

OLARTEMOURE

OLARTE MOURE & ASOCIADOS LTDA.

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 11-019241- -00013-0000

Fecha: 2014-11-24 16:25:11 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 523 RESPUESTA Folios: 36

Carlos R. Olarte
Juan G. Moure

Alexander Agudelo
Carlos A. Parra
Juan F. Acosta
Andrés Rincón
Aónica Guevara
Andrés Alvarado
Aria C. Calderón
Mauricio Ávila
José J. Dangond
Gina Cáceres
Isabella Herrera
Zaava Ravid
Alejandro Moure
Miguel F. Porras
Catalina Amaya
Laura M. Escobar
Lorena Andrade
Sergio Arboleda
Catalina Jiménez
Daniela González
David A. Ramírez
Maritza Sierra
Gina A. Arias
Guillermo Parra

Medellín
María C. Múnera
Eddie Manotas
Adrián Santamaría
Marilza P. Ruiz
Camilo A. Bernal

Barranquilla
Gina De Echeona

Señores

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

DIVISIÓN DE NUEVAS CREACIONES

E. S. D.

Re: Expediente N° 11-019.241

Solicitud de Patente de Invención Titulada: "POLIPÉPTIDOS G-CSF
BOVINOS MODIFICADOS"

Solicitante: AMBRX INC. AND ELI LILLY AND COMPANY.

Nuestra Ref.: P2011/000228

RESPUESTA AL OFICIO No. 10828 NOTIFICADO POR LA FIJACIÓN EN LISTA No. 035

DE FECHA 02 DE SEPTIEMBRE DE 2014

(SEGUNDO EXAMEN DE FONDO)

CARLOS R. OLARTE, mayor de edad y vecino de Bogotá, identificado tal como aparece al pie de mi firma, en mi carácter de apoderado de la sociedad solicitante, me permito manifestar lo siguiente ante el auto emanado por ese Despacho, dentro del plazo legal previsto en el artículo 45 de la Decisión 486, en los siguientes términos:

1. En primer lugar, con el ánimo de brindar mayor claridad a la materia reivindicada, me permito presentar un **Capítulo Reivindicatorio Modificado**, que reemplaza en su totalidad al Capítulo Reivindicatorio presentado el 16 de diciembre de 2013, el cual contiene los siguientes cambios:

- se eliminan las Reivindicaciones 2, 3 y 10 a 15;
- en la Reivindicación 1 se especifica que el polipéptido reclamado es un polipéptido pegilado y se limita el peso molecular del poli(etilen)glicol entre 10kDa y 40kDa de acuerdo con la Reivindicación 9 original;
- las nuevas Reivindicaciones 5 y 6, que corresponden a las antiguas Reivindicaciones 7 y 8, ya no mencionan el peso del polímero soluble en agua;

- la nueva Reivindicación 7, que corresponde a la antigua Reivindicación 9, menciona que el polipéptido reclamado induce un conteo máximo absoluto de neutrófilos (ANC) de 7 a 10 veces mayor que el ANC previo al tratamiento de 12 a 96 horas, de acuerdo con la página 269, y las Figuras 16 y 17 del Capítulo Descriptivo;
- la nueva Reivindicación 8, que corresponde a la antigua Reivindicación 16, menciona que la concentración adecuada para la administración del activo va desde 0.01 a 1000µg/kg de peso corporal, así como los posibles vehículos que pueden estar presentes en la composición reclamada de acuerdo con las páginas 305 y 296 del Capítulo Descriptivo, respectivamente.

Valga anotar que la modificación efectuada al Capítulo Reivindicatorio cumple a cabalidad con lo estipulado en el Artículo 34 de la Decisión 486, ya que éstas de ninguna manera implican una ampliación de la protección correspondiente a la divulgación contenida en la solicitud inicial.

2. a- El Examinador afirma que las Reivindicaciones 8 y 9 reclaman materia no patentable, dado que hacen referencia a métodos de tratamiento.

b- Considero superada la anterior objeción, en la medida que las actuales Reivindicaciones 6 y 7 no se refieren a un método de tratamiento sino a un resultado a alcanzar toda vez que dichas reivindicaciones hacen referencia a la capacidad del polipéptido reclamado para incrementar el conteo máximo absoluto de neutrófilos, así, me permito señalar que el Manual Andino de Patentes (MAP)¹ y la Guía para Examen de Patentes de la SIC² establecen que un parámetro y un resultado a alcanzar no es una característica técnica de la invención, **pero puede aparecer en la reivindicación acompañado de las características técnicas que definen la invención**. En este sentido, las características técnicas esenciales del polipéptido a las que se hace referencia en las Reivindicaciones dependientes 6 y 7 están debidamente definidas en las Reivindicación independiente 1 y, evidentemente, al ser dichas reivindicaciones completamente dependientes de la Reivindicación 2, que a su vez es dependiente de la Reivindicación 1, es claro que las características reclamadas por las actuales Reivindicaciones 6 y 7 corresponden y obedecen únicamente a una limitación particular que, por definición, se encuentra *per se* dentro del alcance de la reivindicación independiente que caracteriza completamente a la presente invención, como lo demostraré más adelante.

¹ Manual Andino de Patentes. (2004). Sección 4.6.6 (disponible en <http://www.comunidadandina.org/public/patentes.pdf>).

² Guía para Examen de Solicitudes de Patentes de Invención y Modelo de Utilidad. Superintendencia de Industria y Comercio (2014). Sección 2.6.5.8 Reivindicaciones definidas por el resultado a alcanzar.

3. a- El Examinador presenta las siguientes observaciones frente a las reivindicaciones:
- la Reivindicación 1 no es clara debido a que un polipéptido debe estar caracterizado por su secuencia de aminoácidos, adicionalmente el Examinador solicita que la Reivindicación 1 comience con la expresión “un polipéptido pegilado” con el fin de describir de una manera más precisa el objeto reclamado en la presente solicitud;
 - las Reivindicaciones 2 a 6 deben definir de manera específica el peso del polímero soluble en agua, a su vez se debe eliminar el término “aproximadamente” ya que es un término que no permite diferenciar a la presente invención del estado de la técnica con precisión;
 - la Reivindicación 7 reclama un resultado a alcanzar;
 - la Reivindicación 10 no es concisa al mencionar “un ácido nucleico”, porque existen muchas secuencias posibles que puedan contener los codones que codifican a los aminoácidos que permiten obtener los polipéptidos de la presente solicitud, así, el Examinador solicita que se mencionen las secuencias específicas que permiten obtener los polipéptidos de la presente invención;
 - las Reivindicaciones 10-11 y 13-14 no son concisas porque al mencionar “un vector” y “una célula hospedera” respectivamente existen muchas posibilidades que cumplen con estas funciones, ante esto, el Examinador solicita que se mencionen de manera específica los vectores y las células hospederas empleadas para obtener a los polipéptidos de la presente invención;
 - la Reivindicación 15 o es clara en vista que un proceso para obtención de polipéptidos se debe describir en función de los medios para obtener el gen e identificar su origen, de la secuencia nucleotídica empleada, los medios para obtener los vectores y el huésped, los medios y métodos para la purificación e identificación del polipéptido y del proceso para la pegilación del polipéptido, por lo cual se debe describir de manera completa el proceso de obtención de polipéptido pegilado de la presente invención;
 - se deben mencionar de manera específica las cantidades y/o proporciones porcentuales del polipéptido pegilado y de los excipientes empleados que describen de forma específica a las composiciones que contienen a los polipéptidos reclamados en la presente solicitud.

Adicionalmente, el Examinador indica que en las Figuras 15, 16, 17 y 18 se debe mencionar el significado de las siglas ANC, MFI, LBS, así como las unidades en las que se expresan los valores. De igual manera, para las Tablas 6 y 7, los compuestos Neupogen y Neulastim corresponden a los

nombres comerciales de los compuestos que tienen su DCI como Filogastrim (Neupogen) y Filgastrim pegilado (Neulastim), por lo que solicita cambiar los nombres comerciales por su respectiva DCI.

b- En respuesta a las observaciones del Examinador me permito considerar superadas las objeciones relacionadas con las Reivindicaciones 10 a 15 en la medida que las mismas fueron eliminadas del Capítulo Reivindicatorio Modificado adjunto, adicionalmente me permito señalar lo siguiente:

i) La Reivindicación 1, de acuerdo con la amable sugerencia del Examinador, ahora reclama un polipéptido pegilado; adicionalmente, me permito llamar la atención del Examinador acerca de la mención específica de las secuencias de aminoácidos, así como las posiciones de las mismas donde se realizan las modificaciones reemplazando el aminoácido natural por para-acetilfenilalanina (posición 133 de la SEQ ID NO: 1 o la posición 134 de la SEQ ID NO: 2), con lo cual es evidente que para cualquier persona medianamente versada en la materia resulta comprensible el alcance de la presente solicitud, toda vez que el polipéptido está plenamente definido mediante sus características estructurales esenciales (secuencia de aminoácidos, posición de la secuencia donde se hace la sustitución y el peso molecular del polímero soluble en agua unido al polipéptido).

ii) Respecto a la claridad del Capítulo Reivindicatorio, frente a la objeción del Examinador en referencia a la expresión “aproximadamente” de las actuales Reivindicaciones 2 a 4, me permito indicar que esta expresión fue eliminada del Capítulo Reivindicatorio Modificado. No obstante, y sin perjuicio de lo anterior, quisiera resaltar respetuosamente que un inventor está en su derecho de reclamar protección para la totalidad de un concepto inventivo desarrollado, sin importar la amplitud de éste, siempre y cuando se conserve el principio de unidad de invención y no haya arte previo que afecte su patentabilidad. De esta manera, una reivindicación puede ser tan amplia como se requiera para así poder enmarcar la totalidad del concepto inventivo desarrollado, y no solamente las modalidades preferidas o específicamente ejemplificadas en la solicitud.

En el presente caso, la invención aquí reclamada consiste en un polipéptido que contiene un aminoácido no natural en una determinada posición de su estructura junto con un polímero soluble en agua, en donde la expresión “aproximadamente” hace referencia a un valor relativo que puede tener el peso molecular del polímero asociado al polipéptido, entendiéndose que el valor exacto del peso de un polímero es difícil de determinar y que cualquier polímero con un peso molecular

cercano al que se menciona en dichas reivindicaciones va a ser funcional de acuerdo a la reclamación de la presente solicitud.

iii) En cuanto a la actual Reivindicación 5 (antigua Reivindicación 7) me permito recordarle al Examinador que el Manual Andino de Patentes (MAP)³ y la Guía para Examen de Patentes de la SIC⁴ establecen que un parámetro y un resultado a alcanzar no es una característica técnica de la invención, pero puede aparecer en la reivindicación acompañado de las características técnicas que definen la invención. En este sentido, las características técnicas esenciales del polipéptido a las que se hace referencia en dicha reivindicación están debidamente definidas en la reivindicación independiente 1, al ser la Reivindicación 5 completamente dependiente de la Reivindicación 2, que a su vez es dependiente de la Reivindicación 1, es claro que las características reclamadas por la actual Reivindicación 5 corresponden y obedecen únicamente a una limitación particular que, por definición, se encuentra *per se* dentro del alcance de la reivindicación independiente que caracteriza completamente a la presente invención, lo cual es plenamente admitido a la luz del Manual Andino de Patentes (MAP) y la Guía para Examen de Patentes de la SIC.

iv) Finalmente, en cuanto a la actual Reivindicación 8 (la antigua Reivindicación 16), en cuanto a las definiciones de las proporciones de los excipientes y demás componentes de la composición reclamada, no puede olvidarse que un inventor está en su derecho de reclamar protección para la totalidad de un concepto inventivo desarrollado, sin importar la amplitud de éste, siempre y cuando se conserve el principio de unidad de invención y no haya arte previo que afecte su patentabilidad. Es decir, un Capítulo Reivindicatorio puede ser tan amplio como se requiera para así poder enmarcar la totalidad del concepto inventivo desarrollado y todas sus posibles modalidades, más no solamente las modalidades preferidas o específicamente ejemplificadas en la solicitud. Incluir elementos y condiciones específicas no esenciales en las reivindicaciones, como lo sugiere el Examinador limitaría injustamente el derecho del inventor a reclamar el alcance del concepto inventivo desarrollado. Así el incluir elementos no esenciales, el alcance de la protección se ve limitado en tanto que, para que existiera una infracción, dichos elementos no esenciales deberían necesariamente estar presentes en el producto o proceso infractor, lo que fácilmente podría lograrse al tratarse de elementos que pueden ser cambiados o modificados. Por tanto, incluir características adicionales que no son esenciales limitaría injustificadamente la protección del concepto inventivo desarrollado. Asimismo, y como lo reconoce el Despacho, la

³ Manual Andino de Patentes. (2004). Sección 4.6.6 (disponible en <http://www.comunidadandina.org/public/patentes.pdf>).

⁴ Guía para Examen de Solicitudes de Patentes de Invención y Modelo de Utilidad. Superintendencia de Industria y Comercio (2014). Sección 2.6.5.8 Reivindicaciones definidas por el resultado a alcanzar.

Guía de la SIC establece que “Si el objeto de la solicitud es un nuevo compuesto, el solicitante tiene derecho a reivindicar las composiciones farmacéuticas caracterizadas por contener ese nuevo compuesto, pero no está obligado a utilizar ejemplos de cómo diseñar y preparar composiciones concretas que contengan el nuevo compuesto”⁵ (énfasis fuera del texto), por lo que la materia reclamada en la actual Reivindicación 8 (antigua Reivindicación 16) es a todas luces patentable.

v) Adicionalmente, me permito presentar las hojas sustitutas de las página 329 del Capítulo Descriptivo donde se aprecian los nombres según las DCI de los activos, solicitada por el Examinador, así como de las Figuras 5 a 18 en donde se especifica el significado de las siglas empleadas, así, ANC (conteo absoluto de nucleótidos), MFI (Intensidad media de la fluorescencia) y LBS (libras de leche producidas/día).

En conclusión, considero superadas las anteriores objeciones, en tanto que el Capítulo Reivindicatorio Modificado adjunto se encuentra redactado de manera tal que cumple con los requisitos de claridad establecidos en la Decisión 486 y la Guía de la SIC.

4. a- El Examinador argumenta que las patentes US 2008/0146781 (D1) junto con US2005/0142102 (D2), afectan el nivel inventivo de la presente solicitud.

El Examinador señala que D1 divulga polipéptidos modificados por lo menos con un aminoácido no natural, unido a un polímero funcional como polietilenglicol. A su vez, D2 enseña divulga una secuencia correspondiente a un polipéptido estimulador de G-CSF. El siguiente cuadro comparativo presenta las características esenciales de la Reivindicación 1 de la presente invención frente a las del arte previo citado:

Solicitud	D1	D2
Un polipéptido hG – CSF	Un polipéptido hG – CSF	Un polipéptido G – CSF proveniente de <i>Bos taurus</i>
Corresponde a las secuencias SEQ ID No 1 y SEQ ID No 2	Corresponde a la secuencia SEQ ID No 29 (Pág. 119 y 120)	Corresponde a la secuencia SEQ ID No 32 (Pág. 65)
La posición 133 de SEQ ID N° 1 o la posición 134 de SEQ ID N° 2 (correspondientes a Thr) esta sustituido por el aminoácido p – acetilfenilalanina	El polipéptido puede sustituirse en la posición (la posición 134 corresponde a Thr) (Pág. 8, párr. [0053]) esta sustituido por un aminoácido no natural de la Fórmula: $\begin{array}{c} \text{(CH}_2\text{)}_n\text{R}_1\text{COR}_2 \\ \\ \text{R}_3\text{HN} - \text{C} - \text{COR}_4 \end{array}$ Donde R ₁ es arilo, n es 1. R ₂ y R ₃ son H (Pág. 9, párr. [0061])	No se menciona
El polipéptido se une a un polímero soluble en agua que comprende una fracción de polietilenglicol	El polipéptido se une a un polímero soluble en agua bifuncional que comprende polietilenglicol que pesa entre 0.1 kDa – 50kDa (Pág. 6, párr. [0041] – [0042]; pág. 11, párr. [0082] – [0083]; pág. 62, párr. [0460])	No se menciona

De acuerdo con lo anterior, el Examinador sostiene que D1 conforma el estado de la técnica más cercano a la Reivindicación 1, debido a que enseña divulga un polipéptido estimulante de Factor de

⁵ Superintendencia de Industria y Comercio, *op. cit.*, 2014, sección 2.7.3, página 94

Crecimiento de Colonias de granulocitos (G-CSF) de la SEQ ID N° 29 que puede ser modificado por sustitución o inserción de aminoácidos no naturales, donde este polipéptido se une a polímeros hidrosolubles como polietilenglicol, donde la sustitución se puede llevar a cabo en la Thr de la posición 134, donde el dominio en el cual se encuentra este aminoácido se describe por la secuencia APALQPTIQGAMP.

El Examinador además concluye que la presente invención se diferencia de D1 en que el polipéptido de la presente solicitud es un polipéptido G-CSF de origen bovino de secuencia APAVCQPTIQGAMP (teniendo la posición 134), mientras que el polipéptido divulgado en el documento D1 es un polipéptido C-CSF de origen humano, teniendo ambos péptidos un porcentaje de homología del 81%, lo que permite al polipéptido reclamado en la presente solicitud, incrementar la cantidad de neutrófilos en un bovino, cifra que se mantiene durante más de 72 horas.

A su vez, el Examinador indica que la sustitución en la posición 134 por p-acetil fenilalanina no modifica la actividad de manera significativa con respecto a la proteína nativa, mientras que la pegilación de la proteína nativa de la solicitud también disminuye su actividad, como es evidente en la página 357, tablas 7 y 8 de D1.

El Examinador finalmente establece que el problema técnico que resuelve la solicitud es ¿Cómo modificar el polipéptido divulgado en el documento D1 para obtener un polipéptido pegilado alternativo para el tratamiento de infecciones en bovinos? Para el Examinador, la persona medianamente versada en la materia habría utilizado las enseñanzas de D2 (polipéptido correspondiente a la secuencia SEQ ID N° 32, el cual procede de *Bos Taurus* de origen bovino, que contiene el dominio de la secuencia APAVCQPTIQGAMP (I es tirosina en la posición 154) para modificar el polipéptido de la SEQ ID No. 1 divulgado por D1, obteniendo así el objeto de la Reivindicación 1, por lo cual la presente solicitud no tiene nivel inventivo.

b- i) En respuesta a las observaciones planteadas por el Examinador, me permito señalar que D1 es una divulgación muy general que da a conocer un género de proteínas que comprende un gran número de variables posibles del péptido humano G-CSF. No obstante, dicho documento no demuestra o divulga de manera explícita un ejemplo funcional, toda vez que D1 describe métodos y procedimientos que permitan estudiar dichos péptidos en el futuro (respecto a la fecha de divulgación de D1) para producir variables del péptido humano G-CSF que sean funcionales y que presenten “una actividad superior y una vida media prolongada”.

Así, las enseñanzas y los datos proporcionados por D1 no son suficientes para proporcionarle a una persona medianamente versada en la materia una motivación tal que le permita desarrollar un péptido alternativo al péptido humano G-CSF que fuese biológicamente activo, lo cual ni siquiera hubiese llevado a dicho experto a extrapolar el contenido de D1 para desarrollar variables de polipéptidos que tuviesen actividad en organismos de otras especies (como lo hace la presente solicitud) sin probabilidades de éxito.

Sumado a lo anterior, D1 no llevaría a una persona medianamente versada en la materia a desarrollar un polipéptido G-CSF que sea activo en bovinos que comprendiera una para-acetilfenilalanina sustituida en la posición 133 de la SEQ ID No: 1 o en la posición 134 de la SEQ ID No: 2, por un polímero soluble en agua que comprenda una fracción de poli(etilen)glicol, que consiste en la materia reclamada de manera específica por la presente solicitud.

ii) Por su parte D2 no motivaría a dicha persona hacia la obtención de la presente invención debido a que este documento se refiere a un método de tratamiento de una condición neurológica mediante la administración de un factor de crecimiento hematopoyético (véase el resumen de D2). Si bien D2 menciona polipéptidos para varias especies (incluidos los bovinos), hay una marcada preferencia hacia los polipéptidos humanos G-CSF como se puede apreciar en los documentos 1, 4 y 6 de D2.

En su divulgación D2 menciona variables puntuales que permiten modificar los polipéptidos B-CSF, a saber:

- a) reemplazar la cisteína de los residuos 17, 36, 42, 64 o 74 con serina;
- b) reemplazar el residuo número 1 con alanina;
- c) reemplazar el ácido aspártico del residuo 27 con serina;
- d) sustituir los residuos 50-56 con la secuencia Leu-Gly-His-Ser-Leu-Gli-Ile; y
- e) mutar la histidina de las posiciones 43, 79 156 o 170.

En este orden de ideas, es evidente que D2 no instaría a una persona medianamente versada en la materia a aislar el polipéptido bovino G-CSF y sustituirlo con un aminoácido no natural en una posición determinada y conjugarlo con un polímero soluble en agua sobre dicho aminoácido no natural para tratar la mastitis en bovinos, dado que la motivación real de D2 es buscar polipéptidos activos para tratar condiciones neurológicas, lo cual no se relaciona con el objetivo de la presente invención.

9

iii) Adicionalmente y en vista que las afirmaciones del Examinador para desestimar el nivel inventivo de la presente solicitud no son suficientes para establecer un juicio certero, me permito a continuación presentar las siguientes afirmaciones que respaldan la no obviedad de la invención que se reclama por medio de la presente solicitud:

La premisa del Examinador carece de fundamento

La conclusión del Examinador en la que establece que la presente solicitud no tiene nivel inventivo parte de las similitudes estructurales que existen entre los polipéptidos G-CSF humano y bovino. No obstante, estas semejanzas no dan soporte a las aseveraciones del Examinador en la medida que no demuestran con claridad de qué manera la presente invención es obvia frente a D1 o D2.

En concordancia con lo anterior, me permito mencionar que el documento WO2007/011166⁶ (en adelante el artículo de Kang) da a conocer mutantes del polipéptido humano G-CSF en donde se sustituyen residuos de posiciones específicas con cisteína, la cual está conjugada a un polímero biocompatible como el PEG. Así el artículo de Kang menciona que los aminoácidos próximos al residuo pegilado en el polipéptido pueden crear un “microentorno” que puede causar diferencias considerables mediante el impedimento estérico generado por las cadenas laterales (véase el párrafo 41 del artículo de Kang). Esto es debido a que la configuración tridimensional que tienen las proteínas se ve afectada de manera directa por las posiciones específicas que tienen los aminoácidos dentro de la secuencia de la proteína, por lo que se puede afirmar que los aminoácidos que se encuentran alrededor de la tirosina en la posición 133 de la secuencia del polipéptido bovino G-CSF crean un microentorno particular para dicho polipéptido, haciéndolo completamente distinto de otros polipéptidos G-CSF de otras especies.

Más aún, a pesar que se conoce la estructura cristalina de la proteína bovina G-CSF, la configuración específica del bucle CD (región del polipéptido donde se encuentra la tirosina de la posición 133) no se conoce de manera precisa, tal y como lo señala el artículo “Crystal Structure of Canine and Bovine Granulocyte-Colony Stimulating Factor (G-CSF)”⁷ en la página 645 (archivo adjunto). Este artículo demuestra que no existe un modelo de la estructura cristalina de los residuos 131-136 del bucle CD debido a la gran movilidad del mismo, haciendo que esta región se represente como un espacio en las Figuras 1 y 3 de las páginas 641 y 645 del artículo citado.

⁶ Solicitud internacional de patente WO2007/011166 disponible en el sitio web:
http://worldwide.espacenet.com/publicationDetails/originalDocument?CC=WO&NR=2007011166A1&KC=A1&FT=D&ND=3&date=20070125&DB=worldwide.espacenet.com&locale=en_EP

⁷ Lovejoy Brett; Cascio Duilio & Eisenberg David; Journal of Molecular Biology (1993) 234; 640-653.

En vista de lo anterior, la afirmación del Examinador se basa en una suposición errada según la cual los polipéptidos G-CSF humano y bovino son iguales en la totalidad de su estructura 3D y en las proximidades del residuo tirosina en la posición 133, toda vez que el modelo estructural de esta región no se conocía al momento en el que se obtuvo a la presente invención, por lo cual la hipótesis del Examinador de acuerdo con la cual una persona medianamente versada en la materia partiría de una información inexistente para pretender modificar un polipéptido existente en una región desconocida esperando tener éxito en dichas modificaciones resulta sin sentido.

El arte previo demuestra no ser predictivo

El arte previo resulta ser insuficiente para enseñar la preparación y la expresión de las variables del polipéptido G-CSF, a su conjugación con PEG y a que dichas variables conjugadas resulten ser activas tanto *in vivo* como *in vitro*. De manera particular, la persona medianamente versada en la materia no habría esperado una actividad *in vivo* superior después de haber reducido la actividad *in vitro* del mismo cuando se comparó con el tipo silvestre del polipéptido G-CSF. Específicamente, el artículo de Kang da a conocer los mutantes puntuales del polipéptido humano G-CSF conjugados con un polímero biocompatible como el PEG, en donde la actividad biológica y la estabilidad de polipéptido mutado no se ven afectadas y los cambios estructurales en la conjugación resultan ser mínimos. Lo anterior debido a que los mutantes contienen un residuo de cisteína sustituido por un residuo de tipo silvestre o contienen una cisteína insertada en la secuencia del polipéptido humano G-CSF. Así, el artículo de Kang sugiere el uso de una cisteína para obtener un polipéptido que se pueda conjugar fácilmente con el PEG (Párrafo 44) que demostraron tener una vida media incrementada *in vivo* sin reducir su actividad biológica que se diseñaron para el tratamiento del cáncer (párrafos 24, 49 y 132 de D3).

A pesar de lo anterior, las mutaciones planteadas en el artículo de Kang deben considerarse como mutaciones conservativas en la medida que la mutación de la treonina de la posición 134 por una cisteína en el polipéptido humano G-CSF (el cual tiene una metionina en su extremo N-terminal) que luego se modifica mediante la conjugación con PEG, sumado al hecho que los aminoácidos cisteína y treonina son aminoácidos polares con cadenas laterales similares, evidencian que en realidad no se puede esperar una actividad diferente, ya que la función y la estructura de los aminoácidos comparados resultar ser prácticamente idéntica.

Adicionalmente, el artículo de Kang no menciona que se pueda hacer una modificación de un aminoácido natural por uno que no se encuentre en la naturaleza tal como lo para-acetilfenilalanina, así como no proporciona datos de actividad de una variable pegilada, ya que sólo muestra resultados

para polipéptidos con una mutación sencilla en la posición 133, haciendo que la presente invención resulte impredecible frente a las enseñanzas de este artículo o de los documentos citados por el Examinador.

La expresión y posterior pegilación de las variables de G-CSF no es predecible

Ninguna de las referencias citadas por el Examinador divulga polipéptidos mutantes del polipéptido bovino G-CSF. De hecho, es el artículo de Kang el que muestra una mutación puntual en la posición 133 en el polipéptido humano G-CSF el cual se encuentra pegilado con un PEG que pesa 20kDa.

Los datos dados a conocer por el artículo de Kang demuestran que la expresión y pegilación de las variantes de los polipéptidos son impredecibles, debido a que el bucle CD (posiciones 126 a 143 en el polipéptido humano G-CSF que tiene una metionina en el extremo N-terminal) es flexible, por lo cual no ejerce un efecto considerable sobre la configuración general del polipéptido, dado que se espera que frente a una pegilación dicha región se vea altamente humectada (véase el párrafo 40 del artículo de Kang), lo cual constituye una motivación para los autores del artículo a desarrollar mutantes de G-CSF que posean una cisteína sustituida en la región del bucle.

En concordancia, en el artículo de Kang se muestran las siguientes mutaciones en la región del bucle CD (posiciones 126 a 143 en el polipéptido humano G-CSF que tiene una metionina en el extremo N-terminal): Ala128, Ala130, Leu131, Gln132, Thr134, Gln135, Gly136, Ala137, Met138, Ala140 y Ala142 (Ejemplos 1 a 11), donde estos aminoácidos son sustituidos de manera conservativa por una cisteína en función del tamaño de la cadena lateral respecto a los aminoácidos Gln, Thr y Gly, debido a que estos aminoácidos son aminoácidos polares. A su vez, el artículo de Kang da a conocer la forma de caracterizar estas 11 variantes incluyendo su expresión, capacidad de conjugarse con PEG sobre la cisteína, y actividad proliferativa celular *in vitro* usando el ensayo M-NFS60.

Los datos demuestran que los mutantes Ala128, Ala140 y Ala142 no pudieron recuperarse siquiera, mientras que los mutantes Ala130 y Leu131 se pegilaron de una manera deficiente y otras dos mutaciones (Gly136 y Met138) no pudieron pegilarse, lo cual demuestra de manera clara que la expresión de los polipéptidos mutantes del péptido humano G-CSF no puede ser predicha, lo cual no hubiese motivado a una persona con conocimientos medios en la materia a predecir en caso de extrapolar la información contenida en el artículo de Kang a otras especies, a esperar que dichas

variantes fuesen efectivas o que se expresaran, fueran purificadas y pegiladas dada la divulgación del artículo de Kang.

La actividad biológica de la variante pegilada de G-CSF es impredecible

Los datos dados a conocer por el artículo de Kang demuestran adicionalmente que la actividad de las variables pegiladas del polipéptido G-CSF es impredecible. Kang sugiere que los mutantes Ala130, Leu 131, Gln132, Thr134, Gln135 y Ala137 pueden ser expresados y pegilados en cierta medida debido a que en el ensayo M-NSF60 5 de estas 6 variables mostraron una reducción considerable en la actividad proliferativa al ser comparados con Filogastim, sólo la variante Thr134 mostró un aumento en la actividad proliferativa *in vivo*. (Véanse los ejemplos 2-6 y 8 en la Tabla 5).

Con lo anterior, queda demostrado que el efecto de la mutación y posterior pegilación de la proteína es impredecible, con lo cual la divulgación de este artículo no motivaría a una persona medianamente versada en la materia a esperar que las variaciones del polipéptido G-CSF en diferentes especies (menos conocidas que el polipéptido humano) pudiesen ser pegiladas y efectivas teniendo actividad antiproliferativa *in vivo* debido a que cada mutación puntual debe ser estudiada rigurosamente en vista que no hay forma alguna que permita predecir los resultados o la actividad de las mismas.

No hay una propiedad descada en el desarrollo de las variantes del polipéptido

Debido a que existe una homología tan solo del 81% entre el polipéptido bovino G-CSF y el humano, la persona medianamente versada en la materia no hubiese esperado que las variantes desarrolladas sobre el polipéptido bovino resultasen activas, toda vez que los cambios puntuales afectaban la actividad de las variantes, permitiéndole al experto en la materia, por el contrario, estimar que dadas las diferencias existentes, los resultados serían impredecibles, por lo cual no habría una motivación fehaciente que le guiara a desarrollar una variante mutada como la que se reclama en la presente solicitud.

Además, los datos de actividad biológica *in vitro*, para cinco variantes 20K PEGiladas de la presente solicitud se presentan en la Tabla 9 del Capítulo Descriptivo. Para mayor conveniencia, parte de dicho cuadro se reproduce a continuación. Los datos reportados para 20kPEG - BT133 se muestra como el promedio de los dos valores en la Tabla 9 original. La EC₅₀ con relación a la proteína silvestre (WT), fue calculada dividiendo el valor de EC₅₀ WT por el valor de la EC₅₀ de cada mutante se agrega a la tabla.

TABLA 9		
	Promedio de valores EC ₅₀	EC ₅₀ Relativa
WT bG-CSF	0.053	1.0
20kPEG-bS62	0.150	0.353
<u>20kPEG-bT133</u>	<u>0.264</u>	<u>0.201</u>
20kPEG-bR7	0.348	0.152
20kPEG-bR166	0.363	0.146
20kPEG-bL3	0.477	0.111

Los datos incluyen las variantes bG –CSF S62, T133 y R166, de las cuales se esperaba que estas posiciones sean las que presenten mayor tolerancia a la sustitución y a la conjugación con una macromolécula, sin afectar negativamente la actividad biológica. Los resultados muestran que las cinco variantes bG -CSF PEGilado conservan la actividad biológica. **Sin embargo todas las variantes muestran una actividad reducida comparadas con el tipo silvestre.** Cabe destacar que la variante que tiene la T133 sustituida con para-acetilfenilalanina y pegilada fue la segunda variante que menos sufrió la reducción de la actividad.

Lo anterior permite determinar que en vista que el efecto no es en realidad el esperado, las materia reclamada no es obvia frente al arte previo dado que los resultados obtenidos no coinciden con lo que se habría esperado a partir de las referencias citadas por el Examinador y del artículo de Kang.

Resultados Inesperados

El artículo de Kang muestra que el mutante pegilado Thr134Cys hG-CSF tiene mejores propiedades cuando se compara con el péptido de tipo silvestre hG-CSF frente a la actividad in vitro, vida media in vivo y activación de neutrófilos (véanse los párrafos 49, 50, 125 y 129 así como las figuras 4 y 5), lo que motivaría a una persona medianamente versada en la materia a esperar que un mutante con una actividad in vitro disminuida también tuviese una vida media disminuida y una menor capacidad para activar neutrófilos, no obstante, el polipéptido bG-CSF T133pAF-20K PEG manifiesta una vida media in vivo superior y una mayor capacidad para activar neutrófilos que la forma silvestre del polipéptido bG-CSF.

Brevemente, la presente solicitud en el Ejemplo 28 compara la eficacia de la variante bG-CSF T133pAF-20K PEG frente a bG-CSF T133pAF en vacas periparturientas, la variable no pegilada fue administrada 14 veces, mientras que la variable pegilada solo fue administrada 2 veces, al día 28 se recogieron los datos que se muestran en la Tabla 10 del Capítulo Descriptivo a continuación:

TABLA 10		
Descripción de Tratamiento	Morbilidad	Mortalidad
Salina Estéril (SIDx2 Día -7 y Día 0)	26/50 (52%)	4
bG-CSF T133-pAF (3 µg/kg, SIDx14, Día-7 – Día 6)	16/49 (33%)	6
bG-CSF T133-pAF 20K PEG (40 µg/kg, SIDx2, Día -7 y Día 0)	6/53 (11%)	2
bG-CSF T133-pAF 20K PEG (20 µg/kg, SIDx2, Día -7 y Día 0)	7/52 (13%)	3

Donde se puede apreciar que la administración de la variable pegilada logra disminuir la morbilidad de la mastitis al 11% (frente al 33% de la variable no pegilada y el 52% del grupo control), con solo dos dosis frente a las 14 que se administraron de la variable no pegilada, demostrando que la variable pegilada del polipéptido tiene una vida media *in vivo* incrementada y una mayor capacidad de activar neutrófilos, lo cual no pudo haber sido predicho a partir de las divulgaciones del arte previo o del artículo de Kang. Lo anterior demuestra, de manera contraria a lo afirmado por el Examinador, que la presente invención es en efecto inventiva dado que el polipéptido aquí reclamado posee un efecto sorprendente e inesperado frente a lo que evidencia el arte previo.

iv) En adición a todas las anteriores consideraciones, el Examinador no proporciona argumentos técnicos que efectivamente desvirtúen el nivel inventivo del polipéptido reclamado, pues no explica cómo a partir de D1 o D2 una persona medianamente versada en la materia podría desarrollar al polipéptido de la presente solicitud de manera obvia. De hecho, en el análisis del Examinador, la persona medianamente versada en la materia NO se vería motivada por D2 a modificar completamente los péptidos de D1, para así llegar a la materia reclamada por la presente solicitud.

El punto a determinar en el análisis de nivel inventivo es si las modificaciones realizadas son motivadas por el arte previo, y por lo tanto obvias, o no. En el presente caso, no es claro cómo la mutación del polipéptido de D2 sería obvia para la persona medianamente versada en la materia con el fin de llegar a la materia reclamada. Es más, la persona medianamente versada en la materia no se vería motivada a combinar los documentos citados, y en caso de que lo hiciera, inevitablemente

obtendría polipéptidos que tuvieran propiedades diametralmente diferentes a las del polipéptido que se reclama en la presente solicitud.

Así las cosas, y a la luz de las evidentes diferencias encontradas entre la materia reclamada en la solicitud frente al arte previo, vale la pena señalar que el Manual Andino de Patentes (MAP) establece que para que pueda negarse la existencia de nivel inventivo, es necesario, no solamente que la combinación de las enseñanzas pueda hacerse, sino también que exista una sugerencia o razón tal, que lleve al técnico con conocimientos medios a combinar las enseñanzas de los documentos⁸. Además, el MAP también establece que las invenciones en el campo de la química tienen nivel inventivo cuando los compuestos tienen una **estructura inesperada** (no derivable por la persona medianamente versada en la materia) o cuando presentan un **efecto técnico inesperado o mejorado impredecible del arte previo**.

Finalmente, es importante destacar que el análisis del nivel inventivo no puede basarse en si la invención era posible, pues claramente es posible, ya que el polipéptido aquí reclamado ahora existe. Aplicar un estándar de "posibilidad de alcanzar" en la evaluación de patentes, haría que ninguna invención fuera susceptible de patentarse, haciendo inviable el sistema. De hecho, no puede olvidarse que el análisis patentabilidad establecido en el Manual Andino de Patentes, implica que: *"...la búsqueda de anterioridades se efectúa a posteriori, tomando como punto de partida la misma invención, ...se debe realizar el esfuerzo intelectual de colocarse en la situación que ha tenido que afrontar el técnico con conocimientos medios en la materia en un momento en que la invención no era conocida, es decir antes de la invención"*⁹ (mi subrayado). Por ende, no es apropiado anticiparse a los resultados presentados por la invención, ya que así, siempre resultará "obvio" todo objeto de una solicitud de patente, pues inevitablemente está parecerá derivada de la teoría en que se basa su desarrollo, sin dar mayor importancia al trabajo, intelecto y creatividad involucrados en la invención para llegar a la solución del problema planteado en la solicitud.

En conclusión, y de acuerdo con el peso de la evidencia proporcionada, la presente invención es inventiva a la luz del arte previo citado, pues siguiendo las enseñanzas de D1 y D2, la persona medianamente versada en la materia no obtendría el polipéptido reclamado, ni podría esperar que al realizar las múltiples modificaciones estructurales para obtenerlo la capacidad de activar neutrófilos y la vida media del mismo se verían mejoradas debido a las modificaciones introducidas.

⁸ Manual Andino de Patentes. (2004). Sección 10.2.1 Ejemplo 1 (disponible en).

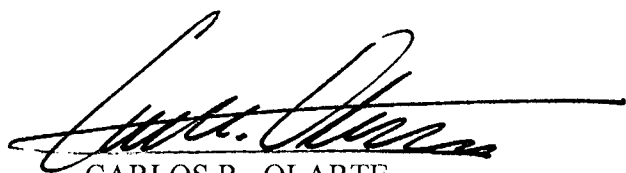
⁹ Manual Andino de Patentes. (2004). Sección 10.2.1 (disponible en <http://www.comunidadandina.org/public/patentes.pdf>).

5. Con un propósito meramente ilustrativo, me permito destacar que esta solicitud de patente se encuentra concedida en la oficina de patentes de Australia (AU2009274076), Eurasia (EA019968B1), México, Nueva Zelanda (NZ591235 y divisional NZ600382), Panamá (PA88368-01), Ucrania (UA103774) y Sudáfrica (ZA2011/01166).

Si bien las concesiones de patentes en oficinas de otros países no son vinculantes en el sentido que obliguen a la SIC a tomar una decisión en el mismo sentido, es a la vez cierto que los criterios de patentabilidad en los que deben fundamentarse el análisis de solicitudes de patente, son análogos globalmente, y concesiones en otras jurisdicciones sugieren el cumplimiento de dichos requisitos, ya que claramente, los criterios de evaluación y estudio aplicados para determinar la novedad y nivel inventivo resultan equivalentes a los que la Superintendencia debe emplear, establecidos de manera semejante por las legislaciones de patentes a lo largo del mundo entero.

6. Por todo lo anterior, considero que las objeciones presentadas por el Examinador en su primer examen de fondo han sido resueltas y por consiguiente solicito respetuosamente que el Despacho tenga por cumplidos los requisitos de patentabilidad y proceda con la concesión de la patente.

Atentamente,



CARLOS R. OLARTE

C.C. No. 79.782.747 de Bogotá

T.P. 74.295

Anexos: Capítulo Reivindicatorio Modificado;

Hoja sustituta Página 329;

Versión corregida Figuras 15-18

ANEXO 1: Artículo "Crystal Structure of Canine and Bovine Granulocyte-Colony Stimulating Factor (G-CSF)";

CAPÍTULO REIVINDICATORIO MODIFICADO

(Segundo Examen de Fondo)

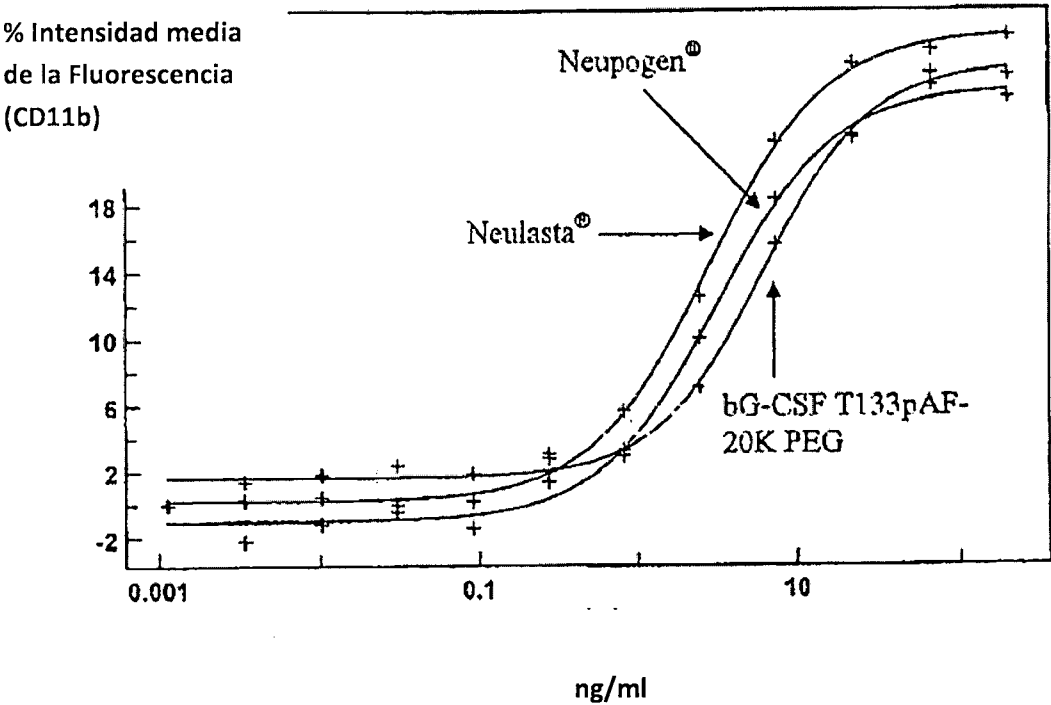
1. Un polipéptido bG-CSF pegilado que comprende la secuencia SEQ ID NO: 1 o SEQ ID NO: 2, donde la posición 133 de la SEQ ID NO: 1 o la posición 134 de la SEQ ID NO: 2 está sustituida con una para-acetilfenilalanina; y en donde la para-acetilfenilalanina está unida a un polímero soluble en agua que comprende una fracción de poli(etilen)glicol, en donde el poli(etilen)glicol tiene un peso molecular de 10kDa a 40kDa.
2. El polipéptido bG-CSF de la Reivindicación 1, donde el polímero soluble en agua tiene un peso molecular de 20 kDa.
3. El polipéptido bG-CSF de la Reivindicación 1, donde el polímero soluble en agua tiene un peso molecular de 30 kDa.
4. El polipéptido bG-CSF de la Reivindicación 1, donde el polímero soluble en agua tiene un peso molecular de 40 kDa.
5. El polipéptido bG-CSF de la Reivindicación 2, en donde dicho polipéptido bG-CSF exhibe el 20% de la actividad de la bG-CSF tipo silvestre *in vitro*, medida mediante el ensayo de proliferación M-NFS60.
6. El polipéptido bG-CSF de la Reivindicación 2, donde dicho polipéptido bG-CSF está caracterizado por inducir un conteo máximo absoluto de neutrófilos (ANC) 10 veces mayor que el ANC previo al tratamiento, medido 72 horas después de una única inyección subcutánea de 40 µg/kg en un ternero saludable.
7. El polipéptido bG-CSF de la Reivindicación 6, donde dicho polipéptido bG-CSF está caracterizado por inducir un conteo máximo absoluto de neutrófilos (ANC) de 7 a 10 veces mayor que el ANC previo al tratamiento de 12 a 96 horas, medido después de una única inyección subcutánea de 40 µg/kg en un ternero saludable.
8. Una composición que comprende el polipéptido bG-CSF de acuerdo con cualquiera de las Reivindicaciones 1 a 7 a una concentración adecuada para su administración desde 0.01 a 1000µg/kg de peso corporal y un portador farmacéuticamente aceptable, comprendiendo el portador uno o más entre solución salina, solución salina amortiguada, dextrosa, agua, glicerol, un amortiguador y etanol.

TABLA 7					
		Filogastrim	Filgastrim pegilado	WT bG-CSF	20K PEG-bT133 [lote NK2]
Sumario	EC50 prom [ng/ml]	0.025	0.077	0.053	0.270
	SD	0.005	0.004	0.005	0.042
	CV	18%	5%	9%	16%
	N	14	5	16	2

TABLA 8					
		Filogastrim	Filgastrim pegilado	WT bG-CSF	20K PEG-bT133 [lote NK2]
Sumario	Dif de Pliegues EC50 [X]	1.0	3.2	1	5.1
	SD	0.0	0.7	0	0.5
	CV	0%	23%	0%	9%
	N	14	5	12	2

TABLA 9			
	Valores EC50 prom	SD	CV
Filogastrim	0.025	0.005	18%
WT bG-CSF	0.053	0.005	9%
Filgastrim pegilado	0.077	0.004	5%
20kPEG-bS62	0.150	0.008	5%
20kPEG-bT133 [NK1]	0.258	0.045	17%
20kPEG-bT133 [NK2]	0.270	0.042	16%
20kPEG-bR7	0.348	0.073	21%
20kPEG-bR166	0.363	0.033	9%
20kPEG-bL3	0.477	0.081	17%

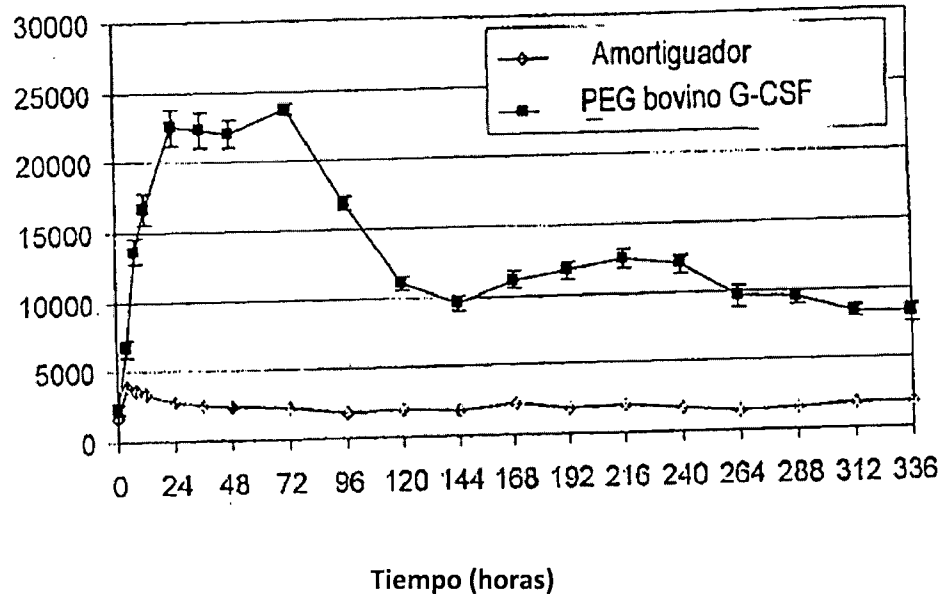
VERSIÓN CORREGIDA FIGURA 15



20

VERSIÓN CORREGIDA FIGURA 16

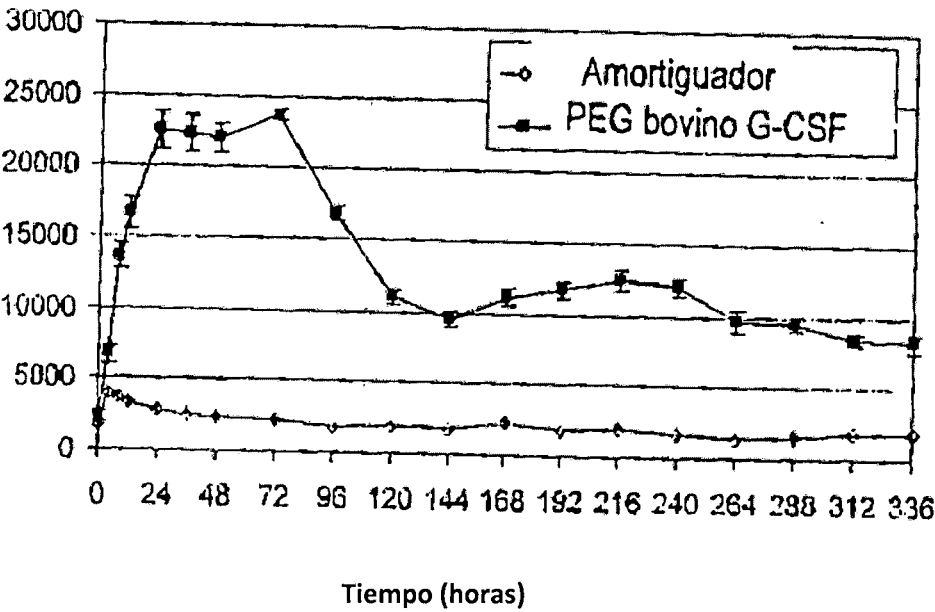
Conteo absoluto de
nucleótidos



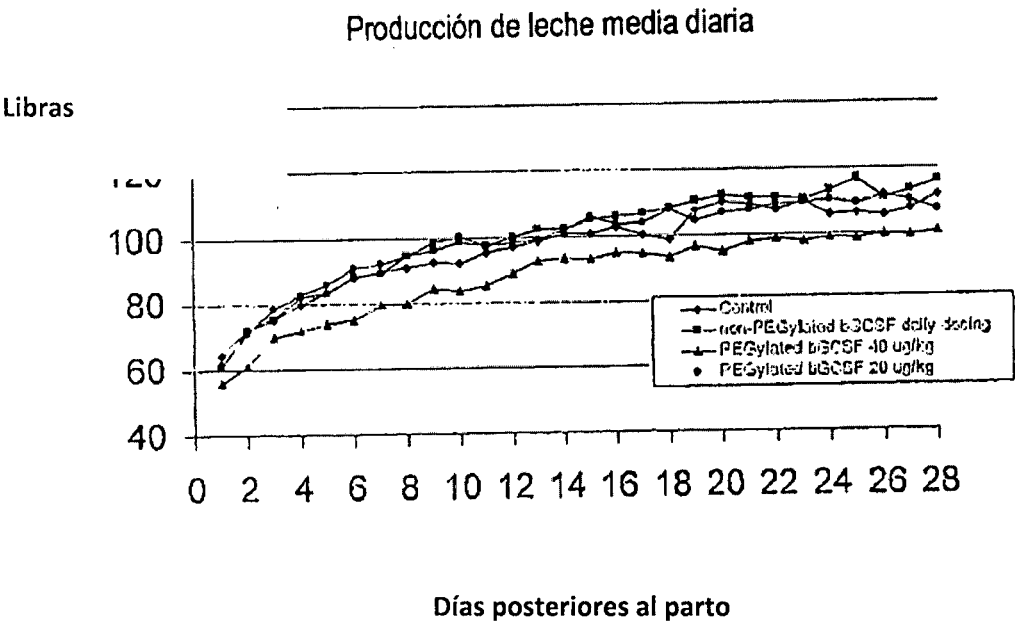
21

VERSIÓN CORREGIDA FIGURA 17

Conteo absoluto de
nucleótidos



VERSIÓN CORREGIDA FIGURA 18



Crystal Structure of Canine and Bovine Granulocyte-Colony Stimulating Factor (G-CSF)

Brett Lovejoy†, Duilio Cascio and David Eisenberg

*Molecular Biology Institute and Department of Chemistry and Biochemistry
University of California, Los Angeles
CA 90024-1570, U.S.A.*

(Received 2 March 1993; accepted 13 May 1993)

The crystal structures of recombinant canine and bovine granulocyte colony stimulating factor (G-CSF) have been determined by X-ray crystallography, using molecular replacement with recombinant human G-CSF as a model. G-CSF is a member of the cytokine family of glycoproteins that stimulate the differentiation and proliferation of blood cells. Human, bovine and canine G-CSF all have a molecular mass of about 19 kDa and share an amino acid sequence identity of about 80%. Two crystal forms of canine G-CSF have been solved. Form I recombinant canine G-CSF (rcG-CSFI; space group *C2*) contains one molecule in the asymmetric unit while form II canine G-CSF (rcG-CSFII; space group *P2₁*) has two molecules in the asymmetric unit and bovine G-CSF (rbG-CSF; space group *P2₁2₁2₁*) contains one molecule in the asymmetric unit. rcG-CSFI has been refined to an *R* factor of 20.7% with data to 2.3 Å resolution and rcG-CSFII has been refined to an *R* factor of 19.3% with data to 2.2 Å resolution. rbG-CSF has been refined to an *R* factor of 21.3% with data to 1.7 Å resolution. The structure of human, canine and bovine G-CSF is an antiparallel 4- α -helical bundle with up-up-down-down connectivity. With the exception of one highly exposed loop (residues 66 to 74), the human, canine and bovine structures are very similar to each other. Using our series of G-CSF crystal structures we developed a function that describes the probability that a particular residue position (*i*) contributes to a G-CSF receptor binding site based on two principles, (1) high sequence conservation in the primary sequence of human, bovine, canine and murine G-CSF and (2) conservation of high solvent accessibility in the human, bovine and canine crystal structures. On the basis of this probability function as well as a comparison of G-CSF to the crystal structure of human growth hormone (hGH) complexed with the extracellular domain of the human growth hormone receptor (hGHbp), residues that contribute to potential G-CSF receptor binding sites are identified.

Keywords: protein structure; cytokines; alpha helical bundles

1. Introduction

Colony stimulating factors (CSFs†) are a family of glycoprotein ligands that promote proliferation and terminal differentiation of multipotential stem cells by binding to membrane-bound receptors presented by these cells (Metcalf, 1985). Each CSF has its own unique high-affinity receptor. The members of the CSF family do not share extensive sequence identity and are classified by the terminal cell type they promote (Metcalf, 1985). Members of the CSF family include: granulocyte colony-stimulating factor (G-CSF), granulocyte-macrophage colony-stimulating factor (GM-CSF), macrophage colony-stimulating factor (M-CSF), interleukin-3, and

erythropoietin (Klingemann, 1989). G-CSF stimulates the proliferation and terminal differentiation

† Abbreviations used: CSF, colony stimulating factor; rbG-CSF, bovine granulocyte-colony stimulating factor; rcG-CSF, canine granulocyte-colony stimulating factor; rcG-CSFI, canine granulocyte-colony stimulating factor (crystal form I); rcG-CSFII, canine granulocyte-colony stimulating factor (crystal form II); GM-CSF, granulocyte-macrophage colony stimulating factor; rhG-CSF, human granulocyte-colony stimulating factor; hGH, human growth hormone; hGHbp, extracellular portion of the human growth hormone receptor; IFN- β , interferon- β ; IFN- γ , interferon- γ ; IL-2, interleukin-2; IL-4, interleukin-4; M-CSF, macrophage colony-stimulating factor; PGH, porcine growth hormone; r.m.s., root-mean-square; PEG, polyethyleneglycol; MIR, multiple isomorphous replacement.

† Present address: Glaxo Research Institute, 5 Moore Drive, Research Triangle Park, NC 27709, U.S.A.

of neutrophilic granulocytes from progenitor cells that present the G-CSF receptor (Nomura *et al.*, 1986).

Clinical trials have shown that G-CSF is effective at treating neutropenia (Scarffe & Kamthan, 1990; Glaspy & Golde, 1992). Neutropenia is a pervasive health problem in which the absolute neutrophil count drops below $1.0 \times 10^9/\text{litre}$; at this concentra-

tion the risk of lethal infections such as sepsis and bacteremia starts to rise. Several important chemotherapeutic agents induce neutropenia but G-CSF can be effective at treating this condition. For example, patients who received G-CSF experienced a decreased incidence of neutropenic fever and culture-confirmed infections compared to patients receiving a placebo (Crawford *et al.*, 1990). Such

Sequence	1				
Human	MTPLGPASSL	PQSFLKCLE	QVRKIQQDGA	ALQEKLCATY	KLCHPEELVL
Canine	MAPLGPTGPL	PQSFLKCLE	QMRKVQADGT	ALQETLCATH	QLCHPEELVL
Bovine	.TPLGPASRL	PQSFLKCLE	QVRKIQADGA	ELQERLCAAH	KLCHPEELML
Structure					
Human	DDDDSS	DDHHHHHHHH	HHHHHHHHHH	HHHHHHHTTS	DDDDGGGGTT
Canine Form I	D	DDHHHHHHHH	HHHHHHHHHH	HHHHHHHHHH	DDDDGGGGTHH
Canine Form II	D	DDHHHHHHHH	HHHHHHHHHH	HHHHHHHHHH	DDDDDDDDHH
Bovine	DD	DDHHHHHHHH	HHHHHHHHHH	HHHHHHHHHH	DDDDHHHHKT
Description			Helix A		E
Sequence	51				
Human	LGHSLGIPWA	PLSSCFSSQAL	QLAGCLSQLH	SGLFLYQGLL	QALAGISPGL
Canine	LGHSLGIPQF	PLSSCFSSQAL	QLMGCLRLH	SGLFLYQGLL	QALAGISPGL
Bovine	LHSLGIPQA	PLSSCFSSQSL	QLRGCLNQLH	GGLFLYQGLL	QALAGISPGL
Structure					
Human	GGGSSDDDDD	DDSDSDSSDD	SSTHHHHHHH	HHHHHHHHHH	HHTTSSSTTH
Canine Form I	HHHHTDDDD	DDGGGSGGGD	DDHHHHHHHH	HHHHHHHHHH	HHTTSSSTTT
Canine Form II	HHHHTDDDD	DDGGGSTTSD	DDHHHHHHHH	HHHHHHHHHH	HHTTSSSTTT
Bovine	THHHHTDDDD	DDGGGSTTSD	DDHHHHHHHH	HHHHHHHHHH	HHTTSSSTTT
Description			<-LOOP->	Helix B	<-TURN->
Sequence	101				
Human	QPTLDTLQLD	VADFATTIWQ	QMEELGMAPA	LQPTQGAMPA	FASAFQRRAG
Canine	APTLDLTLQLD	TTDFAINIWQ	QMEDLGMAPA	VPPTQGTMPA	FTSAFQRRAG
Bovine	APTLDLTLQLD	VTDFATNIWL	QMEDLGMAPA	VQPTQGTMPA	FTSAFQRRAG
Structure					
Human	HHHHHHHHHH	HHHHHHHHHH	HHTTSSDDDD	DDDDSDDDDD	DDSHHHHHHH
Canine Form I	HHHHHHHHHH	HHHHHHHHHH	HHTTSDDDDD	DDDD	DDSHHHHHHH
Canine Form II	THHHHHHHHH	HHHHHHHHHH	HHTTSDSDDD	DDDD	DDSHHHHHHH
Bovine	HHHHHHHHHH	HHHHHHHHHH	HHHHHSDDDD	DDDD	DDSHHHHHHH
Description			Helix C		
Sequence	151				
Human	GVLVASHLQS	FLEVSYRVLR	HLAQP		
Canine	GVLVASNLQS	FLELAYRALR	HPAKP		
Bovine	GVLVASQLHR	FLELAYRGLR	YLAEP		
Structure					
Human	HHHHHHHHHH	HHHHHHHHHH	HTSDD		
Canine Form I	HHHHHHHHHH	HHHHHHHHHH	TDD..		
Canine Form II	HHHHHHHHHH	HHHHHHHHHG	GGD..		
Bovine	HHHHHHHHHH	HHHHHHHHHH	HTD..		
Description			Helix D		

Figure 1. The primary sequence of human, canine and bovine G-CSF along with the secondary structure of the rhG-CSF molecule A (Hill *et al.*, 1993), rcG-CSFI, rcG-CSFII molecule B, and rhG-CSF models. The secondary structure was determined using DSSP (Kabsch & Sander, 1983). H= α -helix, G= 3_{10} helix, S=bend, T=H-bonded turn, D=coil, . = residue missing in the crystallographic model. Residues that are not conserved in human, canine or bovine G-CSF are bold faced. The disulfide bonds between Cys37 and Cys43 and between Cys65 and Cys75 are each shown by a line. The final field in each sequence alignment is a composite description of the secondary structure of G-CSF taken from all 4 models. In this description ■ indicates α -helices, ☼ indicates crossover connections <-Loop-> delineates the loop stabilized by the disulfide bond between Cys65 and Cys75 while <-Turn-> represents the highly conserved hairpin BC turn. In rhG-CSF helix A spans residues 12 to 40, helix B spans residues 72 to 92, helix C spans residues 101 to 125, and helix D spans residues 144 to 171. The AB crossover connection begins at residue 41 and ends at residue 71, the BC crossover connection begins at residue 93 and ends at residue 99 while the CD loop begins at residue 126 and ends at residue 143. Within the AB crossover connection is helix E which is of variable length in each species of G-CSF. For rhG-CSF helix E begins at residue 45 and ends at residue 55.

studies raise the possibility of increasing the dosage of chemotherapeutic agents administered to cancer patients. While the usage of higher dosages of chemotherapeutic agents remains controversial, the finding by Von Hoff *et al.* (1986) that a relatively small increase in drug dose results in a disproportionately large increase in tumour cell kill suggests that higher dosages may be beneficial. G-CSF has also been used in clinical trials to treat neutropenia associated with: autologous bone marrow transplantation (Molineux *et al.*, 1990), neutropenia induced in advanced AIDS patients by the drug zidovudine (Glaspy *et al.*, 1989), congenital (Bonilla *et al.*, 1989) cyclic (Hammond *et al.*, 1989) and idiopathic (Jakubowski *et al.*, 1989) neutropenia, and in patients with hairy-cell leukemia complicated by neutropenia (Glaspy *et al.*, 1988). The possibility is also being studied (Silver *et al.*, 1989) of using G-CSF to treat severe burn victims and antibiotic-resistant infections.

Human G-CSF is a 19.6 kDa glycoprotein consisting of 174 residues. Glycosylation of a single O-linked site at Thr133 contributes to the stability of G-CSF but does not appear to influence receptor binding (Oh-eda *et al.*, 1990). Both bovine and canine G-CSF are 175 residues long and share an amino acid sequence identity of about 80% with human G-CSF (Fig. 1). In addition, they contain the O-linked glycosylation site found in human G-CSF. The structure of recombinant human G-CSF (rhG-CSF) was recently determined to be an all antiparallel 4- α -helical bundle with a left-handed twist (Hill *et al.*, 1993). rhG-CSF is expressed in *Escherichia coli* and consequently is not glycosylated but is biologically active (Hill *et al.*, 1993).

In this paper, we report the determination and refinement of the crystal structures of two crystallographic forms of recombinant canine G-CSF (forms I and II; rcG-CSFI and rcG-CSFII) as well as recombinant bovine G-CSF (rbG-CSF). The molecular structures we find are highly similar to that of rhG-CSF, which was used as a basis for a search model for molecular replacement. The crystal structures of canine, bovine and human G-CSF are analyzed and two potential G-CSF receptor binding sites are identified.

2. Materials and Methods

(a) Crystals

All crystals were obtained by the hanging drop method. Crystals of rcG-CSFI were grown from drops that contained 15 μ l of a 22.8 mg/ml canine G-CSF solution and 8 μ l of reservoir solution. The reservoir solution consisted of 5.80% (v/v) ethanol, 50 mM Hepes buffer adjusted to pH 6.54 with NaOH, and 200 mM ammonium sulfate. Crystals measuring 4.5 mm $\times$ 1.0 mm $\times$ 1.5 mm grew within approximately 4 weeks at 22°C. They belong to space group C2 with unit cell constants $a=99.2$ Å, $b=36.2$ Å, $c=48.6$ Å, $\beta=96.4^\circ$. There is one molecule in the asymmetric unit, corresponding to $V_M=2.33$ Å³/Da⁻¹ (Table 1; Matthews, 1968).

Crystals of rcG-CSFII were obtained from drops that contained 15 μ l of a 18.0 mg/ml canine G-CSF solution and 8 μ l of reservoir solution. The reservoir solution

consisted of 5.85% ethanol, 50 mM Hepes buffer adjusted to pH 6.54 by NaOH and 255 mM ammonium sulfate. Crystals measuring 1.0 mm $\times$ 1.1 mm $\times$ 0.4 mm grew within approximately 4 weeks at 22°C. They belong to space group P2₁ with unit cell constants $a=53.2$, $b=35.9$ Å, $c=91.0$ Å, $\beta=101.8^\circ$. There are 2 molecules in the asymmetric unit, corresponding to a $V_M=2.35$ Å³ Da⁻¹ (Table 1).

Crystals of rbG-CSF were obtained from drops that contained 10 μ l of a 9 mg/ml bovine G-CSF solution (buffered at pH 3.8) and 10 μ l of reservoir solution. The reservoir solution consisted solely of 50 mM Mes buffer adjusted to pH 5.5 with NaOH. In this unusual circumstance, the pH jump from 3.8 to 5.5 was sufficient to precipitate the protein within 2 days of incubation. Crystals measuring 1.0 mm $\times$ 1.0 mm $\times$ 0.5 mm occasionally grew from this precipitated protein within 6 weeks of incubation at 22°C. The addition of 1% (w/v) PEG 8000 to the reservoir solution improved the rate at which crystals nucleated from the precipitated protein. The crystals belong to space group P2₁2₁2₁ with unit cell constants $a=46.1$ Å, $b=53.9$ Å, $c=57.1$ Å. There is one molecule in the asymmetric unit, corresponding to a $V_M=1.87$ Å³ Da⁻¹ (Table 1).

(b) Data collection

rcG-CSF, crystal form I: a single rcG-CSFI crystal was mounted in a thin-wall glass capillary and X-ray data were collected to 2.3 Å resolution using a Xuong-Hamlin area detector. The X-ray source was a Rigaku rotating-anode generator operating at 50 kV/180 mA. A graphite monochromator filtered out CuK α radiation of 1.54 Å wavelength. A total of 18,417 reflections were measured. Symmetry related reflections were scaled and averaged with the program FS (Weissman, 1979). The final data set included 6830 unique reflections representing 88.2% of the unique data in the resolution range 30 to 2.3 Å resolution. The R_{merge} value for all data was 4.4% (Table 2).

rcG-CSF, crystal form II: a single rcG-CSFII crystal was mounted in a thin-wall glass capillary and X-ray data were collected to 2.2 Å resolution using a Rigaku R-Axis IIC imaging plate. The X-ray source and monochromator were identical to that used to collect rcG-CSFI data. For the data set, a total of 70 1.4 deg. consecutive oscillation exposures, representing 98°, were taken. Each exposure was 30 min long. A total of 28,860 reflections were measured and reflections $I/\sigma(I) \geq 1.0$ were processed using the R-Axis software to give 15,536 unique reflections representing 90.2% of the unique data to 2.2 Å resolution. The R_{merge} value for all data was 4.6% (Table 2).

Table 1

X-ray characterization of crystals of rcG-CSFI, rcG-CSFII and rbG-CSF

	rcG-CSFI	rcG-CSFII	rbG-CSF
Unit cell parameters			
<i>a</i> (Å)	99.2	53.2	46.0
<i>b</i> (Å)	36.2	35.9	53.9
<i>c</i> (Å)	48.6	91.0	57.1
α (°)	90.0	90.0	90.0
β (°)	96.4	101.8	90.0
γ (°)	90.0	90.0	90.0
Space group	C2	P2 ₁	P2 ₁ 2 ₁ 2 ₁
Molecules per asymmetric unit	1	2	1
Matthews number (Å ³ Da ⁻¹)	2.33	2.35	1.87

Table 2
Data collection and refinement statistics for rcG-CSFI,
rcG-CSFII and rbG-CSF

	rcG- CSFI	rcG- CSFII	rbG- CSF
Resolution (Å)	2.3	2.2	1.7
R_{merge}^a	0.044	0.046	0.055
Measured reflections (no.)	18,417	28,860	74,290
Unique reflections (no.)	6830	15,536 ^b	15,100 ^b
Completeness of data (%)	88.2	90.2	93.0
Non-hydrogen atoms (no.) ^c	1583	3290	1291
Modeled hydrogen atoms (no.)	358	783	455
Water molecules (no.)	59	134 ^d	97
R factor (%) ^e	20.7	19.3	21.3
r.m.s. bond (Å) ^f	0.014	0.013	0.015
r.m.s. angle (°) ^f	2.57	2.50	2.59

^a $R_{\text{merge}} = \sum_i |I_i - \langle I_i \rangle| / \sum_i I_i$ and $\langle I_i \rangle$ is the average of I_i over all symmetry equivalents.

^b Only reflections $F/\sigma(F) \geq 1.0$ were considered measured in the rcG-CSFII and rbG-CSF data sets.

^c Residues 1 to 9, 130 to 136 and 174 to 175 are missing in rcG-CSFI, residues 1 to 8, 130 to 136 and 175 are missing in rcG-CSFII molecule A, residues 1 to 8, 131 to 136 and 174 to 175 are missing in rcG-CSFII molecule B and residues 1 to 8, 130 to 136 and 174 to 175 are missing in rbG-CSF. The following side-chains were modeled as alanine: rcG-CSFI: Glu20, Glu46, Glu47, Ser67, Gln68, Glu71, Arg167 and His171; rcG-CSFII molecule A: Ser67, Gln68, Arg167, His171; rcG-CSFII molecule B: Glu46, Ser67, Gln68 and Arg167; rbG-CSF: Ser9, Glu46, Arg52, Gln59, Gln68, Glu123, Thr140 and Thr142.

^d There are 63 water molecules included with the rcG-CSFII molecule A and 71 water molecules included with rcG-CSFII molecule B.

^e R factor calculations include all data from 7 Å to the resolution limit of the respective data sets.

^f r.m.s. bond length and angle are the respective root-mean-square deviations from ideal values.

rbG-CSF: a single rbG-CSF crystal was mounted in a thin-wall glass capillary and X-ray data were collected to 1.7 Å resolution under identical conditions used to collect the rcG-CSFII data. For a complete data set, 60 25-min exposures were taken with each exposure measuring 1.8 deg. of data about ψ . A total of 74,290 reflections were measured and reflections having $F/\sigma(F) \geq 1.0$ were processed using the R-Axis software to give 15,100 unique reflections representing 93% of the unique data to 1.7 Å resolution. The R_{merge} value for all data was 5.5% (Table 2).

(c) Molecular replacement

The structures of rcG-CSFI, rcG-CSFII and rbG-CSF were determined by molecular replacement using the program X-PLOR (Brünger, 1990).

rcG-CSFI: the partially refined MIR structure of rhG-CSF was used as a search model (Hill *et al.*, 1993). This rhG-CSF model was missing residues 1 to 9, 67 to 71, 128 to 137 and 173 to 175 corresponding, respectively, to the amino terminus, two loops, and the carboxyl terminus. The rhG-CSF model had been refined to an R factor of 0.25 against 2.5 Å data and had an r.m.s. deviation from ideal bond lengths and angles of 0.020 Å and 3.0°, respectively. To determine the orientation of the molecule in its unit cell, the unmodified rhG-CSF model was used in Patterson searches as follows: a model Patterson map was computed by placing the rhG-CSF model into a 100 Å × 100 Å × 100 Å orthorhombic P₁

box, calculating the structure factors in the resolution range 15 to 4 Å, and fast Fourier transforming their squared amplitudes on a 1.0 Å grid. The crystal Patterson map was computed using all observed rcG-CSFI $F/\sigma(F) \geq 2.0$ data in the resolution range 15 to 4 Å by fast Fourier transformation on a 1.0 Å grid. The model Patterson map was sampled every 2 deg. about the Lattman (1972) angle Θ_2 within the limits of the asymmetric unit of the Patterson function (Rao, 1980). Following Lattman (1972), for a Θ_2 sampling interval Δ , the sampling interval for Θ_1 is $\Delta/\cos(\Theta_2/2)$ and for Θ is $\Delta/\sin(\Theta_2/2)$. At each sampled orientation Ω_i , the values of the crystal Patterson map at the positions of the rotated model Patterson vectors were determined. These values were correlated with the magnitude of the rotated model Patterson vectors using a weighted product correlation function (Brünger, 1990). The rotation solution that had the highest correlation with the crystal Patterson map was subjected to Patterson correlation (PC) refinement (Brünger, 1990) against all $F/\sigma(F) \geq 2.0$ data in the resolution range 15 to 3 Å. Using the refined model, a 2-dimensional translation search against all $F/\sigma(F) \geq 2.0$ data in the resolution range 15 to 4 Å was calculated. The translation search produced a solution that was 5 σ above the mean (score) and twice as high as the second highest peak of the translation function. The R factor for this solution against all $F/\sigma(F) \geq 2.0$ data in the resolution range 8 to 3 Å was 49%. The 2-dimensional translation function solution was sufficient to completely describe the position of rcG-CSFI in its crystallographic unit cell since the remaining dimension, the y -coordinate, is fixed by the choice of origin in space group C2.

rcG-CSFII: the refined model of rcG-CSFI was used as a search model to find the position of the 2 molecules in the asymmetric unit of rcG-CSFII. The rcG-CSFI model had an R factor of 20.7% for all data between 7.0 and 2.3 Å with an r.m.s. deviation from ideal bond distances and angles of 0.014 Å and 3.16°, respectively. A Patterson search failed to identify 2 distinct solutions corresponding to the 2 molecules in the asymmetric unit of rcG-CSFII. Consequently, a direct rotation search, using the PC-refinement target function, was calculated. In this search, the rcG-CSF search model was placed in the crystallographic unit cell of rcG-CSFII and then rotated to every unique orientation Ω , with the same step size used in the rcG-CSFI Patterson searches. For each orientation (Ω_i), the PC-refinement target function in the resolution range 15 to 4 Å was computed. Two distinct solutions, Ω_1 (molecule A) Ω_2 (molecule B), resulted in a PC correlation coefficient 6 σ higher than that of the average PC correlation coefficient for the direct rotation search. These solutions were independently subjected to PC-refinement using data in the resolution range 15 to 3.5 Å. Solutions Ω_1 and Ω_2 differed by only 1.1° when $P2_1$ symmetry was considered, suggesting that the Patterson search failed to find 2 distinct solutions because the cluster analysis, performed by X-PLOR on top solutions of the Patterson search, had grouped these 2 solutions into the same cluster. A 2-dimensional translation search using Ω_1 resulted in a solution that was 14 σ above the mean value for the search. Similarly, a 2-dimensional translation search using Ω_2 resulted in a solution that was 8 σ above the mean value for the search. To determine the relative translation between these 2 solutions, the second translation function solution, corresponding to molecule B, was translated by the fractional unit cell lengths (0,0,0), (1/2,0,0), (1/2,0,1/2) and (0,0,1/2). At each of these locations, a 1-dimensional translation function was computed along the entire unit cell in the y dimension. A

solution was found for the translation function using the molecule that had been translated by the fractional unit cell lengths (1/2,0,1/2). After rigid body refinement, this combined solution resulted in an R factor of 36.9% for all data in the resolution range 7.0 to 2.2 Å.

rbG-CSF: to find the orientation of rbG-CSF, we used a model of rcG-CSFI at an intermediate stage of refinement as the basis for a search model. The rcG-CSFI model had been refined to an R factor of 22.1% against all data in the resolution range 7.0 to 2.3 Å and had an r.m.s. deviation from ideal bond distances and angles of 0.015 Å and 3.34°, respectively. At all residue positions where the sequences of rbG-CSF and rcG-CSF differ, only atoms that were common to both residue types were preserved. For the Patterson search, a model Patterson map was computed from the rcG-CSFI model in a manner identical to that used in the rcG-CSFI Patterson search. The crystal Patterson map was computed from all rbG-CSF data with $F/\sigma(F) \geq 2.0$ in the resolution range 15 to 4 Å by fast Fourier transformation on a 1.0 Å grid. The Patterson search did not result in a clear solution. Consequently the top 77 solutions to the Patterson search were subjected to PC-refinement. This resulted in a clear solution: the score of one PC-refined orientation, the third highest Patterson solution, was 1.8 times that of the second highest score. A 3-dimensional translation search using all data in the resolution range 10 to 3.5 Å resulted in a clear solution that was 11σ above the mean score and 4 times above the second highest score of the search. The R factor for this translation function solution was 49.0% for all data in the resolution range 8 to 3.5 Å.

(d) Refinement

rcG-CSFI: the rcG-CSFI translation function solution was subjected to several cycles of Powell refinement (Brünger *et al.*, 1990), resulting in an R factor of 38.0% within the resolution limits of 8.0 to 3.0 Å. The program FRODO (Jones, 1985) was used to build side-chains that differ in rbG-CSF and rcG-CSF into well-defined $2.5\sigma F_o - F_c$ density maps. The remodeled structure was then subjected to Powell refinement and simulated annealing (Brünger *et al.*, 1990), resulting in an R factor of 34.1% for all data in the resolution range 7.0 to 2.3 Å. Refinement continued for several cycles with each cycle consisting of: (1) an examination of $F_o - F_c$ maps for missing residues and missing side-chains; (2) an examination of $2F_o - F_c$ maps for symptoms of errors in the model, such as weak density; (3) model building with FRODO using information gained from these maps; and (4) stereochemical refinement using X-PLOR with a single isotropic B factor. During this stage of refinement, the loop consisting of residues 67 to 71 was built into $F_o - F_c$ density density maps. The program O (Jones, 1990) was then used to refine exposed loop regions against a database of highly refined structures. In the final stages of refinement, atomic B -factors were refined and 59 water molecules were identified. The final rcG-CSFI model has an R factor of 20.7% for 6532 reflections representing all data measured in the resolution range 7.0 to 2.3 Å (Table 2).

rcG-CSFII: after rigid body refinement, the rcG-CSFII structure, with a single isotropic B factor, was subjected to Powell refinement followed by simulated annealing and atomic B factor refinement. FRODO was used to place Pro9 and Ala173 of molecule A and B into well-defined $2.5\sigma F_o - F_c$ density. In addition, several side-chains that were not defined in rcG-CSFII were modeled and 134 water molecules were identified. The final rcG-CSFII

model has an R factor of 19.3% for 15,402 reflections representing all $F/\sigma(F) \geq 1.0$ data in the resolution range 7.0 to 2.2 Å (Table 2).

rbG-CSF: the rbG-CSF translation function solution was subjected to rigid body refinement and Powell refinement using a single isotropic B factor, resulting in an R factor of 33.4% for all data in the resolution range 20 to 2.5 Å. Subsequent refinement using the simulated annealing protocol (Brünger *et al.*, 1990) brought the R factor down to 29.8%. Several cycles of refinement were conducted. Each cycle was identical to the rcG-CSFI refinement steps (1) to (4) with the addition of atomic B factor refinement. The final rbG-CSF model has an R factor of 21.3% for 14,989 reflections representing all $F/\sigma(F) \geq 1.0$ data in the resolution range 7.0 to 1.7 Å (Table 2).

3. Results

(a) Description of the 4-α-helical bundle

Common to the structures of human, canine and bovine G-CSF is an antiparallel 4-α-helical bundle in which the helices run up-up-down-down (Fig. 2). We refer to the long segments that connect helix A to helix B and helix C to helix D, respectively, as the AB and CD crossover connections. To date, additional examples of up-up-down-down 4-α-helical bundles with long overhand crossover connections include the cytokines human growth hormone (hGH) (de Vos *et al.*, 1992), porcine growth hormone (PGH) (Abdel-Meguid *et al.*, 1987), granulocyte-macrophage colony-stimulating factor (GM-CSF) (Diederichs *et al.*, 1991), macrophage

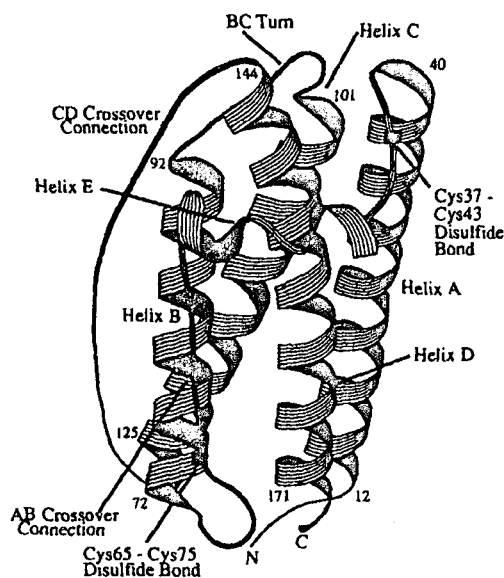


Figure 2. Ribbon diagram of rbG-CSF highlighting its structural elements. G-CSF consists of a 4-α-helical bundle (helices A, B, C and D) with up-up-down-down topology. Two pairs of helices are connected by long crossover connections (AB and CD). The conformational freedom of crossover connection AB is restricted by an additional helix (helix E) and 2 disulfide bridges. The residues at the start and end of helices A, B, C and D are numbered.

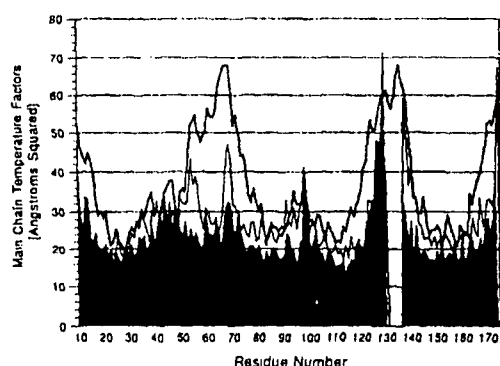


Figure 3. rbG-CSF, rhG-CSF and rcG-CSFII main-chain atomic B factors. rbG-CSF is represented by the shaded area, rcG-CSFII is represented by a thin line and rhG-CSF is represented by a thick line. The B factors of rcG-CSFII molecule A and B are averaged while rhG-CSF is represented by molecule C. The gap in the rbG-CSF and rcG-CSFII lines at residues 130 through 136 reflects the absence of this segment in these models.

colony-stimulating factor (M-CSF) (Pandit *et al.*, 1992), interferon- β (IFN- β) (Senda *et al.*, 1992), interleukin-2 (IL-2) (Mott *et al.*, 1992; Brandhuber *et al.*, 1987) and interleukin-4 (IL-4) (Wlodaver *et al.*, 1992; Walter *et al.*, 1992; Powers *et al.*, 1992). As judged by their atomic B factors (Fig. 3), the G-CSF crossover connections are highly flexible. The segment connecting the B and C helices forms a short hairpin connection and is referred to as the BC loop. A fifth α -helix occurs in the AB crossover connection and is referred to as the E helix.

(b) Quality of the models

The final crystallographic models for rcG-CSFI, rcG-CSFII and rcG-CSF are defined for residues 9 to 130 and for residues 137 to 172 (Fig. 1, Table 2). The amino-terminal region, residues 1 to 8, and the carboxyl terminus, residues 173–174, did not have significant electron density and were not built. The two termini are presumably not essential for the structural integrity of the 4- α -helical bundle of G-CSF. Consistent with this assumption is the finding that deletion of the amino terminus (residues 1 to 11) does not lower the biological activity of G-CSF (Okabe *et al.*, 1990; Hanazono *et al.*, 1990).

The central portion of the CD crossover connection, residues 131 to 136, was absent in all electron density maps and consequently was not included in final models. This region was also not visible in rhG-CSF MTR maps (Hill *et al.*, 1993) indicating that high mobility is a conserved feature of this region. In addition to missing residues, several highly exposed side-chains do not have significant density and consequently were not built (Table 2).

The Ramachandran plots for rcG-CSFI, rcG-CSFII and rbG-CSF show that most residues are in a right-handed α -helical conformation

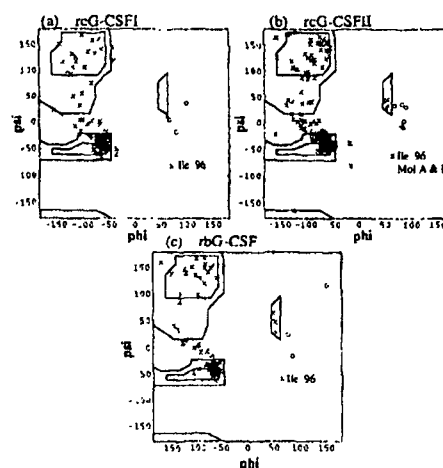


Figure 4. Ramachandran plots (Ramachandran *et al.*, 1963) (a) rcG-CSFI, (b) molecules A and B of rcG-CSFII and (c) rbG-CSF; o = glycine, x = all other residues. The only residue that consistently does not fall in the allowed regions of the Ramachandran plot is Ile96, which has the ϕ ψ angles of a classic γ -turn.

(Fig. 4). Although the AB and CD crossover connections do not form a β -sheet, the ϕ ψ angles of many residues in these segments fall within regions typical of parallel and antiparallel β -sheets. Three non-glycine residues consistently adopt conformations worth noting: Lys41 (Gln41 in rcG-CSF), Ala94 and Ile96. Lys/Gln41 adopt the α_L conformation presumably because it is at the C terminus of helix A (Hill *et al.*, 1993), forming a cap. In human (Hill *et al.*, 1993), canine and bovine G-CSF, Ala94, Gly95 and Ile96 all have positive ϕ angles and together form almost one turn of left-handed helix. These residues are located in the BC hairpin turn and consequently are constrained by helices B and C. The close proximity of helix D and the presence of a Pro98 in the BC turn place additional constraints on the BC hairpin turn. Four hydrogen bonds are stabilizing the left-handed α -helical conformation of residues 94 to 96: (1) Ser97 CO and Ala101 N (length 2.9 Å); (2) Gly95 CO and Ser97 N (2.8 Å); (3) Leu93 CO and Ile96 N (3.1 Å); and (4) Ile96 CO and Arg147 N^H, (3.7 Å). The BC hairpin turn is represented by strong electron density (Fig. 5) Interestingly, Ile57 adopts the α_L conformation in rbG-CSF but not in rcG-CSFI or rcG-CSFII.

The average atomic B factor value for main-chain atoms in the α -helical core of G-CSF is 20.6 Å² for rbG-CSF, 23.7 Å² for rcG-CSFII and 30.6 Å² for rhG-CSF, (Fig. 3). The atomic B factor values in the G-CSF models are most different in the AB crossover connection, residues 41 to 71. In rbG-CSF and rcG-CSFII, the atomic B factors of the main-chain atoms in this segment are 5 to 10 Å² above average values found in the α -helical core of each model while the atomic B factors of main-chain atoms in the AB crossover connection of rhG-CSF are 20 Å² above the average B factor values found in the

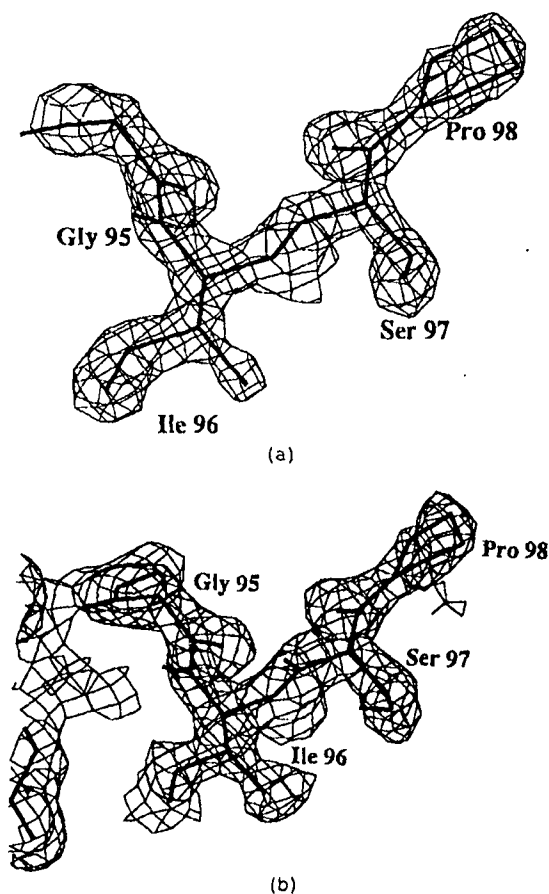


Figure 5. (a) A $2F_o - F_c$ density map of rbG-CSF contoured at 2σ in the BC loop region about Ile96. (b) A SA omit map (Hodel *et al.*, 1992) of rbG-CSF contoured at 1σ . For this map, residues 94 through 98 were omitted and the model was then refined using the SA slow-cooling protocol (Brünger *et al.*, 1990) with a starting temperature of 1000 K. Ala94, Gly95 and Ile96 together form almost one turn of left-handed helix in rhG-CSF, rcG-CSFI, rcG-CSFII and rbG-CSF.

α -helical core of rhG-CSF. The large increase in main-chain B factor values, with the average atomic B factors of side-chains for residues 59 to 72 all exceeding $50 \text{ \AA}^2$. In contrast, the average value for rbG-CSF side-chains of residues 59 to 72 is about $30 \text{ \AA}^2$.

The CD crossover connection, residues 126 to 143, consistently has higher atomic B factors than any other region of G-CSF in all three species of G-CSF (Fig. 3). Another region of high flexibility, represented by high atomic B factors, is the carboxyl terminus, residues 172 to 175. In rhG-CSF, this amino-terminal flexibility extends into the amino half of helix D.

(c) The AB and CD crossover connections

The amino end of the AB crossover connection is bonded to helix A through the Cys37–Cys43 disul-

fide bridge and its carboxyl end is bonded to helix B through the Cys65–Cys75 disulfide bridge (Fig. 2). These disulfide bridges restrict the conformational freedom of the AB crossover connection. The presence of helix E in the amino-terminal half of the AB crossover connection further reduces the conformational freedom of this segment. In contrast, the CD crossover connection does not have the conformational limiting features found in the AB crossover connection and consequently, as indicated by the high atomic B factors of the residues in the CD crossover connection (Fig. 3), is more flexible.

Along with the aligned sequences in Figure 1 is a description of the secondary structure of each model. The secondary structure of each model is nearly identical with the exception of the AB crossover connection which contains helix E. In rhG-CSF helix E is 11 residues long and begins at Pro45 (Fig. 1). The helix has a sharp kink at residues 50 and 51 that disrupts the α -helical $i+4$ main-chain hydrogen bonding pattern. In rcG-CSF helix E begins at Val49 and is only seven residues long while in rhG-CSF molecule A the helix begins at Pro45 but is a nine-residue 3_{10} helix with a sharp kink in the center of the helix, residues 49 and 50, similar to the one found in rhG-CSF. While the quality of the electron density maps for helix E varies, with the electron density for rbG-CSF being the best and rhG-CSF being the worst, the data are of sufficient quality to suggest with high certainty that the differences observed in helix E in each species of G-CSF reflect true structural variations in G-CSF rather than structural uncertainty in the models. As a consequence of its exposed location, helix E is consistently involved in crystal packing interactions. The observation that even canine form I and II crystals (rcG-CSFI and rcG-CSFII) show different structures in this region (Fig. 1) suggests that these packing interactions are likely to be one of the factors that contribute to the structural diversity of helix E. The rest of the AB crossover connection is devoid of secondary structure except for the carboxyl end of the segment beginning with Cys65 where the disulfide bond between Cys65 and Cys75 stabilizes a sharp turn (Fig. 2).

(d) Superposition of the models

Figure 6 shows the superposition of the polypeptide backbones of human, canine and bovine G-CSF by the method of Hendrickson (1979). All superpositions include the backbone atoms of residues 10 to 129 and 137 to 172. Table 3 summarizes the r.m.s. deviation between corresponding atomic positions in each set of superimposed structures, averaged over all the atoms included in the superposition. The three rcG-CSF molecules (rcG-CSFI, rcG-CSFII molecules A and B) are very similar to each other, with an average r.m.s. deviation of about $0.6 \text{ \AA}$. Similarly, canine and bovine G-CSF are in good agreement, with an average r.m.s. deviation of about $1.1 \text{ \AA}$. Superposition of human G-CSF to canine and bovine G-CSF results

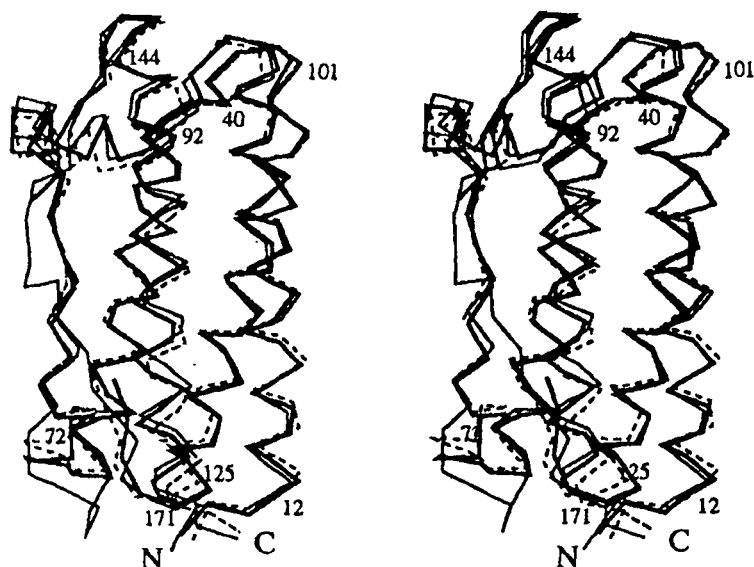


Figure 6. Superposition of the C^α backbones of rhG-CSF molecule B (thin line), rcG-CSFII molecule A (broken line) and rbG-CSF (thick line). The residues at the start and end of helices A, B, C and D are numbered.

in a larger r.m.s. deviation; about 1.7 Å and 1.6 Å, respectively.

The AB crossover connection (residues 41 to 71) is largely responsible for the high r.m.s. deviation between human, canine and bovine G-CSF (Fig. 7). The r.m.s. deviation in the AB crossover connection is highest in the region bounded by the Cys65–Cys75 disulfide bond, residues 66 to 74. Residues 66 to 74 correspond to the C-terminal end of the AB crossover connection and the N terminus of helix B. The superposition of the canine, bovine and human models reveals that this region folds in a very different manner in rhG-CSF than in rcG-CSFII and rbG-CSF (Fig. 8). Residues 66 to 74 were missing in the original search model used to build rcG-CSFI. This segment was built into strong rcG-CSFI $F_o - F_c$ density over the course of several refinement cycles. rcG-CSFI was then used as the starting model for both rcG-CSFII and rbG-CSF so residues 66 to 74 are independent of the rhG-CSF model in these structures as well. Since this region is on the surface

of G-CSF, it is possible that the unique packing environments of each G-CSF structure are responsible for the large r.m.s. separation between the human *versus* the bovine and canine models in this region. However, rcG-CSFI, rcG-CSFII and rbG-CSF each have different packing environments yet maintain very good structural homology (compare Fig. 7a to 7b), suggesting that this possibility is unlikely. Further, the r.m.s. separation of residues 66 to 74 in the human, canine and bovine models is so large (as much as 8 to 9 Å), it cannot be attributed to high B factors. The accuracy of the canine and bovine models is not in doubt: there is strong electron density for this region in rcG-CSFI, rcG-CSFII and rbG-CSF. Interestingly, the atomic B factors for residues 66 to 74 of rhG-CSF are disproportionately high, suggesting that the lack of structural conservation between rhG-CSF and the canine and bovine models may be due to inaccuracies in the rhG-CSF model in this specific region (Fig. 3).

Table 3
Coordinate comparison of G-CSF molecules

	rhG-CSF A	rhG-CSF B	rhG-CSF C	rcG-CSFI	rcG-CSFII A	rcG-CSFII B	rbG-CSF
rhG-CSFA	—	1.31	1.17	1.83	2.03	1.92	1.85
rhG-CSF B	—	—	1.19	1.61	1.80	1.69	1.55
rhG-CSF C	—	—	—	1.47	1.70	1.57	1.48
rcG-CSFI	—	—	—	—	0.63	0.57	1.09
rcG-CSFII A	—	—	—	—	—	0.55	1.20
rcG-CSFII B	—	—	—	—	—	—	1.22

The r.m.s. comparison of rhG-CSF (molecules A, B and C), rcG-CSFI, rcG-CSFII (molecules A and B) and rbG-CSF using all main-chain atoms (N, C^α , C and O) from residues 10 to 129 and 137 to 173. The r.m.s. difference between corresponding atomic positions in each set of superimposed structures, averaged over all 620 atoms included in the superposition is tabulated in Å.

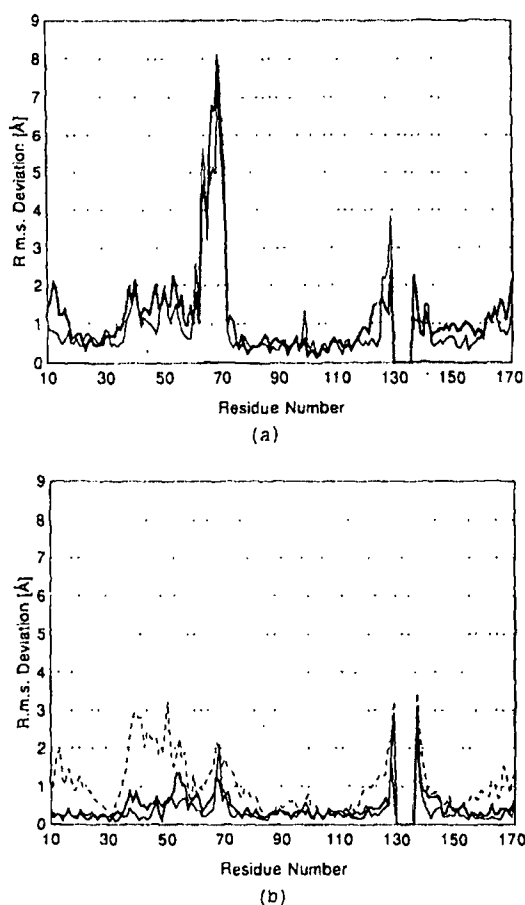


Figure 7. Mean differences between positions of corresponding backbone atoms (N, C α , C and O) after superposition of (a) rhG-CSF and rhG-CSF molecule B (thin line) and rcG-CSFII molecule A and rhG-CSF molecule B (thick line); (b) rcG-CSFII molecule A and B (thin line), rcG-CSFII molecule A (thick line), and rcG-CSFII molecule A and rhG-CSF (broken line)

4. Discussion

Binding of G-CSF to stem cells expressing the G-CSF receptor is an essential step in the cascade of biochemical events resulting in the terminal differentiation of these cells as neutrophilic granulocytes. To elucidate how the binding of G-CSF to its receptor triggers terminal differentiation of the pluripotent stem cells it is first necessary to determine the regions of G-CSF that bind to its receptor. We present two approaches to identify these receptor binding sites (1) a function that describes the probability that a particular residue contributes to a receptor binding site and (2) a structural comparison of the crystal structure of G-CSF and human growth hormone (hGH) complexed with the extracellular domain of the human growth hormone receptor (hGHbp) (de Vos *et al.*, 1992). The hGH·hGHbp complex shows that hGH binds two receptor molecules with two distinct binding sites (de Vos *et al.*, 1992).

(a) Binding probability function

The finding by Hammond *et al.* (1989) that canine and human G-CSF have nearly identical physiological properties in dogs and the finding by Serushago *et al.* (1992) that rhG-CSF enhances resistance to the bacteria *Listeria monocytogenes* in mice raises the possibility that most of the residues constituting the receptor binding sites of G-CSF are conserved in human, canine, bovine and murine G-CSF. If the residues that contribute to receptor binding sites are in fact conserved in the various species of G-CSF, it is highly likely that residues contributing to G-CSF receptor binding sites have a high degree of sequence conservation throughout all species of G-CSF. A second expectation of residues that contribute to G-CSF receptor binding sites is that they have a high degree of solvent accessibility in the human, canine and bovine crystal structures.

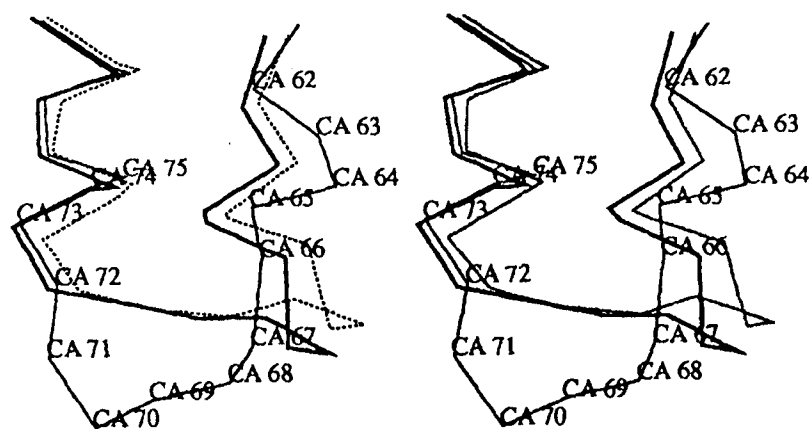


Figure 8. Superposition of the C α backbones of rhG-CSF molecule B (thin line), rcG-CSFII molecule A (broken line) and rhG-CSF (thick line) in region bounded by the Cys65-Cys75 disulfide bond. The C α carbons of the rhG-CSF are numbered. The Figure illustrates the dissimilarity of the human model and the canine and bovine models in the turn bounded by the Cys65-Cys75 disulfide bond.

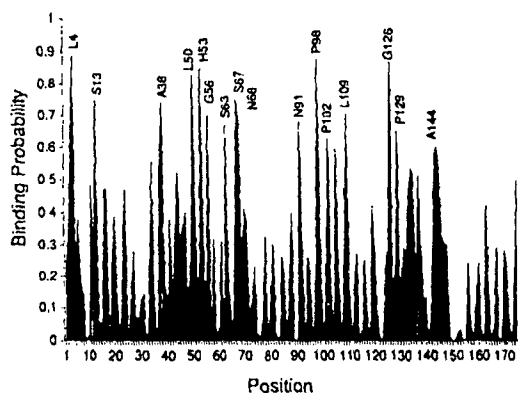


Figure 9. Identification of potential receptor binding sites in G-CSF. The probability for each residue position to be involved in receptor binding is determined with a function that considers accessible surface area and sequence variation (see the text). Residues that score higher than 0.6 are labeled using the primary sequence of rbG-CSF.

Combining the assumptions of (1) high sequence conservation and (2) a high degree of solvent accessibility, we describe the relative probability (binding_i) that a particular residue position (i) in G-CSF contributes to a binding site as:

$$\text{binding}_i = \text{accessibility}_i * 1/\text{variation}_i$$

where accessibility represents the average percent solvent accessibility (Lee & Richards, 1971) of the side-chain i in the rhG-CSF, rcG-CSFI, rcG-CSFII and rbG-CSF crystal structures and variation is the number of different amino acids that occur at residue i . The values for complete exposure of the side-chain are taken from Eisenberg *et al.* (1989). Thus if a side-chain were 100% exposed in the rhG-CSF, rcG-CSFI, rcG-CSFII and rbG-CSF crystal structures and was perfectly conserved in human, canine, bovine and murine G-CSF, the position would receive the largest possible score of one. Plotting this binding probability function against positions reveals several residues that have high scores (Figs 9 and 10a).

(b) Structural comparison of G-CSF to hGH

G-CSF and hGH (de Vos *et al.*, 1992) are both 4- α -helical bundles with up-up-down-down connectivity. The length of the helices that comprise the 4- α -helical bundle of G-CSF and hGH are very similar (Table 4). In both structures, long crossover connections join parallel helices (AB and CD in G-CSF). The length of the crossover connections of hGH are longer than the corresponding crossover connections of G-CSF (Table 4). Despite their differences in lengths, there are many similarities between the G-CSF and hGH crossover connections. In both structures, the first crossover connection (joining helices A and B in G-CSF and helices 1 and 2 in hGH) contains the only disulfide bonds found in

Table 4

The number of residues in corresponding helices and crossover connections in G-CSF and hGH

Element (rbG-CSF/hGH)	Length (rbG-CSF/hGH)
Helix A/helix 1	29/26
Helix B/helix 2	21/21
Helix C/helix 3	25/23
Helix D/helix 4	28/30
Helix E/helix 38-47	11/10
AB segment/12 segment	31/36
CD segment/34 segment	18/25

either structure and conserves an α -helix (corresponding to helix E in G-CSF). In contrast, the second crossover connection (joining helices C and D in G-CSF and helices 3 and 4 in hGH) contains no such conformational limiting features and portions of this segment are missing in the electron density maps of both structures (residues 147 to 153 in hGH and 130 to 136 in rbG-CSF).

Superposition of the C $^{\alpha}$ carbons of hGH onto rbG-CSF reveals five regions that have significant structural similarity. The r.m.s. difference between corresponding C $^{\alpha}$ carbons averaged over all 92 residues in these five regions is 2.3 Å (Table 5). The five regions loosely correspond to G-CSF α -helices A,

Table 5

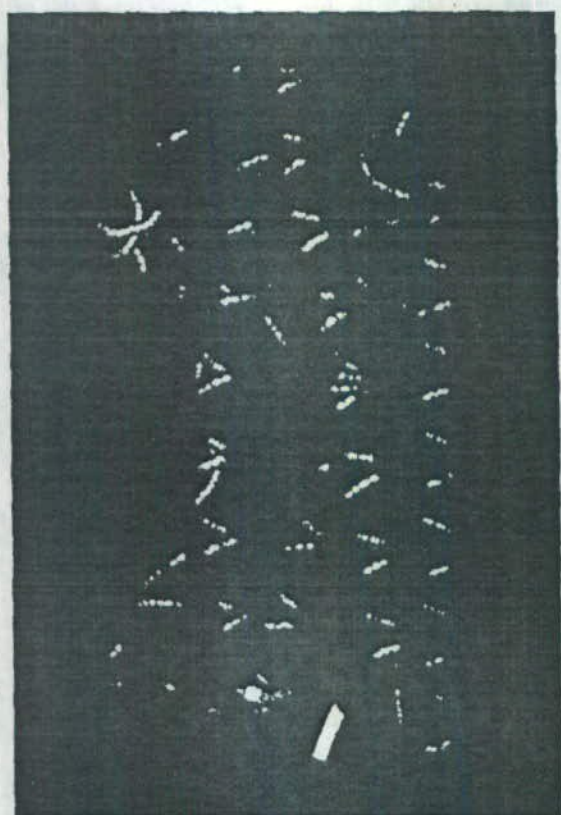
Comparison of 4- α -helical bundles

Structures	Residues used in superposition (no.)	r.m.s. deviation (Å)
G-CSF/hGH	92	2.3
M-CSF/hGH*	66	1.7
M-CSF/GM-CSF*	51	1.8
GM-CSF/IL-4*	63	1.4
G-CSF/M-CSF	57	2.1
IL-2/IL-4	66	2.4
IL-2/IFN- β	62	2.1

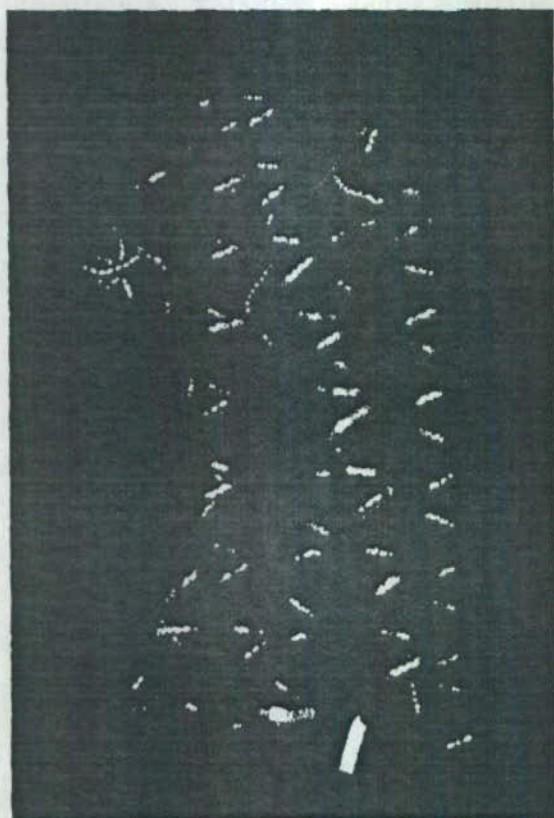
The r.m.s. deviation after the C $^{\alpha}$ superposition of individual members of a structural family of proteins characterized by an up-up-down-down 4- α -helical bundle with long overhand connections. Residues used in each superposition were principally located in the conserved 4- α -helical bundle region of each protein. The C $^{\alpha}$ residues used for the superposition of G-CSF to hGH were: 11 to 33 (10 to 32); 44 to 52 (40 to 48); 67 to 83 (67 to 83); 112 to 126 (111 to 125); and 146 to 173 (154 to 181) where hGH sequences are in parentheses. The C $^{\alpha}$ residues used for the superposition of G-CSF onto M-CSF were: 18 to 32 (12 to 26); 81 to 87 (55 to 61); 108 to 119 (77 to 88); and 150 to 172 (109 to 131) where M-CSF sequences are in parentheses. The C $^{\alpha}$ residues used for the superposition of IL-2 onto IL-4 were: 9 to 23 (5 to 19); 44 to 47 (30 to 33); 55 to 63 (44 to 52); 81 to 95 (75 to 89); 110 to 114 (105 to 109); 116 to 133 (111 to 128) where IL-4 sequences are in parentheses. The C $^{\alpha}$ residues used for the superposition of IL-2 onto IFN- β were: 13 to 29 (8 to 24); 52 to 70 (48 to 66); 81 to 92 (82 to 93); and 113 to 126 (144 to 157) where IFN- β sequences are in parentheses. Unless otherwise noted, superpositions were done using the command LSQ_IMPROVE within the program O (Jones, 1990). This procedure aligns pairs of C $^{\alpha}$ atoms that are within 3.8 Å of each other and does a least-squares minimization of the difference in the positions of these atom pairs to calculate the transformation.

* Superposition reported by Pandit *et al.* (1992).

* Superposition reported by Walter *et al.* (1992).



(a)



(b)

B, C, D and E. The 4- α -helical bundle is structurally conserved in additional cytokines. Pandit *et al.* (1992) found that the C $^{\alpha}$ carbons of 66 residues from the 4- α -helical bundle of M-CSF (Pandit *et al.*, 1992) and hGH (de Vos *et al.*, 1992) can be superimposed with an r.m.s. difference of 1.7 Å. In addition Pandit *et al.* (1992) found that the C $^{\alpha}$ carbons of 51 residues from the 4- α -helical bundle of M-CSF and GM-CSF (Diederichs *et al.*, 1991) can be superimposed with an r.m.s. difference of 1.8 Å. Walter *et al.* (1992) found that the C $^{\alpha}$ carbons of 63 residues in IL-4 (Walter *et al.*, 1992) and GM-CSF can be superimposed with an r.m.s. difference of 1.4 Å. Superposition of rbG-CSF onto M-CSF reveals that the C $^{\alpha}$ carbons of 57 residues from the 4- α -helical bundle of each structure can be superimposed with an r.m.s. difference of 2.1 Å. The C $^{\alpha}$ carbons of 66 residues in a revised 2.5 Å crystal structure of IL-2 (McKay, 1992) can be superimposed onto IL-4 (Wlodaver *et al.*, 1992) with an r.m.s. difference of 2.4 Å and a total of 62 C $^{\alpha}$ carbons from the 4- α -helical bundle region of IL-2 can also be superimposed onto the domain in IFN- β consisting of helices A, B, C and E with an r.m.s. difference of 2.1 Å (Table 5). Finally, Ealick *et al.* (1991) show that IFN- β and interferon- γ (IFN- γ) also have striking similarity in their three-dimensional structures. The structural conservation of the up-up-down-down 4- α -helical bundle in M-CSF, hGH, G-CSF, GM-CSF, IFN- β , IL-2, and IL-4 raises the possibility that additional cytokines can be solved using molecular replacement techniques with the 4- α -helical core of one of these structures as a search model.

hGH receptor binding site I: the majority of the residues that contribute to hGH receptor binding site I (de Vos *et al.*, 1992) map onto three regions of G-CSF: (1) part of helix D, (2) part of helix E, and (3) residues in the AB crossover connection in the vicinity of Cys65 (Fig. 10b). Superposition of hGH on G-CSF, using the same hGH/G-CSF matrix described above, reveals that the average spacial difference between C $^{\alpha}$ carbons of residues in region 1, 157 to 173 of rbG-CSF (helix D) and 165 to 181 of hGH (helix 4), is 2.0 Å. The average spacial difference between C $^{\alpha}$ carbons of residues in region 2,

Figure 10. C $^{\alpha}$ tracing of rbG-CSF depicting residues that may contribute to G-CSF receptor binding sites. Residues 1 to 8 and 174 to 175 are missing in rbG-CSF. Consequently the amino terminus, white, is represented by Ser8 and the carboxyl terminus, yellow, is represented by Ala173. The highly exposed loop that was identified by both the structural comparison of hGH to G-CSF and the binding function is in the lower left-hand corner of both (a) and (b). The structural variation of this region may contribute to the receptor binding specificity of hGH and G-CSF. (a) Residues identified by the binding probability function are red. (b) Residues identified by the structural comparison of hGH to G-CSF. Residues analogous to hGH receptor binding site I are red and residues analogous to hGH receptor binding site II are green.

Table 6
Predicted G-CSF receptor-binding sites

hGH binding site	Corresponding G-CSF site	Residues identified by structural comparison	Residues identified by the binding probability function
Site I (1)	Part of helix D	157, 159, 160, 163, 164, 166, 167, 170	144
Site I (2)	Part of helix E	45, 46, 48, 49	50, 53, 56
Site I (3)	The C-terminal half of the AB crossover connection	63, 66, 67, 68, 69	63, 67, 68
Site II (1)	Part of helix A	13, 15, 17, 20	13
Site II (2)	Part of helix C	117, 120	

Residues predicted to have a high probability of contributing to G-CSF receptor binding sites by (1) our receptor binding function or (2) significant structural homology with residues contributing to hGH receptor binding sites in the hGH·hGHbp complex (de Vos *et al.*, 1992). The structural comparison of hGH and G-CSF revealed 3 regions of G-CSF that correspond to hGH receptor binding site I: (1) part of helix D, (2) part of helix E, and (3) the C-terminal end of the AB crossover connection. In addition, 2 regions of G-CSF that correspond to hGH receptor binding site II were identified: (1) part of helix A and (2) part of helix C (Fig. 10b).

residues 45 to 52 of rhG-CSF (helix E) and residues 41 to 48 of hGH is also 2.0 Å. The topology of region 3 of hGH, residues 60 to 69, and the corresponding region in rhG-CSF, residues 63 to 70, is very similar: a highly exposed loop. However, since the hGH loop is longer than the corresponding G-CSF loop, these two regions cannot be superimposed.

Our binding probability function (Figs 9 and 10a) identifies residues in regions 2 (residues 50, 53 and 56) and 3 (residues 63, 67, 68) as having a high probability of contributing to the G-CSF receptor binding site, but not region 1, with the possible exception of residue 144 (Table 6). However, additional support for residues in region 1 contributing to a receptor binding site in G-CSF is provided by the mutational studies of rhG-CSF, in which Kuga *et al.* (1989) found that deletion of residues 165 to 174 results in substantial loss of biological activity.

hGH receptor binding site II: the hGH binding site II is smaller than hGH binding site I and maps onto two regions of G-CSF: (1) part of helix A and (2) part of helix C of rhG-CSF (Fig. 10b). Upon superposition of hGH to G-CSF using the hGH/G-CSF matrix described above, the average spatial separation between C^α carbons of residues in region 1, 13 to 20 of rhG-CSF (helix A) and 12 to 19 of hGH (helix 1), is 2.5 Å while the average spatial separation C^α carbons of residues in region 2, residues 112 to 120 of rhG-CSF (helix C) and residues 111 to 119 of hGH (helix 3), is 2.3 Å. Our binding probability function identifies only one residue in this region as having a high probability of binding, rhG-CSF Ser13 which is analogous to hGH Asn12 (Table 6).

(c) Conclusion

The structural similarity between hGH, which is required for the regulation of human growth and development, and G-CSF, which stimulates the terminal differentiation of neutrophils, raises the possibility that G-CSF initiates signal transduction by binding to two G-CSF receptors thus causing receptor dimerization. Both the binding probability function and the structural comparison of hGH and

G-CSF indicate that parts of helix E and the amino-terminal half of the AB crossover connection contribute to G-CSF receptor binding sites (Table 6, Figs 9 and 10). The disagreement between the binding probability function and the structural comparison of hGH and G-CSF in the 4-helical bundle region could be caused by the relatively reduced solvent accessibility that residues in this region have. As the primary sequence of G-CSF from additional species are characterized, it may be possible to construct binding probability functions that have better agreement with the G-CSF/hGH structural comparison results.

It is interesting to note that the only region that contributes to a hGH receptor binding site that is not structurally conserved in G-CSF is an exposed region of the crossover connection between helices 1 and 2 (hGH residues 60 to 68), a region that corresponds to the amino-terminal region of the AB crossover connection in G-CSF (G-CSF residues 63 to 69; Fig. 10). If this region does in fact contribute to a G-CSF receptor binding site, the structural variation of this region may confer the specific affinity hGH and G-CSF have for their respective receptors.

We thank NSF and NIH for support, Amgen for their donation of canine and bovine G-CSF and Abraham de Vos for the C^α coordinates of the hGH·(hGHbp)₂ complex. We also thank Matthias Wilmanns, Chris Hill and Tim Osslund for helpful discussions. The coordinates and structure factor amplitudes have been deposited with the Brookhaven Protein Database accession numbers: Bovine G-CSF, 1BGC; Canine G-CSF Form I, 1BGD; Canine G-CSF Form II, 1BGE.

References

- Abdel-Meguid, S. S., Shieh, H.-S., Smith, W. W., Dayringer, H. E., Vieland, B. E. & Bentle, L. A. (1987). Three-dimensional structure of a genetically engineered variant of porcine growth hormone. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 6434-6437.
- Bonilla, M. A., Gillio, A. P., Ruggerio, M., Kernan, N. A., Brochstein, J. A., Abboud, M., Famagalli, L., Vincent, M., Gabrilove, J. L. & Welte, K. (1989).

- Effects of recombinant human granulocyte colony-stimulating factor on neutropenia in patients with congenital agranulocytosis. *New Engl. J. Med.* **320**, 1574-1580.
- Brünger, A. T. (1990). Extension of molecular replacement: a new search strategy based on Patterson correlation refinement. *Acta Crystallogr. sect. A*, **46**, 46-57.
- Brünger, A. T., Krukowski, A. & Erickson, J. W. (1990). Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallogr. sect. A*, **46**, 585-593.
- Crawford, J., Ozer, H. & Johnson, D. (1990). Granulocyte colony stimulating factor: prevention of chemotherapy-induced febrile neutropenia (FN) in patients with small cell lung cancer (SCLC). A randomized double-blind, placebo-controlled trial. *Proc. Amer. Soc. Clin. Onc.* **9**, A884 (abstr).
- de Vos, A. M., Ultsch, M. & Kossiakoff, A. A. (1992). Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science*, **255**, 306-312.
- Diederichs, K., Boone, T. & Karplus, P. A. (1991). Low-resolution structure of recombinant human granulocyte-macrophage colony stimulating factor. *J. Mol. Biol.* **221**, 51-60.
- Ealick, S. E., Cook, W. J., Vijay-Kumar, S., Carson, M., Nagabhushan, T. L., Trotta, P. P. & Bugg, C. E. (1991). Three-dimensional structure of recombinant human interferon- γ . *Science*, **252**, 698-702.
- Eisenberg, D., Wesson, M. & Yamashita, M. (1989). Interpretation of protein folding and binding with atomic solvation parameters. *Chem. Scr.* **29A**, 217-221.
- Glaspy, J. A. & Golde, D. W. (1992). Granulocyte colony-stimulating factor (G-CSF): preclinical clinical studies. *Seminars in Oncology*, **19**, 386-394.
- Glaspy, J. A., Baldwin, G. C., Robertson, P. A., Souza, L., Vincent, M., Ambersley, J. & Golde, D. W. (1988). Therapy for neutropenia in hairy cell leukemia with recombinant human granulocyte colony-stimulating factor. *Ann. Intern. Med.* **109**, 789-795.
- Glaspy, J. A., Moreno, J., Mitsuyasu, R. et al. (1989). Recombinant granulocyte colony stimulating factor (rhG-CSF) recombinant erythropoietin (rHuEPO) and zidovudine (AZT) in patients with AIDS complicated by anemia and neutropenia. *Mol. Biother.* **1**, (suppl. 17).
- Hammond, W. P., IV, Price, T. H., Souza, L. M. & Dale, D. C. (1989). Treatment of cyclic neutropenia with granulocyte colony-stimulating factor. *New Engl. J. Med.* **320**, 1306-1311.
- Hanazono, Y., Hosoi, T., Kuwaki, T., Matsuki, S., Miyazono, K., Miyagawa, K. & Takaku, F. (1990). Structural analysis of the receptors for granulocyte colony-stimulating factor on neutrophils. *Exp. Hematol.* **18**, 1097-1103.
- Hendrickson, W. A. (1979). Transformations to optimize the superposition of similar structures. *Acta Crystallogr. sect. A*, **35**, 158-163.
- Hill, C. P., Osslund, T. D. & Eisenberg, D. (1993). The structure of granulocyte-colony stimulating factor (G-CSF) and its relationship to those of other growth factors. *Proc. Nat. Acad. Sci., U.S.A.* **90**, 5167-5171.
- Hodel, A., Kim, S.-H., Brügger, A. T. (1992). Model bias in macromolecular structures. *Acta Crystallogr. sect. A*, **48**, 851-858.
- Jakubowski, A. A., Souza, L. & Kelley, F. (1989). Effects of human granulocyte-colony stimulating factor in a patient with idiopathic neutropenia. *New Engl. J. Med.* **320**, 38-42.
- Jones, T. A. (1985). Interactive computer graphics: FRODO. *Methods Enzymol.* **115**, 157-171.
- Jones, T. A. (1990). O version 5 users manual, Department of Molecular Biology, BMC, box 590, S-75124, Uppsala, Sweden.
- Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577-2637.
- Klingeman, H.-G. (1989). Clinical application of recombinant human colony-stimulating factors. *Can. Med. Assoc. J.* **140**, 137-142.
- Kuga, T., Komatsu, Y., Yamasaki, M., Sekine, S., Miyaji, H., Nishi, T., Sato, M., Yokoo, Y., Asano, M., Okabe, M., Morimoto, M. & Itoh, S. (1989). Mutagenesis of human granulocyte colony stimulating factor. *Biochem. Biophys. Res. Commun.* **159**, 103-111.
- Lattman, E. E. (1972). Optimal sampling of the rotation function. *Acta Crystallogr. sect. B*, **28**, 1065-1068.
- Lee, B. & Richards, F. M. (1971). The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55**, 379-400.
- Matthews, B. W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491-497.
- McKay, D. B. (1992). Unravelling the structure of IL-2. *Science*, **257**, 410-413.
- Metcalf, D. (1985). The granulocyte-macrophage colony-stimulating factors. *Science*, **229**, 16-22.
- Molincux, G., Pojda, Z., Hampson, I. N., Lord, B. I. & Dexter, T. M. (1990). Transplantation potential of peripheral blood stem cells induced by granulocyte colony-stimulating factor. *Blood*, **76**, 2153-2158.
- Mott, H. R., Driscoll, P. C., Boyd, J., Cooke, R. M., Weir, M. P. & Campbell, I. D. (1992). Secondary structure of human interleukin-2 from 3D heteronuclear NMR experiments. *Biochemistry*, **31**, 7741-7744.
- Nomura, H., Imazeki, I., Oheda, M., Kubota, N., Tamura, M., Ono, M., Ueyama, Y. & Asano, S. (1986). Purification and characterization of human granulocyte colony-stimulating factor (G-CSF). *EMBO J.* **5**, 871-876.
- Oh-oda, M., Hasegawa, M., Hattori, K., Kubonwa, H., Kojima, T., Orita, T., Tomonou, K., Yamazaki, T. & Ochi, N. (1990). O-Linked sugar chain of human granulocyte colony-stimulating factor protects it against polymerization and denaturation allowing it to retain its biological activity. *J. Biol. Chem.* **265**, 11432-11435.
- Okabe, M., Asano, M., Kuga, T., Komatsu, Y., Yamasaki, M., Yokoo, Y., Itoh, S., Morimoto, M. & Oka, T. (1990). In vitro and in vivo hematopoietic effect of mutant human granulocyte colony-stimulating factor. *Blood*, **75**, 1788-1793.
- Pandit, J., Böhm, A., Jancarik, J., Halenbeck, R., Koths, K. & Kim, S.-H. (1992). Three-dimensional structure of dimeric human recombinant macrophage colony-stimulating factor. *Science*, **258**, 1358-1362.
- Powers, R., Garrett, D. S., March, C. J., Frieden, F. A., Gronenborn, A. M. & Clore, G. M. (1992). Three-dimensional solution structure of human interleukin-4 by multidimensional heteronuclear magnetic resonance spectroscopy. *Science*, **256**, 1673-1677.
- Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95-99.
- Rao, S. N., Jih, J. H. & Hartsuck, J. A. (1990).

- Rotation-function space groups (1980). *Acta Crystallogr. sect. A*, **36**, 878-884.
- Scarffe, J. H. & Kamthan, A. (1990). Clinical studies of granulocyte colony stimulating factor (G-CSF). *Cancer Surveys*, **9**, 115-130.
- Senda, T., Shimazu, T., Matsuda, S., Kawano, G., Shimizu, H., Nakamura, K. T. & Mitsui, Y. (1992). Three-dimensional crystal structure of recombinant murine interferon- β . *EMBO J.* **11**, 3193-3201.
- Silver, G. M., Gamelli, R. L. & O'Reilly, M. (1989). The beneficial effect of granulocyte colony-stimulating factor (G-CSF) in combination with gentamicin on survival after *Pseudomonas* burn wound infection. *Surgery*, **106**, 452-456.
- Serushago, B. A., Yoshikai, Y., Handa, T., Mitsuyama, M., Muramori, K. & Nomoto, K. (1992). Effect of recombinant human granulocyte colony-stimulating factor (rhG-CSF) on murine resistance against *Listeria monocytogenes*. *Immunology*, **75**, 475-480.
- Von Hoff, D. D., Clark, G. M., Weiss, G. R., Marshall, M. H., Buchok, J. B., Knight, W. A., III & LeMaistre, C. F. (1986). Use of in vitro dose response effects to select antineoplastics for high-dose or regional administration regimens. *J. Clin. Oncology*, **4**, 1827-1834.
- Walter, M. R., Cook, W. J., Zhao, B. G., Cameron, R. P., Jr, Ealick, S. E., Walter, R. L., Jr, Reichert, P., Nagabhushan T. L., Trotta, P. P. & Bugg, C. E. (1992). Crystal structure of recombinant human interleukin-4. *J. Biol. Chem.* **267**, 20371-20376.
- Weissman, L. J. (1979). Progress in the determination of the structure of crystalline rabbit muscle aldolase. Ph.D. dissertation. University of California, Los Angeles.
- Wlodaver, A., Pavlosvsky, A. & Gustchina, A. (1992). Crystal structure of human recombinant interleukin-4 at 2.25 Å resolution. *FEBS Letters*, **309**, 59-64.

Edited by I. Wilson