

UNIVERSIDAD NACIONAL DEL CALLAO  
FACULTAD DE INGENIERIA QUIMICA



710  
000035  
V35

“ELABORACION DE UN MODELO DE OPTIMIZACIÓN PARA LA  
SINTESIS DE PROCESOS DE PURIFICACIÓN  
DE PROTEÍNAS”

TESIS  
PARA OPTAR EL TITULO DE:  
INGENIERO QUIMICO

ref. 35



PRESENTADO POR  
Bach. ELSA VASQUEZ ALVAREZ

ASESOR: ING. RAYMUNDO CARRANZA NORIEGA

CALLAO - PERU  
ENERO, 1999



A

mis padres

Germán y Fidelina

quienes supieron guiarme

por el camino correcto, y siempre

estuvieron a mi lado con cariño y apoyo.

A mis hermanos Martha y Gilmer por su valiosa

ayuda para superarme; Daniel, Orlando y Félix por su

comprensión; Lily y Mercedes por su fe en mí, confianza y amor

A mis sobrinos Carlitos, los mellizos Dante, Eylen y J. Dieguito nuestra alegría.

## **AGRADECIMIENTOS**

Deseo expresar mi gratitud muy especial a mis asesores:

- Prof. Dr. José Mauricio Pinto del Departamento de Ingeniería Química de la Universidad de Sao Paulo (Brasil), quien ayudo con ideas y sugerencias en la realización de esta tesis.
- Ing. Raymundo Carranza Noriega profesor de la facultad de Ingeniería Química de la Universidad Nacional del Callao quien tuvo la amabilidad de revisar y ayudar en la culminación de esta tesis.

Agradecer a la Dra. M. E. Lienqueo de la Universidad de Chile por proveer de datos experimentales para los ejemplos. Agradezco también a mis compañeros y amigos del Laboratorio de Simulación y Control de Procesos (LSCP) de la Universidad de Sao Paulo (Brasil), quienes me apoyaron con material bibliográfico..

Por último aprecio la ayuda y comprensión de Jorge Casas L., el entusiasmo de toda mi familia y la voluntad de colaboración de todos aquellos que hicieron posible la presente tesis

# ÍNDICE GENERAL

<b>CAP. I.</b>	<b>INTRODUCCIÓN</b>	<b>10</b>
	1.1 Importancia	12
	1.2 Objetivos.	13
	1.3 Resumen	13
<b>CAP. II.</b>	<b>REVISIÓN BIBLIOGRÁFICA</b>	<b>17</b>
	2.1 Procesos Biotecnológicos	18
	2.1.1 Transporte de insolubles	19
	2.1.2 Separación de productos	19
	2.1.2.1 Producto Intracelular	19
	a) Ruptura de la célula	19
	b) Remoción de tejido	20
	c) "Refolding"	20
	2.1.2.2 Producto extracelular	20
	2.1.3 Purificación de alta Resolución.	21
	a) Cromatografía de interacción hidrofóbica	21
	b) Cromatografía de intercambio catiónico	21
	c) Cromatografía de intercambio aniónico	22
	d) Cromatografía por gel filtración	22

<b>2.2 Descripción de la obtención de las proteínas albúmina</b>	
de suero bovino, ecoli_s8 y somatotropina	23
<b>2.2.1 Albúmina de suero bovino</b>	23
<b>2.2.1.1 Método de obtención</b>	23
<b>2.2.2 Ecoli_s8</b>	24
<b>2.2.3 Somatotropina</b>	24
<b>2.2.3.1 Método de obtención</b>	25
<b>2.3 Metodologías para la solución de problemas de síntesis</b>	
de procesos	25
<b>2.3.1 Optimización de Procesos</b>	25
<b>2.3.1.1 Métodos de búsqueda del óptimo</b>	26
a) Representaciones	28
b) Método de Búsqueda	29
c) Estrategia de búsqueda	32
<b>2.3.1.2 Modelos de optimización</b>	34
a) Programación Lineal	34
b) Programación No Lineal	35
c) Programación Lineal Mixta Entera	35
d) Programación Mixta Entera No Linear	37
<b>2.3.1.3 Estrategia de Optimización</b>	37
a) Optimización Unidimensional sin	
Restricciones	38
b) Optimización Multivariable sin	

	Restricciones	39
	2.3.1.4 Softwares de optimización	39
	a) Lindo	40
	b) Gino	41
	c) GAMS	42
	2.3.2 Sistema especialista	45
	a) Elaboración del modelo	46
<b>CAP. III.</b>	<b>ANALISIS DEL PROBLEMA</b>	<b>48</b>
	3.1 Tiempo de retención	48
	3.2 Factor de desvío	49
	3.3 Factor de concentración	49
	3.4 Descripción	50
<b>CAP. IV.</b>	<b>MODELO DE OPTIMIZACIÓN</b>	<b>52</b>
	4.1 Símbolos	52
	4.2 Calculo del factor de concentración	53
	4.3 Modelo matemático	55
	4.4 Estructura del modelo	58
<b>CAP. V.</b>	<b>RESULTADOS</b>	<b>61</b>
	5.1 Purificación de "Serum from Bovine albumin"	61
	5.2 Purificación del "ecoli_s8"	66

5.3	Purificación de la "Somatotropina"	68
5.4	Comparación de Metodologías	71
<b>CAP.VI.</b>	<b>CONCLUSIONES Y RECOMENDACIONES</b>	<b>75</b>
6.1	Conclusiones	75
6.2	Recomendaciones	76

## **REFERENCIAS BIBLIOGRÁFICAS**

## **GLOSARIO**

## **ANEXOS**



## NOMENCLATURA

- $a$  propiedad de la proteína
- $B$  matriz  $m \times n$
- $C^T$  matriz transpuesta
- $C_{dp,i}$  concentración de la proteína deseada después de la técnica  $i$
- $C_{p,k}$  concentración del contaminante  $p$  antes de la técnica cromatográfica en la orden  $k$
- $C_{i,p,k}$  concentración de la proteína  $p$  después de la técnica  $i$  en la orden  $k$
- $C_{dp}^p$  nivel de pureza de la proteína deseada
- $C_p^u$  límite máximo de concentración de la proteína  $p$
- $C_p^o$  concentración inicial de la proteína  $p$
- $CF_{i,p}$  factor de concentración de la proteína  $p$  después de la técnica cromatográfica  $i$
- $dp$  proteína deseada
- $DF_{i,p}$  factor de desvío para la proteína  $p$  en la técnica cromatográfica  $i$
- $f(x)$  Función a ser minimizada
- $g(x)$  Función Lineal
- $h(x)$  Función lineal
- $i$  técnica cromatográfica
- $ip$  índice de la proteína
- $k$  orden en la secuencia en van a ser colocadas los pasos
- $Kd_{i,p}$  tiempo de retención para la proteína  $p$  en la técnica cromatográfica  $i$

- $M$  Limite Superior de la función Lineal
- $p$  proteína
- $P_{a,p}$  valor de la propiedad  $a$  para la proteína  $p$
- $S$  función a ser minimizada
- $y_{i,k}$  variable binaria que indica si la técnica cromatográfica  $i$  es usada en la orden  $k$  siguiendo la secuencia
- $\mathcal{Y}^D$  Conjunto de decisiones discretas
- $Z_k$  variable binaria que indica si la orden  $k$  es la ultima en la secuencia
- $\sigma_i$  Amplitud del pico de la técnica cromatográfica  $i$
- $\eta$  Factor de eficiencia
- $\theta$  Factor de concentración en el sistema especialista

# I. INTRODUCCIÓN

La biotecnología, es una área integrada de las ciencias como la biología, la bioquímica, química, microbiología y ciencias ingenieriles que se encarga de transformar el material biológico. Es una área desafiadora de actuación del Ingeniero de procesos, que se ocupa de desarrollarlo y llevarlo a escala industrial.

La transformación del material de origen biológico enmarca la fermentación y la recuperación (separación y purificación) del producto.

Seleccionar y diseñar un proceso de separación y purificación en la industria de la biotecnología, era muy difícil hace algunos años atrás, debido a la escasa documentación de las propiedades de los productos a ser separados así como también de un sistema de programación matemática que organice tal información.

La separación y purificación de una proteína fermentada son elementos críticos en los procesos biotecnológicos modernos, ya que estos representan el mayor costo de manufactura. Por lo tanto, la producción de proteínas depende no solamente de la innovación en la Biología molecular e inmunología, para la

síntesis y obtención de productos, sino también como innovación en el proyecto de las operaciones de separación y purificación.

Una alternativa de solución del problema de síntesis de procesos de purificación se auxilia de reglas heurísticas (basadas en la experiencia) que establecen las siguientes decisiones. (Asenjo.et. al., 1998):

- Definir los requerimientos (utilización) del producto final,
- Escoger los procesos de separación basado en sus propiedades físicas, químicas o bioquímicas,
- Separar las impurezas que se encuentran en mayor proporción, al inicio,
- Realizar las operaciones unitarias de alta resolución, inicialmente.

Por otro lado, la optimización de procesos se muestra como una herramienta eficiente tanto para el proyecto como para las operaciones unitarias de procesos químicos. En los últimos años, han surgido también importantes aplicaciones en procesos biotecnológicos (Bailey, 1998). Para el problema de síntesis de procesos debe ser definida una superestructura de operación, las cuales son representadas a través de un problema de programación mixta entera (Biegler et. al. , 1997)

En este trabajo, se desarrollará un modelo matemático de optimización de procesos de purificación que considere las propiedades físico-químicas de las proteínas presentes, las características de las etapas de cromatografía, las

decisiones de elegir y la secuencia de pasos de purificación. La solución del modelo será mediante el software de optimización denominado "Sistema de Modelage de Ecuaciones Algebraicas" (GAMS).

El problema trata la selección óptima de pasos de purificación desde una mezcla inicial, compuesta de un conjunto de proteínas, a un producto final de especificación predefinido.

## **1.1 IMPORTANCIA**

La selección de un eficiente proceso para purificar proteínas es uno de los mayores centros de interés en la industria bioquímica, desde que nuevas técnicas de ingeniería de procesos fueron desarrolladas.

La purificación de una proteína es un proceso muy importante y de mucho cuidado, ya que de esta etapa depende la calidad del producto final. Mas aún si se trata de productos medicinales como antibióticos y vacunas, donde se requiere puridades de hasta 99% o 99.9 %, es por eso que se recomienda el uso las operaciones cromatograficas.

## 1.2 OBJETIVOS

El presente trabajo pretende alcanzar los siguientes objetivos:

### Objetivo principal

- Solucionar el problema de síntesis de procesos de purificación de proteínas, mediante un modelo de programación matemática que genere soluciones alternativas a las de un sistema especialista.

### Objetivos secundarios

- Seleccionar las etapas de purificación, basado en valores cuantitativos de propiedades físico-químicas de la proteína de interés y de las proteínas contaminantes.
- Comparar la calidad de solución generada por las dos metodologías.

## 1.3 RESUMEN

La importancia de la etapa de purificación en un proceso biotecnológico sirve para desarrollar la optimización de éste, teniendo en cuenta las técnicas cromatográficas a ser usadas en la purificación de sistemas bioquímicos de multicomponentes.

Los principales descubrimientos en la síntesis de las etapas de purificación de proteínas, es la selección y el ordenamiento apropiado de las etapas del análisis cromatográfico.

En la presente tesis, después de desarrollar el capítulo introductorio, se presenta los procesos biotecnológicos que se llevan a cabo después de la fermentación, los que están divididos en tres grupos:

- a) Transporte de insolubles, donde es seleccionado la parte en la cual se encuentra nuestro producto de interés.
- b) Separación de productos, esta etapa incluye las operaciones de ruptura de la célula (sí el producto es intracelular) remoción de tejido y "refolding" (reflujo; si presenta cuerpos incrustados). El producto remanente es una mezcla de proteínas, incluyendo la proteína de interés, las cuales tiene propiedades físico químicas similares.
- c) En la etapa de purificación, todas las proteínas son sometidas a una secuencia de operaciones cromatográficas a fin de alcanzar el nivel de pureza deseado para el producto.

También en el capítulo II se trataran las metodologías para la solución de problema de síntesis de procesos, la primera de ellas es la optimización de procesos en la que presentamos los diferentes problemas de optimización como son programación lineal(LP), no lineal

(NLP), programación mixta entero lineal (MILP) y programación mixta entera no lineal (MILNP); el segundo método es mediante un sistema especialista, este método combina reglas y relaciones matemáticas que manipula una base de datos para seleccionando y dando una secuencia de operaciones para la purificación de proteínas.

En el capítulo III hacemos un análisis del problema, como éste es resuelto y las comparaciones que se hacen entre las propiedades, hasta obtener el factor de concentración que es el que nos va indicar el grado de separación de cada proteína.

El capítulo siguiente abarca el modelo matemático, que inicia mostrando los índices, las variables y los parámetros, posteriormente se describen con relaciones matemáticas el cálculo del factor de concentración, a continuación se presentan cuatro formas de restricciones y la función objetivo a ser minimizada. El modelo resultante es un problema MILP que es resuelto con la ayuda del software denominado (GAMS).

En el capítulo de resultados, el modelo es probado en tres ejemplos, conteniendo hasta 13 contaminantes y 22 pasos cromatográficos, para luego ser comparados con los resultados obtenidos con el método de un sistema especialista (resultados proporcionados por el centro para la Ingeniería Bioquímica y



Biotecnológica - Universidad de Chile), para finalmente presentar las conclusiones y recomendaciones en el último capítulo.

## II. REVISIÓN BIBLIOGRÁFICA

En la purificación de proteínas de elevados porcentajes de pureza (99 – 99.9) se realizan innumerables ensayos cromatográficos en laboratorio, hasta obtener la secuencia ideal de purificación. La aplicación de múltiples ensayos cromatográficos se debe a que las proteínas son complejas, es por eso que el proceso de purificación está siendo ahora estudiada previamente con herramientas que minimizan el número de técnicas a aplicar, una de ellas es el trabajo de investigación realizado en esta tesis.

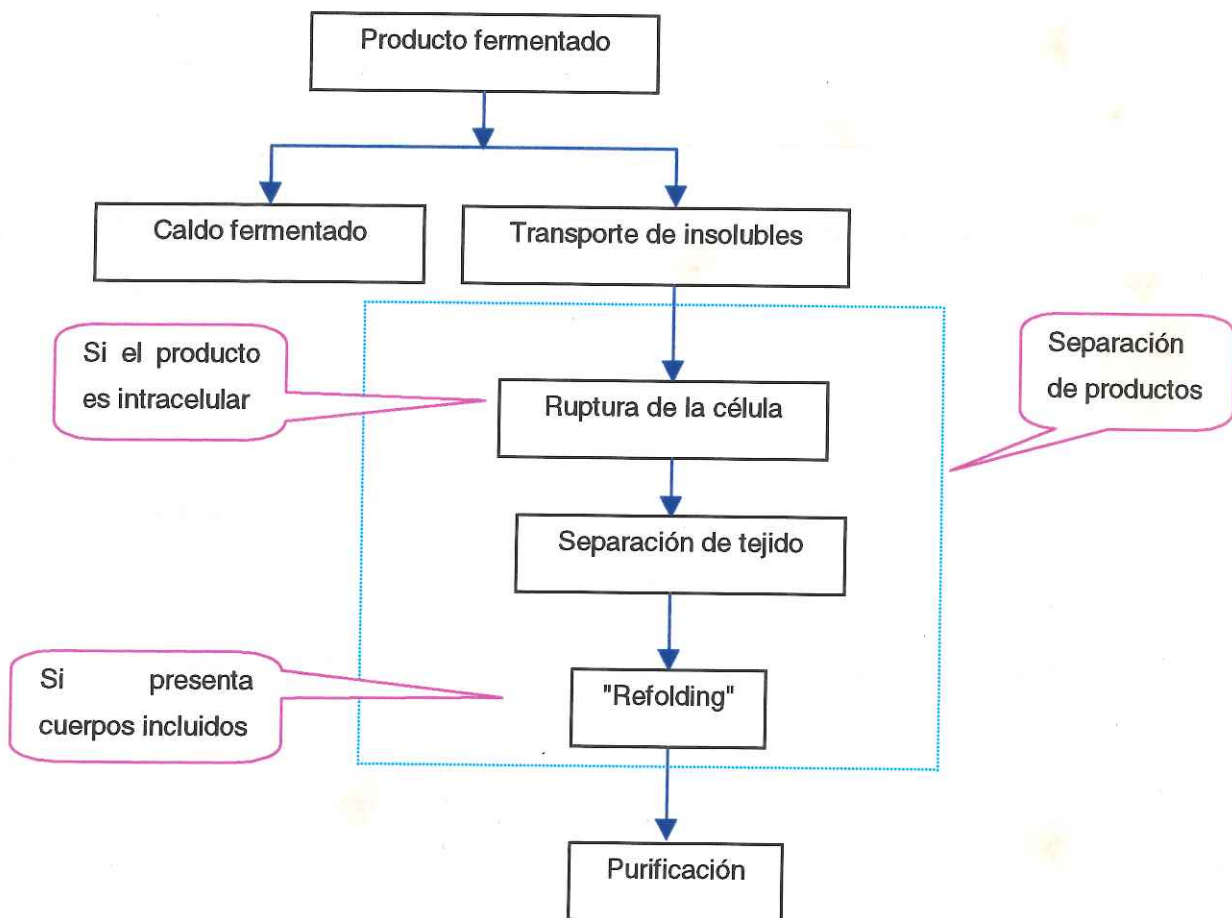
Para un mejor entendimiento definiremos algunos términos usados frecuentemente en la presente tesis.

- **Proteínas.**- Esencialmente constituida de polímeros de aminoácidos (aa) y otras sustancias más o menos complejas. Consta de carbono, hidrógeno, oxígeno y nitrógeno, a veces también de azufre, fósforo y yodo. Pueden ser albúminas o globulinas [4].
  
- **Síntesis.**- Composición de un todo por la reunión de sus partes. Generación y selección de alternativas para la integración de sistemas, compuesto de componentes conocidos o existentes, para especificaciones funcionales de un sistema.

- **Síntesis de procesos.**- Selección sistemática de configuraciones de procesos [10].

## 2.1 PROCESOS BIOTECNOLÓGICOS

Los procesos biotecnológicos denominados "procesos de corriente interna", pueden describirse fácilmente en el diagrama de bloques que se muestra a continuación.



**Figura 1 Procesos Biotecnológicos**

Fuente: (Asenjo, 1990)<sup>[2]</sup>

### **2.1.1 Transporte de insolubles**

Es la etapa previa al proceso de separación en la industria de la biotecnología, en esta etapa es separada del caldo fermentado todo el material sólido al que llamaremos mezcla de proteínas y en la que se encuentra la proteína a ser purificada.

### **2.1.2 Separación de productos**

En este proceso se llevan a cabo en varias etapas dependiendo de cómo la proteína fue obtenida, puede ser intracelular o extracelular. El objetivo de estas etapas es recuperar el producto en solución después de la fermentación.

#### **2.1.2.1 Producto intracelular.**

**a) ruptura de la célula.-** Si la proteína de interés se encuentra dentro de la célula, en otras palabras si el producto es intracelular, la célula es rota, para esto existe mecanismos de rompimiento de célula como: mecánicos o físicos (molino de bolas y homogenizador, para escala industrial; presión osmótica, para escala de laboratorio), químicos (mediante detergentes,

solventes y lisis enzimática). La aplicación de estos mecanismos va a depender del tipo de producto.

- b) **remoción de tejido**.- Una vez que la célula fue rota, partes de tejido celular que se encuentran formando parte de la mezcla tiene que ser separada; dependiendo del producto; existe mecanismos de separación o de remoción que se llevan a cabo, estos pueden ser: filtración (tangencial), precipitación y centrifugación.
  
- c) **"Refolding"**.- (reflujo) Sólo se aplica si después de separar el tejido celular, quedan cuerpos incluidos. La finalidad de este proceso es concentrar la mezcla de proteínas.

#### **2.1.2.2 Producto extracelular.**

Cuando el producto es extracelular, la separación se lleva a cabo de tal manera que no dañe la estructura del producto, se debe evitar usar mecanismos mecánicos.

### 2.1.3 Purificación de alta resolución.

Este proceso es la más importante porque de ella depende la calidad del producto final. Para realizar este proceso se tiene varias alternativas que se usan en secuencia, a fin de obtener una alta pureza con un número mínimo de técnicas.

Las técnicas cromatográficas; que consisten en la separación diferencial entre la fase móvil y la fase estacionaria; son del tipo columna y en la mayoría de aplicaciones, la fase estacionaria consiste de partículas esféricas las cuales están dentro de la columna. Las técnicas cromatográficas que se usan para el desarrollo del modelo se describen a continuación:

- a) **Cromatografía por interacción hidrofóbica (HIC).**- Esta técnica cromatográfica se usa cuando los aminoácidos son hidrofóbicos donde ocurren elevadas fuerzas iónicas. Esta compuesta de una columna empaquetada (puede ser un fenil), buffer y un equipo de filtración. Las proteínas son disueltas en buffer
  
- b) **Cromatografía de intercambio catiónico.**- Grupos "-" donde el punto isoeléctrico (PI) es mayor que el pH. Consiste de una columna y un buffer. En este proceso las moléculas del soluto se

separan basándose en las diferencias de comportamiento ácido base. Para este proceso la columna se llena con una resina sintética de poliestireno sulfonada que contiene grupos cargados fijos. Se hace pasar una serie de disoluciones salinas de diferente fuerza iónica.

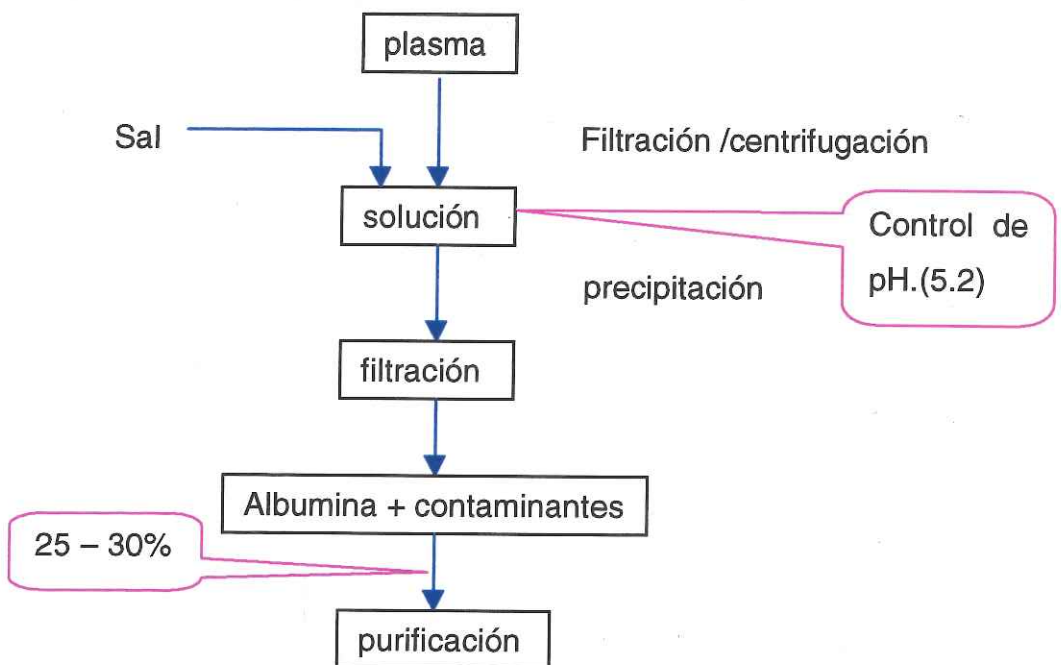
- c) ***Cromatografía de intercambio aniónico.***- Grupos "+" en donde el pH es mayor que el pI, esta compuesto por una columna y un buffer. La columna contiene una resina dieteilaminoetil celulosa que contiene grupos cargados fijos a pH 7.0. Se consigue la resolución de las mezclas de las proteínas y la elución sucesiva de los componentes de la columna, haciendo pasar una serie de tampones a diferente pH.
- d) ***Cromatografía por gel filtración.***- Consiste de una columna empaquetada con gel porosa y dependiente del gradiente del peso molecular ira acumulándose en la columna. Este método se basa en una propiedad biológica de algunas proteínas: capacidad de unión específica, en el enlace covalente con otra molécula llamada ligando. La proteína unida específicamente a las partículas de la columna es después eluida, frecuentemente con una disolución de la molécula de ligando libre.

## 2.2 DESCRIPCIÓN DE LA OBTENCIÓN DE LAS PROTEÍNAS ALBÚMINA DE SUERO BOVINO, ECOLI\_S8 Y SOMATOTROPINA

### 2.2.1 Suero de Albúmina bovino.

Se encuentra en el plasma y en alto porcentaje en el ganado joven, se usa para fabricar vacunas, un estudio reciente indica que produce un síndrome clínico similar a la encefalopatía.

**2.2.1.1 Método de obtención.-** La separación de las proteínas del plasma se muestra en el siguiente diagrama de bloques.



**Figura 2. Obtención de suero de albúmina bovino**

Fuente: (Lienqueo,1998)



El plasma se centrifuga luego se somete a una combinación de pHs, se fracciona con sal, luego se realiza un ultrafiltrado del producto por diafiltración para finalmente someterla a una purificación de alta resolución y posteriormente secado. La preparación final se lleva a cabo bajo estrictas medidas de higiene con la finalidad de evitar la introducción de pyrogenos.

### **2.2.2 Ecoli\_s8.**

Es una proteína proveniente de la fermentación de la Escherichia coli, esta proteína se encuentra en animales y humanos, su obtención y caracterización es similar a la Somatotropina. Actualmente se encuentra en estudio.

### **2.2.3 Somatotropina.**

Es una proteína hormonal (compleja) generalmente producida en el ganado (bST) por la glándula pituitaria localizada en la base del cerebro de los animales. En los humanos también se encuentra presente.

Esta proteína es inyectada en las vacas para aumentar la producción de leche, en los humanos también influye en las glándulas mamarias, pero su mayor aplicación se observa en el crecimiento.

**2.2.3.1 método de obtención.-** Se retira el gen del ganado y se inserta en una bacteria llamada *Escherichia coli*. Esta bacteria que se encuentra en el tracto intestinal de humanos y animales actúa como una fábrica diminuta y produce cantidades grandes de somatotropina en condiciones controladas de laboratorio. La somatotropina producida por la bacteria es purificada, mediante los procesos de purificación de alta resolución para posteriormente ser inyectado en la glándula mamaria de las vacas.

## **2.3 METODOLOGÍAS PARA LA SOLUCIÓN DE PROBLEMAS DE SÍNTESIS DE PROCESOS**

### **2.3.1 Optimización de procesos**

La optimización de procesos se encarga de formular y modelar problemas para la optimización, se aplica en cualquier

proceso productivo, donde se determina la mejor solución de un modelo.

La optimización esta relacionada con la matemática aplicada (computación), investigación operacional, ingeniería y economía.

Los problemas de síntesis de procesos son de naturaleza combinatoria, desde que se ven involucrados decisiones discretas (tienen los valores 0 o 1). Esta compuesta de los siguientes elementos:

- *Función Objetivo.*- indicador cuantitativo de la solución, por ejemplo: costo (inversión, operación), lucro, productividad, etc.

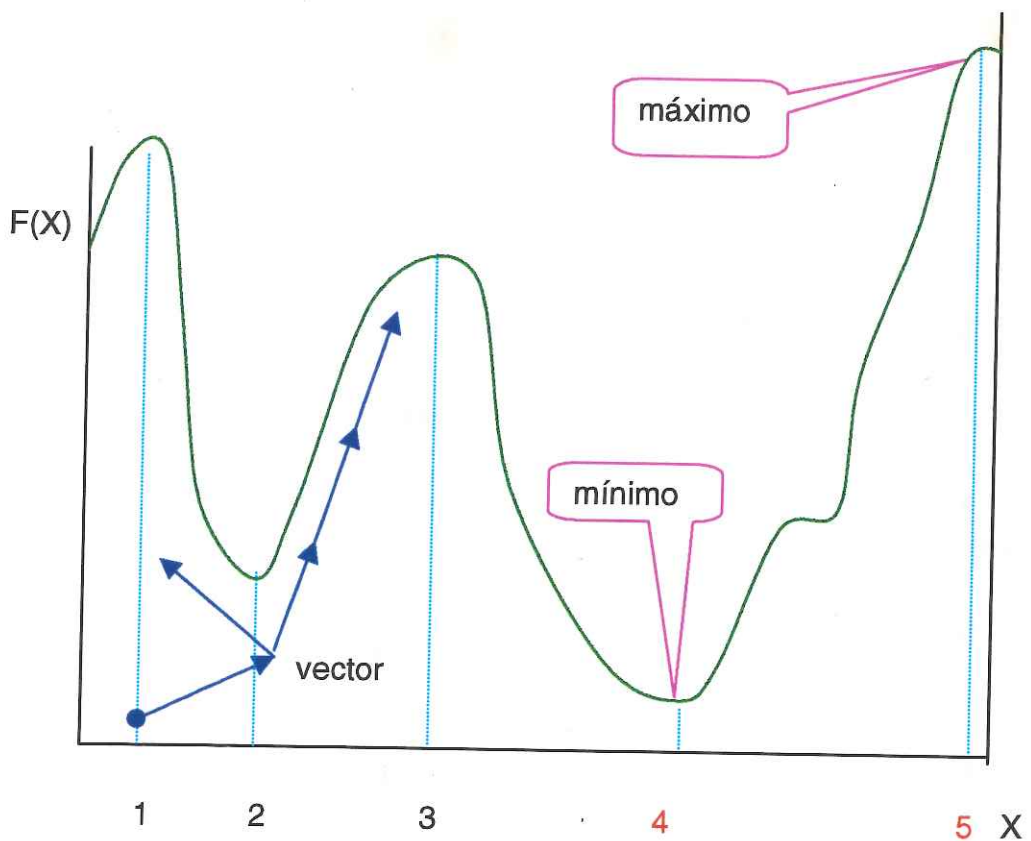
- *Variables.*- independientes o dependientes, por ejemplo: condiciones de operación (temperatura, presión, flujo, etc.), dimensiones del equipo

- *Restricciones.*- son inecuaciones o ecuaciones que indica por ejemplo los límites de operación, pureza (especificación del producto)

#### **2.3.1.1 Métodos de búsqueda del óptimo**

Este método se basa en la búsqueda de una solución óptima o un camino óptimo. En la optimización de procesos, la mejor solución es buscando el punto máximo o

mínimo, esta búsqueda se realiza linealmente (ver figura 3) mediante el método ramificado para problemas mixto enteros. Si la busca es lineal se representa mediante un vector el que tiene una dirección y un tamaño de procura, la cual puede variar si uno lo requiere y através de esta búsqueda el vector puede perderse (no llegar a la mejor solución) es por eso que se le da restricciones. El método de busca ramificada por ser el método aplicado en el desarrollo de esta tesis se estudia en detalle mas adelante.



**Figura 3 Representación del método de busca lineal.**

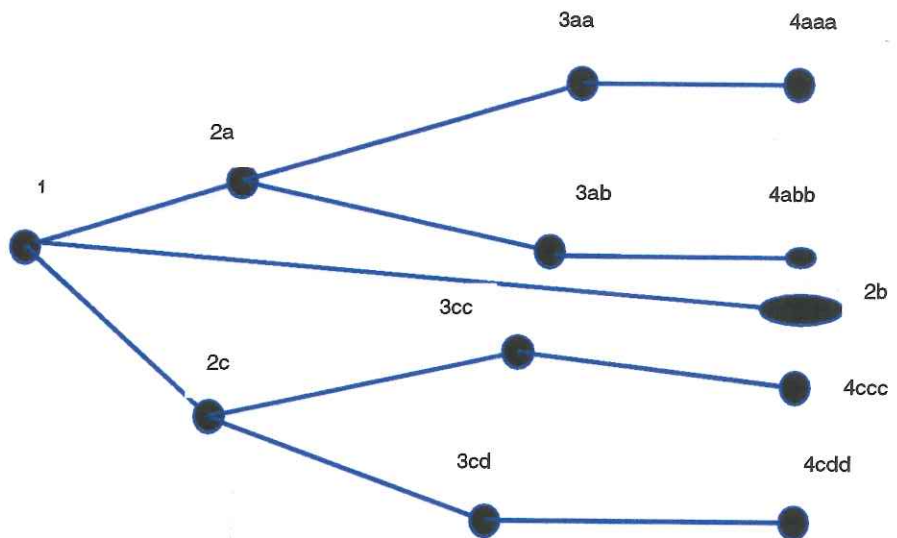
Fuente: (Himmemblau, 1989)<sup>[5]</sup>

**a) Representaciones**

La búsqueda a través de un sistema mixto entero se representa en forma de árbol y red los que tienen los siguientes componentes:

- nudos.- componentes principales y
- arcos.- interconexiones principales

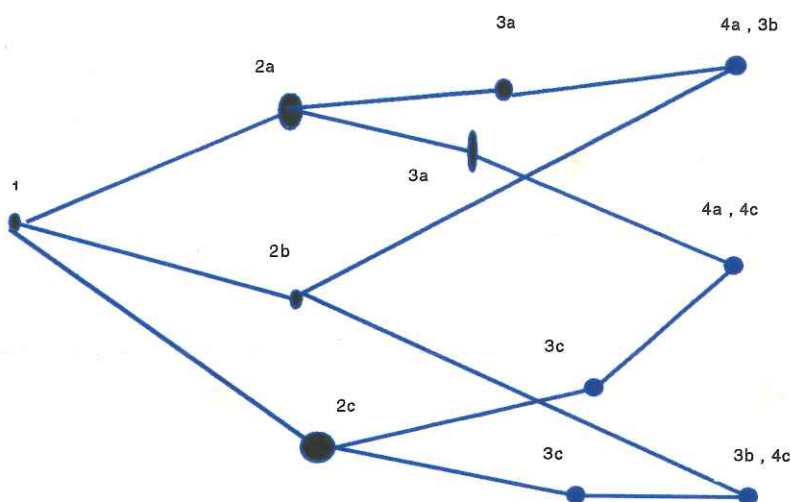
a.1) Árbol (trees).- Representación sin ciclos. Tiene estructura en serie, cada nudo conectado por lo menos a otro nudo. Las principales características son: su representación es implícita de alternativas, la generación de alternativas es por descomposición y la optimización requiere separabilidad y monotonidad.



**Figura 4 Representación en árbol**

Fuente: (Pinto J. M., 1998)<sup>[10]</sup>

a.2) Red (networks).- puede presentar ciclos. Arcos y nudos poseen variables asociados a ellos. Sus principales características son: representación explícita de alternativas, generación de alternativas obtenidas por superposición y optimización hecha simultáneamente.



**Figura 5 Representación en red.**

Fuente: (Pinto J. M., 1998)<sup>[10]</sup>

**b) Método de búsqueda**

La siguiente función representa un problema típico de Optimización:

$$\min Z = \sum_{i=1}^n C_i(y_i) \tag{A}$$

s.a.

$$\sum_{i=1}^n B_{ij}(y_i) \leq \alpha_j \quad j = 1, \dots, m \tag{B}$$

$$y_i \in y_i^D$$

donde:

$y^D$  : conjunto de decisiones discretas

$$a_i \geq 0$$

$$C_i(y_i), B_{ij}(y_i) \geq 0 \quad \text{condición de monotonicidad} \quad (C)$$

Para visualizar este problema se analizara mediante el siguiente ejemplo (ver también estrategia de busca)

Si:

$$y_i = \begin{cases} 1 & \text{escoje el componente } i \text{ del tipo } 1 \\ 2 & \text{escoje el componente } i \text{ del tipo } 2 \end{cases}$$

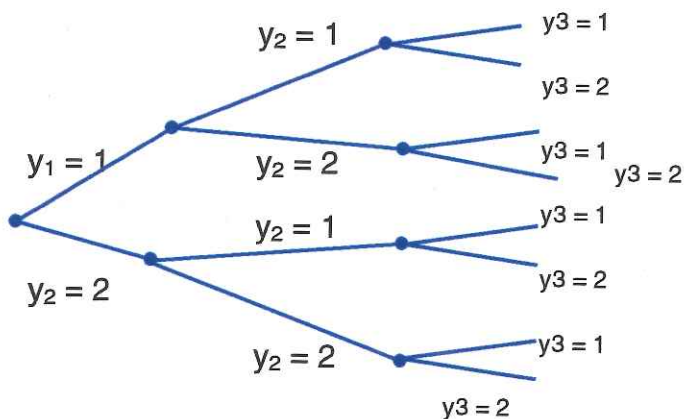
Tenemos que:

$$\text{Niveles de expansión} = n = 3$$

$$\text{El numero de soluciones alternativas} = 2^n = 2^3$$

$$\text{El número de nudos} = 2^{(n+1)} - 1 = 15$$

Esto se observa en la siguiente figura.



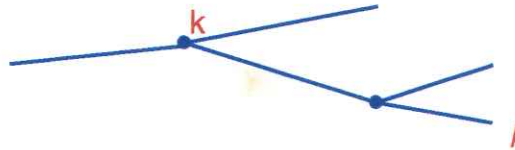
**Figura 6 Método de busca**

Fuente: (Pinto J. M., 1998)<sup>[10]</sup>

como determinamos la solución optima del problema?

Teniendo en cuenta los siguientes puntos:

1. Enumeración exhaustiva: examinar todas las posibilidades. Es un método restringido a problemas de baja dimensión
2. Enumeración dinámica: ("Rama e union"). considera los nudos  $k$  y  $l$ , donde  $l$  es descendente de  $k$  (ver figura 7).



**Figura 7 enumeración dinámica**

Fuente: (Pinto J.M., 1998)<sup>[10]</sup>

Entonces se genera dos sub problemas:

sub problema parcial  $p^k$

$$Z^k = \sum_{i=1}^k C_i(y_i) \quad (D)$$

s.a.

$$\sum_{i=1}^k B_{ij}(y_i) \leq \alpha_j \quad j = 1, \dots, m \quad (E)$$

sub problema parcial  $p^l$

$$Z^l = \sum_{i=1}^l C_i(y_i) \quad (F)$$

s.a.

$$\sum_{i=1}^l B_{ij}(y_i) \leq \alpha_j \quad j = 1, \dots, m \quad (G)$$

Dados estos dos sub problemas  $p^k$  y  $p^l$ , se tiene las siguientes propiedades:

- 1) Si  $p^k$  es inviable, entonces  $p^l$  también es inviable
- 2) Si  $p^l$  fuera viable entonces  $p^k$  es viable y  $Z^k \leq Z^l$

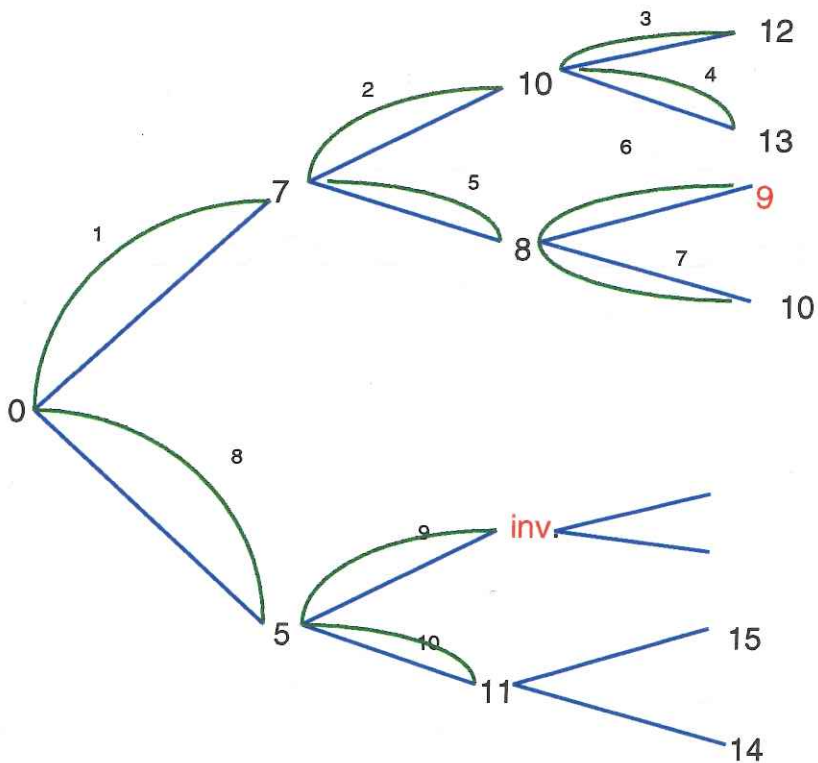


- 3) Si  $p^n$  fuera viable entonces se tiene una solución viable para el problema de valor  $Z^n$  (límite superior)

**c) Estrategia de busca**

Se basa en la busca ramificada, esta estrategia es aplicada por el software para resolver el modelo de optimización generado.

C.1) *Primera longitudinal (Degth First)*. - El principio de estrategia de busca es expandir la rama con el nudo

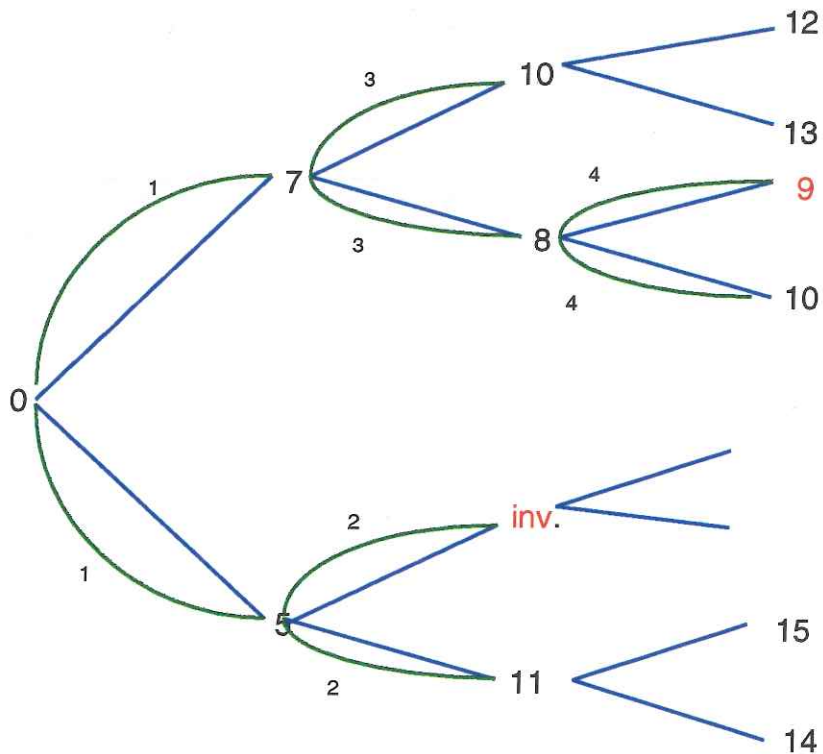


**Figura 8 estrategia de busca "Degth First"**

Fuente: (Pinto J.M, 1998)<sup>[10]</sup>.

abierto mas reciente (7), (10) y (12) luego (13), regresamos hasta (7) avanzamos por (8) y asi sucesivamente, logrando determinar que el nudo (9) es la respuesta por ser el que tiene el menor valor de todos los nudos finales. Notese que fueron examinados 10 nudos, de los 14, debido a que a partir del nudo 11 no se continúa por ser mayor que los nudos analizados y tenemos una inviabilidad (ver (b) enumeración dinámica).

c.2) *Primera Anchura (Breadth First)*.- Normalmente envuelve la busca por el menor número de nudos, es en paralelo, primero analiza el nudo (5) por ser menor, luego



**Figura 9 estrategia de busca "Breadth First"**

Fuente: (Pinto J.M., 1998)<sup>[10]</sup> .

encontramos una inviabilidad y el nudo (11), pero observamos que tenemos un nudo de valor (7) continuamos el análisis por ese lado hasta llegar al mismo valor del nudo (9), notese que sólo fueron examinados 8 nudos; la desventaja de esta estrategia es que exige mayor capacidad de almacenamiento (mayor memoria).

### 2.3.1.2 Modelos de optimización

Los modelos de optimización son: programación lineal (LP), programación no lineal (NLP), programación mixta entera lineal (MILP) y mixta entera no lineal (MILNP).

**a) programación lineal (LP).**- Un problema es llamado de LP si la función objetivo y las restricciones son lineales, para todas las variables continuas. La estructura general de un modelo de LP es:

$$\text{Min } f(x) = C^T X \quad (\text{H})$$

s.a

$$h(x) = A_1 X - b_1 = 0 \quad (\text{I})$$

$$g(x) = A_2 X - b_2 \leq 0 \quad (\text{J})$$

$$\forall x \in R^n$$

La solución de este modelo es aplicando el algoritmo simplex [11].

**b) programación no lineal (NLP).**- Es cuando la función objetivo es lineal y por lo menos una de las restricciones es no lineal, para todas las variables continuas. La solución de este modelo es aplicando algoritmo SIMPLEX, gradiente reducido o el de Newton [5].

**c) programación lineal mixta entera (MILP).**- Se da cuando tiene variables continuas y discretas, la función objetivo y las restricciones son lineales. Se dice entero porque se tiene un número fijo de respuestas (técnicas cromatográficas), mixto lineal porque la búsqueda es mediante ramificaciones (ver pags. 32 y 33), el trabajo de tesis es desarrollado en este modelo y tiene la siguiente estructura:

$$\min f(y,u,x) \quad (K)$$

s.a.

$$g(y,u,x) \leq 0 \quad (L)$$

$$h(y,u,x) = 0 \quad (M)$$

$$y \in y^D \quad D: \text{variable discreta}$$

$$u \in u^P \quad P: \text{problema}$$

f,g,h: son funciones lineales

**c.1 modelaje con variables 0 - 1.**- Generaciones de restricciones lineales

$$y_j = \begin{cases} 1 & \text{si } j \text{ es seleccionado} \\ 0 & \text{caso contrario} \end{cases}$$

### - RESTRICCIONES DE MULTIPLE ELECCIÓN

a) Seleccionar por lo menos 1

$$\sum_{j=1}^n Y_j \geq 1 \quad (\text{N})$$

b) Seleccionar exactamente 1

$$\sum_{j=1}^n Y_j = 1 \quad (\tilde{\text{N}})$$

c) Seleccionar como máximo 1

$$\sum_{j=1}^n Y_j \leq 1 \quad (\text{O})$$

### - IMPLICACIONES

a) Si - entonces (si k seleccionado entonces seleccione j):

$$Y_k \rightarrow Y_j$$

$$Y_k - Y_j \leq 0$$

$$Y_k = 0 \rightarrow Y_j \geq 0$$

$$Y_k = 1 \rightarrow Y_j \geq 1$$

b) Si y solamente si  $Y_k \leftrightarrow Y_j$

$$Y_k - Y_j = 0$$

$$Y_k = Y_j$$

### - RESTRICCIONES MIXTAS

a) Si i seleccionado,  $g(x) \leq 0$  debe ser aplicada.

$$g(x) \leq M(1 - Y_i) \quad (\text{P})$$

$$Y_i = 1 \rightarrow g(x) \leq 0 \quad (\text{Q})$$

$$Y_i = 0 \rightarrow g(x) \leq M, \quad (\text{M limite superior para } g(x))$$

b) Restricciones disyuntivas

$$f_1(x) \leq 0 \quad \text{o} \quad f_2(x) \leq 0 \quad (\text{R})$$

$$E - M \cdot Y_1 \leq f_1(x) \leq M \cdot (1 - Y_1) \quad (\text{S})$$

$$E - M \cdot Y_2 \leq f_2(x) \leq M \cdot (1 - Y_2) \quad (\text{T})$$

$$Y_1 + Y_2 = 1 \quad (\text{U})$$

$$\text{Si } Y_1 = 1 \rightarrow E - M \leq f_1(x) \leq 0 \quad (\text{V})$$

$$E \leq f_2(x) \leq M \quad (\text{W})$$

$$E : \quad \text{número "pequeño"}$$

Teniendo el modelo, la solución se realiza por los métodos: "Branch & Bound" (busca en árbol ver sección **2.3.1.1**) o "Branch & Cut" (este método aun esta en estudio) [11].

**d) Programación mixta entera no lineal.-** Es cuando tiene variables continuas y discretas, la función objetivo es lineal y por lo menos una de las restricciones es no lineal. La solución de este modelo es por descomposición de Benders [11].

### **2.3.1.3 Estrategias de optimización**

Las estrategias de optimización son seis: Optimización unidimensional sin restricciones, estrategia multivariables sin restricciones, estrategia multivariables

com restricciones (programación no lineal), programación lineal, programación mixta entera lineal y programación mixta entera no lineal. Aca sólo trataremos brevemente las dos primeras, las otras cuatro estrategias se enfocaron anteriormente (ver sección **2.3.1.2**).

**a) Optimización Unidimensional sin restricciones (busca en una línea)**

motivación: (ver figura 3)

- \* Reducción de problemas multivariantes a monovariantes (por sustitución de la restricción en la función objetivo)
- \* Problemas inherentemente monovariantes
- \* Aplicación en algoritmos multivariantes con o sin restricción como sub problema

Tipos de métodos: [11]

Los métodos de solución de esta estrategia sólo se indica a manera de información.

a) Método de reducción del intervalo de busca

dividido en dos fases:

1. fase de definición de los límites
2. fase de refinamiento del intervalo
  - Método de la Bisección
  - Método de la sección

b) Método basado en derivadas

1. Newton - Raphson
2. Secante

c) Método de aproximación polinomial

1. Método de aproximación cuadrática
2. Método de aproximación cuadrática sucesiva
3. Método de aproximación cúbica

**b) Optimización multivariable sin restricciones**

basado en dos grandes clases [11]:

*b.1. Método de busca directa.*- Útiles para  $f(x)$  o  $\nabla f(x)$  discontinuas

- Enumeración exhaustiva
- Optimización evolucionaria
- Método de busca simplex
- Método simplex modificado (Nelder & Mead)
- Método de Hooke y Jeeves
- Método de las direcciones conjugadas de Powell

*b.2. Métodos basado en el gradiente.*- Es mas eficiente:

- Método de Newton
- Método genérico
- Método de pasos descendientes
- Método de Marquardt
- Método cuasi-Newtonianos

**2.3.1.4 Softwares de optimización**

Los softwares de optimización son: Programación lineal (MATLAB, Lindo y GAMS), programación no Lineal



(MATLAB, Gino y GAMS), MILP (Lindo y GAMS) y en MINLP GAMS; a continuación se describe alguno de ellos:

**a) Lindo**

El programa lindo es un software iterativo para la solución de problemas LP("linear programming" o programación lineal) y MILP ("mixed integer-linear programming" o programación mixta entera lineal). Para accezar al programa, basta, en el "prompt" del MS-DOS, digitar LINDO <ENTER>. A continuación sigue una lista con los comandos básicos:

- HELP: Muestra informaciones básicas de su uso;
- COM: Lista todos los comandos;
- CAT: Lista el comando por categorías;
- MAX o MIN: Habilita la entrada de la función objetivo. Luego a seguir, el comando ST (sujeto a) debe ser colocado, para habilitar la entrada de las restricciones del problema.
- END: Especifica el fin de la formulación del problema;
- INT: Especifica variables enteras (0 o 1);
- GO: Soluciona el problema previamente especificado
- QUIT: Sale del programa;
- LOOK ALL: Lista la formulación activa en el programa (previamente digitada);
- ALT: Altera una ecuación / variable anteriormente digitada
- APPC: Adiciona una variable o línea;
- DEL: Excluye una línea;

- DIVE: Apunta un archivo como destino de la solución del problema
- TAKE: Efectúa una lectura de un contenido en un archivo (este necesita de un comando "LEAVE" al final del archivo).

### **b) Gino**

Así como el software LINDO, GINO ofrece un ambiente interactivo para la ejecución en plataformas IBM-PC, pero este se destina a la solución de problemas NLP. La solución de los problemas es realizada a través de un método de gradiente reducido generalizado implementado por Abadie e Lasdon. Para acceder al programa, basta, en el "prompt" del MS-DOS, digitar GINO <"ENTER">. Sus principales comandos son listados a continuación:

- HELP: Muestra informaciones básicas sobre la utilización del programa;
- COM: Lista todos los comandos;
- MODEL: Indica el inicio de la entrada de un problema. Después de digitarlo, el signo de "prompt" ? aparecerá. Las expresiones algebraicas a ser utilizadas son: (+) adición, (-) sustracción, (\*) multiplicación, (/) división, (^) exponenciación, (LOG(X)) logaritmo de X, (EXP(X)) función exponencial de X, (=) igual, (<) menor que, (>) mayor que;
- MAX= o MIN=: Habilita la entrada de la función objetivo.
- END: Especifica el fin de la formulación del problema;
- SLB X: Configura el límite inferior de la variable X;

- SUB X: Configura el límite superior de la variable X;
- GUESS ALL: Habilita la entrada de las condiciones iniciales. Si no fueran previamente digitadas, las C.I. serán consideradas nulas para todas las variables.
- LOOK ALL: Lista la formulación activa en el programa (previamente digitada);
- GO: Soluciona el problema previamente especificado;
- QUIT: Sale del programa;
- ALT: Altera una ecuación / variable anteriormente digitada;
- EXT: Habilita la adición de líneas en el modelo;
- TAKE: Efectúa una lectura de un contenido en un archivo (este necesita de un comando "LEAVE" al final del archivo).
- SETP: Configura los parámetros de salida o de la solución. Las opciones disponibles son listadas utilizándose el comando "HELP SETP".

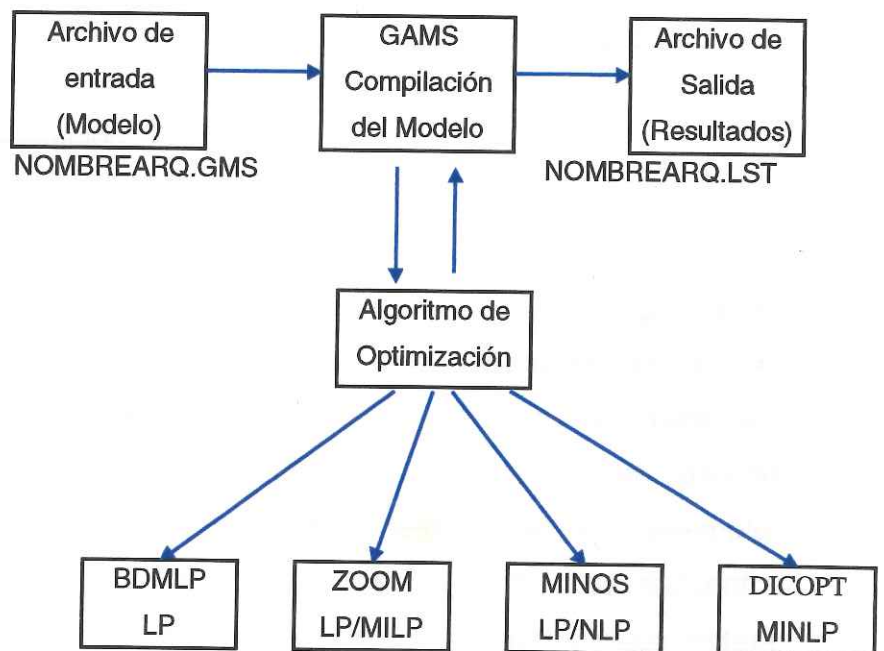
### **c) GAMS**

Al contrario de los otros softwares anteriormente citados, el GAMS no posee interacción con el usuario, por eso este es el sistema para optimización que ofrece una serie de algoritmos y resuelve problemas lineales (LP), problemas no lineales (NLP), problemas mixto entero lineales (MILP) y problemas mixtos entero no lineales (MINLP). Los algoritmos disponibles para cada problema son los siguientes:

- LP: ZOOM, MINOS 5.2, SCICONIC
- MILP: ZOOM, SCICONIC

- NLP: MINOS 5.2, SQP
- MINLP: DICOPT++

Los modelos son dados por el usuario a través de un archivo de entrada en la forma de ecuaciones algebraicas, utilizando un lenguaje adecuado para el programa. GAMS compila las informaciones (detecta posibles errores de sintaxis) e soluciona el problema a través de un algoritmo predefinido, reportando una respuesta en un archivo de salida de los datos. La figura 10 ilustra el proceso.



**Figura 10 Diagrama de bloques del GAMS**

Fuente: (Pinto J. M., 1988)<sup>[10]</sup>

Las convecciones utilizadas por el programa para los archivos (nombres y extensiones) siguen a abajo:

- Archivo de entrada: NOMBREARQ.GMS
- Archivo de salida: NOMBREARQ.LST

Para la compilación ejecución del modelo, el comando es simplemente GAMS NOMBREARQ.GMS. El archivo de entrada es, en general, organizado en las siguientes secciones.

- a) Especificación de índices y de los datos;
- b) Lista de los nombres y tipos de variables y nombres de las ecuaciones (restricciones y función objetivo);
- c) Definición de las ecuaciones (restricciones y función objetivo);
- d) Especificación de los límites (inferiores y superiores) de las variables, condiciones iniciales y opciones de ejecución;
- e) Llamada del algoritmo de resolución.

El formato del archivo de entrada no es rígido, pero su sintaxis debe ser obedecida. Además de eso, el programa verifica el tipo de algoritmo que debe ser utilizado para solucionar el modelo, avisando en caso de especificación incorrecta del algoritmo (ejemplo: especificación de MILP cuando el modelo se encuentra como MINLP, o sea, como función objetivo mixta entera no lineal).

GAMS posee una serie de herramientas que permite el ecuacionamiento de problemas de manera simple a modelos de gran complejidad, pero el programa no acepta

llamadas de funciones o procedimientos (o sea, todos los datos del modelo deberán estar en único archivo de entrada).

Las ecuaciones son definidas listando, primeramente, el respectivo nombre, seguidos de dos puntos (..). Por ejemplo, al citar la EQ1 en el modelo, debe ser digitado en el archivo de entrada: EQ1..<ecuación>. Debe ser destacado que la siguiente sintaxis debe ser utilizada, denotando los signos de igualdad y desigualdad en las ecuaciones:

- =E=: igual (=);
- =G=: mayor o igual que ( $\geq$ );
- =L=: menor o igual que ( $\leq$ ).

Sigue a continuación la sintaxis de los operadores aritméticos básicos:

- +, - : adición, sustracción;
- \*, / : multiplicación división;
- \*\*: exponente

### 2.3.2 Sistema especialista

Es un método que combina reglas heurísticas y relaciones matemáticas para seleccionar la secuencia de procesos de purificación. Este sistema es construido usando una estructura comercialmente disponible, "Nexpert Object" de la empresa "Neuron Data".

**a) Construcción de un sistema especialista.**

Considera dos sub procesos: el primer proceso es el aislamiento (separación), comprendiendo la separación del material de interés, rompimiento de célula, separación de tejido celular, solubilización y "refolding" (reflujo) y si fuera necesario precipitación de ácidos nucleicos.

El segundo proceso es la purificación, este toma la solución conteniendo la mezcla de proteínas y purifica la proteína de interés a una alta pureza de las otras proteínas, muchas de ellas tienen propiedades fisicoquímicas similares.

La construcción para encontrar la selección de operaciones en el proceso de aislamiento se basa sólo en reglas heurísticas [2].

En el proceso de purificación la construcción es difícil, debido a usar operaciones de purificación de alta resolución (técnicas cromatográficas) el cual se basa en propiedades fisico-químicas de todas las proteínas presentes en la mezcla.

El sistema especialista selecciona el mejor proceso usando el criterio de Selección del Coeficiente Separación (SSC)

calculado para cada técnica cromatográfica y para cada contaminante. El mejor proceso será el que tenga mas alto valor de SSC, la que es calculado mediante la siguiente relación.

$$SSC = DF * \eta * \Theta \quad (X)$$

Donde:

DF factor de desvío

$\eta$  factor de eficiencia

$\Theta$  factor de concentración

El factor de eficiencia  $\eta$  toma valores de:

1.0 para intercambio iónico (aniónico y catiónico)

0.86 para cromatografía por interacción hidrofóbica

0.66 para gel filtración

$$\Theta = CF = \frac{CONC.CONTAMINATE.i}{CONC.TOTAL.DE.PROTEINAS}$$



### III. ANÁLISIS DEL PROBLEMA

Antes de analizar el problema se verá algunas definiciones:

#### 3.1 Tiempo de retención ( $K_d$ )

Es un parámetro que se usa para representar el comportamiento de las proteínas en una separación llevada a cabo por las cromatografías de gel filtración, intercambio iónico e hidrofobicidad.

Por ejemplo, si la carga de la proteína es negativa entonces el tiempo de retención para el intercambio iónico es dado como una relación matemática mostrada en la tabla 1, en el otro caso el tiempo de retención es cero.

Por otro lado,  $K_{d,i,p}$  es igual a una relación matemática si la carga ( $P_{a,p}$ ) es positiva, en el caso opuesto  $K_{d,i,p}$  es cero. Para la cromatografía por iteración hidrofóbica el tiempo de retención es una función de la hidrofobicidad y del peso molecular para gel filtración (representado en la tabla 1 como MW).

**Tabla 1. Expresiones y parámetros para las técnicas cromatográficas.**

Técnicas cromatograficas	Tiempo de retención $Kd_{i,p}$	Condición	Amplitud del pico ( $\sigma_i$ )
Intercambio anionico	$Kd_{AE,p} = \frac{7383 *  (P_{a,p} / MW) }{1 * 10^{-25} + 15844 *  (P_{a,p} / MW) }$ $Kd_{AE,p} = 0$	Sí $P_{Ap} < 0$  si $P_{Ap} > 0$	0.15
Intercambio Cationico	$Kd_{CE,p} = \frac{5972 * (P_{a,p} / MW)}{1 * 10^{-25} + 17065 * (P_{a,p} / MW)}$ $Kd_{CE,p} = 0$	Sí $P_{Ap} > 0$  si $P_{Ap} < 0$	0.15
Interacción hidrofóbica	$Kd_{HI,p} = 1 - \frac{[(NH_4)_2SO_4]}{[(NH_4)_2SO_4]_{max}}$	$[(NH_4)_2SO_4]_{max} = 1.5M$	0.22
Gel Filtración	$Kd_{GF,p} = -0.4691 * \log MW + 2.3902$	MW in [Da]	0.46

FUENTE: Lienqueo (1998)

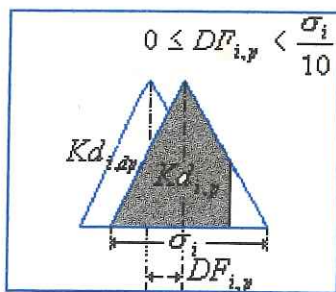
### 3.2 Factor de desvío (DF)

Este parámetro indica la diferencia entre la propiedad de la proteína de interés y la misma propiedad del contaminante (por ejemplo, el tiempo de retención para una determinada técnica cromatográfica) ver fig. 11

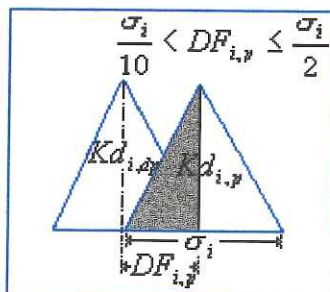
### 3.3 Factor de concentración (CF)

Este valor representa el grado de separación de cada proteína después de aplicar una técnica cromatográfica.

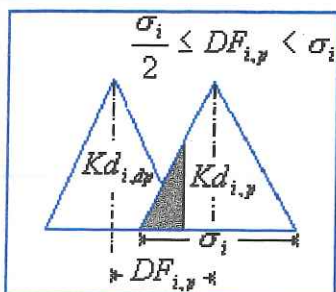
La figura (11e) representa una simplificación de los picos cromatográficos a un triángulo. En las figuras (11a – 11d) se muestra los diferentes casos en que  $DF_{i,p}$  es determinado.



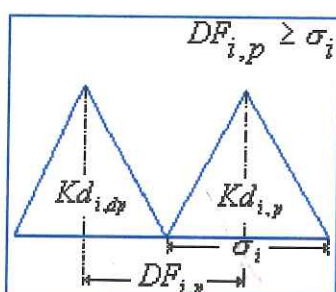
Fig(11a)



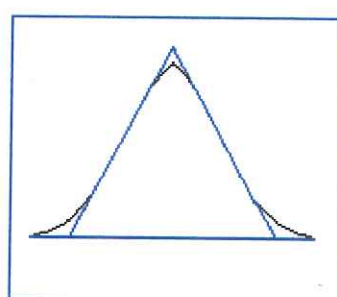
Fig(11b)



Fig(11c)



Fig(11d)



Fig(11e)

**Figura 11. Representación de los picos cromatográficos**

### 3.4 Descripción

Dado una mezcla de proteínas en diferentes niveles de concentración y una especificación del producto deseado en términos de una pureza mínima, el problema consiste en seleccionar las técnicas cromatográficas y

la secuencia a utilizar con la finalidad de minimizar el número total de operaciones.

Cada técnica cromatográfica es adecuado para desarrollar una separación basada en las propiedades fisico-químicas de la mezcla de proteínas, el cual define su tiempo de retención ( $K_d$ ). Por comparación de tiempos de retención de la proteína deseada y de los contaminantes, los factores de desviación son calculados.

Finalmente, de acuerdo al factor de desviación de cada contaminante y el pico cromatográfico de cada técnica se determina el factor de concentración, ésto es, el grado de separación después de cada técnica cromatográfica para cada una de las proteínas.

## IV. MODELO DE OPTIMIZACIÓN

En este capítulo, se describirá el modelo de optimización para sintetizar una secuencia de técnicas de purificación con la finalidad de alcanzar una pureza deseada.

### 4.1 Símbolos

En el modelo de optimización, se presenta a continuación: los índices:

$a$  propiedad de la proteína

$dp$  proteína deseada

$i$  técnica cromatográfica

$k$  orden en la secuencia en van a ser colocadas los pasos

$p$  proteína

Las variables:

$C_{dp,i}$  concentración de la proteína deseada después de la técnica  $i$

$C_{p,k}$  concentración del contaminante  $p$  antes de la técnica cromatográfica en la orden  $k$

$C_{i,p,k}$  concentración de la proteína  $p$  después de la técnica  $i$  en la orden  $k$

$S$  función a ser minimizada

$y_{i,k}$  variable binaria que indica si la técnica cromatográfica  $i$  es usada en la orden  $k$  siguiendo la secuencia

$Z_k$  variable binaria que indica si la orden  $k$  es la última en la secuencia

y los parámetros:

$C_{dp}^p$  nivel de pureza de la proteína deseada

$C_p^u$  límite máximo de concentración de la proteína  $p$

$C_p^o$  concentración inicial de la proteína  $p$

$CF_{i,p}$  factor de concentración de la proteína  $p$  después de la técnica cromatográfica  $i$

$DF_{i,p}$  factor de desvío para la proteína  $p$  en la técnica cromatográfica  $i$

$ip$  índice de la proteína

$Kd_{i,p}$  tiempo de retención para la proteína  $p$  en la técnica cromatográfica  $i$

$Kd_{i,dp}$  tiempo de retención para la proteína deseada  $dp$  en la técnica cromatográfica  $i$

$P_{a,p}$  valor de la propiedad  $a$  para la proteína  $p$

$\sigma_i$  Amplitud del pico de la técnica cromatografica  $i$

## 4.2 Calculo del factor de concentración

Cada técnica cromatográfica es caracterizada por el valor de la amplitud del pico  $\sigma_i$  (ver figura 11). Este parámetro es usado para calcular

el tiempo de retención; valor adimensional; ( $Kd_{i,p}$ ), el cual es una función de las propiedades fisico-químicas y representa el comportamiento de la proteína  $p$  en cada técnica cromatográfica  $i$ .

entonces:

$$Kd_{i,p} = f_i (P_{ap}) \quad \forall i,k \quad (1)$$

también:

$$DF_{i,p} = |Kd_{i,dp} - Kd_{i,p}| \quad \forall i,k \quad (2)$$

Para alguna proteína  $p$  y técnica cromatográfica  $i$ :

$$0 \leq DF_{i,p} < \frac{\sigma_i}{10} \quad \text{entonces} \quad CF_{i,p} = 1 \quad (3a)$$

$$\frac{\sigma_i}{10} \leq DF_{i,p} < \frac{\sigma_i}{2} \quad \text{entonces} \quad CF_{i,p} = 1.02 * \left( \frac{\sigma_i^2 - 2DF_{i,p}^2}{\sigma_i^2} \right) \quad (3b)$$

$$\frac{\sigma_i}{2} \leq DF_{i,p} < \sigma_i \quad \text{entonces} \quad CF_{i,p} = 1.02 * 2 * \frac{(\sigma_i - DF_{i,p})^2}{\sigma^2} \quad (3c)$$

$$DF_{i,p} \geq \sigma_i \quad \text{entonces} \quad CF_{i,p} = 0.02 \quad (3d)$$

Con  $DF_{i,p}$  calculado en (2), y  $\sigma_i$  es posible calcular  $CF_{i,p}$  de (3) el cual es indicado por relaciones obtenidas de las gráficas de los picos simplificados a un triángulo, las cuatro relaciones presentan situaciones diferentes, estos casos se muestran en la figura 11, presentada anteriormente.

### 4.3 Modelo matemático

El modelo para la purificación de proteínas usando las técnicas cromatográficas de alta resolución es compuesta de las siguientes restricciones:

a) restricciones de colocación

$$\sum_i y_{i,k} \leq 1 \quad \forall k \quad (4a)$$

$$\sum_k y_{i,k} \leq 1 \quad \forall i \quad (4b)$$

b) restricciones de orden

$$\sum_i y_{i,k+1} \leq \sum_i y_{i,k} \quad \forall k = 1, \dots, K-1 \quad (5a)$$

$$Z_k \geq \sum_i y_{i,k} - \sum_i y_{i,k+1} \quad \forall k = 1 \dots K-1 \quad (5b)$$

$$\sum_i y_{i,k'} \geq Z_k \quad \forall k, k' \leq k \quad (5c)$$

$$\sum_i y_{i,k'} + Z_k \leq 1 \quad \forall k, k' \geq k \quad (5d)$$

$$\sum_k Z_k = 1 \quad (5e)$$

c) restricciones de los contaminantes

$$C_{p,k} = \sum_i CF_{i,p} C_{p,k-1} y_{i,k} \quad \forall k,p \quad (6)$$



d) restricciones de pureza

$$C_{dp,k+1} \geq C_{dp}^p * \sum_{p'} C_{pp,k+1} - C_p^u * (1 - Z_k) \quad \forall k \quad (7)$$

La ecuación (4a) especifica que por lo menos uno de las técnicas cromatográficas  $i$  es escogida en la orden  $k$ , mientras que (4b) indica que cada técnica es seleccionada a lo mucho una vez en la secuencia.

Con la finalidad de evitar que técnicas vacías sean escogidas inclusive para reducir la búsqueda combinatoria, la ecuación (5a) indica que la técnica  $k$  es asignado incrementando el orden, esto es, el paso  $k+1$  puede solamente ser escogido si el paso  $k$  es también escogido. Las restricciones (5b) - (5c) son usados para representar la ultima técnica en la secuencia, denotado por  $Z_k$ .

La ecuación (6) indica los niveles de concentración de cada proteína en las técnicas subsiguientes. Note, sin embargo que ellos tienen la forma no lineal, desde que la sumatoria contiene un producto de una variable continua y una variable discreta. La ecuación (6) puede ser linealizada si introducimos una variable auxiliar que disagrege proteínas y técnicas. La forma linealizada puede ser escrita como:

$$C_{p,2} = \sum_i CF_{i,p} * y_{i,1} * C_p^0 \quad \forall p,k = 2, \dots, K \quad (8a)$$

$$C_{p,k} - \sum_i CF_{i,p} * C_{i,p,k-1} \geq -C_p^u * (1 - \sum_i y_{i,k-1}) \quad \forall p,k = 3, \dots, K \quad (8b)$$

$$C_{p,k} - \sum_i CF_{i,p} * C_{i,p,k-1} \geq C_p^u * (1 - \sum_i y_{i,k-1}) \quad \forall p,k = 3, \dots, K \quad (8c)$$

$$C1_{i,p,k-1} \leq C_p^u * y_{i,k} \quad \forall i,p,k \quad (8d)$$

$$C_{p,k-1} = \sum_i C1_{i,p,k-1} \quad \forall p,k \quad (8e)$$

Esto es importante para notar que la ecuación (6) es usada para la primera técnica cromatográfica, desde que la concentración inicial es conocida a priori, como expresada en (8a). Las restricciones (8b) y (8c) en las técnicas subsiguientes son impuestas solamente si alguno de ellos es seleccionado. Las variables desagregada  $C1_{i,p,k}$  está vinculados a las variables originales  $y_{i,k}$  y  $C_{p,k}$  en las restricciones (8d) y (8c).

La restricción (7) exige la pureza específica para la proteína de interés. Si  $Z_k$  es 1, esto es  $k$  es la última técnica cromatográfica, entonces la condición siguiente es impuesta.

$$\frac{C_{dp,k+1}}{\sum_p C_{p,k+1}} \geq C_{dp}^p \quad (9)$$

La función objetivo es la minimización del número total de operaciones cromatográficas, dado como:

$$\text{Min } S = \text{Min} \sum_{k=1}^K k * Z_k \quad (10)$$

Donde  $K$  es el máximo número de operaciones permitida.

#### 4.4 Estructura del modelo

El modelo MILP tiene la estructura siguiente:

Minimizar (10)

Sujeto a:

Restricciones de asignación (4)

restricciones de orden (5)

restricciones de contaminantes (8)

restricciones de pureza (7)

$$y_{i,k} \in \{0,1\}$$

$$C_{p,k}, C_{1_{i,p,k}}, Z_k \geq 0$$

$$C_p^u = 2 * \sum_{p=ip} C_p^0$$

A continuación se presenta el modelo generado para la purificación de proteínas, se debe tener en cuenta que el software esta en el idioma ingles y para facilitar el entendimiento se hizo la traducción.

El modelo va digitado en un archivo de entrada llamado "protein1", el que será llamado por el software GAMS, (ver archivos de salida en anexo).

### \* MODELO GENERAL PARA LA PURIFICACIÓN DE PROTEÍNAS

Exhibir kd,df,cf;

Variables S;

Variables positivas C(p,k), C1(i,p,k);

Variables binarias Y(i,k), Z(k);

Ecuaciones OBJ, EQ1, EQ2, EQ3, EQ4, EQ5, EQ6, EQ7, EQ8,eq9,eq4a,eq4b  
eq7a, eq7b;

\*Función Objetivo

c.fx(p,'1') = con(p);

OBJ.. S=E=SUM(k,ord(k)\*z(k));

\*Sujeto a

EQ1(k)..SUM(i,Y(i,k))=L=1;

EQ2(i)..SUM(k,Y(i,k))=L=1;

EQ3(k)\$ (ord(k) lt card(k)).. SUM(i,Y(i,k+1))=L=SUM(i,Y(i,k));

EQ4(p,k)\$ (ord(k) ge 3)..C(p,k)-SUM(i,CF(i,p)\*C1(i,p,k-1)) =g= -U\*(1-sum(i,y(i,k-1)));

EQ4a(p,'2')..C(p,'2')=E=SUM(i,CF(i,p)\*y(i,'1')\*con(p));

EQ4b(p,k) $\$(ord(k) \ge 3)..C(p,k)-SUM(i,CF(i,p)*C1(i,p,k-1)) = U*(1-sum(i,y(i,k-1)))$ ;

EQ5(i,p,k) $\$(ord(k) \ge 2)..C1(i,p,k)=L=con(p)*Y(i,k-1)$ ;

EQ6(p,k) $\$(ord(k) \ge 2)..C(p,k)=E=SUM(i,C1(i,p,k))$ ;

EQ7(k) $\$(ord(k) \le card(k)).. Z(k)=G=SUM(i,Y(i,k))-SUM(i,Y(i,k+1))$ ;

eq7a(k,kk) $\$(ord(kk) \le ord(k)).. SUM(i,Y(i,kk)) =g= z(k)$ ;

eq7b(k,kk) $\$(ord(kk) > ord(k)).. SUM(i,Y(i,kk)) + z(k) = 1$ ;

EQ8..SUM(k,Z(k))=E=1;

EQ9(p,k) $\$(ORD(p) \le EQ IP).. C(p,k+1)=G=0.98*(SUM(pp,C(pp,k+1)))-U*(1-Z(k))$ ;

MODELO PROTEIN1 /OBJ,

EQ1, EQ2, EQ3,

EQ4,

eq4a,eq4b

EQ5,

EQ6,

EQ7, eq7a, eq7b

EQ8

eq9/;

protein1.espacio de trabajo=10;

RESOLVER PROTEIN1 USANDO MIP MINIMIZANDO S;

Exhibir C.I, C1.I, Z.I, Y.I;

## V. RESULTADOS

El modelo descrito en el capítulo IV es aplicado en tres ejemplos diferentes:

### 5.1 Ejemplo 1. Purificación de la Albúmina de suero bovino ("Serum from bovine Albumin" - BSA)

En este ejemplo se considera la purificación de una mezcla que contiene cuatro proteínas, Albúmina de suero bovino (BSA), Inhibidor de tripsina (SBTI), Albumina de huevo y Thaumatina, para facilidad de trabajar denominaremos p1, p2, p3 y p4 respectivamente. Sus propiedades fisico-químicas se muestran en tabla 2. Las técnicas cromatográficas a usarse en este ejemplo se muestran en la tabla 3:

**Tabla 2. Propiedades fisico-químicas de la mezcla de proteínas.**

Proteína	Concentración inicial [mg/ml]	Peso Molecular (Da)	Hidrofobicidad $[(NH_4)_2(SO_4)]$	$P_{a,p} = \text{Carga (C/mol)} \cdot 10^{-25}$				
				pH 4.0	pH 5.0	pH6.0	pH7.0	pH8.0
P1	2	67000	0.86	1.03	-0.14	-1.16	-1.68	-2.05
P2	2	43800	0.54	1.40	-0.76	-1.65	-2.20	-2.36
P3	2	24500	0.90	1.22	-0.76	-1.54	-2.17	-2.13
P4	2	22200	0.89	1.94	1.90	1.98	1.87	0.91

Fuente: Lienqueo (1998) [8]

**Tabla 3. Técnicas cromatográficas para ejemplo 1**

Número de técnica	Nombre de la técnica
1	Intercambio aniónico pH 4.0
2	Intercambio aniónico pH 5.0
3	Intercambio aniónico pH 6.0
4	Intercambio aniónico pH 7.0
5	Intercambio aniónico pH 8.0
6	Intercambio catiónico pH 4.0
7	Intercambio catiónico pH 5.0
8	Intercambio catiónico pH 6.0
9	Intercambio catiónico pH 7.0
10	Intercambio catiónico pH 8.0
11	Interacción hidrofóbica
12	Gel filtración

Fuente: elaboración propia

El cálculo del tiempo de retención,  $K_d$  se realiza empleando la tabla 1 y tabla 2 mostradas anteriormente. Ejemplo: para la proteína 1 en la técnica 1 (intercambio iónico) la carga es positiva ( $>0$ ) por lo tanto  $K_d = 0$ . En la tabla 4 se presenta los valores ya calculados para cada una de las técnicas cromatográficas de las cuatro proteínas.

**Tabla 4. Tiempo de retención (Kd) de la proteína *p* en técnica *i*.**

Técnica	P1	P2	P3	P4
1	0.0	0.0	0.0	0.0
2	0.015	0.100	0.154	0.0
3	0.100	0.174	0.233	0.0
4	0.132	0.206	0.272	0.0
5	0.152	0.215	0.270	0.0
6	0.073	0.124	0.161	0.209
7	0.0	0.0	0.0	0.208
8	0.0	0.0	0.0	0.211
9	0.0	0.0	0.0	0.206
10	0.0	0.0	0.0	0.144
11	0.427	0.640	0.400	0.407
12	0.126	0.213	0.331	0.351

Fuente: Elaboración propia

A continuación con la ecuación 2 se hace el calculo del factor de desvío (DF), por ejemplo para la proteína 2 en la técnica 2 será:

$$Kd_{2,p2} = | kd_{2,p1} - kd_{i,p2} | \text{ como se observa } Kd_{2,p1} = 0.015 \text{ y}$$

$$Kd_{2,p2} = 0.1 \text{ entonces } Kd_{2,p2} = 0.085. \text{ Los resultados completos se}$$

presentan en la tabla 5.

**Tabla 5. Factor de desvío (DF) de la proteína *p* en la técnica *i*.**

Técnica	P1	P2	P3	P4
1	0.0	0.0	0.0	0.0
2	0.0	0.085	0.139	0.015
3	0.0	0.074	0.132	0.100
4	0.0	0.074	0.140	0.132
5	0.0	0.062	0.118	0.152
6	0.0	0.051	0.088	0.137
7	0.0	0.0	0.0	0.208
8	0.0	0.0	0.0	0.211
9	0.0	0.0	0.0	0.206
10	0.0	0.0	0.0	0.144
11	0.0	0.213	0.027	0.02
12	0.0	0.087	0.205	0.225

Fuente: Elaboración propia



Con los valores de DF encontramos el factor de concentración CF, con ayuda del grupo de ecuaciones (3), por ejemplo para la proteína 1 técnica 1  $DF_{1,P1} = 0$ , entonces  $CF_{1,P1} = 1.0$ , la tabla 6 muestra los resultados completos.

**Tabla 6. Factor de concentración (CF) de la proteína  $p$  en la técnica  $i$ .**

Técnica	P1	P2	P3	P4
1	1.000	1.000	1.000	1.000
2	1.000	0.377	0.012	1.000
3	1.000	0.525	0.029	0.224
4	1.000	0.523	0.010	0.028
5	1.000	0.666	0.094	0.020
6	1.000	0.786	0.384	0.016
7	1.000	1.000	1.000	0.020
8	1.000	1.000	1.000	0.020
9	1.000	1.000	1.000	0.020
10	1.000	1.000	1.000	0.003
11	1.000	0.002	0.990	1.000
12	1.000	0.984	0.615	0.532

Fuente: Elaboración propia

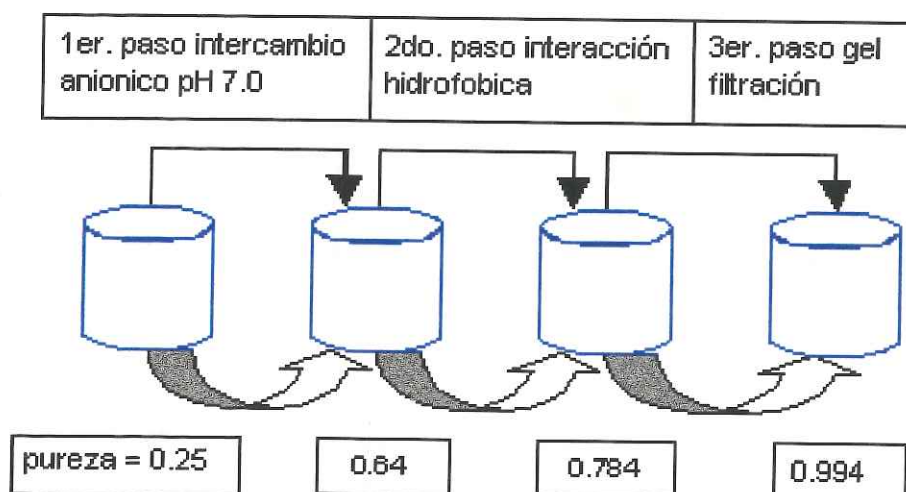
La purificación requerida para proteína p1 es 98% y puede lograrse a través de la selección apropiada entre las técnicas cromatográficas siguientes: el intercambio del anión y catión a diferentes valores de pH, interacción hidrofóbica y gel filtración; en total, hay doce técnicas cromatográficas, la pureza alcanzada es de 99.4. Las concentraciones resultantes se muestran en la tabla 7.

**Tabla 7. Concentraciones resultantes de las proteínas**

Proteína	C <sub>0</sub> (mg/ml)	C <sub>1</sub> (mg/ml)	C <sub>2</sub> (mg/ml)	C <sub>3</sub> (mg/ml)
p1	2.00	2.000	2.000	2.000
p2	2.00	1.047	0.548	0.001
p3	2.00	0.020	1.92E-4	1.91E-4
p4	2.00	0.056	0.002	0.002

Fuente: elaboración propia

Figure 12 muestra que el primer paso seleccionado es intercambio del anión a pH 7.0 en el que la mayoría de p3 y p4 es separado (vea tabla 7). Esto ocurre desde que su peso molecular es más bajo que de p1 y p2 (ver tabla 2). El segundo paso es interacción hidrófoba por el parámetro que es alto, y el tercer y última técnica es gel filtración esta técnica se usa para separar sólo p2 de p1 porque el peso molecular es medianamente bajo que p1, y también la hidrofobicidad es baja.



**Figura 12. Representación de las técnicas cromatograficas**

Fuente: Elaboración propia

## 5.2 Ejemplo 2. Purificación del ecoli\_s8

En este caso nosotros consideramos la purificación de una proteína (ecoli\_s8); generada del proceso de fermentación usando Escherichia coli; en el cual el producto debe separarse de ocho contaminantes. En este caso de las proteínas contaminantes sólo se conocen sus propiedades físico-químicas mas no sus nombres, a las que denominamos p1, p2, p3, p4, p5, p6, p7, y p8. Sus propiedades físico-químicas se muestran en tabla 8.

**Tabla 8. Propiedades físico-químicas de la mezcla de proteínas.**

Proteína	Co G/lt	MW Da	Hidrof.	PH 4,0	pH 4,5	P <sub>a,p</sub> = pH 5,0	Carga (c/mol) pH 5,5	pH 6,0	pH 6,5	*10 <sup>-25</sup> pH 7,0	pH 7,5	PH 8,0	pH 8,5
Producto	0.62	31000	0.00	1.46	0.09	-0.62	-0.66	-1.02	-1.82	-2.33	-2.52	-2.52	-3.51
Cont_1	0.42	62500	0.00	1.46	0.09	-1.06	-0.98	-1.17	-1.71	-2.79	-3.52	-3.32	-3.32
Cont_2	0.25	40600	0.00	1.46	0.09	-0.55	-0.22	-0.22	-0.26	-0.73	-1.26	-1.82	-3.51
Cont_3	0.25	69600	0.00	1.46	0.09	-0.55	-0.22	-0.22	-0.26	-0.73	-1.26	-1.82	-3.51
Cont_4	0.09	40600	0.00	1.46	3.14	1.46	0.28	-0.47	-0.89	-1.06	-1.08	-1.04	-1.01
Cont_5	0.09	69600	0.00	1.46	3.14	1.46	0.28	-0.47	-0.89	-1.06	-1.08	-1.04	-1.01
Cont_6	2.74	41000	1.50	1.46	0.93	0.26	-0.35	-0.87	-1.31	-1.65	-1.90	-2.04	-2.06
Cont_7	2.74	32900	1.50	1.46	0.09	0.00	-1.70	-2.70	-2.90	-3.51	-3.51	-3.51	-3.51
Cont_8	0.25	35500	0.20	1.46	0.09	-0.55	-0.22	-0.22	-0.26	-0.73	-1.26	-1.82	-3.51

Fuente: Lienqueo, 1998-UNIVERSIDAD DE CHILE

La purificación requerida para proteína p1 es de 94% y se logra en cinco pasos de una lista de 22 técnicas cromatográficas las que se indica en la tabla 9.

**Tabla 9. Técnicas cromatográficas usada en el ejemplo 2 y 3**

Número de técnica	Nombre de la técnica cromatográfica
1	Intercambio aniónico ph4.0
2	Intercambio aniónico ph4.5
3	Intercambio aniónico ph5.0
4	Intercambio aniónico ph5.5
5	Intercambio aniónico ph6.0
6	Intercambio aniónico ph6.5
7	Intercambio aniónico ph7.0
8	Intercambio aniónico ph7.5
9	Intercambio aniónico ph8.0
10	Intercambio aniónico ph8.5
11	Intercambio catiónico ph 4.0
12	Intercambio catiónico ph 4.5
13	Intercambio catiónico ph 5.0
14	Intercambio catiónico ph 5.5
15	Intercambio catiónico ph 6.0
16	Intercambio catiónico ph 6.5
17	Intercambio catiónico ph 7.0
18	Intercambio catiónico ph 7.5
19	Intercambio catiónico ph 8.0
20	Intercambio catiónico ph 8.5
21	Interacción hidrofóbica
22	Gel filtración

Fuente: Elaboración propia

En la determinación de  $K_d$  se usa que la relación matemática mostrada en la tabla 1, el de las propiedades en tabla 8 y se usa las ecuaciones mencionadas en el ejemplo 1. Todo el calculo es similar que el ejemplo 1.

Las concentraciones resultantes se muestran en la tabla 10 y la secuencia sugerida en la tabla 17.

**Tabla 10. Concentración resultante de las proteínas del ejemplo 2**

Proteína	C <sub>0</sub> (gr/lt)	C <sub>1</sub> (gr/lt)	C <sub>2</sub> (gr/lt)	C <sub>3</sub> (gr/lt)	C <sub>4</sub> (gr/lt)	C <sub>5</sub> (gr/lt)
Producto	0.620	0.620	0.620	0.620	0.620	0.620
P1	0.420	0.420	0.420	0.155	0.062	0.040
P2	0.250	0.250	0.250	0.235	0.005	0.00000
P3	0.250	0.250	0.250	0.076	0.002	0.00000
P4	0.090	0.090	0.090	0.002	0.00034	0.0000
P5	0.090	0.090	0.090	0.002	0.000002	0.0000
P6	2.740	0.055	0.001	0.00032	0.00019	0.0000
P7	2.740	0.055	0.001	0.001	0.00089	0.0000
P8	0.250	0.004	0.000071	0.000071	0.000001	0.0000

Fuente: Elaboración propia

### 5.3 Ejemplo 3. Purificación del "Somatotropin"

En este ejemplo nosotros tenemos una mezcla de la proteína generada del proceso de fermentación usando *Escherichia coli* recombinante, nuestro interés es purificar solamente una proteína "Somatotropin" de 13 proteínas contaminantes, de las que solo conocemos sus propiedades físico-químicas. La totalidad de técnicas cromatográficas a escoger son 22 y se muestran en la tabla 9. La determinación del tiempo de la retención es usando tabla 1 y 11, el factor de desvío y el de concentración es similar al del ejemplo 1. Las concentraciones finales se muestran en la tabla 12.

**Tabla 11. Propiedades fisico-químicas de la mezcla de proteínas**

Proteína	Co. G/lt	MW Da	Hidro.	PH 4,0	pH 4,5	P <sub>a,p</sub> = pH 5,0	Carga pH 5,5	(coulb pH 6,0	pH 6,5 *10 <sup>-25</sup>	pH 7,0	pH 7,5	pH 8,0	pH 8,5
Somatotropin	25.0	22000	0.93	4.77	3.81	2.42	1.50	1.03	0.67	0.12	0.07	-0.03	-0.50
P1	11.29	18370	0.71	1.94	0.25	-0.80	-1.41	-1.76	-1.97	-2.15	-2.33	-2.45	-2.67
P2	7.06	85570	0.48	2.35	0.29	-1.17	-2.17	-2.83	-3.24	-3.50	-3.63	-3.68	-3.64
P3	4.63	53660	0.76	1.83	0.76	0.04	-0.30	-0.49	-0.65	-0.85	-1.90	-1.34	-1.50
p4	5.58	120000	1.50	3.29	1.38	-0.03	-0.69	-1.07	-1.34	-1.73	-2.30	-2.85	-2.75
p5	4.83	203000	0.36	4.08	1.83	0.04	-1.17	-1.92	-2.46	-3.07	-3.90	-4.98	-5.65
p6	2.48	69380	0.36	5.22	3.17	1.02	-0.72	-1.90	-2.60	-3.05	-3.46	-3.90	-4.24
p7	7.70	48320	0.48	3.96	3.16	1.12	-0.58	-1.36	-1.34	-1.00	-0.95	-1.59	-2.84
p8	6.80	93380	0.93	10.90	5.81	2.78	0.77	-0.81	-2.18	-3.32	-4.12	-4.45	-4.31
p9	7.53	69380	0.00	1.09	0.55	0.26	0.10	-0.03	-0.12	-0.21	-0.28	-0.32	-0.32
P10	6.05	114450	0.63	10.40	5.94	3.15	1.51	0.56	-0.05	-0.53	-0.99	-1.43	-1.72
P11	3.89	198000	0.06	0.33	0.03	0.05	0.05	0.05	0.05	0.05	-0.69	-0.97	-1.57
P12	1.48	30400	0.00	5.17	4.22	3.20	2.25	1.46	0.87	0.50	0.30	0.20	0.08
P13	0.83	94670	0.00	11.70	7.94	5.39	3.73	2.66	1.97	1.50	1.13	0.80	0.51

Fuente: Lienqueo (1996)

**Tabla 12. Concentraciones Resultantes del ejemplo 3**

Proteínas	C <sub>0</sub> (g/lit)	C <sub>1</sub> (g/lit)	C <sub>2</sub> (g/lit)	C <sub>3</sub> (g/lit)
Producto	25.00	25.00	25.00	25.00
P1	11.290	0.226	0.005	0.0000903
P2	7.060	0.141	0.003	0.000564
P3	4.630	0.488	0.051	0.004
P4	5.580	0.112	0.002	0.000213
P5	4.830	0.097	0.002	0.00015
P6	2.480	0.050	0.000992	0.0000198
P7	7.700	0.154	0.003	0.0000
P8	6.800	6.800	6.800	0.136
P9	7.530	0.151	0.003	0.003
P10	6.500	0.121	0.002	0.001
P11	3.890	0.078	0.002	0.002
P12	1.480	0.030	0.000592	0.000592
P13	0.830	0.017	0.000332	0.000332

Fuente: Elaboración propia

La Tabla 13 muestra que los pasos usados son: interacción hidrofóbica, para el primer paso, intercambio aniónico a pH 8,0 para el segundo paso y el paso tres es intercambio aniónico a pH 4,5, este paso es usada sólo para separar p8 que no fue separado en los dos pasos anteriores.

**Tabla 13. Técnicas usadas en la purificación del Somatotropin**

Pasos	paso1	Paso2	Paso3
Técnicas cromatográficas	Interacción Hidrofóbica	Intercambio aniónico pH 8,0	Intercambio aniónico pH 4,5

Fuente Elaboración Propia

#### 5.4 Comparación de Metodología

La comparación entre ambas metodologías se realizó en función del % de pureza =  $\frac{\text{con.proteina.i}}{\sum \text{todas.proteina}}$ , las secuencias y el número de las técnicas sugeridas.

Para este trabajo sólo se comparan los dos primeros ejemplos, no contando con datos del sistema especialista para hacer la comparación del tercer ejemplo.

**Tabla 14. Secuencia sugerida por Sistema Especialista del ejemplo 1**

Técnica Cromatográfica	% Pureza
Intercambio Catiónico a pH 6,0	33.11
Interacción hidrofóbica	49.45
Intercambio aniónico a pH 7,0	97,02

Fuente: Lienqueo (1998)

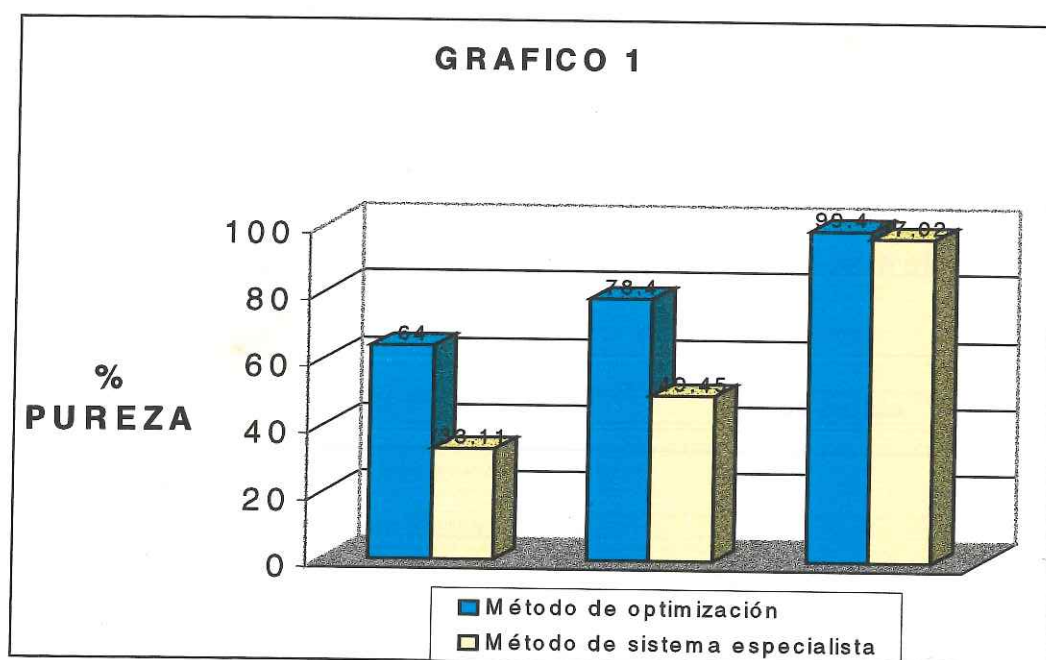
**Tabla 15. Secuencia sugerida por el modelo para el ejemplo 1**

Técnica Cromatográfica	% Pureza
Intercambio aniónico a pH 7,0	64,0
Interacción hidrofóbica	78,4
Gel filtración	99,4

Fuente: Elaboración Propia



De las tablas 14 y 15 se aprecia que la secuencia sugerida por el método de optimización propuesto por esta tesis lleva a un porcentaje mayor de pureza (99.4) que el propuesto por el sistema especialista (97.02), para eso debe comenzar la secuencia de purificación con la técnica 4; (intercambio anionico pH 7.0).



En el gráfico 1 se aprecia mejor que el modelo propuesto es más eficiente que el del sistema especialista. La primera técnica sugerida por nuestro modelo realiza un alto grado de purificación (64%), mientras que la del sistema especialista sólo (33.11%). La ventaja se aprecia mas, en la tercera técnica cuando se obtiene una pureza casi total (99.4%) por el modelo de optimización.

**Tabla 16. Secuencia sugerida por Sistema Especialista para el ejemplo 2**

Técnicas cromatográficas	% de pureza
Cromatografía de interacción hidrofóbica	32
Intercambio aniónico a pH 6,5	69
Cromatografía por gel filtración	72
Intercambio aniónico a pH 8,5	85
Intercambio aniónico a pH 5,0	91
Intercambio catiónico a pH 4,0	94

Fuente: Lienqueo (1998)

**Tabla 17. Secuencia sugerida por el modelo de optimización para el ejemplo 2**

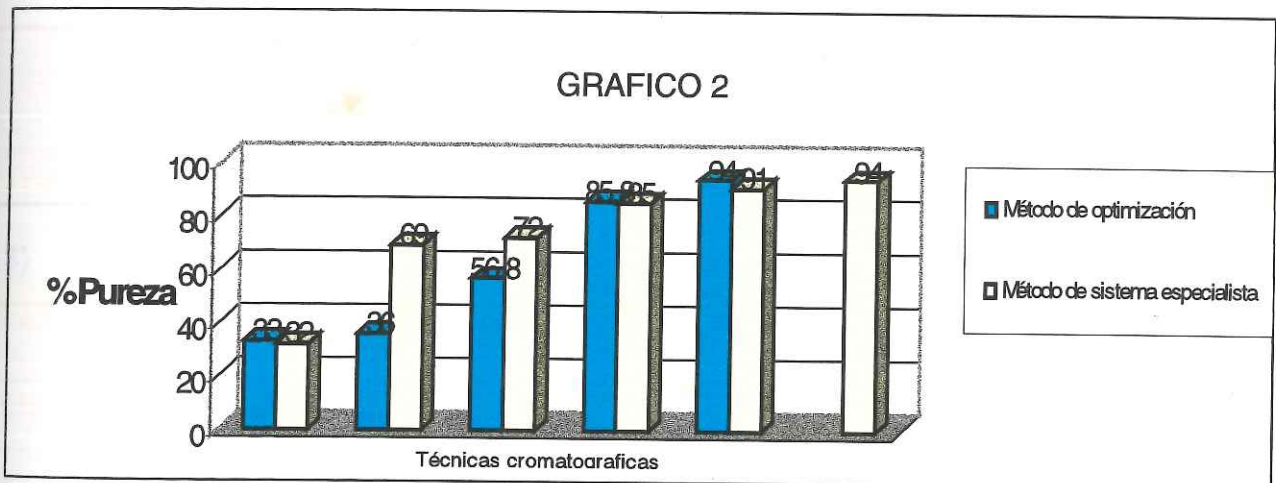
Técnicas cromatográfica	% de pureza
Cromatografía de interacción hidrofóbica	33,8
Intercambio aniónico a pH 8,5	36,0
Intercambio aniónico a pH 6,5	56,8
Intercambio aniónico a pH 7,0	89,8
Intercambio aniónico a pH 6,0	94,0

Fuente: Elaboración Propia

En las tablas 16 y 17 observamos las secuencias sugeridas por ambos métodos y en ellas tenemos como primer paso a la cromatografía por interacción hidrofóbica dando un porcentaje de pureza casi parecidos la ventaja es que se alcanza la pureza requerida de 94 con una técnica menos que por el método del sistema especialista. En la secuencia

sugerida por el método propuesto se llega a la pureza requerida con tan sólo 5 pasos, mientras que con la misma cantidad de pasos el modelo en comparación logra sólo alcanzar un 91% de pureza.

Del grafico se deduce que la purificación lenta, en un inicio y luego elevarlo drásticamente es el mas recomendable para la reducción del número de técnicas a usar.



## VI. CONCLUSIONES Y RECOMENDACIONES

### 6.1 Conclusiones

- 1) El modelo de programación matemática propuesto (MILP) soluciona el problema de síntesis de purificación de proteínas llevadas a cabo en un laboratorio.
- 2) Comparando las dos metodologías, se determina que las soluciones generadas por el modelo propuesto son mejores que las de un sistema especialista
- 3) En el Ejemplo 1, el porcentaje de pureza obtenida por el modelo de optimización es 99.4% en comparación con 97.02% del sistema especialista.
- 4) En el Ejemplo 2, el número de técnica cromatográficas, para una pureza de 94%, calculado por el modelo de optimización es 5 mientras que con el sistema especialista son 6. Este sistema con 5 pasos sólo llega a purificar en un 91%, por lo tanto con el modelo de optimización se minimiza el número de pasos cromatograficos y como consecuencia se reduce el costo.

## 6.2 Recomendaciones

- 1) Es muy importante que las propiedades físico-químicas de las proteínas contaminantes sean muy similares a fin de obtener el menor número de técnicas cromatográficas.
- 2) El valor de la concentración máxima de la proteína de interés en el modelo matemático no debe exceder a lo establecido.
- 3) Para facilitar el modelo es mejor trabajar con modelos mixtos y lineales
- 4) Es muy importante intercambiar información entre los investigadores para la obtención de mejores resultados.

## REFERENCIAS BIBLIOGRAFICAS

- [1] Albert, L.L. *Bioquímica: las bases moleculares de la estructura y función celular* Ed. Omega , Barcelona, España 2ª) edición pag.91-93, 172,173. 1985
- [2] Asenjo, J.A. *Separation processes in biotechnology*, Ed. Marcel Dekker, New York, United States of America. 1990
- [3] Bryant, C.H.; Rowe, R.C. Knowledge discovery in data bases: application to chromatography. *Trac - trends in analytical chemistry*, pag. 17 18 - 24. 1998
- [4] Dorland, *Diccionario de ciencias médicas*, Ed. Ateneo, Argentina, pag. 336, 1154. 1986
- [5] Himmemblau, D. M. ,Edgar,T. F.. *Optimization of chemical processes*, Ed. Mc Graw - Hill. 1989
- [6] Jennings,L.S.; Teo,K.L.; Wang, F.Y.; Yu,Q. Optimal protein separation. *Computes chem. Engng*, pag. 5, 19, 567 - 573. 1995
- [7] Leser,E.W.; Asenjo, J.A. Building an expert system to assist the rational selection of large scale protein purification processes. *Biotechnology and Biochemical Engineering*. Pag. 260 - 266. 1995
- [8] Lienqueo, M.E.; Salgado, J.C.; Asenjo, J.A. Design of an expert system for selection of protein purification processes: comparison between different selection criteria. *Computer applications in biotechnology* pag. 321-326. 1998.

- [9] Lienqueo, M.E.; Leser, E.W.; Asenjo, J.A. An expert system for the selection and synthesis of multistep protein separation processes. *Computer chem. Engng*, 20 (suppl) S189 - S194. 1996
- [10] Pinto J. M., Optimización de procesos (anotaciones de clase), Universidad de São Paulo, Brasil. 1998
- [11] Reklaitis, G. V.; Ravindran, A.; Ragsdell, K. M., *Engineering Optimization Methods and Application*, Ed. Wiley, Ney York. 1983
- [12] [http://www.nalusda.gov/bic/Education\\_res/iastate.info/bio3.html](http://www.nalusda.gov/bic/Education_res/iastate.info/bio3.html);  
*Somatotropina*
- [13] <http://www.tracescientific.com/Biosciences.html#BSA>; *Bovine Serum Albumin*

## GLOSARIO

ALBUMINAS	Proteína soluble en agua y en soluciones salinas sin aminoácidos distintivos.
Bsa,.	Albúmina de suero bovino, proteína clásica.
CROMATOGRAFÍA	Proceso que permite la identificación de mezclas por la separación de alguno o todos sus componentes en zonas concentradas o en fases diferentes a aquellas en que estaban originalmente presentes.
Ecoli_s8.-	Nombre referencial de una proteína producida por la fermentación de la Escherichia coli, tiene 13 contaminantes.
GLOBULINAS	Escasamente solubles en agua pero solvante en soluciones salinas.
LISIS ENZIMÁTICA	Permeabilización de la pared celular a través de la hidrólisis enzimática parcial o total.



OPTIMIZACIÓN	Buscar el máximo de eficiencia y minimizar costos y tiempos, matemáticamente se representa mediante una función.
OVALBUMIN	Albúmina de huevo
"Refolding"	Reflujo del líquido con cuerpos incluidos, tiene la finalidad de concentrar la mezcla de proteínas.
SBTI.-	Inhibidor de tripsina
SIMPLEX	Método de optimización, se representa en una tabla la restricciones y la función objetivo, se usa para encontrar una solución factible.
SOLVER	Algoritmo interno de optimización
THAUMATINA.	Edulzante industrial

## ANEXOS

**ANEXO I**

**INDICE DE TABLAS Y FIGURAS**

## INDICE DE TABLAS

Tabla 1.	Expresiones y parámetros para las técnicas cromatograficas	49
Tabla 2.	Propiedades fisicoquímicas de la mezcla de proteínas para el ejemplo 1	61
Tabla 3.	Técnicas Cromatograficas para Ejemplo 1	62
Tabla 4.	Tiempo de Retención (Kd), para el ejemplo 1	63
Tabla 5.	Factor de Desvío (DF), para el ejemplo 1	63
Tabla 6.	Factor de Concentración (CF), para el ejemplo 1	64
Tabla 7.	Concentraciones resultantes de las proteínas del ejemplo 1	65
Tabla 8.	Propiedades fisicoquímicas de las proteínas del ejemplo 2	66
Tabla 9.	Técnicas cromatográficas usados en el ejemplo 2 y 3	67
Tabla 10.	Concentraciones resultantes del ejemplo 2	68
Tabla 11.	Propiedades fisicoquímicas de las proteínas para el ejemplo 3	69
Tabla 12.	Concentraciones resultantes del ejemplo 3	70
Tabla 13.	Técnicas usadas en la Purificación del "Somatotropin"	70
Tabla 14.	Secuencia sugerida por el sistema especialista en el ejemplo 1	71
Tabla 15.	Secuencia sugerida por el modelo de optimización en ejemplo 1	71
Tabla 16.	Secuencia sugerida por el sistema especialista en el ejemplo 2	73
Tabla 17.	Secuencia sugerida por el modelo de optimización en ejemplo 2	73

## INDICE DE FIGURAS

Figura 1. Procesos Biotecnologicos	18
Figura 2. Obtención del suero de albúmina bovino	23
Figura 3. Representación del método de busca lineal	27
Figura 4. Representación en Arbol	28
Figura 5. Representación en Red	29
Figura 6. Método de Búsqueda	30
Figura 7. Enumeración Dinámica	31
Figura 8. Estrategia de Busca Degth First	32
Figura 9. Estrategia de busca Breadth First	33
Figura 10. Diagrama de bloques del "GAMS"	43
Figura 11. Representación de Picos cromatograficas	50
Figura 12. Representación de Técnicas Cromatograficas	65
Gráfico 1.	72
Gráfico 2	74

## **ANEXO II**

### **DATOS DEL EJEMPLO 1**

#### **PURIFICACION DE SERUM FROM BOVINE ALBUMINA**

Archivo de entrada "datos1".txt

**\$TITLE Purification of protein**

**SET**

I CHROMATOGRAPHY STEPS /1\*12/  
p PROTEINS /p1\*p4/  
A PROPERTY /A1\*A7/  
k ORDER /1\*12/;

**ALIAS** (p,pp)  
(k,kk)

**TABLE PROP(p,A) VALUE OF PROPERTY A FOR FOR PROTEIN p**

	A1	A2	A3	A4	A5	A6	A7
p1	67000	0.86	1.03	-0.14	-1.16	-1.68	-2.05
p2	43800	0.54	1.40	-0.76	-1.65	-2.20	-2.36
p3	24500	0.90	1.22	-0.76	-1.54	-2.17	-2.13
p4	22200	0.89	1.94	1.90	1.98	1.87	0.91;

\*where

\*p1 PROTEIN 'SERUM FROM BOVINE ALBUMIN'

\*p2 PROTEIN 'OVALBUMIN'

\*p3 PROTEIN 'SOYBEAN TRYPSIN INHIBITOR'

\*p4 PROTEIN 'THAUMATIN';

\*AND

\*A1 MOLECULAR WEIGHT (Da)

\*A2 HIDROPHOBICITY

\*A3 pH 4.0

\*A4 pH 5.0

\*A5 pH 6.0

\*A6 pH 7,0

\*A7 pH 8.0;

scalar IP /1/;

**PARAMETER**

SIG(i) VALUE OF PEAK WIDTH

/1	0.15
2	0.15
3	0.15
4	0.15
5	0.15
6	0.15
7	0.15
8	0.15
9	0.15
10	0.15
11	0.22
12	0.46/

**Kd(i,p) RETENTION TIME FOR PRODUCT p IN CHROMATOGRAPHIC STEP i;**

$$\begin{aligned} Kd('1',p) &= ((7383 * ABS(PROP(p,'A3')/PROP(p,'A1')))/(1+15844 * \\ &ABS(PROP(p,'A3')/PROP(p,'A1'))))\$ (PROP(p,'A3') LT 0)+0\$ (PROP(p,'A3') GT 0); \\ Kd('2',p) &= ((7383 * ABS(PROP(p,'A4')/PROP(p,'A1')))/(1+15844 * \\ &ABS(PROP(p,'A4')/PROP(p,'A1'))))\$ (PROP(p,'A4') LT 0)+0\$ (PROP(p,'A4') GT 0); \\ Kd('3',p) &= ((7383 * ABS(PROP(p,'A5')/PROP(p,'A1')))/(1+15844 * \\ &ABS(PROP(p,'A5')/PROP(p,'A1'))))\$ (PROP(p,'A5') LT 0)+0\$ (PROP(p,'A5') GT 0); \\ Kd('4',p) &= ((7383 * ABS(PROP(p,'A6')/PROP(p,'A1')))/(1+15844 * \\ &ABS(PROP(p,'A6')/PROP(p,'A1'))))\$ (PROP(p,'A6') LT 0)+0\$ (PROP(p,'A6') GT 0); \\ Kd('5',p) &= ((7383 * ABS(PROP(p,'A7')/PROP(p,'A1')))/(1+15844 * \\ &ABS(PROP(p,'A7')/PROP(p,'A1'))))\$ (PROP(p,'A7') LT 0)+0\$ (PROP(p,'A7') GT 0); \\ Kd('6',p) &= ((5972 * ((PROP(p,'A3')/PROP(p,'A1')))/(1+17065 * \\ &(PROP(p,'A3')/PROP(p,'A1'))))\$ (PROP(p,'A3') GT 0)+0\$ (PROP(p,'A3') LT 0); \\ Kd('7',p) &= ((5972 * ((PROP(p,'A4')/PROP(p,'A1')))/(1+17065 * \\ &(PROP(p,'A4')/PROP(p,'A1'))))\$ (PROP(p,'A4') GT 0)+0\$ (PROP(p,'A4') LT 0); \\ Kd('8',p) &= ((5972 * ((PROP(p,'A5')/PROP(p,'A1')))/(1+17065 * \\ &(PROP(p,'A5')/PROP(p,'A1'))))\$ (PROP(p,'A5') GT 0)+0\$ (PROP(p,'A5') LT 0); \\ Kd('9',p) &= ((5972 * ((PROP(p,'A6')/PROP(p,'A1')))/(1+17065 * \\ &(PROP(p,'A6')/PROP(p,'A1'))))\$ (PROP(p,'A6') GT 0)+0\$ (PROP(p,'A6') LT 0); \end{aligned}$$



$Kd('10',p) = ((5972 * ((PROP(p,'A7'))/PROP(p,'A1')))/(1+17065 * (PROP(p,'A7'))/PROP(p,'A1')))) * (PROP(p,'A7') > 0) + 0 * (PROP(p,'A7') < 0);$   
 $Kd('11',p) = (1 - PROP(p,'A2')) / 1.5;$   
 $Kd('12',p) = -0.4691 * \log_{10}(PROP(p,'A1')) + 2.3902;$

**PARAMETER**

DF(i,p) DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i OF PROTEIN p;

$$DF(i,p) = ABS(\sum(pp\$ (ORD(pp) EQ IP), Kd(i,pp)) - Kd(i,p));$$

**PARAMETER**

CF(i,p) CONTAMINANT CONCENTRATION AFTER A CHROMATOGRAPHIC STEP;

$$CF(i,p) = 1;$$

$$CF(i,p) * ((DF(i,p) \ge (SIG(i)/10)) \text{ AND } (DF(i,p) < (SIG(i)/2))) = 1.02 * (SQR(SIG(i)) - 2 * SQR(DF(i,p))) / SQR(SIG(i));$$

$$CF(i,p) * ((DF(i,p) \ge (SIG(i)/2)) \text{ AND } (DF(i,p) < (SIG(i)))) = 1.02 * 2 * SQR(SIG(i)) - DF(i,p) / SQR(SIG(i));$$

$$CF(i,p) * (DF(i,p) \ge (SIG(i))) = 0.02;$$

**PARAMETER**

con(p) VALUE DE INITIAL CONCENTRATION

/p1 2.0

p2 2.0

p3 2.0

p4 2.0/

**PARAMETER**

U value maxime of initial concentration;

$$U = 2 * \sum(p\$ (ord(p) EQ IP), con(p));$$

\$ include 'protein1.txt'

## **ANEXO III**

### **MODELOS ESTADÍSTICO (RESULTADOS)**

**MODEL STATISTICS OF EXAMPLE 1.**

BLOCKS OF EQUATIONS 14 SINGLE EQUATIONS 860  
BLOCKS OF VARIABLES 5 SINGLE VARIABLES 729  
NON ZERO ELEMENTS 6460 DISCRETE VARIABLES 156

GENERATION TIME = 1.100 SECONDS

EXECUTION TIME = 1.150 SECONDS VERID MW2-25-087

Purification of protein

Solution Report SOLVE PROTEIN1 USING MIP FROM LINE 166

GAMS 2.25.087

386/486 DOS

**SOLVE SUMMARY**

MODEL PROTEIN1 OBJECTIVE S  
TYPE MIP DIRECTION MINIMIZE  
SOLVER OSL FROM LINE 166

\* SOLVER STATUS 1 NORMAL COMPLETION

\* MODEL STATUS 1 OPTIMAL

\* OBJECTIVE VALUE 3.0000

SOURCE USAGE, LIMIT 39.098 1000000.000

ERATION COUNT, LIMIT 2081 1000000

ading option file C:\GAMS\OSL.OPT

er supplied option file:

ethod simplex

ategy 48

esolve -1

epreproc 3

k space requested by user -- 10.00 Mb

k space requested by solver -- 1.71 Mb

OSL step: Reading data            0.22 Seconds

OSL step: Scale                    0.00 Seconds

Range of matrix coefficients:

before scaling :    0.0019 - 12.0000

after scaling :     0.0016 - 1.0000

OSL step: Crash                   0.05 Seconds

Crash option : 1

OSL step: Simplex                0.05 Seconds

The dual algorithm has been chosen

Status: Successful (optimal)

Iterations:    27

Objective :    1.1995

OSL step: Preprocessor           5.72 Seconds

Iterations:    288

OSL step: Branch&Bound        6.20 Seconds

Status: Successful (optimal)

Iterations:    1486

Nodes :       2

Objective :    3.0000

OSL step: Primal Simplex        0.00 Seconds

Status: Successful (optimal)

Iterations:    0

Objective :    3.0000

## TRADUCCIÓN DEL MODELO ESTADÍSTICO DEL EJEMPLO 1

Bloque de ecuaciones : 14 Ecuaciones solas 860  
Bloque de variables : 5 Variables 729  
Elementos no ceoros : 6460 Variables discretas 156

Objetivo

Tiempo de generación : 1.100 Seg  
Tiempo de ejecución : 1.150 Seg

### Purificación de proteínas

Solución reportada. Resolver protein1 usando MIP de la linea 166

### SUMARIO

Modelo	:	protein1	Objetivo	:	S
Tipo	:	MIP	Dirección	:	minimizar
Solver	:	OSL	Linea	:	166

Estado de Solver : Normal completo  
Estado del modelo : óptimo  
Valor objetivo : 3.00

Estado de trabajo requerido por el usuario 10.0  
Estado de trabajo requerido por el solver 1.71

### Paso OSL : Simplex

Estado : Optimo  
Iteraciones : 27  
Objetivo : 1.1995

### Paso OSL preproceso

Iteraciones : 288

Paso OSL : Branch & Bound 6.20  
Estado : Optimo Nudos : 2

Iteraciones : 1486

Objetivo : 3

Paso OSL : Simplex

Estado : Optimo

Iteraciones : 0

Objetivo : 3.0

**MODEL STATISTICS OF EXAMPLE 2**

BLOCKS OF EQUATIONS    14    SINGLE EQUATIONS    4768  
BLOCKS OF VARIABLES    5    SINGLE VARIABLES    4394  
NON ZERO ELEMENTS    38247    DISCRETE VARIABLES    460

GENERATION TIME    =    7.140 SECONDS

EXECUTION TIME    =    7.300 SECONDS    VERID MW2-25-087  
\_GAMS 2.25.087 386/486 DOS                    10/30/95 19:27:02 PAGE 453  
PURIFICATION OF ECOLI\_S8  
Solution Report    SOLVE PROTEIN1 USING MIP FROM LINE 216

**S O L V E   S U M M A R Y**

MODEL PROTEIN1            OBJECTIVE S  
TYPE MIP                    DIRECTION MINIMIZE  
SOLVER OSL                  FROM LINE 216

\*\*\*\* SOLVER STATUS    1 NORMAL COMPLETION  
\*\*\*\* MODEL STATUS    1 OPTIMAL  
\*\*\*\* OBJECTIVE VALUE            5.0000

RESOURCE USAGE, LIMIT    698.430 1000000.000  
ITERATION COUNT, LIMIT    10189    1000000  
Could not find option file C:\GAMS\OSL.OPT  
Using defaults instead.

Work space requested by user -- 10.00 Mb  
Work space requested by solver -- 9.28 Mb

OSL step: Reading data            2.58 Seconds

OSL step: Scale                    2.75 Seconds

Range of matrix coefficients:  
before scaling :    0.0000 -    20.0000

after scaling : 0.0000 - 1.0000

OSL step: Presolve 1.48 Seconds

Size reduction:

rows : 4768 (old) 4766 (new)  
columns : 4394 (old) 4394 (new)  
nonzeroes: 38247 (old) 38225 (new)

OSL step: Crash 1.32 Seconds

Crash option : 1

OSL step: Primal Simplex 7.25 Seconds

Status: Successful (optimal)

Iterations: 139

Objective : 1.4538

OSL step: Branch&Bound 676.51 Seconds

Status: Successful (optimal)

Iterations: 10050

Nodes : 191

Objective : 5.0000

OSL step: Postsolve 1.76 Seconds

OSL step: Primal Simplex 0.71 Seconds

Status: Successful (optimal)

Iterations: 0

Objective : 5.0000



## TRADUCCIÓN DEL MODELO ESTADÍSTICO DEL EJEMPLO 2

Bloque de ecuaciones	: 14	Ecuaciones solas	4768
Bloque de variables	: 5	Variables	4394
Elementos no ceoros :	: 38247	Variables discretas	460

Tiempo de generación : 7.140 Seg  
 Tiempo de ejecución : 7.300 Seg

Purificación de Ecoli-S8

Solución reportada. Resolver protein1 usando MIP de la linea 216

### SUMARIO

Modelo	:	protein1	Objetivo	:	S
Tipo	:	MIP	Dirección	:	minimizar
Solver	:	OSL	Linea	:	216

Estado de Solver : Normal completo  
 Estado del modelo : óptimo  
 Valor objetivo : 5.00

Estado de trabajo requerido por el usuario 10.0  
 Estado de trabajo requerido por el solver 9.28

Paso OSL : Pre resuelto

Reducción de tamaño		1.48 seg
Filas	4768 (OLD)	4766 New
Columnas	4394 (OLD)	4394 New
Nonzeros	38247 (OLD)	38225 New

Paso OSL : Simplex  
 Estado : Optimo  
 Iteraciones : 139  
 Objetivo : 1.4538

Paso OSL : Branch & Bound 676.51  
Estado : Optimo Nodos : 191  
Iteraciones : 10050 Objetivo : 5

Posterior Resuelto 1.76

Paso OSL : Simplex  
Estado : Optimo  
Iteraciones : 0  
Objetivo : 5.0

### MODEL STATISTICS OF PROBLEM 3

BLOCKS OF EQUATIONS 14 SINGLE EQUATIONS 4050  
BLOCKS OF VARIABLES 5 SINGLE VARIABLES 3819  
NON ZERO ELEMENTS 28250 DISCRETE VARIABLES 276

GENERATION TIME = 6.260 SECONDS

EXECUTION TIME = 6.370 SECONDS VERID MW2-25-087

\_Selection of Multistep Protein Purification Processes

Solution Report SOLVE PROTEIN1 USING MIP FROM LINE 225

### SOLVE SUMMARY

MODEL PROTEIN1 OBJECTIVE S  
TYPE MIP DIRECTION MINIMIZE  
SOLVER OSL FROM LINE 225

\*\*\*\* SOLVER STATUS 1 NORMAL COMPLETION

\*\*\*\* MODEL STATUS 1 OPTIMAL

\*\*\*\* OBJECTIVE VALUE 3.0000

RESOURCE USAGE, LIMIT 219.098 1000000.000

ITERATION COUNT, LIMIT 12787 1000000

Reading option file C:\GAMS\OSL.OPT

User supplied option file:

>method simplex

>strategy 48

>presolve -1

>bbpreproc 3

Work space requested by user -- 10.00 Mb

Work space requested by solver -- 7.09 Mb

OSL step: Reading data 0.82 Seconds

OSL step: Scale 0.21 Seconds

Range of matrix coefficients:

before scaling : 0.0000 - 50.0000

after scaling : 0.0001 - 1.0000

OSL step: Crash 0.11 Seconds

Crash option : 1

OSL step: Simplex 0.50 Seconds

The dual algorithm has been chosen

Status: Successful (optimal)

Iterations: 40

Objective : 1.0579

OSL step: Preprocessor 45.92 Seconds

Iterations: 1052

OSL step: Branch&Bound 142.20 Seconds

Status: Successful (optimal)

Iterations: 11695

Nodes : 10

Objective : 3.0000

OSL step: Primal Simplex 0.11 Seconds

Status: Successful (optimal)

Iterations: 0

Objective : 3.0000

## TRADUCCIÓN DEL MODELO ESTADÍSTICO DEL EJEMPLO 1

Bloque de ecuaciones : 14 Ecuaciones solas 860  
Bloque de variables : 5 Variables 729  
Elementos no ceoros : 6460 Variables discretas 156

Tiempo de generación : 1.090 Seg  
Tiempo de ejecución : 1.20 Seg

Purificación de proteínas

Solución reportada. Resolver protein1 usando MIP de la línea 170

### SUMARIO

Modelo : protein1 Objetivo : S  
Tipo : MIP Dirección : minimizar  
Solver : OSL Línea : 225

Estado de Solver : Normal completo  
Estado del modelo : óptimo  
Valor objetivo : 3.00

Estado de trabajo requerido por el usuario 10.0  
Estado de trabajo requerido por el solver 7.09

Paso OSL : Simplex  
Estado : Optimo  
Iteraciones : 40  
Objetivo 1.0579

Paso OSL : preproceso  
Iteraciones : 1052

Paso OSL : Branch & Bound 142.20  
Estado : Optimo Nudos : 10

Iteraciones : 11695

Objetivo : 3

Paso OSL : Simplex

Estado : Optimo

Iteraciones : 0

Objetivo : 3.0

## **ANEXO IV**

### **CORRIDAS PRELIMINARES DEL PROGRAMA**

## OBSERVACIÓN:

Para simplificar el número de páginas del resultado obtenido para las corridas previas, se vio conveniente dar los comandos de eliminación de ecuaciones que se forman internamente, eso quiere decir, que las ecuaciones no aparecen mas en la pantalla la ventaja es que ocupan menor espacio de memoria y es mas rapido, no existe variación en los resultados.

En archivo de entrada de datos "datos 1", "datos2" y "datos 3" se agrego al inicio:

```
$OFFSYMXREF
```

```
$OFFSYMLIST
```

y en el archivo de entrada del programa "protein1" se cambio  
limrow = 200            por limrow = 0 y se agrego  
limcol = 0

Cabe notar que en este anexo se presentan las corridas preliminares para los tres ejemplos, las respuestas se grabaron dependiendo de lo que uno esta buscando: ejemplo para los archivos "datos1" ou "datos1tesis"  
Resultdatos1p4        proteína a ser purificada p4  
Resultdatos1p170    proteína a ser purificada p1 pero la pureza que se requiere es 70%



purification of protein (ejemplo 1. purificación del suero de albúmina bovino )

10/30/95 10:19:09 PAGE 1

GAMS 2.25

087 386/486 DOS

```

4
5
6 SET
7 I CHROMATOGRAPHY STEPS /1*12/
8 p PROTEINS /p1*p4/
9 A PROPERTY /A1*A7/
10 k ORDER /1*12/;
11
12 ALIAS (p,pp)
13 (k,kk)
14
15 TABLE PROP(p,A) VALUE OF PROPERTY A FOR FOR PROTEIN p
16
17 A1 A2 A3 A4 A5 A6 A7
18 p1 67000 0.86 1.03 -0.14 -1.16 -1.68 -2.05
19 p2 43800 0.54 1.40 -0.76 -1.65 -2.20 -2.36
20 p3 24500 0.90 1.22 -0.76 -1.54 -2.17 -2.13
21 p4 22200 0.89 1.94 1.90 1.98 1.87 0.91;
22
23

```

```

24 *where
25 *p1 PROTEIN 'SERUM FROM BOVINE ALBUMIN'
26 *p2 PROTEIN 'OVALBUMIN'
27 *p3 PROTEIN 'SOYBEAN TRYPSIN INHIBITOR'
28 *p4 PROTEIN 'THAUMATIN';
29

```

```

30 *AND
31 *A1 MOLECULAR WEIGHT (Da)
32 *A2 HIDROPHOBICITY
33 *A3 pH 4.0
34 *A4 pH 5.0
35 *A5 pH 6.0
36 *A6 pH 7,0
37 *A7 pH 8.0;
38

```

```

39 scalar IP /1/;
40

```

```

41 PARAMETER
42 SIG(i) VALUE OF PEAK WIDTH
43 /1 0.15
44 2 0.15
45 3 0.15
46 4 0.15
47 5 0.15
48 6 0.15
49 7 0.15
50 8 0.15
51 9 0.15
52 10 0.15
53 11 0.22
54 12 0.46/
55

```

```

56 Kd(i,p) RETENTION TIME FOR PRODUCT p IN CHROMATOGRAPHIC STEP i;
57 Kd('1',p) = ((7383 * ABS(PROP(p,'A3')/PROP(p,'A1')))/(1+15844 *

```

ABS(PROP(p, 'A3')/PROP(p, 'A1'))))\$(PROP(p, 'A3') LT 0)+0\$(PROP(p, 'A3') GT 0);  
 ~purification of protein

10:19:09 PAGE 2

10/30/95

.087 386/486 DOS

GAMS 2.25

58 Kd('2',p) = ((7383 \* ABS(PROP(p, 'A4')/PROP(p, 'A1')))/(1+15844 \*  
 ABS(PROP(p, 'A4')/PROP(p, 'A1'))))\$(PROP(p, 'A4') LT 0)+0\$(PROP(p, 'A4') GT 0);  
 59 Kd('3',p) = ((7383 \* ABS(PROP(p, 'A5')/PROP(p, 'A1')))/(1+15844 \*  
 ABS(PROP(p, 'A5')/PROP(p, 'A1'))))\$(PROP(p, 'A5') LT 0)+0\$(PROP(p, 'A5') GT 0);  
 60 Kd('4',p) = ((7383 \* ABS(PROP(p, 'A6')/PROP(p, 'A1')))/(1+15844 \*  
 ABS(PROP(p, 'A6')/PROP(p, 'A1'))))\$(PROP(p, 'A6') LT 0)+0\$(PROP(p, 'A6') GT 0);  
 61 Kd('5',p) = ((7383 \* ABS(PROP(p, 'A7')/PROP(p, 'A1')))/(1+15844 \*  
 ABS(PROP(p, 'A7')/PROP(p, 'A1'))))\$(PROP(p, 'A7') LT 0)+0\$(PROP(p, 'A7') GT 0);  
 62 Kd('6',p) = ((5972 \* ((PROP(p, 'A3')/PROP(p, 'A1')))/(1+17065 \*  
 PROP(p, 'A3')/PROP(p, 'A1'))))\$(PROP(p, 'A3') GT 0)+0\$(PROP(p, 'A3') LT 0);  
 63 Kd('7',p) = ((5972 \* ((PROP(p, 'A4')/PROP(p, 'A1')))/(1+17065 \*  
 PROP(p, 'A4')/PROP(p, 'A1'))))\$(PROP(p, 'A4') GT 0)+0\$(PROP(p, 'A4') LT 0);  
 64 Kd('8',p) = ((5972 \* ((PROP(p, 'A5')/PROP(p, 'A1')))/(1+17065 \*  
 PROP(p, 'A5')/PROP(p, 'A1'))))\$(PROP(p, 'A5') GT 0)+0\$(PROP(p, 'A5') LT 0);  
 65 Kd('9',p) = ((5972 \* ((PROP(p, 'A6')/PROP(p, 'A1')))/(1+17065 \*  
 PROP(p, 'A6')/PROP(p, 'A1'))))\$(PROP(p, 'A6') GT 0)+0\$(PROP(p, 'A6') LT 0);  
 66 Kd('10',p) = ((5972 \* ((PROP(p, 'A7')/PROP(p, 'A1')))/(1+17065 \*  
 (PROP(p, 'A7')/PROP(p, 'A1'))))\$(PROP(p, 'A7') GT 0)+0\$(PROP(p, 'A7') LT 0);  
 67 Kd('11',p) = (1-PROP(p, 'A2')/1.5);  
 68 Kd('12',p) = -0.4691 \* log10(PROP(p, 'A1'))+2.3902;

70 PARAMETER

71 DF(i,p) DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i OF PROTEIN p;  
 72 DF(i,p) = ABS(sum(pp\$(ORD(pp) EQ IP),Kd(i,pp))-Kd(i,p))

73

74 PARAMETER

75 CF(i,p) CONTAMINANT CONCENTRATION AFTER A CHROMATOGRAPHIC STEP;

76 CF(i,p)=1;

77 CF(i,p)\$((DF(i,p) GE (SIG(i)/10)) AND (DF(i,p) LT (SIG(i)/2)))=1.02\*(SQR  
 (SIG(i))-2\*SQR(DF(i,p)))/SQR(SIG(i));

78 CF(i,p)\$((DF(i,p) GE (SIG(i)/ 2)) AND (DF(i,p) LT (SIG(i) )))=1.02\*2\*SQ  
 R(SIG(i)-DF(i,p))/SQR(SIG(i));

79 CF(i,p)\$((DF(i,p) GE (SIG(i))))=0.02;

80

81

82

83 PARAMETER

84 con(p) VALUE DE INITIAL CONCENTRATION

85 /p1 2.0

86 p2 2.0

87 p3 2.0

88 p4 2.0/

89 PARAMETER

90 U value maxime of initial concentration;

91 U=2\*sum(p\$(ord(p) EQ IP),con(p));

92

INCLUDE C:\GAMS\PROTEIN1.TXT

94 display kd,df,cf;

95

96

97

98 Variables S;

```

99 positive variables C(p,k), C1(i,p,k);
100 Binary Variables Y(i,k), Z(k);
101 Equations OBJ, EQ1, EQ2, EQ3, EQ4, EQ5, EQ6, EQ7, EQ8,eq9,eq4a,eq4b
102         eq7a
103         eq7b
104         ;
105
106 *Objective function
107
108 c.fx(p,'1') = con(p);
109
110 OBJ.. S=E=SUM(k,ord(k)*z(k));
111
^Purification of protein

```

10:19:09 PAGE 3

10/30/95

.087 386/486 DOS

GAMS 2.25

```

112 *Subject to
113
114 EQ1(k)..SUM(i,Y(i,k))=L=1;
115
116 EQ2(i)..SUM(k,Y(i,k))=L=1;
117
118 EQ3(k)$ (ord(k) lt card(k)).. SUM(i,Y(i,k+1))=L=SUM(i,Y(i,k));
119
120 EQ4(p,k)$ (ord(k) ge 3)..C(p,k)-SUM(i,CF(i,p)*C1(i,p,k-1)) =g= -U*(1-sum(
i,Y(i,k-1)));
121
122 EQ4b(p,k)$ (ord(k) ge 3)..C(p,k)-SUM(i,CF(i,p)*C1(i,p,k-1)) =l= U*(1-sum(
i,Y(i,k-1)));
123
124 EQ4a(p,'2')..C(p,'2')=E=SUM(i,CF(i,p)*y(i,'1')*con(p));
125
126 EQ5(i,p,k)$ (ord(k) ge 2)..C1(i,p,k)=L=con(p)*Y(i,k-1);
127
128 EQ6(p,k)$ (ord(k) ge 2)..C(p,k)=E=SUM(i,C1(i,p,k));
129
130 EQ7(k)$ (ord(k) lt card(k)).. Z(k)=G=SUM(i,Y(i,k))-SUM(i,Y(i,k+1));
131
132 eq7a(k,kk)$ (ord(kk) le ord(k)).. SUM(i,Y(i,kk)) =g= z(k);
133
134 * eq7a(k).. SUM((i,kk)$ (ord(kk) le ord(k)),Y(i,kk)) =g= ord(k)*z(k);
135
136 eq7b(k,kk)$ (ord(kk) gt ord(k)).. SUM(i,Y(i,kk)) + z(k) =l= 1;
137
138 EQ8..SUM(k,Z(k))=E=1;
139
140 EQ9(p,k)$ (ORD(p) EQ IP).. C(p,k+1)=G=0.70*(SUM(pp,C(pp,k+1)))-U*(1-Z(k))
;
141
142 option iterlim = 1000000
143         reslim = 1000000
144         limrow = 0
145         solprint = off
146         optcr = 0
147         limcol = 0
148         optca = 0;
149

```

```

150
151
152 MODEL PROTEIN1 /OBJ,
153     EQ1, EQ2, EQ3,
154     EQ4,
155     eq4a,eq4b
156     EQ5,
157     EQ6,
158     EQ7, eq7a, eq7b
159     EQ8
160     eq9/;
161
162
163 protein1.workspace=10;
164
165 protein1.optfile=1
^Purification of protein
    
```

10:19:09 PAGE 4

10/30/95

.087 386/486 DOS

GAMS 2.25

```

166
167
168
169 SOLVE PROTEIN1 USING MIP MINIMIZING S;
170
171 display C.1, C1.1, Z.1, Y.1;
^Purification of protein
    
```

10:19:09 PAGE 5  
Include File Summary

10/30/95

.087 386/486 DOS

GAMS 2.25

SEQ	GLOBAL TYPE	PARENT	LOCAL	FILENAME
1	1 INPUT	0	0	C:\GAMS\DZTESIS.TXT
2	93 INCLUDE	1	93	.C:\GAMS\PROTEIN1.TXT

COMPILATION TIME = 0.110 SECONDS VERID MW2-25-087  
^Purification of protein

10:19:09 PAGE 6  
E x e c u t i o n

10/30/95

.087 386/486 DOS

GAMS 2.25

```

-----
94 PARAMETER KD          RETENTION TIME FOR PRODUCT p IN CHROMATOGR
APHIC STEP i

      P1          P2          P3          P4
    
```

2	0.015	0.100	0.154	
3	0.100	0.174	0.233	
4	0.132	0.206	0.272	
5	0.152	0.215	0.270	
6	0.073	0.124	0.161	0.209
7				0.208
8				0.211
9				0.206
10				0.144
11	0.427	0.640	0.400	0.407
12	0.126	0.213	0.331	0.351

----- 94 PARAMETER DF OF PROTEIN p DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i

	P2	P3	P4
2	0.086	0.139	0.015
3	0.074	0.132	0.100
4	0.074	0.140	0.132
5	0.062	0.118	0.152
6	0.051	0.088	0.137
7			0.208
8			0.211
9			0.206
10			0.144
11	0.213	0.027	0.020
12	0.087	0.205	0.225

----- 94 PARAMETER CF GRAPHIC STEP CONTAMINANT CONCENTRATION AFTER A CHROMATO

	P1	P2	P3	P4
1	1.000	1.000	1.000	1.000
2	1.000	0.377	0.012	1.000
3	1.000	0.525	0.029	0.224
4	1.000	0.523	0.010	0.028
5	1.000	0.666	0.094	0.020
6	1.000	0.786	0.348	0.016
7	1.000	1.000	1.000	0.020
8	1.000	1.000	1.000	0.020
9	1.000	1.000	1.000	0.020
10	1.000	1.000	1.000	0.003
11	1.000	0.002	0.990	1.000
12	1.000	0.948	0.615	0.532

Purification of protein

10:19:09 PAGE 7  
Model Statistics SOLVE PROTEIN1 USING MIP FROM LINE 169

10/30/95

.087 386/486 DOS

GAMS 2.25

MODEL STATISTICS

BLOCKS OF EQUATIONS	14	SINGLE EQUATIONS	860
BLOCKS OF VARIABLES	5	SINGLE VARIABLES	729

NON ZERO ELEMENTS            6460            DISCRETE VARIABLES            156

GENERATION TIME            =            0.930 SECONDS

EXECUTION TIME            =            0.990 SECONDS            VERID MW2-25-087

^Purification of protein

10:19:09 PAGE            8            10/30/95

Solution Report            SOLVE PROTEIN1 USING MIP FROM LINE 169

.087 386/486 DOS

GAMS 2.25

S O L V E            S U M M A R Y

MODEL    PROTEIN1            OBJECTIVE    S

TYPE     MIP                DIRECTION    MINIMIZE

SOLVER   OSL                 FROM LINE    169

\*\*\*\* SOLVER STATUS            1 NORMAL COMPLETION

\*\*\*\* MODEL STATUS            1 OPTIMAL

\*\*\*\* OBJECTIVE VALUE                            2.0000

RESOURCE USAGE, LIMIT            4.781 1000000.000

ITERATION COUNT, LIMIT            105 1000000

Reading option file C:\GAMS\OSL.OPT

User supplied option file:

>method simplex

>strategy 48

>presolve -1

>bbpreproc 3

Work space requested by user    --    10.00 Mb

Work space requested by solver --    1.71 Mb

OSL step: Reading data                            0.17 Seconds

OSL step: Scale                                    0.06 Seconds

Range of matrix coefficients:

  before scaling :            0.0019 -            12.0000

  after scaling :            0.0016 -            1.0000

OSL step: Crash                                    0.00 Seconds

Crash option : 1

OSL step: Simplex                                    0.05 Seconds

The dual algorithm has been chosen

Status:            Successful (optimal)

Iterations:            26

Objective :            1.0443

OSL step: Preprocessor                            4.23 Seconds

Iterations: 79

OSL step: Branch&Bound 0.00 Seconds

Status: Successful (optimal)  
 Iterations: 0  
 Nodes : 0  
 Objective : 2.0000

OSL step: Primal Simplex 0.06 Seconds

Status: Successful (optimal)  
 Iterations: 0  
 Objective : 2.0000

\*Purification of protein

10:19:09 PAGE 9  
 Solution Report SOLVE PROTEIN1 USING MIP FROM LINE 169

10/30/95

.087 386/486 DOS

GAMS 2.25

\*\*\*\* REPORT SUMMARY :  
 0 NONOPT  
 0 INFEASIBLE  
 0 UNBOUNDED

\*Purification of protein

10:19:09 PAGE 10  
 Execution

10/30/95

.087 386/486 DOS

GAMS 2.25

---- 171 VARIABLE C.L

	1	2	3
P1	2.000	2.000	2.000
P2	2.000	1.047	0.548
P3	2.000	0.020	1.922761E-4
P4	2.000	0.056	0.002

---- 171 VARIABLE C1.L

	2	3
4 .P1	2.000	
4 .P2	1.047	
4 .P3	0.020	
4 .P4	0.056	
11.P1		2.000
11.P2		0.548
11.P3		1.922761E-4
11.P4		0.002

---- 171 VARIABLE Z.L

2 1.000

---- 171 VARIABLE Y.L

1 2

4 1.000

11 1.000

EXECUTION TIME = 0.170 SECONDS VERID MW2-25-087



\*Purification of protein(ejemplo 1. purificación del suero de albúmina bovino)

10/30/95 09:56:05 PAGE 1

.087 386/486 DOS

GAMS 2.25

```

4
5
6 SET
7 I CHROMATOGRAPHY STEPS /1*12/
8 p PROTEINS /p1*p4/
9 A PROPERTY /A1*A7/
10 k ORDER /1*12/;
11
12 ALIAS (p,pp)
13 (k,kk)
14
15 TABLE PROP(p,A) VALUE OF PROPERTY A FOR FOR PROTEIN p
16
17 A1 A2 A3 A4 A5 A6 A7
18 p1 67000 0.86 1.03 -0.14 -1.16 -1.68 -2.05
19 p2 43800 0.54 1.40 -0.76 -1.65 -2.20 -2.36
20 p3 24500 0.90 1.22 -0.76 -1.54 -2.17 -2.13
21 p4 22200 0.89 1.94 1.90 1.98 1.87 0.91;
22
23

```

```

24 *where
25 *p1 PROTEIN 'SERUM FROM BOVINE ALBUMIN'
26 *p2 PROTEIN 'OVALBUMIN'
27 *p3 PROTEIN 'SOYBEAN TRYPSIN INHIBITOR'
28 *p4 PROTEIN 'THAUMATIN';
29

```

```

30 *AND
31 *A1 MOLECULAR WEIGHT (Da)
32 *A2 HIDROPHOBICITY
33 *A3 pH 4.0
34 *A4 pH 5.0
35 *A5 pH 6.0
36 *A6 pH 7,0
37 *A7 pH 8.0;
38

```

```

39 scalar IP /4/;
40

```

```

41 PARAMETER
42 SIG(i) VALUE OF PEAK WIDTH
43 /1 0.15
44 2 0.15
45 3 0.15
46 4 0.15
47 5 0.15
48 6 0.15
49 7 0.15
50 8 0.15
51 9 0.15
52 10 0.15
53 11 0.22
54 12 0.46/;
55

```

```

56 Kd(i,p) RETENTION TIME FOR PRODUCT p IN CHROMATOGRAPHIC STEP i;
57 Kd('1',p) = ((7383 * ABS(PROP(p,'A3')/PROP(p,'A1')))/(1+15844 *

```

ABS(PROP(p, 'A3')/PROP(p, 'A1'))))\$(PROP(p, 'A3') LT 0)+0\$(PROP(p, 'A3') GT 0);  
 'purification of protein

09:56:05 PAGE 2 10/30/95

.087 386/486 DOS GAMS 2.25

```

58      Kd('2',p) = ((7383 * ABS(PROP(p, 'A4')/PROP(p, 'A1')))/(1+15844 *
ABS(PROP(p, 'A4')/PROP(p, 'A1'))))$(PROP(p, 'A4') LT 0)+0$(PROP(p, 'A4') GT 0);
59      Kd('3',p) = ((7383 * ABS(PROP(p, 'A5')/PROP(p, 'A1')))/(1+15844 *
ABS(PROP(p, 'A5')/PROP(p, 'A1'))))$(PROP(p, 'A5') LT 0)+0$(PROP(p, 'A5') GT 0);
60      Kd('4',p) = ((7383 * ABS(PROP(p, 'A6')/PROP(p, 'A1')))/(1+15844 *
ABS(PROP(p, 'A6')/PROP(p, 'A1'))))$(PROP(p, 'A6') LT 0)+0$(PROP(p, 'A6') GT 0);
61      Kd('5',p) = ((7383 * ABS(PROP(p, 'A7')/PROP(p, 'A1')))/(1+15844 *
ABS(PROP(p, 'A7')/PROP(p, 'A1'))))$(PROP(p, 'A7') LT 0)+0$(PROP(p, 'A7') GT 0);
62      Kd('6',p) = ((5972 * ((PROP(p, 'A3')/PROP(p, 'A1')))/(1+17065 * (
PROP(p, 'A3')/PROP(p, 'A1'))))$(PROP(p, 'A3') GT 0)+0$(PROP(p, 'A3') LT 0);
63      Kd('7',p) = ((5972 * ((PROP(p, 'A4')/PROP(p, 'A1')))/(1+17065 * (
PROP(p, 'A4')/PROP(p, 'A1'))))$(PROP(p, 'A4') GT 0)+0$(PROP(p, 'A4') LT 0);
64      Kd('8',p) = ((5972 * ((PROP(p, 'A5')/PROP(p, 'A1')))/(1+17065 * (
PROP(p, 'A5')/PROP(p, 'A1'))))$(PROP(p, 'A5') GT 0)+0$(PROP(p, 'A5') LT 0);
65      Kd('9',p) = ((5972 * ((PROP(p, 'A6')/PROP(p, 'A1')))/(1+17065 * (
PROP(p, 'A6')/PROP(p, 'A1'))))$(PROP(p, 'A6') GT 0)+0$(PROP(p, 'A6') LT 0);
66      Kd('10',p) = ((5972 * ((PROP(p, 'A7')/PROP(p, 'A1')))/(1+17065 * (
PROP(p, 'A7')/PROP(p, 'A1'))))$(PROP(p, 'A7') GT 0)+0$(PROP(p, 'A7') LT 0);
67      Kd('11',p) = (1-PROP(p, 'A2')/1.5);
68      Kd('12',p) = -0.4691 * log10(PROP(p, 'A1')+2.3902);
    
```

```

70  PARAMETER
71      DF(i,p) DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i OF PROTEIN p;
72      DF(i,p) = ABS(sum(pp$(ORD(pp) EQ IP), Kd(i,pp))-Kd(i,p))
    
```

```

73
74  PARAMETER
75      CF(i,p) CONTAMINANT CONCENTRATION AFTER A CHROMATOGRAPHIC STEP;
76      CF(i,p)=1;
77      CF(i,p)$((DF(i,p) GE (SIG(i)/10)) AND (DF(i,p) LT (SIG(i)/2)))=1.02*(SQR
(SIG(i))-2*SQR(DF(i,p)))/SQR(SIG(i));
78      CF(i,p)$((DF(i,p) GE (SIG(i)/ 2)) AND (DF(i,p) LT (SIG(i)  )))=1.02*2*SQ
R(SIG(i)-DF(i,p))/SQR(SIG(i));
79      CF(i,p)$((DF(i,p) GE (SIG(i))))=0.02;
    
```

```

80
81
82
83  PARAMETER
84      con(p) VALUE DE INITIAL CONCENTRATION
85
86      /p1 2.0
87      p2 2.0
88      p3 2.0
89      p4 2.0/
    
```

```

90  PARAMETER
91      U value maxime of initial concentration;
92      U=2*sum(p$(ord(p) EQ IP), con(p));
    
```

```

INCLUDE C:\GAMS\PROTEIN1.TXT
94 display kd,df,cf;
95
96
97
98 Variables S;
    
```

```

99 positive variables C(p,k), C1(i,p,k);
100 Binary Variables Y(i,k), Z(k);
101 Equations OBJ, EQ1, EQ2, EQ3, EQ4, EQ5, EQ6, EQ7, EQ8, eq9, eq4a, eq4b
102     eq7a
103     eq7b
104     ;
105
106 *Objective function
107
108 c.fx(p, '1') = con(p);
109
110 OBJ.. S=E=SUM(k, ord(k)*z(k));
111
Purification of protein

```

09:56:05 PAGE 3

10/30/95

.087 386/486 DOS

GAMS 2.25

```

112 *Subject to
113
114 EQ1(k)..SUM(i,Y(i,k))=L=1;
115
116 EQ2(i)..SUM(k,Y(i,k))=L=1;
117
118 EQ3(k)$ (ord(k) lt card(k))..SUM(i,Y(i,k+1))=L=SUM(i,Y(i,k));
119
120 EQ4(p,k)$ (ord(k) ge 3)..C(p,k)-SUM(i,CF(i,p)*C1(i,p,k-1)) =g= -U*(1-sum(
i,Y(i,k-1)));
121
122 EQ4b(p,k)$ (ord(k) ge 3)..C(p,k)-SUM(i,CF(i,p)*C1(i,p,k-1)) =l= U*(1-sum(
i,Y(i,k-1)));
123
124 EQ4a(p, '2')..C(p, '2')=E=SUM(i,CF(i,p)*y(i, '1')*con(p));
125
126 EQ5(i,p,k)$ (ord(k) ge 2)..C1(i,p,k)=L=con(p)*Y(i,k-1);
127
128 EQ6(p,k)$ (ord(k) ge 2)..C(p,k)=E=SUM(i,C1(i,p,k));
129
130 EQ7(k)$ (ord(k) lt card(k))..Z(k)=G=SUM(i,Y(i,k))-SUM(i,Y(i,k+1));
131
132 eq7a(k,kk)$ (ord(kk) le ord(k))..SUM(i,Y(i,kk)) =g= z(k);
133
134 * eq7a(k)..SUM((i,kk)$ (ord(kk) le ord(k)),Y(i,kk)) =g= ord(k)*z(k);
135
136 eq7b(k,kk)$ (ord(kk) gt ord(k))..SUM(i,Y(i,kk)) + z(k) =l= 1;
137
138 EQ8..SUM(k,Z(k))=E=1;
139
140 EQ9(p,k)$ (ORD(p) EQ IP)..C(p,k+1)=G=0.98*(SUM(pp,C(pp,k+1)))-U*(1-Z(k))
;
141
142 option iterlim = 1000000
143     reslim = 1000000
144     limrow = 0
145     solprint = off
146     optcr = 0
147     limcol = 0
148     optca = 0;
149

```

```

150
151
152 MODEL PROTEIN1 /OBJ,
153     EQ1, EQ2, EQ3,
154     EQ4,
155     eq4a,eq4b
156     EQ5,
157     EQ6,
158     EQ7, eq7a, eq7b
159     EQ8
160     eq9/;
161
162
163 protein1.workspace=10;
164
165 protein1.optfile=1
^Purification of protein
    
```

09:56:05 PAGE 4

10/30/95

.087 386/486 DOS

GAMS 2.25

```

166
167
168
169 SOLVE PROTEIN1 USING MIP MINIMIZING S;
170
171 display C.1, C1.1, Z.1, Y.1;
^Purification of protein
    
```

09:56:05 PAGE 5  
Include File Summary

10/30/95

.087 386/486 DOS

GAMS 2.25

SEQ	GLOBAL TYPE	PARENT	LOCAL	FILENAME
1	1 INPUT	0	0	C:\GAMS\DATOS1\TESIS.TXT
2	93 INCLUDE	1	93	.C:\GAMS\PROTEIN1.TXT

COMPILATION TIME = 0.110 SECONDS VERID MW2-25-087  
^Purification of protein

09:56:05 PAGE 6  
E x e c u t i o n

10/30/95

.087 386/486 DOS

GAMS 2.25

```

-----
94 PARAMETER KD          RETENTION TIME FOR PRODUCT p IN CHROMATOGR
APHIC STEP i
      P1          P2          P3          P4
    
```

2	0.015	0.100	0.154	
3	0.100	0.174	0.233	
4	0.132	0.206	0.272	
5	0.152	0.215	0.270	
6	0.073	0.124	0.161	0.209
7				0.208
8				0.211
9				0.206
10				0.144
11	0.427	0.640	0.400	0.407
12	0.126	0.213	0.331	0.351

----- 94 PARAMETER DF OF PROTEIN p DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i

	P1	P2	P3
2	0.015	0.100	0.154
3	0.100	0.174	0.233
4	0.132	0.206	0.272
5	0.152	0.215	0.270
6	0.137	0.086	0.049
7	0.208	0.208	0.208
8	0.211	0.211	0.211
9	0.206	0.206	0.206
10	0.144	0.144	0.144
11	0.020	0.233	0.007
12	0.225	0.138	0.020

----- 94 PARAMETER CF GRAPHIC STEP CONTAMINANT CONCENTRATION AFTER A CHROMATO

	P1	P2	P3	P4
1	1.000	1.000	1.000	1.000
2	1.000	0.222	0.020	1.000
3	0.224	0.020	0.020	1.000
4	0.028	0.020	0.020	1.000
5	0.020	0.020	0.020	1.000
6	0.016	0.372	0.805	1.000
7	0.020	0.020	0.020	1.000
8	0.020	0.020	0.020	1.000
9	0.020	0.020	0.020	1.000
10	0.003	0.003	0.003	1.000
11	1.000	0.020	1.000	1.000
12	0.532	0.835	1.000	1.000

Purification of protein

09:56:05 PAGE 7  
Model Statistics SOLVE PROTEIN1 USING MIP FROM LINE 169

10/30/95

.087 386/486 DOS

GAMS 2.25

MODEL STATISTICS

BLOCKS OF EQUATIONS	14	SINGLE EQUATIONS	860
BLOCKS OF VARIABLES	5	SINGLE VARIABLES	729



Iterations: 0

OSL step: Branch&Bound 73.54 Seconds

Status: Successful (optimal)
Iterations: 34403
Nodes : 0
Objective : 12.0000

OSL step: Primal Simplex 0.06 Seconds

Status: Successful (optimal)
Iterations: 3
Objective : 1.0000

^Purification of protein

09:56:05 PAGE 9
Solution Report SOLVE PROTEIN1 USING MIP FROM LINE 169

10/30/95

.087 386/486 DOS

GAMS 2.25

\*\*\*\* REPORT SUMMARY :
0 NONOPT
0 INFEASIBLE
0 UNBOUNDED

^Purification of protein

09:56:05 PAGE 10
Execution

10/30/95

.087 386/486 DOS

GAMS 2.25

Table with 3 columns: Variable Name, Value 1, Value 2. Rows include P1, P2, P3, P4.

Table with 3 columns: Variable Name, Value 1, Value 2. Rows include 10.P1, 10.P2, 10.P3, 10.P4.

171 VARIABLE Z.L
1 1.000

171 VARIABLE Y.L

1

10 1.000

EXECUTION TIME = 0.220 SECONDS

VERID MW2-25-087



^GAMS 2.25.087 386/486 DOS

10/30/95 23:17:00 PAGE

1  
PURIFICATION OF ECOLI\_S8

```

4
5 SET
6 I CHROMATOGRAPHY STEPS /1*22/
7 p PROTEINS /p1*p9/
8 A PROPERTY /A1*A12/
9 k ORDER /1*20/;
10
11 ALIAS (p,pp)
12 (K, KK)
13
14
15 TABLE PROP(p,A) VALUE OF PROPERTY A FOR PROTEIN p
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31

```

		A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
2	20	p1	31000	0.00	1.46	0.09	-0.62	-0.66	-1.02	-1.82	-2.33	-2.52	-3.5
1	21	p2	62500	0.00	1.46	0.09	-1.06	-0.98	-1.17	-1.71	-2.79	-3.52	-3.32
2	22	p3	40600	0.00	1.46	0.09	-0.55	-0.22	-0.22	-0.26	-0.73	-1.26	-1.82
1	23	p4	69600	0.00	1.46	0.09	-0.55	-0.22	-0.22	-0.26	-0.73	-1.26	-1.82
1	24	p5	40600	0.00	1.46	3.14	1.46	0.28	-0.47	-0.89	-1.06	-1.08	-1.04
1	25	p6	69600	0.00	1.46	3.14	1.46	0.28	-0.47	-0.89	-1.06	-1.08	-1.04
1	26	p7	41000	1.50	1.46	0.93	0.26	-0.35	-0.87	-1.31	-1.65	-1.90	-2.04
6	27	p8	32900	1.50	1.46	0.09	0.00	-1.70	-2.70	-2.90	-3.51	-3.51	-3.51
1	28	p9	35500	0.20	1.46	0.09	-0.55	-0.22	-0.22	-0.26	-0.73	-1.26	-1.82

```

;
29
30 *prop(p,A)=abs(prop(p,A));
31

```

```

32 *where
33 *P1  PRODUT
34 *P2  CONT_1
35 *P3  CONT_2
36 *P4  CONT_3
37 *P5  CONT_4
38 *P6  CONT_5
39 *p7  CONT_6
40 *P8  CONT_7
41 *P9  CONT_8;
42
43
44
45 *AND
46 *A1  MOLECULAR WEIGHT (Da)
^GAMS 2.25.087 386/486 DOS
2
PURIFICATION OF ECOLI_S8

```

10/30/95 23:17:00 PAGE

```

47 *A2  HIDROPHOBICITY
48 *A3  pH 4.0
49 *A4  pH 4.5
50 *A5  pH 5.0
51 *A6  pH 5.5
52 *A7  pH 6.0
53 *A8  pH 6.5
54 *A9  pH 7.0
55 *A10 pH 7.5
56 *A11 pH 8.0
57 *A12 pH 8.5;
58
59 scalar IP /1/;
60
61 PARAMETER
62     SIG(i)  VALUE OF PEAK WIDTH
63
64                                     /1    0.15
65                                     2     0.15
66                                     3     0.15
67                                     4     0.15
68                                     5     0.15
69                                     6     0.15
70                                     7     0.15
71                                     8     0.15
72                                     9     0.15
73                                     10    0.15
74                                     11    0.15
75                                     12    0.15
76                                     13    0.15
77                                     14    0.15
78                                     15    0.15
79                                     16    0.15
80                                     17    0.15
81                                     18    0.15
82                                     19    0.15
83                                     20    0.15
84                                     21    0.22
85                                     22    0.46/
86
87
88
89

```

$K_d(i,p)$  RETENTION TIME FOR PRODUCT  $p$  IN CHROMATOGRAPHIC STEP  $i$ ;  
 $K_d('1',p) = ((7383 * \text{abs}(\text{PROP}(p, 'A3')) / \text{PROP}(p, 'A1')))) / (1 + 15844 * \text{abs}$

(PROP(p, 'A3')/PROP(p, 'A1'))))\$(PROP(p, 'A3') LT 0) + 0\$(PROP(p, 'A3') GT 0

90 Kd('2',p)=((7383 \* abs(PROP(p, 'A4')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A4')/PROP(p, 'A1'))))\$(PROP(p, 'A4') LT 0) + 0\$(PROP(p, 'A4')  
GT 0)

91 Kd('3',p)=((7383 \* abs(PROP(p, 'A5')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A5')/PROP(p, 'A1'))))\$(PROP(p, 'A5') LT 0) + 0\$(PROP(p, 'A5')  
GT 0)

92 Kd('4',p)=((7383 \* abs(PROP(p, 'A6')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A6')/PROP(p, 'A1'))))\$(PROP(p, 'A6') LT 0) + 0\$(PROP(p, 'A6')

\*GAMS 2.25.087 386/486 DOS

10/30/95 23:17:00 PAGE

3  
PURIFICATION OF ECOLI\_S8

GT 0)

93 Kd('5',p)=((7383 \* abs(PROP(p, 'A7')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A7')/PROP(p, 'A1'))))\$(PROP(p, 'A7') LT 0) + 0\$(PROP(p, 'A7')  
GT 0)

94 Kd('6',p)=((7383 \* abs(PROP(p, 'A8')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A8')/PROP(p, 'A1'))))\$(PROP(p, 'A8') LT 0) + 0\$(PROP(p, 'A8')  
GT 0)

95 Kd('7',p)=((7383 \* abs(PROP(p, 'A9')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A9')/PROP(p, 'A1'))))\$(PROP(p, 'A9') LT 0) + 0\$(PROP(p, 'A9')  
GT 0)

96 Kd('8',p)=((7383 \* abs(PROP(p, 'A10')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A10')/PROP(p, 'A1'))))\$(PROP(p, 'A10') LT 0) + 0\$(PROP(p, 'A10')  
GT 0)

97 Kd('9',p)=((7383 \* abs(PROP(p, 'A11')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A11')/PROP(p, 'A1'))))\$(PROP(p, 'A11') LT 0) + 0\$(PROP(p, 'A11')  
GT 0)

98 Kd('10',p)=((7383 \* abs(PROP(p, 'A12')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A12')/PROP(p, 'A1'))))\$(PROP(p, 'A12') LT 0) + 0\$(PROP(p, 'A12')  
GT 0)

99 Kd('11',p)=((5972 \* ((PROP(p, 'A3')/PROP(p, 'A1')))/(1+17065 \*  
(PROP(p, 'A3')/PROP(p, 'A1'))))\$(PROP(p, 'A3') GT 0) + 0\$(PROP(p, 'A3') LT 0

100 Kd('12',p)=((5972 \* ((PROP(p, 'A4')/PROP(p, 'A1')))/(1+17065 \*  
(PROP(p, 'A4')/PROP(p, 'A1'))))\$(PROP(p, 'A4') GT 0) + 0\$(PROP(p, 'A4') LT 0

```

;
101      Kd('13',p)=((5972 * ((PROP(p,'A5')/PROP(p,'A1')))/(1+17065 *
(PROP(p,'A5')/PROP(p,'A1'))))$(PROP(p,'A5') GT 0) + 0$(PROP(p,'A5') LT 0
)
;
102      Kd('14',p)=((5972 * ((PROP(p,'A6')/PROP(p,'A1')))/(1+17065 *
(PROP(p,'A6')/PROP(p,'A1'))))$(PROP(p,'A6') GT 0) + 0$(PROP(p,'A6') LT 0
)
;
103      Kd('15',p)=((5972 * ((PROP(p,'A7')/PROP(p,'A1')))/(1+17065 *
(PROP(p,'A7')/PROP(p,'A1'))))$(PROP(p,'A7') GT 0) + 0$(PROP(p,'A7') LT 0
)
;
104      Kd('16',p)=((5972 * ((PROP(p,'A8')/PROP(p,'A1')))/(1+17065 *
(PROP(p,'A8')/PROP(p,'A1'))))$(PROP(p,'A8') GT 0) + 0$(PROP(p,'A8') LT 0
)
;
105      Kd('17',p)=((5972 * ((PROP(p,'A9')/PROP(p,'A1')))/(1+17065 *
(PROP(p,'A9')/PROP(p,'A1'))))$(PROP(p,'A9') GT 0) + 0$(PROP(p,'A9') LT 0
)
;
106      Kd('18',p)=((5972 * ((PROP(p,'A10')/PROP(p,'A1')))/(1+17065 *
(PROP(p,'A10')/PROP(p,'A1'))))$(PROP(p,'A10') GT 0) + 0$(PROP(p,'A10')
LT 0)
;
107      Kd('19',p)=((5972 * ((PROP(p,'A11')/PROP(p,'A1')))/(1+17065 *
(PROP(p,'A11')/PROP(p,'A1'))))$(PROP(p,'A11') GT 0) + 0$(PROP(p,'A11')
LT 0)
;
108      Kd('20',p)=((5972 * ((PROP(p,'A12')/PROP(p,'A1')))/(1+17065 *
(PROP(p,'A12')/PROP(p,'A1'))))$(PROP(p,'A12') GT 0) + 0$(PROP(p,'A12')
LT 0)
;
109      Kd('21',p)=PROP(p,'A2');
110      Kd('22',p)=-0.4691 * log10(PROP(p,'A1'))+2.3902;
111
112 PARAMETER
GAMS 2.25.087 386/486 DOS 10/30/95 23:17:00 PAGE
4
PURIFICATION OF ECOLI_S8

113      DF(i,p) DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i OF PROTEIN p;
114      DF(i,p) = ABS(sum(pp$(ORD(pp) EQ IP),Kd(i,pp))-Kd(i,p))
;
115
116 PARAMETER
117      CF(i,p) CONTAMINANT CONCENTRATION AFTER A CHROMATOGRAPHIC STEP;
118      CF(i,p)=1;
119      CF(i,p)$((DF(i,p) GE (SIG(i)/10)) AND (DF(i,p) LT (SIG(i)/2)))
=1.02*(SQR(SIG(i))-2*SQR(DF(i,p)))/SQR(SIG(i))
;
120      CF(i,p)$((DF(i,p) GE (SIG(i)/ 2)) AND (DF(i,p) LT (SIG(i) )))
=1.02*2*SQR(SIG(i)-DF(i,p))/SQR(SIG(i))

```

```

;
121 CF(i,p)$ (DF(i,p) GE (SIG(i)))=0.02;
122 * CF(i,p)=1.02*(SQR(SIG(i))-2*SQR(DF(i,p)))/SQR(SIG(i));
123
124 *SCALAR
125 *      U /3/
126
127 PARAMETERS
128     CON (p) VALUE THE INITIAL CONCENTRATION
129
130                                     /P1      0.62
131                                     P2      0.42
132                                     P3      0.25
133                                     P4      0.25
134                                     P5      0.09
135                                     P6      0.09
136                                     P7      2.74
137                                     P8      2.74
138                                     P9      0.25/
139 parameter
140     U VALUE MAXIME OF INITIAL CONCENTRATION;
141 U=2*sum(p$(ord(p) EQ IP),con(p));
142
INCLUDE      C:\GAMS\PROTEIN1.TXT
144 display kd,df,cf;
145
146
147
148 Variables S;
149 positive variables C(p,k), C1(i,p,k);
150 Binary Variables Y(i,k), Z(k);
151 Equations OBJ, EQ1, EQ2, EQ3, EQ4, EQ5, EQ6, EQ7, EQ8,eq9,eq4a,eq4b
152     eq7a
153     eq7b
154     ;
155
156 *Objective function
157
158 c.fx(p,'1') = con(p);
159
160 * OBJ.. S=E=SUM((i,k),Y(i,k));
161 OBJ.. S=E=SUM(k,ord(k)*z(k));
162
163 *Subject to
^GAMS 2.25.087 386/486 DOS 10/30/95 23:17:00 PAGE
5
PURIFICATION OF ECOLI_S8

164
165 EQ1(k)..SUM(i,Y(i,k))=L=1;
166
167 EQ2(i)..SUM(k,Y(i,k))=L=1;
168
169 EQ3(k)$ (ord(k) lt card(k)).. SUM(i,Y(i,k+1))=L=SUM(i,Y(i,k));
170
171 EQ4(p,k)$ (ord(k) ge 3)..C(p,k)-SUM(i,CF(i,p)*C1(i,p,k-1)) =g=
-U*(1-sum(i,y(i,k-1)))
;
172
173 EQ4b(p,k)$ (ord(k) ge 3)..C(p,k)-SUM(i,CF(i,p)*C1(i,p,k-1)) =l=
U*(1-sum(i,y(i,k-1)))
;

```

```

174
175 EQ4a(p,'2')..C(p,'2')=E=SUM(i,CF(i,p)*y(i,'1')*con(p));
176
177 EQ5(i,p,k)$ (ord(k) ge 2)..C1(i,p,k)=L=con(p)*Y(i,k-1);
178
179 EQ6(p,k)$ (ord(k) ge 2)..C(p,k)=E=SUM(i,C1(i,p,k));
180
181 EQ7(k)$ (ord(k) lt card(k)).. Z(k)=G=SUM(i,Y(i,k))-SUM(i,Y(i,k+1));
182
183 eq7a(k,kk)$ (ord(kk) le ord(k)).. SUM(i,Y(i,kk)) =g= z(k);
184
185 * eq7a(k).. SUM((i,kk)$ (ord(kk) le ord(k)),Y(i,kk)) =g= ord(k)*z(k);
186
187 eq7b(k,kk)$ (ord(kk) gt ord(k)).. SUM(i,Y(i,kk)) + z(k) =l= 1;
188
189 EQ8..SUM(k,Z(k))=E=1;
190
191 EQ9(p,k)$ (ORD(p) EQ IP).. C(p,k+1)=G=0.70*(SUM(pp,C(pp,k+1)))-U*(1-Z(k))
;
192
193 option iterlim = 1000000
194         reslim  = 1000000
195         limrow  = 0
196         solprint = off
197         optcr   = 0
198         limcol  = 0
199         optca   = 0;
200
201
202
203 MODEL PROTEIN1 /OBJ,
204         EQ1, EQ2, EQ3,
205         EQ4,
206         eq4a,eq4b
207         EQ5,
208         EQ6,
209         EQ7, eq7a, eq7b
210         EQ8
211         eq9/;
212
213
214 protein1.workspace=10;
^GAMS 2.25.087 386/486 DOS 10/30/95 23:17:00 PAGE
6
PURIFICATION OF ECOLI_S8

215
216 protein1.optfile=1
217
218
219
220 SOLVE PROTEIN1 USING MIP MINIMIZING S;
221
222 display C.1, C1.1, Z.1, Y.1;
^GAMS 2.25.087 386/486 DOS 10/30/95 23:17:00 PAGE
7
PURIFICATION OF ECOLI_S8
Include File Summary

```

SEQ	GLOBAL TYPE	PARENT	LOCAL	FILENAME
-----	-------------	--------	-------	----------

1 1 INPUT 0 0 C:\GAMS\D21TESPE.TXT  
 2 143 INCLUDE 1 143 .C:\GAMS\PROTEIN1.TXT

COMPILATION TIME = 0.110 SECONDS VERID MW2-25-087  
 GAMS 2.25.087 386/486 DOS 10/30/95 23:17:00 PAGE  
 8  
 PURIFICATION OF ECOLI\_S8  
 Execution

144 PARAMETER KD

RETENTION TIME FOR PRODUCT p IN  
 CHROMATOGRAPHIC STEP i

	P1	P2	P3	P4	P5	P6
3	0.112	0.099	0.082	0.052		
4	0.118	0.093	0.037	0.022		
5	0.160	0.107	0.037	0.022	0.072	0.045
6	0.225	0.141	0.043	0.026	0.120	0.079
7	0.253	0.193	0.103	0.066	0.136	0.091
8	0.262	0.220	0.154	0.104	0.138	0.092
9	0.262	0.213	0.194	0.137	0.135	0.089
10	0.299	0.213	0.269	0.207	0.132	0.087
11	0.156	0.100	0.133	0.092	0.133	0.092
12	0.017	0.008	0.013	0.008	0.199	0.152
13					0.133	0.092
14					0.037	0.022
22	0.283	0.140	0.228	0.119	0.228	0.119
+	P7	P8	P9			
3			0.092			
4	0.056	0.210	0.042			
5	0.117	0.263	0.042			
6	0.157	0.272	0.048			
7	0.181	0.293	0.115			
8	0.197	0.293	0.168			
9	0.205	0.293	0.209			
10	0.207	0.293	0.284			
11	0.132	0.151	0.144			
12	0.098	0.016	0.015			
13	0.034					
21	1.500	1.500	0.200			
22	0.226	0.271	0.256			

144 PARAMETER DF

DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i  
 OF PROTEIN p

	P2	P3	P4	P5	P6	P7
3	0.013	0.030	0.060	0.112	0.112	0.112
4	0.025	0.081	0.095	0.118	0.118	0.062
5	0.053	0.123	0.137	0.087	0.115	0.042
6	0.084	0.182	0.199	0.104	0.146	0.068
7	0.060	0.150	0.187	0.117	0.163	0.072
8	0.043	0.109	0.158	0.124	0.170	0.065
9	0.049	0.069	0.126	0.128	0.173	0.057
10	0.086	0.030	0.092	0.167	0.212	0.093
11	0.056	0.023	0.064	0.023	0.064	0.024
12	0.008	0.004	0.009	0.183	0.136	0.081
13				0.133	0.092	0.034

14  
 ^GAMS 2.25.087 386/486 DOS  
 9  
 PURIFICATION OF ECOLI\_S8  
 Execution

0.037 0.022  
 10/30/95 23:17:00 PAGE

144 PARAMETER DF DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i  
 OF PROTEIN p

	P2	P3	P4	P5	P6	P7
21						1.500
22	0.143	0.055	0.165	0.055	0.165	0.057
+	P8	P9				
3	0.112	0.020				
4	0.092	0.076				
5	0.104	0.118				
6	0.047	0.176				
7	0.039	0.139				
8	0.030	0.095				
9	0.030	0.053				
10	0.006	0.015				
11	0.005	0.012				
12	9.114712E-4	0.002				
21	1.500	0.200				
22	0.012	0.028				

144 PARAMETER CF CONTAMINANT CONCENTRATION AFTER A  
 CHROMATOGRAPHIC STEP

	P1	P2	P3	P4	P5	P6
1	1.000	1.000	1.000	1.000	1.000	1.000
2	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	0.940	0.691	0.130	0.130
4	1.000	0.964	0.435	0.271	0.096	0.096
5	1.000	0.764	0.067	0.014	0.355	0.113
6	1.000	0.399	0.020	0.020	0.188	0.001
7	1.000	0.691	7.434878E-8	0.020	0.099	0.020
8	1.000	0.856	0.155	0.020	0.061	0.020
9	1.000	0.799	0.591	0.053	0.045	0.020
10	1.000	0.369	0.939	0.303	0.020	0.020
11	1.000	0.734	0.973	0.652	0.973	0.652
12	1.000	1.000	1.000	1.000	0.020	0.019
13	1.000	1.000	1.000	1.000	0.026	0.302
14	1.000	1.000	1.000	1.000	0.897	0.974
15	1.000	1.000	1.000	1.000	1.000	1.000
16	1.000	1.000	1.000	1.000	1.000	1.000
17	1.000	1.000	1.000	1.000	1.000	1.000
18	1.000	1.000	1.000	1.000	1.000	1.000
19	1.000	1.000	1.000	1.000	1.000	1.000
20	1.000	1.000	1.000	1.000	1.000	1.000
21	1.000	1.000	1.000	1.000	1.000	1.000
22	1.000	0.823	0.991	0.758	0.991	0.758

GAMS 2.25.087 386/486 DOS  
 10  
 PURIFICATION OF ECOLI\_S8  
 Execution

10/30/95 23:17:00 PAGE



144 PARAMETER CF

CONTAMINANT CONCENTRATION AFTER A CHROMATOGRAPHIC STEP

+	P7	P8	P9
1	1.000	1.000	1.000
2	1.000	1.000	1.000
3	0.130	0.130	0.983
4	0.671	0.303	0.498
5	0.857	0.194	0.093
6	0.601	0.820	0.020
7	0.552	0.879	0.011
8	0.637	0.936	0.278
9	0.726	0.936	0.761
10	0.298	1.000	1.000
11	0.969	1.000	1.000
12	0.430	1.000	1.000
13	0.914	1.000	1.000
14	1.000	1.000	1.000
15	1.000	1.000	1.000
16	1.000	1.000	1.000
17	1.000	1.000	1.000
18	1.000	1.000	1.000
19	1.000	1.000	1.000
20	1.000	1.000	1.000
21	0.020	0.020	0.017
22	0.989	1.000	1.000

^GAMS 2.25.087 386/486 DOS

10/30/95 23:17:00 PAGE

11  
PURIFICATION OF ECOLI\_S8  
Model Statistics SOLVE PROTEIN1 USING MIP FROM LINE 220

MODEL STATISTICS

BLOCKS OF EQUATIONS	14	SINGLE EQUATIONS	4768
BLOCKS OF VARIABLES	5	SINGLE VARIABLES	4394
NON ZERO ELEMENTS	38247	DISCRETE VARIABLES	460

GENERATION TIME = 4.010 SECONDS

EXECUTION TIME = 4.230 SECONDS

^GAMS 2.25.087 386/486 DOS

VERID MW2-25-087  
10/30/95 23:17:00 PAGE

12  
PURIFICATION OF ECOLI\_S8  
Solution Report SOLVE PROTEIN1 USING MIP FROM LINE 220

S O L V E S U M M A R Y

MODEL	PROTEIN1	OBJECTIVE	S
TYPE	MIP	DIRECTION	MINIMIZE
SOLVER	OSL	FROM LINE	220.

\*\*\*\* SOLVER STATUS 1 NORMAL COMPLETION  
\*\*\*\* MODEL STATUS 1 OPTIMAL  
\*\*\*\* OBJECTIVE VALUE 3.0000

RESOURCE USAGE, LIMIT 80.852 1000000.000

ITERATION COUNT, LIMIT 1015 1000000  
 Could not find option file C:\GAMS\OSL.OPT  
 Using defaults instead.

Work space requested by user -- 10.00 Mb  
 Work space requested by solver -- 9.28 Mb

OSL step: Reading data 2.58 Seconds

OSL step: Scale 2.19 Seconds

Range of matrix coefficients:  
 before scaling : 0.0000 - 20.0000  
 after scaling : 0.0000 - 1.0000

OSL step: Presolve 1.70 Seconds

Size reduction:  
 rows : 4768 (old) 4766 (new)  
 columns : 4394 (old) 4394 (new)  
 nonzeros: 38247 (old) 38225 (new)

OSL step: Crash 1.43 Seconds

Crash option : 1

OSL step: Primal Simplex 8.23 Seconds

Status: Successful (optimal)  
 Iterations: 161  
 Objective : 1.3486

OSL step: Branch&Bound 56.19 Seconds

Status: Successful (optimal)  
 Iterations: 854  
 Nodes : 14  
 Objective : 3.0000

OSL step: Postsolve 1.77 Seconds

OSL step: Primal Simplex 0.76 Seconds

Status: Successful (optimal)  
 Iterations: 0  
 Objective : 3.0000

^GAMS 2.25.087 386/486 DOS

10/30/95 23:17:00 PAGE

13  
 PURIFICATION OF ECOLI\_S8  
 Solution Report SOLVE PROTEIN1 USING MIP FROM LINE 220

\*\*\*\* REPORT SUMMARY : 0 NONOPT  
 0 INFEASIBLE  
 0 UNBOUNDED

^GAMS 2.25.087 386/486 DOS

10/30/95 23:17:00 PAGE

14  
 PURIFICATION OF ECOLI\_S8  
 Execution

```

----- 222 VARIABLE C.L
          1          2          3          4
P1      0.620      0.620      0.620      0.620
P2      0.420      0.420      0.420      0.168
P3      0.250      0.250      0.250      0.005
P4      0.250      0.250      0.250      0.005
P5      0.090      0.090      0.090      0.017
P6      0.090      0.090      0.090 1.266353E-4
P7      2.740      0.055      0.001 6.590838E-4
P8      2.740      0.055      0.001 8.985890E-4
P9      0.250      0.004 7.106072E-5 1.421214E-6

```

```

----- 222 VARIABLE C1.L
          2          3          4
6 .P1          0.620
6 .P2          0.420
6 .P3          0.250
6 .P4          0.250
6 .P5          0.090
6 .P6          0.090
6 .P7          0.001
6 .P8          0.001
6 .P9          7.106072E-5
7 .P1          0.620
7 .P2          0.168
7 .P3          0.005
7 .P4          0.005
7 .P5          0.017
7 .P6          1.266353E-4
7 .P7          6.590838E-4
7 .P8          8.985890E-4
7 .P9          1.421214E-6
21.P1         0.620
21.P2         0.420
21.P3         0.250
21.P4         0.250
21.P5         0.090
21.P6         0.090
21.P7         0.055
21.P8         0.055
21.P9         0.004

```

GAMS 2.25.087 386/486 DOS

10/30/95 23:17:00 PAGE

15

PURIFICATION OF ECOLI\_S8  
E x e c u t i o n

```

----- 222 VARIABLE Z.L

```

3 1.000

```

----- 222 VARIABLE Y.L

```

```

          1          2          3
6          1.000

```

7  
21 1.000

1.000

EXECUTION TIME

=

0.390 SECONDS

VERID MW2-25-087

1  
PURIFICATION OF ECOLI\_S8

```

4
5 SET
6 I CRHOMATOGRAPHY STEPS /1*22/
7 p PROTEINS /p1*p9/
8 A PROPERTY /A1*A12/
9 k ORDER /1*20/;
10
11 ALIAS (p,pp)
12 (K, KK)
13
14
15 TABLE PROP(p,A) VALUE OF PROPERTY A FOR PROTEIN p
16
17
18
19          A1      A2      A3      A4      A5      A6      A7      A8      A9
                                     A10     A11     A1
2 20 p1      31000    0.00    1.46    0.09    -0.62    -0.66    -1.02    -1.82    -2.33
                                     -2.52    -2.52    -3.5
1 21 p2      62500    0.00    1.46    0.09    -1.06    -0.98    -1.17    -1.71    -2.79
                                     -3.52    -3.32    -3.3
2 22 p3      40600    0.00    1.46    0.09    -0.55    -0.22    -0.22    -0.26    -0.73
                                     -1.26    -1.82    -3.5
1 23 p4      69600    0.00    1.46    0.09    -0.55    -0.22    -0.22    -0.26    -0.73
                                     -1.26    -1.82    -3.5
1 24 p5      40600    0.00    1.46    3.14    1.46    0.28    -0.47    -0.89    -1.06
                                     -1.08    -1.04    -1.0
1 25 p6      69600    0.00    1.46    3.14    1.46    0.28    -0.47    -0.89    -1.06
                                     -1.08    -1.04    -1.0
1 26 p7      41000    1.50    1.46    0.93    0.26    -0.35    -0.87    -1.31    -1.65
                                     -1.90    -2.04    -2.0
6 27 p8      32900    1.50    1.46    0.09    0.00    -1.70    -2.70    -2.90    -3.51
                                     -3.51    -3.51    -3.5
1 28 p9      35500    0.20    1.46    0.09    -0.55    -0.22    -0.22    -0.26    -0.73
                                     -1.26    -1.82    -3.51
;
29
30 *prop(p,A)=abs(prop(p,A));
31

```

```

32 *where
33 *P1  PRODUT
34 *P2  CONT_1
35 *P3  CONT_2
36 *P4  CONT_3
37 *P5  CONT_4
38 *P6  CONT_5
39 *p7  CONT_6
40 *P8  CONT_7
41 *P9  CONT_8;
42
43
44
45 *AND
46 *A1  MOLECULAR WEIGHT (Da)
^GAMS 2.25.087 386/486 DOS
2

```

10/30/95 23:08:59 PAGE

PURIFICATION OF ECOLI\_S8

```

47 *A2  HIDROPHOBICITY
48 *A3  pH 4.0
49 *A4  pH 4.5
50 *A5  pH 5.0
51 *A6  pH 5.5
52 *A7  pH 6.0
53 *A8  pH 6.5
54 *A9  pH 7.0
55 *A10 pH 7.5
56 *A11 pH 8.0
57 *A12 pH 8.5;
58
59 scalar IP /1/;
60
61 PARAMETER
62     SIG(i)  VALUE OF PEAK WIDTH
63
64                                     /1    0.15
65                                     2     0.15
66                                     3     0.15
67                                     4     0.15
68                                     5     0.15
69                                     6     0.15
70                                     7     0.15
71                                     8     0.15
72                                     9     0.15
73                                     10    0.15
74                                     11    0.15
75                                     12    0.15
76                                     13    0.15
77                                     14    0.15
78                                     15    0.15
79                                     16    0.15
80                                     17    0.15
81                                     18    0.15
82                                     19    0.15
83                                     20    0.15
84                                     21    0.22
85                                     22    0.46/
86
87

```

```

88 Kd(i,p) RETENTION TIME FOR PRODUCT p IN CHROMATOGRAPHIC STEP i;
89 Kd('1',p)=((7383 * abs(PROP(p, 'A3')/PROP(p, 'A1')))/(1+15844 *abs

```

(PROP(p, 'A3')/PROP(p, 'A1'))))\$(PROP(p, 'A3') LT 0) + 0\$(PROP(p, 'A3') GT 0

90 Kd('2',p)=((7383 \* abs(PROP(p, 'A4')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A4')/PROP(p, 'A1'))))\$(PROP(p, 'A4') LT 0) + 0\$(PROP(p, 'A4')  
GT 0)

91 Kd('3',p)=((7383 \* abs(PROP(p, 'A5')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A5')/PROP(p, 'A1'))))\$(PROP(p, 'A5') LT 0) + 0\$(PROP(p, 'A5')  
GT 0)

92 Kd('4',p)=((7383 \* abs(PROP(p, 'A6')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A6')/PROP(p, 'A1'))))\$(PROP(p, 'A6') LT 0) + 0\$(PROP(p, 'A6')

^GAMS 2.25.087 386/486 DOS

10/30/95 23:08:59 PAGE

3  
PURIFICATION OF ECOLI\_S8

GT 0)

93 Kd('5',p)=((7383 \* abs(PROP(p, 'A7')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A7')/PROP(p, 'A1'))))\$(PROP(p, 'A7') LT 0) + 0\$(PROP(p, 'A7')  
GT 0)

94 Kd('6',p)=((7383 \* abs(PROP(p, 'A8')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A8')/PROP(p, 'A1'))))\$(PROP(p, 'A8') LT 0) + 0\$(PROP(p, 'A8')  
GT 0)

95 Kd('7',p)=((7383 \* abs(PROP(p, 'A9')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A9')/PROP(p, 'A1'))))\$(PROP(p, 'A9') LT 0) + 0\$(PROP(p, 'A9')  
GT 0)

96 Kd('8',p)=((7383 \* abs(PROP(p, 'A10')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A10')/PROP(p, 'A1'))))\$(PROP(p, 'A10') LT 0) + 0\$(PROP(p, 'A10')  
GT 0)

97 Kd('9',p)=((7383 \* abs(PROP(p, 'A11')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A11')/PROP(p, 'A1'))))\$(PROP(p, 'A11') LT 0) + 0\$(PROP(p, 'A11')  
GT 0)

98 Kd('10',p)=((7383 \* abs(PROP(p, 'A12')/PROP(p, 'A1')))/(1+15844 \*  
abs(PROP(p, 'A12')/PROP(p, 'A1'))))\$(PROP(p, 'A12') LT 0) + 0\$(PROP(p, 'A12')  
GT 0)

99 Kd('11',p)=((5972 \* ((PROP(p, 'A3')/PROP(p, 'A1')))/(1+17065 \*  
(PROP(p, 'A3')/PROP(p, 'A1'))))\$(PROP(p, 'A3') GT 0) + 0\$(PROP(p, 'A3') LT 0

100 Kd('12',p)=((5972 \* ((PROP(p, 'A4')/PROP(p, 'A1')))/(1+17065 \*  
(PROP(p, 'A4')/PROP(p, 'A1'))))\$(PROP(p, 'A4') GT 0) + 0\$(PROP(p, 'A4') LT 0

```

;
101      Kd('13',p)=((5972 * ((PROP(p,'A5')/PROP(p,'A1')))/(1+17065 *
)      (PROP(p,'A5')/PROP(p,'A1'))))$(PROP(p,'A5') GT 0) + 0$(PROP(p,'A5') LT 0
;
102      Kd('14',p)=((5972 * ((PROP(p,'A6')/PROP(p,'A1')))/(1+17065 *
)      (PROP(p,'A6')/PROP(p,'A1'))))$(PROP(p,'A6') GT 0) + 0$(PROP(p,'A6') LT 0
;
103      Kd('15',p)=((5972 * ((PROP(p,'A7')/PROP(p,'A1')))/(1+17065 *
)      (PROP(p,'A7')/PROP(p,'A1'))))$(PROP(p,'A7') GT 0) + 0$(PROP(p,'A7') LT 0
;
104      Kd('16',p)=((5972 * ((PROP(p,'A8')/PROP(p,'A1')))/(1+17065 *
)      (PROP(p,'A8')/PROP(p,'A1'))))$(PROP(p,'A8') GT 0) + 0$(PROP(p,'A8') LT 0
;
105      Kd('17',p)=((5972 * ((PROP(p,'A9')/PROP(p,'A1')))/(1+17065 *
)      (PROP(p,'A9')/PROP(p,'A1'))))$(PROP(p,'A9') GT 0) + 0$(PROP(p,'A9') LT 0
;
106      Kd('18',p)=((5972 * ((PROP(p,'A10')/PROP(p,'A1')))/(1+17065 *
)      (PROP(p,'A10')/PROP(p,'A1'))))$(PROP(p,'A10') GT 0) + 0$(PROP(p,'A10')
)      LT 0)
;
107      Kd('19',p)=((5972 * ((PROP(p,'A11')/PROP(p,'A1')))/(1+17065 *
)      (PROP(p,'A11')/PROP(p,'A1'))))$(PROP(p,'A11') GT 0) + 0$(PROP(p,'A11')
)      LT 0)
;
108      Kd('20',p)=((5972 * ((PROP(p,'A12')/PROP(p,'A1')))/(1+17065 *
)      (PROP(p,'A12')/PROP(p,'A1'))))$(PROP(p,'A12') GT 0) + 0$(PROP(p,'A12')
)      LT 0)
;
109      Kd('21',p)=PROP(p,'A2');
110      Kd('22',p)=-0.4691 * log10(PROP(p,'A1')+2.3902);
111
112  PARAMETER
^GAMS 2.25.087 386/486 DOS 10/30/95 23:08:59 PAGE
4
PURIFICATION OF ECOLI_S8

113      DF(i,p) DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i OF PROTEIN p;
114      DF(i,p) = ABS(sum(pp$(ORD(pp) EQ IP),Kd(i,pp))-Kd(i,p))
;
115
116  PARAMETER
117      CF(i,p) CONTAMINANT CONCENTRATION AFTER A CHROMATOGRAPHIC STEP;
118      CF(i,p)=1;
119      CF(i,p)$((DF(i,p) GE (SIG(i)/10)) AND (DF(i,p) LT (SIG(i)/2)))
)      =1.02*(SQR(SIG(i))-2*SQR(DF(i,p)))/SQR(SIG(i))
;
120      CF(i,p)$((DF(i,p) GE (SIG(i)/ 2)) AND (DF(i,p) LT (SIG(i) )))
)      =1.02*2*SQR(SIG(i)-DF(i,p))/SQR(SIG(i))

```



```

;
121 CF(i,p)$ (DF(i,p) GE (SIG(i)))=0.02;
122 * CF(i,p)=1.02*(SQR(SIG(i))-2*SQR(DF(i,p)))/SQR(SIG(i));
123
124 *SCALAR
125 *      U /3/
126
127 PARAMETERS
128     CON (p) VALUE THE INITIAL CONCENTRATION
129
130                                     /P1    0.62
131                                     P2    0.42
132                                     P3    0.25
133                                     P4    0.25
134                                     P5    0.09
135                                     P6    0.09
136                                     P7    2.74
137                                     P8    2.74
138                                     P9    0.25/
139 parameter
140     U VALUE MAXIME OF INITIAL CONCENTRATION;
141 U=2*sum(p$(ord(p) EQ IP),con(p));
142
INCLUDE      C:\GAMS\PROTEIN1.TXT
144 display kd,df,cf;
145
146
147
148 Variables S;
149 positive variables C(p,k), C1(i,p,k);
150 Binary Variables Y(i,k), Z(k);
151 Equations OBJ, EQ1, EQ2, EQ3, EQ4, EQ5, EQ6, EQ7, EQ8,eq9,eq4a,eq4b
152     eq7a
153     eq7b
154     ;
155
156 *Objective function
157
158 c.fx(p,'1') = con(p);
159
160 * OBJ.. S=E=SUM((i,k),Y(i,k));
161 OBJ.. S=E=SUM(k,ord(k)*z(k));
162
163 *Subject to
^GAMS 2.25.087 386/486 DOS 10/30/95 23:08:59 PAGE
5
PURIFICATION OF ECOLI_S8

164
165 EQ1(k)..SUM(i,Y(i,k))=L=1;
166
167 EQ2(i)..SUM(k,Y(i,k))=L=1;
168
169 EQ3(k)$ (ord(k) lt card(k)).. SUM(i,Y(i,k+1))=L=SUM(i,Y(i,k));
170
171 EQ4(p,k)$ (ord(k) ge 3)..C(p,k)-SUM(i,CF(i,p)*C1(i,p,k-1)) =g=
                                     -U*(1-sum(i,y(i,k-1)))
;
172
173 EQ4b(p,k)$ (ord(k) ge 3)..C(p,k)-SUM(i,CF(i,p)*C1(i,p,k-1)) =l=
                                     U*(1-sum(i,y(i,k-1)))
;

```

```

174
175 EQ4a(p, '2')..C(p, '2')=E=SUM(i,CF(i,p)*y(i, '1')*con(p));
176
177 EQ5(i,p,k)$ (ord(k) ge 2)..C1(i,p,k)=L=con(p)*Y(i,k-1);
178
179 EQ6(p,k)$ (ord(k) ge 2)..C(p,k)=E=SUM(i,C1(i,p,k));
180
181 EQ7(k)$ (ord(k) lt card(k)).. Z(k)=G=SUM(i,Y(i,k))-SUM(i,Y(i,k+1));
182
183 eq7a(k,kk)$ (ord(kk) le ord(k)).. SUM(i,Y(i,kk)) =g= z(k);
184
185 * eq7a(k).. SUM((i,kk)$ (ord(kk) le ord(k)),Y(i,kk)) =g= ord(k)*z(k);
186
187 eq7b(k,kk)$ (ord(kk) gt ord(k)).. SUM(i,Y(i,kk)) + z(k) =l= 1;
188
189 EQ8..SUM(k,Z(k))=E=1;
190
191 EQ9(p,k)$ (ORD(p) EQ IP).. C(p,k+1)=G=0.80*(SUM(pp,C(pp,k+1)))-U*(1-Z(k))
;
192
193 option iterlim = 1000000
194         reslim  = 1000000
195         limrow  = 0
196         solprint = off
197         optcr   = 0
198         limcol  = 0
199         optca   = 0;
200
201
202
203 MODEL PROTEIN1 /OBJ,
204         EQ1, EQ2, EQ3,
205         EQ4,
206         eq4a,eq4b
207         EQ5,
208         EQ6,
209         EQ7, eq7a, eq7b
210         EQ8
211         eq9/;
212
213
214 protein1.workspace=10;
^GAMS 2.25.087 386/486 DOS 10/30/95 23:08:59 PAGE
6
PURIFICATION OF ECOLI_S8

```

```

215
216 protein1.optfile=1
217
218
219
220 SOLVE PROTEIN1 USING MIP MINIMIZING S;
221
222 display C.l, C1.l, Z.l, Y.l;
^GAMS 2.25.087 386/486 DOS 10/30/95 23:08:59 PAGE
7
PURIFICATION OF ECOLI_S8
Include File Summary

```

```

SEQ    GLOBAL TYPE    PARENT    LOCAL    FILENAME

```

```

1          1 INPUT          0          0 C:\GAMS\DATOS2TE.TXT
2          143 INCLUDE      1          143 .C:\GAMS\PROTEIN1.TXT
    
```

```

COMPILATION TIME      =      0.110 SECONDS      VERID MW2-25-087
^GAMS 2.25.087 386/486 DOS      10/30/95 23:08:59 PAGE
8
    
```

PURIFICATION OF ECOLI\_S8  
Execution

---- 144 PARAMETER KD RETENTION TIME FOR PRODUCT p IN CHROMATOGRAPHIC STEP i

	P1	P2	P3	P4	P5	P6
3	0.112	0.099	0.082	0.052		
4	0.118	0.093	0.037	0.022		
5	0.160	0.107	0.037	0.022	0.072	0.045
6	0.225	0.141	0.043	0.026	0.120	0.079
7	0.253	0.193	0.103	0.066	0.136	0.091
8	0.262	0.220	0.154	0.104	0.138	0.092
9	0.262	0.213	0.194	0.137	0.135	0.089
10	0.299	0.213	0.269	0.207	0.132	0.087
11	0.156	0.100	0.133	0.092	0.133	0.092
12	0.017	0.008	0.013	0.008	0.199	0.152
13					0.133	0.092
14					0.037	0.022
22	0.283	0.140	0.228	0.119	0.228	0.119

+ P7 P8 P9

	P7	P8	P9
3			0.092
4	0.056	0.210	0.042
5	0.117	0.263	0.042
6	0.157	0.272	0.048
7	0.181	0.293	0.115
8	0.197	0.293	0.168
9	0.205	0.293	0.209
10	0.207	0.293	0.284
11	0.132	0.151	0.144
12	0.098	0.016	0.015
13	0.034		
21	1.500	1.500	0.200
22	0.226	0.271	0.256

---- 144 PARAMETER DF DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i OF PROTEIN p

	P2	P3	P4	P5	P6	P7
3	0.013	0.030	0.060	0.112	0.112	0.112
4	0.025	0.081	0.095	0.118	0.118	0.062
5	0.053	0.123	0.137	0.087	0.115	0.042
6	0.084	0.182	0.199	0.104	0.146	0.068
7	0.060	0.150	0.187	0.117	0.163	0.072
8	0.043	0.109	0.158	0.124	0.170	0.065
9	0.049	0.069	0.126	0.128	0.173	0.057
10	0.086	0.030	0.092	0.167	0.212	0.093
11	0.056	0.023	0.064	0.023	0.064	0.024
12	0.008	0.004	0.009	0.183	0.136	0.081
13				0.133	0.092	0.034

14  
 ^GAMS 2.25.087 386/486 DOS  
 9  
 PURIFICATION OF ECOLI\_S8  
 E x e c u t i o n

0.037 0.022  
 10/30/95 23:08:59 PAGE

	144 PARAMETER DF		DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i			
	P2	P3	P4	P5	P6	P7
21						1.500
22	0.143	0.055	0.165	0.055	0.165	0.057
+	P8	P9				
3	0.112	0.020				
4	0.092	0.076				
5	0.104	0.118				
6	0.047	0.176				
7	0.039	0.139				
8	0.030	0.095				
9	0.030	0.053				
10	0.006	0.015				
11	0.005	0.012				
12	9.114712E-4	0.002				
21	1.500	0.200				
22	0.012	0.028				

	144 PARAMETER CF						CONTAMINANT CONCENTRATION AFTER A CHROMATOGRAPHIC STEP
	P1	P2	P3	P4	P5	P6	
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	1.000	1.000	0.940	0.691	0.130	0.130	0.130
4	1.000	0.964	0.435	0.271	0.096	0.096	0.096
5	1.000	0.764	0.067	0.014	0.355	0.113	0.113
6	1.000	0.399	0.020	0.020	0.188	0.001	0.001
7	1.000	0.691	7.434878E-8	0.020	0.099	0.020	0.020
8	1.000	0.856	0.155	0.020	0.061	0.020	0.020
9	1.000	0.799	0.591	0.053	0.045	0.020	0.020
10	1.000	0.369	0.939	0.303	0.020	0.020	0.020
11	1.000	0.734	0.973	0.652	0.973	0.652	0.652
12	1.000	1.000	1.000	1.000	0.020	0.019	0.019
13	1.000	1.000	1.000	1.000	0.026	0.302	0.302
14	1.000	1.000	1.000	1.000	0.897	0.974	0.974
15	1.000	1.000	1.000	1.000	1.000	1.000	1.000
16	1.000	1.000	1.000	1.000	1.000	1.000	1.000
17	1.000	1.000	1.000	1.000	1.000	1.000	1.000
18	1.000	1.000	1.000	1.000	1.000	1.000	1.000
19	1.000	1.000	1.000	1.000	1.000	1.000	1.000
20	1.000	1.000	1.000	1.000	1.000	1.000	1.000
21	1.000	1.000	1.000	1.000	1.000	1.000	1.000
22	1.000	0.823	0.991	0.758	0.991	0.758	0.758

10  
 ^GAMS 2.25.087 386/486 DOS  
 PURIFICATION OF ECOLI\_S8  
 E x e c u t i o n

10/30/95 23:08:59 PAGE

144 PARAMETER CF

CONTAMINANT CONCENTRATION AFTER A CHROMATOGRAPHIC STEP

+	P7	P8	P9
1	1.000	1.000	1.000
2	1.000	1.000	1.000
3	0.130	0.130	0.983
4	0.671	0.303	0.498
5	0.857	0.194	0.093
6	0.601	0.820	0.020
7	0.552	0.879	0.011
8	0.637	0.936	0.278
9	0.726	0.936	0.761
10	0.298	1.000	1.000
11	0.969	1.000	1.000
12	0.430	1.000	1.000
13	0.914	1.000	1.000
14	1.000	1.000	1.000
15	1.000	1.000	1.000
16	1.000	1.000	1.000
17	1.000	1.000	1.000
18	1.000	1.000	1.000
19	1.000	1.000	1.000
20	1.000	1.000	1.000
21	0.020	0.020	0.017
22	0.989	1.000	1.000

^GAMS 2.25.087 386/486 DOS

10/30/95 23:08:59 PAGE

11  
PURIFICATION OF ECOLI\_S8  
Model Statistics SOLVE PROTEIN1 USING MIP FROM LINE 220

MODEL STATISTICS

BLOCKS OF EQUATIONS	14	SINGLE EQUATIONS	4768
BLOCKS OF VARIABLES	5	SINGLE VARIABLES	4394
NON ZERO ELEMENTS	38247	DISCRETE VARIABLES	460

GENERATION TIME = 3.900 SECONDS

EXECUTION TIME = 4.070 SECONDS

VERID MW2-25-087

^GAMS 2.25.087 386/486 DOS

10/30/95 23:08:59 PAGE

12  
PURIFICATION OF ECOLI\_S8  
Solution Report SOLVE PROTEIN1 USING MIP FROM LINE 220

S O L V E S U M M A R Y

MODEL	PROTEIN1	OBJECTIVE	S
TYPE	MIP	DIRECTION	MINIMIZE
SOLVER	OSL	FROM LINE	220

\*\*\*\* SOLVER STATUS 1 NORMAL COMPLETION  
\*\*\*\* MODEL STATUS 1 OPTIMAL  
\*\*\*\* OBJECTIVE VALUE 4.0000

RESOURCE USAGE, LIMIT 178.344 1000000.000



## ----- 222 VARIABLE C.L

	1	2	3	4	5
P1	0.620	0.620	0.620	0.620	0.620
P2	0.420	0.420	0.420	0.155	0.107
P3	0.250	0.250	0.250	0.235	
P4	0.250	0.250	0.250	0.076	0.002
P5	0.090	0.090	0.090	0.002	1.784426E-4
P6	0.090	0.090	0.090	0.002	3.600000E-5
P7	2.740	0.055	0.001	3.266798E-4	1.802938E-4
P8	2.740	0.055	0.001	0.001	9.629643E-4
P9	0.250	0.004	7.106072E-5	7.106072E-5	

## ----- 222 VARIABLE C1.L

	2	3	4	5
7 .P1			0.620	
7 .P2			0.155	
7 .P3			0.235	
7 .P4			0.076	
7 .P5			0.002	
7 .P6			0.002	
7 .P7		3.266798E-4		
7 .P8		0.001		
7 .P9		7.106072E-5		
10.P1		0.620		
10.P2		0.420		
10.P3		0.250		
10.P4		0.250		
10.P5		0.090		
10.P6		0.090		
10.P7		0.001		
10.P8		0.001		
10.P9	7.106072E-5			
12.P1			0.620	
12.P2			0.107	
12.P4			0.002	
12.P5			1.784426E-4	
12.P6			3.600000E-5	
12.P7			1.802938E-4	
12.P8			9.629643E-4	
21.P1	0.620			
21.P2	0.420			
21.P3	0.250			
21.P4	0.250			
21.P5	0.090			
21.P6	0.090			
21.P7	0.055			
21.P8	0.055			
21.P9	0.004			

^GAMS 2.25.087 386/486 DOS

10/30/95 23:08:59 PAGE

15

PURIFICATION OF ECOLI\_S8

Execution

## ----- 222 VARIABLE Z.L

4 1.000

```
----- 222 VARIABLE Y.L
          1          2          3          4
7
10
12
21 1.000          1.000          1.000
```

EXECUTION TIME = 0.440 SECONDS VERID MW2-25-087



## ^Selection of Multistep Protein Purification Processes

AGE 1

10/30/95 17:27:32 P

6 DOS

GAMS 2.25.087 386/48

```

4
5 SET
6 I CRHOMATOGRAPHY STEPS /1*22/
7 p PROTEINS /p1*p14/
8 A PROPERTY /A1*A12/
9 k ORDER /1*12;/
10
11 ALIAS (p,pp)
12 (k,kk)
13
14
15 TABLE PROP(p,A) VALUE OF PROPERTY A FOR PROTEIN p
16
17
18
19
20 A1 A2 A3 A4 A5 A6 A7 A8 A9 A
07 p1 22000 0.93 4.77 3.81 2.42 1.50 1.03 0.67 0.12 0.
21 p2 18370 0.71 1.94 0.25 -0.80 -1.41 -1.76 -1.97 -2.15 -2.
33 p3 85570 0.48 2.35 0.29 -1.17 -2.17 -2.83 -3.24 -3.50 -3.
63 p4 53660 0.76 1.83 0.67 0.04 -0.30 -0.49 -0.65 -0.85 -1.
90 p5 120000 1.50 3.29 1.38 -0.03 -0.69 -1.07 -1.34 -1.73 -2.
30 p6 203000 0.36 4.08 1.83 0.04 -1.17 -1.92 -2.46 -3.07 -3.
90 p7 69380 0.36 5.22 3.17 1.02 -0.72 -1.90 -2.60 -3.05 -3.
46 p8 48320 0.48 3.96 3.16 1.12 -0.58 -1.36 -1.34 -1.00 -0.
95 p9 93380 0.93 10.90 5.81 2.78 0.77 -0.81 -2.18 -3.32 -4.
12 p10 69380 1.09 0.55 0.26 0.10 -0.03 -0.12 -0.21 -0.
28 p11 114450 0.63 10.40 5.94 3.15 1.51 0.56 -0.05 -0.53 -0.
99 p12 198000 0.06 0.33 0.03 0.05 0.05 0.05 0.05 0.05 -0.
69 p13 30400 5.17 4.22 3.20 2.25 1.46 0.87 0.50 0.
30 p14 94670 11.70 7.94 5.39 3.73 2.66 1.97 1.50 1.
13 0.80 0.51;
34
35 *prop(p,A)=abs(prop(p,A));
36
37 *where
38 *P1 PRODUT
39 *P2 CONT_1
40 *P3 CONT_2
41 *P4 CONT_3
42 *P5 CONT_4
43 *P6 CONT_5
44 *P7 CONT_6
45 *P8 CONT_7
46 *P9 CONT_8

```

47 \*P10 CONT\_9  
 48 \*P11 CONT\_10  
 49 \*P12 CONT\_11  
 50 \*P13 CONT\_12  
 51 \*P14 CONT\_13;

52  
 53

54 \*AND

55 \*A1 MOLECULAR WEIGHT (Da)

56 \*A2 HIDROPHOBICITY

57 \*A3 pH 4.0

\*Selection of Multistep Protein Purification Processes

AGE 2

10/30/95 17:27:32 P

6 DOS

GAMS 2.25.087 386/48

58 \*A4 pH 4.5  
 59 \*A5 pH 5.0  
 60 \*A6 pH 5.5  
 61 \*A7 pH 6.0  
 62 \*A8 pH 6.5  
 63 \*A9 pH 7.0  
 64 \*A10 pH 7.5  
 65 \*A11 pH 8.0  
 66 \*A12 pH 8.5;

67

68 scalar IP /1/

69 PARAMETER

70 SIG(i) VALUE OF PEAK WIDTH

71

72 /1 0.15

73 2 0.15

74 3 0.15

75 4 0.15

76 5 0.15

77 6 0.15

78 7 0.15

79 8 0.15

80 9 0.15

81 10 0.15

82 11 0.15

83 12 0.15

84 13 0.15

85 14 0.15

86 15 0.15

87 16 0.15

88 17 0.15

89 18 0.15

90 19 0.15

91 20 0.15

92 21 0.22

93 22 0.46/

94

95

96 Kd(i,p) RETENTION TIME FOR PRODUCT p IN CHROMATOGRAPHIC STEP i;  
 97 Kd('1',p)=((7383 \* abs(PROP(p,'A3')/PROP(p,'A1')))/(1+15844 \* abs(PR

OP(p,'A3')/PROP(p,'A1'))))\$(PROP(p,'A3') LT 0) + 0\$(PROP(p,'A3') GT 0);  
 98 Kd('2',p)=((7383 \* abs(PROP(p,'A4')/PROP(p,'A1')))/(1+15844 \* abs(PR

OP(p,'A4')/PROP(p,'A1'))))\$(PROP(p,'A4') LT 0) + 0\$(PROP(p,'A4') GT 0);  
 99 Kd('3',p)=((7383 \* abs(PROP(p,'A5')/PROP(p,'A1')))/(1+15844 \* abs(PR

OP(p,'A5')/PROP(p,'A1'))))\$(PROP(p,'A5') LT 0) + 0\$(PROP(p,'A5') GT 0);  
 100 Kd('4',p)=((7383 \* abs(PROP(p,'A6')/PROP(p,'A1')))/(1+15844 \* abs(PR

OP(p,'A6')/PROP(p,'A1'))))\$(PROP(p,'A6') LT 0) + 0\$(PROP(p,'A6') GT 0);

```

101      Kd('5',p)=((7383 * abs(PROP(p,'A7')/PROP(p,'A1')))/(1+15844 * abs(PR
OP(p,'A7')/PROP(p,'A1'))))$(PROP(p,'A7') LT 0) + 0$(PROP(p,'A7') GT 0);
102      Kd('6',p)=((7383 * abs(PROP(p,'A8')/PROP(p,'A1')))/(1+15844 * abs(PR
OP(p,'A8')/PROP(p,'A1'))))$(PROP(p,'A8') LT 0) + 0$(PROP(p,'A8') GT 0);
103      Kd('7',p)=((7383 * abs(PROP(p,'A9')/PROP(p,'A1')))/(1+15844 * abs(PR
OP(p,'A9')/PROP(p,'A1'))))$(PROP(p,'A9') LT 0) + 0$(PROP(p,'A9') GT 0);
104      Kd('8',p)=((7383 * abs(PROP(p,'A10')/PROP(p,'A1')))/(1+15844 * abs(P
ROP(p,'A10')/PROP(p,'A1'))))$(PROP(p,'A10') LT 0) + 0$(PROP(p,'A10') GT 0);
105      Kd('9',p)=((7383 * abs(PROP(p,'A11')/PROP(p,'A1')))/(1+15844 * abs(P
ROP(p,'A11')/PROP(p,'A1'))))$(PROP(p,'A11') LT 0) + 0$(PROP(p,'A11') GT 0);
106      Kd('10',p)=((7383 * abs(PROP(p,'A12')/PROP(p,'A1')))/(1+15844 * abs(
PROP(p,'A12')/PROP(p,'A1'))))$(PROP(p,'A12') LT 0) + 0$(PROP(p,'A12') GT 0);
107      Kd('11',p)=((5972 * ((PROP(p,'A3')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A3')/PROP(p,'A1'))))$(PROP(p,'A3') GT 0) + 0$(PROP(p,'A3') LT 0);
108      Kd('12',p)=((5972 * ((PROP(p,'A4')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A4')/PROP(p,'A1'))))$(PROP(p,'A4') GT 0) + 0$(PROP(p,'A4') LT 0);
109      Kd('13',p)=((5972 * ((PROP(p,'A5')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A5')/PROP(p,'A1'))))$(PROP(p,'A5') GT 0) + 0$(PROP(p,'A5') LT 0);
110      Kd('14',p)=((5972 * ((PROP(p,'A6')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A6')/PROP(p,'A1'))))$(PROP(p,'A6') GT 0) + 0$(PROP(p,'A6') LT 0);
111      Kd('15',p)=((5972 * ((PROP(p,'A7')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A7')/PROP(p,'A1'))))$(PROP(p,'A7') GT 0) + 0$(PROP(p,'A7') LT 0);
^Selection of Multistep Protein Purification Processes

```

AGE 3

10/30/95 17:27:32 P

6 DOS

GAMS 2.25.087 386/48

```

112      Kd('16',p)=((5972 * ((PROP(p,'A8')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A8')/PROP(p,'A1'))))$(PROP(p,'A8') GT 0) + 0$(PROP(p,'A8') LT 0);
113      Kd('17',p)=((5972 * ((PROP(p,'A9')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A9')/PROP(p,'A1'))))$(PROP(p,'A9') GT 0) + 0$(PROP(p,'A9') LT 0);
114      Kd('18',p)=((5972 * ((PROP(p,'A10')/PROP(p,'A1')))/(1+17065 * (PROP
(p,'A10')/PROP(p,'A1'))))$(PROP(p,'A10') GT 0) + 0$(PROP(p,'A10') LT 0);
115      Kd('19',p)=((5972 * ((PROP(p,'A11')/PROP(p,'A1')))/(1+17065 * (PROP
(p,'A11')/PROP(p,'A1'))))$(PROP(p,'A11') GT 0) + 0$(PROP(p,'A11') LT 0);
116      Kd('20',p)=((5972 * ((PROP(p,'A12')/PROP(p,'A1')))/(1+17065 * (PROP
(p,'A12')/PROP(p,'A1'))))$(PROP(p,'A12') GT 0) + 0$(PROP(p,'A12') LT 0);
117      Kd('21',p)=PROP(p,'A2');
118      Kd('22',p)=-0.4691 * log10(PROP(p,'A1')+2.3902);
119
120  PARAMETER
121      DF(i,p) DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i OF PROTEIN p;
122      DF(i,p) = ABS(sum(pp$(ORD(pp) EQ IP),Kd(i,pp))-Kd(i,p));
123
124  PARAMETER
125      CF(i,p) CONTAMINANT CONCENTRATION AFTER A CHROMATOGRAPHIC STEP;
126  CF(i,p)=1;
127  CF(i,p)$((DF(i,p) GE (SIG(i)/10) AND (DF(i,p) LT (SIG(i)/2)))=1.02*(SQR(SIG
(i))-2*SQR(DF(i,p)))/SQR(SIG(i));
128  CF(i,p)$((DF(i,p) GE (SIG(i)/ 2) AND (DF(i,p) LT (SIG(i) )))=1.02*2*SQR(SI
G(i)-DF(i,p))/SQR(SIG(i));
129  CF(i,p)$((DF(i,p) GE (SIG(i)))=0.02;
130  * CF(i,p)=1.02*(SQR(SIG(i))-2*SQR(DF(i,p)))/SQR(SIG(i));
131
132
133  PARAMETER
134      con(p) VALUE OF INTIAL CONCENTRATION
135
136
137
138
139
140

```

/p1	25.00
p2	11.29
p3	7.06
p4	4.63
p5	5.58
p6	4.83

```

141
142 p7 2.48
143 p8 7.70
144 p9 6.80
145 p10 7.53
146 p11 6.05
147 p12 3.89
148 p13 1.48
149 p14 0.83/
149 parameter
150 U VALUE MAXIME OF INITIAL CONCENTRATION;
151 U=2*sum(p$(ord(p) EQ IP),con(p));
152
INCLUDE C:\GAMS\PROTEIN1.TXT
154 display kd,df,cf;
155
156
157
158 Variables S;
159 positive variables C(p,k), C1(i,p,k);
160 Binary Variables Y(i,k), Z(k);
161 Equations OBJ, EQ1, EQ2, EQ3, EQ4, EQ5, EQ6, EQ7, EQ8,eq9,eq4a,eq4b
162 eq7a
163 eq7b
164 ;
165

```

Selection of Multistep Protein Purification Processes

AGE 4 10/30/95 17:27:32 P

6 DOS GAMS 2.25.087 386/48

```

166 *Objective function
167
168 c.fx(p,'1') = con(p);
169
170 OBJ.. S=E=SUM(k,ord(k)*z(k));
171
172 *Subject to
173
174 EQ1(k)..SUM(i,Y(i,k))=L=1;
175
176 EQ2(i)..SUM(k,Y(i,k))=L=1;
177
178 EQ3(k)$ (ord(k) lt card(k)).. SUM(i,Y(i,k+1))=L=SUM(i,Y(i,k));
179
180 EQ4(p,k)$ (ord(k) ge 3)..C(p,k)-SUM(i,CF(i,p)*C1(i,p,k-1)) =g= -U*(1-sum(i,y(
i,k-1)));
181
182 EQ4b(p,k)$ (ord(k) ge 3)..C(p,k)-SUM(i,CF(i,p)*C1(i,p,k-1)) =l= U*(1-sum(i,y(
i,k-1)));
183
184 EQ4a(p,'2')..C(p,'2')=E=SUM(i,CF(i,p)*y(i,'1')*con(p));
185
186 EQ5(i,p,k)$ (ord(k) ge 2)..C1(i,p,k)=L=con(p)*Y(i,k-1);
187
188 EQ6(p,k)$ (ord(k) ge 2)..C(p,k)=E=SUM(i,C1(i,p,k));
189
190 EQ7(k)$ (ord(k) lt card(k)).. Z(k)=G=SUM(i,Y(i,k))-SUM(i,Y(i,k+1));
191
192 eq7a(k,kk)$ (ord(kk) le ord(k)).. SUM(i,Y(i,kk)) =g= z(k);
193
194 * eq7a(k).. SUM((i,kk)$ (ord(kk) le ord(k)),Y(i,kk)) =g= ord(k)*z(k);
195
196 eq7b(k,kk)$ (ord(kk) gt ord(k)).. SUM(i,Y(i,kk)) + z(k) =l= 1;

```

```

197
198 EQ8..SUM(k,Z(k))=E=1;
199
200 EQ9(p,k)$ (ORD(p) EQ IP).. C(p,k+1)=G=0.90*(SUM(pp,C(pp,k+1)))-U*(1-Z(k));
201
202 option iterlim = 1000000
203       reslim   = 1000000
204       limrow   = 0
205       solprint = off
206       optcr    = 0
207       limcol   = 0
208       optca    = 0;
209
210
211

```

```

212 MODEL PROTEIN1 /OBJ,
213       EQ1, EQ2, EQ3,
214       EQ4,
215       eq4a,eq4b
216       EQ5,
217       EQ6,
218       EQ7, eq7a, eq7b
219       EQ8

```

Selection of Multistep Protein Purification Processes

AGE 5 10/30/95 17:27:32 P

6 DOS GAMS 2.25.087 386/48

```

220 eq9/;
221
222
223 protein1.workspace=10;
224
225 protein1.optfile=1
226
227
228
229 SOLVE PROTEIN1 USING MIP MINIMIZING S;
230
231 display C.1, C1.1, Z.1, Y.1;

```

Selection of Multistep Protein Purification Processes

AGE 6 10/30/95 17:27:32 P  
Include File Summary

6 DOS GAMS 2.25.087 386/48

SEQ	GLOBAL	TYPE	PARENT	LOCAL	FILENAME
1	1	INPUT	0	0	C:\GAMS\DATOS3\TESIS.TXT
2	153	INCLUDE	1	153	.C:\GAMS\PROTEIN1.TXT

COMPILATION TIME = 0.100 SECONDS VERID MW2-25-087  
Selection of Multistep Protein Purification Processes

AGE 7 10/30/95 17:27:32 P  
Execution

----- 154 PARAMETER KD RETENTION TIME FOR PRODUCT p IN CHROMATOGRAPHI  
C STEP i

P7 P14	P1 P8	P2 P9	P3 P10	P4 P11	P5 P12	P6 P13	
3		0.190	0.083		0.002		
4		0.256	0.134	0.038	0.039	0.039	0
.066	0.074						
5		0.281	0.160	0.059	0.058	0.061	0
.141	0.144	0.056	0.003				
6		0.293	0.175	0.075	0.070	0.075	0
.174	0.142	0.126	0.012	0.003			
7		0.303	0.183	0.093	0.087	0.090	0
.191	0.115	0.168	0.021	0.032			
8		0.311	0.187	0.167	0.109	0.109	0
.206	0.111	0.192	0.028	0.056	0.024		
9	0.010	0.316	0.189	0.132	0.127	0.130	0
.220	0.160	0.200	0.032	0.077	0.034		
10	0.123	0.325	0.188	0.143	0.124	0.143	0
.229	0.225	0.197	0.032	0.090	0.052		
11	0.275	0.225	0.112	0.129	0.112	0.089	0
.197	0.204	0.233	0.074	0.213	0.010	0.260	
0.237							
12	0.261	0.066	0.019	0.061	0.057	0.047	0
.153	0.185	0.180	0.042	0.164	9.025149E-4	0.246	
0.206							
13	0.228			0.004		0.001	0
.070	0.099	0.118	0.021	0.112	0.002	0.225	
0.172							
14	0.188						
		0.043	0.008	0.064	0.002	0.195	
0.141							
15	0.155						
				0.027	0.002	0.158	
0.113							
16	0.120				0.002	0.115	
0.092					0.002	0.115	
17	0.030						
					0.002	0.077	
0.074							
18	0.018						
						0.050	
0.059							
19							
						0.035	
0.044							
20							
						0.015	
0.029							
21	0.930	0.710	0.480	0.760	1.500	0.360	0
.360	0.480	0.930		0.630	0.060		
22	0.353	0.390	0.076	0.172	0.008	-0.100	0
.119	0.193	0.059	0.119	0.017	-0.094	0.287	
0.056							

----- 154 PARAMETER DF DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i OF  
PROTEIN p

P8	P2 P9	P3 P10	P4 P11	P5 P12	P6 P13	P7 P14
----	----------	-----------	-----------	-----------	-----------	-----------

3	0.190	0.083		0.002			
4	0.256	0.134	0.038	0.039	0.039	0.066	0
.074							
5	0.281	0.160	0.059	0.058	0.061	0.141	0
.144	0.056	0.003					
6	0.293	0.175	0.075	0.070	0.075	0.174	0
.142	0.126	0.012	0.003				
7	0.303	0.183	0.093	0.087	0.090	0.191	0
.115	0.168	0.021	0.032				
8	0.311	0.187	0.167	0.109	0.109	0.206	0
.111	0.192	0.028	0.056	0.024			
9	0.306	0.179	0.122	0.118	0.121	0.210	0
.150	0.191	0.022	0.067	0.024	0.010	0.010	
10	0.202	0.064	0.020	7.542337E-4	0.019	0.106	0
.101	0.073	0.092	0.034	0.071	0.123	0.123	
11	0.050	0.164	0.147	0.164	0.186	0.079	0
.071	0.043	0.202	0.063	0.266	0.015	0.038	
12	0.196	0.242	0.200	0.204	0.215	0.108	0
.077	0.081	0.220	0.097	0.261	0.015	0.055	
13	0.228	0.228	0.224	0.228	0.227	0.158	0
.129	0.110	0.207	0.116	0.227	0.004	0.056	
14	0.188	0.188	0.188	0.188	0.188	0.188	0
.188	0.145	0.180	0.124	0.187	0.007	0.048	
15	0.155	0.155	0.155	0.155	0.155	0.155	0
.155	0.155	0.155	0.128	0.154	0.002	0.042	
16	0.120	0.120	0.120	0.120	0.120	0.120	0
.120	0.120	0.120	0.120	0.118	0.005	0.028	
17	0.030	0.030	0.030	0.030	0.030	0.030	0
.030	0.030	0.030	0.030	0.028	0.047	0.045	
18	0.018	0.018	0.018	0.018	0.018	0.018	0
.018	0.018	0.018	0.018	0.018	0.032	0.041	
19							
20					0.035	0.044	
21	0.220	0.450	0.170	0.570	0.015	0.029	
.450		0.930	0.300	0.870	0.570	0.570	0
22	0.037	0.277	0.182	0.346	0.930	0.930	
.160	0.295	0.234	0.336	0.448	0.453	0.234	0
.018	0.018	0.018	0.018	0.018	0.066	0.297	

Selection of Multistep Protein Purification Processes

AGE 8  
Execution

10/30/95 17:27:32 P

6 DOS

GAMS 2.25.087 386/48

---- 154 PARAMETER CF CONTAMINANT CONCENTRATION AFTER A CHROMATOGRAP

HIC STEP

	P1	P2	P3	P4	P5	P6	
P7	P8	P9	P10	P11	P12	P13	
P14							
1	1.000	1.000	1.000	1.000	1.000	1.000	1
.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.000							
2	1.000	1.000	1.000	1.000	1.000	1.000	1
.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.000							
3	1.000	0.020	0.407	1.000	1.000	1.000	1
.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.000							
4	1.000	0.020	0.024	0.890	0.883	0.882	0
.627	0.517	1.000	1.000	1.000	1.000	1.000	

1.000							
5	1.000	0.020	0.020	0.705	0.718	0.686	0
.007	0.004	0.733	1.000	1.000	1.000	1.000	
1.000							
6	1.000	0.020	0.020	0.510	0.575	0.509	0
.020	0.005	0.053	1.000	1.000	1.000	1.000	
1.000							
7	1.000	0.020	0.020	0.290	0.364	0.326	0
.020	0.111	0.020	0.979	0.928	1.000	1.000	
1.000							
8	1.000	0.020	0.020	0.020	0.156	0.154	0
.020	0.140	0.020	0.949	0.734	0.966	1.000	
1.000							
9	1.000	0.020	0.020	0.070	0.095	0.079	0
.020	2.526080E-6	0.020	0.977	0.611	0.969	1.000	
1.000							
10	1.000	0.020	0.646	0.985	1.000	0.986	0
.177	0.215	0.531	0.309	0.917	0.558	0.064	
0.064							
11	1.000	0.789	0.020	9.544235E-4	0.020	0.020	0
.460	0.557	0.856	0.020	0.663	0.020	0.999	
0.888							
12	1.000	0.020	0.020	0.020	0.020	0.020	0
.159	0.484	0.428	0.020	0.254	0.020	0.998	
0.741							
13	1.000	0.020	0.020	0.020	0.020	0.020	0
.020	0.039	0.142	0.020	0.102	0.020	1.000	
0.737							
14	1.000	0.020	0.020	0.020	0.020	0.020	0
.020	0.020	0.002	0.020	0.062	0.020	1.000	
0.815							
15	1.000	0.020	0.020	0.020	0.020	0.020	0
.020	0.020	0.020	0.020	0.042	0.020	1.000	
0.860							
16	1.000	0.083	0.083	0.083	0.083	0.083	0
.083	0.083	0.083	0.083	0.083	0.092	1.000	
0.949							
17	1.000	0.939	0.939	0.939	0.939	0.939	0
.939	0.939	0.939	0.939	0.939	0.947	0.821	
0.839							
18	1.000	0.991	0.991	0.991	0.991	0.991	0
.991	0.991	0.991	0.991	0.991	0.991	0.925	
0.866							
19	1.000	1.000	1.000	1.000	1.000	1.000	1
.000	1.000	1.000	1.000	1.000	1.000	0.907	
0.844							
20	1.000	1.000	1.000	1.000	1.000	1.000	1
.000	1.000	1.000	1.000	1.000	1.000	0.999	
0.941							
21	1.000	0.020	0.020	0.105	0.020	0.020	0
.020	0.020	1.000	0.020	0.020	0.020	0.020	
0.020							
22	1.000	1.000	0.324	0.702	0.126	5.115227E-4	0
.492	0.772	0.264	0.492	0.148	0.001	0.978	
0.255							

~Selection of Multistep Protein Purification Processes

AGE 9  
 Model Statistics SOLVE PROTEIN1 USING MIP FROM LINE 229

10/30/95 17:27:32 P

6 DOS GAMS 2.25.087 386/48

MODEL STATISTICS



BLOCKS OF EQUATIONS	14	SINGLE EQUATIONS	4050
BLOCKS OF VARIABLES	5	SINGLE VARIABLES	3819
NON ZERO ELEMENTS	28250	DISCRETE VARIABLES	276

GENERATION TIME = 5.940 SECONDS

EXECUTION TIME = 6.050 SECONDS VERID MW2-25-087  
 ^Selection of Multistep Protein Purification Processes

AGE 10 10/30/95 17:27:32 P  
 Solution Report SOLVE PROTEIN1 USING MIP FROM LINE 229

6 DOS

GAMS 2.25.087 386/48

S O L V E S U M M A R Y

MODEL	PROTEIN	OBJECTIVE	S
TYPE	MIP	DIRECTION	MINIMIZE
SOLVER	OSL	FROM LINE	229

\*\*\*\* SOLVER STATUS 1 NORMAL COMPLETION  
 \*\*\*\* MODEL STATUS 1 OPTIMAL  
 \*\*\*\* OBJECTIVE VALUE 2.0000

RESOURCE USAGE, LIMIT 29.879 1000000.000  
 ITERATION COUNT, LIMIT 690 1000000  
 Reading option file C:\GAMS\OSL.OPT

User supplied option file:

>method simplex  
 >strategy 48  
 >presolve -1  
 >bbpreproc 3

Work space requested by user -- 10.00 Mb  
 Work space requested by solver -- 7.09 Mb

OSL step: Reading data 0.22 Seconds

OSL step: Scale 0.17 Seconds

Range of matrix coefficients:

before scaling :	0.0000 -	50.0000
after scaling :	0.0001 -	1.0000

OSL step: Crash 0.16 Seconds

Crash option : 1

OSL step: Simplex 0.44 Seconds

The dual algorithm has been chosen

Status: Successful (optimal)

Iterations: 40

Objective : 1.0154

OSL step: Preprocessor 28.40 Seconds

Iterations: 650

OSL step: Branch&Bound 0.11 Seconds

Status: Successful (optimal)

Iterations: 0

Nodes : 0

Objective : 2.0000

OSL step: Primal Simplex 0.11 Seconds  
 Status: Successful (optimal)  
 Iterations: 0  
 Objective : 2.0000

~Selection of Multistep Protein Purification Processes

AGE 11  
 Solution Report SOLVE PROTEIN1 USING MIP FROM LINE 229

10/30/95 17:27:32 P

6 DOS

GAMS 2.25.087 386/48

\*\*\*\* REPORT SUMMARY :  
 0 NONOPT  
 0 INFEASIBLE  
 0 UNBOUNDED

~Selection of Multistep Protein Purification Processes

AGE 12  
 Execution

10/30/95 17:27:32 P

6 DOS

GAMS 2.25.087 386/48

---- 231 VARIABLE C.L

	1	2	3
P1	25.000	25.000	25.000
P2	11.290	0.226	0.005
P3	7.060	0.141	0.003
P4	4.630	0.093	0.002
P5	5.580	0.112	0.002
P6	4.830	0.097	0.002
P7	2.480	0.050	9.920000E-4
P8	7.700	0.304	0.012
P9	6.800	0.965	0.137
P10	7.530	0.151	0.003
P11	6.050	0.616	0.063
P12	3.890	0.078	0.002
P13	1.480	1.480	1.480
P14	0.830	0.612	0.451

---- 231 VARIABLE C1.L

	2	3
9 .P1	25.000	
9 .P2	0.005	
9 .P3	0.003	
9 .P4	0.002	
9 .P5	0.002	
9 .P6	0.002	
9 .P7	9.920000E-4	
9 .P8	0.012	
9 .P9	0.137	
9 .P10	0.003	
9 .P11	0.063	
9 .P12	0.002	
9 .P13	1.480	

9 .P14		0.451
13.P1	25.000	
13.P2	0.226	
13.P3	0.141	
13.P4	0.093	
13.P5	0.112	
13.P6	0.097	
13.P7	0.050	
13.P8	0.304	
13.P9	0.965	
13.P10	0.151	
13.P11	0.616	
13.P12	0.078	
13.P13	1.480	
13.P14	0.612	

Selection of Multistep Protein Purification Processes

AGE 13  
Execution

10/30/95 17:27:32 P

6 DOS

GAMS 2.25.087 386/48

---- 231 VARIABLE Z.L

2 1.000

---- 231 VARIABLE Y.L

1 2

9 1.000 1.000  
13

EXECUTION TIME

=

1.100 SECONDS

VERID MW2-25-087

## Selection of Multistep Protein Purification Processes

AGE 1

10/30/95 16:51:14 P

6 DOS

GAMS 2.25.087 386/48

```

4
5 SET
6 I CRHOMATOGRAPHY STEPS /1*22/
7 p PROTEINS /p1*p14/
8 A PROPERTY /A1*A12/
9 k ORDER /1*12;/
10
11 ALIAS (p,pp)
12 (k,kk)
13
14
15 TABLE PROP(p,A) VALUE OF PROPERTY A FOR PROTEIN p
16
17
18
19
10 A11 A12 A1 A2 A3 A4 A5 A6 A7 A8 A9 A
20 p1 22000 0.93 4.77 3.81 2.42 1.50 1.03 0.67 0.12 0.
07 -0.03 -0.50
21 p2 18370 0.71 1.94 0.25 -0.80 -1.41 -1.76 -1.97 -2.15 -2.
33 -2.45 -2.67
22 p3 85570 0.48 2.35 0.29 -1.17 -2.17 -2.83 -3.24 -3.50 -3.
63 -3.68 -3.64
23 p4 53660 0.76 1.83 0.67 0.04 -0.30 -0.49 -0.65 -0.85 -1.
90 -1.34 -1.50
24 p5 120000 1.50 3.29 1.38 -0.03 -0.69 -1.07 -1.34 -1.73 -2.
30 -2.85 -2.75
25 p6 203000 0.36 4.08 1.83 0.04 -1.17 -1.92 -2.46 -3.07 -3.
90 -4.98 -5.65
26 p7 69380 0.36 5.22 3.17 1.02 -0.72 -1.90 -2.60 -3.05 -3.
46 -3.90 -4.24
27 p8 48320 0.48 3.96 3.16 1.12 -0.58 -1.36 -1.34 -1.00 -0.
95 -1.59 -2.84
28 p9 93380 0.93 10.90 5.81 2.78 0.77 -0.81 -2.18 -3.32 -4.
12 -4.45 -4.31
29 p10 69380 1.09 0.55 0.26 0.10 -0.03 -0.12 -0.21 -0.
28 -0.32 -0.32
30 p11 114450 0.63 10.40 5.94 3.15 1.51 0.56 -0.05 -0.53 -0.
99 -1.43 -1.72
31 p12 198000 0.06 0.33 0.03 0.05 0.05 0.05 0.05 0.05 -0.
69 -0.97 -1.57
32 p13 30400 5.17 4.22 3.20 2.25 1.46 0.87 0.50 0.
30 0.20 0.08
33 p14 94670 11.70 7.94 5.39 3.73 2.66 1.97 1.50 1.
13 0.80 0.51;
34
35 *prop(p,A)=abs(prop(p,A));
36
37 *where
38 *P1 PRODUT
39 *P2 CONT_1
40 *P3 CONT_2
41 *P4 CONT_3
42 *P5 CONT_4
43 *P6 CONT_5
44 *p7 CONT_6
45 *P8 CONT_7
46 *P9 CONT_8

```

47 \*P10 CONT\_9  
 48 \*P11 CONT\_10  
 49 \*P12 CONT\_11  
 50 \*P13 CONT\_12  
 51 \*P14 CONT\_13;

52  
 53

54 \*AND  
 55 \*A1 MOLECULAR WEIGHT (Da)  
 56 \*A2 HIDROPHOBICITY  
 57 \*A3 pH 4.0

~Selection of Multistep Protein Purification Processes

AGE 2

10/30/95 16:51:14 P

6 DOS

GAMS 2.25.087 386/48

58 \*A4 pH 4.5  
 59 \*A5 pH 5.0  
 60 \*A6 pH 5.5  
 61 \*A7 pH 6.0  
 62 \*A8 pH 6.5  
 63 \*A9 pH 7.0  
 64 \*A10 pH 7.5  
 65 \*A11 pH 8.0  
 66 \*A12 pH 8.5;

67  
 68 scalar IP /3/  
 69 PARAMETER  
 70 SIG(i) VALUE OF PEAK WIDTH

71		
72		
73	/1	0.15
74	2	0.15
75	3	0.15
76	4	0.15
77	5	0.15
78	6	0.15
79	7	0.15
80	8	0.15
81	9	0.15
82	10	0.15
83	11	0.15
84	12	0.15
85	13	0.15
86	14	0.15
87	15	0.15
88	16	0.15
89	17	0.15
90	18	0.15
91	19	0.15
92	20	0.15
93	21	0.22
94	22	0.46/
95		

96 Kd(i,p) RETENTION TIME FOR PRODUCT p IN CHROMATOGRAPHIC STEP i;  
 97 Kd('1',p)=((7383 \* abs(PROP(p,'A3')/PROP(p,'A1')))/(1+15844 \* abs(PR  
 OP(p,'A3')/PROP(p,'A1'))))\$(PROP(p,'A3') LT 0) + 0\$(PROP(p,'A3') GT 0);  
 98 Kd('2',p)=((7383 \* abs(PROP(p,'A4')/PROP(p,'A1')))/(1+15844 \* abs(PR  
 OP(p,'A4')/PROP(p,'A1'))))\$(PROP(p,'A4') LT 0) + 0\$(PROP(p,'A4') GT 0);  
 99 Kd('3',p)=((7383 \* abs(PROP(p,'A5')/PROP(p,'A1')))/(1+15844 \* abs(PR  
 OP(p,'A5')/PROP(p,'A1'))))\$(PROP(p,'A5') LT 0) + 0\$(PROP(p,'A5') GT 0);  
 100 Kd('4',p)=((7383 \* abs(PROP(p,'A6')/PROP(p,'A1')))/(1+15844 \* abs(PR  
 OP(p,'A6')/PROP(p,'A1'))))\$(PROP(p,'A6') LT 0) + 0\$(PROP(p,'A6') GT 0);

```

101      Kd('5',p)=((7383 * abs(PROP(p,'A7')/PROP(p,'A1')))/(1+15844 * abs(PR
OP(p,'A7')/PROP(p,'A1'))))$(PROP(p,'A7') LT 0) + 0$(PROP(p,'A7') GT 0);
102      Kd('6',p)=((7383 * abs(PROP(p,'A8')/PROP(p,'A1')))/(1+15844 * abs(PR
OP(p,'A8')/PROP(p,'A1'))))$(PROP(p,'A8') LT 0) + 0$(PROP(p,'A8') GT 0);
103      Kd('7',p)=((7383 * abs(PROP(p,'A9')/PROP(p,'A1')))/(1+15844 * abs(PR
OP(p,'A9')/PROP(p,'A1'))))$(PROP(p,'A9') LT 0) + 0$(PROP(p,'A9') GT 0);
104      Kd('8',p)=((7383 * abs(PROP(p,'A10')/PROP(p,'A1')))/(1+15844 * abs(P
ROP(p,'A10')/PROP(p,'A1'))))$(PROP(p,'A10') LT 0) + 0$(PROP(p,'A10') GT 0);
105      Kd('9',p)=((7383 * abs(PROP(p,'A11')/PROP(p,'A1')))/(1+15844 * abs(P
ROP(p,'A11')/PROP(p,'A1'))))$(PROP(p,'A11') LT 0) + 0$(PROP(p,'A11') GT 0);
106      Kd('10',p)=((7383 * abs(PROP(p,'A12')/PROP(p,'A1')))/(1+15844 * abs(
PROP(p,'A12')/PROP(p,'A1'))))$(PROP(p,'A12') LT 0) + 0$(PROP(p,'A12') GT 0);
107      Kd('11',p)=((5972 * ((PROP(p,'A3')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A3')/PROP(p,'A1'))))$(PROP(p,'A3') GT 0) + 0$(PROP(p,'A3') LT 0);
108      Kd('12',p)=((5972 * ((PROP(p,'A4')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A4')/PROP(p,'A1'))))$(PROP(p,'A4') GT 0) + 0$(PROP(p,'A4') LT 0);
109      Kd('13',p)=((5972 * ((PROP(p,'A5')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A5')/PROP(p,'A1'))))$(PROP(p,'A5') GT 0) + 0$(PROP(p,'A5') LT 0);
110      Kd('14',p)=((5972 * ((PROP(p,'A6')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A6')/PROP(p,'A1'))))$(PROP(p,'A6') GT 0) + 0$(PROP(p,'A6') LT 0);
111      Kd('15',p)=((5972 * ((PROP(p,'A7')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A7')/PROP(p,'A1'))))$(PROP(p,'A7') GT 0) + 0$(PROP(p,'A7') LT 0);
~Selection of Multistep Protein Purification Processes

```

AGE 3

10/30/95 16:51:14 P

6 DOS

GAMS 2.25.087 386/48

```

112      Kd('16',p)=((5972 * ((PROP(p,'A8')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A8')/PROP(p,'A1'))))$(PROP(p,'A8') GT 0) + 0$(PROP(p,'A8') LT 0);
113      Kd('17',p)=((5972 * ((PROP(p,'A9')/PROP(p,'A1')))/(1+17065 * (PROP(
p,'A9')/PROP(p,'A1'))))$(PROP(p,'A9') GT 0) + 0$(PROP(p,'A9') LT 0);
114      Kd('18',p)=((5972 * ((PROP(p,'A10')/PROP(p,'A1')))/(1+17065 * (PROP
(p,'A10')/PROP(p,'A1'))))$(PROP(p,'A10') GT 0) + 0$(PROP(p,'A10') LT 0);
115      Kd('19',p)=((5972 * ((PROP(p,'A11')/PROP(p,'A1')))/(1+17065 * (PROP
(p,'A11')/PROP(p,'A1'))))$(PROP(p,'A11') GT 0) + 0$(PROP(p,'A11') LT 0);
116      Kd('20',p)=((5972 * ((PROP(p,'A12')/PROP(p,'A1')))/(1+17065 * (PROP
(p,'A12')/PROP(p,'A1'))))$(PROP(p,'A12') GT 0) + 0$(PROP(p,'A12') LT 0);
117      Kd('21',p)=PROP(p,'A2');
118      Kd('22',p)=-0.4691 * log10(PROP(p,'A1')+2.3902);
119
120  PARAMETER
121      DF(i,p) DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i OF PROTEIN p;
122      DF(i,p) = ABS(sum(pp$(ORD(pp) EQ IP),Kd(i,pp))-Kd(i,p));
123
124  PARAMETER
125      CF(i,p) CONTAMINANT CONCENTRATION AFTER A CHROMATOGRAPHIC STEP;
126  CF(i,p)=1;
127  CF(i,p)$((DF(i,p) GE (SIG(i)/10)) AND (DF(i,p) LT (SIG(i)/2)))=1.02*(SQR(SIG
(i))-2*SQR(DF(i,p)))/SQR(SIG(i));
128  CF(i,p)$((DF(i,p) GE (SIG(i)/ 2)) AND (DF(i,p) LT (SIG(i) )))=1.02*2*SQR(SI
G(i)-DF(i,p))/SQR(SIG(i));
129  CF(i,p)$DF(i,p) GE (SIG(i))=0.02;
130  * CF(i,p)=1.02*(SQR(SIG(i))-2*SQR(DF(i,p)))/SQR(SIG(i));
131
132
133  PARAMETER
134      con(p) VALUE OF INTIAL CONCENTRATION
135
136      /p1 25.00
137      p2 11.29
138      p3 7.06
139      p4 4.63
140      p5 5.58
141      p6 4.83

```

```

141
142
143
144
145
146
147
148
149 parameter
150 U VALUE MAXIME OF INITIAL CONCENTRATION;
151 U=2*sum(p$(ord(p) EQ IP),con(p));
152
INCLUDE C:\GAMS\PROTEIN1.TXT
154 display kd,df,cf;
155
156
157
158 Variables S;
159 positive variables C(p,k), C1(i,p,k);
160 Binary Variables Y(i,k), Z(k);
161 Equations OBJ, EQ1, EQ2, EQ3, EQ4, EQ5, EQ6, EQ7, EQ8,eq9,eq4a,eq4b
162 eq7a
163 eq7b
164 ;
165
^Selection of Multistep Protein Purification Processes
-----
AGE 4 10/30/95 16:51:14 P

6 DOS GAMS 2.25.087 386/48

166 *Objective function
167
168 c.fx(p,'1') = con(p);
169
170 OBJ.. S=E=SUM(k,ord(k)*z(k));
171
172 *Subject to
173
174 EQ1(k)..SUM(i,Y(i,k))=L=1;
175
176 EQ2(i)..SUM(k,Y(i,k))=L=1;
177
178 EQ3(k)$ (ord(k) lt card(k)).. SUM(i,Y(i,k+1))=L=SUM(i,Y(i,k));
179
180 EQ4(p,k)$ (ord(k) ge 3)..C(p,k)-SUM(i,CF(i,p)*C1(i,p,k-1)) =g= -U*(1-sum(i,Y(
i,k-1)));
181
182 EQ4b(p,k)$ (ord(k) ge 3)..C(p,k)-SUM(i,CF(i,p)*C1(i,p,k-1)) =l= U*(1-sum(i,Y(
i,k-1)));
183
184 EQ4a(p,'2')..C(p,'2')=E=SUM(i,CF(i,p)*y(i,'1')*con(p));
185
186 EQ5(i,p,k)$ (ord(k) ge 2)..C1(i,p,k)=L=con(p)*Y(i,k-1);
187
188 EQ6(p,k)$ (ord(k) ge 2)..C(p,k)=E=SUM(i,C1(i,p,k));
189
190 EQ7(k)$ (ord(k) lt card(k)).. Z(k)=G=SUM(i,Y(i,k))-SUM(i,Y(i,k+1));
191
192 eq7a(k,kk)$ (ord(kk) le ord(k)).. SUM(i,Y(i,kk)) =g= z(k);
193
194 * eq7a(k).. SUM((i,kk)$ (ord(kk) le ord(k)),Y(i,kk)) =g= ord(k)*z(k);
195
196 eq7b(k,kk)$ (ord(kk) gt ord(k)).. SUM(i,Y(i,kk)) + z(k) =l= 1;

```

```

197
198 EQ8..SUM(k,Z(k))=E=1;
199
200 EQ9(p,k)$ (ORD(p) EQ IP).. C(p,k+1)=G=0.98*(SUM(pp,C(pp,k+1)))-U*(1-Z(k));
201
202 option iterlim = 1000000
203         reslim  = 1000000
204         limrow  = 0
205         solprint = off
206         optcr   = 0
207         limcol  = 0
208         optca   = 0;
209
210
211
212 MODEL PROTEIN1 /OBJ,
213         EQ1, EQ2, EQ3,
214         EQ4,
215         eq4a,eq4b
216         EQ5,
217         EQ6,
218         EQ7, eq7a, eq7b
219         EQ8

```

^Selection of Multistep Protein Purification Processes

AGE 5 10/30/95 16:51:14 P

6 DOS GAMS 2.25.087 386/48

```

220         eq9/;
221
222
223 protein1.workspace=10;
224
225 protein1.optfile=1
226
227
228
229 SOLVE PROTEIN1 USING MIP MINIMIZING S;
230
231 display C.l, C1.l, Z.l, Y.l;

```

^Selection of Multistep Protein Purification Processes

AGE 6 10/30/95 16:51:14 P  
Include File Summary

6 DOS GAMS 2.25.087 386/48

SEQ	GLOBAL	TYPE	PARENT	LOCAL	FILENAME
1	1	INPUT	0	0	C:\GAMS\DATOS3\TESIS.TXT
2	153	INCLUDE	1	153	.C:\GAMS\PROTEIN1.TXT

COMPILATION TIME = 0.110 SECONDS VERID MW2-25-087  
^Selection of Multistep Protein Purification Processes

AGE 7 10/30/95 16:51:14 P  
E x e c u t i o n



6 DOS

GAMS 2.25.087 386/48

---- 154 PARAMETER KD RETENTION TIME FOR PRODUCT p IN CHROMATOGRAPHY  
C STEP i

P7 P14	P1 P8	P2 P9	P3 P10	P4 P11	P5 P12	P6 P13	
3		0.190	0.083		0.002		
4		0.256	0.134	0.038	0.039	0.039	0
.066	0.074						
5		0.281	0.160	0.059	0.058	0.061	0
.141	0.144	0.056	0.003				
6		0.293	0.175	0.075	0.070	0.075	0
.174	0.142	0.126	0.012	0.003			
7		0.303	0.183	0.093	0.087	0.090	0
.191	0.115	0.168	0.021	0.032			
8		0.311	0.187	0.167	0.109	0.109	0
.206	0.111	0.192	0.028	0.056	0.024		
9	0.010	0.316	0.189	0.132	0.127	0.130	0
.220	0.160	0.200	0.032	0.077	0.034		
10	0.123	0.325	0.188	0.143	0.124	0.143	0
.229	0.225	0.197	0.032	0.090	0.052		
11	0.275	0.225	0.112	0.129	0.112	0.089	0
.197	0.204	0.233	0.074	0.213	0.010	0.260	
0.237							
12	0.261	0.066	0.019	0.061	0.057	0.047	0
.153	0.185	0.180	0.042	0.164	9.025149E-4	0.246	
0.206							
13	0.228			0.004		0.001	0
.070	0.099	0.118	0.021	0.112	0.002	0.225	
0.172							
14	0.188						
		0.043	0.008	0.064	0.002	0.195	
0.141							
15	0.155						
				0.027	0.002	0.158	
0.113							
16	0.120						
					0.002	0.115	
0.092							
17	0.030						
					0.002	0.077	
0.074							
18	0.018						
						0.050	
0.059							
19							
						0.035	
0.044							
20							
						0.015	
0.029							
21	0.930	0.710	0.480	0.760	1.500	0.360	0
.360	0.480	0.930		0.630	0.060		
22	0.353	0.390	0.076	0.172	0.008	-0.100	0
.119	0.193	0.059	0.119	0.017	-0.094	0.287	
0.056							

---- 154 PARAMETER DF DEVIATION FACTOR FOR CHROMATOGRAPHY STEP i OF  
PROTEIN p

P8	P1 P9	P2 P10	P4 P11	P5 P12	P6 P13	P7 P14
----	----------	-----------	-----------	-----------	-----------	-----------

3	0.083	0.107	0.083	0.081	0.083	0.083	0
.083	0.083	0.083	0.083	0.083	0.083	0.083	0
4	0.134	0.122	0.096	0.095	0.095	0.068	0
.059	0.134	0.134	0.134	0.134	0.134	0.134	0
5	0.160	0.121	0.101	0.103	0.099	0.019	0
.017	0.104	0.157	0.160	0.160	0.160	0.160	0
6	0.175	0.119	0.100	0.105	0.100	0.001	0
.032	0.049	0.162	0.172	0.175	0.175	0.175	0
7	0.183	0.119	0.090	0.097	0.093	0.008	0
.068	0.015	0.162	0.151	0.183	0.183	0.183	0
8	0.187	0.124	0.020	0.079	0.079	0.018	0
.077	0.004	0.159	0.131	0.163	0.187	0.187	0
9	0.179	0.127	0.057	0.061	0.058	0.031	0
.029	0.012	0.157	0.112	0.155	0.189	0.189	0
10	0.064	0.137	0.045	0.063	0.045	0.042	0
.037	0.009	0.156	0.098	0.136	0.188	0.188	0
11	0.164	0.113	0.017	1.280650E-4	0.022	0.085	0
.092	0.121	0.038	0.101	0.102	0.149	0.126	0
12	0.242	0.047	0.042	0.038	0.028	0.134	0
.165	0.161	0.023	0.145	0.018	0.227	0.187	0
13	0.228		0.004		0.001	0.070	0
.099	0.118	0.021	0.112	0.002	0.225	0.172	0
14	0.188						
	0.043	0.008	0.064	0.002	0.195	0.141	
15	0.155						
			0.027	0.002	0.158	0.113	
16	0.120						
				0.002	0.115	0.092	
17	0.030						
				0.002	0.077	0.074	
18	0.018						
					0.050	0.059	
19							
					0.035	0.044	
20							
					0.015	0.029	
21	0.450	0.230	0.280	1.020	0.120	0.120	
	0.450	0.480	0.150	0.420	0.480	0.480	
22	0.277	0.313	0.095	0.069	0.176	0.043	0
.116	0.018	0.043	0.059	0.171	0.211	0.021	0

Selection of Multistep Protein Purification Processes

AGE 8  
Execution

10/30/95 16:51:14 P

6 DOS

GAMS 2.25.087 386/48

---- 154 PARAMETER CF CONTAMINANT CONCENTRATION AFTER A CHROMATOGRAPHIC STEP

	P1	P2	P3	P4	P5	P6	
P7	P8	P9	P10	P11	P12	P13	
P14							
1	1.000	1.000	1.000	1.000	1.000	1.000	1
.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.000							
2	1.000	1.000	1.000	1.000	1.000	1.000	1
.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.000							
3	0.407	0.165	1.000	0.407	0.430	0.407	0
.407	0.407	0.407	0.407	0.407	0.407	0.407	
0.407							
4	0.024	0.070	1.000	0.268	0.278	0.279	0

.604	0.703	0.024	0.024	0.024	0.024	0.024	
0.024							
5	0.020	0.078	1.000	0.215	0.204	0.231	0
.987	0.995	0.193	0.020	0.020	0.020	0.020	
0.020							
6	0.020	0.089	1.000	0.229	0.186	0.230	1
.000	0.924	0.803	0.020	0.020	0.020	0.020	
0.020							
7	0.020	0.084	1.000	0.329	0.259	0.293	1
.000	0.599	0.999	0.020	0.020	0.020	0.020	
0.020							
8	0.020	0.062	1.000	0.984	0.460	0.463	0
.989	0.488	1.000	0.020	0.032	0.020	0.020	
0.020							
9	0.020	0.046	1.000	0.728	0.678	0.711	0
.935	0.943	1.000	0.020	0.132	0.020	0.020	
0.020							
10	0.646	0.015	1.000	0.840	0.655	0.836	0
.863	0.895	1.000	0.020	0.245	0.019	0.020	
0.020							
11	0.020	0.121	1.000	0.994	1.000	0.975	0
.382	0.301	0.075	0.891	0.217	0.209	1.774240E-4	
0.053							
12	0.020	0.821	1.000	0.857	0.887	0.951	0
.023	0.020	0.020	0.974	0.002	0.990	0.020	
0.020							
13	0.020	1.000	1.000	1.000	1.000	1.000	0
.573	0.234	0.093	0.980	0.132	1.000	0.020	
0.020							
14	0.020	1.000	1.000	1.000	1.000	1.000	1
.000	1.000	0.851	1.000	0.645	1.000	0.020	
0.008							
15	0.020	1.000	1.000	1.000	1.000	1.000	1
.000	1.000	1.000	1.000	0.954	1.000	0.020	
0.121							
16	0.083	1.000	1.000	1.000	1.000	1.000	1
.000	1.000	1.000	1.000	1.000	1.000	0.112	
0.308							
17	0.939	1.000	1.000	1.000	1.000	1.000	1
.000	1.000	1.000	1.000	1.000	1.000	0.487	
0.517							
18	0.991	1.000	1.000	1.000	1.000	1.000	1
.000	1.000	1.000	1.000	1.000	1.000	0.789	
0.702							
19	1.000	1.000	1.000	1.000	1.000	1.000	1
.000	1.000	1.000	1.000	1.000	1.000	0.907	
0.844							
20	1.000	1.000	1.000	1.000	1.000	1.000	1
.000	1.000	1.000	1.000	1.000	1.000	0.999	
0.941							
21	0.020	0.020	1.000	0.020	0.020	0.421	0
.421	1.000	0.020	0.020	0.207	0.020	0.020	
0.020							
22	0.324	0.207	1.000	0.933	0.974	0.721	1
.000	0.889	1.000	1.000	0.986	0.738	0.591	
1.000							

Selection of Multistep Protein Purification Processes

AGE 9 10/30/95 16:51:14 P  
 Model Statistics SOLVE PROTEIN1 USING MIP FROM LINE 229

6 DOS GAMS 2.25.087 386/48

## MODEL STATISTICS

BLOCKS OF EQUATIONS	14	SINGLE EQUATIONS	4050
BLOCKS OF VARIABLES	5	SINGLE VARIABLES	3819
NON ZERO ELEMENTS	28250	DISCRETE VARIABLES	276

GENERATION TIME = 5.930 SECONDS

EXECUTION TIME = 6.040 SECONDS VERID MW2-25-087  
 ^Selection of Multistep Protein Purification Processes

AGE 10 10/30/95 16:51:14 P  
 Solution Report SOLVE PROTEIN1 USING MIP FROM LINE 229

6 DOS GAMS 2.25.087 386/48

## S O L V E S U M M A R Y

MODEL	PROTEIN1	OBJECTIVE	S
TYPE	MIP	DIRECTION	MINIMIZE
SOLVER	OSL	FROM LINE	229

\*\*\*\* SOLVER STATUS 1 NORMAL COMPLETION  
 \*\*\*\* MODEL STATUS 1 OPTIMAL  
 \*\*\*\* OBJECTIVE VALUE 5.0000

RESOURCE USAGE, LIMIT 1535.879 1000000.000  
 ITERATION COUNT, LIMIT 151872 1000000  
 Reading option file C:\GAMS\OSL.OPT

User supplied option file:

>method simplex  
 >strategy 48  
 >presolve -1  
 >bbpreproc 3

Work space requested by user -- 10.00 Mb  
 Work space requested by solver -- 7.09 Mb

OSL step: Reading data 0.21 Seconds

OSL step: Scale 0.16 Seconds

Range of matrix coefficients:  
 before scaling : 0.0002 - 25.0000  
 after scaling : 0.0016 - 1.0000

OSL step: Crash 0.17 Seconds  
 Crash option : 1

OSL step: Simplex 1.21 Seconds  
 The dual algorithm has been chosen  
 Status: Successful (optimal)  
 Iterations: 126  
 Objective : 1.4736

OSL step: Preprocessor 45.81 Seconds  
 Iterations: 744

OSL step: Branch&Bound 1487.87 Seconds  
 Status: Successful (optimal)  
 Iterations: 151002

Nodes : 93  
Objective : 5.0000

OSL step: Primal Simplex 0.17 Seconds  
Status: Successful (optimal)  
Iterations: 0  
Objective : 5.0000

~Selection of Multistep Protein Purification Processes

AGE 11 10/30/95 16:51:14 P  
Solution Report SOLVE PROTEIN1 USING MIP FROM LINE 229

6 DOS GAMS 2.25.087 386/48

\*\*\*\* REPORT SUMMARY : 0 NONOPT  
0 INFEASIBLE  
0 UNBOUNDED

~Selection of Multistep Protein Purification Processes

AGE 12 10/30/95 16:51:14 P  
E x e c u t i o n

6 DOS GAMS 2.25.087 386/48

---- 231 VARIABLE C.L

	1	2	3	4	5	6
P1	25.000	0.612	0.015	3.000051E-4	6.000102E-6	
P2	11.290	0.794	0.056	0.003	0.003	5.689106E-5
P3	7.060	7.060	7.060	7.060	7.060	7.060
P4	4.630	1.240	0.332	0.327	0.280	0.006
P5	5.580	1.550	0.430	0.198	0.176	0.004
P6	4.830	1.345	0.375	0.173	0.165	0.070
P7	2.480	1.497	0.904	0.894	0.020	0.009
P8	7.700	5.415	3.808	1.859	0.037	0.037
P9	6.800	0.167	0.004	0.004	8.160139E-5	1.632028E-6
P10	7.530	0.184	0.005	9.036154E-5	8.799605E-5	1.759921E-6
P11	6.050	0.148	0.004	1.170805E-4		
P12	3.890	0.095	0.002	4.668080E-5	4.620780E-5	
P13	1.480	0.036	8.880152E-4	1.776030E-5		
P14	0.830	0.020	4.980085E-4	9.960170E-6		

---- 231 VARIABLE C1.L

	2	3	4	5	6
4 .P1	0.612				
4 .P2	0.794				
4 .P3	7.060				
4 .P4	1.240				
4 .P5	1.550				
4 .P6	1.345				
4 .P7	1.497				
4 .P8	5.415				
4 .P9	0.167				
4 .P10	0.184				
4 .P11	0.148				
4 .P12	0.095				

```

4 .P13      0.036
4 .P14      0.020
8 .P1              0.015
8 .P2              0.056
8 .P3              7.060
8 .P4              0.332
8 .P5              0.430
8 .P6              0.375
8 .P7              0.904
8 .P8              3.808
8 .P9              0.004
8 .P10           0.005
8 .P11           0.004
8 .P12           0.002
8 .P13           8.880152E-4
8 .P14           4.980085E-4

```

```

5.689106E-5
7.060

```

Selection of Multistep Protein Purification Processes

10/30/95 16:51:14 P

AGE 13  
Execution

GAMS 2.25.087 386/48

6 DOS

231 VARIABLE C1.L

	2	3	4	5	6
11.P4					0.006
11.P5					0.004
11.P6					0.070
11.P7					0.009
11.P8					0.037
11.P9					1.632028E-6
11.P10					1.759921E-6
12.P1		3.000051E-4			
12.P2		0.003			
12.P3		7.060			
12.P4		0.327			
12.P5		0.198			
12.P6		0.173			
12.P7		0.894			
12.P8		1.859			
12.P9		0.004			
12.P10		9.036154E-5			
12.P11		1.170805E-4			
12.P12		4.668080E-5			
12.P13		1.776030E-5			
12.P14		9.960170E-6			
21.P1				6.000102E-6	
21.P2				0.003	
21.P3				7.060	
21.P4				0.280	
21.P5				0.176	
21.P6				0.165	
21.P7				0.020	
21.P8				0.037	
21.P9				8.160139E-5	
21.P10				8.799605E-5	
21.P12				4.620780E-5	

---- 231 VARIABLE Z.L

5 1.000

---- 231 VARIABLE Y.L

	1	2	3	4	5
4	1.000				
8		1.000			
11					1.000
12			1.000		
21				1.000	

~Selection of Multistep Protein Purification Processes

AGE 14  
E x e c u t i o n

10/30/95 16:51:14 P

6 DOS GAMS 2.25.087 386/48

EXECUTION TIME = 1.150 SECONDS VERID MW2-25-087