producir una severa destrucción tisular cuando IgG o IgM se unen a células propias o a partículas externas, ya que el reconocimiento de antígenos en eritrocitos circulantes, plaquetas, leucocitos o bacterias puede ocasionar la activación del complemento, el depósito de fragmentos de C3b en la superficie de estas células y la producción de C3a, C4a y C5a, moléculas conocidas como anafilatoxinas (Gallin, 1993). Durante la activación del complemento los extremos amino terminal de las cadenas  $\alpha$  de C3, C4 y C5 son cortadas en un enlace arginil-X-péptido, lo que origina péptidos de 70-80 aminoácidos designados C3a, C4a y C5a. Estos péptidos tienen propiedades anafilácticas, aunque difieren en su potencia relativa; C5a es 200 y 3000 veces más potente que C3a y C4a, respectivamente (Frank, 1991). El control de la actividad de las anafilatoxinas es ejercido por la enzima plasmática Carboxipeptidasa-N que elimina el residuo arginina carboxilo terminal de cada péptido y de este modo los inactiva. C3a, C4a y C5a se unen a receptores en la superficie de células cebadas y basófilos e inician la liberación de aminas vasoactivas contenidas en gránulos intracelulares. La histamina y la serotonina liberadas en este proceso son factores proinflamatorios que causan la contracción del músculo liso y

aumentan la permeabilidad de capilares. C3a, C4a y C5a son también importantes factores quimiotácticos que se difunden del sitio de infección o destrucción tisular donde se activó el complemento y forman un gradiente de concentración que atrae a los neutrófilos. La potencia quimiotáctica relativa de estos factores es similar a la proinflamatoria. Solo se han caracterizado los receptores para las formas activa e inactiva de C5a. Estos receptores se encuentran implicados en anafilaxis, median la respuesta celular a C5a y son responsables de la internalización y degradación de la anafilatoxina dentro de las células; son principalmente expresados por neutrófilos, monocitos y macrófagos.

La activación del complemento en el sitio de inflamación desencadena la inserción del TCC en la membrana de organismos invasores y de algunas células del huésped. Debido a que este complejo causa lisis de algunas células bacterianas, así como de eritrocitos en ensayos de fijación de complemento, durante un largo tiempo se ha considerado que su principal función es de ataque; de ahí su denominación marcial como "complejo de ataque a membranas". Sin embargo, la evidencia más reciente indica que este complejo también induce la producción y liberación de una gran variedad de mediadores inflamatorios. La evidencia a favor de la participación del TCC en la transducción de señales es creciente, por lo que en el presente trabajo de tesis se prefiere utilizar la denominación de "complejo terminal del complemento" (TCC).

El proceso de limpieza posterior a la fase aguda de la reacción inflamatoria es mediado por las moléculas C3b y C4b, que se generan durante la activación del complemento en la superficie activadora y funcionan como opsoninas\* cuando cubren al organismo o tejido dañado y facilitan su unión a receptores en la superficie de distintas células fagocíticas. La molécula C3b y el producto de degradación iC3b tienen actividad de opsoninas siendo la primera más potente por su alta afinidad por los receptores fagocíticos.

Hasta el momento se han descubierto cuatro receptores para fragmentos del complemento. Estos receptores se les conoce como Receptores de Complemento tipo 1-4 (CR1-4). Difieren en sus ligandos específicos y se encuentran en la superficie de las células fagocíticas. CR1 se describió originalmente en la membrana de eritrocitos y posteriormente en neutrófilos,

---

\* Opsonización es el proceso por el cual un organismo o partícula extraña es preparado para facilitar su ingestión por fagocitos. La opsonización implica el recubrimiento de las partículas blanco con proteínas que la célula fagocítica reconoce y une a través de receptores específicos de membrana.

monocitos/macrófagos, linfocitos B/T, eosinófilos, células NK entre otros tipos. Este receptor de complemento es el más ampliamente distribuido en distintos tipos celulares. El número de receptores CR1 en un eritrocito es grande y constante a lo largo del tiempo, sin embargo, en otras células se encuentra normalmente en cantidad reducida y puede elevarse hasta diez veces cuando las células son estimuladas por factores quimiotácticos. Su ligando primario es C3b y con menor afinidad iC3b y C4b. Este receptor también actúa como cofactor para que el Factor I degrade a C3b (Morgan, 1990b).

CR2 está presente solo en linfocitos B, donde une partículas con los fragmentos C3d, C3dg, iC3b y con una afinidad muy reducida, a C3b. Los linfocitos B no son fagocíticos, por lo tanto este receptor no está implicado en la remoción de partículas extrañas. Su papel aún no es claro pero se le ha relacionado con el aumento de la citotoxicidad celular dependiente de anticuerpos y la proliferación linfocitaria. CR2 es también la molécula que permite la entrada del virus Epstein-Barr a los linfocitos B (Liszewski, 1993).

CR3 únicamente se encuentra en neutrófilos, macrófagos y linfocitos implicados en citotoxicidad celular dependiente de anticuerpos. Tiene como ligando principal a iC3b y su presencia en fagocitos es importante para la remoción normal de partículas extrañas. Los individuos deficientes en este receptor tienen una capacidad fagocítica disminuida y sufren de infecciones bacterianas recurrentes (Liszewski, 1993).

CR4 es estructuralmente similar a CR3 y está presente en los mismos tipos celulares y tiene la misma especificidad de ligandos, también está implicado en la remoción de partículas opsonizadas (Liszewski, 1993).

En el Anexo I se ofrece una descripción detallada de otras funciones en que participa la vía del complemento.

## 2. Vías de activación del complemento.

### 2a. La vía alterna

La vía alterna del complemento representa un sistema natural de defensa contra agentes infecciosos, ya que en ausencia de anticuerpos puede neutralizar una variedad de patógenos

monocitos/macrófagos, linfocitos B/T, eosinófilos, células NK entre otros tipos. Este receptor de complemento es el más ampliamente distribuido en distintos tipos celulares. El número de receptores CR1 en un eritrocito es grande y constante a lo largo del tiempo, sin embargo, en otras células se encuentra normalmente en cantidad reducida y puede elevarse hasta diez veces cuando las células son estimuladas por factores quimiotácticos. Su ligando primario es C3b y con menor afinidad iC3b y C4b. Este receptor también actúa como cofactor para que el Factor I degrade a C3b (Morgan, 1990b).

CR2 está presente solo en linfocitos B, donde une partículas con los fragmentos C3d, C3dg, iC3b y con una afinidad muy reducida, a C3b. Los linfocitos B no son fagocíticos, por lo tanto este receptor no está implicado en la remoción de partículas extrañas. Su papel aún no es claro pero se le ha relacionado con el aumento de la citotoxicidad celular dependiente de anticuerpos y la proliferación linfocitaria. CR2 es también la molécula que permite la entrada del virus Epstein-Barr a los linfocitos B (Liszewski, 1993).

CR3 únicamente se encuentra en neutrófilos, macrófagos y linfocitos implicados en citotoxicidad celular dependiente de anticuerpos. Tiene como ligando principal a iC3b y su presencia en fagocitos es importante para la remoción normal de partículas extrañas. Los individuos deficientes en este receptor tienen una capacidad fagocítica disminuida y sufren de infecciones bacterianas recurrentes (Liszewski, 1993).

CR4 es estructuralmente similar a CR3 y está presente en los mismos tipos celulares y tiene la misma especificidad de ligandos, también está implicado en la remoción de partículas opsonizadas (Liszewski, 1993).

En el Anexo I se ofrece una descripción detallada de otras funciones en que participa la vía del complemento.

## 2. Vías de activación del complemento.

### 2a. La vía alterna

La vía alterna del complemento representa un sistema natural de defensa contra agentes infecciosos, ya que en ausencia de anticuerpos puede neutralizar una variedad de patógenos

potenciales. Esta formada por los Factores B, D, H, I, properdina (P) y C3\* , los cinco primeros son elementos de amplificación y regulación, mientras C3 juega una papel central en las cuatro vías del complemento. (Fig 1)

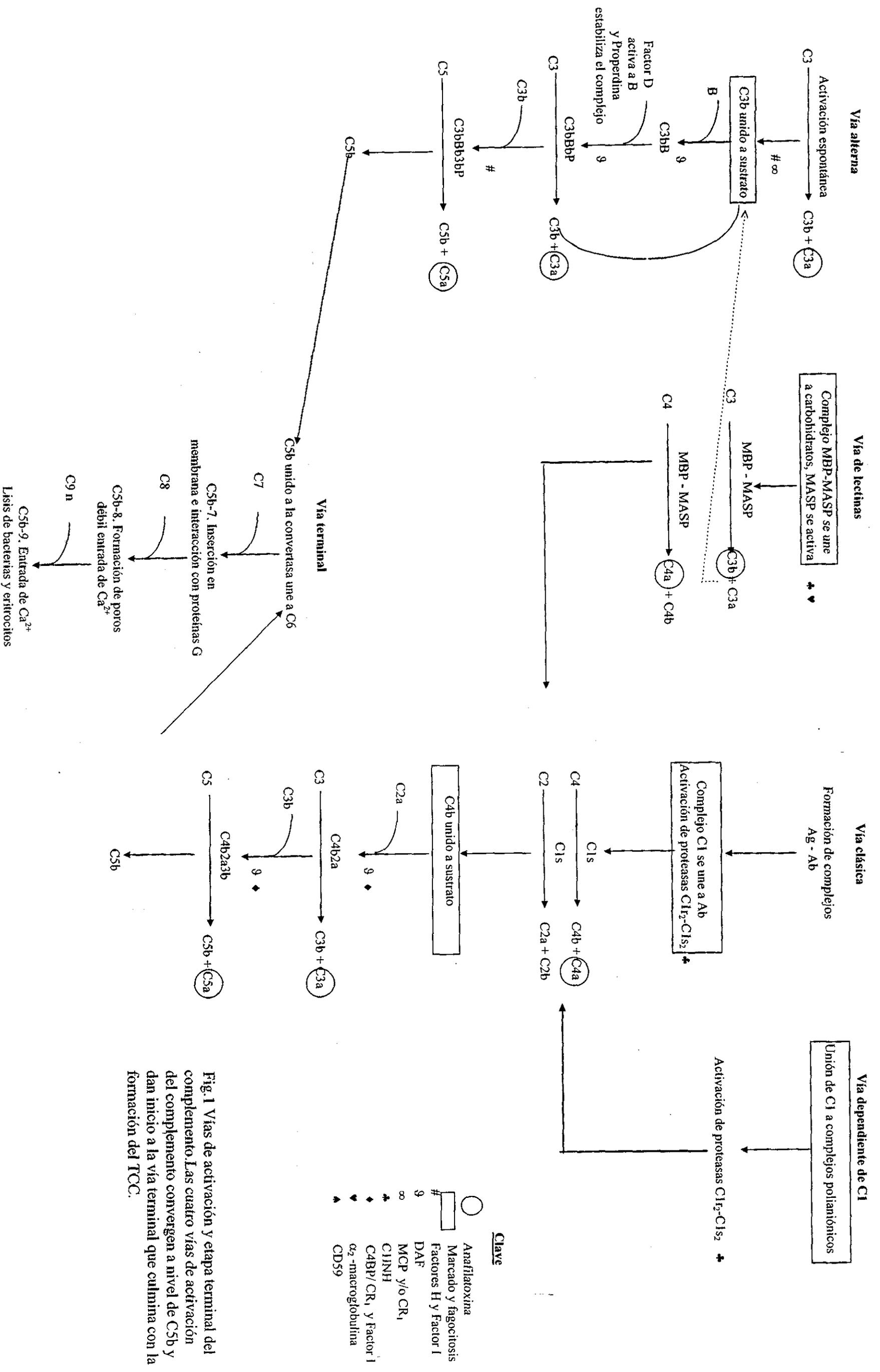
La molécula C3 inicia la reacción en cascada de la vía alterna. Esta glicoproteína formada por dos cadenas unidas por un enlace disulfuro, así como las proteínas del complemento C4, y  $\alpha_2$ -macroglobulina pertenecen a una familia caracterizada por una modificación post-traduccional en su cadena  $\alpha$ , conocida como enlace tioéster (Becherer, 1989). Este enlace es producto de la transacilación intramolecular entre el grupo tiol de una cisteína y el grupo  $\gamma$ -amida de una glutamina que se encuentran en el motivo altamente conservado Gly-Cys<sub>988</sub>-Gly-Glu-Gln<sub>991</sub>-Asn. Esa unión es inestable y el grupo carbonilo de la glutamina se vuelve susceptible al ataque de grupos nucleofílicos cercanos, lo que permite que la glutamina se una covalentemente con los grupos amino o hidroxilo de moléculasceptoras, con las que forma enlaces amida o éster.

El enlace tioéster de C3 (Fig. 2) se hidroliza hacia una forma en que la cadena  $\alpha$  está intacta, pero el tioéster está roto. De este modo, una pequeña cantidad de C3 es activado espontáneamente en solución y puede unirse directamente a cualquier substrato o unirse con una molécula de agua [C3(H<sub>2</sub>O)], capaz de unirse al Factor B que después será activado por Factor D para formar C3(H<sub>2</sub>O),Bb. El complejo es una C3-convertasa y activa más C3, que dispara la vía alterna. El C3b generado de esta forma se une a una superficie y forma, junto con la properdina, el complejo C3b,Bb,P que es una C3-convertasa estable (Pangburn, 1986).

La formación y activación de la C3-convertasa constituye un ciclo de amplificación donde C3 pasa a C3b, formándose nueva C3-convertasa en la superficie del antígeno. El paso final de esta vía es la unión de otro fragmento C3b a C3-convertasa que se incorpora y activa al complejo C3b,Bb,C3b,P con actividad de C5-convertasa. Esta enzima genera C5a y C5b a partir de C5, que es el primer elemento de la vía terminal de la cascada del complemento (Pangburn, 1986).

---

\* Diversos componentes del complemento son cortados proteolíticamente durante su activación , a los fragmentos se les denomina con letras minúsculas: C3a, C3b, Ba y Bb. Los complejos formados por la asociación no covalente de dos o más proteínas se escriben usando una coma entre los símbolos: C3b,Bb. Las enzimas activas no tienen un nombre distintivo. Estas reglas de nomenclatura también se aplican a las demás vías.



- Clave**
- Anafilatoxina
  - Marcado y fagocitosis
  - ∞ Factores H y Factor I
  - ∞ DAF
  - ∞ MCP y/o CR<sub>1</sub>
  - ◆ C1INH
  - ◆ C4BP/CR<sub>1</sub> y Factor I
  - ◆ α<sub>2</sub>-macroglobulina
  - ◆ CD59

Fig. 1 Vías de activación y etapa terminal del complemento. Las cuatro vías de activación del complemento convergen a nivel de C5b y dan inicio a la vía terminal que culmina con la formación del TCC.

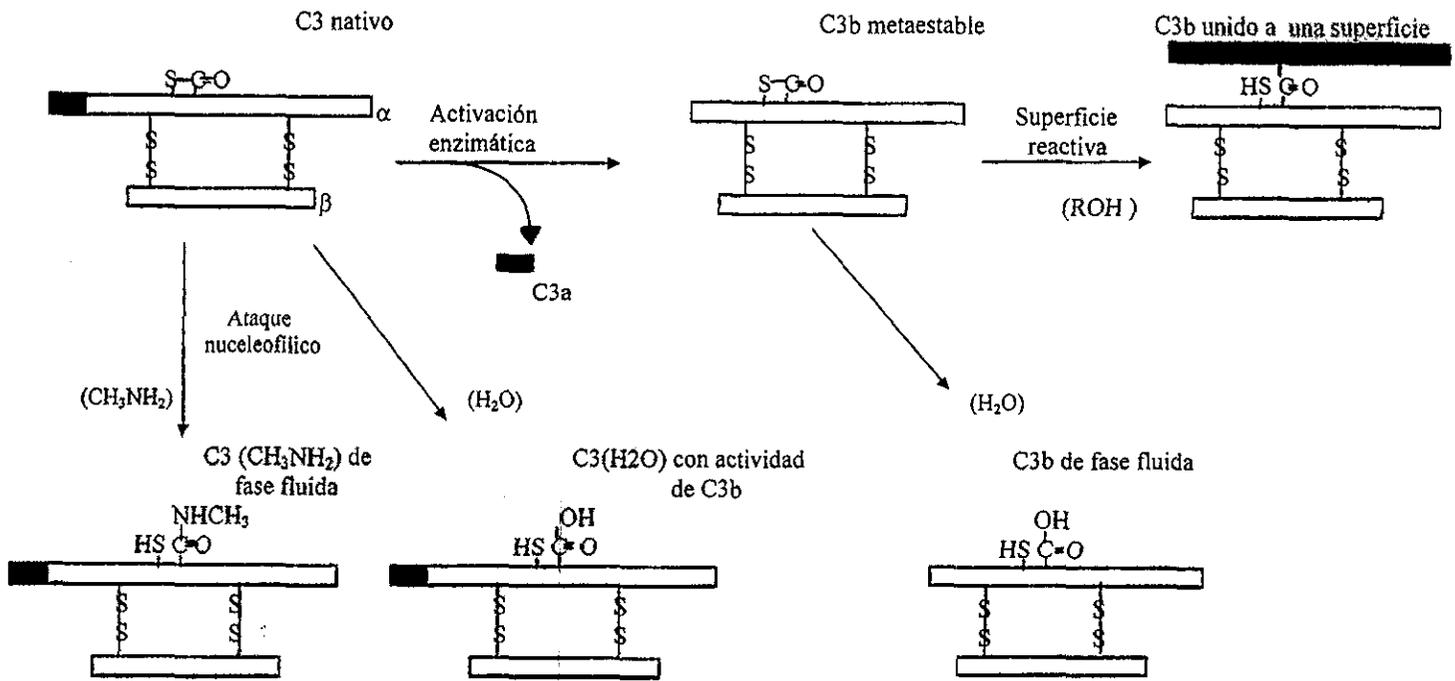


Fig.2 Reacciones del enlace tioester con grupos amino, hidroxilo o agua.

La capacidad para distinguir entre lo propio y lo extraño está dada en la cascada del complemento por factores presentes en el suero y la membrana del huésped. Esta distinción es un asunto crucial, ya que la activación espontánea de C3 o la unión de C3b a células propias constituyen un peligro potencial. Entre los reguladores séricos se encuentra el Factor H, que por su capacidad para unir C3b y C4b controla las convertasas de C3/C5 de las vías alterna y clásica. En la vía alterna, H compite con B por la unión a C3b antes de que se forme la C3-convertasa. Cuando las convertasas de C3 y C5 ya están activas, H acelera la tasa de disociación de Bb uniéndose a C3b. Aunque el Factor H no tiene actividad de proteasa, es necesario para la inactivación permanente de C3b por el Factor I, una serin proteasa que circula en forma funcional por la sangre e inactiva permanentemente a C3b y C4b asociadas a H (Pangburn, 1986).

El complemento es ineficiente para lisar eritrocitos y células nucleadas propias, proceso que se conoce como restricción homóloga, debido a las proteínas membranales MCP (Membrane Cofactor Protein), DAF (Decay-accelerating factor), CD59 y CR1 receptor de C3b/C4b. Mientras que los tres primeros factores están distribuidos en una amplia variedad de tejidos, el CR1 se

encuentra especialmente en células sanguíneas periféricas. Las características y funcionamiento de CD59 se detallan en la sección dedicada a la regulación de la vía terminal. En la vía alterna, la molécula de membrana DAF impide la asociación de C3b con el Factor B, mientras que en la vía clásica hace lo mismo con C4b y C2. También disocia a estos complejos una vez formados por lo que reduce la activación de nuevo C3 por la convertasa de las vías alterna y clásica. MCP y CR1 participan en la hidrólisis de C3b a iC3b por el Factor I. CR1 participa además en la ruptura subsecuente de iC3b a C3dg, aunque su principal función es la limpieza de complejos inmunes (Atkinson, 1987; Lachmann, 1991).

## 2b. Vía clásica

La vía clásica conecta al sistema del complemento con la respuesta inmune humoral y es el principal mecanismo efector activado por complejos antígeno-anticuerpo (clase IgG o IgM). Esta conexión le da al complemento la capacidad de activación por una respuesta inmune, por lo que requiere el desarrollo de una respuesta de anticuerpos antes de que se dé la posibilidad de activación por un antígeno dado (Liszewski, 1993).

La vía clásica consiste de cuatro componentes C1, C4, C2 y C3. El complejo C1 está compuesto por C1q, molécula de reconocimiento y unión a anticuerpos y la serín proteasa C1r<sub>2</sub>-C1s<sub>2</sub> en forma proenzimática. El subcomponente C1q está compuesto de 18 cadenas polipeptídicas que se organizan como seis unidades. Cada unidad está formada por tres cadenas cuyos extremos amino son de tipo colagénico e interactúan en una hélice. En cada subunidad existe un punto donde la estructura de hélice se rompe por discontinuidades de diversa naturaleza en los tripletes Gly-X-Y, esto genera una especie de quiebre o bisagra a la mitad de la región colagénica (Reid, 1981). En los últimos 110 residuos del extremo carboxilo de estas cadenas existen dominios globulares que se unen a complejos inmunes (Perlmutter, 1987). Los dímeros de C1r y C1s interactúan en una cadena C1s-C1r-C1r-C1s que se inserta entre las fibras colagénicas de C1q, de manera que los dos extremos libres del complejo C1r<sub>2</sub>-C1s<sub>2</sub> se enredan hacia el interior del cono formado por C1q, esta disposición permite que los sitios catalíticos de C1s estén en contacto con los de C1r (Müller-Eberhard, 1988). El C1q se une al tercer dominio constante de la cadena pesada de la IgM o de los isotipos IgG3, IgG1, IgG2 e IgG4A. Al interactuar con C1q

unido a un complejo inmune, las dos moléculas de C1r adquieren actividad proteolítica específica hacia un enlace peptídico en la subunidad adyacente, esta actividad es accesible cuando C1q se une a complejos inmunes en presencia de C1s y  $Ca^{++}$ . En estas condiciones, el C1r adquiere actividad de esterasa, por lo que actúa de inmediato sobre otra molécula de C1r que posteriormente activa a una molécula de C1s que a su vez activará a una segunda C1s que inicia la vía clásica. El mecanismo molecular del estímulo transmitido por la unión de C1q a C1r es aún controvertido. Se sugiere que la activación del complejo C1 se debe a un cambio conformacional en C1q debido a que esta molécula debe unir por lo menos dos de sus regiones globulares a anticuerpos para inducir la activación de C1r (Ziccardi, 1976; Kovacovics, 1987).

La siguiente molécula en la cascada es el C4 que se activa cuando el C1s le corta un fragmento para producir C4a y C4b. El C4b expone su enlace tioester interno y por un mecanismo similar al de C3b, se une covalentemente sobre algún sitio cercano. Posteriormente, se incorpora el C2 que también es activado por C1s e interactúa con C4b. El complejo C4bC2a es la C3-convertasa de la vía clásica. La subunidad C2a de la C3-convertasa escinde C3 y el fragmento C3b se une a la superficie del antígeno. Cuando C3b se une a la C3-convertasa, ésta adquiere un sitio de unión para C5 y actividad de C5-convertasa (Hughes-Jones, 1986; Morgan, 1990b).

El control de la actividad de C1 es el primer punto de regulación de la vía clásica. Dos moléculas del inhibidor de C1 (C1INH) se combinan con el complejo al reaccionar con el sitio catalítico de C1r y C1s, esto ocasiona que ambas se disocien de C1q y el complejo quede inactivo. En esta reacción se libera un complejo formado por una molécula de C1r y C1s y dos moléculas de C1INH. Este inhibidor no enzimático impide la activación espontánea de C1 en plasma, a través de una combinación inestable con  $C1r_2-C1s_2$ . Se ha observado que la activación de C1 mediante complejos inmunes no es afectada por C1INH, probablemente por la rápida tasa de activación de C1 en presencia de dichos complejos (Ziccardi, 1982). Además de regular al C1 activado, el C1INH impide en una etapa previa, la estimulación de C1 por activadores débiles e inespecíficos como DNA de doble cadena y heparina. De esta manera C1INH restringe la activación no-específica del complejo C1. El C1INH también posee la función de inhibir calicreína, plasmina y los Factores XI y XII de la cascada de coagulación.

Las C3/C5-convertasas de esta vía son inactivadas por la disociación de C2a promovida por C4BP (C4-binding protein) y DAF. La C4BP es una proteína plasmática con una configuración que le permite combinarse con hasta seis moléculas de C4, esta unión ocurre cerca o en el sitio de unión de C4 con C2, de modo que se evita la asociación entre dichas moléculas o acelera su separación. El DAF es una proteína presente en las membranas de distintos tipos celulares y su función es similar a la de C4BP en cuanto ocasiona la disociación de C2 de C4. Una etapa posterior de la inactivación de las C3/C5-convertasas está dada por la degradación de C4b en C4c y C4d por el Factor I, lo que ocurre en forma similar a la degradación de C3b donde CR1 y C4BP funcionan como factores auxiliares (Hughes-Jones, 1986).

### 2c. Vía dependiente de C1

La vía dependiente de C1 es una variación de la activación de la vía clásica. La evidencia experimental indica que en ausencia de inmunoglobulinas, el C1 puede activarse mediante la interacción con algunas superficies como retrovirus, *Mycoplasma*, y complejos de polianiones entre los que se encuentran heparina, protamina, polisacáridos de pneumococos (Liszewski, 1993), DNA de doble hélice, cardiolipina y fosfatidilglicerol (Kovacsovics, 1987). Estos activadores no inmunes pueden unir a C1 en múltiples sitios y aparentemente simulan los complejos antígeno-anticuerpo que inducen los cambios conformacionales que activan al C1. La activación del C1 que es independiente de anticuerpos tiene características que dependen de los activadores no inmunes. Se han descrito los llamados activadores "fuertes" tales como vesículas de cardiolipina, fragmentos de mitocondrias cardíacas humanas, o bacterias de la cepa semi-rugosa de *E. coli* J5. Estas partículas pueden activar a C1qC1s<sub>2</sub> en forma independiente de C1r sin que C1-INH tenga un papel regulador (Kovacsovics, 1987). En presencia de estos activadores unidos al C1q, la proteasa C1s<sub>2</sub> es activada en forma específica, corta a C4 y C2 e inicia la cascada del complemento como se describió para la vía clásica.

En el caso de los activadores "débiles", tales como las cepas no-rugosas de *E. coli*, DNA, heparina y mitocondrias cardíacas humanas intactas, la activación de C1 se da únicamente en unión a C1r<sub>2</sub>-C1s<sub>2</sub> y es anulada por C1-INH (Peitch, 1987). Aún se discute acerca de la relevancia fisiológica de esta vía.

## 2d. Vía de las lectinas

La vía de activación del complemento que se ha descrito más recientemente es la vía de las lectinas (Ikeda, 1987). Las lectinas son proteínas o glicoproteínas con capacidad de unir monosacáridos, polisacáridos o glicoproteínas. El origen de estas proteínas puede ser vegetal, animal, bacteriano o de hongos (Pereira, 1994). Las lectinas animales han sido principalmente involucradas en funciones de comunicación intracelular, fagocitosis y destrucción de microorganismos (Holmskov, 1994). Entre las lectinas animales caracterizadas hasta el momento se encuentran las proteínas séricas MBP (Mannan o Mannose-Binding Protein), colectina 43 (CL-43), conglutinina y los surfactantes pulmonares SP-A, SP-D. Estas moléculas forman la familia de las colectinas, lectinas animales dependientes de calcio (Malhotra, 1994).

Las colectinas son oligómeros donde cada subunidad está compuesta de tres polipéptidos idénticos, con excepción de SP-A. En la parte N-terminal de cada polipéptido existe una secuencia breve que contiene dos cisteínas conservadas, estas son seguidas por un grupo de tripletes repetidos Gly-X-Y que permiten a los tres polipéptidos de cada unidad, adoptar una forma de triple hélice similar a la de colágena. Las subunidades de cada monómero se unen covalentemente a través de enlaces disulfuro para formar oligómeros de hasta seis subunidades. En MBP y SP-A las interacciones entre subunidades se dan en la sección N-terminal de la región colagénica, lo que le da a la molécula una apariencia sertiiforme (como ramo de tulipanes). En el extremo C-terminal se localiza el dominio de lectinas donde se da la unión a carbohidratos; en el caso de MBP esta zona forma una estructura globular independiente en cada polipéptido. Los dominios de lectinas unen carbohidratos de la superficie de microorganismos, mientras la región colagénica es el ligando de un receptor de colectinas en fagocitos (Holmskov, 1994). Las colectinas son sintetizadas por hepatocitos y han sido aisladas del plasma de diversos mamíferos.

El subcomponente C1q del C1 es estructural y funcionalmente relacionado con las colectinas que también tienen actividad de opsoninas. A pesar de que el C1q también posee una región colagénica un poco más larga formada por tripletes Gly-X-Y, sus seis regiones globulares no son homólogas a las de MBP. El C1q reconoce sus ligandos por una agrupación determinada de cargas y no une amino-azúcares o residuos neutrales como lo hace MBP (Malhotra, 1994).

Solo la lectina MBP activa la cascada del complemento (Ikeda, 1987). MBP está formada por trímeros, tetrámeros o pentámeros en los que cada subunidad posee tres dominios de unión a carbohidratos. En cada especie de organismo el sitio de unión a azúcares es estructuralmente similar aunque uno distintos tipos de oligosacáridos en presencia de  $Ca^{++}$ . En el caso de la MBP humana, la especificidad de estos sitios es hacia N-acetilglucosamina > manosa, N-acetilmannosamina y fucosa > maltosa > glucosa > galactosa y N-acetilgalactosamina. El arreglo en grupo de los sitios de unión a carbohidratos permite alcanzar una avidéz funcional significativa cuando la proteína se une a muchos grupos azúcar repetidos. Ya que MBP se puede unir a muchos oligosacáridos diferentes se puede considerar que es una especie de anticuerpo universal y específico contra blancos que no se encuentran normalmente en las superficies de células de mamífero (Turner, 1996).

MBP puede asociarse *in vitro* al complejo  $C1r_2-C1s_2$ , activarlo e iniciar la vía clásica del complemento al emplear C4 en la hemólisis de eritrocitos de cordero, sin la participación de C1q. Se ha encontrado que el MBP se une al MASP *in vivo* (Mannan-Binding Protein Serin Proteases) un par de serín proteasas 37 y 36% homólogas a C1r y C1s (Matsushita, 1996). En forma equivalente a lo que sucede en C1, las MASP se autoactivan cuando se unen al MBP; mostrando una actividad similar a la de C1s contra C4 y C2 y por lo tanto generan complejos C4b2a (Matsushita, 1992). A diferencia de C1s, MASP también tiene la capacidad de cortar C3 para generar C3b, de esta manera el complejo MBP-MASP también puede activar la vía alterna del complemento, por lo que tiene un papel relevante en la primer línea de defensa inmune innata (Schweinle, 1989). MBP activa el sistema del complemento al unirse con bacterias gram-positivas, gram-negativas, levaduras, virus de la influenza, HIV, hongos, micobacterias, parásitos y una gran variedad de bacterias (Epstein, 1996). La actividad de vía de las lectinas es regulada a nivel de MASP por  $\alpha_2$ -macroglobulina y C1INH, de manera similar a como sucede con C1r/C1s. Las demás etapas de esta cascada son controlados como las vías clásica y alterna.

Además de las funciones en conjunto con el sistema del complemento, MBP tiene actividad de opsonina al unirse a un receptor específico de colectinas expresado en células fagocíticas. El receptor de colectinas es una proteína relacionada con la calreticulina (Turner, 1996). Se ha identificado en muchos tipos celulares, incluyendo a la mayoría de los leucocitos, plaquetas, células endoteliales, fibroblastos y tejido especializado (Holmskov, 1994). La MBP,

conglutinina y la proteína SP-A también se unen al mismo receptor que C1q (Malhotra, 1990). Otra forma en que la vía de las lectinas puede mediar fagocitosis es a través de activar el complemento y depositar grandes cantidades de C3b en las superficies activadoras. La vía de las lectinas es importante fisiológicamente, puesto que a través de MBP, C3b y sus respectivos receptores se produce la opsonización en ausencia de anticuerpos. Esto es apoyado por la observación de que recién nacidos que carecen o tienen alguna deficiencia de MBP sufren infecciones bacterianas recurrentes (Super, 1989).

### 3. Etapa terminal de la cascada

La etapa terminal es una serie de reacciones compartida por todas las vías de activación del complemento. En esta etapa, se ensambla el complejo terminal del complemento (TCC) a partir de C5, C6, C7, C8 y C9. El TCC participa en procesos de transducción de señales y lisis celular. Una diferencia importante entre las etapas de activación, y la terminal es que la primera etapa se lleva a cabo en la fase fluida y en la superficie hidrofílica de las membranas, mientras que los efectos funcionales de la vía terminal se deben a la interacción de sus proteínas con la fase hidrófoba de las membranas. Aunque los precursores inactivados de la vía terminal del complemento se encuentren como moléculas hidrofílicas en circulación, pueden pasar por una etapa hidrofílica-anfifílica para insertarse en la fase hidrófoba y desestabilizar la membrana. Los mecanismos moleculares de esta transición son similares para todas las proteínas de la etapa terminal: interacciones proteicas de alta afinidad permiten un drástico rearrreglo conformacional que resulta en la exposición e inserción de dominios hidrófobos previamente ocultos (Podack, 1984a).

#### 3a. Elementos de la fase terminal.

El vínculo entre las cuatro cascadas de activación y la vía terminal es el C5. A diferencia del resto de las moléculas del TCC, la secuencia de aminoácidos de C5 es homóloga a la de C3, C4 y  $\alpha_2$ -macroglobulina, sin embargo C5 no posee un enlace tioéster interno ni forma uniones covalentes con membranas (Podack, 1986a). El precursor del C5 humano comprende 1676 aminoácidos, incluyendo un péptido señal de 18 residuos. Tanto el C5 como el C3 y el C4 son sintetizados como un precursor de una sola cadena que antes de ser secretado al plasma es cortado

conglutinina y la proteína SP-A también se unen al mismo receptor que C1q (Malhotra, 1990). Otra forma en que la vía de las lectinas puede mediar fagocitosis es a través de activar el complemento y depositar grandes cantidades de C3b en las superficies activadoras. La vía de las lectinas es importante fisiológicamente, puesto que a través de MBP, C3b y sus respectivos receptores se produce la opsonización en ausencia de anticuerpos. Esto es apoyado por la observación de que recién nacidos que carecen o tienen alguna deficiencia de MBP sufren infecciones bacterianas recurrentes (Super, 1989).

### 3. Etapa terminal de la cascada

La etapa terminal es una serie de reacciones compartida por todas las vías de activación del complemento. En esta etapa, se ensambla el complejo terminal del complemento (TCC) a partir de C5, C6, C7, C8 y C9. El TCC participa en procesos de transducción de señales y lisis celular. Una diferencia importante entre las etapas de activación, y la terminal es que la primera etapa se lleva a cabo en la fase fluida y en la superficie hidrofílica de las membranas, mientras que los efectos funcionales de la vía terminal se deben a la interacción de sus proteínas con la fase hidrófoba de las membranas. Aunque los precursores inactivados de la vía terminal del complemento se encuentran como moléculas hidrofílicas en circulación, pueden pasar por una etapa hidrofílica-anfifílica para insertarse en la fase hidrófoba y desestabilizar la membrana. Los mecanismos moleculares de esta transición son similares para todas las proteínas de la etapa terminal: interacciones proteicas de alta afinidad permiten un drástico rearrreglo conformacional que resulta en la exposición e inserción de dominios hidrófobos previamente ocultos (Podack, 1984a).

#### 3a. Elementos de la fase terminal.

El vínculo entre las cuatro cascadas de activación y la vía terminal es el C5. A diferencia del resto de las moléculas del TCC, la secuencia de aminoácidos de C5 es homóloga a la de C3, C4 y  $\alpha_2$ -macroglobulina, sin embargo C5 no posee un enlace tioéster interno ni forma uniones covalentes con membranas (Podack, 1986a). El precursor del C5 humano comprende 1676 aminoácidos, incluyendo un péptido señal de 18 residuos. Tanto el C5 como el C3 y el C4 son sintetizados como un precursor de una sola cadena que antes de ser secretado al plasma es cortado

en dos cadenas,  $\alpha$  y  $\beta$  unidas por un enlace disulfuro (Podack, 1984a). La anafilatoxina C5a consta de los primeros 74 residuos de la cadena  $\alpha$ , el resto de la molécula es conocido como C5b (Hughes, 1994). La activación de C5 expone un sitio de unión para C6 en C5b, además C8 $\beta$  posee un dominio de unión a C5b que media su asociación al complejo C5b-C7.

Existe una clara relación estructural, funcional y filogenética entre C6, C7, C8 $\alpha$ , C8 $\beta$  y C9 (C6-C9). Todas son proteínas mosaico formadas por 3-8 módulos ricos en cisteínas y cuya longitud va de 30 a 70 residuos, estos módulos se encuentran alrededor de una región pobre en cisteínas (CP) de aproximadamente 350 residuos. Los módulos de C6-C9 son homólogos a los que se encuentran individualmente en otras proteínas que no pertenecen al sistema del complemento (González, 1996). Se han identificado y nombrado en los componentes terminales a los módulos TSP, LDLr, EGF, CFI y SCR.

El módulo TSP está repetidos tres y seis veces en trombospondina y properdina respectivamente (Goundis, 1988) estos módulos tienen una longitud de 60 aminoácidos y contienen un consenso de seis cisteínas y tres triptófanos (Reid, 1989). El módulo LDLr tiene una longitud de 40 aminoácidos contiene seis cisteínas consenso. En el consenso de los módulos LDLr del receptor de lipopolisacáridos se ha observado que predominan los residuos ácidos, todos los residuos cargados que están conservados tienen una carga negativa, especialmente en la secuencia Asp-Cys-X-Asp-Gly-Ser-Asp-Glu. De los 19 residuos conservados en el módulo LDLr, 14 también están en C9 humano, incluyendo la zona de cargas negativas (Südhof, 1985). Este módulo está repetido siete y seis veces en los receptores de lipopolisacáridos de baja densidad, y en el de apolipoproteína E respectivamente, además se encuentra en el Factor I, regulador del complemento (Reid, 1989). El módulo EGF primeramente se descubrió en el Factor de crecimiento epidermal, aunque también se encuentra en el receptor clase B de lipopolisacáridos de baja densidad, en las proteasas C1r, C1s, en Factor IX de la coagulación y en la perforina. En el módulo EGF también se distinguen seis cisteínas consenso (González, 1996).

En C6 y C7 existen además los módulos CFI y SCR. El módulo CFI es homólogo a los que existen en el Factor I del complemento, tiene 70 aminoácidos de longitud y 8-10 cisteínas consenso. Las secuencias consenso repetidas (Short Consensus Repeats) (SCR), se encuentran en el grupo de genes RCA, donde están incluidos los reguladores del complemento con capacidad de mediar interacciones con C3b/C4b, tales como C4BP, el Factor H, CR1, y CR2. Además se

encuentran SCR en serin proteasas asociadas con la activación del sistema del complemento como B, C2, C1r, C1s, se ha establecido que los SCR están implicados en interacción de estas moléculas con C3b, C4b y C5b. Los módulos SCR contienen 60-70 aminoácidos, y cuatro cisteínas conservadas que forman enlaces disulfuro intramódulo en orden 1-3 y 2-4 (Reid, 1989). El módulo EGF y parte de la región CP también están presentes en las perforinas (Shinkai, 1988), proteínas de linfocitos citotóxicos y NK, que al igual que C9, tienen la capacidad de formar poros (Tschopp, 1986).

En la Tabla II se muestran las características de los módulos TSP, LDLr y EGF.

Tabla II. Secuencias de los módulos TSP, LDLr y EGF de las proteínas donde se caracterizaron inicialmente, así como de las proteínas C6-C9 y perforina. Información a partir de las referencias (Goundis, 1988; Lawler, 1986; Südhof, 1985; Lowrey, 1989).

Tabla II A. Módulo de Trombospondina

Homología tipo 1	DSADDGWSFWSEWTSCSTSCNGIQORGRSCDSLNNR.....	CEGSSVQTRT..	CHIQE.CDK	(361-337)			
Homología tipo 1	FKQDGGWSHWSPWSSCSVTCDGVTIRILCNPSPP.QMNGKPCGEARETKA..	CKKDA.CPI		(417-473)			
Homología tipo 1	NGGWGFWSPWDICSVTCGGGVQKRSRLCENNPTP.QFGGKDCVGDVTENQI..	CNKQD.CPI		(474-530)			
C6 HUM	TSP I	CDHYAWTQWTS...CSKTCNSGTQSRHRQIVVDKYYQENF..	CEQICSKQETRECNWQR.CPI				
C9.FUG		CV...WSRWAPWSSCDP.CTNTRR.RSRGVEVFG..	QFAGIACQGSV..GDREYCITNAKCNL				
C9.TRO		CV...WSRWSEWTPCNS.CTKIRH.RSRSVEVFG..	QFGGKPCQG.QPIGEQQRCSTDAVCEQ				
C9.RAT		CRM...STWSQWSQCDP.CLKQRF.RSRSMEVFG..	QFQKSCA..DALGDRQHCEPTQECEE				
C9.MUS		CRM...SPWSNWSECDP.CLKQRF.RSRSILAFG..	QFNGKSCV..DVLGDRQGCETTQECEE				
C9.RAB		CRM...SPWSEWSECDP.CLRQMF.RSRSIEVFG..	QFHGKSCV..DALGDRRACIPTTEACED				
C9.HOR		CRM...SSWSEWSECDP.CLRQMF.RSRSIEVFG..	QFNGQRCV..DAVGDRRQCVPTTEACEE				
C9.HUM		CRM...SPWSEWSQCDP.CLRQMF.RSRSIEVFG..	QFNGKRCT..DAVGDRRQCVPTPEPCED				
C8A.RAB		CQL...SSWSEWTDCCP.CQDTKY.RHRSLLQPN..	KFGGTICSG.DIW.DRASCYSPTACL				
C8A.HUM		CQL...SNWSEWTDCCP.CQDKKY.RHRSLLQPN..	KFGGTICSG.DIW.DQASCSSSTTCVR				
C8B.RAB		CEL...SSWSSWTMCDP.CQKKRY.RHAYLLRPS..	QFNGEPCNFSDKEVED..CATSRPCRS				
C8B.HUM		CEL...SSWSSWTTCDP.CQKKRY.RYAYLLQPS..	QFHGEPCNFSDKEVED..CVTNRPCGS				
C7.HUM		CQWDFYAPWSE...CNG.CTKTQT.RRRSVAVYG..	QYGGQPCVNAF.ETQS..CEPTRGCPT				
C6.HUM	TSP II	CLLGDGFGWSD...CDP.CIEKQS.KVRSVLRPS..	QFGGQPCTEPLVAFQP..CIPSKLCKI				
C9FUGU		CQ...WSCWSSWSNCQ...GQKRSRTRYCNTEGVL...GAECRGEIRSEEY..	C				
C9TROUT		CQ...WSCWAAWSGCS...GGKRIRTRSCNTQGLS...DATCRGDIVTEDY..	C				
C8ARABB		CE...WSCWGSWSPCTA...GTRE.RRRECNNPAP.QNGGAPCPGWRVQTOA..	C				
C8AHUMAN		CE...WSCWSSWVCRA...GIQE.RRRECNDPAP.QNGGASCPGRKVQTOA..	C				
C8BRABB		CE...WSCWSRWSSCSG...GQKT.RRRQCNNPAP.QDGGSPCSGPASETLD..	C				
C8BHUMAN		CE...WNCWSNWSSCSG...RRKT.RQRQCNNPPP.QNGGSPCSGPASETLD..	C				
C6HUMAN	TSP III	CE...WGCWSSWSTCDA...TYKRSRTRECNDPAP.QRGGKRCEGEKQREED..	C				
C7HUMAN							
Consenso		C	W W C C	R R	Q G C	C	C

Tabla II B. Módulo de Receptor de Lipopolisacáridos de Baja Densidad

Módulo Y	VGDR.C.ERNEFQCQD..GKCISYKWVCDGSAECQDGS.DESQETC
Módulo II	LSVT.C.KSGDFSCGGRVNRCPQFWRCQVDCDNGS.DEQG..C
Módulo III	PKPT.C.SQDEFRCHE..GKCISRQFVCDSDRDCLDGS.DEAS..C
Módulo IV	PVLT.C.GPASFCQNS..STCIPQLWACDNDPCEDGS.DEWPQRC
Módulo V	DSSP.C.SAFEFHCLS..GECIHSSWRCDGGPDCCKDKS.DEEN..C
Módulo VI	AVAT.C.RPDEFQCS..GNCIHGSRQCDREYDCKDMS.DEVG..C
Módulo VII	VNVTLCCEGPNKFKCHS..GECITLDKVCNMARDCRDWS.DEPIKEC
C9.FUG	PRE..CSDS.EFQC.ES.GSICKLRLKCNQDCEDEGS.DE.D..CE..PLRKT.CPPTVLDTNEQ
C9.TRO	PSE..CSSI.EFTC.ES.GACIKLRLSCNGDYDCEDGS.DE.D..CE..PVRKP.CGTKLYDTNEQ
C9.RAT	EN...C.GN.DFQC.ET.GRCIKRLLCNGDNDCEGDFS.DESD..CESDP.RLP.CRDRVVEESEL
C9.MUS	EN...C.GN.DFQC.ET.GRCIKRLLCNGDNDCEGDFS.DEND..CDDDP.RTP.CRDRVAESEL
C9.RAB	E.D..CEKD.EFHCG.T.GRCIKRLLCNGDNDCEGDFS.DEDD..CETEP.RLT.CRNREVQESL
C9.HOR	D.D..C.GN.DFQCG.T.GRCIKRLLCNGDNDCEGDFS.DEDD..CENDP.RPP.CRERVVESEL
C9.HUM	D.D..C.GN.DFQCS.T.GRCIKMRLCNGDNDCEGDFS.DEDD..CESEP.RPP.CRDRVVEESEL
C8A.RAB	Q....C.GQ.DFQCKET.GRCLKRHLVCGENDCLDGS.DEDN..CEDIRATESDCAQYD..PIPG
C8A.HUM	Q....C.GQ.DFQCKET.GRCLKRHLVCGENDCLDGS.DEDD..CEDVRAIDEDCSQYE..PIPG
C8B.RAB	R....CEG...FVCAQT.GRCVNRLLCNGDNDCEGDFS.DEAN..CRKIYKK...CHHEMEQYWAI
C8B.HUM	R....CEG...FVCAQT.GRCVNRLLCNGDNDCEGDFS.DEAN..CRRIYKK...CQHEMDQYWGI
C6.HUM	EAD..C.KN.KFERCD.S.GRCIARKLECGENDCEGDFS.DERD..CGRTKAV...CTRKN.PIPS
C7.HUM	E.G..C.GE.RFERCF.S.GQCISKSLVCGDSDCEDSADE.DR.CEDSE.RRPSCIDK..PPP
Consenso	C DF C G CI C DC DGS DE C
	E E

Zona negativa

Tabla II C. Módulo de Factor de Crecimiento Epidermal

EGF MUS repeat 1	..CSSPDNGGCSQICLPLRPGSWECDCFPGYDLQSDRKSCAASG
EGF RAT repeat 1	..CSFSDNGGCSQICLPLSLASWECDCFPGYDLQLDRKSCAASM
EGF HUM repeat 1	..CSSPDNGGCSQICVPLSPVSWECDCFPGYDLQLDEKSCAASG
C9.FUG	CKCRPCHNGGTLALLDGKICMCSNLFEGLG.CQ
C9.TRO	CKCKPCHNGGSLALLDGKCLCLLPQFEGLA.CQ
C9.RAT	..CYPQNGGTAILLDGQCMCSCTIKFKGIA.CE
C9.MUS	..CYPCLNGGTIILLDGQCLCSCPMFRGMA.CE
C9.RAB	..CSPCQNGGTALLMDGQCLCTCPFMFEGIA.CE
C9.HOR	..CHPCQNGGTVIQIDGQCLCSCPIAFEGIA.CE
C9.HUM	..CHTCQNGGTVILMDGKCLCACPFKFEFEGIA.CE
C8A.RAB	CRCGPCFNNGKPILEGTSCRCQCSLGLQGPA.CE
C8A.HUM	CRCGPCFNNGVPILEGTSCRCQCRGLGSLGAA.CE
C8B.RAB	CRCAPCQNGVVPVQKGSRCDCICPVGFQGS.A.CE
C8B.HUM	CHCAPCQNGVVPVLKGSRCDCICPVGSQGLA.CE
C6.HUM	CQCAPCPNNGRPTLSGTECLCVCQSGTYGEN.CE
C7.HUM	CHCRPCQNGGLATVEGTHCLCHCKPYTFGAA.CE
PER.MUS	CS.RPCRS.GQHKSSHDSQCCEQDSKVTNQDC.
PER.RAT	CN.RPCRA.GQHKSSRDSCQCVCQDSNVTNQDC.
PER.HUM	CS.RPCPP.GRQKSPRDPCQCVCCHGSAVTTQDC.

### 3b. Características de las proteínas C6-C9.

En la Tabla III se concentra la información general sobre las proteínas caracterizadas hasta la fecha. En la Figura 3 se muestra la información sobre la longitud y módulos que las componen<sup>†</sup>.

#### C6

Mediante la técnica de *dicroísmo circular*, se ha observado que la estructura secundaria del C6 presenta un 29% de estructura  $\beta$  plegada, 21% de giros  $\beta$  y un 12% de hélices  $\alpha$ ; éstas se encuentran principalmente en la región CP donde se han detectado varias zonas hidrófobas que pudieran ser los sitios de interacción con la membrana (DiScipio, 1989). La apariencia tridimensional de C6 ha sido observada por microscopía electrónica y consiste de una región globular que contiene los módulos compartidos con C8 $\alpha$ , C8 $\beta$  y C9, y una región alargada donde se ubican dos pares de módulos SCR y CFI (DiScipio 1989). En estos últimos se encuentra el sitio de unión de C6 con C5b. Los módulos SCR y CFI del C6, son buenos candidatos para la unión con el C5b pues se ha demostrado que se unen específicamente al C5b y a sus homólogos C3b y C4b. Esto se deduce del hecho de que el Factor I inactiva a C3b en cooperación con el Factor H, que únicamente consiste de módulos SCR, por lo tanto es posible que el módulo CFI ubicado en la cadena pesada del Factor I ayude a reconocer un sitio en C3b (Haeffliger, 1989).

#### C7

El análisis de la estructura secundaria de C7 por *dicroísmo circular* ha revelado que contiene 38% de estructura  $\beta$  plegada, 24% de giros  $\beta$  y 7% de hélices  $\alpha$  (DiScipio, 1988). Análisis de hidropatía indica que las posiciones 1-325 son hidrofílicas mientras la zona ubicada entre los residuos 326-821 contiene zonas intercaladas de hidrofobicidad, especialmente en la región central y entre las posiciones 607-785 (DiScipio, 1988). El C7 es una molécula soluble pero cuando forma monómeros o dímeros con C5b-6, el complejo se vuelve anfifílico. Los monómeros que forma con estas proteínas consisten de una hoja y un tallo flexible, mientras el dímero tiene dos hojas unidas por los tallos superenrollados. El tallo de esta estructura está

<sup>†</sup> Abreviaturas de las proteínas usadas en este proyecto: C9.fug = C9 de *Fugu rubripes* (pez globo); C9.tro = C9 de trucha arcoiris; C9.mus = C9 de ratón; C9.rat = C9 de rata; C9.rab = C9 de conejo; C9.hor = C9 de caballo; C9.hum = C9 de humano; C8 $\alpha$ .rab = C8 $\alpha$  de conejo; C8 $\alpha$ .hum = C8 $\alpha$  de humano; C8 $\beta$ .rab = C8 $\beta$  de conejo; C8 $\beta$ .hum = C8 $\beta$  de humano; C6.hum = C6 de humano; C7.hum = C7 de humano; per.mus = perforina de ratón; per.rat = perforina de rata; per.hum = perforin de humano.

Tabla III. Características generales de las secuencias proteicas deducidas de los genes C6-C9.

ORGANISMO	PROTEÍNA	LONGITUD*	SEÑAL	GLICOSILACIÓN	CISTEÍNAS*	FUENTE	CLAVE	REFERENCIA
<i>Fugu rubripes</i>	C9 fugu	586	1-26	246, 274, 354, 545*	28	DNA genómico	g1845349	Yeo, 1997
<i>Oncorhynchus mykiss</i>	C9 trucha	574	?	3, 231, 261, 533*	28	?	g116616	Gene Bank
<i>Mus musculus</i>	C9 ratón	528 (parcial)	?	38, 243, 397*	22	?	g755764	Gene Bank
<i>Rattus norvegicus</i>	C9 rata	567	?	274, 419, 422, 436, 539*	22	?	g1256828	Gene Bank
<i>Oryctolagus cuniculus</i>	C9 conejo	557	1-21	261, 282, 424*	22	hígado	g1352109	Gene Bank
<i>Equus caballus</i>	C9 caballo	547	1-21	260, 277, 451*	24	cDNA hígado	g1352107	Gene Bank
<i>Homo sapiens</i>	C9 humano	559	1-21	256	24	cDNA hígado	g1352108	DiScipio, 1984
<i>Oryctolagus cuniculus</i>	C8 $\alpha$ conejo	585	1-20	408	29	cDNA hígado	g1352105	White, 1994
<i>Homo sapiens</i>	C8 $\alpha$ humano	584	1-20	43, 437	29	cDNA hígado	g729167	Rao, 1987
<i>Oryctolagus cuniculus</i>	C8 $\beta$ conejo	590	1-26	47	28	cDNA hígado	g1352106	White, 1994
<i>Homo sapiens</i>	C8 $\beta$ humano	591	1-26	47, 189, 499*	28	cDNA hígado	g116612	Howard, 1987
<i>Homo sapiens</i>	C6 humano	934	1-21	303, 834*	64	cDNA hígado	g116609	Haefliger, 1989
<i>Homo sapiens</i>	C7 humano	820	1-22	180, 732	54	cDNA hígado	g87197	DiScipio, 1988
<i>Homo sapiens</i>	perforina humana	555	1-21	184, 528	20	cDNA CTL	g129819	Lichtenheld, 1988
<i>Mus musculus</i>	perforina ratón	554	1-20	184, 351, 528 *	20	cDNA NK	g110800	Shinkai, 1988
<i>Rattus norvegicus</i>	perforina rata	554	1-20	184, 380, 528*	20	cDNA LGL	g548477	Ishikawa, 1989

\* Incluye péptido señal

\* Determinación no experimental por identificación de potenciales sitios de glicosilación con la secuencia Asn-X-Ser/Thr.

\* No incluye péptido señal

Abreviaturas: CTL: Linfocitos T citotóxicos. NK: Linfocitos Natural Killer. LGL: Linfocitos Granulares Grandes.

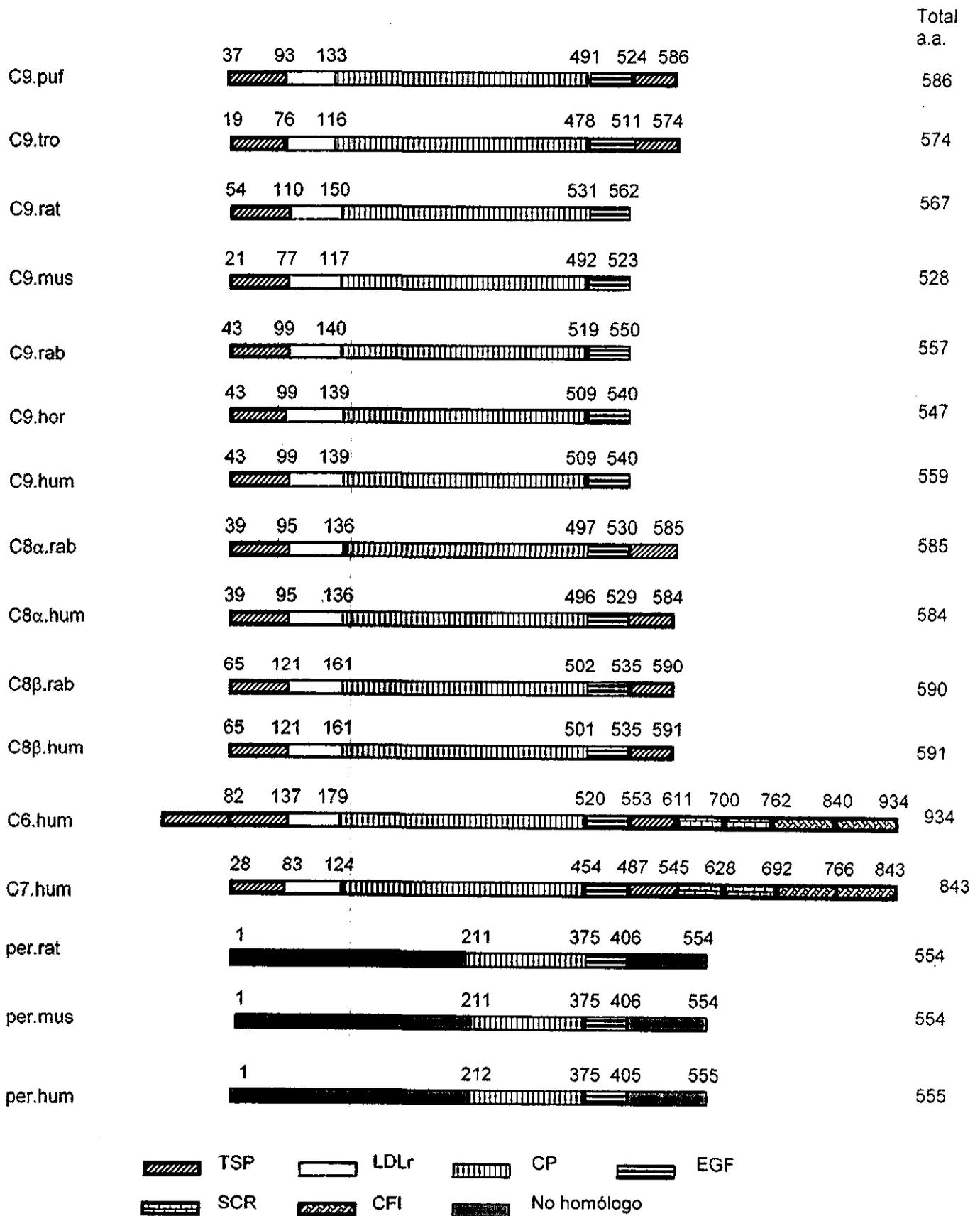


Fig. 3 Módulos que forman a las proteínas de los genes C6-C9 conocidos. TSP: Trombospondina, LDLr: receptor de lipopolisacáridos de baja densidad, CP: Zona pobre en cisteínas, EGF: Factor de crecimiento epidérmico, SCR: Short Consensus Repeat, CFI: Factor I del complemento.

formado principalmente por C6 y C7, su función es la de mediar la unión con la membrana sin penetrar lo suficiente como para formar canales transmembranales. Cualquiera de las formas del complejo C5b-7 insertado en membrana puede activar la formación del TCC en presencia de C8 y C9 (DiScipio, 1988).

## C8

El complejo C8 consiste de las subunidades  $\alpha$ ,  $\beta$  y  $\gamma$ . Estas proteínas son codificadas por tres genes separados, sin embargo los genes  $\alpha$  y  $\beta$  de humanos están ligados genéticamente entre sí y con PGMI, en el locus 1p32 (Michelotti, 1995). Además, C8 $\alpha$  y C8 $\beta$  exhiben una estructura modular idéntica.

En el complejo C8 las subunidades  $\alpha$  y  $\gamma$  están unidas por un enlace disulfuro. En cambio, C8 $\beta$  se asocia directamente con C8 $\alpha$  en forma no covalente. En presencia de SDS las subunidades de C8 nativo pueden separarse en las subunidades  $\alpha$ - $\gamma$  y  $\beta$ , al retirarse el detergente estas pueden combinarse nuevamente y exhibir actividad hemolítica normal (Monahan, 1980). La subunidad C8 $\beta$  contiene el dominio estructural que media las interacciones de C8 con el complejo C5b-C7 (Monahan, 1981). Este sitio reconoce específicamente a C5b. Un par de dominios independientes en C8 $\beta$  median las interacciones con los lípidos de membrana y con un sitio particular de  $\alpha$ - $\gamma$ . Por otra parte, C8 $\alpha$  posee un dominio que le permite insertarse en la bicapa membranal y contribuir a la actividad lítica de C5b-C9 (Steckel, 1983). Otros dominios en C8 $\alpha$  unen por separado al C8 $\gamma$  en forma covalente (Brickner, 1985) y a la primera molécula de C9 que participa en la formación del TCC (Podack, 1982a). En el complejo C5b-8, la molécula C8 $\alpha$  es la que principalmente resulta identificada por marcadores fotosensibles hidrófobos, lo mismo sucede con esta molécula y C9 en el complejo C5b-9, lo que quiere decir que tanto C8 $\alpha$  como C9 son las que median la interacción del complejo C5b-9 con membranas (Steckel, 1983). En la Figura 4 se representan las asociaciones entre las moléculas que forman el complejo C8.

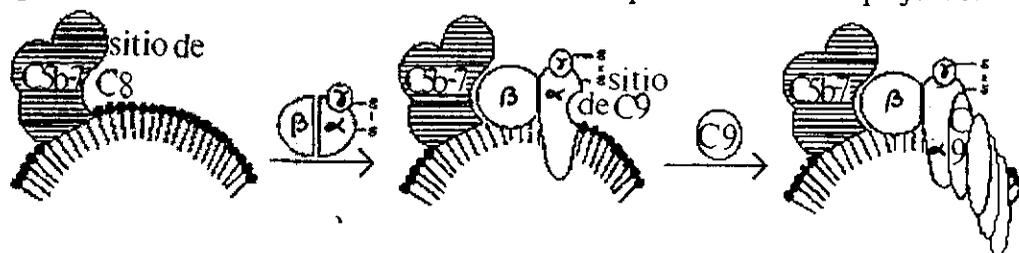


Fig.4 Organización del complejo C8 e interacción con las demás moléculas que forman el TCC. Figura modificada de ( Sodetz, 1988)

No existe una reserva intracelular de C8 $\alpha$  y C8 $\gamma$  libres lo que sugiere que son sintetizadas a una misma velocidad y que la formación del enlace disulfuro entre  $\alpha$ - $\gamma$  ocurre en forma simultánea a la traducción o durante una etapa temprana del procesamiento postraduccional. La biosíntesis de C8 $\beta$  aparentemente es más lenta que la de  $\alpha$  y  $\gamma$  de modo que todo el C8 $\beta$  intracelular se une al complejo C8 antes de ser secretado (Sodetz, 1988). El complejo C8 es esencial para la unión del C9 al módulo de polimerización C5b-C8.

### C8 $\alpha$

C8 $\alpha$  contiene 29 cisteínas, la mayor parte de las cuales se encuentran formando enlaces disulfuro dentro de los módulos TSP, LDL y EGF, mientras la cisteína en la posición 164 forma un enlace disulfuro con la cisteína 40 de la subunidad C8 $\gamma$ . Mediante análisis de la secuencia de aminoácido se han identificado diversos segmentos de la región libre de cisteínas que pueden formar estructuras anfífilicas que interactúan con la membrana (Rao, 1987). Experimentos con marcadores fotosensibles hidrófobos han demostrado que C8 $\alpha$  se inserta en la bicapa lipídica (Steckel, 1983). Adicionalmente C8 $\alpha$  tiene un sitio de interacción con C9 de modo que media la unión entre C5b-C8 y C9.

### C8 $\beta$

El perfil hidropático de C8 $\beta$  humano sugiere que es una proteína predominantemente hidrofílica, con siete segmentos que se unen a la superficie de la membrana sin atravesarla (Howard, 1987) (65-75, 147-162, 196-216, 236-249, 320-347, 405-418 y 489-499). Esta interacción marginal con la bicapa lipídica ha sido confirmada mediante marcadores fotosensibles (Sodetz, 1988). Análisis de carbohidratos revelan que esta proteína contiene una composición muy similar a la de C8 $\alpha$  (Haefliger, 1987).

Entre C8 $\alpha$ , C8 $\beta$  y C9, además de los módulos homólogos y la región pobre en cisteínas, existen otras porciones relevantes como los sitios únicos a C8 $\alpha$  y C8 $\beta$  humanos (posiciones 379-395 y 504-543 de C8 $\alpha$ ) y una región que existe solamente entre C8 $\beta$  y C9 humanos (posiciones 157-181 de C8 $\beta$ ) (Howard, 1987).

La asociación de C8 al complejo C5b-7 está dada por un sitio específico en C8 $\beta$  que se une específicamente y con alta afinidad a la molécula C5b. Tanto C5b como C8 $\beta$  forman un

complejo que se distingue al microscopio electrónico como un prolongación que sobresale del túbulo formado por C6, C7, C8 $\alpha$ - $\gamma$  y C9. Dicho apéndice se puede separar del resto del complejo por tratamiento con SDS (Stewart, 1987).

### C8 $\gamma$

La secuencia de C8 $\gamma$  humano se ha derivado de una clona de cDNA. Contiene tres cisteínas, una de las cuales se une con C8 $\alpha$ . Adicionalmente, C8 $\alpha$  y C8 $\gamma$  pueden interactuar en forma no covalente y con una alta afinidad mediante un dominio distinto a los ya mencionados, que no es bloqueado por la unión de C8 $\beta$  a C8 $\alpha$  o por la integración de éstos al complejo C5b-C7 (Brickner, 1985). En esta proteína no existen regiones homólogas con los otros componentes del TCC, tampoco tiene segmentos que interactúan con la membrana. El C8 $\gamma$  en el complejo C5b-C9, se localiza en un ambiente hidrofílico en la periferia de C5b-8 y C5b-9 (Sodetz, 1988). En contraste con C8 $\alpha$  y C8 $\beta$ , no se ha determinado cual es el papel de C8 $\gamma$  en la cascada del complemento o la respuesta inmune, pues su ausencia no afecta la función de C8. Por su estructura de intrones y exones C8 $\gamma$  se ha relacionado con las lipocalinas, proteínas con funciones de identificación y transporte de ligandos hidrofóbicos (Kaufman y Sodetz, 1994).

### C9

Diversas líneas de evidencia indican que el C9 tiene un papel directo en la interacción del TCC con los lípidos de membrana. El C9 aislado y sometido a temperaturas de 37-46°C tiene capacidad de formar polímeros lineales en solución (poli-C9), o polímeros en forma de poros en presencia de liposomas o eritrocitos de carnero (Podack, 1982b; Tschopp, 1982). El C9 polimeriza circularmente en presencia de iones de Zn<sup>++</sup>, Ca<sup>++</sup> y Mg<sup>++</sup> (Young, 1986). La polimerización espontánea del C9 en presencia de membranas lipídicas es un proceso lento que toma 2-3 días (Tschopp, 1982). En presencia de Zn<sup>++</sup> la polimerización espontánea ocurre en 2-3 horas. En contraste, la polimerización mediada por C5b-8 insertado en una membrana requiere 2-3 minutos (Podack, 1986b). Esto sugiere que el complejo C5b-8 induce cambios de orientación entre dos o más moléculas de C9, lo que desencadena el desdoblamiento, inserción y polimerización de C9 en la bicapa lipídica. El complejo C5b-8 funciona como un catalizador pues reduce la energía de activación necesaria para que C9 polimerice.

El Poli-C9 insertado en alguna membrana por iones o temperatura exhibe cambios conformacionales y en sus determinantes antigénicos, equivalentes a los del TCC formado por C6, C7, C8 $\alpha$ - $\gamma$  y C9<sub>10-18</sub> (Tschopp, 1982). Al igual que el TCC, el poli-C9 es resistente a ebullición en presencia de detergentes iónicos y su tamaño aparente en condiciones reductoras y no reductoras es de  $1.1 \times 10^6$  kDa (Podack, 1982c). La imposibilidad de reproducir la polimerización de C9 en células nucleadas indica que existen mecanismos de regulación (presentes en las membranas y en el suero sanguíneo) que limitan la actividad membranolítica de C9. El C9 es esencial para la formación de lesiones membranales inducidas por el TCC. Se requieren entre 1-18 monómeros de C9 para que el TCC ocasione citólisis.

La enzima  $\alpha$ -trombina hidroliza la unión del residuo 245 de C9, así se obtienen C9a hidrofílico y C9b hidrófóbico, los cuales no pierden su actividad hemolítica (Biesecker, 1982, Shiver 1986). Experimentos con sondas fotoactivas hidrófóbicas, han mostrado que cuando C9a y C9b forman poros en vesículas de fosfolípidos, ya sea solos o con los demás componentes del TCC, únicamente el C9b resulta marcado (Ishida, 1982). Puesto que solo C9b penetra en la fase hidrocarbonada de la membrana, se ha propuesto que este fragmento contiene la o las regiones de C9 que interactúan directamente con membrana. Análisis por métodos computacionales han revelado que C9b no contiene regiones hidrófóbicas continuas cuya longitud pueda atravesar la membrana, sin embargo se han detectado varios fragmentos candidatos a plegarse en hélices  $\alpha$  con capacidad de formar un canal de centro hidrofílico. Por estos mismos métodos se han identificado en C9a regiones con capacidad de unión superficial a membrana (membrane-seeking) que pueden adoptar un plegamiento de láminas anfipáticas  $\beta$  plegadas (Shiver, 1986; DiScipio, 1984; Stanley, 1985).

La ausencia de largas zonas hidrófóbicas en esta proteína evita que el C9 se comporte como una proteína integral de membrana, además de posibilitar su secreción. Es posible que los fragmentos cortos de estructura secundaria anfipática son los que median la interacción de C9 con membrana. El uso de marcadores hidrófóbicos ha permitido ubicar las regiones de interacción con membrana entre los residuos 176-345, en esta zona posiblemente se conforma una helice-giro-helice que media las interacciones con membrana (Peitch, 1990).

En esta década se han propuesto varios modelos moleculares de la estructura terciaria del C9 basados en estudios de microscopía electrónica, dispersión de neutrones y cristalografía por

rayos X. Las dimensiones obtenidas por estos métodos van de 7.7 x 7 x 5.2 a 11 x 6.8 x 1.4 nm, la apariencia tridimensional de esta molécula por microscopía electrónica recuerda una letra "C".

### 3c. Formación del Complejo Terminal

En la fase terminal de la cascada del complemento participan cinco proteínas homólogas que dan origen al TCC, formado por más de 20 subunidades. Esta fase de la vía puede dividirse en dos etapas: 1. Etapa de inserción, incluye la formación del complejo C5b-7 el cual contribuye con funciones de marcado de la superficie activadora y de activación celular mediada por proteínas G. 2. Etapa de permeabilización de la membrana, incluye la unión de C8 y la polimerización de C9, lo que puede ocasionar la lisis de algunas células (bacterias, eritrocitos), mientras que en las células propias la permeabilidad de la membrana permite la entrada de  $Ca^{++}$ , lo que resulta en la amplificación de vías de traducción de señales iniciadas previamente por C5b-7 (Müller-Eberhard, 1988).

El ensamble del TCC comienza en el medio hidrofílico cercano a membranas u otras superficies activadoras, donde la C5-convertasa se une y corta al C5. Este evento proteolítico libera un péptido llamado C5a con propiedades de anafilatoxina, mientras que el C5b se une al dominio C3b de la convertasa y retiene actividad aceptora de C6 durante 2.3 minutos (a 37 °C). El complejo C5b-C6 es hidrosoluble y se mantiene frágilmente unido a C3b en la superficie membranal. A continuación se le une C7 que al entrar en contacto con su sitio de unión en C5b-C6, expone dominios que le permiten interactuar con la parte iónica de la bicapa membranal, en esta etapa el complejo pasa irreversiblemente de un estado hidrofílico a uno anfifílico. El complejo C5b-C7 es una estructura alargada de que sobresale 210 Å de la superficie membranal, tiene una base estrecha insertada 40 Å hacia dentro de la bicapa lipídica donde C6 y C7 forman el dominio hidrofóbico mientras C5b se ubica fuera del sitio de unión a membrana. C5b-C7 no causa daño a la célula pero la marca para un ataque futuro. C5b-C7 es liberado a la fase fluida si la superficie activadora no es una membrana fosfolipídica, pues el dominio de unión a membrana de C7 no tiene sustrato para su inserción. En la fase fluida C5b-C7 puede insertarse en células adyacentes y continuar con la cascada terminal, a este proceso se le llama lisis reactiva y es controlado por proteínas del suero (ver más adelante) (Müller-Eberhard, 1988; Podack, 1986a).

En la etapa de permeabilización de la membrana, el C8 se une al complejo receptor C5b-7 a través de los sitios de unión de las subunidades C8 $\beta$  y C5b. Posteriormente, el C8 se inserta en la región hidrofóbica de la membrana a través de C8 $\alpha$ . Es posible que en esta etapa se formen poros de 30 Å de diámetro cuando dos o más complejos C5b-8 se agregan en la superficie membranal. Aunque estos complejos alteran la integridad estructural de las membranas y permiten la entrada de iones bivalentes, generalmente no lisan células nucleadas.

El complejo C5b-8 se relaciona con el C9 de tres distintas maneras: 1. Como receptor de C9 dirigiendo la acción citolítica a la célula blanco 2. Facilitando la interacción de la primera molécula de C9 con el C8 $\alpha$ , promoviendo el acceso e inserción de mas C9 en la membrana y 3. Catalizando la inserción y polimerización circular de 12-21 moléculas de C9. La polimerización de C9 va acompañada de un aumento de su estructura  $\beta$  laminar, de modo que cada subunidad pasa de 8 a 16 nm de longitud y expresa dominios hidrofóbicos. El producto final de esta cascada es poli-C9 en forma de poro, unido firmemente a la unidad de polimerización C5b-C8 (Podack, 1982a). En la última etapa de la vía terminal se expresa la capacidad lítica del TCC dependiendo de las características de la membrana blanco (Müller-Eberhard, 1986; Podack, 1985).

Existe evidencia de que durante el proceso de ensamble del TCC se forman enlaces disulfuro entre moléculas de C9 (Ware, 1981), de modo que en cada TCC existen aproximadamente 4 moles de C9 monomérico por cada mol de C9 dimérico. Aparentemente la formación de dímeros de C9 depende de que la polimerización sea catalizada por C5b-C8 (Yamamoto, 1982). Se ha propuesto que estos enlaces disulfuro tienen que ver con la estabilidad estructural y las propiedades líticas del TCC. Por ejemplo, se ha observado que los complejos hemolíticamente inactivos de C5b-C9 unidos a la proteína inhibidora S, no contienen dímeros de C9. Esto se debe probablemente a que la proteína S no permite que se incorporen mas de 4 moléculas de C9 al complejo.

### 3d. Características del TCC

El complejo terminal del complemento tiene un peso molecular de  $1.2-1.3 \times 10^6$  kDa y muestra un poro formado por C6-C7-C8 $\alpha$ - $\gamma$ - C9<sub>10-18</sub> y un apéndice constituido por C5b y C8 $\beta$ , que se disocia del TCC mediante ebullición en SDS (Podack, 1984b). A pesar de que la incubación del TCC en SDS a 100 °C deja intacta la estructura tubular, el borde de ésta se

despliega hasta alcanzar los 22 nm de longitud. Las dimensiones del TCC se muestran en la Figura 5. Estudios con marcadores fotoactivables hidrofóbicos han demostrado que el anclaje del TCC en la membrana está dado fundamentalmente por C8 $\alpha$  y C9 (Steckel, 1983; Hu, 1981; Podack, 1981). El poro es permeable a solutos como: iones monovalentes y bivalentes, colorantes, sacarosa, etc. El diámetro funcional del TCC vá de 30-70 Å dependiendo del número de moléculas de C9 que contenga (Young, 1986, Podack 1986a ) Dicho diámetro permite que enzimas como la lisozima (30x30x45 Å) puedan ingresar a la célula donde se ha insertado el TCC.

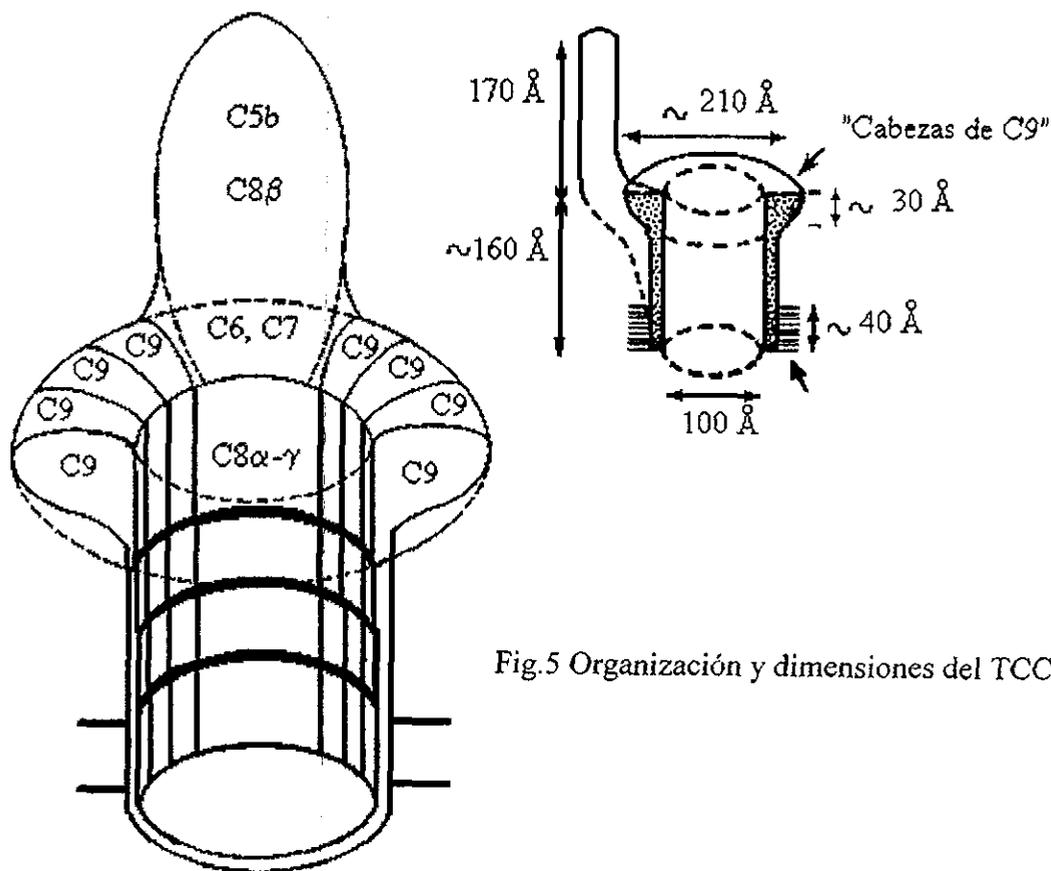


Fig.5 Organización y dimensiones del TCC.

### 3e. Regulación de la vía terminal.

El requerimiento de señales específicas de activación y la corta vida media de C5b-C9n son algunos de los mecanismos de regulación intrínseca en la formación del TCC. También existe un grupo de moléculas que regulan las etapas críticas de formación del TCC en la fase fluida y en la membrana de las células propias. Por ejemplo, los lípidos del suero compiten por los sitios

hidrofóbicos de las moléculas del TCC disminuyendo las posibilidades de unión a la membrana. Un efecto similar resulta de la unión del C8 al C5b-7 en solución. Otro regulador en la fase soluble es la proteína S o vitronectina. Esta molécula es parte de una familia de proteínas de adhesión, forma el 0.1-0.5 % del total de proteínas plasmáticas y también se le encuentra en el líquido amniótico y la orina (Preissner, 1991). Se sintetizada principalmente en los hepatocitos, pero las plaquetas, megacariocitos, monocitos y macrófagos, expresan una molécula con propiedades inmunoquímicas de proteína S. La proteína S también compite por los sitios de unión a membrana del complejo C5b-7 (Milis, 1993). Aunque SC5b-7 puede formar SC5b-8 y SC5b-9 (al unir C8 y C9 en solución), estos complejos carecen de propiedades líticas (Milis, 1993).

La incapacidad del C5b-9 para lisar eritrocitos y células nucleadas propias se conoce como restricción homóloga, y se debe principalmente a los factores DAF y CD59. Estas moléculas están ancladas a la membrana celular mediante un enlace glicosilfosfatidil-inositol. Como ya se había mencionado, el DAF controla la formación de las convertasas de C3/C5 (Scheren, 1993) mientras que el CD59 regula el ensamble del TCC.

El CD59 es una glicoproteína de 18-20 kDa presente en la superficie de células hematopoyéticas y de algunas no hematopoyéticas tales como células epiteliales, endoteliales, espermatozoides, células de riñón, de cerebro (astrocitos) y de tiroides entre otras. Además se ha descrito una forma soluble de CD59 en orina, saliva, lágrimas, sudor, líquido cefalorraquídeo, leche, líquido amniótico y plasma seminal, aunque no se sabe si estas formas son secretadas o se separan de la membrana por algún otro mecanismo (Lachman, 1991).

El CD59 inhibe la formación del TCC ya que actúa en la etapa de incorporación del C8 y C9 para formar C5b-9. El CD59 se une a C5b-8 e interfiere con la unión y subsecuente polimerización del C9 dentro del TCC. Los sitios específicos de unión han sido identificados en C8 $\alpha$  y el dominio b de C9. Estos sitios son los que participan en la interacción entre C8 y C9, siendo el dominio C9b el que se une a la membrana (Davies, 1993). En presencia de CD59, el número de moléculas de C9 incorporadas al TCC se restringe a 1.5 (C9/C8), en vez del 3.5 normal. La única molécula de C9 unida en estas condiciones se incorpora al complejo CD59-C5b-8 pues no puede ser desplazada por otras moléculas de C9 (Meri, 1990). Posteriormente, el complejo CD59-C5b-8 es eliminado de la superficie celular por mecanismos que se detallan en la sección "Efectos del TCC".

A diferencia de otros reguladores de la cascada del complemento, el CD59 no se relaciona estructuralmente y no está codificado por un gen situado en la región RCA del cromosoma 1 donde se ubican los genes para los reguladores del complemento. El CD59 muestra un 25% de similitud con la molécula Ly-6 humana y su gen ha sido clonado en rata, cerdo y una amplia variedad de géneros de primates. Además de ser un factor de regulación homóloga, el CD59 ha sido implicado en el proceso de transducción de señales y proliferación de células T.

### 3f. Efectos del TCC

Durante el último cuarto del siglo pasado, diversos investigadores describieron una actividad lítica sensible al calor, presente en el suero, que destruía bacterias y eritrocitos. Estas observaciones llevaron al descubrimiento del sistema del complemento. Hasta la década de los sesenta se identificó que la actividad lítica es mediada por el TCC formado durante la etapa final de la cascada del complemento. En este contexto, no resulta sorprendente que hasta hace poco tiempo el TCC fuera considerado por la mayoría de los inmunólogos como una entidad cuya única función era lisar bacterias y otras células (Ross, 1986).

Sin embargo, cuando los complejos C5b-7, C5b-8 y C5b-9 se encuentran en cantidades reducidas, muestran una capacidad lítica limitada, y en cambio inducen una gran variedad de actividades biológicas en las células propias, que incluyen la producción de eicosanoides, radicales libres de oxígeno, factor de necrosis tumoral e interleucina 1, así como un aumento en la colágena y colagenasas. También se induce hidrólisis de la proteína básica de la mielina, activación de plaquetas, activación de la eliminación de los TCCs potencialmente líticos e inducción de mitosis. El concepto de que el TCC ocasiona respuestas celulares distintas a la lisis comienza a ganar amplia aceptación gracias a evidencia obtenida en células nucleadas y metabólicamente activas, aunque también se conocen cambios no letales en células no nucleadas.

Se sabe que una sola lesión funcional del TCC puede ocasionar la lisis de un eritrocito y durar varios días insertada en una membrana. Sin embargo, esto no sucede en células nucleadas donde las lesiones son transitorias y consecuentemente se requieren muchos poros funcionales para ocasionar la lisis. Las células nucleadas poseen mecanismos de escape no presentes en eritrocitos que permiten contender con un ataque a membrana (Morgan, 1992).

La primera célula donde se describieron estos mecanismos de escape fue el neutrófilo humano. En estas células durante los ataques no letales de la cascada del complemento la vida media del TCC en membrana es de 3.5 minutos a 37°C (Campbell, 1985). En el proceso de escape se forman y desprenden de la célula vesículas membranales que contienen los TCC. Se ha estimado que en promedio 25,000 TCCs pueden formarse en una célula sin que ocurra lisis, ya que el 65% de estos complejos son rápidamente eliminados en vesículas membranales. Dichas vesículas constituyen sólo el 2% de la superficie celular, lo que implica un proceso selectivo de acumulación del TCC, en una pequeña región de la superficie celular. Por otra parte, una pequeña proporción de los TCC unidos a la célula son endocitados y degradados (Morgan, 1987). La eliminación endocítica de TCC implica la acumulación de complejos en pozos cubiertos de clatrina<sup>†</sup>, así como su internalización y degradación en lisosomas (Carney, 1986). Los procesos de desprendimiento del TCC por ectocitosis y/o endocitosis ha sido también demostrados para oligodendrocitos, células de epitelio glomerular, monocitos, entre otras.

El proceso de eliminación del TCC es dependiente del C8 y del C9. Pequeñas cantidades de vesículas pueden formarse en ausencia de C9, lo que sugiere que C5b-8 es suficiente para causar esta respuesta. Esta evidencia y a la disminución de los procesos de escape en ausencia de  $Ca^{++}$  extracelular, han llevado a proponer que los mecanismos específicos de escape en células nucleadas y su recuperación de un ataque por TCC dependen de procesos y señales intracelulares, inducidos por la entrada de  $Ca^{++}$  y por la liberación de  $Ca^{++}$  en compartimientos intracelulares (Morgan, 1987).

La mayor parte de las consecuencias de la elevación del  $Ca^{++}$  intracelular son mediadas por calmodulina<sup>§</sup> e implican la modulación de la actividad de protein-cinasas que regulan procesos celulares. La activación de las protein-cinasas mediada por TCC se ha demostrado en diversos tipos de células nucleadas. Adicionalmente, se ha observado que si estas enzimas son inhibidas, la lisis celular aumenta, lo que sugiere que el proceso de escape y recuperación está mediado al menos parcialmente a través de protein-cinasas (Carney, 1990). Otros estudios han implicado al c-AMP y su efecto modulador de protein-cinasas en los procesos de recuperación y escape de las células nucleadas (Campbell, 1985). El aumento en la concentración de este

<sup>†</sup> Proteína que constituye una especie de armazón que facilita la formación de vesículas endocíticas y exocíticas

<sup>§</sup> Proteína por su capacidad de unir calcio participa en numerosos procesos de regulación celular donde están

segundo mensajero puede ser ocasionada como resultado de la activación por  $\text{Ca}^{++}$ -calmodulina de enzimas involucradas en la producción de c-AMP. Un segundo mecanismo que ocasiona la elevación del c-AMP es el contacto directo del TCC con proteínas G. Los inhibidores de las proteínas G disminuyen la producción de c-AMP, diacilglicerol (DAG) y aumentan la lisis de células nucleadas atacadas por el TCC, lo cual implica que la función de estas enzimas es significativa para la respuesta a un ataque subléptico por complemento (Daniels, 1990). Se han coimmunoprecipitado las proteínas  $\text{G}_{i\alpha}/\text{G}_{o\alpha}$  en asociación con el TCC en células B linfoblastoides humanas. La actividad de estas GTPasas se observa inicialmente al insertarse el complejo C5b-7 y se incrementa con C5b-8 y C5b-9 (Niculescu, 1994). En células ascíticas de Ehrlich también se ha propuesto que el contacto del complejo C5b-7 con proteínas G activa cascadas donde se produce c-AMP<sup>84</sup>. Presumiblemente, la eficiencia relativa del complejo C5b-7 para estimular la iniciación de estas vías, depende de la no inhibición por el factor CD59. Por otra parte el complejo C5b-7 en células B linfoblastoides genera la liberación de DAG y ceramida, que son reguladores endógenos de la protein-quinasa C (PKC) que está relacionada con la liberación de  $\text{Ca}^{++}$  de reservas internas (Nicholson-Weller, 1993).

El ataque no letal por TCC además de desencadenar procesos de escape, activa vías de traducción de señales que culminan con la liberación de factores proinflamatorios. En consecuencia, los ataques sub-lépticos por complemento son relevantes para la patogénesis de la inflamación *in vivo* (Campbell, 1985; Daniels, 1990). Una vez que el  $\text{Ca}^{++}$  entra a la célula puede iniciar la liberación de más  $\text{Ca}^{++}$  de las reservas internas. Este proceso se puede dar por dos vías: El  $\text{Ca}^{++}$  externo puede activar a la fosfolipasa C (PLC) y permitir la generación de IP<sub>3</sub> que se une a receptores en las reservas internas de  $\text{Ca}^{++}$  y promueve su liberación; otra posibilidad es que el  $\text{Ca}^{++}$  extracelular se una a los receptores de IP<sub>3</sub> y reduzca su umbral de activación, de modo que se dispare la liberación de  $\text{Ca}^{++}$  intracelular. El aumento del  $\text{Ca}^{++}$  intracelular está relacionado con la producción de DAG que estimula a la PKC que a su vez activa a la fosfolipasa A<sub>2</sub> para iniciar las vías de metabolismo y liberación de mediadores de la reacción inflamatoria como prostaglandinas, tromboxanos, leucotrienos, metabolitos tóxicos de oxígeno e IL-6. Se ha demostrado que la liberación de estos factores en células reumatoides sinoviales humanas no solo

---

implicadas enzimas cuyo funcionamiento depende de calcio citoplasmica de eucariontes.

depende de la presencia de  $Ca^{++}$  extracelular sino también de la iniciación de síntesis proteica *de novo* (Daniels, 1990). Al igual que en los mecanismos de eliminación del TCC, también entre las vías que inician la producción de factores proinflamatorios las proteínas  $G_i$  están involucradas de manera importante (Niculescu, 1994). En la Figura 6 se esquematizan los mecanismos de transducción de señales activados por la formación del TCC, en la Tabla IV se enlistan algunos efectos no líticos del TCC y su dependencia de  $Ca^{++}$ .

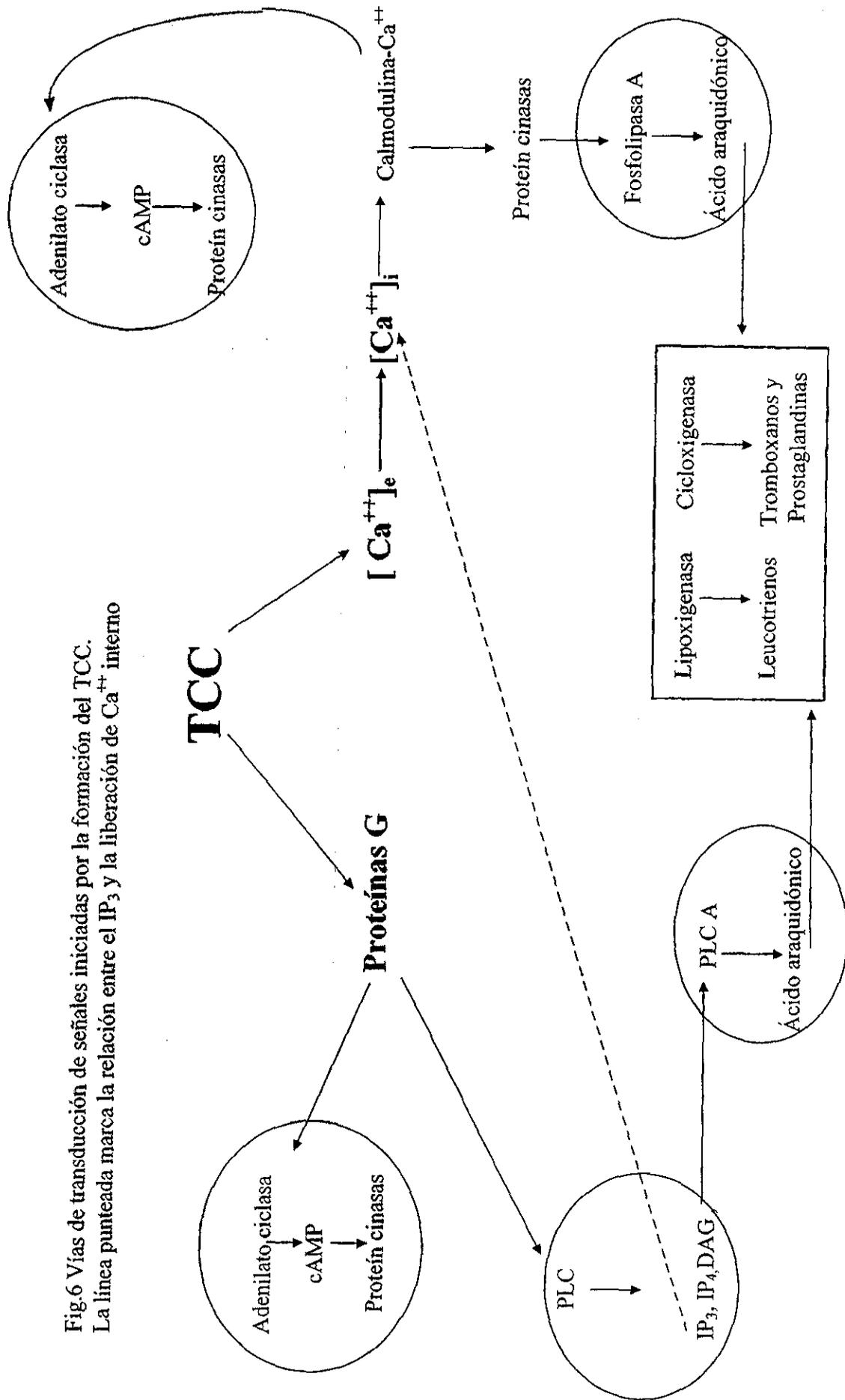
Tabla IV. Efectos no letales del TCC y su relación con  $Ca^{++}$ . Morgan, 1990b.

Tipo Celular	Efectos	Dependencia de $Ca^{++}$
Neutrófilo (rata)	Producción de ROM y $LTB_4$	✓ ✓
Neutrófilo (humano)	Vesiculación	✓
	Producción de ROM y $LTB_4$	✓ ✓
Plaqueta (humano)	Activación de protrombinasa	✓
	Vesiculación	✓
	Producción de $TXB_2$	✓ a
Monocito/macrófago (humano)	Producción de $PGE_2$ y $TXB_2$	✓
	Producción de ROM	✓
Célula del epitelio glomerular (rata)	Producción de $PGE_2$ y $TXB_2$	NI
	Vesiculación	NI
	Activación de PL y producción de $IP_3$	✓
Célula glomerular mesangial (rata)	Producción de $PGE_2$ e $IL-1$	NI
	Producción de ROM	NI
Oligodendrocito (rata)	Producción de $LTB_4$	NI
Sinoviocito (humano)	Producción de ROM	✓
	Producción de $PGE_2$	NI
Líneas celulares tumorales	Producción de $LTB_4$	✓
	Vesiculación	✓

a : No se requiere  $Ca^{++}$  intracelular pero la liberación de  $Ca^{++}$  de reservas internas es necesaria  
 Abreviaturas: ROM: Metabolitos reactivos de oxígeno;  $LTB_4$ : Leucotrieno  $B_4$ ;  $TXB_2$ : Tromboxano  $B_2$ ;  $PGE_2$ : Prostaglandina  $E_2$ ;  
 PL: Fosfolipasa;  $IL-1$ : Interleucina 1; NI: No investigado.

Las principales consecuencias de la inserción del TCC en cantidades líticas en las membranas de células bacterianas y eritrocitos son la reorganización de la bicapa lipídica y el incremento del área total de membrana, lo que tiene un efecto negativo en la conservación de la integridad estructural y funcional de la membrana blanco (Müller-Erberhard, 1986). Sin embargo, estas reacciones pueden variar dependiendo de si la membrana es bacteriana, viral, o de eritrocito. También, estos poros permiten la degradación de la capa de peptidoglicanos bacterianos por

Fig.6 Vías de transducción de señales iniciadas por la formación del TCC. La línea punteada marca la relación entre el  $IP_3$  y la liberación de  $Ca^{++}$  interno



lisozima, pérdida de elementos intracelulares menores a 10 nm, dispersión del potencial eléctrico de membrana, hinchamiento osmótico de las células, desplazamiento de los componentes membranales y activación inespecífica de vías intracelulares de señalamiento por entrada de calcio (Podack, 1986a). Estos efectos se presentan en relación directa con la cantidad de TCC insertado y a las características intrínsecas de la membrana. El gasto energético que implican estas vías ocasiona el agotamiento de las reservas de fosfatos de alta energía y algunas alteraciones metabólicas. La muerte celular es inevitable si no se ponen en marcha los mecanismos de reparación que contrarresten las lesiones y sus efectos (Podack, 1986a). Sin embargo, se han descrito mecanismos de evasión al ataque por TCC homólogo en eritrocitos (Iida, 1991). El escape de los eritrocitos por vesiculación, también es dependiente de  $Ca^{++}$  extracelular (Halperin, 1993). Se considera que el desenlace de un ataque por complemento homólogo en eritrocitos depende de la cantidad de complejos líticos que se inserten en la membrana. De este modo se puede distinguir entre un ataque sublítico que activará mecanismos de escape y un ataque que por su magnitud resulta lítico (Halperin, 1988).

La mayoría de los virus con cubierta lipoproteica son susceptibles a lisis mediada por complemento con o sin la intervención de anticuerpos. El ataque mediante TCC ocasiona la fragmentación de la cubierta viral, pérdida de la identidad estructural de la bicapa lipídica, y desintegración de la nucleocápside (Cooper, 1986).

## II. ANTECEDENTES DEL PROYECTO

*The history of life is a story of massive removal followed  
by differentiation within a few surviving stocks,  
not the conventional tale of steady increasing excellence,  
complexity and diversity.  
Prefacio. Wonderful Life.  
Stephen Jay Gould.*

### 1. Evolución molecular de la vía terminal

Existen dos líneas de explicación acerca de la evolución de la vía terminal; ambas aceptan la existencia de un gen ancestral común para perforinas y para las proteínas del TCC (C6-C9). El primer planteamiento está basado en la homología molecular, estructural y funcional entre perforinas y las proteínas de la vía terminal del complemento, especialmente C9. Se sugiere que la perforina y el C9 surgieron inicialmente a partir del gen ancestral (Morgan, 1990c). Las otras proteínas del TCC (C8 $\alpha$ , C8 $\beta$ , C7 y C6) se originaron posteriormente a través de duplicaciones génicas e inserción de módulos. La segunda hipótesis basada en observaciones sobre los límites de intrones/exones de los genes C6-C9, sugiere que la proteína ancestral de la vía terminal era una molécula parecida a C6/C7 (Hobart, 1993; 1995).

Cuando una célula con actividad citotóxica como las células NK y los linfocitos T citotóxicos se une a su célula blanco algunos gránulos citoplásmicos que contienen perforina, granzimas, fragmentinas, leucalexina y TIA-1\* se liberan en el área de contacto. La perforina al unirse al Ca<sup>++</sup> expone sitios de inserción a membrana, así como regiones de contacto con otras moléculas de perforina. Este cambio conformacional inicia un proceso de polimerización que culmina en la formación de un poro de 3-20 monómeros, que además de incrementar la permeabilidad membranal y ocasionar lisis osmótica, facilita la entrada de otros factores citotóxicos que inician un programa de apoptosis en la célula blanco. Aunque la función citolítica de linfocitos T y células NK puede darse a través del antígeno Fas, la expresión de la perforina es fundamental para el mecanismo de citólisis por exocitosis de gránulos (Liu, 1995).

---

\* Otras proteínas contenidas en los gránulos en los que se secreta la perforina. Las granzimas y las fragmentinas forman parte de una familia de serin esterases. TIA-1 es una enzima aislada de células NK que induce la fragmentación de DNA.

Los componentes de la vía terminal C6-C9 son homólogos a la perforina<sup>†</sup>. A nivel de secuencia comparten una zona de aproximadamente 200 aminoácidos en el extremo carboxilo terminal. Esta región de la perforina contiene 159 residuos equivalentes a la región pobre en cisteínas, donde supuestamente se encuentra la región de unión a membrana (Liu, 1995), además de 37 aminoácidos del módulo EGF (ver sección I 3a). Las seis cisteínas conservadas en este último módulo pueden formar enlaces disulfuro intramoleculares que contribuyen a mantener un dominio funcional en perforina y C6-C9 (Shinkai, 1988). En la región de 168 residuos de longitud que puede alinearse con más certeza, las proteínas del complemento y la perforina muestran aproximadamente un 21% de similitud. En cambio, los residuos 1-210 y 407-555 (numeración de perforina humana) de la perforina no son homólogos a las proteínas del complemento. La estructura primaria de las perforinas es altamente conservada, por ejemplo, la perforina de rata tiene 69.3% y 84.9% de identidad y similitud<sup>‡</sup> a las perforinas humana y de ratón, respectivamente (Ishikawa, 1989).

La similitud de las proteínas del complemento y la perforina también se ha demostrado a nivel de sus determinantes antigénicos. Anticuerpos policlonales generados contra perforina y C9 puras, reaccionan en forma cruzada; estos resultados apoyan la existencia de una homología estructural y de secuencia entre las dos moléculas (Young, 1986). En otros experimentos con anticuerpos contra C9 y perforina, también se han observado reacciones cruzadas entre sus antígenos y contra los complejos C5b-7, C5b-8 y C5b-9. Asimismo anticuerpos contra regiones conservadas del C9 también reaccionan con C8 $\alpha$ , C7, C6 y perforina (Tschopp, 1986).

Mediante análisis de hidropatía se concluyó que la perforina de ratón es una proteína predominantemente hidrofílica que posee un dominio hidrofóbico flanqueado por zonas altamente hidrofílicas entre las posiciones 191-251. Adicionalmente, su longitud le permite atravesar la membrana al menos dos veces; por lo que esa región parece ser un dominio transmembranal. Las moléculas C7, C8 y C9 presentan zonas anfipáticas cortas que podrían tener la misma función (Kwon, 1989). Estudios posteriores de identificación de los sitios de contacto del C9 con la membrana, así como de modelaje molecular, sugieren que los sitios de interacción de C6-9 y de la

---

<sup>†</sup> Shinkai 1988, Haefliger 1989, Rao 1987, Howard 1987, Haefliger 1987, DiScipio 1984, Liu 1995, Chakravarti 1989.

<sup>‡</sup> Identidad a nivel de secuencias de aminoácidos alineadas se define como el porcentaje de residuos iguales que existen entre dos secuencias, mientras que similitud es el porcentaje de residuos iguales y de la misma polaridad que comparten ambas secuencias.

perforina con membranas, son homólogos y están formados por una hélice  $\alpha$ , seguido por un giro y una segunda hélice  $\alpha$  no presente en todas las proteínas analizadas. En las moléculas de perforina y C9 humanas estas estructuras se encuentran entre los residuos 135-223 y 266-345 (Peitch, 1990). El resultado coincide con los obtenidos con perforina por análisis de hidropatía. Sin embargo, pruebas de la capacidad lítica de los distintos péptidos de perforina en liposomas han revelado que la región fundamentalmente implicada en la actividad lítica reside en los primeros 19 aminoácidos y forman una estructura  $\beta$  laminar parcialmente anfifílica, no similar a ninguna región de C6-C9. Los péptidos de los residuos 189-218 de perforina y los correspondientes 311-340 de C9 pueden también adoptar una conformación de hélice  $\alpha$  y formar poros *in vitro*, sin embargo tienen un papel secundario en la función lítica de perforina (Perechini, 1992).

La capacidad del C9 para formar poros citolíticos *in vitro*, ha permitido compararla funcionalmente con la perforina. El peso molecular de ambas proteínas es de 70-75 kDa. Cuando la perforina y el C9 polimerizan en bicapas lipídicas pueden formar estructuras tubulares con un diámetro interno de 100 y 160 Å, respectivamente.

La polimerización del C9 puede acelerarse mediante  $Zn^{++}$ ; se sabe que este ión facilita la unión del C9 a vesículas lipídicas pues induce la exposición de dominios hidrofóbicos. La ausencia de  $Ca^{++}$ , también afecta la polimerización del C9, sin embargo, en presencia de C5b-8 la polimerización del C9 ocurre aún en ausencia de cationes divalentes (Podack, 1986b). En forma similar, el  $Ca^{++}$  es esencial e insustituible en la formación de poros por perforina.

Las estructuras constituidas por C9 y perforina son canales estables que permiten el paso de agua y iones. Los canales únicamente son selectivos al tamaño de los solutos y no son sensibles a cambios en el potencial de membrana (Young, 1986). Sin embargo, el C9 sólo polimeriza *in vivo* en presencia del complejo receptor C5b-8, mientras que la perforina no requiere de una estructura equivalente (ver arriba). Los mecanismos de procesamiento y transporte postraduccionales son diferentes en ambas moléculas, puesto que el C9 es liberado al plasma y la perforina es almacenada en gránulos citoplásmicos.

En forma similar a lo que ocurre durante la vía terminal del complemento, la formación de poros de perforina también está controlada por lipoproteínas y proteína S del suero. El mecanismo de control de la perforina y del C9 puede ser a través del bloqueo de sus regiones

cargadas negativamente. Estos sitios son necesarios para la formación de los poros líticos. A pesar de que existen mecanismos comunes de regulación para perforina y C6-9 en el suero, no sucede lo mismo en la superficie membranal. La falta de factores de restricción homóloga para perforina se explica por la función de las células citotóxicas, que al atacar a las células blanco interactúan específicamente mediante receptores. La manera en que las células citotóxicas se protegen de su propia perforina, es a través de moléculas de adhesión en las vesículas secretoras que aseguran la unidireccionalidad de la liberación de perforina (Podack, 1992).

En resumen, la primer línea de explicación acerca de la evolución de la vía terminal del complemento se basa en la similitud bioquímica y funcional de la perforina y de las proteínas C6-C9. Esta explicación propone que el C9 y la perforina surgieron a partir de la duplicación de un gen ancestral común, que tenía una estructura similar a la perforina actual. Las otras proteínas del TCC (C8 $\alpha$ , C8 $\beta$ , C7 y C6) se originaron posteriormente a través de duplicaciones génicas e inserción de módulos. Este orden de aparición es retrógrado con respecto al orden en que actúan en la cascada del complemento y esta sustentado en la capacidad del C9 para polimerizar *in vitro*. Puesto que C9 es capaz de formar poros sin la intervención de los demás elementos de la vía terminal (ver sección I 3b), es factible que a partir de la duplicación de este gen se originaran las proteínas C8, C7 y C6, que facilitan su función lítica. Posteriormente, este sistema lítico se asocia a la cascada del complemento que tenía únicamente funciones de marcado y opsonización. Esta asociación se dio a través del C5, que surgió de una duplicación génica de C3 (Stanley, 1988). La aparición de moléculas con funciones especializadas, hizo que el C9 perdiera su capacidad de unirse a membranas (Morgan, 1990c). En resumen, este orden de aparición propone que las proteínas del complemento surgieron a partir de proteínas modularmente simples y su evolución se dirigió hacia la formación de proteínas mas complejas (Fig. 7).

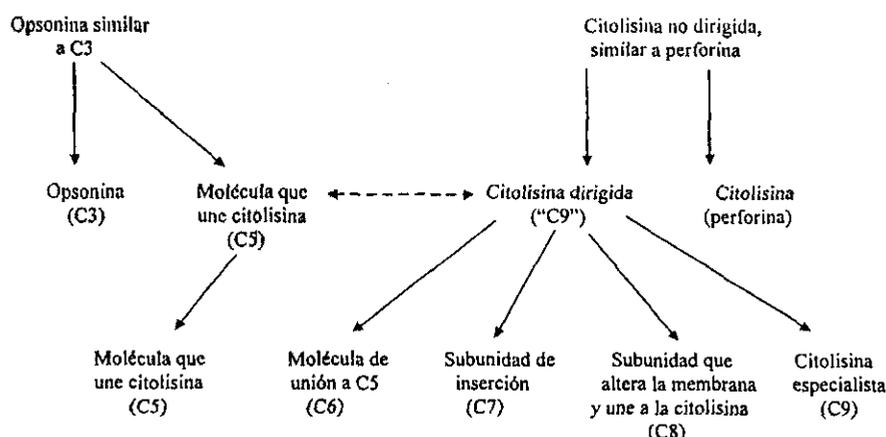


Fig.7 Evolución de los componentes terminales del complemento deducida a partir de su estructura modular. La línea punteada marca la unión entre el sistema primitivo de opsoninas y la naciente cascada terminal mediante la interacción de C5 y "C9". Dibujo adaptado de Morgan 1990c .

Recientemente se ha planteado una segunda hipótesis acerca de la evolución del TCC. Observaciones sobre los límites de intrones/exones de los genes C6-C9 han sugerido que la proteína ancestral de la vía terminal era una molécula parecida a C6/C7 (Hobart, 1993; 1995). Al respecto, se han definido como intrones de fase 0 a los que se inician en el primero nucleótido de un codón, los intrones de fase 1 se inician en el segundo y los de fase 2 se inician en el tercer nucleótido del codón. Un exón simétrico es aquél que tiene intrones de la misma fase de lectura en ambos extremos, de esta manera hay exones simétricos del tipo 1-1, 2-2 ó 0-0. Los exones simétricos poseen características que facilitan su intercambio entre genes ya que solo esta clase de exones pueden duplicarse en tándem, ser eliminados o insertarse por recombinación intrónica sin que se altere el marco de lectura. En algunas proteínas mosaico hay módulos que se ubican en exones simétricos. Consecuentemente, en genes ensamblados a partir de módulos los intrones de una sola fase pueden predominar. La presencia no aleatoria de intrones de una misma fase y la correlación de los límites de un dominio o módulo con los límites de un intron sugieren que algunos genes mosaico se ensamblan por intercambio de exones (Patthy, 1987). La segunda hipótesis acerca de la evolución de la vía terminal, basada en la estructura de intrones/exones de los genes humanos, propone que el orden de aparición de las proteínas homólogas del TCC fue el mismo en el que procede la vía terminal de la cascada: primero el C6 y el C7, seguidos por el C8 y el C9 (Hobart, 1993; 1995).

Los tipos de fase en intrones de los genes C6-C9 están conservados (Fig. 8). La única excepción es el intron entre los exones 5, 6 y 10, en extremo 3' del gen C9. Esta conservación incluye también la longitud de los exones de los cinco genes. Las excepciones están en ambos extremos donde las proteínas difieren más, así como en el exón 6 de la región CP. Lo anterior refuerza la idea de que estos genes tienen un ancestro común, con características que se pueden deducir de la información en genes actuales. Los límites de los exones que codifican las regiones ricas en cisteína están cerca de la mitad de los módulos protéicos. Un ejemplo es el primer módulo SCR del C6 codificado por dos exones. La única excepción de un módulo codificado en un exón es el segundo SCR de C6 y C7. Adicionalmente, la mayor parte de los exones de estos genes son asimétricos, lo que impide duplicaciones internas de exones o deleciones como un mecanismo de adaptación posterior a la divergencia del ancestro común (puesto que la mayoría de las duplicaciones internas o deleciones cambiarían la fase de lectura del mensajero).

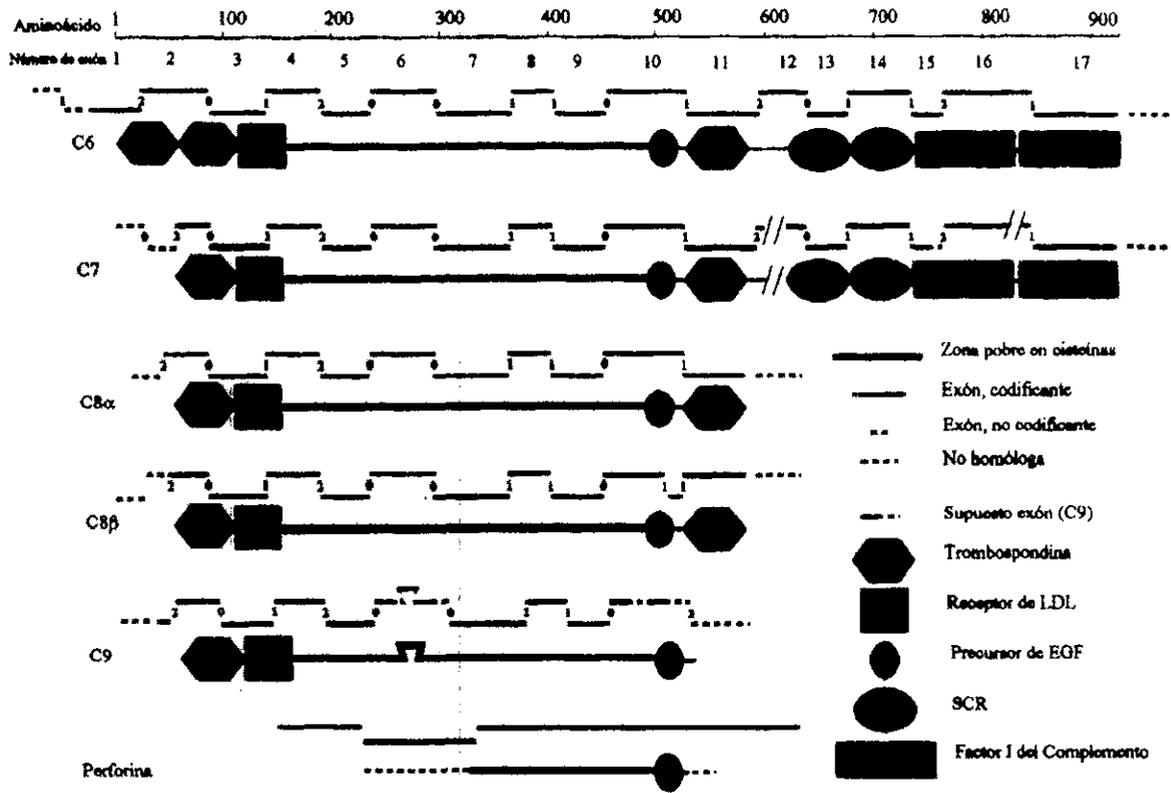


Fig.8 Relación entre la estructura de exones y módulos de las proteínas C6-C9 humanas. Las líneas sobre los esquemas de cada proteína representan la longitud y fases de los exones que conforman los genes de estas proteínas. Figura adaptada de Hobart 1995.

El extremo 3' del exón 2 de C6, C7 y C8α, C8β y C9 codifica para la porción amino terminal de un módulo TSP, que es el segundo en el C6 y el primero en C7, C8α, C8β y C9 (Fig. 8). Esta región es altamente conservada y presumiblemente tiene un origen común. En C6, el extremo 5' del exón 2 codifica para la porción carboxilo terminal del primer módulo de trombospondina, mientras que en C7, C8α, C8β y C9, el extremo es más corto y codifica para segmentos protéicos cuya estructura y función son desconocidos. Este exón tiene en su extremo 5' un intron de fase 2 en los genes de C6-C9. Se ha propuesto que los cinco exones homólogos al exón 2 de C6, se derivan de un ancestro común mas parecido al de C6, que esta formado por una parte de dos módulos TSP en el extremo 5'. En contraste, en C7, C8α, C8β y C9, esa zona ha sufrido cambios internos. El módulo TSP amino terminal del C6 está ausente en C7, C8α, C8β y

C9. Esta diferencia es consistente con la idea de que el ancestro común de los genes terminales era más parecido al moderno gen de C6. Los genes de C7-C9 evolucionaron por cambios en los extremos de los genes duplicados. Esta idea es apoyada por el hecho de que el C9 de trucha y del pez globo tienen una estructura más similar al C8 $\alpha$  y C8 $\beta$  que al C9 humano. A partir de estas observaciones se dedujo que el ancestro común de los componentes terminales del complemento debió extenderse desde la región amino terminal del exón 1 del C6 hasta el exón 10 u 11, cerca del carboxilo terminal del C9 o C8. De nuevo, parece más probable que el ancestro común se extendiera hasta el extremo carboxilo del C8 puesto que el C9 de trucha y del pez globo son muy parecidos al C8 (Hobart, 1993; 1995).

La conservación de las fases de intrones/exones, muestra que muy pocos intercambios de módulos han ocurrido desde que la actual estructura de intrones/exones fue establecida. Probablemente éste patrón se fijó antes de que ocurrieran las duplicaciones que originaron los genes modernos de las proteínas del TCC. Adicionalmente, la conservación de la estructura de intrones/exones implica que la integridad de dicho arreglo es importante para su expresión. Las diferencias en la estructura modular entre los genes C6-C9 sugieren que por lo menos parte del proceso que ha guiado su divergencia a lo largo de la evolución, es a través de pérdidas de los extremos de los genes ancestrales más complejos, puesto que proceden de una estructura compleja hacia una simple. Este esquema reta la noción general de que la evolución de las proteínas procede de lo simple hacia lo complejo y de que un módulo protéico puede tener un valor predictivo en relación a la estructura del gen (Hobart, 1993; 1995).

Una descripción detallada de la evolución de las vías de activación del sistema del complemento se presenta en el Anexo 2.

### III. OBJETIVO

Se ha reportado más de una docena de secuencias de genes que codifican para proteínas C6-C9 de diversas especies. Sorprendentemente, hasta ahora no se ha reportado un análisis filogenético para determinar cuál de las dos teorías sobre la evolución de estas proteínas (descritas en la sección II) es la correcta. La información para determinar el orden de aparición evolutivo de estas proteínas esta contenida en las secuencias de aminoácidos y puede ser utilizada mediante métodos de inferencia filogenética. La similitud de C6-C9 con la secuencia de aminoácidos de la perforina permite emplearla como grupo externo, es decir, como el punto de referencia para establecer el orden de aparición de estas proteínas.

El objetivo del presente proyecto de tesis fue determinar el orden evolutivo de aparición de los genes C6, C7, C8 $\alpha$ , C8 $\beta$  y C9 por medio de métodos de inferencia filogenética.

## IV. MATERIALES Y MÉTODOS

### 1. Secuencias

En el presente estudio se emplearon los módulos TSP, LDLr, CP, y EGF de las secuencias de aminoácidos de C6, C7, C8 $\alpha$ , C8 $\beta$ , C9 de humano; C8 $\alpha$ , C8 $\beta$  y C9 de conejo; así como C9 de rata, ratón, caballo, *Fugu rubripes* (pez globo) y trucha arco iris (*Oncorhynchus mykiss*). Además se alinearon la región CP y el módulo EGF de las perforinas de rata, ratón y humano (Fig.9)<sup>§</sup>. Estas secuencias son todas las que se encontraban disponibles en GeneBank para las proteínas C6-C9 y perforina (hasta junio de 1998). Solo los módulos TSP, LDLr, EGF y la región CP región fueron incluidas en el análisis pues son las únicas regiones presentes en todas las proteínas empleadas y se pueden alinear de manera confiable. Se decidió emplear secuencias de aminoácidos en lugar de secuencias de nucleótidos con la finalidad de facilitar el proceso de alineación lo que disminuye el riesgo de obtener alineaciones espurias. Esto es especialmente crítico en la detección de los módulos ricos en cisteínas y otras zonas comparables entre las moléculas C6-C9 y perforina.

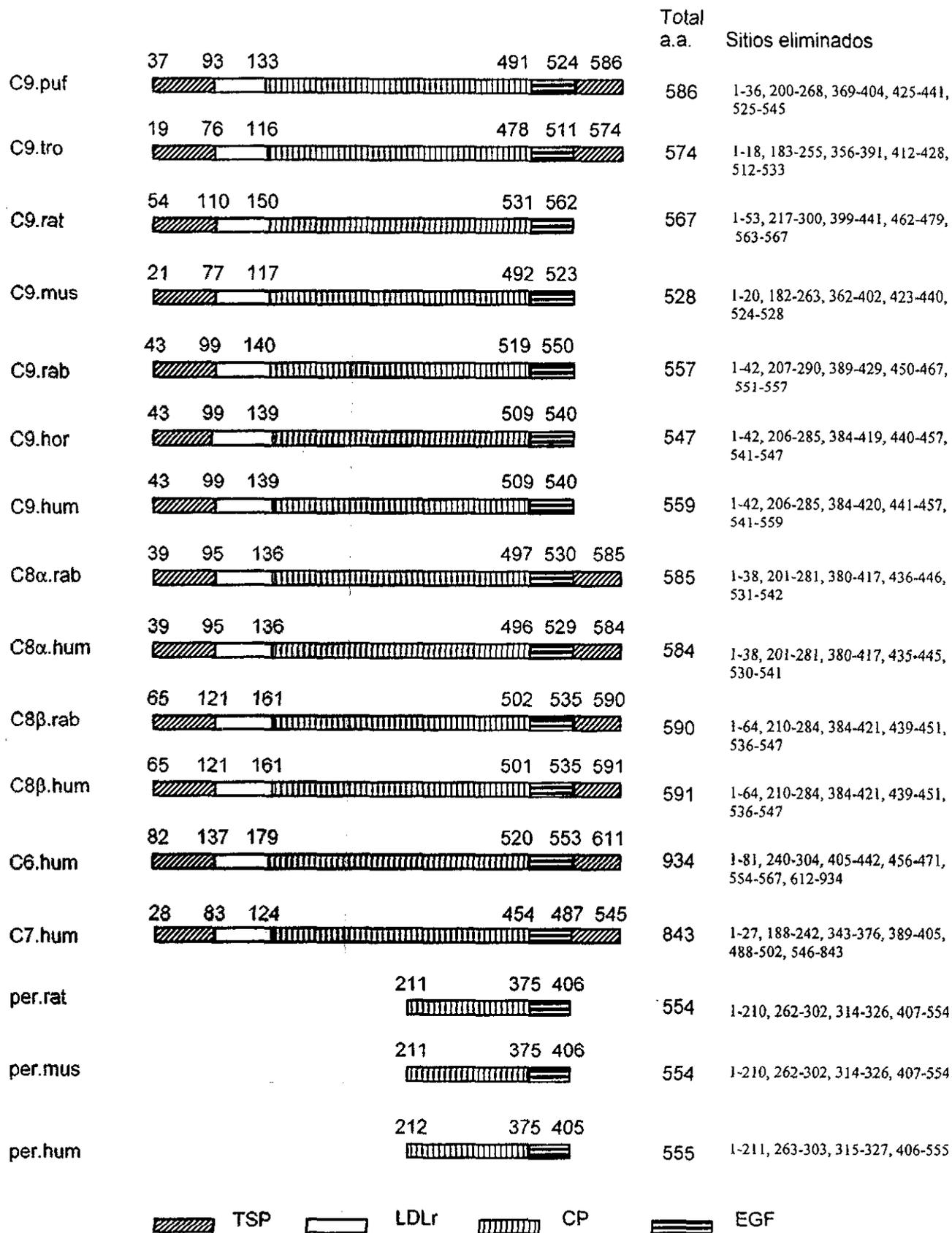
### Alineación

Las secuencias mencionadas arriba se alinearon inicialmente con el algoritmo Clustal (Higgins, 1988) implementado en DNAMAN (Lynnon BioSoft, 1994-97), posteriormente la alineación se refinó manualmente. Las regiones marcadas en la figura 9 fueron excluidas de la alineación por que la alta variabilidad en dichas zonas impide su alineación confiable y por ende, su utilidad en el análisis.

En el Anexo 3 se proporciona una explicación detallada de los métodos de inferencia filogenética empleados en este proyecto.

---

<sup>§</sup> Fig. 9 (En la siguiente página) Representación esquemática de las secuencias de aminoácidos de C6-C9 usadas en los análisis de inferencia filogenética. Los números sobre cada figura son las posiciones incluidas en los módulos. Los números a la izquierda de cada secuencia señalan el número total de residuos reportados para cada proteína, así como las posiciones que fueron excluidas por no generar una alineación confiable. Los números de cada secuencia se emplearon de acuerdo a las secuencias originales publicadas en GenBank.



 TSP  
  LDLr  
  CP  
  EGF

### 3. Proceso de inferencia filogenética

Un análisis con el método de Parsimonia con el algoritmo de "Branch-and-Bound" se llevó a cabo con 100 réplicas de "Bootstrap" con la versión 4.0.0d61a del programa PAUP (Swofford, 1990), se emplearon las perforinas como grupo externo. Las otras condiciones empleadas en PAUP fueron las preestablecidas (default) en el programa. (Tabla V)

Tabla V. Búsqueda con el método de Branch-and-Bound por Bootstrap

<b>Número de caracteres muestreados en cada réplica : 444</b>
<b>Réplicas: 100</b>
<b>Criterio de optimización: Máxima Parsimonia</b>
<b>Características de los caracteres empleados: 444 caracteres del tipo no ordenado con igual valor, 75 caracteres son constantes, 55 caracteres son variables y no informativos para parsimonia</b>
<b>Número de caracteres informativos para parsimonia: 314</b>
<b>Límite superior inicial: Calculado progresivamente (stepwise)</b>
<b>Adición de secuencias: Más próxima (Furthest)</b>
<b>Ramas de longitud igual a cero son colapsadas para crear politomías</b>
<b>Tiempo empleado: 95 h. 52 min. 48 seg.</b>
<b>Árbol obtenido por mayoría del 50%</b>

Análisis filogenéticos por métodos de distancia fueron realizados mediante búsquedas con el método de "Neighbor-Joining" implementado en PAUP, con un "Bootstrap" de 100 réplicas. En este caso la medida de distancia es la diferencia media entre los caracteres (Distancia P). El resto de las condiciones empleadas fueron las preestablecidas en el programa. (Tabla VI). En el cálculo de las distancias se empleó la Distancia P ya que es la única manera de obtener estos estimados en PAUP; además, el empleo de otros programas de inferencia filogenética (vgr. MEGA) no permite hacer comparaciones válidas entre los valores obtenidos en la prueba de bootstrap.

Tabla VI. Búsqueda con el método de Neighbor-Joining por Bootstrap

<b>Número de caracteres muestreados en cada réplica: 444</b>
<b>Réplicas: 100</b>
<b>Si se encuentran vínculos estos se rompen aleatoriamente</b>
<b>Medida de distancia: Promedio de caracteres diferentes</b>
<b>Tiempo empleado: 21.60 seg.</b>
<b>Árbol obtenido por mayoría del 50%</b>

#### 4. Prueba de la tasa relativa

Las tasas de cambio de C6-C9 y perforinas se evaluaron mediante la prueba de la tasa relativa (RRT) (Takezaki, 1995) aplicada a las distancias por corrección de Poisson que se obtienen al comparar entre sí todas las secuencias empleadas para los análisis filogenéticos, así como a la región CP y el módulo EGF por separado. Para el cálculo de distancias por corrección de Poisson que se emplearon en todos los RRT de este trabajo, no se consideraron las posiciones donde existen espacios. Se eligió la corrección de Poisson como medida de distancia por que este método supone que el número de sustituciones de aminoácidos en cada sitio sigue una distribución de Poisson. Este tipo de distribución considera que las mutaciones se presentan en el tiempo en forma aleatoria y con la misma probabilidad para cada aminoácido en todas las posiciones de una alineación. Es ventajoso emplear esta clase de medida ya que es un modelo que describe de manera más detallada las diferencias entre secuencias que el solo cálculo del promedio de sus diferencias (Distancia P).

Para realizar los cálculos del RRT se empleó el programa Phyltest (Kumar, 1996). En esta prueba,  $L_a$  y  $L_b$  son el número observado de sustituciones por sitio con respecto al ancestro común de un par de secuencias comparadas, por lo tanto si entre dos secuencias A y B existe una tasa de cambio constante desde su divergencia, se dice que entre ellas existe un reloj molecular, en este caso la hipótesis nula es que el número de sustituciones por sitio con respecto a un ancestro común son iguales para las dos secuencias ( $L_a = L_b$ ). En caso de que se cumpla la hipótesis nula, la diferencia es  $\delta = L_a - L_b = 0$ . La desviación de  $\delta$  de 0 se puede examinar mediante un two-tailed normal deviate test, donde la Z estadística no debe ser mayor a 1.96 para aceptar la constancia de las tasas de evolución a un nivel del 5%.

#### 6. Análisis de Hillis-Huelsenbeck.

Se realizó un análisis de Hillis-Huelsenbeck (Hillis y Huelsenbeck, 1992) a la topología del árbol obtenido por el análisis de distancia, con la finalidad de averiguar si la misma topología puede generarse aleatoriamente. El árbol al que se aplicó este análisis se obtuvo con el mismo método y parámetros ya explicados en "Análisis de distancia" con la diferencia de que no se hizo el "Bootstrap". Se evaluaron 10,000 árboles muestreados aleatoriamente de entre todas las topologías sin raíz que se pueden generar con la información disponible.

## V. RESULTADOS

### 1. Características de la alineación generada

Todas las moléculas conocidas de C6, C7, C8 $\alpha$ , C8 $\beta$  y C9 comparten los módulos TSP, LDLr, CP, y EGF. La alineación elaborada en base a estos módulos comunes cuenta con un total de 444 posiciones de las cuales 314 son informativas según el criterio de Parsimonia (ver Anexo 3). Las perforinas tienen en común con C6-C9 una parte de la región CP y el módulo EGF completo. Esta región homóloga comprende 168 posiciones totales, con 133 informativas, que se pueden alinear con confianza con C6-C9.

### 2. Análisis filogenético por Parsimonia

Inicialmente se llevó a cabo un análisis de inferencia filogenética por Parsimonia ya que además de ser un método ampliamente validado, considera un modelo de evolución con filogenias basadas en el menor número de supuestos, es decir, la explicación más sencilla o parsimoniosa es la que posiblemente describe la relación entre un grupo de secuencias o taxa. (para una explicación detallada de los métodos de inferencia filogenética ver Anexo 3).

El análisis por el método de Parsimonia con "Branch-and-Bound" resultó en un árbol donde la mayoría de los valores aceptables de "Bootstrap" superiores al 70% (Fig.10). Este árbol muestra que todas las proteínas del complemento tienen un origen monofilético, ya que se ubican en un grupo con un origen evolutivo común. Las proteínas de genes ortólogos\*\* forman cuatro grupos bien definidos: las secuencias de los C9 constituyen uno solo grupo; en otro clado se ubican por separado los C8 $\alpha$  y los C8 $\beta$ , mientras las secuencias únicas de C6 y C7 forman un tercer grupo. Finalmente, las perforinas se ubican en un cuarto grupo distinto al de todas las proteínas del TCC. Como era de esperarse, el arreglo interno de cada conjunto de proteínas ortólogas está de acuerdo con la filogenia de las especies de procedencia, con la excepción del C9 equino que aparece más cercano al C9 humano que el del ratón y el del conejo (ver Anexo 5).

Sin embargo, este árbol no resuelve el orden de aparición evolutiva de las proteínas C6-C9 que parecen emerger simultáneamente del nodo basal en la rama que agrupa a las proteínas del

---

\*\* Proteínas codificadas por genes ortólogos, que son loci en distintas especies cuyas secuencias son suficientemente similares como para sugerir que se originaron de un gen ancestral común.

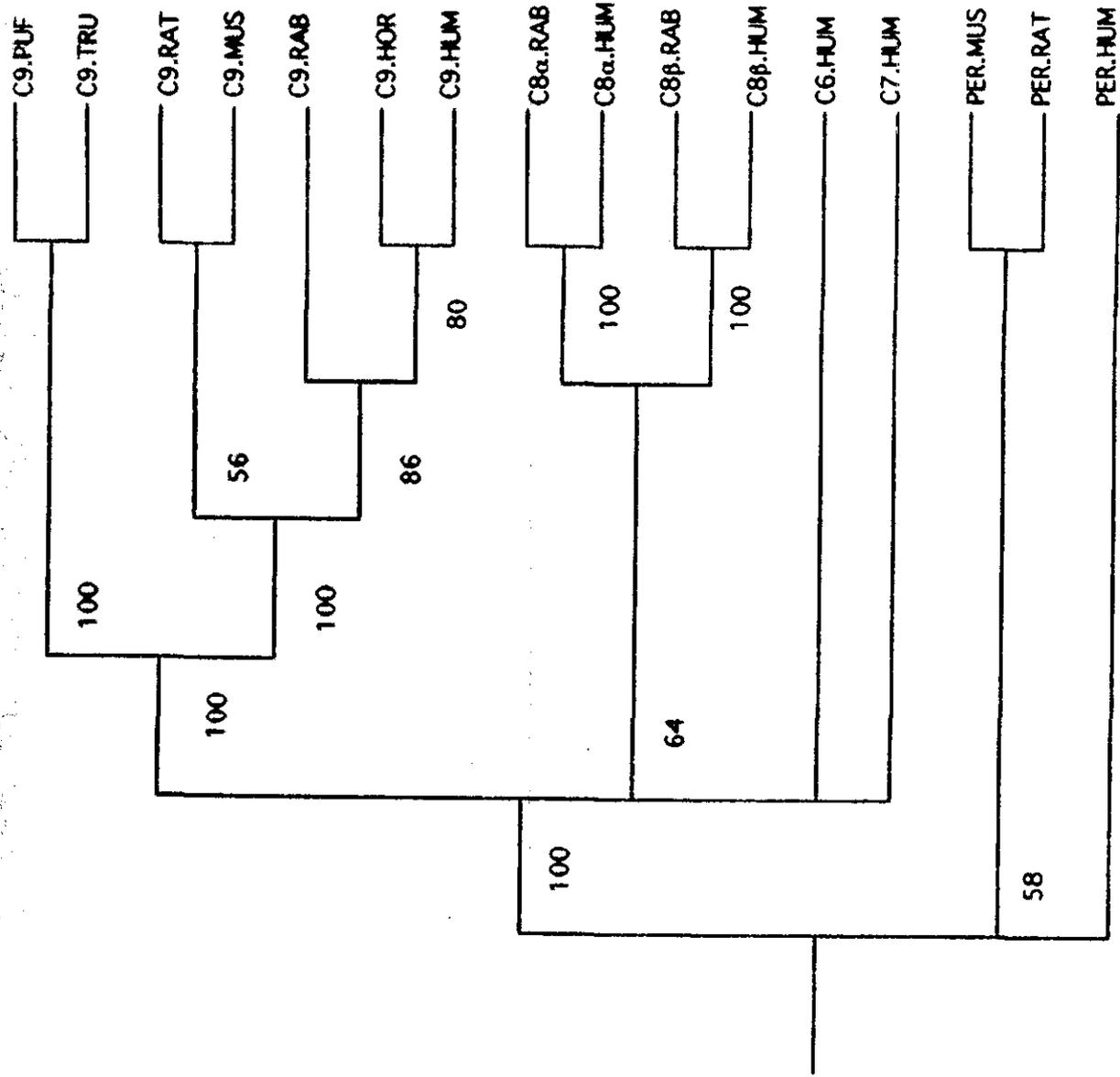


Fig. 10 Resultados del análisis por Parsimonia. Filogenia consenso del 50%, resultado de un análisis con Branch-and-Bound. Las perforinas se emplearon como grupo externo. Los valores en cada clado son los resultados de la prueba de "Bootstrap" de 100 iteraciones.



complemento (Fig. 10). La falta de resolución del análisis por el método de Parsimonia se puede deber a la estructura de mosaico de las proteínas C6-C9: si los módulos de cada proteína cambian a velocidades distintas, es posible que algunas proteínas aparezcan fuera de lugar en un análisis de Parsimonia. El análisis de Parsimonia sin prueba de Bootstrap genera dos árboles más parsimoniosos de 1444 pasos. Estas dos topologías además de ser monofiléticas, agrupan en forma consistente a las proteínas ortólogas de la misma manera que el árbol obtenido con la prueba de Bootstrap. En uno de estos árboles más parsimoniosos el C9 de ratón es el primero en divergir del ancestro común de los C9 de mamíferos, seguido por el C9 de rata (Fig. 11). La diferencia más notable entre estos árboles y el obtenido con la prueba de Bootstrap es que, ambos árboles están completamente resueltos. Las proteínas C6 es la primera en divergir a partir del ancestro común de las proteínas del complemento, es seguida por C7 y el ancestro común de los C8 y C9.

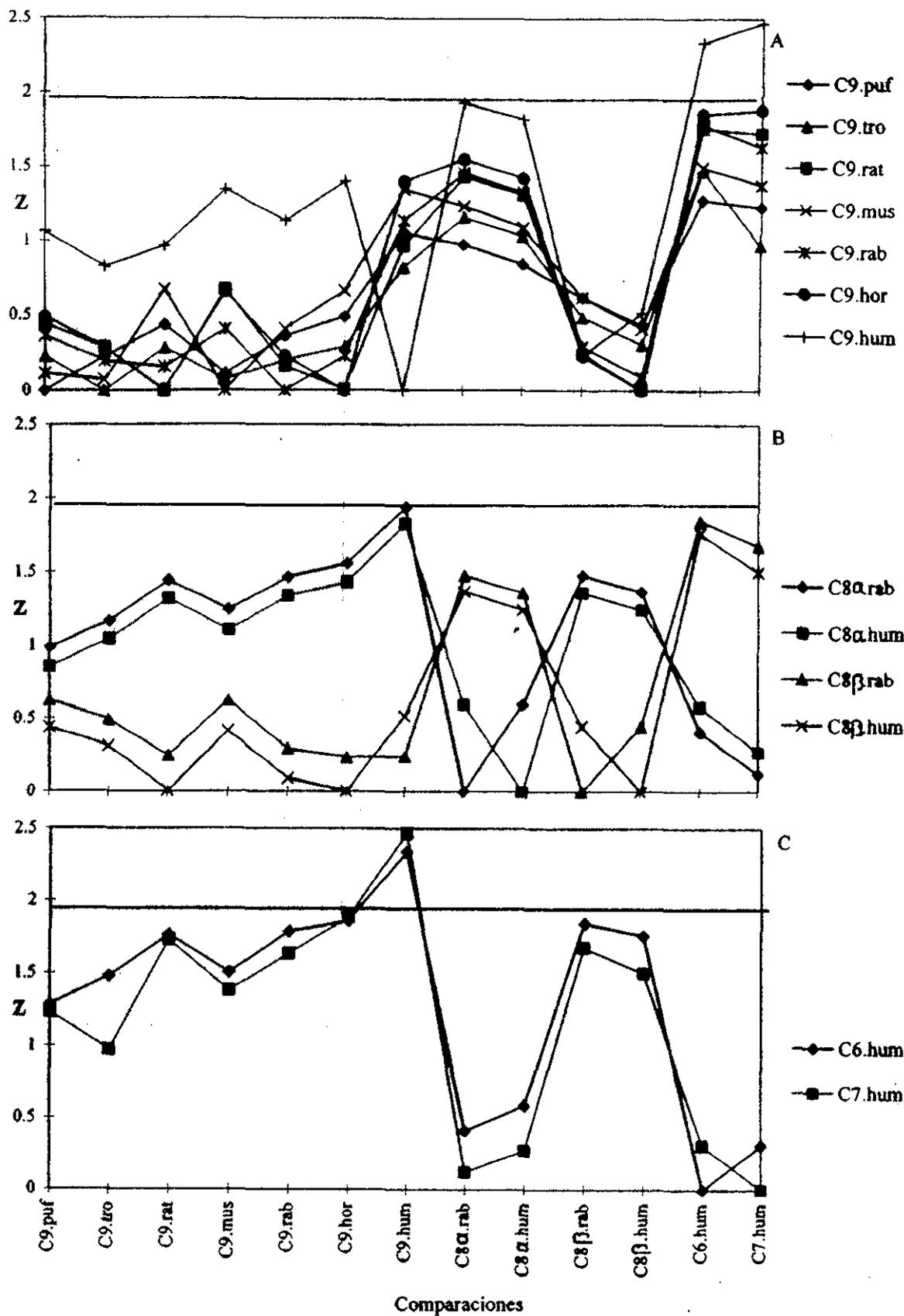
### 3. Prueba de la tasa relativa

Para determinar la velocidad de cambio evolutivo entre las secuencias de aminoácidos empleadas en la presente tesis, se aplicó la RRT a comparaciones entre pares de todas las proteínas (Fig. 12)<sup>††</sup>. Cada punto en la gráfica representa el valor de Z obtenido al comparar la distancia entre un par de secuencias con respecto a las perforinas.

Una divergencia significativa de la tasa constante entre los valores comparados (al 5% de confianza) se observa exclusivamente cuando el C9 humano se comparó con el C6 y el C7 (Fig. 12A), aunque también se observaron divergencias cercanas al valor de corte de (1.96) entre C9 humano y los dos C8 $\alpha$  (Fig. 12B). De acuerdo con esto, la mayor parte de las demás comparaciones con el C9 humano también resultaron en valores Z más altos que las mismas comparaciones con otras moléculas de C9. Las proteínas C9 de rata y de caballo muestran un traslape de sus valores en la mayoría de las comparaciones realizadas. Los valores de C8 $\beta$  se

---

<sup>††</sup> Fig. 12 (En la siguiente página) Tasas de cambio heterogéneas en la evolución de C6-C9. Valores de Z-estadística obtenidos al comparar mediante la prueba de la tasa relativa (Takezaki, 1995), pares a pares de secuencias: A, los C9; B, los C8 $\alpha$  y C8 $\beta$  y C, C6 y C7. Cada punto en la gráfica representa el valor de Z obtenido de comparar la distancia de dos secuencias a las perforinas. La línea horizontal marca el valor de corte (arriba de 1.96, se considera que las secuencias comparadas cambian con una tasa no constante).



agrupan con todos los C9 no humanos, mientras los dos C8 $\alpha$  siguen una tendencia opuesta (Fig. 12 B). Finalmente, las tasas de cambio de C6 y C7 siguen tendencias paralelas que son más similares a las de los C8 $\alpha$  que a la de la mayoría de los C9 y ambos C8 $\beta$  (Fig. 12C).

También se estudiaron las tasas de cambio de las proteínas C6-C9, mediante el cálculo porcentual de la aceleración o desaceleración de cada proteína con respecto a las demás. Este análisis confirmó que el C9 humano es la molécula que tiene una tasa de cambio más rápida en todas las comparaciones (Tabla VIIA). También se observa que los C9 cambian a una velocidad más alta con respecto a C8 $\alpha$ , C6 y C7 (Tabla VIIB). En contraste, la velocidad relativa de cambio de los C9 con respecto a C8 $\beta$  es más lenta, con excepción de C9 humano, que tiene una tasa de cambio más rápida que las de los C8 $\beta$  (Tabla VIIC). Por otra parte, las moléculas de C8 $\alpha$  y C8 $\beta$  cambian con mayor velocidad que C6 y C7. Finalmente, el C7 cambia a una velocidad 6.94 % más rápido que el C6 (Tabla VIID). Este análisis muestra una tendencia de cambio acelerada en las proteínas C6-C9 humanas que parecen incrementar progresivamente su velocidad de cambio en un orden  $C6 < C7 < C8 < C9$  (Tabla VIII). Estos resultados dan una aproximación para entender la notoria divergencia de las tasas de cambio obtenidas en el análisis de RRT descrito anteriormente.

Los resultados del RRT aplicado a comparaciones entre pares de módulos individuales CP o EGF indican que la desviación de la tasa constante que se observó al hacer esos análisis con el total de la secuencias alineadas, se debe principalmente al módulo CP. En la figura 13 se observa que las tendencias de cambio del CP son muy similares a las obtenidas para toda la secuencia alineada, con la diferencia de que ninguna de las comparaciones supera el valor de corte de 1.96. En cambio, el análisis del módulo EGF (Fig. 14) muestra patrones de cambio muy diferentes a los encontrados para la totalidad de la alineación así como con la obtenida con las secuencias del módulo CP, que superan casi todas el valor de corte de 1.96.

La desviación de una tasa que se considera constante y las distintas tasas de cambio observadas en los análisis por RRT para la totalidad de la secuencia alineada y en los módulos individuales indican que para determinar el orden evolutivo de aparición de estas moléculas es necesario emplear un método de análisis que acepte heterogeneidad en las tasas de evolución de las proteínas.

Tabla VII. Porcentajes de aceleración relativa de las proteínas C6-9. En las tablas se muestra la cantidad porcentual por la que las proteínas enlistadas en la primera columna (en negritas) superan en tasa de cambio a las moléculas de la primera fila.

VIIA

	<b>C9.fug</b>	<b>C9.tru</b>	<b>C9.rat</b>	<b>C9.mus</b>	<b>C9.rab</b>	<b>C9.hor</b>	<b>C8α.rab</b>	<b>C8α.hum</b>	<b>C8β.rab</b>	<b>C8β.hum</b>	<b>C6.hum</b>	<b>C7.hum</b>
<b>C9.hum</b>	35.19	28.61	70.48	88.65	102.73	181.89	38.39	36.46	5.99	11.6	42.09	40.59

VIIIB

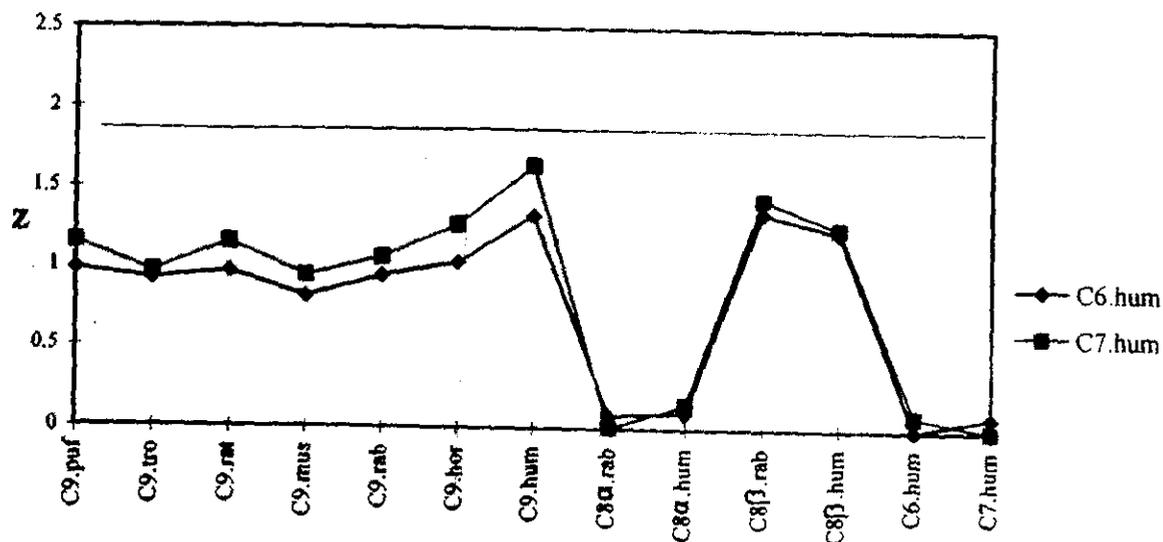
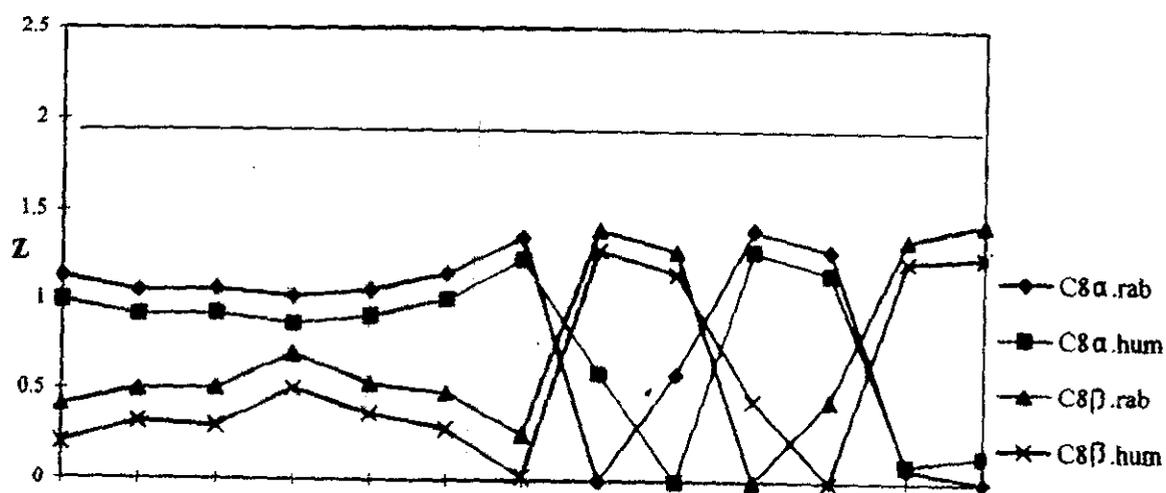
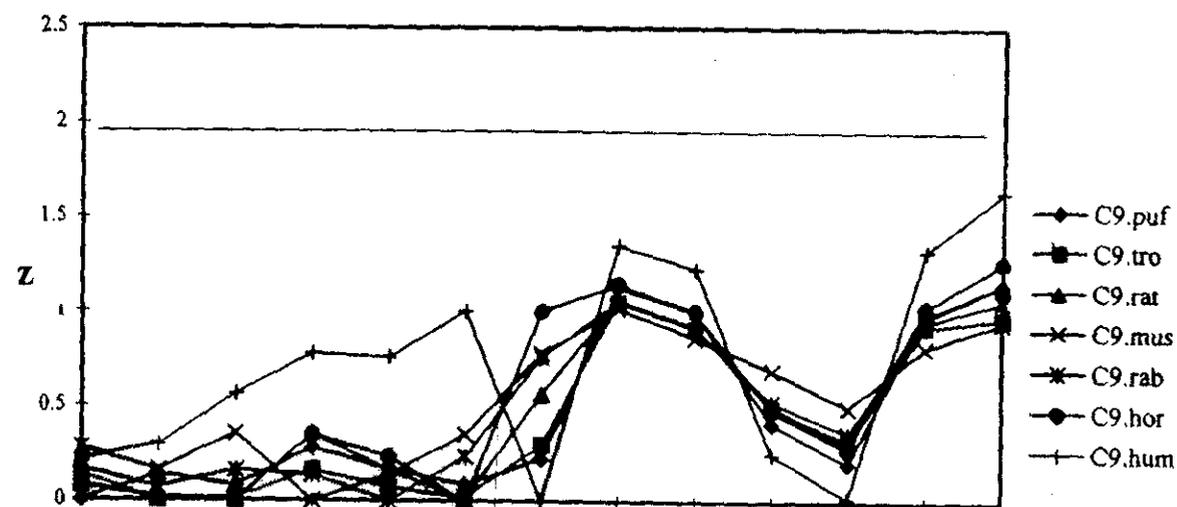
	<b>C8α.rab</b>	<b>C8α.hum</b>	<b>C6.hum</b>	<b>C7.hum</b>
<b>C9.fug</b>	19.73	16.76	26.39	21.63
<b>C9.tru</b>	22.56	20.07	28.6	22.44
<b>C9.rat</b>	29.57	32.89	33.8	31.81
<b>C9.mus</b>	24.59	21.34	29.32	25.24
<b>C9.rab</b>	28.01	26.13	32.98	30.62
<b>C9.hor</b>	31.37	28.42	33.67	31.69

VIIIC

	<b>C9.fug</b>	<b>C9.tru</b>	<b>C9.rat</b>	<b>C9.mus</b>	<b>C9.rab</b>	<b>C9.hor</b>
<b>C8β.rab</b>	15.23	11.89	6.41	15.48	8.06	6.09
<b>C8β.hum</b>	9.8	6.14	0.17	8.83	2.19	0.1

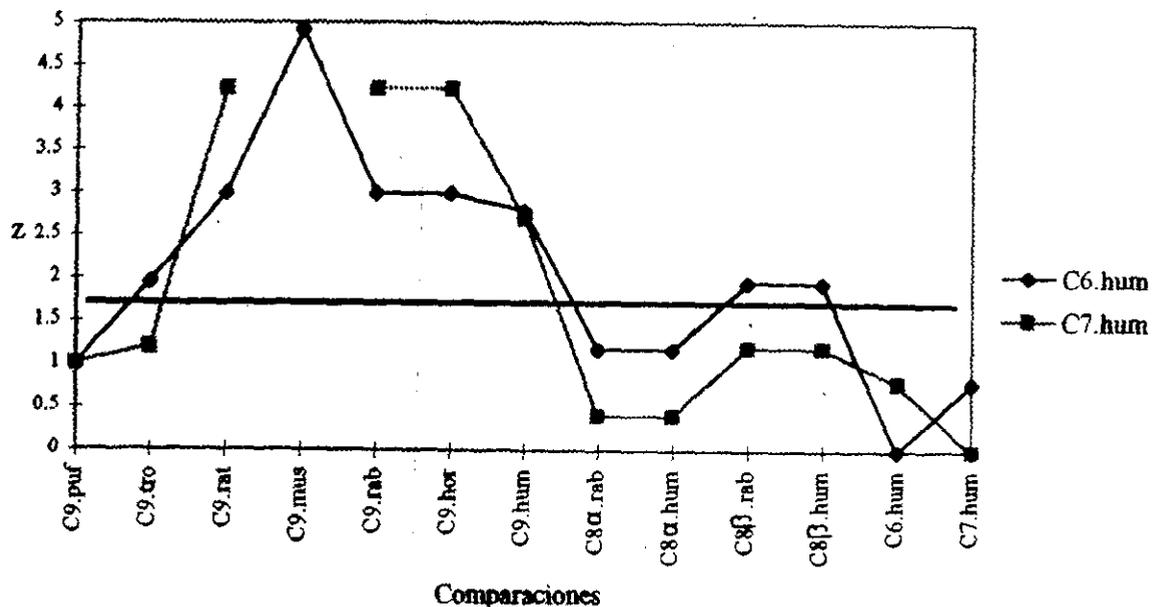
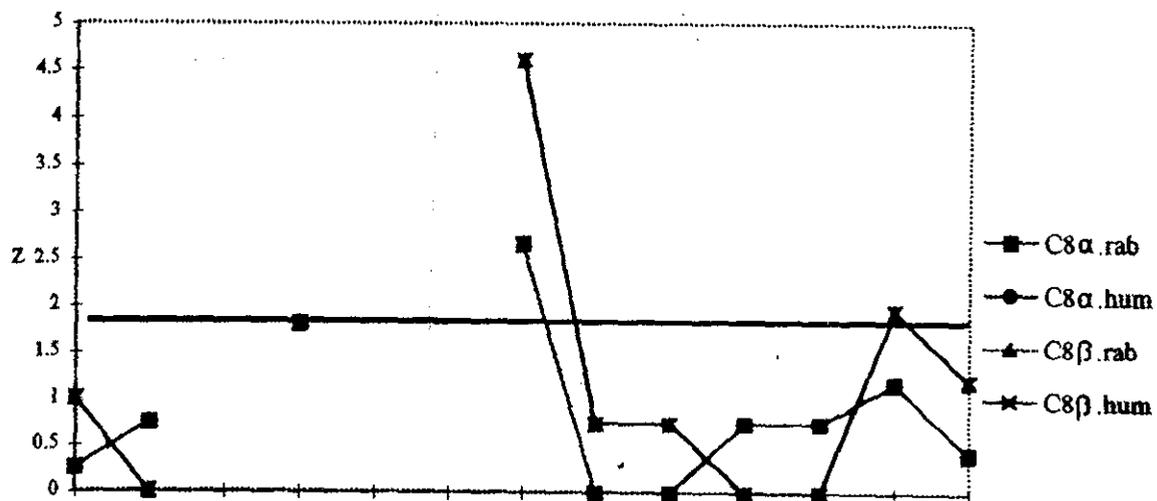
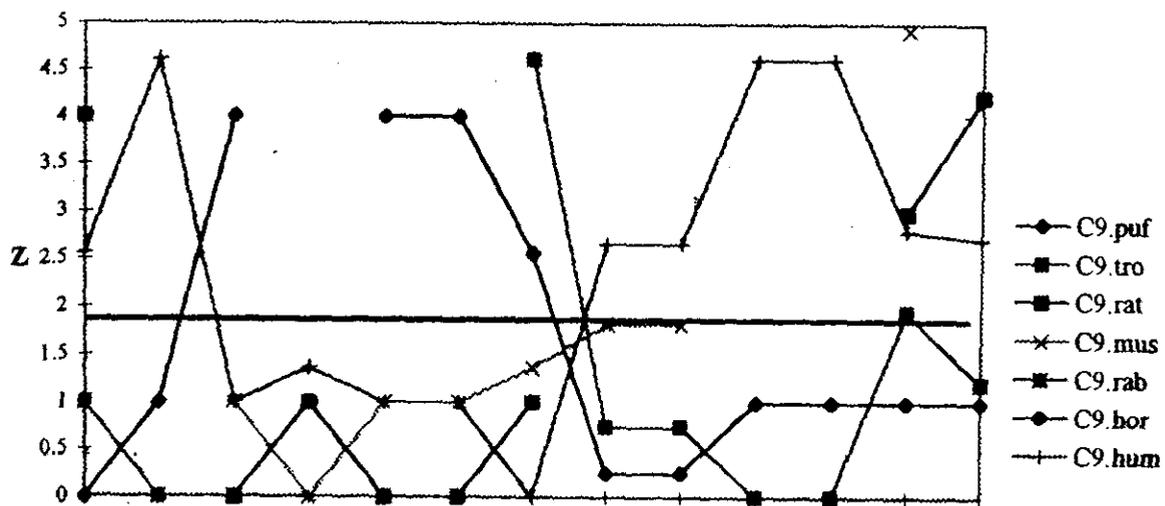
VIIID

	<b>C6.hum</b>	<b>C7.hum</b>
<b>C8α.rab</b>	9.06	2.85
<b>C8α.hum</b>	12.48	7.34
<b>C8β.rab</b>	39.16	34.19
<b>C8β.hum</b>	35.36	30.52



Comparaciones

Fig. 13 Tasas de cambio de la región pobre en cisteínas.



Comparaciones

Fig. 14 Tasas de cambio altamente heterogéneas e irregulares en el módulo EGF de C6-C9. valores no calculados se deben a varianzas negativas que impiden el cálculo de Z por el prog Phyltest. Nótese que en la gráfica 13B las tendencias de los C8a así como los de C8b se sobre

Tabla VIII Aceleración de las proteínas C6-9 con respecto a C6 ó C7. En la tabla se muestran los porcentajes por los que las moléculas C6-9 superan en tasa de cambio a las proteínas C6 y C7.

	C6.hum	C7.hum
C7.hum o C6	6.92	6.49
C8 $\alpha$ .hum	12.45	7.34
C8 $\beta$ .hum	35.36	30.52
C9.hum	42.09	40.57

#### 4. Análisis por método de distancia

Un análisis con el método de "Neighbor-Joining" usando PAUP resultó en un árbol completamente resuelto (Fig. 15) y con mejores valores de "Bootstrap" (Fig. 16) que el árbol obtenido en el análisis de Parsimonia. Al igual que en el análisis por el método de Parsimonia, el origen de todas las proteínas del complemento aparece como monofilético y todas las proteínas del TCC se reúnen en grupos de ortólogos. En este árbol un clado formado por C6 y C7 aparece como grupo hermano del resto de las proteínas terminales del complemento sugiriendo que estas proteínas emergieron antes que C8 y C9. A su vez, el grupo de C8, subdividido en los tipos C8 $\alpha$  y C8 $\beta$ , es el grupo hermano de los C9.

#### 5. Análisis de Hillis-Huelsénbeck

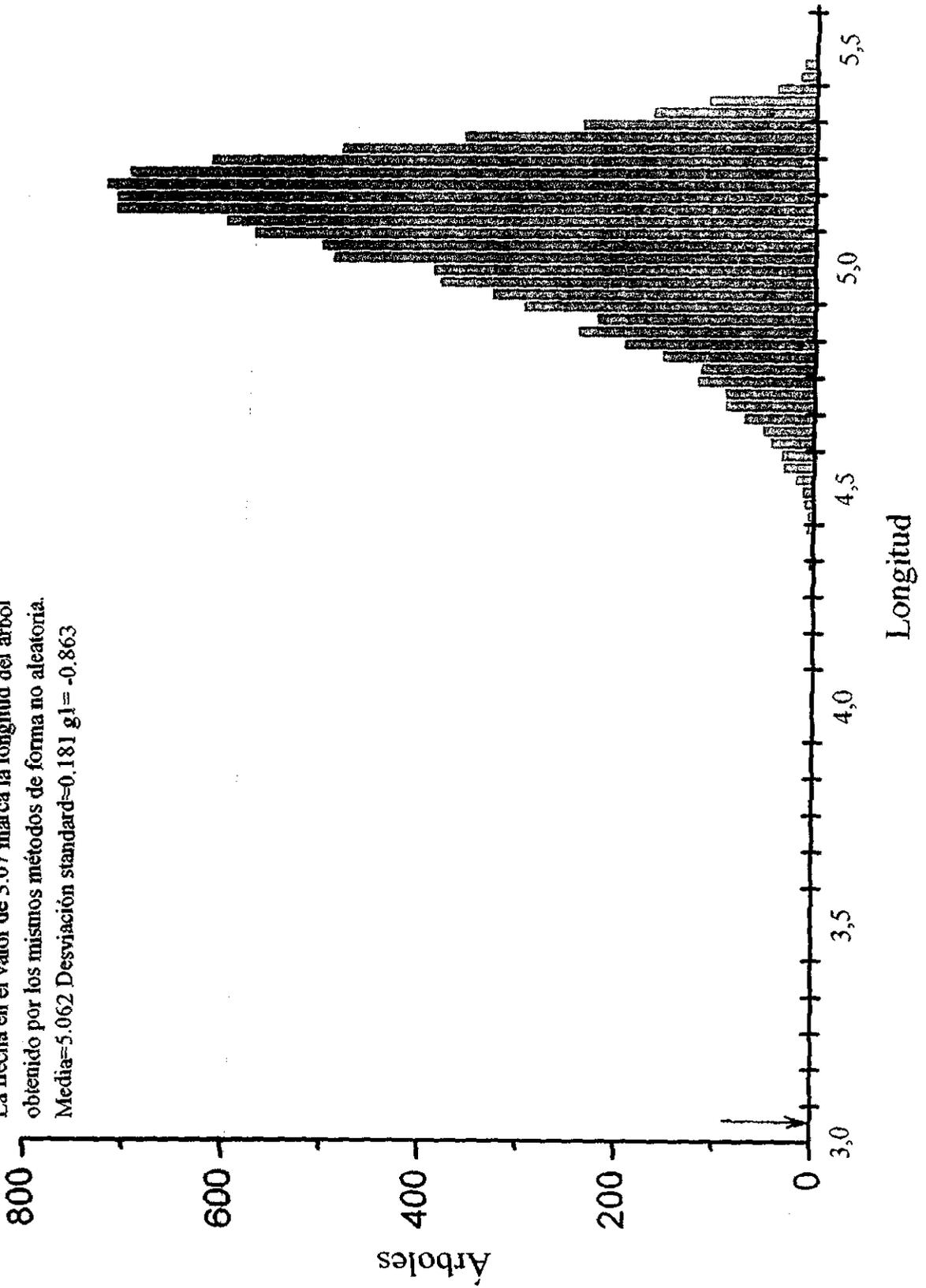
Con la finalidad de averiguar si la misma topología del árbol obtenido en el análisis de distancia pudiera generarse aleatoriamente, se realizó un análisis de Hillis-Huelsénbeck (ver materiales y Métodos). Los resultados mostraron que la media de todos los árboles generados al azar es 5.062085 unidades. En contraste, el árbol obtenido por el método de distancia tiene una longitud de 3.07631 por lo que dista 1.985775 unidades de la media de distancia obtenida aleatoriamente, siendo claro que se trata de topologías significativamente distintas. En la gráfica se muestra la frecuencia de distribución de las longitudes de los 10,000 árboles generados aleatoriamente (Fig. 17). Esta gráfica no se puede interpretar en base a los criterios de un análisis convencional de Hillis-Huelsénbeck debido a que se trata de información obtenida a partir de secuencias de aminoácidos. Adicionalmente, resultados de un análisis de decaimiento de las topologías obtenidas por el método de parsimonia, indican que alrededor de la mitad de los árboles de hasta 1450 pasos poseen la misma topología que el más parsimonioso. Lo anterior indica que cada una de las relaciones representadas en el árbol están sólidamente sustentadas.





Fig. 17 Resultado de análisis Hillis-Huelsensbeck.

Se muestran el número de árboles de una longitud determinada obtenidos al azar. Estos árboles se generaron a partir de la alineación empleada para los análisis de inferencia filogenética. Los árboles se calcularon empleando Distancia P y NJ. La flecha en el valor de 3.07 marca la longitud del árbol obtenido por los mismos métodos de forma no aleatoria. Media=5.062 Desviación standard=0.181  $g1 = -0.863$



## VI. DISCUSIÓN

*"...for a biologist to ignore the evolutionary history  
either of an organism or a complex system  
is as much an impediment to comprehending  
that organism or system as it is for a student  
of society to ignore the history of mankind."  
Introducción de Phylogenesis of immune  
functions. Warr G.W*

Los resultados del análisis de inferencia filogenética usando todas las secuencias de aminoácidos disponibles para perforinas y C6-C9 sugieren que C6 y C7 fueron las primeras proteínas en aparecer durante la evolución.

Otro grupo ha obtenido conclusiones similares a las nuestras mediante un análisis de la estructura de intrones/exones de los genes C6-C9 humanos (Hobart, 1993; 1995). Nuestras conclusiones también coinciden con las de este grupo en que las proteínas C6-C9 tienen un origen único, así como que el ancestro común de los componentes terminales del complemento debió haber tenido una estructura modular compleja (Hobart, 1993; 1995). La diferencia entre ambos acercamientos estriba en el procedimiento empleado. Nosotros usamos un método computacional estándar de inferencia filogenética con todas las secuencias de aminoácidos disponibles para estas proteínas. Además, el orden de aparición de C6-C9 se estableció empleando como punto de referencia la secuencia de la perforina, con mayor similitud con estas moléculas.

Entre los componentes terminales del complemento, el C9 es la proteína más simple, formada por los módulos TSP, LDLr, la región CP y el módulo EGF. Estos dos últimos módulos también se encuentran en las perforinas (Fig.2). Además de estos módulos, las moléculas C8 $\alpha$  y C8 $\beta$  tienen un segundo módulo TSP en su región carboxilo terminal (también presente en C9 de peces). Las estructuras modulares más complejas son las de C6 y C7 que tienen dos módulos SCR y dos módulos CFI adicionales, en sus extremos carboxilo terminal. Finalmente, el C6 además posee un tercer módulo TSP en su extremo amino terminal.

Considerando la complejidad de la estructura modular de estas proteínas, el árbol obtenido en el análisis de distancia sugiere que la duplicación de un gen ancestral con una composición modular compleja produjo a los grupos de las perforinas y de las proteínas de la vía terminal del complemento. La estructura simple de la perforina pudo haberse originado a partir del ancestro complejo por medio de una serie de deleciones modulares. Por otro lado, el ancestro de las proteínas del TCC sufrió una serie de duplicaciones que en combinación con la misma tendencia

a perder módulos manifiesta en el lado de las perforinas, dio origen inicialmente al C6 y C7, seguidos sucesivamente por el C8 y el C9 que poseen una menor complejidad modular.

Los módulos perdidos durante este proceso evolutivo son los que se encontraban en los extremos amino y carboxilo de la proteína ancestral, habiéndose conservado la región central. Evidencia de esto, es que los exones 1 del C7, C8 y C9 son homólogos al exón 2 del C6 puesto que tienen en su extremo 5' un intron de tipo dos. En el C7, C8 y C9 ya no existe un exón homólogo al 1 de C6, sin embargo, la evidencia de que existió un exón es que el primer intron de C7, C8 y C9, conserva la fase del que tiene el gen del C6 en su extremo 5' (Hobart, 1993; 1995). Otra evidencia de la tendencia a perder módulos posterior a la duplicación del gen ancestral más complejo es que los C9 del pez globo y de la trucha son mas parecidos en su composición modular a las moléculas C8 $\alpha$  /C8 $\beta$  del humano y del conejo (Yeo, 1997; Tomlinson 1993), que a los propios C9 de mamíferos. Puesto que ambas especies de peces tienen un origen evolutivo más temprano, se puede suponer que en los genes de las especies de mamíferos, el módulo TSP en el extremo carboxilo se ha perdido durante la evolución del C9 (Fig.9). No se ha encontrado una pérdida equivalente de módulos al comparar la estructura de las perforinas caracterizadas que pertenecen exclusivamente a especies de mamíferos. Al respecto, la caracterización de perforinas de reptiles y anfibios podría aclarar si la misma tendencia ha guiado la evolución de las perforinas. Cabe hacer notar que no se dispone de evidencia que demuestre la existencia de linfocitos citotóxicos y células NK en peces.

La capacidad del C9 para formar poros citolíticos ha llevado a compararla funcionalmente con la perforina. Sin embargo, las funciones de perforina y C9 dependen de regiones distintas entre sí, lo que sugiere la ocurrencia de algunos eventos evolutivos adicionales a la duplicación y pérdida de módulos. Esto quiere decir que a partir de la duplicación del gen ancestral común, los linajes de perforinas y C9s siguieron un proceso de divergencia, evidenciado por el hecho de que regiones no homólogas en ambas proteínas son aparentemente las responsables de su inserción a membranas.

¿Cuál fue el origen del ancestro de C6-C9 y perforina que aparentemente poseía una estructura modular compleja?. Es posible que la molécula ancestral se formó por eventos de intercambio de exones ("exon shuffling"). Se ha propuesto que los genes formados por exones de la misma fase tienden a intercambiar o duplicar sus exones. Una vez que este intercambio ha

ocurrido, los genes participantes comparten regiones extensas, facilitando futuros intercambios. De esta manera se pueden integrar redes de familias génicas en las que el intercambio de exones ocurre con relativa frecuencia (Patthy, 1987). Se ha planteado que los genes de trombospondina, receptor de lipopolisacáridos de baja densidad y factor de crecimiento epidermal con exones de fase 1, integraron junto con los genes de las proteínas del TCC una red de intercambio de exones (Patthy, 1988). Sin embargo, los genes de C6-C9 conocidos no están formados por exones de fase 1 ni la longitud de la mayor parte de sus módulos coincide con la de sus exones. Es decir, que el patrón de intrones/exones en los genes actuales de las proteínas del TCC no parece ser el adecuado para propiciar un intercambio de exones propuesto. A pesar de ello, no se puede excluir la posibilidad de que la historia evolutiva reciente de estos genes, en particular algunos eventos de inserción/delección hayan alterado el patrón original de intrones (Patthy, 1987). También se ha propuesto que el intercambio de exones solo ha jugado un papel importante en la formación del gen ancestral de C6-C9 (Hobart, 1995). El ancestro común a C6-C9 y a perforina pudo haberse formado por intercambio de exones de fase 1 con los genes con que comparte módulos. Posteriormente, este gen mosaico adquirió estabilidad cuando sus exones adoptaron la estructura de fases actual que impidió la continuación de los intercambios. Esta estructura se ha conservado en C6-C9 humanos y en C9 de pez globo. En contraste, las fases y longitudes de intrones y exones en las perforinas no guardan relación con las de los componentes terminales del complemento.

¿Qué significado evolutivo tienen las características moleculares y funcionales compartidas por C9 y perforina?. A pesar de que C9 y perforina comparten la capacidad de insertarse en membrana y formar poros, las condiciones selectivas que guiaron la evolución de las dos proteínas debieron haber sido diferentes. Por parte de las perforinas, el desarrollo de una capacidad lítica pudo conferir una ventaja selectiva por su función efectora, mientras que en el caso del C9 fueron quizá las funciones de señalamiento celular, a través del aumento de la permeabilidad membranal a cationes, las que impulsaron su historia evolutiva. Se trata del surgimiento de estructuras en forma de poros membranales por dos vías evolutivas independientes a partir de un ancestro común que no tenía esta función\*. El término convergencia

---

\* Esto lo inferimos por la semejanza que se ha propuesto tiene con el actual C6 que no forma canales.

implica cambios adaptativos que hacen que entidades no relacionadas aparezcan como más cercanas mientras que paralelismo es el surgimiento de una característica funcional similar en entidades relacionadas genéticamente, pero cuyo ancestro común es distinto (Doolittle, 1994). En las macromoléculas se pueden considerar fenómenos de convergencia o paralelismo en el plano de lo estructural, mecanístico, funcional y de secuencia. En el actual proyecto de tesis, únicamente se consideraron los dos últimos planos debido a la carencia de información respecto a los otros. Se considera que existe convergencia funcional cuando en forma independiente surge una determinada función en más de una ocasión. Como ejemplos se pueden mencionar las enzimas que catalizan una misma reacción y cuyos genes tienen un origen diferente (p. ej. serin proteasas, alcohol deshidrogenasas). Asimismo, se puede definir convergencia a nivel de secuencias cuando las proteínas muestran regiones parecidas, supuestamente originadas por condiciones selectivas similares (Doolittle, 1994), a pesar de que el resto de su secuencia indica claramente que no están relacionadas genéticamente. Las secuencias de aminoácidos empleadas en el presente estudio son homólogas entre sí puesto que su extensa similitud indica que provienen de un ancestro común. Sin embargo, la topología del árbol obtenido por el método de distancia revela que tanto C9 como perforina están más cercanamente relacionadas con otras moléculas del complemento que entre sí mismas. En otras palabras, si la aparente convergencia funcional se reflejara a nivel de las secuencias de aminoácidos de C9 y perforina, el mayor número de sitios idénticos las habrían ubicado en un mismo clado en el árbol obtenido. Por consiguiente, se trata de un fenómeno de paralelismo, es decir, del surgimiento de una característica funcional similar en proteínas relacionadas genéticamente, pero cuyo ancestro común es distinto (Strickberger, 1996a). En base a estos argumentos se puede concluir que a nivel de secuencias las similitudes observadas en C9 y perforina están dadas por que ambas proteínas tienen un ancestro común, mientras que su capacidad para formar poros es resultado de un fenómeno de paralelismo.

Aunque las dos vías independientes que parecen haber originado a perforina y C9 sean representadas como una sucesión lineal de ancestro-descendiente, esto no significa que la evolución de ambas moléculas haya ocurrido a través de una vía directa hacia el perfeccionamiento de funciones de lisis y señalamiento. Durante la evolución de C9 y perforina a partir de su ancestro común pudieron existir genes que al duplicarse perdieron su función

(pseudogenes), o adquirieron otras funciones que los distanciaron de la familia formada por C6-C9 y perforina. La historia completa de las proteínas terminales del complemento posiblemente involucra tanto formas que permanecieron como otras que desaparecieron, o que por su estructura y función específica no son fácilmente identificables (Fig. 18).

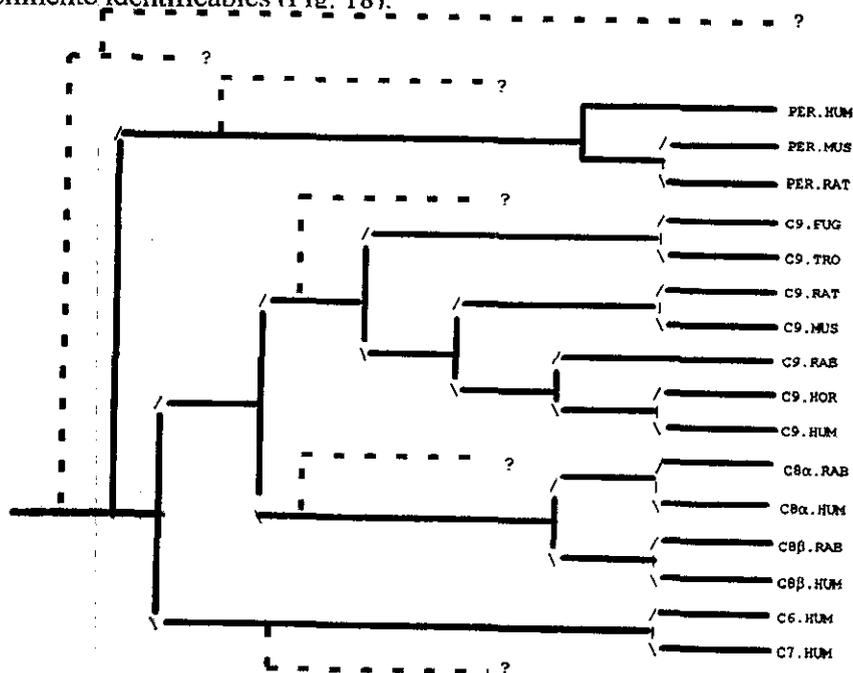


Fig. 18 Representación ideal de la filogenia de C6-C9. En este árbol se incluyen las posibles formas intermedias o extintas durante la evolución de las proteínas de la vía terminal.

Las condiciones que favorecieron la aparición de moléculas formadoras de canales con funciones de señalamiento más que de lisis, posiblemente están relacionadas con el desarrollo de mecanismos de evasión de las propiedades citolíticas del complemento por organismos patógenos. Las bacterias gram negativas y gram positivas evitan la lisis por complemento al interponer como barrera física diversas moléculas de su cubierta externa, que además de evitar el reconocimiento del C3b depositado en la superficie bacteriana por fagocitos, impide la formación de C3 convertasa en ciertas cepas de *Streptococcus*, *Campylobacter* (así como en el protozoario *Trypanosoma cruzi*). Los tripomastigotes de *T. brucei*, *T. gambiense* y *T. lewisi* degradan al C3b que se une a su superficie evitando la posterior inserción del TCC (Joiner 1988). Además, estos patógenos han desarrollado mecanismos directos que les permiten resistir la destrucción citolítica por complemento. La gruesa capa de peptidoglicanos de las bacterias gram positivas restringe la inserción de C5b-9 a la membrana citoplásmica mientras que en las bacterias gram negativas

como *P. aeruginosa*, *C. fetus* y *E. coli* esta misma el TCC no se inserta en la región hidrófoba de la membrana externa debido a la presencia de lipopolisacáridos con largas cadenas laterales de O-polisacáridos. Casos similares de inserción deficiente se presentan en diversas cepas resistentes de *Neisseria* y amastigotes de *T. cruzi* (Cooper, 1991; Morgan, 1990d). En promastigotes metacíclicos de *Leishmania major* los complejos C5b-9 son desprendidos de la superficie del parásito por exocitosis activa (Joiner, 1988). De manera similar, la *Entamoeba histolytica* es capaz de eliminar TCC de sus membranas por endocitosis (Fuhrman, 1989). Estos ejemplos de evasión y el hecho de que la deficiencia homocigótica del gen para C9 en uno de cada 1000 japoneses no ocasiona padecimiento alguno (Morgan, 1990d), desafían la noción de la relevancia de la función lítica del TCC.

Por otra parte, la aparición temprana de C6 y C7 sugerida por nuestro análisis tiene implicaciones con respecto a la relevancia de las funciones de señalamiento de la vía terminal, puesto que C5b y C5b6 no participan en señalización mientras que C5b67 se inserta en la membrana y transmite señales mediante proteínas G. La subsecuente incorporación del C8 permite la formación de un canal estrecho y poco estable, que aumenta la función señalizadora mediada por proteínas G, a través de un modesto flujo de  $Ca^{++}$ . Finalmente, la incorporación del C9 permite la formación de un canal estable y un mayor flujo de  $Ca^{++}$ . Aunque el C9 puede polimerizar por si solo formando poros, las condiciones necesarias para que esto suceda no son las fisiológicas. Por lo tanto, el complejo C5b67 es la mínima unidad funcional de señalización, lo que está de acuerdo con el orden de emergencia evolutiva de los componentes terminales del complemento que sugieren nuestros resultados. La emergencia tardía del C9 también se apoya por el hecho de que solo funciona en condiciones fisiológicas cuando se une al complejo C5b-8. La capacidad lítica del complemento parece perder relevancia ante las funciones transductoras de señales, en el contexto de la historia evolutiva de la vía terminal y de los mecanismos de evasión descritos anteriormente.

Tanto el árbol obtenido por el método de Parsimonia como el generado por distancia muestran un solo ancestro común para las proteínas del TCC, es decir que tuvieron un origen monofilético. Este resultado confirma que C6-C9 son proteínas cuyas características compartidas se deben a un origen evolutivo común. ¿Qué evidencia apoya las filogenias de donde derivan estas conclusiones?. En los árboles obtenidos por ambos métodos de inferencia filogenética se

observa el agrupamiento consistentemente de las proteínas ortólogas que se ordenan ( con excepción del C9 de caballo) de acuerdo a la filogenia de las especies que representan. Lo anterior además de confirmar la identidad de las secuencias empleadas, facilita la inferencia del orden de aparición de las proteínas parálogas<sup>†</sup> C6-C9. La topología de los árboles obtenidos también es apoyada por los resultados de la prueba de "Bootstrap". La mayoría de las relaciones filogenéticas en el árbol obtenido por distancia, muestran valores superiores al 70% . Estos altos valores de "Bootstrap" indican que es muy probable obtener las mismas relaciones entre las proteínas analizadas si realizáramos el proceso de inferencia filogenética un número infinito de veces con una muestra nueva e independiente de secuencias. Los resultados obtenidos en la prueba de "Bootstrap" también proporcionan un estimado conservador sobre la probabilidad de haber inferido correctamente las relaciones filogenéticas entre un grupo de secuencias (Ver anexo 3). Un cuarto argumento de apoyo al árbol obtenido por distancia es que la longitud de su topología es 1.6455 veces más pequeña que la obtenida más frecuentemente al azar, lo que quiere decir que este resultado está lejos de ser obtenido de forma aleatoria.

Con respecto a la alineación empleada en este estudio y al análisis filogenético por Parsimonia, las 444 posiciones incluidas en el análisis de Parsimonia permitieron generar un solo árbol con el mínimo número de cambios evolutivos que explican las diferencias entre las proteínas examinadas, sin embargo, los 314 sitios informativos resolvieron únicamente las conexiones internas de los grupos C9, C8, C6/C7 y perforina sin que las relaciones entre dichos grupos quedaran establecidas. Si bien es claro que se requieren más sitios informativos para resolver el orden de aparición de estas proteínas mediante Parsimonia, cabe señalar que se emplearon todas las secuencias de aminoácidos disponibles en GenBank, y que se les comparó en un proceso iterativo de alineación/inferencia/alineación basado en sus regiones más conservadas, donde se eliminaron 167 posiciones cuya alineación tenía algún grado de incertidumbre. Aunque las zonas eliminadas contienen 131 sitios informativos, su inclusión en otros análisis de Parsimonia no descritos en el presente trabajo de tesis tampoco generó árboles que resolvieran el orden de aparición de C6-C9.

El método de Parsimonia (Ver Anexo 3), tiene como una de sus suposiciones el que la

---

<sup>†</sup> Proteínas codificadas por genes parálogos. Esta clase de genes pueden ser dos o más loci en el mismo organismo cuya similitud a nivel de secuencia indica que se originaron de una o más duplicaciones de un gen ancestral común. Generalmente los genes parálogos se pueden agrupar en familias.

tasa de sustitución de nucleótidos o aminoácidos en las secuencias analizadas es constante. Esto puede ser cierto si el número de sustituciones por sitio es pequeño. Sin embargo, si este número es tan grande que ocurren mutaciones paralelas y reversas frecuentemente, el método de Parsimonia tiende a hacer reconstrucciones erróneas (Nei, 1987). El que éste método no considere tasas de sustitución desiguales ocasiona conflictos cuando entre las distintas ramas de una topología hay una gran cantidad de sustituciones múltiples, pues los árboles obtenidos serán erróneos debido a secuencias evolutivamente más cercanas que aparecen como lejanas y viceversa (Swofford *et al.*, 1996). Mediante la prueba de la tasa relativa encontramos que las proteínas C6-C9 muestran tasas de cambio variables, lo que indica que han ocurrido mutaciones paralelas y reversas frecuentes. Es posible que la falta de resolución del árbol obtenido por Parsimonia se deba al conflicto entre las suposiciones del método y las características de las secuencias empleadas. Para sortear la dificultad que representan los supuestos del método de Parsimonia, empleamos métodos de matriz de distancia que indirecta o directamente consideran las tasas de cambio variables en C6-C9 a través de distinguir las mutaciones paralelas y reversas que pueden presentarse en cada sitio. Una ventaja adicional de emplear métodos de distancia en vez de Parsimonia, es que el método de Parsimonia descarta para el análisis a los sitios que considera no informativos, mientras los métodos de distancia generalmente emplean todos los sitios de la alineación (Nei, 1987).

En el proceso de inferencia filogenética con métodos de distancia se calcularon las relaciones de similitud entre todos los pares de secuencias mediante la suma del número total de diferencias entre cada par de secuencias dividido entre el número total de posiciones, las distancias calculadas de este modo son por lo tanto el promedio de los caracteres diferentes entre un par de secuencias o Distancia P (Swofford *et al.* 1996). A partir de estas distancias el algoritmo de Neighbor-Joining (NJ) construye un árbol filogenético.

El cálculo de las distancias por el método de Distancia P no considera las sustituciones múltiples, por lo que no toma en cuenta las diferentes tasas de cambio entre estas secuencias. Sin embargo el algoritmo de NJ si considera que cada linaje ha evolucionado con una tasa distinta. A partir de los resultados obtenidos por Distancia P, en el método de NJ se construye una matriz de distancias corregidas, donde se ajustan las longitudes de las ramas que relacionan un par de secuencias en base a la divergencia media de estas con respecto a todas las demás (Avice, 1994).

Estudios sobre la consistencia de distintos métodos de distancia cuando se analizan secuencias con tasas de cambio variables han concluido que si las distancias se corrigen en función de las sustituciones múltiples, el método de NJ es el que produce la mejor reconstrucción filogenética (DeBry, 1992). Aunque el método para calcular Distancia P no corrige las sustituciones múltiples, se realizó un tercer análisis filogenético considerando tasas de cambio variables al calcular las distancias por Corrección de Poisson y el árbol se construyó nuevamente por NJ en el programa MEGA (Kumar, 1993a) resultando en un árbol con la misma topología al obtenido por el método de Distancia P y NJ (Fig. 16).

La prueba de la tasa relativa nos ha proporcionado información sobre la variabilidad de las tasas de cambio de las proteínas C6-C9 con respecto a las perforinas. Estos datos no solo han ayudado a tomar decisiones durante el proceso de inferencia filogenética, también han subrayado tendencias de cambio que consideramos relevantes en la evolución de estas proteínas. Resulta interesante observar que mientras el C9 humano muestra tendencias de cambio distintas y con valores más altos que el resto de los C9, éstos se agrupan de manera muy similar y sus patrones de cambio son paralelas a las de ambos C8 $\beta$  (Fig. 12). Por otra parte C8 $\alpha$  y C8 $\beta$  que son proteínas con una estructura modular idéntica muestran tendencias de cambio opuestas, en contraste con los patrones de cambio paralelos de C6 y C7 que también son similares modularmente (Fig.3). Es difícil interpretar estos resultados, sin embargo es posible que las velocidades de cambio tengan relación con los contactos intermoleculares dentro del TCC o de la interacción de éstas proteínas con la membrana. En el TCC, C6 y C7 forman el complejo C5b-7; el complejo C8 se les une mediante C5b. La subunidad C8 $\alpha$  ayuda a insertar el complejo en la membrana y une la primer molécula de C9. Si una o varias moléculas se encuentran en contacto estrecho y de la estabilidad de su unión depende el funcionamiento del TCC, es posible que las subunidades del complejo que están en contacto entre sí muestren tendencias de cambio similares, pues si una subunidad cambiara más rápido o más lento que sus vecinas, el ensamble perdería estabilidad. Aparentemente, las tendencias de cambio de estas moléculas no guardan relación con la manera en que se realizan las uniones durante la formación del TCC, sin embargo se ha demostrado que las relaciones de estas moléculas entre sí y con membrana, cambian una vez que se ha ensamblado el TCC. Estudios de la interacción de C9 con el complejo C5b-8 en la membrana, usando marcadores hidrofóbicos, indican que mientras la incorporación de marca en

el C9 se mantiene constante, en el complejo C5b-8 disminuye a casi cero en cuando se aplica el marcador 30 minutos después de agregar el C9 (Hu, 1981). Estos patrones de marcado demuestran que el C9 expone sitios de unión a lípidos lo que induce un rearrreglo de los componentes tempranos C5b-8, lo que es confirmado con la observación de que C5b-8 en ausencia de C9 pueden ser marcados por las sondas que señalan a ésta molécula. De acuerdo a esta evidencia es posible que en el TCC insertado en la membrana, las subunidades se rearreglen de manera que tanto C6, C7 y C8 $\alpha$ , así como C8 $\beta$  y C9 entren en contacto o se relacionen de la misma manera con la membrana y por esta razón muestran las mismas tendencias de cambio evolutivo.

De acuerdo a las tendencias de cambio mostradas por los módulos CP y EGF es posible concluir que mientras el módulo CP marca las tendencias generales de cambio, el módulo EGF evoluciona a mayor velocidad lo que ocasiona que las tasas globales de cambio aumenten y en el caso de C6-C7 y C9 humanos se desvíen de una tasa constante. Estas velocidades locales de cambio sugieren que de acuerdo a su función, la región CP y el módulo EGF están sujetas a condiciones de selección que se ven reflejados en la manera en que dichas zonas evolucionan. En la región CP se encuentran la mayor parte de los sitios que se supone hacen contacto con la membrana, por lo tanto es posible que dicha zona requiera de mayor estabilidad estructural para conservar su función y por lo tanto su tasa de cambio local en C6-C9 es constante. Para confirmar un posible proceso de selección purificadora en la región CP es necesario analizar la relación entre las tasas de sustitución sinónima y no sinónima en los genes de esta proteínas. Si una selección purificadora está ocurriendo en la región CP la tasa de sustitución sinónima debiera ser mayor que la no sinónima (Li, 1997a).

En las comparaciones porcentuales de las velocidades relativas de C7, C8 $\alpha$ , C8 $\beta$  y C9 con respecto al C6, se observó que las tasas de cambio han tendido a acelerarse. Esto podría ser ocasionado por un relajamiento gradual de las condiciones de selección sobre alguna copia después de las duplicaciones que originaron a dichas secuencias. Para determinar las posibles causas de este aumento en la velocidad relativa es necesario analizar el tipo de sustituciones que se han presentado en la secuencia del gen, con la finalidad de detectar si se está llevando a cabo un proceso de selección positiva o si las mutaciones son de tipo neutral. Adicionalmente, es necesario contar con más moléculas de C6-C9 para confirmar si estas tendencias son particulares

del complemento humano o tienen un significado más amplio.

La RRT permite observar diferencias en las tasas. Con esta información es posible eliminar la secuencia o secuencias que cambian de modo desigual y hacer la inferencia con las secuencias restantes sin requerir correcciones para sustituciones múltiples. En el caso de las secuencias empleadas en esta tesis, eso significaría eliminar de los análisis al C9, C6 y C7 humanos lo que impediría, por exclusión inferir el orden de aparición de C6-C9.

La RRT también se emplea para determinar el tiempo en que divergieron las secuencias analizadas (Takezaki, 1995; Avise, 1994). Esto no se puede hacer con los datos actuales puesto que la tasa de cambio de las secuencias indican que no se comportan como reloj molecular cuya regularidad permita hacer las estimaciones.

La ubicación del C9 equino en los árboles generados por Parsimonia y distancia no coincide con el sitio que ocupa este organismo en un árbol de especies (Ver Anexo 5). Este hecho es independiente del uso de nucleótidos o aminoácidos y del método de inferencia empleado. Sin embargo, la filogenia de los mamíferos es aún motivo de controversia puesto que se trata de un grupo que sufrió una diversificación en muy corto tiempo (Benton, 1988. Kuma y Miyata, 1994). Acercamientos recientes para resolver las relaciones filogenéticas de Perissodactyla, Carnivora, Paenungulata, Primates y Cetartiodactyla sugieren que con la información disponible hasta el momento no es posible resolver el orden de aparición evolutivo de estos grupos de mamíferos. (Graur et al. 1997)

Es posible que el C9 equino y humano estén evolucionando con las mismas tendencias debido a las condiciones intrínsecas del sistema del complemento de ambas especies. Esta hipótesis se puede evaluar con la información ya disponible. Un análisis de la proporción de sustituciones sinónimas y no sinónimas de los C9 equino y humano permitirían distinguir si los cambios que han hecho que estas proteínas sean más cercanas se deben a mutaciones neutrales o a la existencia de un fenómeno de selección positiva en ambas moléculas.

Los resultados obtenidos en este proyecto permiten descartar el orden de aparición de los genes determinado con base en los módulos compartidos entre C6-C9 (Morgan, 1990c; Podack, 1988) (ver sección II). Esta propuesta tiene como idea principal la duplicación de un ancestro común a partir del cual se generaron inicialmente perforina y C9 y es el resultado de un momento en el estudio del sistema del complemento (1986-1992), donde no se habían establecido

claramente las funciones de señalamiento del TCC y sus complejos intermedios. Se desconocían las características de los C9 de peces y a pesar de que la estructura modular de los genes C9-C7 era conocida, no se había estudiado detalladamente su estructura de intrones/exones. Adicionalmente, la visión de la época con respecto a la evolución del complemento estaba influida por el entonces reciente descubrimiento de similitudes entre C9 y perforina y el estudio de las similitudes bioquímicas y funcionales de ambas moléculas. Posiblemente, estos factores hicieron que se sobrestimara la cercanía entre ambas moléculas. A esto se agrega la tendencia a pensar en los procesos evolutivos como una progresión a partir de formas sencillas hacia formas más complejas (Gould, 1989). De acuerdo a los propios autores de la propuesta, este concepto guió las interpretaciones e incluso resultó difícil de abandonar para los que proponen un orden de aparición que inicia con moléculas modularmente complejas (Hobart, 1993).

Otras funciones del sistema del complemento.

- Complemento y solubilización de complejos inmunes.

El sistema del complemento también tiene un papel importante en la solubilización de complejos inmunes y la respuesta óptima a antígenos. A causa de que las moléculas de anticuerpo tienen dos sitios de unión, la interacción de antígeno y anticuerpo muy concentrado resulta en la formación de una capa que crece hasta precipitarse. El sistema del complemento a través de C3b inhibe la formación de estos agregados insolubles en los tejidos, ya que esta molécula se intercala y rompe la capa de antígeno-anticuerpo.

Ya que más del 90% de CR1 en circulación se encuentra en eritrocitos, los complejos inmunes cubiertos de C3b rápidamente se unen a estas células, se evita su precipitación, el tamaño del complejo se reduce y se hace más soluble. Los eritrocitos también actúan como un sistema de transporte para complejos inmunes, llevándolos hasta macrófagos fijos en el sistema reticuloendotelial, donde éstos se internalizan y son degradados (Liszewski, 1993; Morgan, 1990a).

- Papel del complemento en el desarrollo de la memoria inmune.

Los linfocitos B expresan grandes cantidades de CR1 y CR2. El entrecruzamiento de estos receptores mediante agregados de C3b y C3d estimula la proliferación mientras que la molécula C5b activa la secreción de anticuerpos por estas células.

La anamnesis o memoria inmune es la capacidad del sistema inmune vertebrado para expresar una respuesta de anticuerpos aumentada ante una segunda o subsecuente administración de un antígeno en particular. Esta respuesta resulta de la generación de una población de linfocitos B de memoria durante la primera exposición al antígeno. El papel del complemento en este proceso ha sido sugerido por experimentos con animales timectomizados y depletados de C3. Estos organismos son incapaces de producir células B de memoria en respuesta a un antígeno. La respuesta primaria mediante IgM fue poco afectada por la falta de C3, pero durante la segunda

exposición en vez de generar una respuesta IgG, nuevamente se estimuló una respuesta con IgM, lo mismo sucede cuando C4 no está presente. De esta manera se prueba que el desarrollo de memoria inmune requiere de activación del complemento, no únicamente la presencia de C3. Aunque el papel preciso del complemento en la inducción de linfocitos B de memoria no se ha definido por completo, se ha mostrado que C3 facilita la localización de antígeno y la retención del mismo en los folículos linfoides. Las células foliculares dendríticas en los folículos linfoides del bazo y nódulos linfoides unen complejos inmunes que contienen fragmentos de C3 en su superficie. El antígeno de estos complejos es atrapado por receptores de complemento y Fc presentes en estas estructura y puede estimular la maduración de linfocitos B y la proliferación de una población de linfocitos B de memoria. En ausencia de C3, no ocurre la captura de complejo antígeno-anticuerpo y por lo tanto no se desarrolla memoria inmune (Liszewski, 1993; Morgan, 1990a).

#### - Sistema del complemento y coagulación

Entre las respuestas de plaquetas humanas mediadas directa o indirectamente por componentes activados del sistema del complemento están la secreción de enzimas proteolíticas y el ensamblaje de enzimas clave en las vías fibrinolítica y de la coagulación. Las moléculas del complemento que intervienen en estas vías son principalmente C1q y el complejo de ataque a membrana (Sims, 1991).

C1q modula la agregación de plaquetas inducida por colágeno ya que esta molécula inhibe la respuesta de las plaquetas a colágeno de forma competitiva. A su vez, polímeros de C1q agregada inician la agregación plaquetaria y desencadenar reacciones de forma similar al colágeno. Estas respuestas son mediadas por la unión de C1q a receptores de colágeno o de receptores específicos para C1q en la superficie de las plaquetas. A través de estas interacciones C1q puede inhibir la unión de plaquetas a colágeno vascular o puede funcionar como un activador de plaquetas cuando se une a complejos inmunes de forma polivalente (Sims, 1991).

Como ya se mencionó anteriormente, la visión de que el TCC tiene únicamente funciones líticas comienza a ser desafiada por evidencia experimental de que esta estructura puede modular las propiedades de permeabilidad de la membrana plasmática iniciando vías de transducción de señales que resultan en la modificación de las funciones celulares sin que esto desencadene la

muerte celular por lisis.

La participación de las proteínas terminales del complemento en la activación no lítica de plaquetas se estableció inicialmente en conejo donde se ha demostrado que la aceleración de las reacciones de coagulación o la inducción de agregación plaquetaria ocasionados por varios activadores del complemento se anulan en suero deficiente de C3, C5, C6 o C7 (Zimmerman TS, 1971). Estas y otras observaciones han sugerido que la activación del complemento puede evocar respuestas que imitan la estimulación plaquetaria por trombina y otros agonistas, estos efectos son mediados principalmente por la inserción de proteínas terminales del complemento en la superficie de plaquetas (Sims, 1991).

La inserción de C5b-9 en la membrana de las plaquetas en cantidades sublíticas (menos de  $10^3$  complejos) ocasiona un aumento en la conductividad al  $\text{Na}^+$  a través de la membrana plasmática y depolarización del potencial de membrana. La repolarización y posterior eliminación del TCC dependen de la entrada de  $\text{Ca}^{++}$  extracelular. Este ion dispara la fosforilación de substratos de protein cinasa C, miosin cinasa y otras cinasas intracelulares, además de la secreción de sustancias almacenadas en gránulos plaquetarios (Sims, 1991).

Las plaquetas activadas no solo forman un tapón hemostático, también proporcionan una superficie catalítica para el ensamblaje y activación de enzimas del sistema de la coagulación. La presencia de elementos de la vía alterna del complemento son fundamentales para la activación plaquetaria y secreción de gránulos, sin embargo aún no se conoce el mecanismo exacto por el que la activación del complemento inicia la activación de trombina y acelera el proceso de coagulación. Aparentemente el complejo C3bBb puede activar directamente a protrombina, lo cual representa un mecanismo potencial para la activación directa de la actividad procoagulante de plaquetas a través de trombina (Sims, 1991).

Por otra parte, la hemostasis vascular es regulada por interacciones entre plaquetas, leucocitos, endotelio vascular y proteínas de los sistemas del complemento, coagulación, de la quinina y fibrinólisis. Existe evidencia experimental de que parte de la relación entre mecanismos hemostáticos e inflamatorios es mediada directamente por interacciones complemento-plaquetas. El sistema del complemento es activado durante la coagulación sanguínea, lo cual sugiere la participación de las proteínas de este sistema en la formación de trombos.

La participación directa del complemento en reclutar plaquetas durante la hemostasis

normal ha sido demostrada en conejo. Aunque los mecanismos hemostáticos normales en humanos no parecen depender de la activación del complemento, existe evidencia de su activación puede alterar la función de plaquetas y contribuir al inicio de una enfermedad vascular (Sims, 1991).

## Anexo 2

*The history of life is endlessly fascinating,  
endlessly curious but scarcely the stuff of  
our usual thoughts and hopes.  
Cap. 1 Wonderful Life. Stephen Jay Gould.*

### Evolución del sistema del complemento

El sistema del complemento puede activarse por el reconocimiento de complejos antígeno anticuerpo. Sin embargo esta cascada tiene una capacidad intrínseca de reconocimiento entre propio y extraño dada por la vía alterna. El sistema de reconocimiento de la vía alterna funciona mediante la constante generación de productos activados derivados de C3. Normalmente en el suero y superficie de células homólogas existen moléculas que evitan la amplificación de esta cascada. En partículas extrañas que no tienen estos reguladores se forman C3 convertasas estables que por retroalimentación positiva amplifican la reacción. Una vez cubierto por C3b, la partícula ajena puede ser fagocitada o lisada por la vía terminal.

Aunque la vía alterna provee de inmunidad innata contra algunos microorganismos, mediante el reconocimiento de estructuras poliméricas simples, no es activada por estímulos debajo de cierto umbral, carece de memoria para contender futuros ataques y los organismos patógenos evolucionan mecanismos de evasión que superan rápidamente sus capacidades de reconocimiento (Farries, 1991).

Por lo tanto la vía alterna parece ser funcionalmente más simple que la clásica por lo cual la postura prevaleciente sobre la evolución del sistema del complemento sostiene que la vía de activación alterna es evolutivamente anterior a la vía clásica. Probablemente la vía alterna del complemento se originó como un sistema inmune independiente que en forma posterior adquirió capacidades de memoria y reconocimiento mediante la unión con anticuerpos y otras moléculas. Esta hipótesis está apoyada en que no existe en invertebrados un sistema inmune innato ya que no generan anticuerpos ni un equivalente funcional. Adicionalmente la molécula C3 se ha identificado en aves, reptiles, anfibios, peces óseos (osteichthyes), peces cartilaginosos (chondrichthyes) y Lampreas (ciclostomados) (Nonaka, 1992). A lo largo de estas especies se conservan la estructura de cadena y el enlace tioester de C3. Evidencia de la conservación funcional de esta molécula es que en lampreas, C3 se deposita sobre zimosan (paredes celulares

de levadura) y es reconocida como ligando para opsonización (Linthicum, 1970), lo que sugiere que el primer componente de la vía alterna ancestral pudo haber sido homólogo al C3 actual.

En estudios de inmunología comparada se ha buscado la generación de actividad hemolítica en respuesta a polisacárido o factor de veneno de cobra (CVF, una proteína derivada del C3 de cobra que se combina con factor B heterólogo para formar una C3 convertasa de la vía alterna). Mediante esta técnica se ha detectado actividad similar a la de la vía alterna en los vertebrados más tempranos como *Petromyzon marinus* (lamprea) y *Eptatrautus stroutii* (hagfish), así como en la hemolinfa de algunos invertebrados como la estrella de mar, el cangrejo *Lymulus polyphemus* y *Golfingea* (gusano cipuncúlido). Sin embargo no se han identificado componentes del sistema que produce esta actividad en invertebrados. La detección de actividad hemolítica en estas especies no implica la existencia de un sistema equivalente al del complemento (Day, 1970). Para comprender la evolución temprana del sistema del complemento es necesario elucidar cómo el C3 primitivo se asoció con una proteasa como el Factor B, y si existen proteínas homólogas a los receptores de C3 humanos en vertebrados tempranos.

En todos los vertebrados a partir de los chondrichthyes (tiburones) se observan rearrreglos génicos parecidos a los que ocurren en la formación de las inmunoglobulinas (Bartl, 1994; Litman, 1997). Probablemente el reconocimiento mediante anticuerpos apareció en una etapa temprana de la evolución de los vertebrados, pero en forma posterior a el sistema del complemento. Por lo tanto este sistema de reconocimiento altamente específico pudo haberse conectado al sistema de opsonización y liberación de factores inflamatorios representado por la vía alterna ancestral y generar la vía clásica del complemento. Por consideraciones funcionales se piensa que el puente que se unió estos dos sistemas pudo haber sido una molécula parecida a C1q. El organismo más antiguo donde se ha demostrado una vía clásica es el tiburón *Ginglymostoma cirratum* (Nurse shark) donde se han detectado componentes funcionalmente relacionados a C1, C2, C3, C4, C8 y C9 (Jensen, 1981).

La vía alterna y la clásica comparten una serie de componentes que ponen en evidencia su origen común. Ambos sistemas generan una C3 convertasa dependiente de cationes y que está integrada por dos subunidades, C3bBb o C4b2a que corta a C3 en C3a y C3b. Se sabe que C3 y C4 son proteínas que surgieron por duplicación de un ancestro común, las dos se caracterizan por tener una misma estructura de cadena y la capacidad de formar enlaces tioester, lo que les permite

unirse covalentemente a superficies biológicas, estas dos proteínas tienen virtualmente los mismos mecanismos de activación y degradación. Así mismo C2 y B los precursores de C2a y Bb son proteínas homólogas cuyos genes se encuentran ligados dentro de la región MHC de humanos y presumiblemente son resultado de una duplicación génica. Los reguladores negativos C4BP y factor H también son homólogos, están genéticamente ligados y se unen a las proteínas homólogas C4b y C3b respectivamente. Los Factores D y B de la vía alterna y C1q y C1r<sub>2</sub>C1s<sub>2</sub> de la vía clásica no son homólogos, sin embargo el corte del Factor B por el Factor D es funcionalmente equivalente al corte de C2 por C1s (Farries, 1991).

Se han propuesto una serie de eventos que pudieron darse durante la evolución de la vía clásica (Farries, 1990; 1991). El primer componente de la vía clásica debió ser capaz de aumentar la inmunidad preexistente. La molécula C1q satisface este criterio pues acopla en su estructura el reconocimiento de antígeno por IgM o IgG a receptores de C1q en células sanguíneas periféricas (Tenner, 1981). C1q está compuesto de dominios globulares de adhesión reunidos a través de tallos colagénicos. Por lo tanto C1q y su receptor o alternativamente C1q membranal, pudieron haber evolucionado de moléculas de adhesión celulares a receptores de IgM. C1q también se puede unir a superficies de lipopolisacáridos en ausencia de anticuerpo. La unión de C1q a anticuerpos pudo acoplarse a la vía alterna si esta actividad estaba asociada a una actividad capaz de cortar C3. *In vitro*, proteasas plasmáticas como tripsina y plasmina convierten a C3 en un fragmento parecido a C3b. Por lo tanto una proteasa plasmática fue probablemente el progenitor del complejo C1r<sub>2</sub>C1s<sub>2</sub>. La adición de actividad proteolítica de C3 a el complejo C1 hace que C3b se deposite alrededor de los sitios de reconocimiento por anticuerpo. Este C3b se une a Bb e inicia la vía alternativa que aumenta la deposición de C3b. Posteriormente a través de la aparición de C4 a partir de una duplicación de C3 surgió un mecanismo de amplificación independiente de la vía alterna. Probablemente C4 resultó ser un substrato más específico para la acción proteolítica de C1s y aumentó la eficiencia de activación del complemento. Finalmente la amplificación y selectividad de la naciente vía clásica aumentaron por la evolución de C2 a través de duplicación génica del factor B. El efecto de C4b en la generación de C3b fue aumentado por su combinación con C2 para formar una C3 convertasa de la vía clásica: C4b2a. Posteriormente estos componentes coevolucionaron para llegar a ser la cascada clásica moderna. Evidencia que apoya esta teoría es que se ha observado *in vitro* que se puede activar la vía alterna del

Con respecto a la evolución de la vía de las lectinas, se puede proponer en base a una serie de puntos relevantes en común entre MBP y C1q, que la vía dependiente de lectinas se ubica evolutivamente en una fase intermedia en el desarrollo de la vía clásica a partir de la vía alterna. C1q y MBP tienen una estructura cuaternaria terciaria que les permite interactuar con superficies polivalentes e indirectamente activar sus respectivas serin proteasas C1r<sub>2</sub>-C1s<sub>2</sub> y MASP. La enzima MASP en humanos guardan 37 y 36% de homología con C1r y C1s por lo que se presume estas proteasas son el resultado de la duplicación de un solo gen ancestral, adicionalmente C1r<sub>2</sub>-C1s<sub>2</sub> y MASP pueden proteolizar los mismos precursores de C4b y C2a, sin embargo C3 también es sustrato de MASP, lo que hace que a diferencia de la vía clásica la vía de las lectinas, pueda iniciar la cascada del complemento a través de la formación de C3b, como sucede en la vía alterna. Tanto C1q como MBP pueden funcionar como opsoninas al interactuar con el mismo receptor. Sin embargo ambas moléculas tienen capacidades de reconocimiento diferentes, MBP está limitado al reconocimiento de una amplia variedad de oligosacáridos presentes en superficies bacterianas, mientras C1q a través de sus seis sitios de unión interactúa con IgG e IgM y de este modo amplía su capacidad y magnitud de reconocimiento a todos el repertorio de antígenos para los que existen anticuerpos en el sistema inmune. Tanto la vía de las lectinas como la clásica están reguladas a nivel de MBP y C1q por la esterasa C1INH.

Por lo anterior la vía clásica representa un refinamiento de los mecanismos de reconocimiento y un aumento en la especificidad de activación de la vía de las lectinas. Mientras la vía de las lectinas puede activar la vía alterna y la clásica sin requerir anticuerpos, la vía clásica se sirve de un mecanismo especializado de reconocimiento a través de anticuerpos, que inicia una sola vía efectora. El organismo más antiguo en que se ha identificado la proteasa MASP de la vía de las lectinas es *Halocynthia roretzi* (Ji, 1997). Los tunicados a los que pertenece esta especie ocupan un sitio intermedio entre los invertebrados y los vertebrados. En análisis filogenéticos de las secuencias de aminoácidos de MASP humano, murino y de *H. roretzi*, así como C1r y C1s de mamíferos, se propone que por duplicación de un gen ancestral común, se generaron MASP de tunicados y de mamíferos, una duplicación posterior de MASP mamífero originó a C1r y C1s de la vía clásica. El orden de aparición de las serin proteasas apoya la idea de que la vía clásica

surgió en forma posterior y como una especialización de la vía de las lectinas. Sin embargo la identificación de MASP en *H. roretzi* ha servido para proponer que la vía de las lectinas es el mecanismo original de activación del sistema del complemento, este planteamiento se basa en que hasta el momento no se han encontrado componentes de otras vías del complemento en organismos aún mas antiguos. Este planteamiento no se puede considerar válido por que la identificación de un componente no está relacionado directamente con la existencia de una cascada completa. La identidad de 28% a nivel de residuos de aminoácido entre MASP tunicado y mamífero, no implica una función común, solo manifiesta la existencia de un ancestro común. Asimismo no se puede suponer que en organismos más antiguos o en el mismo *H. roretzi* no existan otras vías de activación del complemento pues los métodos empleados para detectar actividad de complemento (por lisis de eritrocitos heterólogos activada mediante veneno de cobra) están basados en sistemas mamíferos por lo que pueden dar falsos negativos al no considerar las diferencias funcionales, estructurales y de regulación intrínsecas a otras especies.

La vía dependiente de C1 también representa una etapa evolutiva previa al reconocimiento mediante anticuerpos. Resulta más cercana funcionalmente a la vía clásica que la vía de las lectinas, pues ya incluye a la proteasa C1r<sub>2</sub>-C1s<sub>2</sub>. Hasta el momento la vía dependiente de C1 no ha sido descrita con detalle en mamíferos y se desconoce si existe un sistema equivalente en otras clases de vertebrados.

En este contexto la vía terminal coevolucionó con las vías de activación desde una etapa muy temprana del desarrollo de éstas. El desarrollo de la vía terminal fue posible gracias a duplicaciones del gen ancestro de C3, C4 y C5 (Hughes, 1994). Mediante C5b, la molécula C6 se vinculó con las cascadas de activación del complemento. Mediante posteriores duplicaciones y pérdidas de regiones del gen de C6 surgieron C7, C8 $\alpha$ , C8 $\beta$  y finalmente C9. La duplicación de estos genes dio origen a moléculas con nuevas funciones y con capacidad de unión a los duplicados que surgieron previamente. Posiblemente la capacidad de C7 para insertarse en membrana y fue resultado de la pérdida del módulo TSP amino terminal, esto propició que el complejo C5b-7 adquiriera la capacidad de transducir señales por su interacción con proteínas G. La pérdida de módulos que se dio en los posteriores duplicaciones de estas familias de genes conservó la capacidad de interacción entre estas proteínas y permitió que se insertaran en membranas de forma estable y paulatinamente generaran poros que aumentaron la capacidad de

señalamiento del complejo. Las proteínas de la vía terminal posiblemente no formaron un sistema previo e independiente al de las vías de activación del complemento, ya que su inserción en membrana y polimerización es catalizado por C5. Por esta razón la evolución de la vía terminal fue un proceso simultáneo y dependiente del desarrollo de las vías de activación, esto justifica el uso del término coevolución para describir la manera en que ambos procesos se dieron de forma paralela. Cabe subrayar que la evolución de la vía terminal, así como cualquier otro proceso evolutivo no estuvo guiado por una función que debía mejorarse, las pérdidas de módulos se dieron como parte de la dinámica de un proceso de duplicación de genes, donde las copias no estaban sujetas a las mismas presiones funcionales y mutacionales del gen original (Fig. 19\*).

---

\* Fig. 19 (En la siguiente página) Evolución del sistema del complemento. Relación entre los eventos claves que dieron origen a las cuatro vías de activación y a la cascada terminal. Nótese que posiblemente muchos de los componentes que surgieron en la vía alterna sirvieron a la de las lectinas, Asimismo, elementos de ésta última fueron la base para la formación de la vía clásica. El mecanismo fundamental que ha guiado la evolución de estos sistemas es la duplicación de genes y posterior cambio de funciones (diversificación). La evolución de las proteínas que regulan este sistema sale de los alcances de este proyecto.

Sistema de marcado, fagocitosis y anafilaxis  
C3 ancestral, proteasa y receptores de C3.

Vía alterna  
Formación de C5 por  
C3 ancestral, Factor  
B y receptores de C3  
de C3 y C4.

Se origina la C3 convertasa  
por unión de C3b a la convertasa  
de C3 que ya existía, así se forma  
el complejo C3bBbC3b

C5 es cortado a C5b

Unión de C5b a C6.

Coevolución de la vía terminal con esta y las  
las vías de activación del complemento que  
surgieron posteriormente. Comienza el  
desarrollo de un sistema de señalamiento  
basado inicialmente en proteínas G y  
posteriormente por la entrada de Ca.

Duplicaciones del Factor B

y C3 ancestral originan a C2 y C4  
respectivamente. Estas moléculas  
tienen funciones de C3 convertasa  
homólogas a las de las proteínas de  
donde surgieron

Surge la proteasa ancestral  
de MASP, se une a la lectina  
MBP y puede cortar C3  
y a su duplicado C4

Formación de la Vía de las lectinas  
por asociación de C4, C2, C3, MBP  
y MASP.

La convertasa de C3 de esta vía  
es el complejo C4b2a, el cual como  
su homólogo de la vía alterna une  
C3b y forma una C5 convertasa

El C5b generado en esta  
vía se une con la cascada terminal.

MASP se duplica y genera  
a C1s y C1r, serin proteasas.  
con capacidad de cortar C4

C1q, una proteína estructuralmente  
similar a MBP une y activa a C1s y C1r.  
Se forma el complejo C1

Vía dependiente de C1.

El complejo C1 corta C4 y C2  
de forma similar a como lo hace  
MASP, sin embargo C1q reconoce  
complejos polianiónicos por lo que  
es un mecanismo de activación  
diferente unido a la vía terminal.

Aparición de moléculas de  
anticuerpos en alguna etapa de la  
evolución de los ciclostomados.

Aparece la Vía Clásica. Se integran  
en un sistema de reconocimiento alta  
mente específico, las moléculas de  
anticuerpo, con el sistema formado por  
C1, C4, C2, C3 y la cascada terminal.

## Conceptos básicos sobre Evolución Molecular .

### - ¿Qué es Evolución Molecular?

A principios de la década de los sesenta comenzaron a aplicarse modelos conceptuales y técnicas de Biología Molecular al estudio de la evolución biológica. Este nuevo acercamiento reveló aspectos novedosos de los procesos evolutivos y métodos experimentales para la reconstrucción de la historia evolutiva. A estas tendencias se sumaron una creciente cantidad de información sobre secuencias proteicas y una serie de importantes estudios comparativos sobre la tasa y patrón de sustitución de aminoácidos en diversos organismos. La convergencia de estas circunstancias ocasionó una revolución en Genética de Poblaciones, ya que la disponibilidad de secuencias de aminoácidos eliminó la barrera de especies que tradicionalmente había limitado los estudios en esta área a grupos de organismos que podían cruzarse entre sí. El descubrimiento de que el árbol reconstruido a partir del Citocromo c era similar al árbol de especies conocido entonces estimuló tanto el desarrollo de métodos para la reconstrucción de árboles filogenéticos, como el interés por estudiar las relaciones evolutivas entre secuencias proteicas. A esta nueva disciplina cuyo objeto de estudio es la evolución biológica a nivel macromolecular se le llamó Evolución Molecular (Ayala, 1976).

En las últimas dos décadas el desarrollo de la Evolución Molecular se ha acelerado gracias al surgimiento de técnicas como el análisis de restricción de secuencias de DNA, la clonación de genes, la reacción en cadena de la polimerasa, la secuenciación de DNA, así como el rápido desarrollo de computadoras con mayor capacidad y velocidad para procesar información, las cuales han permitido realizar análisis de inferencia filogenética cada vez más rigurosos.

En forma general la Evolución Molecular comprende dos áreas: el estudio de la evolución de macromoléculas y la reconstrucción de la historia evolutiva de genes y organismos. En la primer área se estudian las tasas y patrones de cambio de DNA, RNA y proteínas a través del tiempo evolutivo y los mecanismos responsables de dichos cambios. La segunda área, conocida como Filogenética Molecular estudia la historia evolutiva de los organismos y macromoléculas a partir de información molecular. Estas dos áreas están relacionadas estrechamente por que el conocimiento del orden de aparición de un grupo de organismos es necesario para ordenar

temporalmente los cambios en los caracteres moleculares estudiados y posteriormente inferir su causa. En forma similar, el conocimiento sobre los patrones y tasas de cambio de una molécula dada es indispensable para diseñar métodos de reconstrucción de la historia evolutiva de un grupo de organismos o moléculas (Li, 1997c).

#### - Supuestos fundamentales de la Filogenética molecular

Un proceso de inferencia filogenética se basa en los siguientes supuestos:

- Todos los organismos están unidos por relaciones de ancestro-descendiente. Esta secuencia de eventos es históricamente única. Por lo tanto uno solo uno de todos los posibles árboles que se pueden construir con un grupo dado de secuencias, representa la verdadera historia evolutiva.
- En los organismos actuales existen caracteres en común con un ancestro y estos caracteres son heredables.
- Existe un proceso de evolución definido por cambios en las características ancestrales

La transmisión estable de los caracteres ancestrales está mediada por el genoma. Es decir, las diferencias debidas a factores epigenéticos, o ambientales no proporcionan información sobre las relaciones ancestro-descendiente entre organismos.

A partir de lo anterior se deriva que la historia evolutiva de un grupo de organismos es un registro detallado de las mutaciones que se fijaron en cada linaje genómico. Si un método de inferencia filogenética contara con la información completa sobre la historia evolutiva, obtendría en forma consistente el árbol que representa la filogenia real. Como no se dispone de toda esta información, la representación obtenida por un determinado método se llama árbol inferido, que puede o no ser idéntico al árbol real. Es decir, el inferir una filogenia es hacer la mejor estimación de una historia evolutiva en base a información incompleta contenida en las secuencias de aminoácidos o nucleótidos. En ausencia de la información completa sobre el proceso evolutivo estudiado, se pueden emplear modelos hipotéticos de dicho proceso para derivar métodos de inferencia que obtengan consistentemente la filogenia real, suponiendo que el modelo empleado es el correcto. Sin embargo los datos empleados para obtener una filogenia no tienen que apegarse estrictamente a un modelo para que los resultados sean válidos. Bajo circunstancias especiales ciertos supuestos no son necesarios y el método puede resistir que éstos sean transgredidos (Swofford, 1996).

## - Alineación de secuencias

El primer paso en la obtención de una filogenia es la comparación de los sitios comunes entre las secuencias de proteínas ó aminoácidos analizadas. Este proceso de alineación implica el reconocimiento de las posiciones donde durante la evolución se han dado eventos de inserción ó deleción. Estas posiciones son representadas como espacios ó "indels" (insertions-deletions) (Li, 1997d). En la presente descripción de este proceso se da por sentado que las secuencias a alinear son homólogas, es decir se derivan de un ancestro común. La identificación de residuos homólogos entre dos o más secuencias así como los problemas metodológicos relacionados con su alineación son independientes del proceso de inferencia filogenética. Sin embargo la obtención de las relaciones evolutivas entre un grupo de secuencias depende fundamentalmente del proceso de alineación. Si la alineación de dos secuencias se considera como una serie de pares de residuos, uno de cada secuencia, se pueden distinguir pares de sitios con residuos iguales, con residuos desiguales ó con un espacio y un residuo. Los pares iguales manifiestan que en ese sitio no han ocurrido sustituciones desde la divergencia de las dos secuencias, un par desigual implica que por lo menos ha ocurrido una sustitución y un espacio representa una inserción o deleción en dicho sitio. Por lo tanto el proceso de inferencia filogenética depende de la alineación ya que ésta representa una hipótesis acerca de la evolución de un grupo de secuencias (Li, 1997d).

Existen diversos métodos para alinear secuencias de aminoácidos y nucleótidos: el método de matriz de puntos, el de distancia, el de similitud, etc. El principio fundamental de estos procedimientos consiste en obtener el máximo número de pares iguales entre dos secuencias o minimizar el número de pares desiguales, haciendo que en ambos casos el número de los espacios sea mínimo. El método que se empleó para alinear las secuencias C6-C9 y perforina es el conocido como CLUSTAL (Higgins, 1988). El algoritmo que sigue el programa donde se implementó este método es el siguiente:

1. Se introducen las secuencias
2. Cálculo de las similitudes entre secuencias.
3. Las distancias obtenidas se analizan por el método UPGMA que forma grupos con las secuencias que tienen menor número de diferencias entre si.
4. Se toman los dos grupos o secuencias (restantes) más similares y se les alinea.
5. Se establece un consenso y se muestran el par de secuencias alineadas y con espacios.

6. ¿Hay más secuencias que alinear? Si = Ir a paso 4, No = Ir a 6

7. Mostrar la alineación final.

Este algoritmo procede en tres etapas. a) Primero se calculan las similitudes entre todos los pares de secuencias mediante el algoritmo de Wilbur y Lipman. Los valores de similitud son iguales al número de residuos idénticos entre dos secuencias en la mejor de las alineaciones menos un valor fijo por cada espacio que se inserte (pasos 1 y 2). b) Con la matriz generada en el paso anterior se construye un árbol por el método de UPGMA, que maneja matrices muy grandes rápidamente. Este método busca en una matriz de distancia el valor más pequeño y une las dos secuencias AB implicadas y ubica esta relación a lo largo de un eje que contiene los valores de distancia. Posteriormente este primer par es eliminado de la matriz. En la nueva matriz se busca el valor más pequeño, si este valor relaciona una secuencia que ya se unió con otra por ejemplo el par BC, la secuencia C se unirá a AB mediante una rama cuyo extremo se ubica en el eje de los valores de distancia al nivel de la media aritmética de las distancias entre C y AB. Esta forma de agrupar las secuencias procede hasta que se ha examinado el último valor de la matriz (Avisé, 1994) (paso 3). c) Finalmente se alinean pares de secuencias siguiendo el orden en que estas se agrupan en el árbol obtenido en la etapa anterior. Se emplea nuevamente el algoritmo de Wilbur y Lipman. Después de alinear cada par de secuencias se obtiene un consenso que representará a este par en posteriores alineaciones entre grupos (pasos 4-7) (Avisé, 1994).

Las ventajas de emplear el algoritmo CLUSTAL implementado en el programa DNAMAN son que además de ser un procedimiento rápido que genera resultados consistentes, posee herramientas para visualizar tanto el árbol de la etapa 2 como la alineación final, además permite corregir la secuencia manualmente. Este programa se empleó en una estrategia donde en cada ocasión se agrega una secuencia más a la alineación, se corrige el resultado mediante alineaciones ya publicadas para los módulos ricos en cisteínas, y se agrega otra secuencia al banco de datos. La alineación final se sometió a otra etapa de refinamiento donde después de emplearla en programas de inferencia filogenética se le inspecciona visualmente, se hacen las correcciones necesarias y se vuelve a obtener la filogenia. Este proceso alineación-inferencia-alineación finaliza cuando en forma consistente se obtiene el mismo árbol.

## 2. Métodos de inferencia filogenética

### - Distancia y Parsimonia

En filogenia molecular no se tiene información directa del pasado, únicamente hay acceso a especies y moléculas contemporáneas. Con esta información se pueden proponer condiciones evolutivas en las que cualquier filogenia posible produce las secuencias actuales. Los métodos de inferencia filogenética buscan establecer una base para elegir de entre todos los árboles posibles, el o los que mejor describan la historia evolutiva de un grupo de secuencias. Lo anterior se establece a través de dos criterios: A) Definir una secuencia específica de pasos (algoritmo) para generar un árbol. B) Precisar un criterio para comparar distintas filogenias y determinar cual es la mejor o si existen varias igualmente buenas (Swofford, 1996).

Los métodos algorítmicos combinan en un solo paso la inferencia del árbol y la definición del criterio para determinar cuál es la mejor filogenia. En estos métodos están incluidos los llamados de distancia. Generalmente este tipo de métodos tienden a ser computacionalmente rápidos pues buscan directamente la solución sin evaluar una gran cantidad de árboles.

Los métodos basados en el criterio B proceden en dos pasos: Primeramente definen un criterio de lo que es óptimo (función objetiva) para evaluar un árbol dado, esto es un valor numérico que se asigna inicialmente y que después se usa para comparar un árbol con otro. El segundo paso es usar un algoritmo específico para calcular el valor de la función objetiva y encontrar los árboles con el mejor valor de acuerdo con esta función. De esta manera los supuestos evolutivos hechos inicialmente se separan del procedimiento computacional hecho en el segundo paso. La consecuencia de buscar el árbol que optimice una función objetiva es que estos métodos son más lentos que los de distancia. Entre estos métodos se encuentra el de Máxima Parsimonia. La diferencia entre el uso de algoritmos por el primer tipo de métodos y el segundo es que en los primeros el algoritmo define el criterio de selección de árboles, mientras en el segundo tipo los algoritmos son usados para evaluar la función objetiva y buscar árboles que la optimizen (Swofford, 1996).

### - Máxima Parsimonia

En el método de Máxima Parsimonia (MP) las secuencias de nucleótidos o aminoácidos para las especies ancestrales son inferidas a partir de las especies actuales y se produce un árbol

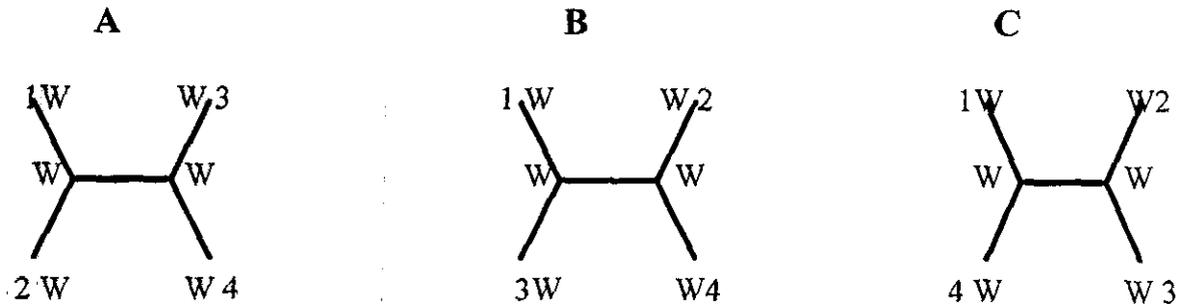
en donde se minimizan el número de cambios evolutivos (Nei, 1987). En MP los sitios de una alineación de secuencias se dividen en informativos y no informativos. Un sitio informativo es el que favorece una reconstrucción filogenética sobre otras. En esta clase de sitios por lo menos dos tipos diferentes de aminoácidos están representados por lo menos dos veces. Otros sitios variables no se usan para reconstrucción filogenética por MP, aunque son informativos para los métodos de distancia o máxima verosimilitud (Kumar, 1993b). En la figura 20 se ejemplifica el concepto de sitio informativo.

Cuando se habla sobre MP se debe distinguir entre el criterio de lo óptimo (longitud mínima del árbol) y el algoritmo para buscar el árbol óptimo (Swofford, 1996). En el proyecto aquí presentado, se calculó la longitud mínima del árbol con el criterio de Máxima Parsimonia, este es el criterio de lo óptimo. Posteriormente se empleó el algoritmo de "Branch and Bound" para buscar de entre los 23,614,056,982 árboles posibles el que satisface el criterio de lo óptimo. En el cálculo de la longitud mínima del árbol bajo el criterio de parsimonia, cada posición de la alineación se considera en forma independiente de las demás, por lo tanto la longitud total del árbol se calcula al sumar las longitudes mínimas obtenidas para cada posición informativa. En el método de MP se infieren los aminoácidos de las secuencias ancestrales de cada sitio para una topología dada, y se cuenta el número mínimo de sustituciones que se requieren para explicar cada cambio. La suma de este número para todos los sitios informativos de las secuencias para una topología dada es la longitud del árbol. Consideremos el cálculo de la longitud mínima de dos posiciones informativas en cuatro secuencias de aminoácidos (Fig.20). La topología con menor número de sustituciones o pasos es la que requiere de un menor número de supuestos para explicar las relaciones filogenéticas entre estas secuencias. Esto no quiere decir que las explicaciones no parsimoniosas sean falsas, sino que en ausencia de más información sobre la evolución de estas proteínas, se elige la explicación más sencilla. En las posiciones 2 y 4, el árbol A es el que requiere de menor número de sustituciones y por lo tanto el más parsimonioso. La suma del número de sustituciones por cada sitio informativo nos da la longitud total de este árbol, que es de 2 pasos. El árbol mas parsimonioso es el apoyado por un mayor número de sitios informativos.

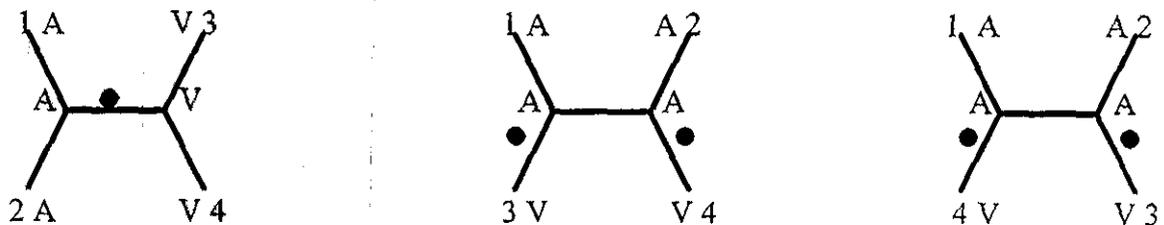
Cuando el número de secuencias es menor a 10, es factible examinar todos los árboles posibles mediante una búsqueda exhaustiva. En nuestro caso no es factible examinar todas las

- 1 WAFWC
- 2 WAVWC
- 3 WVTDC
- 4 WVPDC
- \* \*

Posición 1



Posición 2



Posición 4

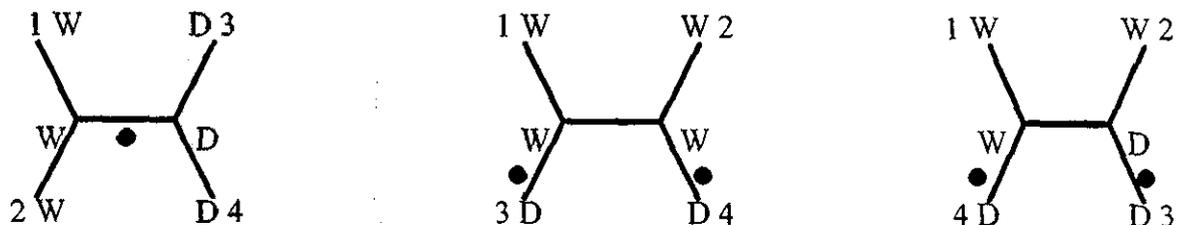


Fig. 20. Ejemplo de reconstrucción por MP. Se tiene una alineación de cuatro secuencias con cinco posiciones. La relación filogenética entre estas secuencias puede describirse por tres filogenias distintas. Sólo las posiciones 2 y 4 son informativas para el criterio de parsimonia ya que en ellas sólo un árbol es el marcado como más parsimonioso. La posición 1 no se considera informativa por que cualquiera de las tres filogenias tiene el mismo número de sustituciones (0), por lo tanto esta posición no proporciona información sobre cual de las tres filogenias es la más parsimoniosa. Las posiciones marcadas en los extremos de las ramas de los árboles de esta figura representan las secuencias actuales. En los dos nodos internos se representan solo algunas de los posibles estados ancestrales. Se marca con un punto la rama donde tuvo que darse una sustitución de los estados ancestrales para generar las secuencias actuales.

topologías posibles para 16 secuencias (en condiciones óptimas este cálculo llevaría 300 años). La estrategia que seguimos para buscar el árbol que cumpliera con el criterio de lo óptimo fue emplear el método de "Branch-and-Bound". En este método todos los árboles con una longitud mayor al examinado en un determinado paso son ignorados. El árbol que satisface el criterio de MP se encuentra a través de evaluar las longitudes de un grupo de árboles que potencialmente tienen una longitud menor.

Una búsqueda con el método de "Branch-and-Bound" comienza con un árbol de tres secuencias, una a una, otras secuencias son agregadas a este árbol inicial de acuerdo a el orden de adición "furthest". Cada vez que se agrega una secuencia se calcula la longitud total del árbol. Si la adición de una secuencia a una rama en particular del árbol inicial resulta en un árbol cuya longitud es mayor a el criterio de lo óptimo previamente determinado, esta topología y las subsecuentes que se generen por adición de secuencias a este árbol inicial serán eliminadas del análisis. Si se encuentra un árbol con longitud menor al criterio de lo óptimo empleado hasta el momento, este criterio es reemplazado por el del árbol con menor número de pasos. Este algoritmo ahorra una cantidad considerable de tiempo al eliminar árboles que no es necesario examinar, especialmente si el árbol inicial tiene una longitud cercana a la del árbol más parsimonioso (Kumar, 1993b).

A pesar de sus ventajas, el método de MP está sustentado en supuestos que no son válidos en todos los casos. El método de MP no toma en cuenta las sustituciones múltiples, lo cual puede ser un problema si las tasas de cambio entre secuencias son muy distintas y existen linajes que han acumulado más cambios que otros a lo largo de la evolución. Estas diferencias en tasas de cambio puede hacer que proteínas más cercanas parezcan como más lejanas y viceversa.

Por otra parte, el método de MP al prescindir de los sitios que considera no informativos, siempre elimina parte de la información suministrada para el análisis, mientras los métodos de distancia emplean todos los sitios de una alineación.

#### - Métodos de Distancia.

En este proyecto se calculó el número de sustituciones de aminoácidos por sitio, o distancias entre todos los pares de secuencias mediante los métodos de Distancia P y Corrección de Poisson, implementados en PAUP y MEGA respectivamente. La Distancia P es el número de

aminoácidos distintos entre dos secuencias ( $n_d$ ) dividido entre el número total de aminoácidos comparados ( $n_t$ ). Lo anterior se expresa matemáticamente como:

$$p = n_d / n_t \quad (1)$$

Esta proporción no considera las sustituciones múltiples que se pueden dar en cada posición, por lo tanto considera que las tasas de sustitución entre secuencias son iguales.

En el cálculo de distancia por Corrección de Poisson se supone que el número de sustituciones de aminoácidos en cada sitio sigue una distribución de Poisson. La expresión matemática del cálculo de distancia por Corrección de Poisson es:

$$d = -\log_e(1-p) \quad (2)$$

donde  $p$  es estimada por la ecuación (1)

Se considera que un modelo matemático complejo se ajusta mejor a los datos que un modelo sencillo. Sin embargo un modelo complejo requiere de la estimación de más parámetros lo cual aumenta la varianza de la distancia estimada (Kumar, 1993c). Fundamentalmente la elección de los métodos para el cálculo de distancias estuvo dada por su disponibilidad en PAUP y por que dadas las tasas de cambio variables de C6-C9 y perforina, resultó de interés hacer los análisis de inferencia filogenética con un modelo que corrigiera las distancias considerando las sustituciones múltiples.

En las dos maneras de calcular la distancia se empleó el modo de delección por parejas, esto significa que en la comparación entre dos secuencias se eliminan los sitios que en cualquiera de las dos contienen un espacio.

A partir de las matrices de distancias obtenidas mediante los métodos mencionados se reconstruyó el árbol filogenético con el método de "Neighbor-Joining". Con este método en base a la matriz de distancia se construye en cada paso una matriz modificada en la que la separación entre cada par de nodos es ajustada en base a su divergencia promedio de todas las demás secuencias. En este método la topología inicial es un árbol tipo estrella a partir del cual se construye progresivamente un árbol final a través de unir el par de nodos menos distante definidos por la matriz corregida (Swofford, 1996). El par de secuencias unidas en el paso anterior son sustituidas en una nueva matriz, por el valor del nodo que las une. El proceso continúa por  $n-1$  ciclos hasta que todas las secuencias son sustituidas por nodos internos (Avice, 1994). El

“Neighbor-Joining” está clasificado como un método algorítmico por que construye solo un árbol y no optimiza ninguna función objetiva (Swofford, 1996).

#### - Prueba de “Bootstrap”

La técnica de “Bootstrap” se usa para estimar el nivel de confianza de una reconstrucción filogenética. La hipótesis nula a probar es la topología de una determinada filogenia. Para estimar los niveles de confianza de cada una de las relaciones descritas en un árbol se generan un número  $N_b$  de muestras aleatorias con reemplazo de las secuencias originales. Cada muestra se genera tomando al azar una posición de la alineación original, hasta que la muestra tiene el tamaño de la alineación original. La muestra se emplea para construir un árbol con el método elegido. A este procedimiento se le llama una repetición y se ejecuta  $N_b$  de veces. Cada vez que la relación entre dos o más secuencias aparece se le dará un punto. Una vez que la etapa de muestreo termina se combinan los resultados de todas las repeticiones en un árbol consenso, en cada clado de esta filogenia se muestra el porcentaje de repeticiones en que apareció dicha relación (Li, 1997b). Un consenso del 50% como el empleado en este proyecto significa que un patrón de enramado que ocurre con una frecuencia  $>50\%$  será considerada para el cálculo del consenso final.

Uno de los supuestos de la prueba de Bootstrap es que la alineación de la que se parte representa el espacio de muestra original, de modo que cada muestra puede considerarse como parte del espacio original. Para que este supuesto se cumpla, la alineación original tiene que ser lo suficientemente grande. El muestreo aleatorio se hace con reemplazo, de lo contrario todas las muestras serían iguales pero con un orden distinto, lo cual no produciría ninguna variación entre las muestras (Li, 1997b).

Estudios sobre el “Bootstrap” revelan que los resultados de este método proporcionan una medida conservadora sobre la probabilidad de inferir correctamente las relaciones filogenéticas representadas en un árbol. Se considera que valores de “Bootstrap”  $\geq 70\%$  usualmente corresponden a una probabilidad  $\geq 95\%$  de que el clado al que pertenecen sea real, siempre y cuando el método que se empleó para hacer la reconstrucción filogenética sea consistente. Sin embargo en condiciones de tasas de cambio altamente desiguales entre secuencias y tasas de cambio internodal aleatorias, los valores de Bootstrap  $> 50\%$  sobrestiman la exactitud de la reconstrucción filogenética (Hillis, 1993). Sin embargo aún se debaten las consideraciones sobre

los resultados que la prueba de bootstrap debe generar para considerar que cada relación filogenética está firmemente establecida (Efron *et al.* 1996; Brown, 1994 y Sitnikova *et al.* 1995)

En el presente proyecto se aplicó esta prueba por que permite evaluar cuantitativamente la confiabilidad de los árboles obtenidos y posibilita la comparación de las filogenias generadas con diferentes métodos. Adicionalmente el procedimiento y significado de los resultados son fáciles de abordar por un público diverso.

#### - Enraizado y grupos externos

Una filogenia con raíz es un árbol donde se ha identificado el punto más temprano en el tiempo, es decir, el sitio donde se ubica el ancestro común de todas las secuencias analizadas. Para enraizar un árbol usualmente se incluye en el análisis una o más secuencias a las que se denomina grupo externo. Las secuencias del grupo externo previsiblemente se ubican fuera del resto de las secuencias ó grupo interno. A su vez se presume que el grupo interno tiene un solo origen evolutivo, es decir que es monofilético. La razón de incluir más de una secuencia en el grupo externo es que de esta manera se puede evaluar mediante un análisis filogenético si el grupo interno efectivamente es monofilético. Si las secuencias del grupo externo no se ubican junto con las del grupo interno se acepta la hipótesis de que el grupo interno es monofilético (Swofford, 1996).

En un análisis de inferencia filogenética se incluye un grupo externo como punto de referencia temporal con respecto al que se ordenan las secuencias del grupo interno. En este sentido no es relevante que el grupo externo este formado por secuencias que hayan surgido antes o después de las moléculas del grupo interno. Las secuencias que forman el grupo externo de una filogenia de moléculas generalmente son secuencias homólogas a las del grupo interno pero que han divergido lo suficiente para constituir un grupo separado y que conserva una estructura primaria que permite alinearlas confiablemente con el grupo interno. En contraste para obtener una filogenia de especies se emplea como grupo externo la misma molécula con la que se está realizando el análisis, pero perteneciente a especies que se sabe están separadas evolutivamente del grupo estudiado.

El objetivo del presente estudio es conocer el orden de aparición evolutivo de las proteínas C6-C9. Como punto de referencia temporal se empleó a las perforinas de rata, ratón y humano. La

perforina es la única proteína disponible que comparte una amplia región de su estructura primaria con C6-C9, por lo que se ha concluido que tienen un ancestro común. Esta homología descubierta nivel de secuencia está apoyada por características inmunológicas, fisicoquímicas y funcionales compartidas por ambos grupos de secuencias.

### 3. Reloj molecular

Como fundamento de un proceso de inferencia filogenética está el concepto de que las diferencias evolutivas entre organismos se deben a diferencias ocasionadas por mutaciones. Por lo tanto entre mas diferencias debidas a mutaciones haya entre dos organismos, la distancia evolutiva entre ambos será mayor. Algunos evolucionistas han empleado esta noción para construir escalas temporales evolutivas, es decir, han supuesto que las mutaciones se incorporan en los genes (son fijadas), en tasas regulares durante el tiempo y por lo tanto los cambios mutacionales entre dos secuencias correlacionan con la longitud de tiempo en que éstos se dieron. Esto significa que existe un reloj evolutivo a nivel molecular que determina las tasas en que las mutaciones serán fijadas (Strickberger, 1996b). La hipótesis del reloj molecular dice que para cualquier secuencia de DNA o proteína, la tasa de cambio o evolución es aproximadamente constante a través del tiempo y en todo linaje. Si las moléculas de DNA y proteínas evolucionan a tasas constantes pueden emplearse para ubicar en el tiempo a los eventos de divergencia de especies, duplicaciones de genes y otros hechos evolutivos, de manera similar a como se obtienen las fechas de los tiempos geológicos mediante el uso de elementos radioactivos. Adicionalmente, un proceso de reconstrucción filogenética resulta más sencillo si las tasas de cambio son constantes, pues en este caso no resulta necesario emplear métodos que consideren las sustituciones múltiples. Por otra parte, las variaciones importantes en la tasa de cambio de distintos linajes pueden evidenciar cambios substanciales en las condiciones evolutivas. Si hay un incremento notable en la tasa de evolución de una proteína en un linaje particular esto podría indicar evolución adaptativa, pérdida de función, duplicaciones génicas, etc (Li, 1997e).

Para calcular las tasas de cambio con respecto al tiempo se necesita información proveniente del registro fósil o de evidencia biogeográfica. Desafortunadamente estos datos nos son siempre confiables, tienen un rango de variación muy grande, no existen para la mayoría de los organismos y a veces son difíciles de interpretar. Para evitar estos problemas existe la Prueba

de la Tasa Relativa (RRT) que permite evaluar tasas de cambio sin tener información sobre los tiempos absolutos de divergencia. Esta prueba requiere de por lo menos dos secuencias A y B y de una secuencia C de referencia.

La RRT empleada en este proyecto es la prueba de los dos grupos de Takezaki, Rzhetsky y Nei (Takezaki, 1995) implementada en el programa Phyltest. A diferencia de otras pruebas en que puede incluir múltiples secuencias en cada uno de los linajes A, B y C. En esta prueba  $b_A$  y  $b_B$  son los promedios del número de sustituciones por sitio de A y B a su ancestro común N respectivamente. Si la tasa de cambio de A y B son constantes, es decir existe un reloj molecular, se espera que la diferencia ( $\delta$ ) entre  $b_A$  y  $b_B$  sea igual a cero ( $\delta = b_A - b_B = 0$ ), ésta es la hipótesis nula a evaluar. Sean  $L_{AB}$ ,  $L_{AC}$  y  $L_{BC}$  las distancias promedio entre los grupos de secuencias o las secuencias A y B, A y C, y B y C respectivamente. Esto es:

$$L_{AB} = \sum_{i \in A; j \in B} \frac{d_{ij}}{n_{ANB}}, \quad L_{AC} = \sum_{i \in A; j \in C} \frac{d_{ij}}{n_{ANC}}, \quad L_{BC} = \sum_{i \in B; j \in C} \frac{d_{ij}}{n_{BNC}},$$

Donde  $d_{ij}$  es la distancia entre las secuencias i y j que en este programa se obtienen por corrección de Poisson en el modo de delección total, es decir, en la alineación se eliminan todas las posiciones donde por lo menos haya un espacio. Las variables  $n_A$ ,  $n_B$  y  $n_C$  son el número de secuencias que pertenecen a los grupos A, B y C respectivamente, en nuestro caso en A y B se encontraban pares de las secuencias C6-C9 y en C siempre se ubicaron las secuencias de perforinas de rata, ratón y humano. Las variables  $b_A$  y  $b_B$  se estiman con estas ecuaciones:

$$b_A = \frac{L_{AB} + L_{AC} - L_{BC}}{2}$$

$$b_B = \frac{L_{AB} + L_{BC} - L_{AC}}{2}$$

La diferencia  $\delta$  se calcula :

$$\delta = b_A - b_B = L_{AC} - L_{BC}$$

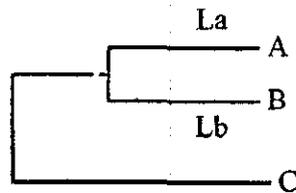
Se puede probar la desviación de  $d$  de 0 mediante un two-tailed normal deviate test con la Z estadística dada por:

$$Z = \frac{|\delta|}{\sqrt{V(\delta)}}$$

Donde  $V(\delta)$  es la varianza de  $d$  que se calcula con información de las varianzas y covarianzas de las distancias, esta información es calculada por el programa. Si  $Z$  es menor a 1.96, la hipótesis de tasas constantes de evolución es aceptada a un nivel del 5%. En la sección de resultados de este proyecto se graficaron los valores de  $Z$  en vez de los de  $\delta$  por que además de dar información sobre la constancia en las tasas de cambio también manifiesta si la diferencia  $d$  es significativamente distinta de cero.

Este es un ejemplo de la manera en que Phyltest presenta los resultados del RRT aplicado a una alineación de la región pobre en cisteínas.

- A: c9fug (1)
- B: c9tru (1)
- C: perfmus+perfhum+perfrat (3)



$L_a = 0.1666$

$L_b = 0.1848$

$L_a - L_b = -0.0182511 \hat{=} 0.0788395$  ( $Z = 0.231496$  rate constancy is NOT rejected at 5% level).

Anexo 4

CLAVE

Negrita = Posiciones idénticas  
 Subrayado = Aminácidos del mismo grupo

□ = Posición informativa para Parsimonia

↓ = Punto donde se eliminaron sitios hipervariables.

— = Módulos TSP

■ = Módulo LDlr

- - - = Módulo CP

■ = Módulo EGF

C9FUGU	1111111112	2222222223	3333333334	4444444445	5555555556
C9TROUT	CV... <u>WSRWA</u>	<u>FWSSCDPCTN</u>	<u>TRRRSRGVEV</u>	<u>EGQFAGIACQ</u>	GSV..GDREY
C9RAT	CV... <u>WRWS</u>	<u>EWTFPCNSCTK</u>	<u>IRHRSRSVEV</u>	<u>EGQFGKPCQ</u>	G.QPIGEQOR
C9MUS	CRM... <u>STWS</u>	<u>QWSQCDPCLK</u>	<u>QRFRSRSMEV</u>	<u>EGQFQKSCA</u>	..DALGDRQH
C9RABB	CRM... <u>SPWS</u>	<u>NWSECDPCLK</u>	<u>QRFRSRSILA</u>	<u>EGQFNGKSCV</u>	..DVLGDRQG
C9HORS	CRM... <u>SPWS</u>	<u>EWSHCDPCLR</u>	<u>QMFERSRSIEV</u>	<u>EGQFHGKSCV</u>	..DALGDRRA
C9HUMAN	CRM... <u>SPWS</u>	<u>EWSECDPCLR</u>	<u>QMFERSRSIEV</u>	<u>EGQFNGRCV</u>	..DAVGDRRQ
C8ARABB	COL... <u>SPWS</u>	<u>EWSDCDPCLR</u>	<u>QMFERSRSIEV</u>	<u>EGQFNGKRCV</u>	..DAVGDRRQ
C8AHUMAN	COL... <u>SNWS</u>	<u>EWTDCCFPQD</u>	<u>TKYRHRSLLO</u>	<u>PNKFGGTCIS</u>	G.DIW.DRAS
C8BRABB	CEL... <u>SNWS</u>	<u>EWTDCCFPQD</u>	<u>KKYRHRSLLO</u>	<u>PNKFGGTCIS</u>	G.DIW.DQAS
C8BHUMAN	CEL... <u>SNWS</u>	<u>SWTMCDFPCQK</u>	<u>KRYRWAYLLR</u>	<u>PSQFHGEPEN</u>	FSDKEVED..
C6HUMAN	CLLGDGFGWS	D... <u>CDPCIE</u>	<u>KQSKVRSVLR</u>	<u>PSQFGGQPCV</u>	EPLVAFQP..
C7HUMAN	CWDFEYAPWS	E... <u>CNGCTK</u>	<u>TQTRRRSVAV</u>	<u>YGQYGGQPCV</u>	NAF.ETQS..
PERFUS1	.....	.....	.....	.....	.....
PERFORAT	.....	.....	.....	.....	.....
PERFOHUM	.....	.....	.....	.....	.....

C9FUGU	6666666667	7777777778	8888888889	9999999990	1111111111
C9TROUT	PPRECSDEF	QC.ESGCIK	LRLKNGDYD	CEGDS.DE.D	.CE..PLRKT
C9RAT	LPSECSIEF	TC.ESGACIK	LRLSCNGDYD	CEGDS.DE.D	.CE..PVRKP
C9MUS	QEN.C.GNDF	QC.ETGRCIK	RLLCNGDND	CGDFS.DESD	.CESDP.RLP
C9RABB	EE.DCEKDEF	HCG.TGRCIK	RLLCNGDND	CGDYS.DEND	.CDDDP.RTP
C9HORS	ED.DC.GNDF	QCG.TGRCIK	KRLCNGDND	CGDFS.DEDD	.CETEP.RLT
C9HUMAN	ED.DC.GNDF	QCS.TGRCIK	MRLCNGDND	CGDFS.DEDD	.CENDP.RPP
C8ARABB	AQ..C.GQDF	QCKETGRCIK	RHLVCGEND	CLDGS.DEDN	.CEDIRATES
C8AHUMAN	AQ..C.GQDF	QCKETGRCIK	RHLVCGDOD	CLDGS.DEDD	.CEDVRAIDE
C8BRABB	VR..CEG..F	VCAQTGRGVN	RRLCNGDND	CGDQS.DEAN	.CRKIYKK..
C8BHUMAN	VR..CEG..F	VCAQTGRGVN	RRLCNGDND	CGDQS.DEAN	.CRRIYKK..
C6HUMAN	EEADC.KNKF	RCD.SGRCIA	RLECGEND	CGDNS.DERD	.CGRTKAV..
C7HUMAN	EE.GC.GERF	RCF.SGQCIS	KSLVCGDSD	CDEDSADE.D	RCEDESE.RRP
PERFUS1	.....	.....	.....	.....	.....
PERFORAT	.....	.....	.....	.....	.....
PERFOHUM	.....	.....	.....	.....	.....

C9FUGU EQGRTA. .GY GINIL.GADE RMNPFND.F FNG.RCDKVR NPNTLQLDRL PWNIGVLNYQ  
 C9TROUT EQGRTA. .GY GINIL.GMEP RINPFND.Y FNG.MCNKVK NINNEYNRL PWNVGLNVE  
 C9RAT ELGRTA. .GY GINIL.GMDP LGTFPDNE.F YNG.LCDVR DGNTLYYRK PWNVAFLAYE  
 C9MUS ELGLRA. .GY GINIL.GMEP LRTFPDNE.F YNG.LCDVR DGNTLYYRK PWNVVS LIYE  
 C9RABB ELARTA. .GY GINIL.GMDP LATFPDNE.Y YHG.LCDVR DGNTLTHYRK PWNVAVLAYE  
 C9HORS ELARTA. .GY GINIL.GMDP LSTFPDNE.Y YNG.LCDVR DGNTLYYRK PWNLASLAYE  
 C9HUMAN ELARTA. .GY GINIL.GMDP LSTFPDNE.F YNG.LCNRDR DGNTLYYRR PWNVASLIYE  
 C8ARABB PGSEKAAALGY .NLTQEEA QS. .VYDARY Y.GGRCETVY NGEWRHRYD PVCER.LHHG  
 C8AHUMAN PGSQKAAALGY .NLTQEDA QS. .VYDASY Y.GGOCETVY NGEWRELRYD STCER.LYYG  
 C8BRABB AIGSLAS. .GIN.LFTNSL EG.PVLDRY YAGG.CNPHY ILDMR. ....  
 C8BHUMAN GIGSLAS. .GIN.LFTNSF EG.PVLDRY YAGG.CSPHY ILNTR. ....  
 C6HUMAN PSVQLM. .GN GEHFLAGE. P RGEVLDNS.F T.GGICKTVK SSRTSNPYRV PAN. .LENV  
 C7HUMAN PNIELT. .GN GYNELTGQFV NR.VINTKSF . .GGQCRKVF SGDGKDFYRL SGN. .VLSYT  
 PERFMUS1  
 PERFORAT  
 PERFOHUM

C9FUGU LTNI.KNKSF MRVKGRLQLS TYMRSHQ. L QVADEFVAHV KSLPLEYEKG I.YYAFLE.D  
 C9TROUT YTTI.KNKSE MRVNGRVQLS TYMRSRD. L QVAGEFLEHV KSLPLEYEKG Q.YFSFLE.D  
 C9RAT YLSQTK.KMF LHVRCMIQLG RFVMRNRG. V MLTTFELDDV KALPVSYEKG E.YFGFLE.T  
 C9MUS YLSQSK.KMF VHLRGVVQLG RFVMRNRD. V VLRSTELDDV KALPTSYEKG E.YFGFLE.T  
 C9RABB YSSQ.KEKMF LHVKGVIQLG RFVMRNRD. V MLTNTFLODI KSLPTTYEKG E.YFAFLE.T  
 C9HORS YSSK.KEKMF LHVKGVIQLG RFVMRNRD. V VLTTFVDDI KALPTAYEKG E.YFAFLE.T  
 C8AHUMAN YNEK.KYN.F MRIFTKVQTA HFKMRDD. I VLDEGMLOAL VELPEQNYG M.YSKFIN.D  
 C8ARABB YNEK.KFI.F TRIFTKVQTA HFKMRDD. I MLDEGMLOSL MELPDQNYG M.YAKFIN.D  
 C8BRABB FSHT.KSK.F LHARSALEVA HYKL.KPRSL MLHYDELQRV QRVPLEYSYG E.YRD.LFRD  
 C8BHUMAN FSHT.KSV.F LHARSDLEVA HYKL.KPRSL MLHYEFLORV KRLPLEYSYG E.YRD.LFRD  
 C6HUMAN ASHK.KDSSE IRIHKVMKVL NETT.KAKDL HLSDFELKAL NHLPLEYN. S ALYSR.IFDD  
 C7HUMAN . .HKGKSYQL LVVENTVEVA QFINNPEFL QLAEPFWKEL SHLPSLYDYS A.YRR.LIDQ  
 PERFMUS1  
 PERFORAT  
 PERFOHUM

C9FUGU	YGTHYTKNGK	SGGEYE.LVY	.VLNQDTIK	AKN.LTE.RK	IQECUKIGIV	VDNVMTSVKG	VDKVITVVRG
C9TROUT	YGTHYTRNGK	SGGEHQ.LVY	.VLNQDKAMK	EK.GVE.LSD	VKRCL..GFI	IDDVLSFIRG	IDDVLSFIRG
C9RAF	YGTHYSSSGS	LGGLYE.LIY	.VLDKASMK	EK.GVD.LND	VKHCL..GFI	IDDVLSFIRG	IDDVLSFIRG
C9MUS	YGTHYSSSGS	LGQOYE.IVY	.VLDKASMK	EK.GIE.LND	IKKCL..GFL	IDDVLSLIRG	IDDVLSLIRG
C9RABB	YGTHYSSSGS	IGGRYE.LIY	.VLDKASMK	OK.GVE.LRD	IQRCL..GFL	IDDVLSLIRG	IDDVLSLIRG
C9HORS	YGTHYSSSGS	LGGLYE.LIY	.VLDKASMK	RK.GVE.LKD	IKRCL..GYL	IDDVLSLIRG	IDDVLSLIRG
C9HUMAN	YGTHYSSSGS	IGGLYE.LIY	.VLDKASMK	RK.GVE.LKD	IKRCL..GYL	IDDVLSLIRG	IDDVLSLIRG
C8ARABB	YGTHYITSGS	MGGTYE...Y	ILVNLTEKME	SL.GVTS.ED	ISSCF.GGF.	.KDIISRVRG	.DIISRVRG
C8AHUMAN	YGTHYITSGS	MGGIYE...Y	ILVIDKAKME	SL.GITS.RD	ITTCF.GGS.	.DIISRVRG	.DIISRVRG
C8BRABB	FGHFITEAV	LGGIYE...Y	TLIMNKEAME	R..ADYSLND	VQACAKNDF.	.DLVVLVVRG	.DLVVLVVRG
C8BHUMAN	FGTHYITEAV	LGGIYE...Y	TLVMNKEAME	R..GDYTLNN	VHACAKNDF.	.DLVVLVVRG	.DLVVLVVRG
C6HUMAN	FGTHYFTSGS	IGGVYD.LLY	.QFSSEEL.K	NS.GLTEE.E	AKHCVRIET.	.....LIRG	.....LIRG
C7HUMAN	YGTHYLQSGS	IGGEYRVLFY	.VDSEKL.K	QONDF..NSVE	EKKCKSSGW.	.....IRG	.....IRG
PERFUMUS1	YGTHFITAVD	IGG..RISVL	TALRTCQLTL	N..GLTAD.E	VGDCINVEA.	.....LG	.....LG
PERFORAT	YGTHFITAVD	IGG..RVSVL	TALRTCQLTL	D..GLTAD.E	VGDCLSVEA.	.....LG	.....LG
PERFOHUM	YGTHFIRAVE	IGG..RISAL	TALRTCELAL	E..GLTDN.E	VEDCLTVEA.	.....VG	.....VG

C9FUGU	GSLEAVTMW	ARTIASA.PA	LINSEPEPIY	M.LIPTDIPG	A..NSRIANL	KQATADYVAE	KQATADYVAE
C9TROUT	GTLEAAVAMW	ARTYGDA.PA	LISSEPEPIQ	T.LIPLSMPD	A..NTRRLNM	QRATQEYEA	QRATQEYEA
C9RAT	GTRKQAVLW	ASSLDDA.PA	LISQKLSPIY	N.LIPLTMKD	A..YAKKQNM	EKAIEDYVNE	EKAIEDYVNE
C9MUS	GTRKQAVLW	ASSLANA.PA	LISQRMSPY	N.LIPLKIKD	A..YIKKQNL	EKAVEDYIDE	EKAVEDYIDE
C9RABB	GTQKFAFELW	ASSLSDA.PV	LINQKLSPIY	N.LIPVKIKD	A..HQKRQNL	ERGIEDYINE	ERGIEDYINE
C9HORS	GTRKYAFELW	ASSLND.A.PV	LISQRLSPIY	N.LIPVKMKD	A..HQKRQNL	ERAIEDYINE	ERAIEDYINE
C8AHUMAN	GSSGWGGGLW	GRSLKYN.PV	LISQKLSPIY	N.LVPVKMKN	A..HLKKQNL	ERAIEDYINE	ERAIEDYINE
C8RABB	GSSGWGGGLW	GRSLKYN.PV	VIDFEMQPIH	EVLJHTNLGH	V..EAKRQNL	RRAIDQYLME	RRAIDQYLME
C8AHUMAN	GSSGWGGGLW	GRSLKYN.PV	VIDFEMQPIH	EVLRHTSLGP	L..EAKRQNL	RRAIDQYLME	RRAIDQYLME
C8BRABB	GASEHITAJW	GDAVQYN.PA	LIKIKVEPLY	E.LVTATDV.	AYSSTVKQNM	RQALEEFQGE	RQALEEFQGE
C8BHUMAN	GASEHITTLW	GDAVQYN.PA	LIKVKVEPLY	E.LVTATDF.	AYSSTVRQNM	KQALEEFQKE	KQALEEFQKE
C6HUMAN	GRSEYGAALW	LESVKEN.PA	VIDFELAPIV	D.IVRNIPC.	AV..TKRNNL	RKALEQYAAK	RKALEQYAAK
C7HUMAN	GGAGFISGLW	AESVT.NLPQ	VIKQKLTPLY	E.LVKEVPC.	AS..VKKLYL	KWALEEYL..	KWALEEYL..
PERFUMUS1	GPLDSTHDLW	TASLPSN.PG	LVDYSLEPLH	T.LLEEONP.	.....KREAL	QQAISHYIMS	QQAISHYIMS
PERFORAT	GPLDSSNDLW	IASLPTR.PD	VVDYSLEPLH	I.LLESDP.	.....KREAL	QQAISHYVMS	QQAISHYVMS
PERFOHUM	GHHTSINDLW	VNSLPGS.PG	LVDYTLLEPLH	V.LLDSQDP.	.....RREAL	RRAALSQYLTID	RRAALSQYLTID

C9FUGU	YNV. . CKCR	PCHNGGTIAL	LDGKICMCS	NLFEGLG.CQ	WSCWSSWNC	Q.GOKRSRTR
C9TROUT	YSV. . CKCK	PCHNGGSLAL	LDGKCLCLCL	POFEGLA.CQ	WSCWSSWNC	Q.GOKRIRTR
C9RAT	FSARK. . . CY	PCQNGGTAIL	LDGQCMCSCT	IKFKGIA.CE		
C9MUS	FSTKR. . . CY	PCLNGGTIIL	LDGQCLCSCP	MMFRGMA.CE		
C9RABB	FSTKK. . . CS	PCQNGGTALL	MDGQCLCTCP	EMFEGIA.CE		
C9HORS	FSVRK. . . CH	PCQNGGTVIQ	IDGQCLCSOP	IAFEGIA.CE		
C9HUMAN	FSVRK. . . CH	TCQNGGTVIL	MDGKCLCACP	EKFEKIA.CE		
C8ARABB	F. . . NACRCG	PCFNNGKPIL	EGTSCRCQCS	IGLQGPA.CE	WSCWGSWSPC	TAGTRRRRE
C8AHUMAN	F. . . NACRCG	PCFNNGVPIL	EGTSCRCQCR	IGSILGAA.CE	WSCWSSWVC	RAGIQERRRE
C8BRABB	VSP. . . CRCA	PCQNGVPVQ	KGSRDCICP	VGFQGSA.CE	WSCWSSWSSC	SGGQKTRRRQ
C8BHUMAN	VSS. . . CHCA	PCQNGVPVVL	KGSRDCICP	VGSQGLA.CE	WSCWSSWSSC	SGRRKTRRRQ
C6HUMAN	FDP. . . CQCA	PCPNNGRPTL	SGTECLCVQ	SGTYGEN.CE	WSCWSSWSTC	DATYKRSRTR
C7HUMAN	DP. . . CHCR	PCQNGGLATV	EGTHCLCHK	PYTFGAA.CE	WSCWSSWSPC	VQKKTRSR.
PERFMUS1	RARWQNC.S.R	PCRS.GQHK	SHDSCQCEQ	DSKVTNQDC.		
PERFORAT	RARWRDCN.R	PCRA.GQHK	SRDSCQVCQ	DSNVTNQDC.		
PERFOHUM	RARWRDCS.R	PCPP.GROKS	PRDPCQCVCH	GSAVTTQDC.		

C9FUGU	YCNTQGLS..	GAEGRGEIRS	EEYC
C9TROUT	SCNTQGLS..	DATCRGDIVT	EDYC
C9RAT			
C9MUS			
C9RABB			
C9HORS			
C9HUMAN			
C8ARABB	CNNPAPQNG	GAPCPGWRVQ	TQAC
C8AHUMAN	CDNPAPQNG	GASCPGRKVQ	TQAC
C8BRABB	CNNPAPQDG	GSPCSGPASE	TLAC
C8BHUMAN	CNNPPPPQNG	GSPCSGPASE	TLDC
C6HUMAN	ECNNPAPQNG	GKRCEGEKRO	EEDC
C7HUMAN	ECNNPPPPSGG	GRSCVGETTE	STQC
PERFMUS1			
PERFORAT			
PERFOHUM			

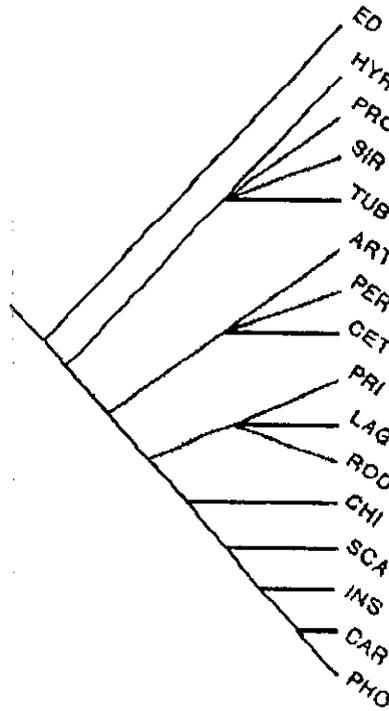
## Anexo 5

### Filogenia de mamíferos.

A nivel molecular, el análisis de las relaciones filogenéticas entre mamíferos que se considera más completo (Benton, 1988) genera el siguiente cladograma:

Clave:

ARTiodáctila  
CARNívora  
CETácea  
CHIróptera  
EDentata  
HYRacoidea  
INSectívora  
LAGomorfa  
PERisodáctila  
PHOlidota  
PRImates  
SCAndentians  
SIRenia  
TUBilodontata



En este esquema se puede observar que la rama que dio origen a los perisodáctilos, entre los que se encuentra el caballo, divergió antes que la que agrupa a roedores, lagomorfos y primates. El conflicto de esta filogenia con las obtenidas a partir de C6-C9 y perforina es que en nuestros árboles los C9 de rata, ratón y conejo divergieron antes que el C9 de caballo, que aparece como el más cercano a C9 humano.

## VIII. BIBLIOGRAFÍA.

1. Atkinson JP y Farries T (1987) Separation of self from non-self in the complement system. *Immunol. Today* 8: 212-215
2. Avise JC (1994) Interpretive Tools. Cap. 4. En Avise JC. *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York. p. 92.
3. Ayala FJ (1976) Preface. En Ayala FJ (Ed.) *Molecular Evolution*. Sinauer Associates Inc. Sunderland Massachusetts. p. vii.
4. Bartl S. *et al.* (1994) Molecular evolution of the vertebrate immune system. *Proc. Natl. Acad. Sci. USA* 91:10769-10770.
5. Becherer JD *et al.* (1989) Molecular aspects of C3 Interactions and structural/functional analysis of C3 from different species. *Curr. Top. Microbiol. Immunol.* 153: 45-65.
6. Benton MJ (1988) The relationships of the major group of mammals: New approaches. *Trends. Ecol. Evol.* 3:40-45.
7. Biesecker G *et al.* (1982) An amphiphilic structure of the ninth component of human Complement. *J. Biol. Chem.* 257:2584-2590.
8. Brickner A y Sodetz JM (1985) Functional Domains of the a subunit of the eighth component of human complement: Identification and characterization of a distinct binding site for the g chain. *Biochem.* 24:4603-4607.
9. Brown JK (1994) Bootstrap hypothesis tests for evolutionary trees and other dendrograms. *Proc. Natl. Acad. Sci. USA.* 91:12293-12297.
10. Campbell AK y Morgan BP (1985) Monoclonal antibodies demonstrate protection of polymorphonuclear leucocytes against complement attack. *Nature* 317:164-166.
11. Carney DF *et al.* (1986) Elimination of terminal complement complexes in the plasma membrane of nucleate cells: influence of extracellular  $Ca^{2+}$  and association with cellular  $Ca^{2+}$ . 137:263-270. *J. Immunol.* 145:623-629.
12. Carney DF *et al.* (1990) Multiple signal messengers generated by terminal complement complexes and their role in terminal complement complex elimination. *J. Immunol.* 145:623-629.
13. Cooper NR y Nemerow GR (1986) Complement-dependent mechanisms of virus neutralization. Cap. 6. En Ross GD (Ed.) *Immunobiology of the complement system*.

Academic Press Inc., Orlando, p. 139.

14. Cooper N. (1991) Complement evasion strategies of microorganisms. *Immunol. Today* 12:327-331
15. Chakravarti DN *et al.* (1989) Structural homology of complement protein C6 with other channel-forming proteins of complement. *Proc. Natl. Acad. Sci. USA* 86:2799.
16. Daniels RH *et al.* (1990a) Human rheumatoid synovial cell stimulation by the membrane attack complex and other pore-forming toxins *in vitro*: the role of calcium in cell activation. *Immunology* 71:312-316.
17. Daniels RH *et al.* (1990b) Stimulation of human rheumatoid synovial cells by non-lethal complement membrane attack. *Immunol.* 69:237-242.
18. Davies A y Lachmann PJ (1993) Membrane defence against complement lysis: The Structure and Biological Properties of CD59. *Immunol. Res.* 12: 258-275.
19. Day NKB *et al.* (1970) Complement and complement-like activity in lower vertebrates and invertebrates. *J. Exp. Med.* 132:941-950.
20. DeBry RW (1992) The consistency of several phylogeny-inference methods under varying evolutionary rates. *Mol. Biol. Evol.* 9:537-551.
21. DiScipio RG *et al.* (1984) Nucleotide sequence of cDNA and derived amino acid sequence of human complement component C9. *Proc. Natl. Acad. Sci. USA.* 81:7298-7302.
22. DiScipio RG *et al.* (1988) The structure of human complement component C7 and the C5b-7 complex. *J. Biol. Chem.* 263:549-560.
23. DiScipio RG y Hugli TE (1989) The molecular architecture of human Complement component C6. *J. Biol. Chem.* 264:16197-16206.
24. Dolittle R (1994) Convergent evolution: The need to be explicit. *TIBS* 19:15-18.
25. Efron B *et al.* (1996) Bootstrap confidence levels for phylogenetic trees. *Proc. Natl. Acad. Sci. USA.* 93:13429-13434.
26. Epstein J *et al.* (1996) The collectins in innate immunity. *Curr. Opin. Immunol.* 8:29-35.
27. Farries TC *et al.* (1990) Evolutionary implications of a new bypass activation pathway of the complement system. *Immunol. Today* 11:78-80.
28. Farries TC y Atkinson JP (1991) Evolution of the complement system. *Immunol. Today*

12:295-300.

29. Frank MM y Fries LF. (1991) The role of complement in inflammation and phagocytosis. *Immunology Today* 12:322-326
30. Fuhrman SA y Joiner KA (1989) Complement evasion by protozoa. *Exp. Parasitol.* 68:474-481.
31. Gallin JI (1993) Inflammation. Cap 29. En Paul WE (Ed.) *Fundamental Immunology*. Tercera edición. Raven Press Ltd., New York, p. 1015.
32. González S *et al.* (1996) Genetic structure and organization of the membrane attack complement components. *Eur. J. Immunogen.* 23:181-197.
33. Gould SJ (1989) The ichonography of an expectation. Cap.1. En Gould SJ. *Wonderful Life: the Burgess Shale and the nature of history*. WW Norton & Company, New York. p.23.
34. Goundis D y Reid KBM (1988) Properdin, the terminal complement components, thrombospondin and the circumsporozoite protein of malaria parasites contain similar sequence motifs. *Nature* 335:82-85.
35. Graur D *et al.* (1997) Evolutionary affinities of the order Perissodactyla and the phylogenetic status of the superordinal taxa Ungulata and Altungulata. *Mol. Phylogenet. Evol.* 7:195-200
36. Haefliger JA *et al.* (1987) Complementary DNA cloning of complement C8b and its sequence homology to C9. *Biochem.* 26:3551-3556.
37. Haefliger JA *et al.* (1989) Complete primary structure and functional characterization of the sixth component of the human Complement system. *J. Biol. Chem.* 264:18041-18051.
38. Halperin JA (1988) Complement induces a transient increase in membrane permeability in unlysed erythrocytes. *J. Clin. Invest.* 82: 594-600.
39. Halperin JA *et al.* (1993) Transient changes in erythrocyte membrane permeability are induced by sublytic amounts of the complement membrane attack complex (C5b-9) *Blood* 81:200-205.
40. Higgins DG y Sharp PM (1988) CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. *Gene* 73:237-244.

41. Hillis DM y Huelsenbeck JP (1992) Signal, noise and reliability in molecular phylogenetic analyses. *J Hered* 83:189-195.
42. Hillis DM y Bull JJ (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst. Biol.* 42:182-192.
43. Hobart *et al.* (1993) Structure of the human C6 gene. *Biochem.* 32:6198-6205.
44. Hobart *et al.* (1995) Structure of the human C7 gene and comparison with the C6, C8A, C8B and C9 genes. *J. Immunol* 154:5188-5194.
45. Holmskov U *et al.* (1994) Collectins: collagenous C-type lectins of the innate immune defense system. *Immunol. Today.* 15:67-74.
46. Howard *et al.* (1987) Complementary DNA and derived amino acid sequence of the b subunit of human complement protein C8: Identification of a close structural and ancestral relationship to the a subunit and C9. *Biochem.* 26:3565-3570.
47. Hu VW *et al.* (1981) *J. Immunol.* The membrane attack mechanism of complement: Photolabeling reveals insertion of terminal proteins into target membrane. 127:380-386
48. Hughes AL (1994) Phylogeny of the C3/C4/C5 Complement-Component gene family indicates that C5 diverged first. *Mol. Biol. Evol.* 11:417-425.
49. Hughes-Jones NC (1986) The classical pathway. Cap. 1 En Ross GD (Ed.) *Immunobiology of the complement system.* Academic Press Inc., Orlando, p. 21.
50. Iida K *et al.* (1991) Membrane vesiculation protects erythrocytes from destruction by complement. *J. Immunol* 147:2638-2642.
51. Ikeda K *et al.* (1987) Serum lectin with known structure activates Complement through the classical pathway. *J. Biol. Chem.* 262 7451-7454.
52. Ishida B *et al.* (1982) Photolabeling of a hydrophobic domain of the ninth component of human complement. *J. Biol. Chem.* 257: 10551-10553.
53. Ishikawa H. *et al.* (1989) Molecular cloning of rat cytolysin. *J. Immunol.* 143: 3069-3073.
54. Jensen JA *et al.* (1981) The complement system of the nurse shark: hemolytic and comparative characteristics. *Science* 214:566-569.
55. Ji X. *et al.* (1997) Ancient origin of the complement lectin pathway revealed by molecular cloning of mannan binding protein-associated serine protease from a

- urochordate, the Japanese ascidian, *Halocynthia roretzi*. Proc.Natl.Acad.Sci. USA 94: 6340-6345.
- 56.Joiner K. (1988) Complement evasion. Ann. Rev. Microbiol. 42: 201-30.
- 57.Kaufman KM y Sodetz JM. (1994) Genomic structure of the human complement protein C8 $\gamma$ : Homology to the lipocalin gene family.
- 58.Kovacsovics TJ *et al.* (1987) Antibody-independent activation of C1. J. Immunol. 138:1864-1870.
- 59.Kuma K y Miyata T (1994) Mammalian phylogeny inferred from multiple protein data. Jpn. J. Genet. 69:555-566.
- 60.Kumar S. *et al.* (1993a) MEGA: Molecular Evolutionary Genetics Analysis, versión 1.0. The Pennsylvania State University, University Park, PA 16802.
- 61.Kumar S *et al.* (1993b) Phylogenetic inference. Cap. 5. En Kumar *et al.* MEGA Molecular Evolutionary Genetics Analysis, version 1.02. Manual del usuario. The Pennsylvania State University. p31.
- 62.Kumar S *et al.* (1993c) Distance estimation. Cap. 4. En Kumar *et al.* MEGA Molecular Evolutionary Genetics Analysis, version 1.02. Manual del usuario. The Pennsylvania State University. p15.
- 63.Kumar S. (1996) PHYLTEST: a program for testing phylogenetic hypothesis. versión 2.0. The Pennsylvania State University, University Park Pennsylvania, USA.
- 64.Kwon BS *et al.* (1989) The structure of the mouse lymphocyte pore-forming protein perforin. Biochem. Biophys. Res. Commun. 158:1-10.
- 65.Lachmann PJ (1991) The control of homologous lysis. Immunol. Today 12: 312-315.
- 66.Lawler J y Hynes RO (1986) The structure of human thrombospondin, an adhesive glycoprotein with multiple Calcium-binding sites and homologies with several different proteins. J. Cell. Biol. 103:1635-1648.
- 67.Li W-H (1997a) Rates and patterns of nucleotide substitution. Cap. 7. En Li W-H. Molecular Evolution. Sinauer Associates. Sunderland Massachusetts. p.177.
- 68.Li W-H (1997b) Molecular phylogenetics. Cap. 5. En Li W-H. Molecular Evolution. Sinauer Associates. Sunderland Massachusetts. p.99.
- 69.Li W-H (1997c). Molecular evolution: A brief history of the Pre-DNA Era. Introduction.

- En Li W-H. Molecular Evolution. Sinauer Associates. Sunderland Massachusetts. p.1.
70. Li W-H (1997d). Estimating the Number of nucleotide substitutions. Cap. 4. En Li W-H. Molecular Evolution. Sinauer Associates. Sunderland Massachusetts. p.79.
71. Li W-H (1997e). Molecular clocks. Cap. 8. En Li W-H. Molecular Evolution. Sinauer Associates. Sunderland Massachusetts. p.215.
72. Lichtenheld MG *et al.* (1988) Structure and function of human perforin. *Nature* 335:448-451.
73. Linthicum DS (1970) Immunologic response of Pacific hagfish. 3 Serum antibodies to cellular antigens. *J. Immunol.* 105:912-918.
74. Liszewski MK. y Atkinson JP (1993) The Complement System. Cap. 26. En Paul WE (Ed.) *Fundamental Immunology*. Tercera edición. Raven Press Ltd., New York, p. 917.
75. Litman GW (1997) Origen de la inmunidad de los vertebrados. *Inv. y Cien.* 244:10-15.
76. Liu CC *et al.* (1995) Perforin: structure and function. *Immunol. Today.* 16:194-201.
77. Lowrey DM *et al.* (1989) Cloning, analysis, and expression of murine perforin 1 cDNA, a component of cytolytic T-cell granules with homology to complement component C9. *Proc. Natl. Acad. Sci. USA.* 86:247-251.
78. Lynnon BioSoft (1994-97) DNAMAN for Windows, version 2.6. Montreal.
79. Malhotra R *et al.* (1990) Human leucocyte C1q receptor binds other soluble proteins with collagen domains *J. Exp. Med.* 172: 955-959
80. Malhotra R *et al.* (1994) Collectins, collectin receptors and the lectin pathway of complement activation. *Clin. Exp. Immunol.* 97: Suppl 2, 4-9.
81. Matsushita M y Fujita T (1992) Activation of the classical Complement Pathway by Monnose-binding Protein in association with a novel C1s-like serine protease. *J. Exp. Med.* 176: 1497-1502.
82. Matsushita M (1996) The lectin pathway of the Complement System. *Microbiol. Immunol.* 40:887-893
83. Meri S. *et al.* (1990) Human protectin (CD59), and 18,000-20,000 MW complement lysis restricting factor, inhibits C5b-8 catalysed insertion of C9 into lipid bilayers. *Immunol.* 71:1-9.

84. Michelotti *et al.* (1995) Genomic organization of human complement protein C8a and further examination of its linkage to C8b. *Hum. Genet.* 95:513-518.
85. Milis L *et al.* (1993) Vitronectin-mediated inhibition of complement: evidence for different binding sites for C5b-7 and C9. *Clin. Exp. Immunol.* 92:114-119.
86. Monahan JB y Sodetz JM (1980) Binding of the eighth component of human complement to the soluble cytolytic complex is mediated by its b subunit. *J. Biol. Chem.* 255:10579-10582.
87. Monahan JB y Sodetz JM (1981) Role of the b-subunit in the interaction of the eighth component of human Complement with the membrane-bound cytolytic complex. *J. Biol. Chem.* 256:3258-3262.
88. Morgan BP *et al.* (1987) Recovery of human neutrophils from complement attack: Removal of the membrane attack complex by endocytosis and exocytosis. *J. Immunol.* 138:246-253.
89. Morgan BP (1990a) The Complement System. Cap. 1. En Morgan BP. *Complement: Clinical aspects and relevance to disease.* Academic Press Inc., San Diego, p. 1.
90. Morgan BP (1990b) The biological effects of Complement activation. Cap. 2. En Morgan BP. *Complement: Clinical aspects and relevance to disease.* Academic Press Inc., San Diego, p. 36.
91. Morgan BP (1990c) The genetics of complement. Cap. 3. En Morgan BP. *Complement: Clinical aspects and relevance to disease.* Academic Press Inc., San Diego, p. 57.
92. Morgan BP (1990d) Complement and Infectious diseases. Cap. 4. En Morgan BP. *Complement: Clinical aspects and relevance to disease.* Academic Press Inc., San Diego, p. 78.
93. Morgan BP (1992) Effects of the membrane attack complex of complement on nucleated cells. *Curr. Top. Microbiol. Immunol.* 178:115-140.
94. Müller-Eberhard H. (1986) The membrane attack complex of complement. *Ann. Rev. Immunol.* 4:503-528.
95. Müller-Eberhard HJ (1988) Molecular organization and function of the complement system. *Annu. Rev. Biochem.* 57:321-347.
96. Nei M (1987) Phylogenetic trees. Cap. 11. En Nei M. *Molecular Evolutionary Genetics.*

Columbia University Press New York. p. 287.

97. Niculescu F *et al.* (1994) Receptor-independent activation of guanine nucleotide-binding regulatory proteins by terminal complement complexes. *J. Biol. Chem.* 269:4417-4423.
98. Nicholson-Weller A y Halperin JA (1993) Membrane signaling by complement C5b-9, the Membrane Attack Complex. *Immunol. Res.* 12:244-257.
99. Nonaka M y Takahashi M (1992) Complete complementary DNA sequence of the Third component of complement of lamprey. *J. Immunol.* 148:3290-3295.
100. Pangburn MK. (1986) The alternative pathway. Cap. 2. En Ross GD (Ed.): *Immunobiology of the complement system.* Academic Press Inc., Orlando, p. 45
101. Pathy L (1987) Intron-dependent evolution: preferred types of exons and introns. *FEBS Letters* 214: 1-7.
102. Pathy L (1988) Detecting distant homologies of mosaic proteins. *J. Mol. Biol.* 202:689-696.
103. Peitch MC *et al.* (1987) Antibody-independent activation of C1. *J. Immunol.* 138:1871-1876
104. Peitch MC *et al.* (1990) Localization and molecular modelling of the membrane-inserted domain of the ninth component of human complement and perforin. *Mol. Immunol.* 27: 589-602.
105. Perechini PM *et al.* (1992) Channel-forming activity of the perforin N-terminus and a putative  $\alpha$ -helical region homologous with complement C9. *Biochem.* 31:5017-502.
106. Pereira M (1994) Lectins. Cap. 11. En van Oss CJ y van Regenmortel MHV (eds) *Immunochemistry.* Marcel Decker Inc. New York. p. 259.
107. Perlmutter DH y Colten HR (1987) Structure and expression of the complement genes. *Pharmac. Ther.* 34:247-270
108. Podack ER *et al.* (1981) Membrane attack complex of complement: distribution of subunits between the hydrocarbon phase of target membrane and water. *Proc. Natl. Acad. Sci. USA* 78: 4544-4548.
109. Podack ER *et al.* (1982a) Molecular organization of C9 within the membrane attack complex of Complement. *J. Exp. Med.* 156: 268-282.

- 110.Podack ER y Tschopp J (1982b) Polymerization of the ninth component of complement (C9): Formation of poly C9 with a tubular ultrastructure resembling the membrane attack complex of complement. *Proc.Natl. Acad. of Sci. USA* 79:574-578.
- 111.Podack ER y Tschopp J (1982c) Circular polymerization of the ninth component of complement. Ring closure of the tubular complex confers resistance to detergent dissociation and to proteolytic degradation. *J. Biol. Chem.* 257: 15204-15212.
- 112.Podack ER y Tschopp J (1984a) Membrane attack by Complement. *Mol. Immunol.* 21:589-693.
- 113.Podack ER (1984b) Molecular composition of the tubular structure of the membrane attack complex of complement. *J. Biol. Chem.* 259:8641-8647.
- 114.Podack ER y Tschopp J (1985) Membrane attack and complement. *J. Mol. Immunol.* 21:589-603.
- 115.Podack ER (1986a) Assembly and functions of the terminal components. Cap. 5. En Ross GD (Ed.) *Immunobiology of the complement system.* Academic Press Inc., Orlando, p. 115.
- 116.Podack ER (1986b) Molecular mechanisms of cytolysis by complement and by cytolytic lymphocytes. *J. Cell. Biochem.* 30:133-170.
- 117.Podack ER *et al.*(1988) Structure and function of perforin. *Curr Top. Microbiol. Immunol.* 140:11-17
- 118.Podack ER (1992) Perforin: structure, function and regulation. *Curr. Top. Microbiol. Immunol.* 178:175-184.
- 119.Preissner K (1991) Structure and biological role of vitronectin. *Ann. Rev. Cell. Biol.* 7:275-310.
- 120.Rao G *et al.* (1987) Complementary DNA and derived amino acid sequence of the a subunit of human complement protein C8: Evidence for the existence of a separate a subunit messenger RNA. *Biochem.* 26:3556-3564.
- 121.Reid KBM y Porter RR (1981) The proteolytic activation systems of the complement. *Ann. Rev. Biochem* 50: 433-64
- 122.Reid KBM y Day AJ (1989) Structure-function relationships of the complement components. *Immunol. Today* 10:177-180.

123. Ross GD (1986) Introduction and history of complement research. En Ross GD (Ed.) *Immunobiology of the complement system*. Academic Press Inc., Orlando, p. 1.
124. Scheren G y Hansch GM (1993) Membrane-associated proteins regulating the complement system: Functions and deficiencies. *Intern. Rev. Immunol.* 10:87-101.
125. Schweinle JE *et al.* (1989) Human Mannose-binding Protein activates the alternative Complement pathway and enhances serum bactericidal activity on a mannose-rich isolate of *Salmonella*. *J. Clin. Invest.* 84:1821-1829
126. Shinkai Y *et al.* (1988) Homology of perforin to the ninth component of complement (C9). *Nature* 334:525-527.
127. Shiver JW *et al.* (1986) The ninth component of human Complement (C9) *J. Biol. Chem.* 261:9629-9636.
128. Sims PJ y Weidmer T (1991) The response of human platelets to activated components of the complement system. *Immunol. Today* 12:338-341.
129. Sitnikova T *et al.* (1995) Interior-branch and bootstrap tests of phylogenetic trees. *Mol. Biol. Evol.* 12:319-333.
130. Sodetz JM (1988) Structure and functions of C8 in the membrane attack sequence of Complement. *Curr. Top. Microbiol. Immunol.* 140:19-31.
131. Stanley KK *et al.* (1985) The sequence and topology of human complement component C9. *EMBO J.* 4, 375-382.
132. Stanley K y Luzio P (1988) A family of killer proteins. *Nature* 334:475-476.
133. Steckel EW (1983) Evidence of direct insertion of terminal complement proteins into cell membrane bilayers during cytolysis. *J. Biol. Chem.* 258: 4318-4324.
134. Steuer KLK *et al.* (1989) Lysis of sensitized sheep erythrocytes in human sera deficient in the second component of complement. *J. Immunol.* 143:2256-2261.
135. Stewart JL (1987) Evidence that C5b recognizes and mediates C8 incorporation into the cytolytic complex of Complement. *J. Immunol.* 139:1960-1964.
136. Strickberger MW (1996a) Systematics and classification. Cap. 11. En Strickberger MW. *Evolution*. Segunda edición. Jones and Bartlett Publishers, Boston. p. 227.
137. Strickberger MW (1996b) Molecular phylogenies and evolution. Cap. 12. En Strickberger MW. *Evolution*. Segunda edición. Jones and Bartlett Publishers, Boston. p.

138. Südhof TC *et al.* (1985a) Cassette of eight exons shared by genes for LDL receptor and EGF precursor *Science* 228:893-895.
139. Südhof TC (1985b) The LDL receptor gene: A mosaic of exons shared with different proteins. *Science* 228:815-822.
140. Super M. *et al.* (1989) Association of low levels of mannan-binding protein with a common defect of opsonisation. *Lancet*. 2: 1236-1239. 1236-1239.
141. Swofford DL (1990) PAUP: Phylogenetic Analysis Using Parsimony, versión 4.0.0d61a. Beta version computer program. Laboratory of Molecular Systematics and pre-release testers.
142. Swofford DL *et al.* (1996) Phylogenetic inference. Cap. 11. En *Molecular Systematics*. Hillis DM, Moritz CM y Mable BK (Eds). Segunda edición. Sinauer Associates, Sunderland Massachusetts. p. 407.
143. Takezaki N *et al.* (1995) Phylogenetic test of the molecular clock and linearized trees. *Mol. Biol. Evol.* 12: 823-833.
144. Tenner AJ y Cooper NR (1981) Identification of types of cells y human peripheral blood that bind C1q. *J. Immunol.* 126:1174-1179.
145. Tomlinson S. *et al.* (1993) Domain structure, Functional activity and polymerization of trout complement protein C9. *Dev. Comp. Immunol.* 17:67-76.
146. Tschopp J *et al.* (1982) Formation of transmembrane tubules by spontaneous polymerization of the hydrophilic complement protein C9. *Nature* 298:534-538.
147. Tschopp J *et al.* (1986) Structural/functional similarity between proteins involved in Complement- and cytotoxic T-lymphocyte-mediated cytotoxicity. *Nature* 322:831-834.
148. Turner MW (1996) Mannose-binding lectin: the pluripotent molecule of the innate immune system. *Immunol. Today* 11:532-540;
149. Ware CF y WP Kolb (1981) Assembly of the functional membrane attack complex of human complement: Formation of disulfide linked C9 dimers. *Proc. Natl. Acad. of Sci.* 10: 6426-6430.
150. Yamamoto K *et al.* (1982) Glutathione-catalyzed Disulfide-linking of C9 in the membrane attack complex of complement. *J. Biol. Chem.* 257:8573-8576.

151. Yeo GSH *et al.* (1997) Cloning and sequencing of complement component C9 and its linkage to DOC-2 in the pufferfish *Fugu rubripes*. *Gene* 200:293-211.
152. Young ZD-E (1986) The ninth component of complement and the pore-forming protein (perforin 1) from cytotoxic T cells: Structural, Immunological, and Functional similarities. *Science* 233:184-190.
153. Ziccardi RJ (1982) A new role for C1-inhibitor in homeostasis: Control of activation of the first component of human complement. *J. Immunol* 128:2505-2508.
154. Ziccardi RJ y Cooper NR (1976) Activation of C1r by proteolytic cleavage. *J. Immunol.* 116:504-509
155. Zimmerman TS *et al.* (1971) A blood coagulation abnormality in rabbits deficient in the sixth component of complement (C6) and its correction by purified C6. *J. Exp. Med.* 134:1591-1600.
156. White RV *et al.* (1994) Characterization of rabbit complement component C8. *J. Immunol.* 152:2501-2508.

Tabla I. Características de los componentes de la cascada del complemento, obtenida a partir de tablas modificadas e información de Liszewski, 1993; Morgan, 1990a; Turner, 1996; Matsushita, 1996.

	Componente	Peso molecular (kDa)	Subunidades	Concentración en plasma (mg/L)	Función
Vía Clásica	C1q	460	6 cadenas A (26 kDa) 6 cadenas B (26 kDa) 6 cadenas C (24 kDa)	80	Une inmunoglobulinas; inicia la activación de la vía clásica
	C1r	83	Individual, se une a C1q	50	Hidroliza y activa C1s
	C1s	83	En plasma forma el complejo C1qC1r <sub>2</sub> C1s <sub>2</sub>	50	Hidroliza y activa C4 y C2
	C4	205	1 cadena α (97 kDa) 1 cadena β (75 kDa) 1 cadena γ (33 kDa)	600	Une a C2 a la membrana durante activación
	C2	102	Una cadena	20	Une y activa C3 y C5
Vía alterna	Factor B	93	Una cadena	210	Une y activa C3 y C5
	Factor D	24	Una cadena	2	Une y activa Factor B
	Properdina	220	4 subunidades idénticas (55 kDa)	26	Estabiliza convertasas
Vía de las lectinas	MBP (Mannan Binding Protein)	500-650	2-4 subunidades formadas por trimeros de cadenas idénticas	2	Activa las vías clásica y alternativa por reconocimiento de oligosacáridos. Funciona como opsonina
	MASP (Mannan Binding Protein Associated Serine Protease)	97	cadena pesada (66 kDa) cadena ligera (31 kDa)	?	Serín proteasa asociada con MBP. Corta a C4 y C2.
Común a todas las vías	C3	185	1 cadena α (110 kDa) 1 cadena β (75 kDa)	1300	Subunidad de C3 y C5 convertasas, une C5 en convertasa, opsonina, factor quimiotáctico.
Fase terminal	C5	190	1 cadena α (115 kDa) 1 cadena β (75 kDa)	70	Inicia la formación del TCC; C5a principal péptido anafiláctico y quimiotáctico
	C6	120	Una cadena	65	Se asocian con C5b unido a membrana para formar sitios de contacto con membrana
	C7	110	Una cadena	55	a los que C9 se pueda unir
	C8		1 cadena α (64 kDa) 1 cadena β (64 kDa) 1 cadena γ (22 kDa)	55	
	C9	69	Una cadena	60	Principal componente del TCC
Regulación en suero	C1-INH	110	Una cadena	200	Une C1r y C1s, disocia C1
	C4bp	500	8 subunidades iguales (70 kDa)	250	Acelera la degradación de C4b2a, auxilia al factor I a degradar C4b
	Factor H	150	Una cadena	450	Acelera degradación de C3bBb, facilita la hidrólisis de C3b por factor I; corta e inactiva C4b y C3b
	Factor I	80	1 cadena α (50 kDa) 1 cadena β (38 kDa)	35	Serín Proteasa
	proteína S	83	Una cadena	500	Se une a C5b-7 en el suero
Regulación en membrana	CR <sub>1</sub>	160-250	Una cadena	?	Aceleran la degradación de las convertasas de C3 y C5
	DAF	70	Una cadena	?	auxiliares en la degradación de C3b
	MCP	45-70	Una cadena	?	Control de la formación del TCC
	CD59	20	Una cadena	?	
Receptores	CR <sub>2</sub>	140	?	?	Une C3dg/C3bi Activación de células B, localización de complejos inmunes, invasión de EBV a la célula
	CR <sub>3</sub>	160/90	?	?	Une C3bi; función fagocítica
	CR <sub>4</sub>	150/90	?	?	Une C3bi; función fagocítica
	C3aR	70	?	?	Se le une C3a, activación celular
	C5aR	80	?	?	Se le une C5a; implicado en activación celular y quimiotaxis
	C1qR	80	?	?	Receptor de C1q y MBP; fagocitosis
	MBPR	60	?	?	Receptor de MBP, media fagocitosis