

**DETERMINANTES ESTRUCTURALES DEL
PCI QUE PUEDEN EXPLICAR SU
ACTIVIDAD ANTAGONISTA DEL EGF**

*Descripción de las claves estructurales comunes
entre el PCI y el EGF y su implicación sobre su
relación funcional*

CAPÍTULO VI

VI-1 Introducción

El conjunto de algoritmos y programas presentados hasta ahora permite realizar profundos análisis estructurales sobre proteínas ricas en puentes de azufre. Con tal finalidad fue desarrollado el KNOT-MATCH, programa que nos permite alinear estructuralmente proteínas utilizando la topología de sus puentes de azufre como elemento estructural común. Mediante el uso de este programa se pudo inferir en la relación estructural y funcional entre proteínas hasta entonces no relacionadas como el PCI y el EGF. El FOLD-SCOPE se desarrolló para disponer de una herramienta potente y flexible para la comparación del esqueleto polipeptídico de proteínas. Los resultados obtenidos mediante el uso de estas herramientas fueron contrastados, siempre que fue posible, con datos experimentales publicados por otros autores o por otros miembros del grupo.

El trabajo que aquí se presenta forma parte de un proyecto más amplio dedicado a rediseñar el Inhibidor de Carboxipeptidasa de patata (PCI) y al diseño de proteínas derivadas del mismo para su aplicación como herramienta antitumoral. Este proyecto consta de tres líneas paralelas y complementarias de trabajo:

- a) una línea de biocomputación, cuyos resultados se muestran en esta tesis,
- b) trabajos de biología molecular, dedicados al rediseño experimental del PCI y,
- c) una línea encargada de los cultivos y análisis *in vivo*.

A partir del trabajo desarrollado en nuestro grupo ha podido establecerse una relación funcional entre el PCI y el EGF (Blanco-Aparicio et al., 1998; Mas et al., 1998). En estos artículos se describe la capacidad del PCI para unirse a los receptores del EGF, comportándose como un antagonista y mostrando un comportamiento interesante para el desarrollo aplicado al PCI como inhibidor de la actividad tumoral. Para ello ha sido necesario establecer la relación estructural y funcional de las proteínas miembro de la familia *EGF-like* (ver capítulo IV), especialmente del EGF (*Epidermal Growth Factor*), y el PCI (*Potato Carboxypeptidase Inhibitor*).

Los resultados experimentales obtenidos describen la unión del PCI a los receptores del EGF, produciéndose una disminución de la capacidad de duplicación de las células tumorales. Se ha podido determinar en cultivos tumorales de células humanas que el PCI actuaría como el primer antagonista conocido del EGF en mamíferos (Blanco-Aparicio et al., 1998). Experimentos posteriores mostraban la capacidad del PCI para ser internalizado en la célula. Paralelamente se llevaron a cabo estudios in vivo sobre ratones transgénicos que confirmaban los resultados obtenidos en cultivos celulares. Se observó además que el suministro de PCI a estos ratones transgénicos alargaba la vida de los animales, producía una disminución significativa del tamaño de los tumores y provocaba una disminución de la irrigación de las zonas tumorales. Tales resultados experimentales eran paralelamente utilizados en el laboratorio de biocomputación para la interpretación estructural de los fenómenos observados, y la predicción y diseño de nuevos experimentos (Blanco-Aparicio et al., 1998).

El mecanismo de funcionamiento del PCI era y es desconocido, del mismo modo que no está claro la mecánica de funcionamiento del EGF. Se han descrito varias moléculas receptoras de EGF (EGFR), las cuales muestran variaciones de afinidad entre los distintos miembros de la familia *EGF-like*. Diversos autores han descrito un carácter dimérico en la funcionalidad del receptor, pudiéndose formar homo y heterodímeros. Así, para que se produzca la funcionalidad natural del EGF mayoritariamente es necesario que una molécula de receptor se una a una molécula de EGF. La pareja EGF-EGFR se une a otra pareja similar y ambas son entonces internalizadas al interior de la célula, generalmente mediante vesículas de clatrina. Este hecho desencadena la entrada en la fase mitótica de la célula. Otros autores describen un comportamiento distinto para la pareja EGF-EGFR.

El desconocimiento de los mecanismos estructurales y funcionales del EGF complica y contribuye a dificultar los análisis de los resultados experimentales obtenidos sobre el PCI. Estos motivos nos obligan a trabajar sobre hipótesis que se basan en datos no siempre probados. Estas hipótesis, como en toda línea de investigación abierta, son contrastadas experimentalmente y rechazadas o modificadas por otras cuando las evidencias experimentales o teóricas así lo indican.

En este capítulo se muestran los resultados más relevantes que se desprenden del trabajo desarrollado en esta tesis, si bien se requerirá la posterior comprobación experimental de los mismos.

VI-B Objetivos

El objetivo último de este proyecto es diseñar una molécula o moléculas antagonista(s) al EGF que pueda(n) ser utilizadas con finalidades médicas en el tratamiento del cáncer. Este ambicioso propósito lleva consigo otros propósitos más asequibles, al menos en el estado actual del proyecto son los siguientes:

- a) desvelar las claves estructurales que relacionan al PCI con el EGF, dicho de otro modo, determinar los residuos o regiones del PCI que presentan homología estructural con el EGF y que lo hacen responsable de los efectos funcionales que se han descrito, y
- b) proponer una serie de mutantes que hagan al PCI un mejor antagonista, aumentando su afinidad por los receptores de EGF potenciando la competencia PCI-EGF y promover en lo posible la acción de apoptosis celular y la inhibición de la vascularización tumoral
- c) en un futuro sería interesante el diseño de moléculas peptidomiméticas que sean fáciles de obtener y que mejoren la farmacocinética y la farmacodinámica del PCI.

VI-C Métodos

Del desconocimiento del mecanismo de funcionamiento del EGF surge la necesidad de plantear ciertas hipótesis. Es lícito, o cuanto menos precavido, reconocer la provisionalidad de algunos resultados experimentales que nos van a servir de partida del análisis. Es también previsible que en cierto tiempo algunas de las consideraciones que tomamos aquí como hipótesis sean definitivamente contrastadas o bien definitivamente descartadas.

Partimos de las siguientes premisas experimentales e hipótesis de partida:

1. El PCI compite con el EGF en la unión por su receptor. De este modo el PCI se puede comportar como un antagonista de la acción del EGF.
2. Conocemos varios residuos del EGF implicados directa o indirectamente en la unión con su receptor (Tabla IV.3). Esta información, publicada por diversos autores, suele ser contradictoria o incompleta.
3. Conocemos ciertos datos del funcionamiento del receptor de EGF. Desde la llegada de la señal, llegada de una molécula de EGF, la dimerización, la transducción de la señal al interior de la célula y la posterior internalización de los receptores para su recirculación. De modo análogo al punto anterior esta información es muy contradictoria entre los diversos autores.

Se asigna al EGF un importante papel en el proceso de dimerización. Para ello cabe concebir el EGF como una molécula en la que se pueden diferenciar dos partes, una **cabeza** y una **cola** (Figura VI.1).

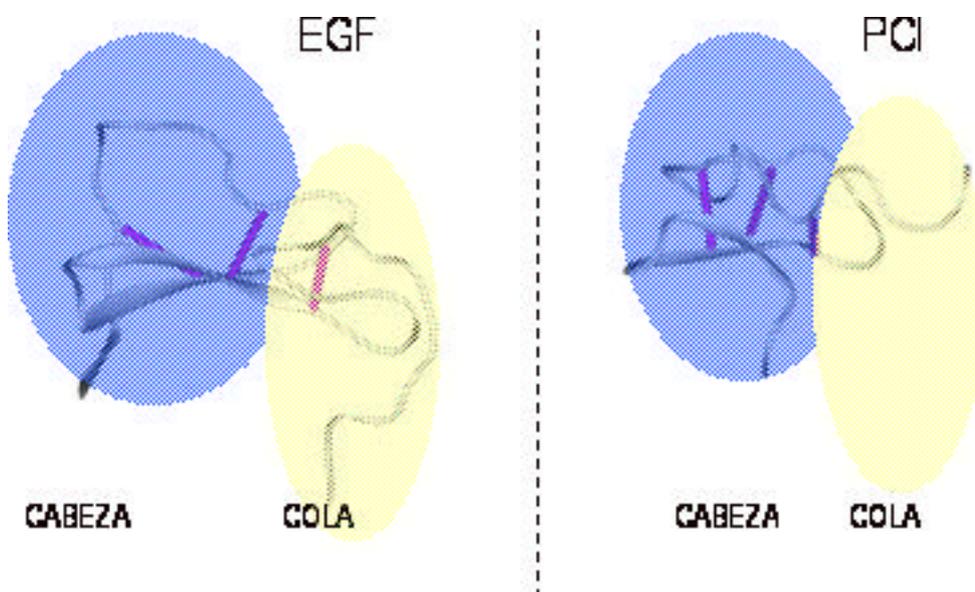


Figura VI.1 Se marcan las regiones funcionales en las que se han dividido las moléculas EGF y PCI. Las sombras no se corresponden a ningún posicionamiento concreto determinado, siendo meramente indicativas de regiones de las moléculas.

Consideramos dos posibilidades no excluyentes que justifican este comportamiento observado experimentalmente de la dimerización:

La cabeza del EGF interaccionaría con el EGFR, mientras la cola, formada principalmente por el extremo N-terminal del EGF (parte de la cual carece el PCI), quedaría expuesta, pudiendo unirse al otro heterodímero EGF-EGFR. La formación del tetrámero es esencial para la consiguiente transmisión de la entrada en fase mitótica al interior de la célula. El PCI, como se puede ver en la región amarillenta de la Figura VI.1 carece de una parte importante de la cola, sin embargo, un *cluster* de residuos hidrofóbicos¹ (muchos de ellos aromáticos) son comunes a ambas proteínas, justificando estos residuos que el PCI pueda dimerizar en los receptores de EGF y comportarse como un antagonista, aunque menos potente.

El PCI con la región N-terminal de menor tamaño se une a los monómeros de EGFR por la cabeza, de estructura similar, pero al no tener parte de la cola la dimerización se produce sin la consiguiente transmisión de la señal de entrada en fase M al interior de la célula. En este caso la cola del EGF, su extremo N-terminal, es de vital importancia. Un resumen del papel atribuido a la cola sería el siguiente:

- a) de facilitar la formación del tetrámero (EGFR-EGF)-(EGF-EGFR), y
- b) producir la transmisión de la señal al interior de la célula.

La unión del PCI a los receptores de EGF sin producir la transmisión de señal de entrada en mitosis produce una inutilización de los EGFR, que nos lleva a observar efectos de funcionamiento antagónico.

¹ El importante papel que se otorga a este cluster hidrofóbico será tratado ampliamente a continuación

El cluster hidrofóbico

En la literatura se describen un grupo de proteínas pequeñas, ricas en puentes de azufre, y con funciones extracelulares (como se detalla en los capítulos III y IV). En algunas de estas proteínas se ha descrito la presencia de *clusters* hidrofóbicos similares a los descritos aquí para el EGF y el PCI. Los distintos autores asignan a estos residuos una importancia crucial para la dimerización, ya que un elevado número de estas proteínas presentan su actividad tras sufrir procesos de dimerización. Este *cluster* hidrofóbico común al EGF y al PCI ha podido ser detectado para otras moléculas de la familia EGF-*like* como el TGF.

A partir de datos descritos por otros autores y por datos obtenidos bajo análisis propios (Mas et al., 1998) se ha podido establecer el siguiente patrón:

- es frecuente encontrar proteínas de diversas familias que actúan de forma oligomérica unidas por un cluster hidrofóbico similar al cluster que tiene el EGF, el TGF y el PCI.
- cuando estas proteínas mensajeras (hormonas, venenos, factores de crecimiento, inhibidores, etc) no están transportándose, sino que están realizando su función, lo suelen hacer en forma de complejos generalmente díméricos, es decir, el EGF, pese a ser un transmisor de señal, precisa de la presencia de otra molécula de EGF para que se pueda obtener respuesta funcional (dimerización del receptor).

Aunque estos patrones sean comunes para un gran número de proteínas, las funciones y orígenes evolutivos de estas moléculas son muy diversos.

Un dato significativo es que el cluster hidrofóbico está formado por las cadenas laterales de residuos del cuerpo de ambas proteínas, PCI y EGF. Sin embargo, este cluster hidrofóbico se ve enriquecido por la presencia de varios residuos de la cola N-terminal para el EGF, lo que liga este cluster hidrofóbico a la cola N-terminal del EGF.

La hipótesis de partida

El EGF se une al receptor por su cabeza. La cola del EGF y especialmente el extremo N-terminal quedan expuestos, del mismo modo quedan expuestos los residuos que conforman el *cluster* hidrofóbico. Finalmente, tanto la cola del EGF como la presencia del cluster permiten la consiguiente dimerización del par (EGFR-EGF)-(EGF-EGFR) y la consiguiente transmisión de la señal al interior de la célula para que ésta entre en fase M.

El *cluster* hidrofóbico podría además tener un papel en la disposición del EGF correctamente dentro de su receptor. El PCI, de mucho menos volumen, podría interactuar con los EGFR por la cabeza, tal y como se ha propuesto. Su menor volumen podría hacer que no estuviera suficientemente bien anclado al receptor, de modo que o bien el *cluster* hidrofóbico o la falta de un extremo terminal equivalente al EGF dificultara su posicionamiento en el EGFR. Una posible consecuencia de esta falta de volumen o imposibilidad de tener un anclaje más sólido pueden hacer que la afinidad del PCI por los EGFR sea significativamente menor que la del EGF.

Consecuentemente podemos suponer que si añadimos al PCI la cola del EGF éste debe aumentar su afinidad por el receptor del EGF. Sin embargo, las diferencias estructurales de ambas moléculas, tanto en la cabeza, como en la relación estructural cabeza-cola del EGF respecto del PCI-ColaEGF (PCI quimérico) podrían impedir que se activara la transmisión de la señal al interior de la célula. En el caso en que la señal de entrada en fase M fuera transmitida cabría pensar en la posibilidad de mutar la cola del EGF implantada sobre el PCI y eliminar aquellos residuos que otorgan tal funcionalidad.

VI-D Resultados

Análisis de la familia *EGF-like*

La familia *EGF-like* ya ha sido tratada a lo largo de esta tesis en los capítulos III y IV. Su importancia estructural, al margen de la relación existente entre sus miembros y el PCI, yace esencialmente en el hecho de ser una familia de proteínas ampliamente estudiada por diversos autores y desde muy diversos puntos de vista. No obstante, como ya ha sido comentado, estos datos muchas veces contradictorios o faltos de detalle impiden extraer información concluyente.

La superposición de todos los miembros de la familia *EGF-like* que presentamos se ha llevado a cabo utilizando el programa KNOT-MATCH, es decir, basando la superposición en datos puramente estructurales comunes a todos ellos. Recapitulando sobre estos datos presentados en el capítulo IV de este trabajo podemos decir que :

- es posible alinear las topologías de los puentes de azufre de todos los miembros de esta familia,
- el alineamiento de la topología de puentes de azufre permite el correcto alineamiento del andamiaje de las proteínas, observándose también un correcto alineamiento de lazos y estructuras secundarias regulares,
- el análisis de la disposición de las cadenas laterales de ciertos residuos en el espacio permite extraer un *consensus* característico de la familia. Este *consensus*, junto con la información publicada de estas proteínas, nos permite determinar ciertos residuos como los principales responsables de la unión a los receptores.

De este modo se pudo extraer un patrón característico de la familia *EGF-like* el cual se puede ver en la Figura IV.3, siendo los resultados obtenidos contrastados con datos publicados, tal y como muestra la Tabla IV.3.

Se pueden apreciar 15 posiciones, de las 53 posibles en el EGF, a las que se les ha asignado algún papel funcional y/o estructural. De estas posiciones, siete son presentadas como residuos implicados en la unión con el receptor, concretamente Y13, N16, I23, Y29, Y37, R41 y R45. Además de estas posiciones se han descrito otras, no incluidas en el *consensus*, que podrían tener un papel importante, como es la L47 y el extremo C-terminal del EGF.

Relaciones estructurales entre el EGF y el PCI

El PCI y el EGF son moléculas muy alejadas evolutivamente, siendo el PCI una proteína de origen vegetal y el EGF una factor de crecimiento de mamíferos. Estas proteínas, que secuencialmente tampoco están relacionadas, han podido ser relacionadas estructuralmente por el programa KNOT-MATCH. En el trabajo que se presenta en este capítulo es resultado de análisis concretos realizados tras el uso de los programas KNOT-MATCH (capítulos III y IV) y FOLD-SCOPE (capítulo V), cuyos métodos fueron descritos en los capítulos correspondientes. En el capítulo IV se muestra una clasificación de proteínas ricas en puentes de azufre y en ella el PCI y el EGF forman parte de un mismo **cluster**, indicando que están fuertemente relacionadas por la topología de sus puentes de azufre.

En un principio las regiones con similitud estructural fueron determinadas por inspección visual y analizadas mediante el cálculo de RMSD (Mas et al., 1998). Para agilizar este cálculo y correspondiente análisis se creó un programa (FRAZER) especialmente adaptado para hacer este tipo de cálculos. Dos regiones fueron determinadas como similares estructuralmente entre el PCI y miembros de la familia *EGF-like*, especialmente el EGF. Así, el fragmento 27-34 del PCI se relacionaba con el fragmento H22-Y29 del EGF con un RMSD de 0,7 Å mientras para el resto de la familia de los EGF se podían determinar valores próximos a 1 Å. Análogamente, la región 18-24 en el PCI mostraba regiones equivalentes con miembros de la familia EGF, sirva como ejemplo la relación existente con el 1ccf con 0.5 Å (fragmento C55-C61).

El desarrollo del programa FOLD-SCOPE ha permitido establecer relaciones más profundas entre el EGF y el PCI, por lo que se refiere al andamiaje de ambas proteínas. La Figura VI.2 muestra los detalles de la superposición.

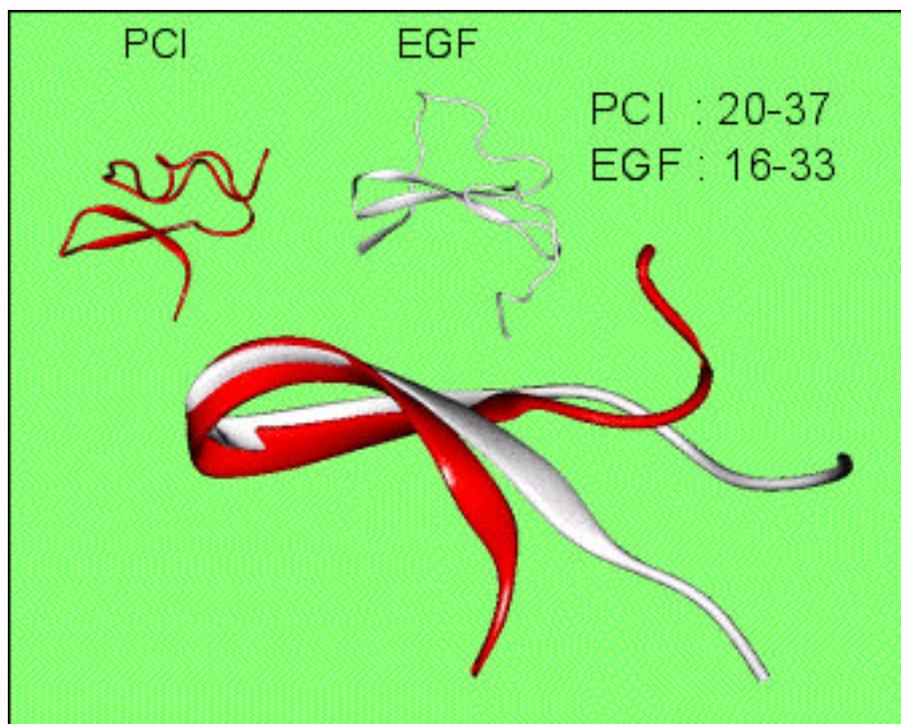


Figura VI.2. Representación en cinta de los esqueletos de las proteínas EGF y PCI.

Como se puede ver en la Figura VI.2 ha sido posible relacionar una amplia región de ambas moléculas, en concreto 17 residuos con un RMSD de aproximadamente 3,5 Å, siendo esta región significativamente más grande que la determinada por inspección visual (Mas et al., 1998). De este modo ha sido posible relacionar el 45% de la estructura del PCI con el EGF. Gracias al sistema de auto-focus disponible en el FOLD-SCOPE fue posible determinar la región de mayor similitud estructural, concretamente la región 25-33 en el EGF podía superponerse con la región 29-36 del PCI con un RMSD de aproximadamente 1,2 Å.

Los resultados indican que las proteínas EGF y PCI, pese a tener una homología secuencial muy baja, están claramente relacionadas estructuralmente.

Comparación espacial de grupos químicos equivalentes.

Estas proteínas, como ya hemos visto, tienen similar andamiaje, elemento necesario para realizar la comparación de grupos químicos equivalentes. El uso del KNOT-MATCH ha sido muy útil para la superposición de las estructuras de ambas proteínas y el análisis de los grupos químicos equivalentes en el espacio y ha permitido establecer importantes relaciones estructurales entre ellas. Esta superposición alinea correctamente las estructuras secundarias regulares de ambas proteínas, como ya ha sido expuesto. Los elementos utilizados para tal fin forman un patrón estructural común no tan sólo de estas dos proteínas o familias, sino de un heterogéneo grupo de familias de proteínas (Mas et al., 1998).

Cuando se realiza tal superposición es posible observar que hay nueve grupos químicos equivalentes situados en el espacio, es decir, hay nueve aminoácidos cuyas cadenas laterales están orientadas en el espacio de un modo similar y las cuales aportan a esa posición del espacio unas propiedades fisicoquímicas similares (Figura VI.3). Algunas de estas posiciones han sido descritas como estructural o funcionalmente importantes para el PCI o para el EGF (Rees & Lipscomb, 1982; Brown et al., 1989; Brown & Wüthrich, 1992; Ullner et al., 1992; Groenen et al., 1994, Picot et al., 1994; Molina et al., 1994; Barbacci et al., 1995; Jacobsen et al., 1996; Blanco-Aparicio et al., 1998; Mas et al., 1998). La Figura VI.3 muestra los residuos alineados de las dos proteínas. Cinco de estos residuos son además conservados para todo el *consensus* extraído de la familia EGF-like que se muestra en la Figura IV.3. En concreto los residuos D11, V19, S25, T30 y R41 se conservan para la mayoría de los miembros de la familia del *EGF-like*.

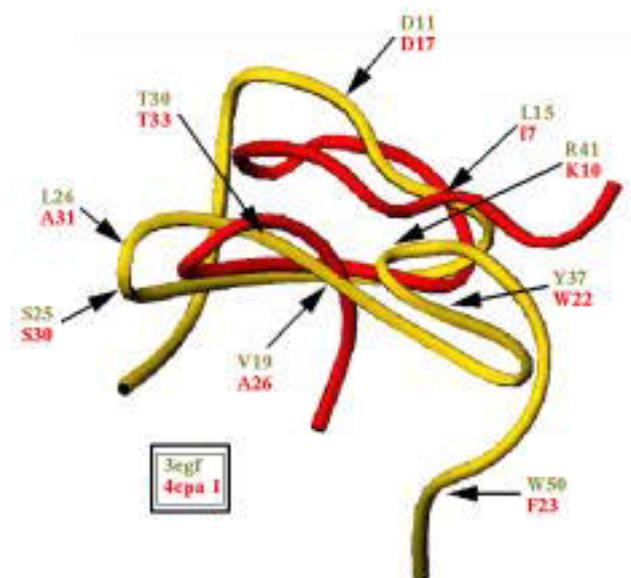


Figura VI.3 Representación en cinta de las proteínas EGF y PCI tras la superposición por la topología de sus puentes de azufre. Se señalan además los residuos cuyas cadenas laterales ofrecen al espacio similares propiedades fisicoquímicas.

De las nueve posiciones señaladas en la Figura IV.3 cabe destacar siete residuos por su especial similitud. Estas siete posiciones han sido agrupadas en tres regiones (Figura VI.4). **La Región 1** contiene los aminoácidos L15 y R41 del EGF y los I7 y K10 del PCI. **La Región 2** está compuesta por los residuos V19, S25, L26 y T30 del EGF y sus equivalentes del PCI que son A26, S30, A31 y T33. Finalmente la **Región 3** que incluye los residuos aromáticos Y37 y W50 del EGF y W22 y F23 del PCI.

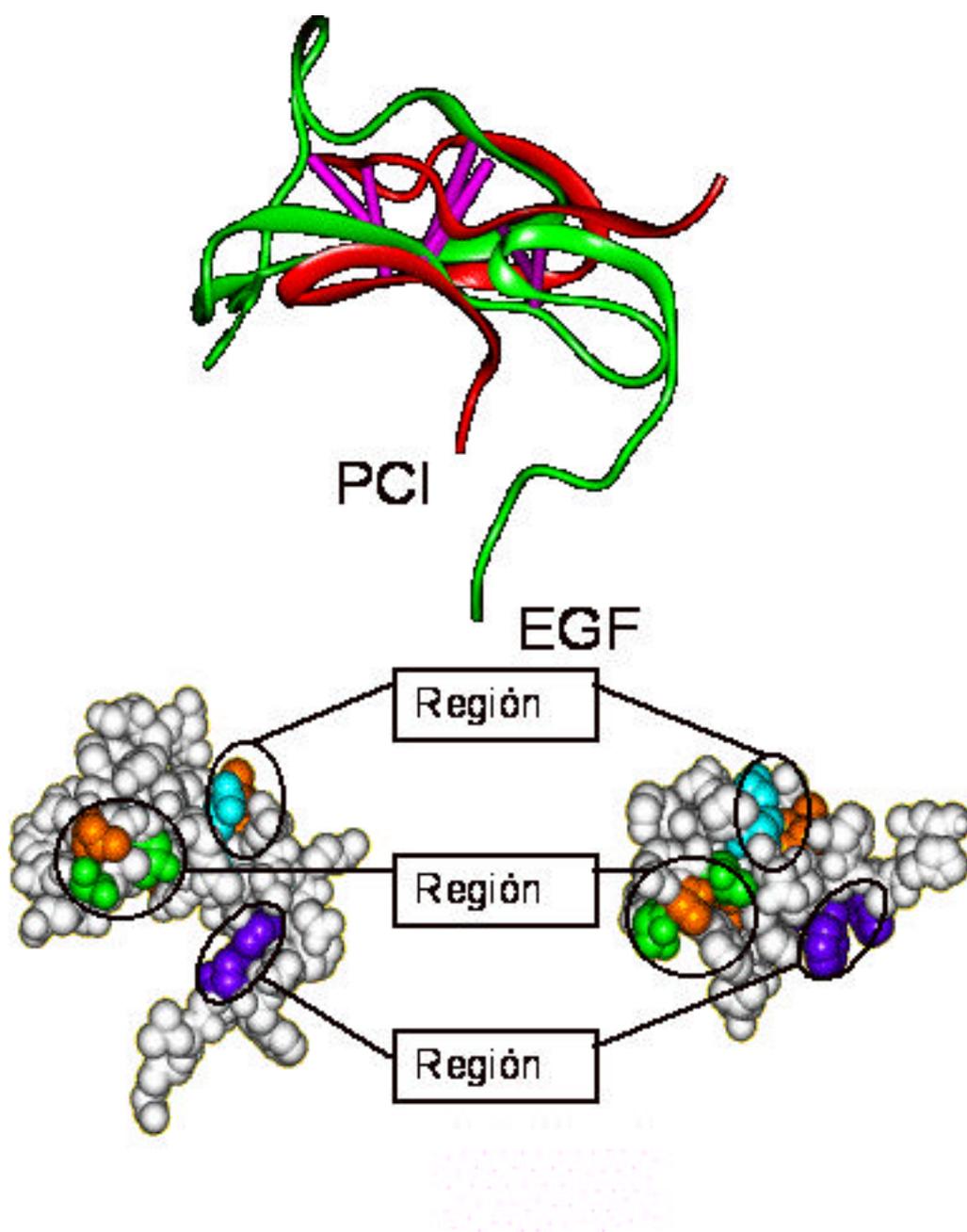


Figura VI.4 : EGF y PCI superpuestas por la topología de sus puentes de azufre. En la parte superior se observa una representación en cinta de ambas moléculas con los puentes de azufre destacados. En la parte inferior se observa una representación en esferas de van der Waals. Se destacan en este caso con colores el tipo de residuo. Así el color verde indica residuo polar sin carga, el marrón es apolar, el azul es cargado positivo y el violeta son residuos aromáticos.

La Región 1 y la Región 3 contienen el mayor número de residuos descritos en la literatura como implicados en la unión con el receptor. Por su parte, a la Región 2 se le asigna una función más estructural, según los residuos implicados en esta región.

Estudios para el diseño de un PCI quimérico. La movilidad de las colas N y C-Terminal del PCI

El análisis de la movilidad de las colas N y C-Terminal del PCI han de permitir una mejor comprensión del efecto de esta proteína en los EGFR. El PCI y EGF son moléculas cuya estructura nativa depende esencialmente de la presencia de los puentes de azufre. Estos enlaces fijan por tres puntos las proteínas y hacen que éstas puedan mantener su estructura en condiciones cambiantes del medio. Consecuentemente las regiones centrales de estas proteínas (*core*) tienen muy poca movilidad. Sin embargo, los lazos más grandes y los extremos N y C terminales de ambas proteínas quedan mucho menos anclados al *core* de la proteína, observándose en estos casos una cierta movilidad.

Cuando analizamos cada caso en concreto observamos que los extremos N y C terminal del EGF son muy móviles, tal y como se describe en la literatura. Los datos disponibles sobre la movilidad de los extremos N y C terminales del PCI fueron obtenidos por investigadores de nuestro grupo (Oliva et al. 1995). Se realizó un estudio por **Dinámica Molecular** partiendo de la estructura resuelta por Cristalografía de Rayos X del PCI en un medio acuoso durante 500 ps. En la Figura VI.4 se puede observar que el cuerpo central del PCI (*core*) mantiene su estructura tras los 500 ps de proceso, tal y como cabría esperar, mientras los lazos sufren una ligera deformación. Del mismo modo el extremo C-terminal del PCI, con tan sólo 4 residuos no sufre variación estructural significativa. El extremo N-terminal del PCI, aquél que se corresponde estructuralmente con el C-terminal del EGF, sufre una variación estructural significativa.

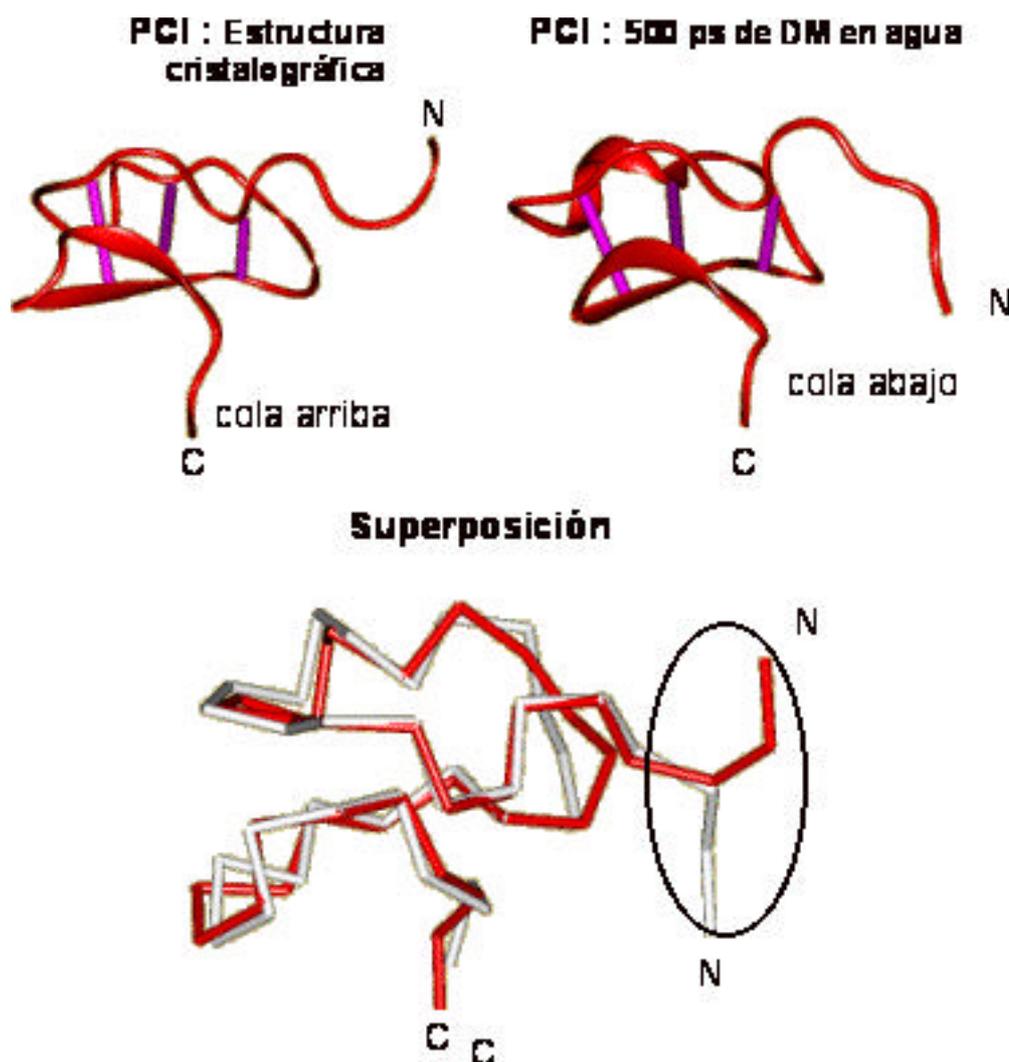


Figura VI.5 Representación en cinta de la estructura cristalina y tras 500 ps de Dinámica Molecular del PCI en agua.

En la Figura VI.5 se puede observar el movimiento que ha sufrido el extremo N-terminal del PCI. El motivo por el cual el PCI cristalizado muestra la **cola arriba** es porque las coordenadas disponibles del PCI corresponden al PDB 4cpaI. En este archivo, el PCI fue cristalizado unido a la carboxipeptidasa, a quien está inhibiendo, y aparecen interacciones entre el extremo C-terminal del PCI y la carboxipeptidasa. Al poner en disolución esta molécula se observa una clara tendencia a adoptar la conformación **cola abajo**.

VI-E Conclusiones

De los resultados obtenidos se pueden extraer varias conclusiones:

1. Según datos experimentales del propio grupo y de información de otros autores podemos decir que la cola N-terminal del EGF tiene un importante papel en la funcionalidad del EGF.
2. El extremo C-terminal del EGF se corresponde topológicamente con el extremo N-terminal del PCI.
3. Los extremos N-terminal del PCI y el C-terminal del EGF son muy flexibles.
4. La superposición por puentes de azufre del PCI y del EGF, en el cuerpo central de ambas proteínas, permiten un correcto alineamiento de la cola C-terminal del EGF colocada sobre el extremo N-terminal del PCI.
5. El residuo L47, definido como residuo funcionalmente muy importante para el EGF por varios autores, se corresponde mediante alineamiento espacial, al residuo Q1 del PCI.
6. A la vista de los resultados descartamos utilizar la cola C-terminal del PCI para emular la cola C-terminal del EGF.

CONCLUSIONES

Conclusiones generales de la tesis

CAPÍTULO VII

Podemos destacar las siguientes conclusiones globales obtenidas durante el desarrollo de este trabajo de tesis doctoral:

1. Se han establecido los métodos matemáticos y algoritmos que permiten superponer estructuralmente proteínas ricas en puentes de azufre. Se ha desarrollado el programa KNOT-MATCH, accesible desde internet en la dirección <http://luz.uab.es>, que agrupa todos los algoritmos mencionados, permitiendo al usuario seleccionar el algoritmo y parámetros que considere oportunos en cada caso.
2. Se han aplicado métodos de cálculo utilizados en visión por computación para el análisis estructural de esqueletos polipeptídicos de proteínas. Estos algoritmos permiten eliminar errores clásicos de las metodologías utilizadas para la superposición de estructuras, aumentar la eficiencia de cálculo para este cometido y disponer de un algoritmo eficiente para comparar elementos subestructurales y esqueletos polipeptídicos de proteínas.
3. Se ha podido establecer que la topología de puentes de azufre de una proteína, es decir, la arquitectura que dibujan sus puentes en el espacio muestran ciertas regularidades entre proteínas aparentemente no relacionadas. Esto nos ha permitido relacionar estructuralmente proteínas que por los métodos clásicos no hubiera sido posible.
4. Se ha establecido una clasificación preliminar de proteínas basada en la topología de sus puentes de azufre. Se ha propuesto la redifinición del grupo de proteínas conocido como T-Knot agrupando las definiciones dadas por diversos autores.

5. Se han predicho los determinantes estructurales que permiten explicar el efecto antagonista del PCI al EGF al unirse a los receptores EGFR y se han realizado los análisis previos a la construcción de una quimera PCI-EGF.

REFERENCIAS BIBLIOGRÁFICAS

- Alexandrov, N. N., Takahashi, K., Go, N. (1992) Common spatial arrangements of backbone fragments in homologous and non-homologous proteins. *J. Mol. Biol.* 225: 5-9.
- Avilés FX, Vendrell J, Guash A, Coll M, Huber J. (1993). Advances in metallo-procarboxypeptidases. Emerging details on the inhibition mechanism and the activation process. *Eur.J. Biochem.*, 211:381-389.
- Barbacci EG, Guarino BC, Stroh JS, Singleton DH, Rosnack JK, Moyer JD. (1995). The structural basis for the specificity of epidermal growth factor and heregulin binding. *J. Biol. Chem.* 16:9585-9589.
- Bernstein FC, Koetzle TF, Williams GJB, Meyer EF, Brice MD, Rodgers JR, Kennard O, Shimanovich T, Tasumi M. (1977). The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535-542.
- Betz, S. F. (1993). Disulfide bonds and the stability of globular proteins. *Protein Sci.* 2:1551-1558
- Blanco-Aparicio C, Molina MA, Fernandez-Salas E, Frazier ML, Mas JM, Querol E, Avilés FX, de Llorens R. (1998). Potato Carboxypeptidase Inhibitor, a T-knot protein, is an epidermal growth factor antagonist that inhibits tumor cell growth. *J. Bio. Chem.* 20:12370-12377.
- Bork P, Ouzounis C, Sander C. (1994). From genome sequences to protein function. *Curr. Opin. Struct. Biol.* 4:393-403.
- Bork, P., Sander, C. , Valencia (1992), A. An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proc. Natl. Acad. Sci. USA* 89:7290-7294.
- Brown, L.R. & Wüthrich, K. (1992). Nuclear magnetic resonance solution structure of the a-neurotoxin from Black Mamba. *J. Mol. Biol.* 227
- Brown, S.C., Mueller, L. & Jeffs, P.W. (1989). ¹H assignment and secondary structural of human transforming growth factor a. *Biochemistry*, 28
- Chang J, Shindler P, Ramseier U, Lai P. (1995). The disulfide folding pathway of human epidermal growth factor. *J. Biol. Chem.* 270:9207-9216.
- Chang, J., Canals, F., Shindler, P., Querol, E. & Avilés, F.X. (1994) The disulfide folding pathway of potato carboxipeptidase inhibitor. *J. Biol. Chem.* 269:33087-33094.
- Chatrenet, B. & Chang, J. (1992). The folding of hirudin adopts a mechanism of trial and error. *J. Biol. Chem.* 267:3038-3043.
- Chothia C. Proteins. One thousand families for the molecular biologist. *Nature.* (1992) 357:543-544.
- Chothia, C. & Lesk, A.M. (1986). The relation between the divergence of

- sequence and structure in proteins. *EMBO J.* 5:823-826.
- Cohen, F. E., Gregoret, L., Presuell, S. B. & Kuntz, I. D. (1989). Protein structure prediction: new theoretical approaches. *289:75-85.*
- Everit, B. (1974) A review of clustering techniques in Cluster analysis. ed Heinemann educational books Ltd..
- Groenen, L.G., Nice, E.C. & Burgess, A.W. (1994). Structure-function relationships for the EGFT/TGF- β family of mitogens. *Growth Factors*, 11:235-257.
- Harrison P. & Sternberg M.J.E. (1996) The disulphide β -cross: from cysteine geometry and clustering to classification of small disulphide-rich protein folds. *J. Mol. Biol.* 264:603-623.
- Harrison, P. M. & Sternberg, M. J. E. (1994). Analysis and classification of disulphide connectivity in proteins. The entropic effect of cross-linkage. *J. Mol. Biol.* 244:448-463.
- Hill, C.P., Yee, J., Slested, M.E. & Eisenberg, D. (1991). Crystal structure of defensin HNP-3, an amphiphilic dimer: Mechanism of membrane permeabilization. *Science*, 251:1481-1485.
- Hobohm, U. & Sander, C. (1994). Enlarged representative set of protein structures. *Protein Sci.* 3:522-524.
- Hobohm, U., Scharf, M., Schneider, R., Sander C. (1992). Selection of a representative set of structures from the Brookhaven Protein Data Bank. *Protein Sci.* 1:409-417.
- Holm L, Murzin AG, Sander C. (1994). Three sisters, different names: 3 α ,20 β -dihydroxysteroid dehydrogenase, dihydropteridine reductase and UDP-galactose 4-epimerase. *Nature Struct. Biol.* 1:46-147.
- Holm L, Ouzounis C, Sander C, Tuparev G, Vriend G. (1992). A database of protein structure families with common folding motifs. *Prot. Sci.* 1:1691-1698.
- Holm L, Sander C, Schnarr M, Rüterjans H, Fogh R, Boelens R, Kaptein R. (1993). LexA repressor and iron uptake regulator from Escherichia coli: new members of the CAP-like DNA binding domain superfamily. *Prot. Eng.* 7:1449-1453.
- Holm L, Sander C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* 223:123-138.
- Holm L, Sander C. (1994). The FSSP database of structurally aligned protein fold families. *Nucleic Acid Res.* 22:3600-3609.
- Holm, L. & Sander, C. (1996). The FSSP database: fold classification based on structure-structure alignment of proteins. *Nucleic Acid Res.* 24:206-209.

- Hrabal R, Chen Z, James S, Bennet HPJ, Feng N. (1996). The hairpin stack fold, a novel protein architecture for a new family of protein growth factors. *Nature Struct. Biol.* 3:747-752.
- Isaacs, N.W. (1995). Cysteine knots. *Curr. Opin. Struct. Biol.* 5:391-395.
- Jacobsen, N.E., Abadi, N., Sliwkowski M.X., Reilly, D., Skelton N.J. & Fairbrother, W.J. (1995). High-resolution structure of the EGF-like domain of heregulin- β *Biochemistry*, 35:3402-3417.
- Johnson MS, Srinivasan N, Swdhanini R, Blundell TL. (1994). Knowledge-based protein modeling. *Crit. Rev. Biochem. Mol. Biol.* 29:1-68.
- Levitt, M. & Chothia. C. (1976). Structural patterns in globular proteins. *Nature*, 261:552-558.
- Lin SL, Nussinov R. (1995). A disulphide-reinforced structural scaffold shared by small proteins with diverse functions. *Nature Struct. Biol.* 2:835-837.
- Madej, T., (1995). Threading analysis suggests that the obese gene product may be a helical cytokine. *FEBS Lett.* 373:356-359.
- Mas JM, Aloy P, Marti-Renom MA, Oliva B, Blanco-Aparicio C, Molina MA, de Llorens R, Querol E, Aviles FX (1998). Protein similarities beyond disulphide bridge topology. *J Mol Biol* 284:541-548.
- Matthews, B. W. , Rossmann, M. G. (1985). Comparison of protein structures. *Meth. Enzymol.* 115:397-420.
- McInnes C, Hoyt DW, Harkins RN, Pagila RN, Debanne MT, O'Connor-McCourt M, Sykes D. (1996). NMR study of the transforming growth factor TGF- β Epidermal growth factor complex. *J. Biol. Chem.* 271:32204-32211.
- Michie, A. D., Orengo, C. A., Thornton, J. M. (1996). Analysis of domain structural class using an automatic class assignment protocol. *J. Mol. Biol.* 262:168-185.
- Molina MA, Marino C, Oliva B, Avilés FX, Querol E. (1994). The C-tail valine is a key residue for the stabilization of the complex between potato inhibitor and carboxypeptidase A. *J. Biol. Chem*, 269:21467-21472.
- Montelione, G.T., Wüthrich, K., Burgess, A.W., Nice, E.C., Wagner, G., Gibson, K.D. & Sheraga, H.A. (1992). Solution structure murine epidermal growth factor determined by NMR spectroscopy and refined by energy minimization with restraints. *Biochemistry*, 31:236-249.
- Murray-Rust J, McDonald NQ, Blundell TL, Hosang M, Oefner C, Winkler F, Bradshaw RA. (1993). Topological similarities in TGF-beta2, PDGF-BB and NGF define a superfamily of polipeptide growth factors. *Structure* 1:153-159.
- Murray-Rust, J. (1993). Topological similarities in TGF-beta, 2PDGB-BB and

- NGF define a superfamily of polipeptide growth factors. *Structure*. 1:153-159.
- Murzin AG, Brenner SE, Hubbard T, Chothia C. (1995). SCOP a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247:536-540.
- Narasimhan, L., Singh, J., Humblet, C. & Blundell, T. L. (1994). Snail and spider toxins share a similar tertiary structure and cystine motif. *Nature Struct. Biol.* 1:850-852.
- Nogata K., Kohda, D., Hatanaka, H., Saori, I., Matsuda, S., Yamamoto, T., Suzuki, A. & Inagaki, F. (1994). Solution structure of the epidermal growth factor-like domain of heregulin- β , a ligand for p180erbB-4. *EMBO J.* 13:3517-3523.
- Oliva B, Bates PA, Querol E, Aviles FX, Sternberg MJ (1998). Automated classification of antibody complementarity determining region 3 of the heavy chain (H3) loops into canonical forms and its application to protein structure prediction. *J Mol Biol.* 279:1193-210
- Oliva B, Daura X, Nilsson O, Querol E, Avilés FX, Tapia O. (1995). Structure and fluctuation patterns of potato carboxypeptidase A inhibitor protein in aqueous solution. A molecular dynamics study. *Eur. Biophys. J.* 24:1-12.
- Oliva, B., Daura, X., Querol, E., Avilés, F.X. & Tapia, O. (1995) Structure and atomic fluctuation pattern of Potato Carboxypeptidase A inhibitor. A molecular Dynamic Study in water. *Eur. Biophys, J.* 24:1-11.
- Oliva B, Wastlund M, Cardenas R, Querol E, Avilés FX, Tapia O. (1991). On the stability and fluctuations of the potato carboxypeptidase inhibitor fold: A molecular dynamics study. *Biochem. Biophys. Res. Commun.* 176:616-621.
- Orengo C. A., Michie, A. D., Jones, D. T., Swindells, M. B. & Thornton, J. M. (1997). CATH- A hierarchical classification of protein domain structures. *Structure*. 8:1093-1108.
- Orengo CA, Brown NP, Taylor WT. (1992). Fast structure alignment for protein databank searching. *Proteins* 14:139-167.
- Orengo, C. A. and Taylor, W. R. (1996). SSAP: sequential structure alignment program for protein structure comparison. In: *Computer methods for macromolecular sequence analysis*. Edited by R. F. Doolittle. Orlando, USA, Academic Press, 266:617-635.
- Pascarella, S. & Argos, P. (1992). A data bank merging related protein structures and sequences. *Prot. Eng.* 5:121-137.
- Picot, D., Loll, P.J. & Garavito, R.M. (1994). The X-ray crystal structure of the

- membrane protein prostaglandin H2 synthase-1. *Nature*, 367:243-249.
- Rees DC, Lipscomb WN. (1982). Refined crystal structure of the potato carboxypeptidase inhibitor complex of carboxypeptidase A at 2.5 Å resolution. *J. Mol. Biol.* 160:475-498.
- Richardson, J. S. (1981). The anatomy and taxonomy of protein structure. *Advan. Prot. Chem.* 34:167-339.
- Rossmann MG, Argos P. (1976). Exploring structural homology of proteins. *J. Mol. Biol.* 105:75-95.
- Rost B. (1997) Protein structures sustain evolutionary drift. *Folding Design*, 2: S19-20.
- Rousell, A., & Cambillau, C. (1991). Silicon Graphics directory (Silicon Graphics Mountain View): California, USA.
- Roussel A, Inisan A, Knoops-Mouthy E. (1994). *TURBO FORBO version 5.0a Manual, BIOGRAPHICS*, Marseille, France: Technopole de Chateaux-Gombert.
- Rufino S. D., Donat L. E., Canard L. H. J. & Blundell T. L. (1997) Predicting the conformational class of short and medium size loops connecting regular secondary structures: application to comparative modelling. *J. Mol. Biol.* 267:352-367.
- Rusell, R. B. & Barton, G. J. (1994). Structural features can be unconserved in proteins with similar fold: An analysis of side-chain to side-chain contacts, secondary structure and accessibility. *J Mol. Biol.* 244:332-350.
- Rusell, R. B., Saqi, M. A. S., Sayle, R. A., Bates, P. A. & Sternberg, M. J. E. (1997). Recognition of analogous and homologous protein folds: Analysis of sequence and structure conservation. *J. Mol. Biol.* 269:423-439.
- Sali A, Blundell TL. (1990). Definition of general topological equivalence in protein structures. *J. Mol. Biol.* 212:403-428.
- Srinivasab N, Bax B, Blundell TL, Parker PJ. (1996). Structural aspect of the functional modules in C-alpha deduced from comparative analysis. *Proteins: Struct., Func. and Gen.* 26:217-235.
- Srinivisan N, Sowdhamini R, Ramakrishnan C, Balaram P. (1990). Conformations of disulfide bridges in proteins. *Int. J. Pept. Prot. Res.* 36:147-155.
- Subbarao N, Haneef I. (1991). Defining topological equivalences in macromolecules. *Protein Eng.* 4:877-884.
- Sun PD, Davies DR. (1995). The cystine-knot growth factor superfamily. *Annu. Rev. Biophys. Biomol. Struct.* 24:269-291.

- Taylor WR, Orengo C. (1989). Protein structure alignment. *J. Mol. Biol.* 208:1-22.
- Thornton JM, (1981). Disulphide bridges in globular proteins. *J. Mol. Biol.* 151:261-287.
- Ullner M, Selander M, Persson E, Stenflo J, Drakenberg T, Teleman O. (1992). Three-dimensional structure of the apo form of the N-terminal EGF-Like module of blood coagulation factor X as determined by NMR spectroscopy and simulated folding. *Biochemistry* 31:5974-5983.
- Vita C, Toma F, Menez A. 1995. Scorpion toxins as natural design scaffolds. *Proc. Natl Acad. Sci.* 92:6404-6408.
- Vriend, G. , Sander, C (1991). Detection of common three-dimensional substructures in proteins. *Proteins* 11:52-58.
- Watanabe Y, Barbashov SF, Komatsu S, Hemmings AM, Miyagi M, Tsunasawa S, Hirano H. (1994). A peptide that stimulates phosphorylation of the plant insulin-binding protein. Isolation, primary structure and cDNA cloning. *Eur. J. Biochem.* 224:167-172.
- Wintjens R, Wodak SJ, Rooman M (1998) Typical interaction patterns in alphabeta and betaalpha turn motifs. *Protein Eng.* 11:505-522
- Zuker M, Somorjai R. L. (1989). The alignment of protein structures in three dimensions. *Bull. Math. Biol.* 51:55-78.

Ya está, creo!. Bueno, quienes me conocen saben que ya estaba... casi!!!. Bueno, "un coffee más y ahora sí que lo acabo!!!". Pero ahora sí. AL FIN! Ahora los agradecimientos.

a la familia...

Pero para acabar un trabajo de tesis es preciso haberlo empezado. Una perogrullada similar tiene una gran importancia en mi caso. Poder hacer una tesis necesitas haber hecho una carrera y luego dedicar varios años. En ese tiempo, ¡algo tienes que comer! Y como dice la Lumi: "donde comen tres comen cuatro, si le pones". El esfuerzo personal necesario para poder llegar a presentar este trabajo no es comparable en lo más mínimo al sacrificio que mi familia ha tenido. La tesis de un individuo no es tan sólo el fruto del estrujar de sus meninges, ni de su sacrificio personal (si lo hubiese), como tampoco es el fruto de largos años de trabajo de muchas personas en una misma línea de investigación. Es mucho más que todo eso, al menos bajo mi entender. Para lograr algo, incluso lo más mínimo, se precisa un apoyo, un amor como el que yo he disfrutado a raudales.

Cuando uno repasa el grupo de personas a los que debe todo o parte de su trabajo se da cuenta qué larga puede ser la lista. Encima de esa lista tres personas: mi madre, mi hermano y mi mujer. Sin ellos nada hubiera sido posible. Más aun, sin ellos nada hubiera merecido la pena. Pero agradecerles lo que me dieron no lo puedo hacer sólo con palabras. Necesitaré al menos esta vida para poderles mostrar mi agradecimiento. No puedo dejar de nombrar de un modo especial al verdadero artífice de esta tesis, a la persona que lo hizo posible: mi madre. Desde luego es mi máxima aspiración ser algún día tan buen padre para mis hijos como ella lo es para mí. El apoyo incondicional de mi hermano y su temprana madurez combinado con el cabronismo que alberga su ser me hicieron crecer y crecer, y ya ves ¡'70!!!. Los tres en casa (la Lumi, el Oscar y yo) pasamos momentos muy malos y bajo la supervisión de la superLumi lo sacamos para adelante. Ele nuestros cojones!!! Gracias Lumi.

Creo que no había antes, pero un día conocí a la madre de mis hijos. Su apoyo, su sacrificio paralelo, su aguante y su cariño han sido y son esenciales para mí. Si ella no estuviera obviamente yo no sería yo. Sería otro, igual de feo, pero otro. A ella le debo mucho, mucho... ya le pagaré con carne!!!

Pero hay más. Todo el soporte de una familia cuando pasamos nuestro peor momento es algo que siempre tendremos que agradecer. Ellos nos dieron lo que nos faltaba y nos apoyaron mientras nos hizo falta. La lista es larga y siempre corres el riesgo de dejarte alguien pero no puedo por menos que destacar a Genaro, a Jesús, a Paco, a Antonio, a Pepi, a Mari, a Poli, a Juani, y a mis primos (que no cabrían).

Y en los últimos años mi nueva familia. Maravillosos. Otra madre, un nuevo padre y otra hermana pequeña. A ellos también las gracias de adoptarme como uno más. Conociéndolos a ellos sabes por qué tengo una mujer tan increíblemente maravillosa.

Ya habréis deducido que para mí el trabajo de investigación es el fruto de un estado de ánimo. A vosotros os debo estar ahí. Muchas gracias.

... A la peny del IBF

Cuando llegué al centro encontré unos compañeros, me fui y ya no estaban todos. Algunos eran ya mis amigos. Sabían mucho de proteínas o no tenían ni puta idea ni falta que les hace. Los Patrick, Baldo, Nati, Pepe, Marc y un largo etcétera hicieron que fueran unos años maravillosos. Aprendimos juntos el verdadero arte de la tauromaquia, del autoeslome, espurreamos sin miramiento, comprendimos el concepto cap-cua y disfrutamos de lo lindo con la teoría de las colas.

Pronto fuimos capaces de grandes hazañas. Como no va a haber cojones por parte del resto, voy a ser yo quien deje para la posteridad algunas anécdotas con las que hemos disfrutado de verdad. Sólo voy a explicar dos que guardo en mi mente con un cariño especial y sin ánimo de hacer daño a nadie, va por vosotros!!!!

Tema tauromaquia: Corrían malos tiempos. los revenimientos estaban a la orden del día y los potes y soluciones no se hacían solos. Fuimos al Bar. Entre fenomenales hamburguesas y servestitas frías se coció la tragedia. Volvimos al lugar del crimen. Los potes seguían vacíos y las soluciones sin hacer. Todavía recuerdo las paredes repletas de alusiones. Apuntaban al culpable. No tenía culpa de nada, seguro, pero necesitábamos carnaza. Fue entonces cuando entre vítores asistimos a los mejores pases de pexo. De rodillas. Sin miedo. Muchas tardes de gloria dieron los sin motivos para reír. A ellos y a ese público. Gracias.

Tema espurreo: Mu malos tiempos. El Emperador marcius en el exilio había estado dispuesto a contar las verdades del barquero. La situación había llegado al borde del

abismo. Darth Balder apenas volvía de su encuentro con el mal. Malnutrido y cansado no respondía como humano. Fue entonces cuando llegó la apuesta: Patrick Skywalker, Xouwaka y el Gran Olvidado. El reto estaba servido. Xouwaka se retiró a un grito de la Princesa. Se oyó un "tú quitas y yo pongo". Luego la nada. Una nube nos cubrió a todos. Cuando todo volvió a la normalidad vimos que no había muertos. ¡De poco! Luego el exilio y la condena indefinida. Todavía hoy hay quien sufre las consecuencias, para él, un abrazo.

Quiero hacer una especial mención de Toni, Ángel, Teresa y de Juan Checa, de quienes guardo un gratísimo recuerdo.

... a las aportaciones científicas

Las dos primeras instituciones científicas a las que tengo que agradecer su aportación son mis directores de tesis. Tanto Enrique como Avilés (tanto monta, monta tanto) han puesto a mi servicio su inmenso conocimiento. Hemos emprendido cosas interesantes y nos hemos estrellado en otras, investigación le llaman. Guardaré de ellos un recuerdo en color: el amarillo Avilés y el verde Enrique. También quiero nombrar la aportación de Rafa y su relación entre el cáncer y todo cuanto le rodea.

Me cabe destacar una de las personas con mayores conocimientos en biocomputación en España, mi amigo Baldo. También las aportaciones de mis compañeros y amigos: Patrick, quien seguro acabará teniendo un nombre en el campo, y Marc. Quiero hacer una mención especial para el futuro Dr. Albert Pujol especialmente por su aportación en un capítulo de esta tesis, posiblemente el mejor.

Pero hay personas: Quico, Piñol y Pepo entre otros me han ayudado muy por encima de lo que nuestra convivencia exigía. Quería dar especiales gracias a Pepo quien me dio el toque mágico para hacer la tesis y al que tengo especial estima.

... a INFOCIENCIA

Quiero hacer especial mención de mi compañero de aventuras y amigo Jordi. Con él he comenzado mi '*después del IBF*'. Durante los últimos años hemos creado y compartido nuestra INFOCIENCIA. Hemos partido de nada, hemos hecho algo. Todavía nos queda hacerla muy grande y hundirla estrepitosamente. Con un par de cojones!!!