

00582



UNIVERSIDAD NACIONAL AUTÓNOMA DE  
MÉXICO

**PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS  
QUÍMICAS**

**“BÚSQUEDA DE BLANCOS TERAPEÚTICOS EMPLEANDO UN MODELO DE SEÑALES  
INTRACELULARES MEDIADAS POR RAS”**

TESIS  
PARA OPTAR POR EL GRADO DE  
**DOCTOR EN CIENCIAS**

PRESENTA

M en C MAURA CÁRDENAS GARCÍA



**Posgrado**  
**Ciencias**  
**Químicas**

TUTOR Dr. JAIME LAGUNEZ OTERO

2005

m341244



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



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

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

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

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

## DEDICATORIA

Dedico este trabajo realizado a la memoria de mi papá el Ing. Daniel Cárdenas

Chimal con todo cariño a mi familia:

mi mamá Maura García de Cárdenas

mis hermanas y hermanos Daniel, Güerin, Ramón Peque y Lili

a mis cuñadas Gaby y Lina, sobrinos Sergito, Danielito, Betito y Martín, y

sobrinas Linita, Dianis y Gaby.

A mi esposo y al mejor de los hijos José Boris.

Los químicos son  
una extraña clase de mortales motivados  
por un impulso cercano a la locura para buscar  
placer entre humo y vapor, hollín y  
fuego, venenos y pobreza, pero en  
todos esos males encuentro tal felicidad  
que moriría si cambiara mi lugar  
con el del Rey de Persia.

Johann Joachim Becker 1669

## AGRADECIMIENTOS

Agradezco a mi tutor Jaime Lagunez Otero, y a Pedro Pablo González Pérez, Alejandro Zentella Dehesa y Alejandro García Carrancá por su ayuda desinteresada durante el desarrollo de esta tesis.

Gracias a Alejandro García Carrancá y a Marco Antonio Cerbón Cervantes por sus comentarios y sugerencias y por ser parte de mi comité tutorial.

Agradezco a todos los miembros del jurado por sus indicaciones, comentarios y sugerencias Jorge Vázquez Ramos, Alejandro Zentella Dehesa, Enedina Jiménez Cardoso, Alejandro García Carranca, Roberto Martínez, Pedro Pablo González Pérez y Rogelio Pereda Miranda.

Muchas gracias a mi amiga Rocío Alcántara Hernández

Gracias a todas y cada una de las personas que contribuyeron a la realización de este trabajo.

Agradezco el apoyo experimental sin el cual la mitad de esta tesis no existiría a Rocío Alcántara del IFC, a Enrique Pérez del INCAN, a Alejandro Zentella, a José Luis, a Francisco Javier del IFC y a Federico Martínez de la FM.

Esta tesis se realizó en la UNAM en el Instituto de Química, en el Instituto de Fisiología Celular y en la Facultad de Medicina. Con el apoyo económico de CONACyT la beca de Doctorado con registro No. 121824, el Programa de Apoyo a los Estudiantes de Posgrado de la UNAM (PAEP) proyectos 108306 y 207307.

**CONTENIDO**

<b>Indice</b>	<b>i</b>
<b>Resumen</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>I. Introducción</b>	<b>1</b>
<b>II. Objetivos</b>	<b>38</b>
<b>III. Material y métodos</b>	<b>39</b>
<b>IV. Resultados y discusión</b>	<b>47</b>
<b>V. Conclusiones</b>	<b>74</b>
<b>VI. Referencias</b>	<b>76</b>
<b>VII. Artículos</b>	<b>SP</b>

## INDICE

Resumen	iii
Abstract	v
<b>I. INTRODUCCIÓN</b>	
I.I Modelos computacionales de sistemas biológicos	1
1.0 Sistemas dinámicos	2
2.0 Sistemas expertos	2
2.1 Elementos básicos de los sistemas expertos	2
2.2 Areas de aplicación de la inteligencia artificial	3
2.3 Sistemas basados en el conocimiento	4
2.4 Sistemas expertos	4
2.4.1 Estructura de un sistema experto	4
2.4.2 Desarrollo de un sistema experto	7
2.5 Xpert rule KBS	7
2.5.1 La plataforma xpert rule KBS	8
2.5.2 Herramientas de xpert rule KBS	10
2.5.3 Form painter	11
2.5.4 Xpert rule KBS forms	11
<b>I.II Señalización intracelular</b>	13
1. Señalización	13
2. Proteínas Ras	17
2.1 Activación de Ras	20
2.2 Vías de señalización en las que participa Ras	21
2.3 Proteínas efectoras de Ras	22
2.3.1 Raf	22
2.3.2 PI3K	22
2.3.3 Ral-GDS	23
2.4 Regulación del ciclo celular y apoptosis por la vía Ras/Raf/MEK/ERK	23
2.4.1 Efecto proliferativo y antiproliferativo de la vía Ras/Raf/MEK/ERK	23
2.4.2 Regulación de pRb	24
2.4.3 Inducción de ARF y p53	24
2.5 Proteínas 14-3-3	25
<b>I.III Oligonucleótidos que forman una triple hélice (triplex)</b>	32
1. Oligonucleótidos terapéuticos	32
<b>II. OBJETIVOS</b>	38
<b>III. MATERIALES Y MÉTODOS</b>	
III.I Diseño de oligonucleótidos que forman una triple hélice	39
III.II Formación de triplex <i>in vitro</i> .	41
III.III Experimentos en líneas celulares	43

<b>IV. RESULTADOS Y DISCUSIÓN</b>	
<b>1.0 MODELOS COMPUTACIONALES</b>	47
1.1 Primer modelo	49
1.2 Sistema experto	54
<b>2.0 RESULTADOS EXPERIMENTALES</b>	
2.1 Formación de triplex <i>in vitro</i>	62
2.1.1 Desnaturalización térmica	62
2.1.2 Dicroísmo circular	63
2.1.3 Movilidad en gel	64
2.2 Experimentos en líneas celulares	65
2.2.1 Internalización del oligonucleótido	67
2.2.2 Efecto de la concentración del oligonucleótido MCG3 sobre el crecimiento de la línea celular MCF7 en cultivo.	68
2.2.3 Efecto del oligonucleótido MCG3 sobre el ciclo celular de la línea celular MCF7 en cultivo.	71
2.2.4 Efecto del oligonucleótido MCG3 sobre la expresión de la proteína 14-3-3 $\sigma$ en la línea celular MCF7.	72
<b>V. CONCLUSIONES</b>	74
<b>VI. REFERENCIAS</b>	76
<b>VII. ARTICULOS</b>	SP



**RESÚMEN**

El trabajo de esta tesis está dedicado a la creación de modelos computacionales de las vías de señalización intracelular para la búsqueda de blancos terapéuticos, potencialmente útiles en el control de enfermedades.

A las vías de señalización intracelular las conforman elementos de diferente naturaleza química (proteínas, iones, hormonas lipídicas entre otros). Las reacciones en las que participan los diferentes elementos sirven para transmitir señales exteriores a través de la membrana plasmática hacia el interior de la célula; o señales internas a través de receptores nucleares. En ambos casos finalmente se provocan respuestas celulares como por ejemplo diferenciación, envejecimiento y muerte celular. Generalmente, los elementos de las vías forman una complicada red de interacciones y el comportamiento del sistema completo es difícil de comprender y predecir. Por ejemplo, no es fácil determinar la ruta que seguirá una señal para dar la respuesta celular final.

El progreso de las técnicas de computación en las últimas décadas hace factible el desarrollo de programas que ayudan a representar y modelar el comportamiento de complicados sistemas biológicos. En esta tesis hemos estudiado las posibilidades de la aplicación de estas técnicas para modelar las vías de señalización con la esperanza de obtener información que puede ser útil para la identificación de posibles blancos terapéuticos. Consideramos un blanco terapéutico a una proteína esencial para la célula.

Las vías seleccionadas para modelar son aquellas en las que participa la proteína Ras (producto del protooncogen *ras*. Los protooncogenes son genes normales que generalmente están relacionados con la regulación de la proliferación celular, que por mutación pueden convertirse en oncogenes o genes promotores de cáncer). Se eligió a esta proteína, ya que es un punto de convergencia para varias vías de señalización, implicadas en el desarrollo y control de la proliferación celular, así como en el control del ciclo celular.

Se han desarrollado dos modelos. El primero se hizo en base al cálculo tradicional para sistemas dinámicos, donde el estado del sistema en el siguiente momento del tiempo depende de su estado en el momento anterior. Para la computación los momentos de tiempo son discretos y no continuos. Finalmente el sistema llega a algún estado de equilibrio, o a la repetición continua de cierto número de estados llamado ciclo límite. En este caso a cada uno de los elementos que participa en las vías de señalización se le asigna un valor inicial (activo o inactivo). El problema, entonces, se reduce matemáticamente a la determinación del estado final del sistema, dado su estado inicial. La dificultad consiste en el gran número de posibles combinaciones iniciales, porque típicamente el número de combinaciones depende de manera exponencial del tamaño de sistema. Usualmente las vías de señalización incluyen más de 100 elementos, por lo que el cálculo directo se hace imposible. Hemos desarrollado y aplicado la técnica que permite reducir de manera drástica el volumen de cálculo, utilizando el hecho de que existen ciertos grupos de elementos de las vías de señalización que tienen interacciones entre sí. En este caso se consideran posibles blancos terapéuticos a aquellos elementos que no

importando la combinación de estados (activo o inactivo), se requiere que siempre estén activos; si no están activos el sistema no funciona.

El segundo modelo es un sistema experto, un programa basado en técnicas derivadas de la inteligencia artificial. Los sistemas expertos emplean una base de conocimientos y un programa que infiere nuevos hechos a partir de esta base. En este caso cada uno de los elementos que participa en las vías de señalización cambia de estado activo a inactivo empleando una serie de reglas de inferencia, almacenadas en la base de conocimientos; esta base se realizó empleando la información experimental reportada en la literatura. El motor de inferencia determina qué reglas de inferencia deben cumplirse para llegar a una combinación final de estados (activo o inactivo). En este caso una proteína blanco será aquella cuya ausencia llevará a una combinación final de estados, que represente a la muerte celular.

Como un ejemplo particular, nuestro sistema experto indica que la inhibición de la proteína 14-3-3 $\sigma$  puede causar la muerte celular. Realizamos experimentos en la línea celular MCF 7 de cáncer de mama, utilizando oligonucleótidos que forman una triple hélice de ADN o triplex, con una secuencia de ADN para inhibir la producción de esta proteína. Nuestros resultados muestran que existe una disminución de los niveles de la proteína 14-3-3 $\sigma$  y disminución en la viabilidad celular.

Nuestra experiencia con los sistemas computacionales nos indica, que aunque ellos muestran algunas dependencias útiles y pueden ser de ayuda, todavía se necesitan mucho más datos sobre las vías de señalización y refinamiento de técnicas computacionales para llegar a modelos confiables.

El comportamiento real de las células y sus vías de señalización es más complicado de lo que actualmente se puede con seguridad predecir utilizando las técnicas de computación. Si embargo, los modelos computacionales de sistemas biológicos se desarrollan rápidamente y tienen gran potencial para el futuro.

**ABSTRACT**

The work of this thesis is related to the creation of computer models of intercellular signal transduction pathways to the purpose of finding the therapeutic targets, which can be potentially useful for illness control.

The intercellular signal transduction pathways contain agents of different chemical nature (proteins, ions, and hormones, among them). The reactions in which different elements participate serve to transmit exterior signals through the plasmatic membrane into the cell; or to transmit internal signals through the receptors of nuclei. In both cases the cellular response, for example cell differentiation, aging and death. Normally, the pathway elements form the complicated network of interactions, and the behavior of the system as a whole is difficult to understand and to predict. For example, it is not easy to determine the route of the signal to obtain the final cellular response.

The recent progress of computer techniques makes possible the development of the programs, which help to represent and model the behavior of complicated biological systems. In this thesis we have studied the possibility of the application of these techniques to model the signal pathways with a hope to obtain the information useful for determining possible therapeutic targets. We consider as the therapeutic target the protein, which is vital for the cell.

The pathways selected for modeling are those in which the Ras protein is involved. The Ras protein is the product of ras protooncogen (protooncogenes are normal genes, usually related to the cell proliferation regulation, which by mutation can be converted to oncogenes, or cancer promoting genes.) We have chosen this protein, because it is the convergence points for various signal pathways related to the development and cellular proliferation control, as well as in the cellular cycle control.

Two models have been developed. The first one was made on the basis of the traditional calculation for the dynamic systems in which the System State at the next moment depends on its state in the previous moment. For the computation the time is discrete, and not continuous. Finally, the system arrives to some equilibrium state, or to the repetition of a certain number of states called the limit cycle. In this case to every one of the elements participating in the signal transduction pathways the initial value (active or non-active is assigned). The problem thus is reduced mathematically to the determination of the final state of the system with a given initial state. The difficulty is in the great number of possible initial states because typically the number of combinations exponentially depends on the system size. Usually the signal pathways include more than 100 elements this makes the direct calculation impossible. We have developed and applied the method, which reduces drastically the amount of calculation, using the fact that there are some groups of elements, which do not interact one with another. In this case the possible therapeutic targets are those elements which always remain active independently on the initial state of the system; if they are not active, the system does not work.

The second model is an expert system, the program based on the techniques derived from the artificial intelligence methods. Expert systems contain the knowledge database and the program, which extracts new knowledge from the

database. In this case every element participating in the signal transduction pathways change the state from active to inactive through the series of inference rules stored in the knowledge database; this base was made with experimental data reported in the literature. The inference engine determines which rules must be fulfilled to arrive to some final combinations of states (active or inactive). In this case the target protein is the one, which, if absent, produces the cellular death.

As a particular example, our expert system indicates that the inhibition of 14-3-3 $\sigma$  protein may produce the cellular death. We have made experiments with cellular line MCF 7 of the breast cancer using oligonucleotides, which form the triple DNA helix or triplex with a DNA sequence to inhibit the protein production. Our results demonstrate that there exists reduction in 14-3-3 $\sigma$  protein levels and cell viability as well.

Our experience with computational systems indicate that though they demonstrate some useful dependencies and can be helpful, much more data on signal pathways and refinement of computational techniques are necessary to obtain good models.

The real behavior of cells and signal transduction pathways is more complicated than the behavior, which can be well predicted with actual computational techniques. Nevertheless, the computer models for biological systems are developing rapidly and have a great promise for the future.

## INTRODUCCION

**I.1 Modelos computacionales de sistemas biológicos.**

La célula vista como un sistema contiene un gran número de elementos que se relacionan de diferente manera. El comportamiento del sistema completo es difícil de comprender y predecir, sin embargo el progreso de las técnicas de computación en las últimas décadas hace factible el desarrollo de programas que ayuden a representar y modelar el comportamiento de complicados sistemas biológicos con cierta exactitud.

A la fecha se han empleado diferentes aproximaciones computacionales para este fin. En la tabla 1 se resumen las técnicas empleadas para modelar diferentes vías de señalización.

**Tabla 1.** Aproximaciones computacionales empleadas para modelar sistemas biológicos.

Aproximación computacional	Idea relacionada con la aproximación	Forma de procesar la información	Sistema biológico modelado	Referencias
Redes booleanas	Se consideran dos estados para cada gen, y la interacción entre ellos. El estado de cada gen depende de una función booleana particular.	Paralelo	Redes genéticas, interacción entre genes.	Kauffman, 1991; Edwards, 1995; Karp y Paley, 1994; Amas et al 2000
Sistemas expertos	Las interacciones entre los componentes de las redes de señalización se modelan usando reglas de producción.	Paralelo secuencial	Señalización intracelular	Lagunez-Otero, 1998; Takai-Igarashi y Kaminuma, 1998
Autómatas celulares	La interacción entre las proteínas se modela como matriz, donde el estado de un elemento de la matriz depende de los estados de los elementos vecinos.	Paralelo	Redes proteicas, señalización intracelular	Marijuan, 1994; Levy, 1992; Wurthner et al., 2000
Redes de Petri	La célula se considera como gráfico conectado con dos tipos de nodos. Un tipo representa elementos o moléculas de señalización y el otro representa transiciones como la activación.	Secuencial concurrente	Vías metabólicas	Holcombe, 1994 Doi et al, 2004
Redes neuronales artificiales	Las redes proteicas de señalización se ven como neuronas artificiales. Como una neurona artificial, una proteína recibe un valor de entrada, produce una salida, y un valor de activación.	Paralelo distribuido, emergente.	Redes proteicas, señalización intracelular	Bray, 1990; Bray y Lay, 1994; Bray, 1995; Pritchard y Dufton, 2000; Paton, 1993
Sistemas distribuidos (agentes)	La célula se considera como una colección de agentes que trabajan en paralelo y se comunican entre ellos a través de mensajes.	Paralelo distribuido, emergente	Señalización intracelular	Paton et al., 1995; Fisher, et al., 2000 González-Pérez et al, 2003.
Modelos continuos	El comportamiento de las proteínas depende de sus variables fisicoquímicas como la concentración y la afinidad. Los cambios de estas variables se representan con respecto al tiempo continuo, mediante ecuaciones diferenciales.	Paralelo distribuido, emergente	Activación por EGF y NGF, y su relación con MAPK cinasas. Ca <sup>2+</sup> , PKC, PKA, CaM y sistema CaMKII.	Brighman y Fell, 2000; Bhalla y Iyengar 1999

En esta tesis hemos aplicado técnicas computacionales para modelar las vías de señalización intracelular con la esperanza de obtener información que puede ser útil para la identificación de posibles blancos terapéuticos.

### **1.0 Sistemas dinámicos.**

Sistemas dinámicos en general son aquellos en los cuales el estado del sistema se determina por su estado inicial e influencias externas que lo modifican. El espacio de tiempo en que ocurra esto puede ser continuo, o una secuencia discreta de momentos. En el primer caso la descripción matemática usualmente se lleva a cabo con las ecuaciones diferenciales ordinarias o en derivadas parciales. En el segundo caso la descripción se produce por medio de "mapeo". Los parámetros que definen el estado del sistema modelado también pueden ser continuos o discretos. Por ejemplo, el movimiento de una partícula se describe con 3 coordenadas que cambian con el tiempo. En este caso tanto los parámetros, como el tiempo son continuos. Por otro lado, la computadora como sistema dinámico tiene el estado que se describe con los números 0 y 1, y la unidad del tiempo se define por el pulso de reloj, entonces es un sistema discreto. Los sistemas dinámicos se utilizan en gran escala en física, ingeniería y áreas afines.

### **2.0 Sistemas expertos.**

Los sistemas expertos tratan de reproducir el razonamiento de los expertos humanos en la resolución de problemas.

Los sistemas expertos constituyen una de las principales aplicaciones de la inteligencia artificial, el término fue utilizado por McCarthy, Minsky, Newell y Simon quienes definieron sus principales características, en la "Dartmouth computer conference" en 1956, [http://livinginternet.com/i/ii\\_ai.htm](http://livinginternet.com/i/ii_ai.htm).

La Inteligencia Artificial es el conjunto de técnicas que se aplican en el diseño de programas para computadora con la capacidad de razonar, en el sentido de inferir nueva información, y que por la dificultad del problema a resolver requieren una solución "inteligente".

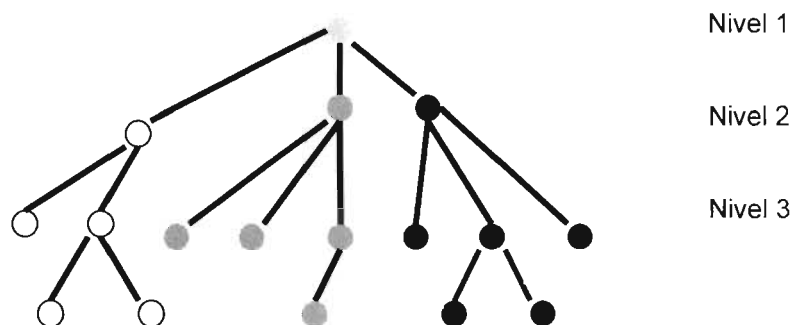
Esta cualidad de inteligencia implica una serie de elementos distintivos y fundamentales de la inteligencia, tales como la capacidad de aprendizaje, de autocorrección y de razonamiento.

#### **2.1 Elementos básicos de los sistemas expertos.**

Los elementos básicos que caracterizan a las aplicaciones de la inteligencia artificial son:

- I. Búsqueda heurística: Es un conjunto de técnicas que hacen posible el proceso de selección de la solución de un determinado problema, ya que en cada etapa de decisión la situación cambia, abriendo nuevas opciones y haciendo que el número de posibles ramas sea tan elevado que es preciso establecer ciertas reglas empíricas (heurísticas) que permitan elegir las ramas que tienen la máxima probabilidad de llevar a la solución y optimizar así el proceso de búsqueda (figura 1).

- II. Lógica y razonamiento: Es el grupo de procedimientos que permiten deducir nueva información a partir del conocimiento de un conjunto de hechos o deducir conclusiones a partir de determinadas premisas. Son los métodos y algoritmos de trabajo, y con ellos se intenta reproducir los mecanismos habituales en el razonamiento humano.
- III. Lenguajes y herramientas de programación específicos (por ejemplo los lenguajes LISP y PROLOG) para poder desarrollar las diversas aplicaciones de la inteligencia artificial.
- IV. Representación del conocimiento: cuyo propósito es organizar la información requerida por los métodos y algoritmos, ya que el comportamiento inteligente se debe no solamente a métodos adecuados de razonamiento sino también depende de un conocimiento suficiente con el cual razonar.



**Figura 1.** Niveles de toma de decisiones. Los puntos representan objetos o elementos, y las líneas la relación o flujo de información entre ellos, el color de los puntos representa diferentes opciones de elección. La información no se transmite de manera lineal del nivel 1 al nivel 3, ya que en el nivel 2 hay varias posibilidades. Es necesario establecer ciertas reglas empíricas (heurísticas) que permitan elegir la rama con mayor probabilidad de llevar a la solución.

## 2.2 Áreas de aplicación de la inteligencia artificial.

La inteligencia artificial tiene aplicación en casi todas las áreas tecnológicas, aunque tradicionalmente se han considerado cuatro:

- I. El procesamiento del lenguaje natural. Emplean cualquiera de los idiomas utilizados por los humanos para la traducción automática, la comprensión y generación de textos, etc.
- II. La visión por computadora, que incluye las técnicas de registro y de proceso de imágenes, de reconocimiento de la realidad que rodea a la computadora, por medio de la captación de las imágenes del mundo exterior.
- III. La resolución de problemas y la planificación automática. Incluye la mayor parte de los trabajos de inteligencia artificial. Según la inteligencia artificial, un problema se caracteriza por la existencia de una situación inicial, un objetivo o situación deseada y un grupo de métodos o acciones que permiten pasar de un estado a otro.
- IV. Los sistemas basados en el conocimiento. Dentro de los que se encuentran los sistemas expertos.

### 2.3. Sistemas basados en el conocimiento.

Es una aplicación de la inteligencia artificial que intenta construir programas que imiten el razonamiento de los expertos humanos, o bien que utilicen conocimientos disponibles en publicaciones especializadas; estos programas deben tener algunas propiedades asociadas a la inteligencia humana y en ellos se distinguen dos características fundamentales.

- I. Tienen separados los mecanismos de operación (algoritmos, métodos), y
- II. El conocimiento (información).

Los sistemas basados en el conocimiento se dividen en:

- I. Sistemas expertos, en los que el conocimiento requerido para resolver un problema viene dado por la experiencia adquirida por un técnico en la materia al que se denomina experto, y enriquecidos con el conocimiento públicamente disponible en la documentación técnica y científica sobre la materia.
- II. Bases de datos Inteligentes. La diferencia entre los sistemas expertos y los incluidos en este segundo grupo radica en que en estos últimos, el conocimiento suele estar basado no en la experiencia de los técnicos, sino en el conocimiento públicamente disponible en la documentación técnica y científica sobre la materia.

### 2.4 Sistemas expertos.

Un sistema experto es un programa inteligente de computadora, que usa procedimientos de conocimiento e inferencia para solucionar problemas que son suficientemente difíciles como para requerir experiencia humana en su solución. El conocimiento necesario para operar en tal nivel, añadido a los procedimientos de inferencia utilizados, puede ser considerado como un modelo de la experiencia de los mejores operadores en dicho campo.

De esta definición se obtienen las características principales de un sistema experto:

- I. Es un programa de computadora.
- II. Tiene una gran cantidad de conocimiento de hechos (público) y heurístico (experto humano) sobre un problema.
- III. Realiza un razonamiento similar al que haría el experto.
- IV. Puede operar con datos cualitativos y cuantitativos.
- V. Puede obtener conclusiones a partir de datos incompletos o inciertos.

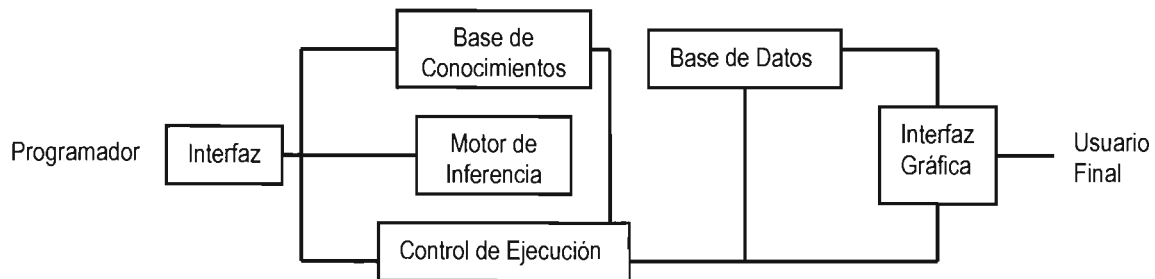
#### 2.4.1 Estructura de un sistema experto.

Un sistema experto está formado por:

- I. Base de conocimientos.
- II. El motor de inferencia.
- III. Las interfaces gráficas con el usuario del sistema experto y con el programador.
- IV. El control de ejecución del sistema.



El siguiente diagrama representa las distintas partes que forman el sistema experto y las relaciones que unen a estos componentes.



**Figura 2.** Diagrama de los componentes de un sistema experto. El programador construye la base de conocimientos, motor de inferencia y el control de ejecución, del sistema experto. El usuario final a través de la interfaz gráfica, alimenta a la base de datos, y controla la ejecución.

### I. La base de conocimientos.

Está constituida por todo el conocimiento e información que el programador haya podido reunir, a partir no solamente del experto, sobre el problema a tratar, sino también de la información científica publicada sobre el área. Todos los demás bloques estarán condicionados a una serie de decisiones tomadas al realizar la base de conocimientos. De ahí la importancia de su desarrollo dentro del tiempo dedicado al desarrollo del sistema experto.

La base de conocimientos contiene dos tipos de conocimiento acerca del problema que trata: Los hechos y las relaciones heurísticas entre los hechos.

I.I Los hechos son datos, situaciones presentes o relaciones admitidas entre las distintas variables del sistema. Forman la realidad que rodea al problema que trata el sistema experto. El programador, además de plantearlos, debe organizar dichos hechos de forma eficaz para que puedan ser manejados y comprendidos por el sistema experto; por otra parte, debe también conseguir que la estructura de los hechos tenga una gran capacidad de crecimiento y modificación.

I.II Las relaciones heurísticas representan el conocimiento de las relaciones causales entre los distintos tipos de hechos: son las relaciones admitidas entre los valores de las distintas variables, las consecuencias que tendrán las modificaciones de ciertos elementos respecto a otros, etc.

La forma más utilizada para la representación del conocimiento en la base de conocimiento, es la de las reglas de producción, conocidas también como reglas de situación-acción o reglas SI-ENTONCES-ACCION (IF-THEN-AND DO)

Una regla está formada por un conjunto de condiciones iniciales o antecedentes que deben cumplirse para que la regla pueda aplicarse. La aplicación de la regla da como resultado un conjunto de acciones sobre la base de conocimientos.

La base debe poseer un conocimiento muy amplio sobre la materia, pero no redundante, y estructurado adecuadamente para que sea rápidamente accesible, evitando así tiempos de búsqueda largos.

Hay que considerar como parte integrante de la base de conocimientos a las bases de datos. Estas bases son módulos en los que se almacena toda la información general que no depende explícitamente del problema tratado.

II. El motor de inferencia.

Constituye con la base de conocimiento el otro bloque fundamental de un sistema experto. Su misión principal es la de utilizar toda la información disponible en la base de conocimiento para llegar a la solución buscada; los motores de inferencia utilizan mecanismos de razonamiento lo más parecido posible al de los expertos humanos.

El motor de inferencia realiza dos tipos de actividades:

II.I El examen de los datos existentes y de las reglas, añadiendo hechos a la base de conocimiento, en caso necesario.

II.II La decisión sobre el orden en el que se ejecutan las reglas cuando hay varias disponibles (prioridad de reglas) .

Cuando el motor de inferencia examina una regla, comprueba si sus condiciones iniciales se cumplen. En ese caso, ejecuta o afirma el conjunto de acciones de dicha regla. Si hay varias reglas ejecutables en el mismo instante (bajo una serie de situaciones y condiciones), el motor de inferencia debe elegir a la que tiene mayor prioridad atendiendo a diversos criterios: aquella cuyos antecedentes hayan sido activados recientemente, la de un coeficiente de prioridad más alto, la más específica, etc. La ejecución de una regla puede aportar nuevos datos a la base de conocimiento, y puede hacer que haya más reglas ejecutables. El proceso se repite hasta que no queda ninguna regla por ejecutar con las condiciones dadas normalmente, el motor de inferencia puede presentar dos tipos de mecanismos de razonamiento:

A. Razonamiento hacia atrás o deductivo (Backward Reasoning). El motor de inferencia debe evaluar una hipótesis, por ello intenta aplicar todas aquellas reglas que apunten a dicha hipótesis. A su vez, para aplicar estas reglas, debe evaluar sus respectivos conjuntos de condiciones, entre los que puede haber nuevas hipótesis que hagan que el proceso se repita.

B. Razonamiento hacia adelante o inductivo (Forward Reasoning). A partir de los valores de las propiedades de objetos concretos (datos), el Motor de inferencia evalúa aquellas reglas que contienen dichos datos.

III. Las interfaces del sistema experto.

Estos bloques se encargan de traducir las representaciones de conocimiento exterior a una forma interna que pueda ser comprendida por el código del sistema.

Se pueden clasificar en cuatro grupos:

III.I Interfaz gráfica con el usuario.

III.II Interfaz con el programador.

III.III Dos interfaces con las bases de datos, una entre el sistema experto y la base de datos, y otra de conexión de éstas con el mundo exterior.

De estas cuatro, la más importante es la interfaz con el usuario. Permite la comunicación bidireccional entre éste y el sistema experto, es decir, es capaz de enviar información en los dos sentidos, de tal forma que el usuario puede tener información sobre lo que está realizando el sistema experto, y éste a su vez,

puede obtener más datos acerca del problema gracias al usuario. Se debe procurar que esta interfaz sea lo más amigable para el usuario, utilizando todos los medios posibles (gráficos interactivos, menús, pantallas de texto, histogramas, etc) .

La interfaz del programador facilita la labor de construcción del sistema experto. Suele disponer de todas las herramientas de la Interfaces de usuario y tiene además editores de gran potencia, así como mecanismos de seguimiento del razonamiento y de detección y corrección de errores. Una buena Interfaz del programador supone una reducción importante en el tiempo de desarrollo del sistema experto.

#### **2.4.2 Desarrollo de un sistema experto.**

La herramienta de programación debe ser elegida después de un riguroso estudio de las necesidades y tareas que planteará el desarrollo del sistema experto concreto. Se deben considerar.

- I. La complejidad del problema a tratar.
- II. Las personas que intervienen en el proceso de desarrollo: los expertos, el programador, el usuario final. La adquisición del conocimiento por parte del ingeniero a partir de los expertos puede ser muy compleja y problemática.
- III. La flexibilidad del propio sistema, es decir, la capacidad del sistema experto de trabajar con un universo de respuestas (entradas) amplio y a la vez ambiguo, donde todas las situaciones posibles tengan cabida. Un sistema experto debe ser muy generoso con los errores del usuario, si se considera flexible.

En esta tesis desarrollamos nuestro sistema experto en una plataforma comercial Xpert rule KBS (Knowledge-Based System), en la que creamos la base de conocimientos a partir de información recabada por nosotros.

#### **2.5 Xpert rule KBS.**

Para la realización de este modelo, se ha utilizado el programa Xpert rule KBS (versión 3.6), de la compañía Attar Software Ltd.

Es una aplicación elaborada en lenguaje C, y existen diferentes versiones para trabajar en diversos sistemas operativos, tales como DOS, UNIX, MacOS, VM y VMS, etc., y en sus respectivos entornos hardware (PC, AT, Macintosh, VAXstations, Mainframes, etc.) .

La aplicación Xpert rule KBS se ejecuta bajo los entornos de ventanas de cada sistema operativo, ya sea Microsoft Windows, X Windows 11 o Presentation Manager. La Interfaz de usuario es totalmente gráfica, facilitándose el acceso a cualquier tipo de información relacionada con el sistema experto y la labor de desarrollo y creación del sistema experto.

Xpert rule KBS y sus programas pueden llamar y ser llamados por subrutinas externas, lo que permite la creación de arquitecturas complejas (sistemas expertos embebidos, sistemas expertos integrados), y la mejor resolución de los problemas.

### 2.5.1 La plataforma Xpert rule KBS.

En la plataforma Xpert rule KBS existen tres programas distintos que son complementarios:

- I. Xpert rule KBS, orientado a la creación de la base de conocimiento del sistema experto, y que incorpora el motor de inferencia que será utilizado en los sistemas expertos desarrollados.
- II. Form Painter, programa que se ejecuta bajo Windows, y que sirve para desarrollar las Interfaces de ventanas para el usuario final.
- III. Xpert rule KBS Forms, programa que desarrolla el proceso de razonamiento del sistema experto para el usuario final. Ofrece la posibilidad de automatizar y dar coherencia a las distintas sesiones de trabajo del usuario con el sistema experto, limitando el uso de información relacionada con la base de conocimiento, evitando así posibles modificaciones no intencionales.

#### I. Xpert rule KBS

Xpert rule KBS se ocupa principalmente de crear la Base de Conocimiento del sistema experto. La base tiene las siguientes características. Presenta las dos representaciones de conocimientos más habituales:

- I.I Reglas (dimensión del razonamiento) y
- I.II Objetos (dimensión de la realidad).

#### I.I La dimensión del razonamiento

La dimensión del razonamiento está formada por las reglas. Estas tienen una estructura bien definida y están formadas por:

1. El conjunto de condiciones que se debe cumplir para que se aplique una regla, también llamado parte izquierda de la regla (LHS: Left Hand Side). Estas condiciones pueden ser de dos tipos:
  - 1.1 Las que comprueban, verifican, dan valor; en definitiva, trabajan, con objetos, propiedades e hipótesis de la base, se les llama datos.
  - 1.2 Las que ejecutan diversas tareas no relacionadas directamente con datos de la base, como ejecutar rutinas externas, mostrar ficheros y ventanas de Interfaz.
  - 1.3 Las condiciones de la LHS de una regla pueden tomar los siguientes valores:
    - 1.3.1 Desconocido (Unknown), cuando dicha condición no ha sido evaluada por el Motor de Inferencia.
    - 1.3.2 Verdadero (TRUE), cuando la condición ha sido evaluada y verificado su cumplimiento.
    - 1.3.3 Falso (FALSE) , cuando la condición ha sido evaluada y ha sido verificado que no se cumple.
    - 1.3.4 No conocido (Notknown) , cuando se ha evaluado la condición, pero el sistema experto (y el usuario) no dispone de información suficiente sobre los datos implicados para dar los valores falso o verdadero.

Este conjunto de condiciones de la LHS están unidas por operaciones lógicas de tipo AND, es decir, se deben cumplir todas las condiciones de una LHS, para que la regla correspondiente se pueda aplicar.

2. La parte derecha de la regla (RHS: Right Hand Side); está formada por dos partes bien diferenciadas:

2.1.1 La hipótesis a la que apunta cada regla.

2.1.2 El conjunto de acciones que se ejecutan si se aplica la regla.

3. La aplicación de las reglas en el proceso de inferencia puede darse en los dos sentidos que permite el razonamiento hacia adelante y atrás (Forward y Backward Reasoning).

4. Las relaciones y dependencias entre las reglas se establecen por la utilización de datos (hipótesis, propiedades, etc) comunes (llamados eslabones o links). Este mecanismo se implementa de dos formas diferentes (strong links o enlaces fuertes):

4.1 Enlaces fuertes basados en condiciones (Condition based strong links), en el que el dato que se comparte está en la LHS de las dos reglas.

4.2 Enlaces fuertes basados en acciones (Action based strong links), en el que el dato compartido ocupa la RHS y la LHS de dos reglas (el caso más frecuente es la utilización de la hipótesis de una regla en un condición de la LHS de otra regla).

5. Otra forma de relación entre reglas es el enlace débil (weak link). En este caso, la investigación de una regla puede hacer necesario que se explore otra, aun cuando no se comparta ningún dato con ella. Este mecanismo se establece cuando no existe el conocimiento suficiente para establecer una regla de manera determinante, pero se da cierta correlación en la que la investigación de una regla bajo ciertas condiciones puede ser interesante. A este tipo de relación también se le denomina contexto (una regla está en el contexto de otra cuando la primera debe ser investigada si se dispara la segunda. El proceso inverso de inferencia no se da).

I.II Objetos. La representación de la realidad.

El mundo de los objetos representa la realidad a la que se le aplican los mecanismos de razonamiento. La unidad elemental de esta dimensión es el objeto. Un objeto está constituido principalmente por:

1. Su nombre.
2. Sus propiedades o características.
3. Sus relaciones con otros objetos de la base de conocimiento.
4. Su influencia en el curso del razonamiento en el caso de que algunas de sus propiedades sean consultadas o modificadas.

La dimensión de la representación de la realidad o de los objetos tiene jerarquías. Existen varios niveles:

1. Objetos, que constituyen la unidad elemental.

2. Subobjetos. Considerados separadamente, son también objetos, pero a su vez son componentes de otro objeto; los Subobjetos no heredan necesariamente todas las propiedades de objeto progenitor.
3. Clases. Son conjuntos de objetos que tienen algunas propiedades o aspectos comunes. Un objeto perteneciente a una clase es un ejemplo o un caso de dicha clase.
4. Subclases. Son clases que comparten propiedades con otra clase de orden superior.

Los objetos se pueden clasificar al momento de su creación:

- A. Objetos estáticos, que son aquellos objetos que son creados al mismo tiempo que se diseña la base de conocimientos.
- B. Objetos dinámicos, que son los objetos diseñados junto con la base de conocimientos, pero creados durante el periodo de inferencia por el sistema experto es decir, su existencia se limita al tiempo del proceso de razonamiento del sistema experto. Se consigue de esta forma ahorrar memoria ante un número indefinido de objetos que pueden ser útiles sólo en algunas ocasiones:

### **2.5.2 Herramientas de Xpert rule KBS.**

Para desarrollar la bases de conocimiento, Xpert rule KBS cuenta con los mecanismos de creación de las partes fundamentales de un sistema experto:

- I. El componente de adquisición del conocimiento.
- II. El motor de inferencia (que está incorporado).
- III. El componente explicativo o interfaz con el usuario.

I. El componente de adquisición del conocimiento viene dado por los editores y las listas de elementos que forman la base de conocimiento. Hay seis editores: de reglas, de propiedades, de clases, de objetos, de "meta-slots" y de contexto. Presentan como principales características la comprobación automática de la sintaxis, y la existencia de un conjunto de funciones que permiten la optimización de las sesiones de trabajo (creación automática de objetos, clases y propiedades, que están implícitos en las condiciones de una regla). También se muestran, seis agendas con listas de hipótesis, reglas, datos, objetos, propiedades y clases.

II. El motor de inferencia de Xpert rule KBS realiza el razonamiento sobre la base de conocimiento que se encuentra cargada en la memoria.

Para iniciar el proceso de inferencia, se puede actuar de dos maneras:

- A. Sugerir la investigación de una o varias hipótesis (Suggest).
- B. Dar valor a unas propiedades determinadas de unos objetos (Volunteer), para que a partir de esta información el motor de Inferencia pueda trabajar sobre la base de conocimiento.

Otra de las características del motor de inferencia de Xpert rule KBS es la posibilidad de efectuar razonamientos hacia adelante y hacia atrás, de manera indistinta, según cuál sea la situación.

III. La interfaz con el usuario de Xpert rule KBS, incluye una serie de herramientas que facilitan la creación de la base de conocimiento.

III.I Los espacios de navegación para la dimensión del razonamiento y de la realidad (Rule Network y Object Network). En ellos se representan de manera gráfica las relaciones existentes entre las reglas (u objetos y clases) de una determinada base, los enlaces fuertes y los débiles, así como el estado de evaluación de cada una de las condiciones e hipótesis que componen las reglas representadas, durante el proceso de inferencia. Se consigue de esta forma dar coherencia a la base de conocimiento, estableciendo los caminos posibles de razonamiento ante un determinado problema.

### **2.5.3 Form painter.**

Este programa está destinado a la creación de la interfaz del usuario final.

Las ventanas de interfaz permiten la comunicación entre el sistema experto y el usuario, que a través de ellas podrá responder a las cuestiones que le formule el sistema experto, introduciendo en el sistema los valores de los datos que mediante el razonamiento no han podido ser inferidos directamente.

Cada pantalla de la interfaz contiene varios de los siguientes elementos:

- I. Recuadro de texto. Presenta mensajes en pantalla, sin recibir información por el teclado.
  - II. Campo de entrada de un dato. Recibe datos por el teclado.
  - III. Lista de selección única. Se puede elegir una opción entre las que aparecen en una lista. Puede combinarse con el campo de entrada de un dato.
  - IV. Lista de selección múltiple. En este caso, se puede elegir más de una opción de la lista establecida de posibles respuestas. Puede combinarse con el campo de entrada de un dato.
1. Tabla de selección. En ella se ofrecen distintos objetos con sus respectivas propiedades entre los que el usuario puede elegir los que desee.
  2. Tabla de entrada de datos. El usuario puede introducir valores de distintos objetos y sus propiedades en una tabla para estructurar la información.

Estas ventanas pueden tener otras opciones como texto constante, un color de fondo, determinados tipos de recuadro de borde, teclas de función. etc. Con ello se debe conseguir que el proceso de comunicación establecido sea lo más fluido y fácil posible. El usuario final podrá contestar a las preguntas que el sistema experto le plantea con las teclas del cursor y por medio del teclado.

### **2.5.4 Xpert rule KBS forms.**

Es el único programa de la aplicación Xpert rule KBS que se ejecuta en pantalla de tipo texto. Su tarea es hacer posible el funcionamiento de un sistema experto desarrollado con Xpert rule KBS en cualquier computadora personal con pantalla de texto. Además permite al autor de la base de conocimientos añadir una interfaz con el sistema experto de diseño propio, de forma que impida el acceso

del usuario final a la información de la base de conocimientos, y evitar así una modificación fortuita tanto de la base como del motor de inferencia.

Las posibilidades que ofrece Xpert rule KBS Forms son las mismas que las de Xpert rule KBS, respecto a los mecanismos de inferencia y al componente explicativo (aunque no a los editores de la Base de Conocimiento); por ello, también le es aplicable todo lo dicho en el apartado anterior sobre estas aplicaciones.



## **I.II Señalización intracelular.**

En las vías de señalización intracelular están involucrados un gran número de elementos que se interrelacionan y dan lugar a diferentes respuestas celulares ante un mismo estímulo.

Durante la señalización intracelular se debe tomar en cuenta la presencia inicial al estímulo de un gran número de elementos, las interacciones posteriores que estos efectúen llevarán a una respuesta celular final. Los resultados experimentales reportados en la literatura permiten una aproximación real a la respuesta celular final, y a la inversa a partir de una respuesta celular final se pueden aproximar los elementos iniciales al estímulo que genera esa respuesta. Las vías seleccionadas para el modelaje son las vías de señalización en las cuales participa la proteína Ras. Se eligió a esta proteína, ya que es un punto de convergencia para varias vías de señalización, está implicada en el control de la proliferación celular y en el control del ciclo celular. A continuación se describe la señalización intracelular y las vías de señalización seleccionadas para modelar.

### **1. Señalización.**

Cada célula en un organismo pluricelular es programada durante el desarrollo para responder específicamente a un grupo de señales químicas o mensajes intra y extracelulares. Estas señales químicas actúan en varias combinaciones para regular mediante la señalización intracelular el desarrollo de la célula y determinar cuando vive o muere, cuando prolifera o se mantiene en un estado de quiescencia (no se divide). La mayoría de estas señales químicas son mediadas por el sistema paracrino, este sistema permite la comunicación entre células relativamente cercanas, sin que para ello exista una estructura especializada (como en el caso de la sinapsis, que se explica más adelante). Generalmente estas señales químicas actúan únicamente en células vecinas a las células que las producen. Para evitar que las señales salgan de la zona en la que se producen, existen proteínas que las atrapan, ya sea, para destruirlas o para almacenarlas.

El control central de las señales paracrinas está dado por las señales endocrinas u hormonales. Las células de las glándulas de secreción interna, como la hipófisis, la tiroides, los islotes de páncreas, las suprarrenales, los ovarios y los testículos, producen diferentes hormonas y las vierten al torrente circulatorio. Una vez en la sangre, las hormonas circulan por todo el organismo e interactúan con algunas células que son "receptoras" para una hormona determinada, a las cuales se les llama células blanco. La señal química (en este caso la hormona) es selectiva. La selectividad se debe a los receptores específicos (Raff, 1992).

En la neurotransmisión o comunicación química a través de las células nerviosas, las neuronas con sus largos axones están muy cercanas a las células con las que se comunican. Entre la membrana externa de la neurona y la membrana externa de la célula, se encuentra una estructura especializada que se llama sinapsis, en ella hay un espacio (el espacio sináptico). El flujo o sentido de la información es unidireccional y va de la neurona, o célula presináptica, a la célula receptora o postsináptica. En esta forma de comunicación, la célula presináptica vierte su "mensaje" o neurotransmisor al espacio sináptico, y éste

viaja e interacciona con la célula postsináptica, la cual lo recibe y responde (Changeux, 1993).

La señalización celular requiere de una señal extracelular y de un grupo complementario de proteínas receptoras en cada célula que sea capaz de unirse y responder a la señal de manera específica.

El receptor tiene dos características fundamentales:

1. Reconoce a la señal química para interactuar con ella.
2. Activa la secuencia de eventos que conducen a la respuesta celular.

Las señales químicas extracelulares pueden tener una naturaleza química muy variada; sin embargo, se pueden agrupar en cuatro clases fundamentales: los lípidos (entre los que se encuentran los esteroides y las prostaglandinas), los de naturaleza peptídica (como algunos factores de crecimiento), las aminas y los iones.

Las moléculas de señalización hidrofóbicas, que incluyen hormonas esteroides y tiroideas, y los retinoides, se difunden a través de la membrana plasmática de las células blanco y activan receptores intracelulares, los cuales regulan directamente la transcripción de genes específicos. Algunos gases disueltos como el óxido nítrico y el monóxido de carbono, actúan como mediadores locales por difusión a través de la membrana plasmática de la célula blanco, activando a una enzima intracelular, usualmente la guanilato ciclasa, la cual produce GMP cíclico en la célula blanco (GMP 5'-monofosfato de guanosina). El GMP cíclico es un segundo mensajero. La mayoría de los receptores asociados a proteínas G activan una cadena de eventos que alteran la concentración de una o más moléculas de señalización intracelular. A estas moléculas se les conoce como mediadores intracelulares o segundos mensajeros, los segundos mensajeros transmiten la señal, modificando la actividad de las proteínas intracelulares. Las proteínas G son una gran familia de proteínas que unen 5'-trifosfato de guanosina (GTP), (Jans, 1994).

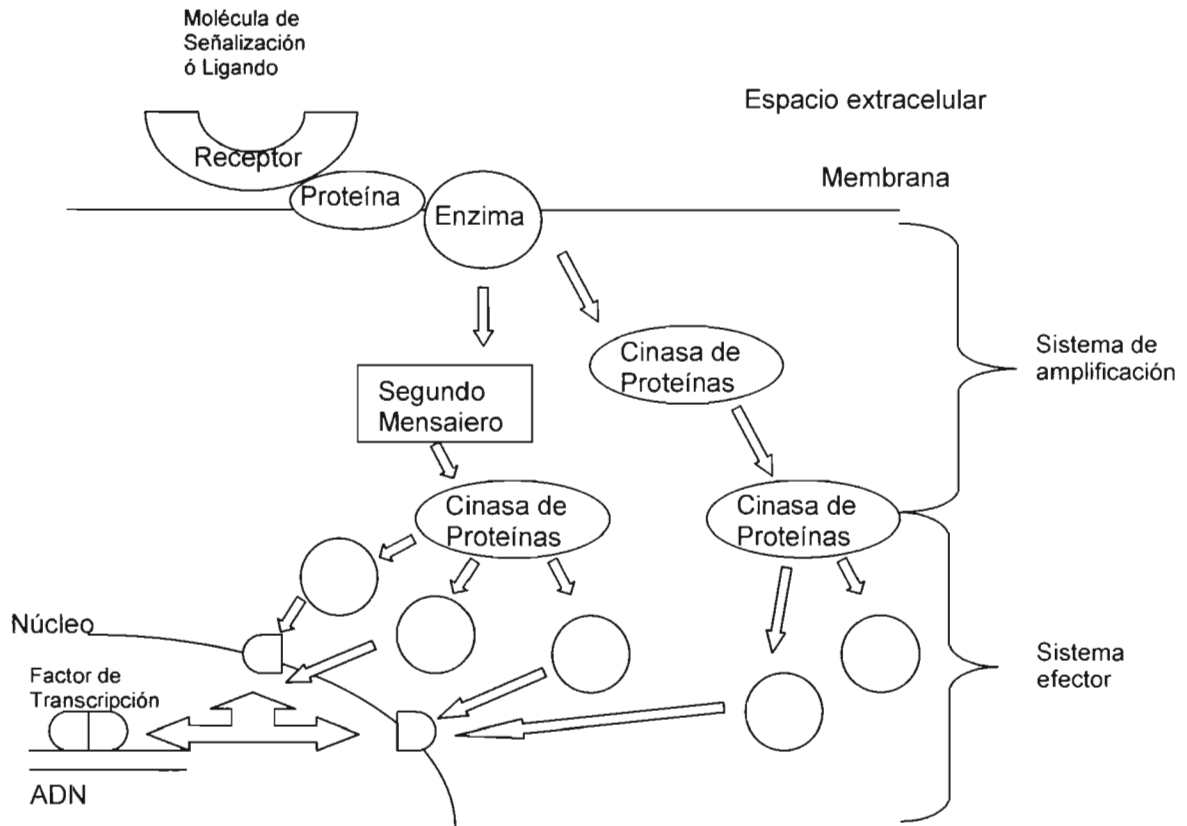
Pero la mayoría de las moléculas de señalización intracelular son hidrofílicas y capaces de activar receptores únicamente en la membrana plasmática de células blanco. Estos receptores actúan transduciendo la señal, es decir convierten el evento de unión extracelular en señales intracelulares que alteran el desarrollo de la célula blanco. Son tres las familias de receptores de la superficie celular, cada una de las cuales transduce señales extracelulares por una vía diferente. Los receptores ligados a canales iónicos abren o cierran los canales iónicos en respuesta a la unión de un neurotransmisor. Los receptores ligados a proteínas G activan o inactivan de manera indirecta a enzimas ancladas en la membrana plasmática o a canales iónicos mediante proteínas G heterotriméricas (Milligan, 1993). Receptores ligados a enzimas o con actividad enzimática usualmente son cinasas que fosforilan proteínas específicas en las células blanco. A través de cascadas de señalización mediadas por fosforilaciones de proteínas específicas se altera el patrón de expresión genética de la célula y como consecuencia su desarrollo. Las diferentes cascadas de señalización se entrecruzan por lo que la célula es capaz de integrar las múltiples señales que recibe del exterior (Nishizuka, 1992; Kato et al. 2000).

Para que la señal extracelular se transmita desde el exterior a través de su receptor específico hasta los factores de transcripción, que activarán la transcripción de las proteínas involucradas en la respuesta celular, se requiere principalmente la “comunicación” entre proteínas mediante fosforilación o desfosforilación de residuos de tirosina, treonina y serina; estos aminoácidos fosforilados son reconocidos por otras proteínas mediante dominios específicos como SH2 (*Src homology region 2*). Otros dominios proteicos que participan en la transducción de señales son dominios ricos en prolina y los dominios SH3 (*Src homology region 3*). Los dominios proteicos permiten la interacción específica entre proteínas. En la tabla 2 se encuentran ejemplos de las secuencias de aminoácidos de los dominios proteicos que interactúan entre sí, para que se efectúe la transducción de señales.

**Tabla 2.** Dominios proteicos importantes en la comunicación entre proteínas involucradas en transducción de señales y ejemplos de las secuencias de aminoácidos reconocidas por cinasas y fosfatasa de residuos de serina, treonina y tirosina.

Proteína	Ejemplo	Dominio	Dominios importantes	Interactúa con	Proteína que la contiene
Cinasas de Ser y Thr	PAK $\gamma$		C Terminal (36 Kda)	KKRKS <sup>619</sup> GL	cAbl
Cinasas de Ser	Raf1		CR3	GLS <sup>338</sup> YLHENQ	MEK
Cinasas de Tyr	Abl1		GMEYLESKNFVHRDLAARNCLVG Dominio rico en prolina y dominio catalítico 362-625	Y <sup>221</sup> AGP	Crk
Fosfatasa Ser y Thr	PP2A		LKVRYPERITILRGNHESRQITQVY GFYDECLRKYGNANVWKYFTDLF DYLPLT C Terminal (36 Kda) Pph21p/Pph22p	T S <sup>378</sup> RH KKL	p53
Fosfatasa de Tyr	SHP-1		RYKNIVEKGRNKQMVRAQRSGMV Q Dominio catalítico 271-514	EY <sup>394</sup> TARE	Lck
Fosfatasa de Tyr, Thr y Ser	Cdc25A		GKRVIVVFHCEFSSEGRPMCRY VRERDRLGNEYPKLHYPELYVLKG GYEFFMKCQSYC DominioN25A (27Kda) C terminal	G T <sup>14</sup> YGVV	Cdc2
Adaptadoras que unen Tyr fosforiladas	Grb2	SH2	HRWFELTKRTFSYFDVENVERRR ERGRIHLKGVRLVEEATVSGEGG DPFAPDGYPFQVGYCEISASANSH QLENGNGGGSG	QRY <sup>1068</sup> SSDPT EY <sup>1086</sup> INQSV	EGFR
Adaptadoras que unen Ser fosforiladas	14-3-3	LR-NES	KGDYQKALLYLCGGDD	RSTS <sup>259</sup> TP RSAS <sup>621</sup> EP	Raf1
Que unen dominios ricos en Prolina	Crk	SH3	MMGTKHRNSHVNGSIKSSSSLRS SSKSFQAKMDLMSERLYDVVKSG SMVKRAQNKK	P <sup>299</sup> PALPPK	PI3K (p85)
Que se unen a ADN	AP1	Dominio que se une a ADN	RACCACGTGCTY	ATGACTCAT	c-jun

En la figura 3, se presenta un esquema general de la transducción de una señal desde el exterior hasta el núcleo.



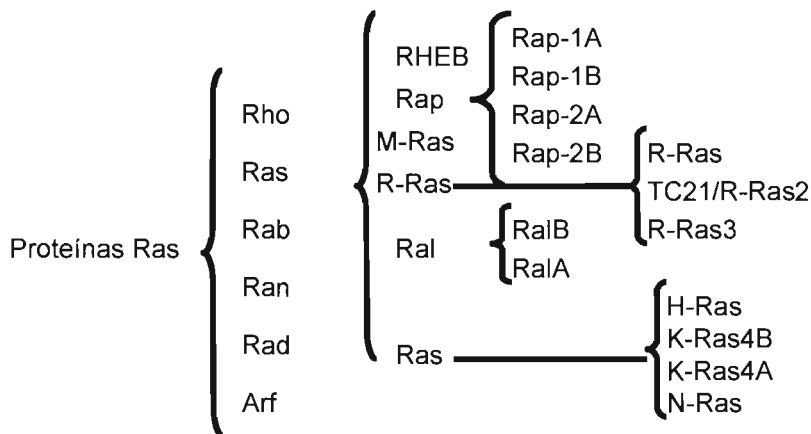
**Figura 3.** En esta figura se ejemplifican los procesos generales que se observan durante la transducción de señales iniciadas por receptores de membrana. La molécula señal se une a su receptor específico localizado en la cara externa de la membrana plasmática. Al unirse a su ligando el receptor modifica su conformación, por lo que proteínas localizadas en el citoplasma próximas a la membrana plasmática o en el espacio yuxtamembranal interactúan con él. Esta interacción permite el reclutamiento de otras proteínas y la activación de las mismas. Estas proteínas pueden generar segundos mensajeros o activar mediante fosforilación a otras proteínas. Tras una cascada de activaciones finalmente se llega a la activación de factores de transcripción en el citoplasma. Los factores de transcripción son translocados al núcleo. En el núcleo se unen a la región promotora de genes que codifican para proteínas que participan en la función biológica específica que da respuesta a la señal recibida.

El trabajo de esta tesis está dedicado a la creación de modelos computacionales de las vías de señalización intracelular para la búsqueda de blancos terapéuticos, potencialmente útiles en el control de enfermedades. Seleccionamos las vías en las que participa la proteína Ras por ser un punto de convergencia de varias vías involucradas en crecimiento y diferenciación celular. Estas funciones celulares cuando no son correctamente controladas pueden dar lugar a enfermedades, como por ejemplo el cáncer. A continuación se describe la proteína Ras y sus funciones.

## 2. Proteínas Ras.

Las proteínas Ras participan en la señalización intracelular mediada por receptores asociados a enzimas, en la señalización intracelular mediada por receptores ligados a proteínas G y en el control del ciclo celular.

Las proteínas Ras se localizan en la cara interna de la membrana plasmática, forman parte de la familia de proteínas G, por tener actividad intrínseca de GTPasa (hidrolizan nucleótidos de guanina), pero a diferencia de otras proteínas G que son triméricas, son monoméricas, por lo que forman parte de la subfamilia de proteínas G, las proteínas G pequeñas.



**Figura 4.** La familia de proteínas Ras, está formada por un grupo de proteínas de entre 20 y 25 KDa. En mamíferos, la familia de proteínas G pequeñas contiene a más de 65 miembros. Basados en su función y estructura, esta familia se ha subdividido en 6 familias: Ras, Rho, Rab, Ran, Rad y Arf. La familia Ras se subdivide en base a su homología en el dominio efector en 6 subfamilias: RAP (Rap-1A, Rap-1B, Rap-2A y Rap-2B), R-RAS (R-Ras, TC21/R-Ras2 y R-Ras3), RAL (RalA y RalB), RHEB, M-RAS, y la subfamilia RAS que incluye a H-Ras, K-Ras4A, K-Ras4B y N-Ras.

En este trabajo, al referirnos a las proteínas Ras estaremos considerando únicamente a la subfamilia Ras o p21Ras (H-Ras, K-Ras4A, K-Ras4B y N-Ras), a los demás miembros de la familia se le referirá con su nombre directamente.

En el humano se conocen tres formas del proto-oncogen *ras*: *Hras*, *Nras* y *Kras*, cada uno codifica para las proteínas de 188 amino ácidos H-Ras, N-Ras y K-Ras4A respectivamente; *Kras* produce una de 189 amino ácidos, K-Ras4B, (Lacal y McCormick, 1993; Rodríguez Viciano, 1994 y 1996; Crespo y León, 2000; McMahon y Woods, 2001).

Los productos proteicos de los oncogenes juegan un papel central en el desarrollo de la mayoría, si no de todos los tipos de cáncer. Los proto-oncogenes celulares son genes que bajo ciertas condiciones son capaces de inducir transformación neoplásica en las células, la cual puede llevar a las células a la muerte o derivar en cáncer.

La modificación por mutación o cambios en el control de la expresión de genes normales o proto-oncogenes da lugar a un oncogen. Además de los oncogenes en el genoma celular, algunos virus pueden llevar oncogenes como parte de su genoma. Los oncogenes de virus de ARN tumorigénicos son derivados de proto-oncogenes celulares, estos virus pueden transformar rápidamente células en cultivo e inducen tumores en animales. Algunos virus de ADN también son

tumorigénicos pero no existe evidencia que indique que expresen oncogenes derivados de progenitores celulares.

Algunas células son más susceptibles a ser transformadas, esto depende de su capacidad para reparar daño a ADN, del potencial mutagénico de sustancias con las que estén en contacto y de la combinación de proteínas producto de oncogenes que expresen.

A las células neoplásicas en cultivo *in vitro* se les llama células transformadas. Las células transformadas difieren de las normales en cultivo por ser “inmortales”, ya que las células normales tienen un crecimiento limitado y un número finito de divisiones celulares. Las células transformadas tienen una menor dependencia a los factores de crecimiento extracelulares con respecto a las células normales.

Se han reportado más de 100 proto-oncogenes cuyos productos proteicos participan en las vías de señalización intracelular a diferentes niveles, por ejemplo hay receptores, cinasas de tirosina, proteínas que unen nucleótidos de guanina, factores de transcripción, etc. En la tabla 3 se encuentra una lista de productos proteicos de proto-oncogenes involucrados en las vías de señalización intracelular, con su función y ejemplos de cáncer en el que se han encontrado con alguna mutación.

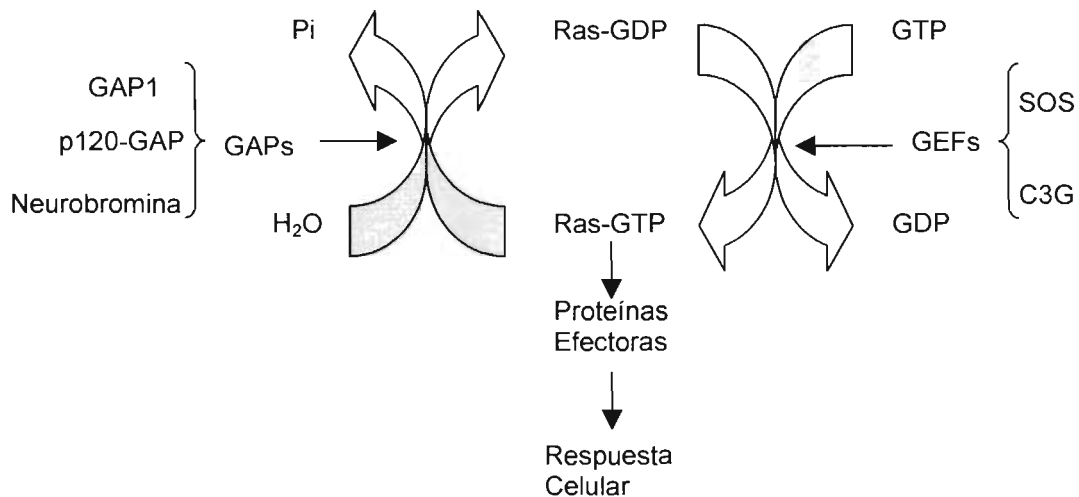
Durante la “evolución” de células normales a células transformadas ocurren muchos cambios genéticos. Las células transformadas pierden fidelidad durante la replicación, reparación y segregación de su genoma.

**Tabla 3.** Productos proteicos de proto-oncogenes involucrados en las vías de señalización intracelular, su función y ejemplos de cáncer en el que se han encontrado con alguna mutación. Los proto-oncogenes codifican para proteínas que participan en la transducción de señales a diferentes niveles. Como se ve en esta tabla hay ligandos ó moléculas señal, receptores, cinasas, proteínas G y factores de transcripción.

Proteína	Función	Cáncer
PDGFB/Sis	Ligando	Fibrosarcoma (Koveker, 2000)
Int2	Ligando	Sarcoma de Kaposy (Kiuru-Kuhlefeldt et al., 2000)
HSTF1/HST1	Ligando	Tumor en células germinales (Yoshida et al., 1988)
EPH	Receptor con actividad de cinasas de Tirosina	Tumor en células germinales (Koveker, 2000)
EGFR/ERB	Receptor con actividad de cinasas de Tirosina	Adenocarcinomas y adenomas (Porebska et al., 2000)
Fms	Receptor con actividad de cinasas de Tirosina	Cáncer de mama (Hung et al., 2001)
Kit	Receptor con actividad de cinasas de Tirosina	Meduloblastoma (Chilton-Macneill et al. 2004)
Met	Receptor con actividad de cinasas de Tirosina	Cáncer de mama (Shaharabany et al., 2001)
HER2/NEU	Receptor con actividad de cinasas de Tirosina	Carcinoma de mama (Lee et al, 2001)
TRK	Receptor con actividad de cinasas de Tirosina	Adenocarcinoma de páncreas (Sakamoto et al., 2001)
Abl	Cinasa de Tirosina	Leucemia mieloide crónica (Gorre et al., 2001)
Src	Cinasa de Tirosina asociadas a membrana	Cáncer de mama (García et al., 2001)
Fgr	Cinasa de Tirosina asociadas a membrana	Cáncer de pulmón (Zarn et al., 1996)
Fyn	Cinasa de Tirosina asociadas a membrana	Neuroblastoma (Davy et al., 2000)
Hck	Cinasa de Tirosina asociadas a membrana	Cáncer en células B (Taguchi et al., 2000)
Lck	Cinasa de Tirosina asociadas a membrana	Relacionado con metástasis (Harashima et al, 2001)
Yes	Cinasa de Tirosina asociadas a membrana	Cáncer de pulmón (Jeremic, 2001; Hansen et al, 2001)
Más	Receptor sin actividad de cinasa	Cáncer en fibroblastos (Zohn et al, 1998)
H-Ras	Proteína G asociada a membrana	Células tiroideas (Cobellis et al., 2001)
K-Ras2	Proteína G asociada a membrana	Cáncer colorectal (Etoh et al., 2001)
N-Ras	Proteína G asociada a membrana	Mesotelioma (Papp et al., 2001)
Gsp	Proteína G asociada a membrana	Adenomas pituitarios (Shi et al., 1998)
Gip2	Proteína G asociada a membrana	Cáncer de ovario (Shen et al., 1996)
Bcr	Cinasa de Serina	Leucemia mieloide (Gorre et al., 2001)
Mos	Cinasa de Serina	Cáncer de pulmón (Gorgoulis et al, 2001)
Pim1	Cinasa de Serina	Linfoma (Feldman et al., 1997)
Raf/Mil	Cinasa de Serina	Cáncer de mama (Pratt et al., 1998)
Crk	Regulador citoplásmico	Cáncer de células epiteliales (Donaldson et al, 2000)
Fos	Factor de transcripción	Cáncer en fibroblastos (Van Dam y Castellazzi, 2001)
Jun	Factor de transcripción	Cáncer en fibroblastos (Van Dam y Castellazzi, 2001)
Ets	Factor de transcripción	Cáncer de mama (Behrens et al, 2001)
Myc	Factor de transcripción	Melanoma (D'Agnano, 2001)
Myb	Factor de transcripción	Cáncer colon (Funato et al., 2001)
Rel	Factor de transcripción	Leucemia (Madruga et al, 2000)
ErbA	Factor de transcripción	Eritroleucemia y sarcomas (Thormeyer y Baniahmad, 1999)
p53	Factor de transcripción	Carcinoma de colon (Mahyar-Roemer y Roemer, 2001)

## 2.1 Activación de Ras.

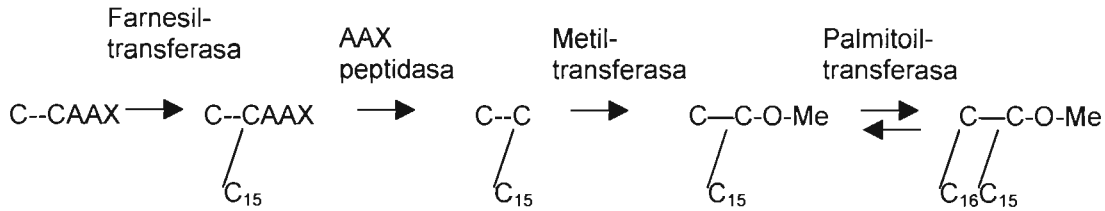
Las proteínas Ras existen en un equilibrio entre su forma activa o unida a trinucleótidos de guanina y su forma inactiva o unida a dinucleótidos de guanina. El ciclo de transición entre un estado y otro está regulado por dos tipos de proteínas: los factores intercambiadores de nucleótidos de guanina (GEFs por sus siglas en inglés) y los factores que incrementan la actividad de GTPasa (del inglés, GAPs). En la figura 5 se esquematiza el ciclo de activación de las proteínas Ras.



**Figura 5.** Ciclo de activación de las proteínas Ras. El ciclo de transición entre la forma activa (Ras-GTP) y la inactiva (Ras-GDP) de las proteínas Ras está regulado por factores intercambiadores de guanina (GEFs) y por factores que activan trifosfatasas de guanina (GAPs). Ras-GTP activa a sus proteínas efectoras (Adjei, 2001).

Las proteínas H-Ras, N-Ras y K-Ras son modificadas post-traduccionalmente en su región amino terminal en la secuencia CAAX (figura 6). En la secuencia CAAX (C=cisteína 186, A= aminoácido alifático y X aminoácido no alifático), la cisteína es alquilada por un grupo farnesilo, los amino ácidos AAX son removidos por proteólisis. En el carboxilo  $\alpha$  de la cisteína terminal se da una metilesterificación. El producto modificado es más hidrofóbico que el no modificado y se asocia débilmente con la membrana celular. Posteriormente las proteínas N-Ras, H-Ras y K-Ras4A son palmitoiladas en las cisteínas de la región hipervariable (amino ácidos 165-185), incrementándose su unión a la membrana figura 6. En la proteína K-Ras4B no ocurre la palmitoilación, ya que posee 6 lisinas en los amino ácidos 175-180 que le permite unirse a la membrana, por interacciones electrostáticas con fosfolípidos ácidos y otros grupos membranales cargados negativamente en la cara interna de la membrana plasmática (Hancock et al., 1990; Ghomashchi et al., 1995).

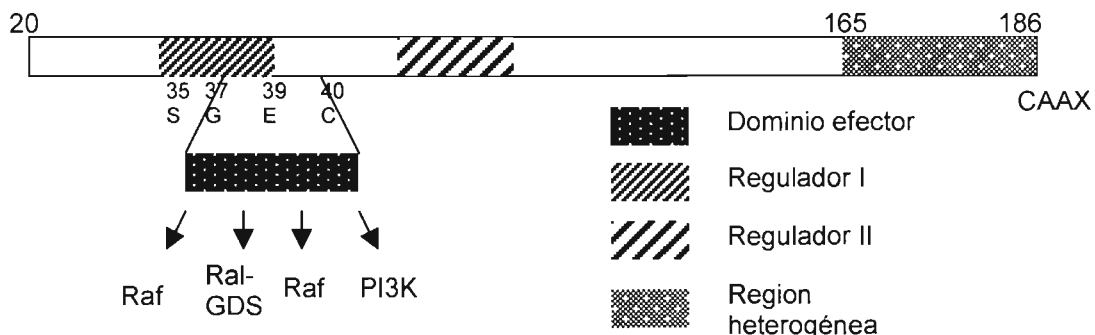




**Figura 6.** Modificaciones post-traduccionales en la región amino terminal de las proteínas Ras. En la secuencia CAAX (C corresponde a cisteína 186, A corresponde a aminoácido alifático y X corresponde a aminoácido no alifático), las proteínas Ras son modificadas post-traduccionamente, estas modificaciones incrementan su hidrofobicidad y consecuentemente su unión a membrana. El nombre de las enzimas que catalizan la reacción correspondiente se encuentra sobre las flechas. Las modificaciones que se dan son: primero a la cisteína 186 se le une un grupo farnesilo, segundo los aminoácidos AAX son removidos por proteólisis, tercero en el carboxilo  $\alpha$  de la cisteína terminal se da una metilesterificación y cuarto las proteínas N-Ras, H-Ras y K-Ras4A son palmitoiladas en las cisteínas de la región hipervariable. En la proteína K-Ras4B no ocurre la palmitoilación, ya que posee una región de poli-lisinas que le permite unirse por atracciones electrostáticas con fosfolípidos negativos a la membrana plasmática (Hesketh, 1994, Hancock et al., 1990; Ghomashchi et al., 1995).

## 2.2 Vías de señalización en las que participa Ras.

Diferentes tipos de señales extracelulares, que estimulan a la célula a través de receptores localizados en la membrana plasmática, con o sin actividad de cinasa de tirosina, o acoplados a proteínas G, activan a Ras. Una vez activada como Ras-GTP, activa a su vez a proteínas intermediarias de la señalización o proteínas efectoras. Se requiere de la unión específica entre las diferentes proteínas y Ras para que éstas sean activadas. En la figura 7 se encuentra un esquema de la proteína Ras en donde se identifican los diferentes dominios y los sitios en los cuales se unen las diferentes proteínas. Las tres proteínas efectoras de Ras mejor caracterizadas son: Raf, PI3K y Ral-GDS. Existen otras proteínas que interactúan con Ras como MEKK1, Rin1, AF-6 y Nore-1, pero aún no se ha determinado si son efectoras de Ras, por lo que para nuestros modelos únicamente consideramos las vías en las que participan Raf, PI3K y Ral-GDS.



**Figura 7.** Dominios importantes en la actividad de las proteínas Ras. El dominio efector es el sitio en el que interactúan las proteínas efectoras de Ras. El número y letra indican el aminoácido importante en la interacción con Raf, Ral-GDS o PI3K. La región heterogénea como su nombre lo indica, varía en las diferentes proteínas Ras. La región heterogénea contiene la secuencia CAAX, esta secuencia sufre modificaciones post-traduccionales importantes para el anclaje de Ras en la cara interna de la membrana plasmática. El dominio regulador I está relacionado con la interacción de GAPs, factores que activan trifosfatasas de guanina (GTPase Activating Proteins). El dominio regulador II está relacionado con la interacción con factores intercambiadores de guanina GEFs (Guanine Exchange Proteins). Las GAPs

mantiene a Ras unido a GTP, que es su forma activa y las GEFs unido a GDP o inactivo, (Crespo y León, 2000).

## 2.3 Proteínas efectoras de Ras.

### 2.3.1 Raf.

La fosforilación catalizada por proteínas cinasas permite la transmisión de las señales extracelulares. Las vías de las cinasas activadas por mitógenos MAPK (mitogen-activated protein kinases), se dividen en tres módulos involucrados en la transmisión de la señal.

Un módulo está formado por la MAP cinasa o cinasa regulada por señales extracelulares, ERK (extracellular signal-regulated protein kinase). Esta es activada por fosforilación por una cinasa de MAP cinasa o también llamada cinasa de ERK (MEK o MAPKK). Esta cinasa es activada por fosforilación por una cinasa de MEK (MEKK o MAPKKK). Raf es una cinasa de residuos de serina y treonina, y es una MEKK.

Raf se une a Ras de manera GTP-dependiente. Esta asociación permite que Raf se localice en la cara interna de la membrana plasmática, en donde es activada mediante un mecanismo en el que participa la proteína 14-3-3 $\zeta$ . Una vez activa, la proteína Raf activa a MEK, por fosforilación en residuos de serina.

La función de MEK tiene una especificidad dual ya que fosforila residuos de tirosina y de treonina, sus sustratos son MAPKs como ERK 1 y ERK 2. La activación de MAPKs da como resultado la fosforilación y activación de la cinasa ribosomal S6 (p70S6K), de factores de transcripción como c-Jun, c-Myc y c-Fos, que se unen a genes cuyos productos proteicos están involucrados en la proliferación celular.

Las ERKs tienen diferentes tipos de sustratos incluyendo proteínas citoplásmicas como: la proteína SOS que es un factor intercambiador de nucleótidos de guanina, la cinasa supresora de Ras Rsk, PHAS-1, la fosfolipasa A2 (PLA2), y otras proteínas que al ser fosforiladas por ERK son translocadas al núcleo como los factores de transcripción Elk-1, Ets-2, C/EBP y SMADs. (Kosako et al., 1992; Marshall, 1994; Mansour et al., 1994).

### 2.3.2 PI3K.

Otra proteína efectora de Ras es la fosfatidil inositol-3 cinasa (PI3K Phosphatidylinositol 3-kinase). PI3K es una cinasa de lípidos que cataliza la fosforilación de fosfoinosítidos en la posición 3' del anillo de inositol en respuesta a diferentes estímulos, mediados por receptores asociados a proteínas G. PI3K está compuesta de dos subunidades, la reguladora p85, de la cual se han identificado 5 isoformas y la catalítica p110 también con 5 isoformas descritas, con diferente especificidad por sustrato. Dos de estas isoformas, p110a y p110b se unen a Ras-GTP. Esta interacción da lugar a un aumento de la actividad de PI3K, y consecuentemente al incremento en la concentración de inositol fosforilado en 3', como el fosfatidilinositol 3,4,5-trifosfato (PIP3). PIP3 activa a Rac. Rac induce la producción de fosfatidil 4,5-bifosfato (PIP2) por activación de las cinasas PI4 y PI5, (Vanhaesebroeck et al., 1997; Fruman et al., 1998; Rodríguez-Viciano et al., 1996; Downward, 1998). Los sustratos de PI3K son PKB/Akt y la proteína ribosomal p70S6K, los que participan en el control

transcripcional. La activación de PKB genera señales antiapoptóticas y la regulación de vías metabólicas que llevan a la fosforilación de GSK3 y otros sustratos. La activación de PI3K es suficiente para inducir la transcripción del promotor de *fos*, la proteína Fos es un factor de transcripción que se activa por la vía de las MAPK, pero se requiere de la activación PI3K para la estimulación de la síntesis de ADN (Hu et al., 1995; McIlroy et al., 1997; Toker y Cantley, 1997).

### **2.3.3 Ral-GDS.**

La tercera proteína efectora de Ras es la proteína Ral-GDS. Ral es un miembro de la superfamilia de proteínas G pequeñas y consiste en RalA y RalB. Así como otras proteínas G pequeñas, la forma inactiva de Ral está unida a GDP y la forma activa unida a GTP. La conversión de Ral unido a GDP a Ral unido a GTP la realiza Ral-GDS, e inversamente la forma unida a GTP cambia a la unida a GDP por RalGAP (Hinoi et al., 1996; Shirouzu et al., 1998).

Ral-GDS se une a Ras en respuesta a señales de estimulación de crecimiento. Rap1 regula la interacción de Ras con Ral-GDS. Rap fue identificado por revertir la transformación oncogénica inducida por Ras.

La proteína cinasa A inhibe la activación de Raf, pero no la interacción entre Ras y Ral-GDS. (Spaargaren y Bischoff, 1994).

Se conoce poco sobre la actividad de Ral-GDS, pero se sabe que inhibe la miogénesis del esqueleto, lo que explica por ejemplo el hecho de que Ras inhiba la diferenciación de las células musculares independientemente de que las células sigan proliferando. Ya que por un lado Ras activa a Raf que a través de las MAPKs activa la proliferación y por otro lado activa a Ral-GDS que participa en la inhibición de la diferenciación (Ramocki et al, 1998).

## **2.4 Regulación del ciclo celular y apoptosis por la vía Ras/Raf/MEK/ERK.**

Las vías de transducción de señales que involucran a Ras/Raf/MEK/ERK regulan el ciclo celular y la apoptosis en diferentes tipos celulares. Mutaciones en estas vías se han observado en líneas de células transformadas y frecuentemente están relacionadas con cáncer. Estas vías pueden inducir tanto eventos asociados con proliferación, como con el arresto del ciclo celular. Estas vías también están involucradas en mantener la supervivencia celular por modular la actividad de moléculas que participan en la apoptosis como Bad y Bcl<sub>2</sub>.

### **2.4.1 Efecto proliferativo y antiproliferativo de la vía Ras/Raf/MEK/ERK.**

Las proteínas Ras tienen diferentes habilidades para activar Raf/MEK/ERK y la vía PI3K/Akt. K-Ras activa preferentemente Raf/MEK/ERK y H-Ras activa la vía PI3K/Akt. En células tumorales es más frecuente encontrar mutado K-Ras, lo que sugiere que la vía Raf/MEK/ERK es un blanco más para la generación de tumores

El control del ciclo celular es un proceso complejo, controlado esencialmente por las siguientes proteínas:

Las cinasas dependientes de ciclina (CDKs cyclin-dependent kinases) y sus subunidades reguladoras, las ciclinas.

Un grupo de proteínas inhibidoras de las CDKs, las CKIs.

La familia Ras participa en el control del ciclo celular al regular a las proteínas pRb, p53, ARF y Mdm2.

La regulación de estas proteínas la hace en general por medio de su proteína efectora Raf. Raf induce la expresión de CKIs, de las familias INK4 o Cip/Kip. La inducción de CKIs da lugar a la inhibición de CDKs, deteniendo a las células en la fase G<sub>1</sub> y/o G<sub>2</sub>.

#### **2.4.2 Regulación de pRb.**

pRb es un supresor de tumores, evita que las células entren en la fase S del ciclo celular. Ras influye en la actividad de las CDKs que fosforilan a pRb. La fosforilación de pRb promueve la actividad de los factores de transcripción E2F que se requieren para la expresión de genes que promueven la entrada en la fase S.

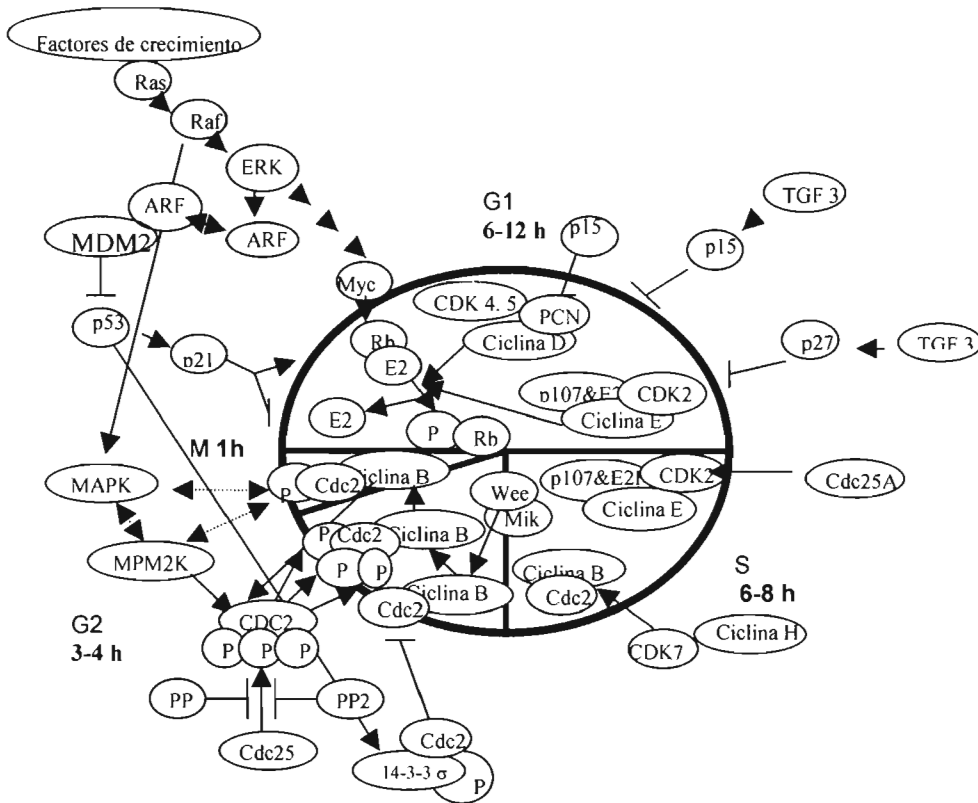
#### **2.4.3 Inducción de ARF y p53.**

El mecanismo por el cual Ras induce p53 es a través de la vía de las MAPK.

En esta vía, como se mencionó, Ras activa a la cinasa Raf. Raf fosforila y activa a la cinasa ERK. ERK activa a ARF. ARF se une a Mdm2 y lo secuestra en el núcleo, Mdm2 es un regulador negativo de p53. El secuestro de Mdm2 provoca la acumulación de p53. p53 induce a los genes que codifican para p21<sup>CIP</sup>, 14-3-3  $\sigma$  y GADD45. p21<sup>CIP</sup> inhibe la actividad del complejo Ciclina/cdk2 y detiene el ciclo celular. La proteína 14-3-3 $\sigma$  se une a Cdc2 manteniéndola en el citoplasma, evitando su entrada al núcleo y la formación del complejo Cdc2/CiclinaB, y con ello el inicio de la mitosis, y GADD45 induce senescencia, (McMahon y Woods, 2001 y Wilkinson y Millar, 2000)

El mecanismo por el cual Ras induce la expresión de CKIs en células transformadas, es a través de la activación de Raf que induce la expresión CKIs, de la familia de INK4 o Cip/Kip. La inhibición de las CDKs lleva a la detención de las fases G<sub>1</sub> y/o G<sub>2</sub> del ciclo. Este hecho sugiere que en algunas células neoplásicas, el control del ciclo celular está en estado de dormancia, pero bajo ciertas circunstancias puede ser activado. Aparentemente la habilidad de Ras para provocar senescencia prematura se debe a un mecanismo de defensa natural para evitar transformación neoplásica (Nurse, 2000; Blagosklonny y Pordee, 2001).

La figura 8 es un esquema de las diferentes fases del ciclo celular con algunas proteínas que participan en su control.



**Figura 8.** Fases del ciclo celular. Cada fase se realiza en un tiempo determinado, aquí se presenta el tiempo promedio de duración en células normales. También se muestran algunas de las proteínas que participan en las diferentes fases y en los puntos de control (Wilkinson, 2000).

### 2.5 Proteínas 14-3-3.

A continuación se describe a la familia de proteínas 14-3-3. Los miembros de esta familia no son proteínas efectoras de Ras, pero algunas de ellas participan en las vías mediadas por Ras y en el control del ciclo celular.

#### Generalidades.

Las proteínas 14-3-3 se expresan en la mayoría de los tejidos, aunque en 1967 fueron identificados por primera vez dentro de un grupo de proteínas ácidas de 30 KDa abundantes en el cerebro (Moore y Perez, 1967). El nombre 14-3-3 corresponde a su posición de migración en una cromatografía de DEAE-celulosa de las proteínas del cerebro con 30 KDa. Se determinó posteriormente por HPLC que 14-3-3 es una mezcla de 7 isoformas a las que se les identifica con letras griegas.

En el humano existen 7 isoformas, productos de 7 genes diferentes, aunque se reportan 9 isotipos α, β, δ, ε, η, γ, σ, θ, ψ y ζ, de los cuales α y δ son formas fosforiladas de β y θ respectivamente.

Más de 15 isoformas están presentes en plantas y se han identificado dos isoformas en levadura, en *Drosophila melanogaster* y en *Caenorhabditis elegans*.

(Rosenquist et al, 2001; Wang y Shakes, 1996). En la tabla 4 se resumen las propiedades de las proteínas 14-3-3.

**Tabla 4.** Propiedades generales de las proteínas 14-3-3. R corresponde a arginina, S a serina, pS a serina fosforilada, X a cualquier residuo de aminoácidos, P a prolina y T a treonina.

Propiedades físicas	Sitios de unión	Actividad
Proteínas ácidas 28-33 KDa	Proteínas que se unen a pS/pT Consenso: RSXpSXp RXXXpSXP No consenso	Modula la actividad enzimática Modifica la localización
Estructura conservada	Unión independiente de fosforilación	Previene la desfosforilación
Forma homo y heterodímeros		Promueve la estabilidad protéica
En mamíferos al menos 7 isoformas, 15 en plantas y 2 en <i>Drosophila</i> y <i>C.elegans</i> .		En algunos casos inhibe interacciones proteicas y en otras media interacciones proteicas

### Relación entre la estructura y la función.

Las proteínas 14-3-3 forman homo y heterodímeros, lo que les permite interactuar con una gran variedad de proteínas celulares. Se han reportado más de 50 proteínas de señalización que pueden unirse a 14-3-3; en la tabla 5 se presentan algunas de ellas (Dougherty y Morrison, 2004). Las proteínas 14-3-3 se unen a residuos de fosfoserinas/fosfotreoninas; reconocen dos grupos de secuencias RSXpSXP y RXXXpSXP, donde X es cualquier aminoácido, R arginina, P prolina y pS serina fosforilada (Rittinger et al, 1999; Yaffe et al, 1997; y Aitken, 2002). También se han encontrado algunas interacciones de las proteínas 14-3-3 que no requieren de ningún residuo fosforilado, por ejemplo la unión a la exoenzima S, p190RhoGEF y al péptido inhibidor R18 (Henriksson, 2002; Masters et al 1999; Petosa et al, 1998; Wang et al 1999 y Zhain et al, 2001). Independientemente de que se requiera fosforilación o no, la interacción se da en el mismo dominio de las proteínas 14-3-3 (Wang et al, 1999).

La primera función conocida de las proteínas 14-3-3 fue su relación con hidroxilasas de tirosina y triptofano, enzimas involucradas en la síntesis de neurotransmisores. Posteriormente se describieron como proteínas chaperonas, adaptadoras, localizadoras o translocadoras y como reguladoras de las vías de transducción de señales; en la tabla 5 se enlistan algunas proteínas que unen las diferentes isoformas de la proteína 14-3-3.

La estructura de las proteínas 14-3-3 se presenta en la figura 9, como ejemplo se presenta un heterodímero de las isoformas  $\theta$  y  $\psi$  y un homodímero de la isoforma  $\zeta$ , las flechas señalan los sitios de unión a péptidos con residuos de fosfoserina. Cada monómero está compuesto de un arreglo de 9 alfa hélices antiparalelas organizadas en dos dominios estructurales (Rittinger et al, 1999).

La formación de heterodímeros facilita la regulación de proteínas con funciones diferentes, ya que una isoforma puede unir a una proteína y la otra isoforma a otra proteína (Praseuth et al, 1999).

**Tabla 5.** En esta tabla se enlistan algunas proteínas que se unen a las diferentes isoformas de la proteína 14-3-3. Se presentan las secuencias de unión y el efecto biológico de la misma.

Ligando	Actividad	Secuencia	Efecto después de la unión	Referencia
cRaf	Cinasa de S y T	RS <sup>202</sup> TS <sup>204</sup> TP	La mantiene en su forma activa e inactiva	McPherson et al, 1999
Raf1	Cinasa de S y T	RS <sup>619</sup> AS <sup>621</sup> EP	La mantiene en su forma activa e inactiva	Muslin et al, 1996; Morrison y Cutler, 1997
Bad	Homólogo de Bcl2	RS <sup>97</sup> RS <sup>99</sup> AP	La retiene en el citoplasma e inhibe su función proapoptótica	Zha et al, 1996; Zha et al, 1997
Antígeno mediano T	Producto proteico del oncogene mta	RS <sup>257</sup> HS <sup>259</sup> YP	Promueve tumores en algunos tejidos	Cullere et al, 1998; Senften et al, 1997
ASK1	Cinasa de S y T	RS <sup>964</sup> IS <sup>966</sup> LP	Inhibe su función proapoptótica	Zhang et al, 1999
Cdc25C	Fosfatasa de Y y T	RS <sup>216</sup> PS <sup>218</sup> MP	La retiene en el citoplasma y evita la entrada en mitosis	Peng et al, 1997; Kumagai et al, 1998; Yang et al, 1999; López-Girona et al, 1999
Receptor de glucocorticoides	receptor	RS <sup>132</sup> TS <sup>134</sup> VP	Regula su actividad por relocalización del receptor rip 140	Zilliaccus et al, 2001; Wakui et al, 1997
KSR1	Cinasa de S y T	RS <sup>297</sup> PS <sup>299</sup> FP	Favorece su interacción con Raf y la vía de las MAPK	Cacace et al, 1999; Xing et al, 1997
PLCγ	Fosfolipasa	RS <sup>338</sup> ES <sup>340</sup> SP	Regula su resistencia a luz UV	Andoh et al, 1998
PI3K (p85)	Cinasa	RS <sup>231</sup> PS <sup>233</sup> IP	Permite la unión de GP1bIX	Munday et al, 2000
PTPH1	Fosfatasa de Y	RS <sup>358</sup> LS <sup>361</sup> VE RVDS <sup>653</sup> EP	Participa en su regulación	Zhang et al, 1997
WPK4	Cinasa	RS <sup>388</sup> RS <sup>388</sup> YP	Represión catabólica	Ikeda et al, 2000
Cbl	Adaptadora	RHS <sup>619</sup> LPFS <sup>623</sup> RLGS <sup>639</sup> TFS <sup>642</sup>	Formación de complejos	Liu et al, 1997; Liu y Altman 1998
Queratina 18	Componente de citoesqueleto	RPVSS <sup>33</sup> AAS <sup>36</sup>	Distribución subcelular	Ku et al, 1998
PKC μ	Cinasa de S y T	RRLS <sup>205</sup> NVS <sup>208</sup> LT IRTSS <sup>219</sup> AELS <sup>223</sup> T	Inhibe su actividad de cinasa	Hausser et al, 1999
Receptor de IGF1	Cinasa de Y	SVPLDPS <sup>1272</sup> A SSS <sup>1283</sup> SLP	Posiblemente regula la diferencia entre la señal por unión de IGF o de IR	Craparo et al, 1997; Furlanetto et al, 1997; Rivera et al, 2000
5-Fosfatasa p43	Fosfatasa	RS <sup>364</sup> ES <sup>366</sup> EE	Estimula su actividad de fosfatasa	Waterman et al, 1998
p53	Factor de transcripción	KGQS <sup>376</sup> TS <sup>378</sup> RH	Incrementa su unión a ADN	Campbell et al, 1997
IRS-1	Proteína de unión	RSKS <sup>270</sup> QS HSRS <sup>374</sup> IP KSVS <sup>641</sup> AP	Regulación de sensibilidad a insulina	Ogihara et al, 1997
Lip35	Proteína asociada a MHC	RS <sup>8</sup> RS <sup>8</sup> CR	Permite su salida de RE	Kuwana et al, 1998
FKHRL1	Factor de transcripción	RS <sup>30</sup> CTWP RAVS <sup>253</sup> MD	Lo retiene en el citoplasma e inhibe su función proapoptótica	Brunet et al, 1999
Slob	Proteína que une canales de K <sup>+</sup>	RS <sup>52</sup> NS <sup>54</sup> AI RS <sup>77</sup> AS <sup>79</sup> SE	Modula su sensibilidad al voltaje de sus canales de K <sup>+</sup>	Zhou et al, 1999



**Figura 9.** La imagen (A) corresponde al heterodimero de la proteína 14-3-3 isoformas  $\theta$  y  $\zeta$ . La imagen (B) corresponde al homodimero  $\zeta$ ; en ambos casos las flechas indican el sitio de unión a péptidos con residuos de fosfoserina (Protein data bank ID acceso 1QJB).

### **El papel de las proteínas 14-3-3 en la transducción de señales.**

Las proteínas 14-3-3 participan en las vías de señalización al unirse a las proteínas PKC, Raf1, Bad, Bax, KSR1, Rin1, Byr2, RGS3, RGS7 y p90Rsk.

#### **Proteínas 14-3-3 y su relación con PKC.**

En este punto existe un conflicto, ya que hay grupos que presentan a las proteínas 14-3-3 como regulador positivo y otros como regulador negativo de PKC. Aparentemente la interacción de las diferentes isoformas de 14-3-3 con PKC da lugar a diferentes resultados. Algunas isoformas pueden unir lípidos o unirse a la región amino terminal de la membrana plasmática, por lo que las proteínas 14-3-3 en algunos casos median la interacción de PKC con proteínas de membrana. PKC fosforila a algunas isoformas de 14-3-3 (Van der Hoeven et al, 2000; Tan et al, 1999; Hausser et al, 1999).

#### **Proteínas 14-3-3 y su interacción con Raf.**

La cinasa Raf1 tiene un papel importante en la transducción de señales inducidas por factores de crecimiento. Después de la activación de Ras, Ras interactúa directamente con Raf1 y lo relocaliza en la membrana plasmática. La activación de Raf1 es un mecanismo que no se ha explicado completamente. Raf se mantiene en su estado activo al unirse una proteína 14-3-3 a los residuos de serina 259 y 621 fosforiladas. Se ha propuesto que la interacción de las proteínas 14-3-3 con Raf1 evita su desfosforilación y en consecuencia prolonga su activación.

Las isoformas  $\beta$  y  $\zeta$  están relacionadas con la regulación de Raf. Se ha demostrado que las proteínas 14-3-3 no se requieren para inducir la dimerización de Raf. Posiblemente se requieren para mantener a Raf en su forma inactiva en ausencia de Ras-GTP y estabilizar la conformación activa de Raf.

Ras-GTP desplaza a la proteína 14-3-3 del residuo de serina 259 fosforilada. En la Figura 10 se presenta un esquema general de la interacción entre las proteínas 14-3-3 y Raf1 (Weber et al, 2001; Zhang et al, 2000; Widen et al, 2000; Yip-Schneider et al, 2000; Qiu et al, 2000; Hayne et al, 2000; Yuryev et



al, 2000; MacNicol et al, 2000; Aoki et al, 2000; Van der Hoeven et al, 2000; Improta-Brears et al, 1999).

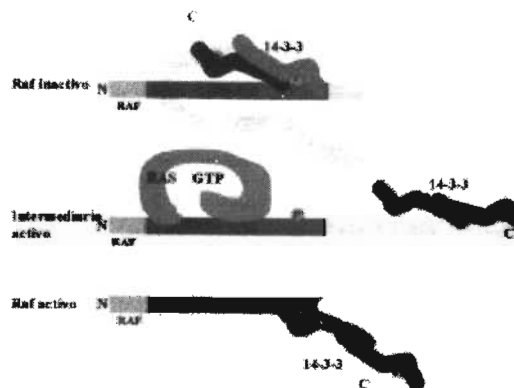


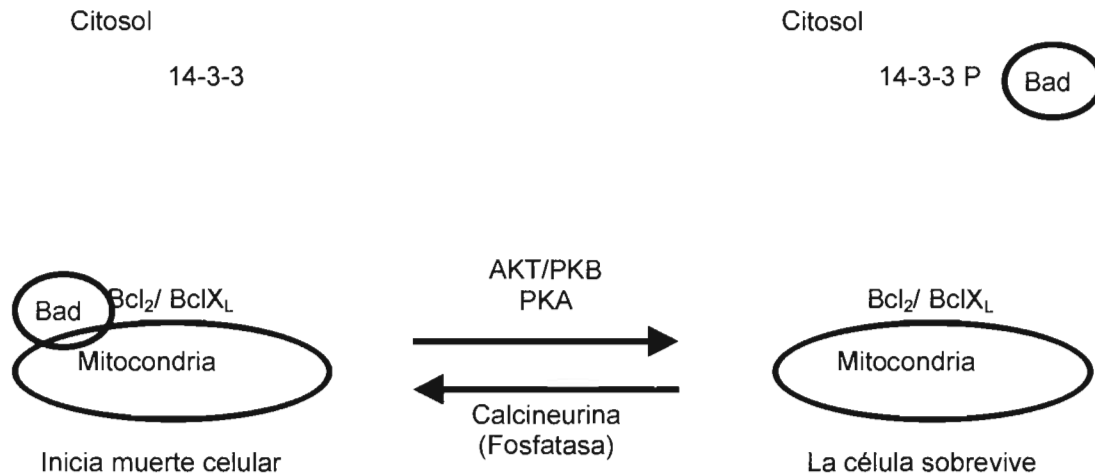
Figura 10. Modelo de interacción y regulación de la actividad de Raf1 por la proteína 14-3-3.

### Proteínas 14-3-3 y su relación con otras cinasas.

Las proteínas 14-3-3 se unen a diferentes proteínas involucradas en la transducción de señales. MEKK1; además de participar en la vía de las MAP cinasas, es un sustrato de la caspasa 3 (proteína involucrada en la muerte celular). Cuando MEKK1 está anclada, el dominio C terminal se relaciona con las proteínas 14-3-3; eso cambia la distribución de MEKK1 en la célula. El dominio de cinasa libre es proapoptótico, sin embargo si 14-3-3 está unida, pierde esa propiedad. Las proteínas 14-3-3 funcionan en parte como localizador específico de MEKK1 en la célula y cuando esta interacción es destruida con la unión de la caspasa 3 a MEKK1, ocurre la redistribución que puede alterar la función de la cinasa (Fanger et al, 1998).

### Proteínas 14-3-3, su relación con Bad y la muerte celular.

La apoptosis es un proceso de muerte celular que juega un papel crítico en el desarrollo de la patofisiología de una variedad de enfermedades incluyendo el cáncer. Un componente importante en regulación de este proceso es la familia de proteínas pro- y antiapoptóticas Bcl<sub>2</sub>. Las proteínas 14-3-3 interactúan con un miembro de esta familia, la proteína Bad. Bad puede dimerizarse con otros miembros de la familia, es un mediador de la muerte celular y sus niveles son dinámicamente regulados por estímulos apoptóticos; su expresión forzada lleva a la inhibición de la población celular. Bad causa muerte celular por unirse e inhibir la acción antiapoptótica de Bcl-X<sub>L</sub> y Bcl<sub>2</sub>. Las proteínas 14-3-3 evitan la interacción entre Bad y Bcl<sub>2</sub> o Bcl-X<sub>L</sub>. Al unirse las proteínas 14-3-3 a Bad, la última se mantiene en el citoplasma, lo que evita que se acerque a la membrana mitocondrial y se una a Bcl<sub>2</sub>/BclX<sub>L</sub>, de tal manera que es incapaz de inducir apoptosis. En la figura 11 se presenta un esquema de esta interacción (Korsmeyer, 1999).



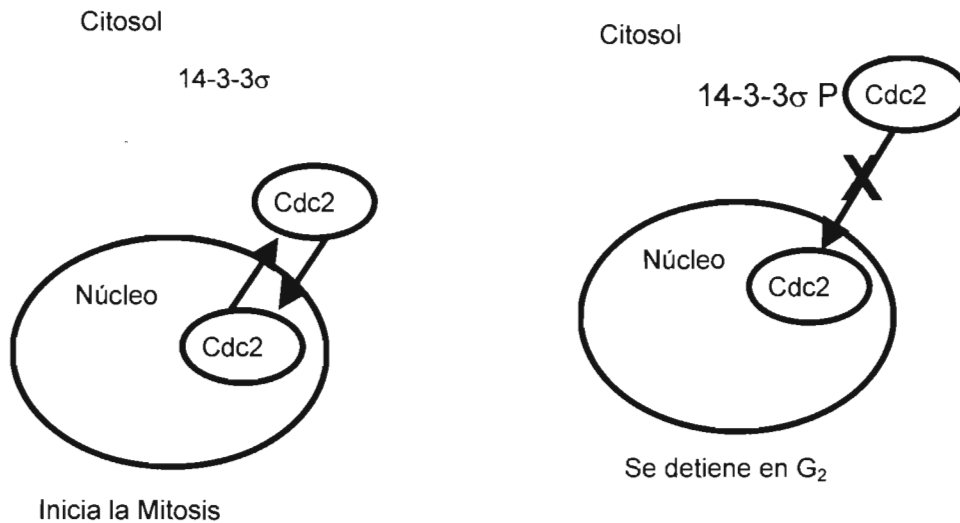
**Figura 11.** En este dibujo se representa como la proteína 14-3-3 secuestra a BAD en el citosol y evita su acceso a la mitocondria, y su unión a Bcl<sub>2</sub> y BclX<sub>L</sub>. La calcineurina puede revertir este proceso y causar muerte celular, ya que desfosforila a BAD y evita su unión con 14-3-3.

### Las proteínas 14-3-3, su relación con p53 y el control del ciclo celular.

Diferentes proteínas participan en el control de la transición de una fase del ciclo celular a otra.

Las proteínas p21 y 14-3-3 $\sigma$  son productos de algunos de los genes inducidos por p53 que es un regulador del ciclo celular.

Cdc25 es un regulador del ciclo celular, desfosforila y activa a la proteína cinasa Cdc2 para dar inicio a la mitosis. La inhibición de la desfosforilación de Cdc2 es clave para detener la entrada a mitosis en respuesta a daño al ADN. CDC25 al estar fosforilada y unida a 14-3-3 $\theta$  se encuentra en su forma inactiva. Cdc2 activo forma un complejo con la ciclina B, entra y se acumula en el núcleo, lo que inicia la mitosis. La proteína 14-3-3 $\sigma$  juega un papel importante en la localización de Cdc2. Después de que hay daño al ADN, la proteína 14-3-3 $\sigma$  se incrementa considerablemente por su inducción por p53. La proteína 14-3-3 $\sigma$  se une a Cdc2 fosforilada, manteniéndola en el citoplasma, evitando su entrada al núcleo y la formación del complejo Cdc2/CiclinaB, y con ello el inicio de la mitosis. En la figura 12 se encuentra un dibujo que representa estas interacciones (Chan et al, 1999 y 2000; Laronga et al, 2000; Hermeking et al, 1997).



**Figura 12.** En este dibujo se representa como 14-3-3 $\sigma$  secuestra a Cdc2 en citosol, evita su entrada a núcleo y desfosforilación por Cdc25. Al no entrar Cdc2 al núcleo no se da el inicio de la mitosis.

### Regulación de las proteínas 14-3-3 por fosfatasa.

Las fosfatasa desfosforilan los sitios de unión de las proteínas 14-3-3, regulando así su actividad. La fosfatasa PP2A desfosforila al residuo de serina 112 de BAD (14-3-3 requiere para su unión que estén fosforilados los residuos 112 y 136). Desfosforila al residuo de serina 259 de Raf (14-3-3 requiere que estén fosforilados los residuos 259 y 621). Desfosforila al residuo de serina 392 de KSR (14-3-3 requiere que estén fosforilados los residuos 297 y 392). La fosfatasa PP1 desfosforila al residuo de serina 216 de CDC25C y al residuo de serina 1394 de Ron, en ambos casos se requiere de esta fosforilación para la unión de 14-3-3 (Janssens y Goris, 2001; Cohen, 2002).

### I.III OLIGONÚCLEOTIDOS QUE FORMAN UNA TRIPLE HÉLICE (TRIPLEX).

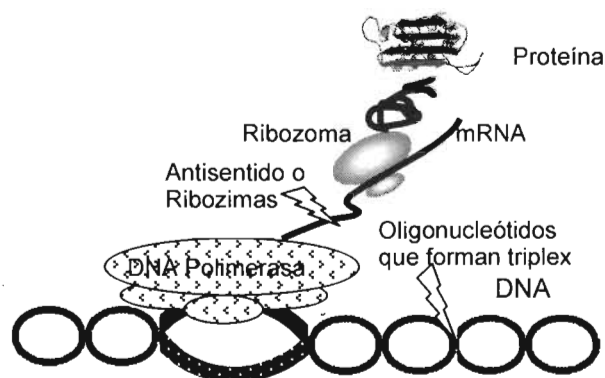
En 1968 se observó el potencial biológico de los oligonucleótidos que forman una triple hélice (triplex), al encontrar en experimentos *in vitro* que la transcripción mediada por la ARN polimerasa de *Escherichia coli* era inhibida por una tercera hebra de ARN, (Morgan y Wells, 1968).

En 1974 se propuso el uso de oligonucleótidos con fines terapéuticos, ya que al ser necesario probar un promedio de 10,000 compuestos para identificar una nueva sustancia biológicamente activa, se vio la necesidad de diseñar fármacos de una manera racional con varios propósitos y con la ventaja de ser específicos por estar dirigidos a ácidos nucleicos, (Arnott y Selsing 1974).

#### 1. Oligonucleótidos terapéuticos.

Existen tres diferentes estrategias para la regulación genética mediante el uso de oligonucleótidos.

Los oligonucleótidos antisentido, dirigidos a ARN mensajero, los oligonucleótidos formadores de triplex o estrategias antígeno, dirigidos a una secuencia de ADN de doble cadena y las ribozimas (endoribonucleasa de secuencia específica) o fragmentos de ARN catalíticos, dirigidos a un ARN mensajero blanco. Tanto el antisentido como las ribozimas se basan en la formación de puentes de hidrógeno Watson-Crick entre un blanco específico de ARN y el oligonucleótido o la ribozima, figura 13.



**Figura 13.** Oligonucleótidos terapéuticos. Los oligonucleótidos antisentido y las ribozimas están dirigidos al mRNA, los oligonucleótidos que forman triplex están dirigidos al DNA de doble cadena.

La inhibición de la expresión de un gen de manera específica en células en cultivo empleando oligonucleótidos que forman una triple cadena con fines terapéuticos fue reportada por primera vez por Zamecnik y Stephenson, 1978.

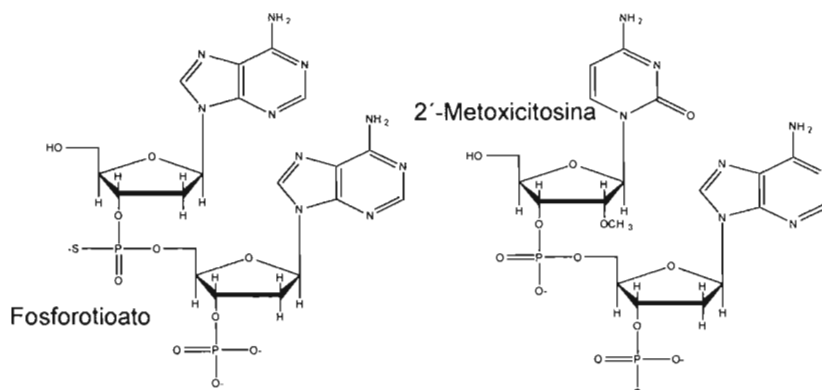
Los oligonucleótidos que forman una triple hélice, presentan varias ventajas con respecto a los oligonucleótidos antisentido o a las ribozimas que se dirigen al ARN mensajero (mARN), ya que los TFO están dirigidos al gen, del cual solamente hay dos copias (dos alelos) mientras que puede haber miles de copias de un mARN. Por lo que en teoría se requiere una menor concentración de oligonucleótidos para bloquear la expresión de un gen específico en la célula. Al bloquear la traducción del mARN o incluso induciendo una modificación secuencia-específica de la cadena del ARN, no se evita que el gen

correspondiente sea transcrito, de tal modo que se puede incrementar el mRNA a modificar o inhibir. En contraste, se espera que la prevención de la transcripción del gen disminuya la concentración del mRNA de una manera más eficiente y más duradera, dependiendo del tiempo que dure el antígeno unido a su secuencia blanco. Por estas razones seleccionamos a este tipo de oligonucleótidos terapéuticos para inhibir a la proteína blanco propuesta por nuestro modelo computacional.

Para el caso de los oligonucleótidos que forman una triple hélice inicialmente se emplearon oligonucleótidos con uniones fosfodiéster para inhibir la expresión de genes. Sin embargo, aparecieron varios problemas como:

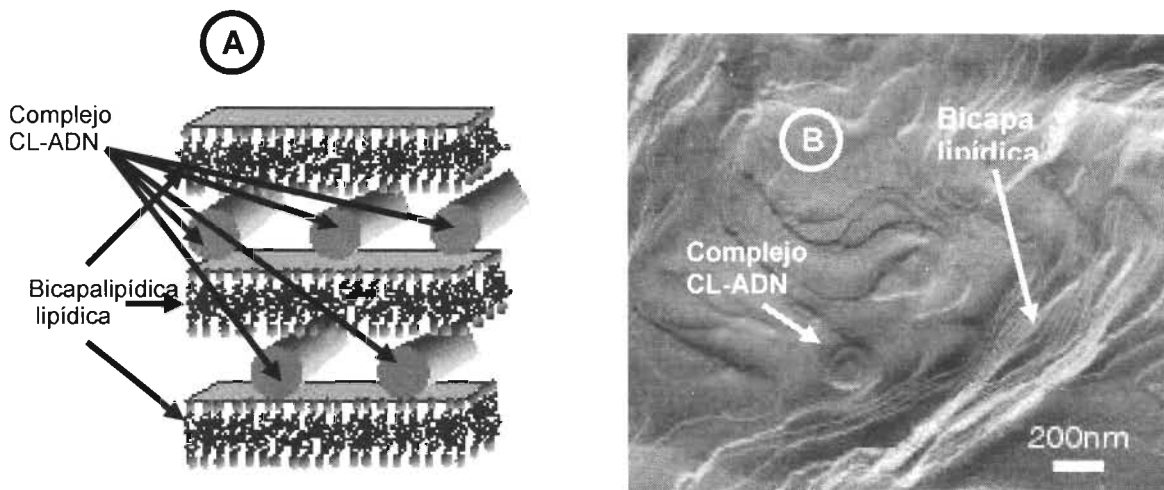
1. Rápida degradación por nucleasas celulares.
2. Transporte de los oligonucleótidos dentro de las células y compartimientos celulares.
3. La dirección de los oligonucleótidos al tipo celular apropiado dentro del organismo vivo y la selección de sitios de reconocimiento apropiados como blanco.

Para resolver estos problemas se inició el empleo de análogos de ADN que tienen las uniones fósforo internucleótido modificadas lo que les da una resistencia a las nucleasas celulares mientras mantienen su especificidad. La sensibilidad del oligonucleótido a nucleasas, disminuye también al emplear 2'-metil citosina (Leitner et al, 2000; Wang y Kool, 1995; Beban y Miller, 2000; Torigoe et al, 1999; Ushijima et al, 1999) figura 14.



**Figura 14.** Modificaciones que pueden evitar la acción de nucleasas. Fosforotioato y 2'-metoxicitosina, las nucleasas reconocen enlaces fosfodiéster.

La entrada hasta el núcleo se facilita con policones, ligandos hidrofóbicos o acarreadores, liposomas, lipoproteínas o nanoesferas, como los agentes transfectantes. Únicamente el ADN B, pero no el A o el Z, al mezclarse con vesículas de lípidos catiónicos forma de manera espontánea complejos (CL-ADN), los cuales se intercalan de paralelamente a la bicapa lipídica de las membranas. La carga positiva neta del CL-ADN es atraída por la carga negativa de la membrana plasmática, lo que favorece su entrada a la célula, figura 15.



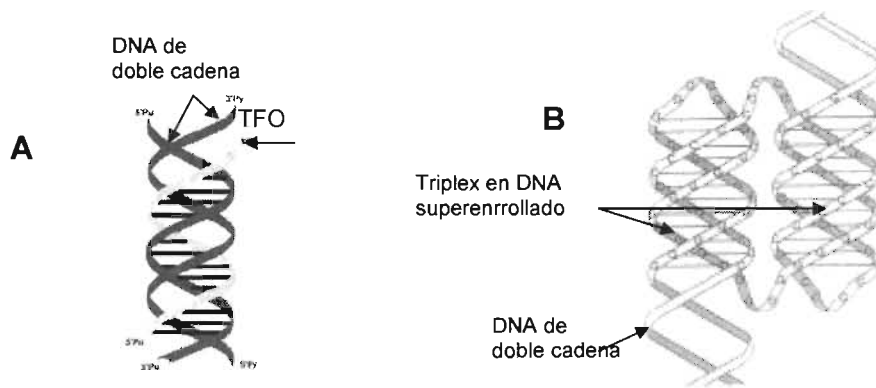
**Figura 15.** Complejo Lípido catiónico-ADN, (CL-ADN). A) Modelo de una serie de bicapas lipídicas y complejos CL-ADN. B) Micrografía electrónica de bicapa lipídica y complejo CL-ADN, <http://www.cmm.upenn.edu/research/cdna.html>.

Los oligonucleótidos pueden ser empleados como inhibidores de genes específicos, para investigar su función y regulación. Un oligonucleótido se puede utilizar también como trampa para atrapar proteínas que se unen a ADN, evitando que se asocien a su blanco normal y como agentes terapéuticos para modular de manera selectiva la expresión de genes.

Los oligonucleótidos que forman una triple hélice, pueden reconocer ADN de doble cadena en secuencias específicas formando enlaces tipo Hoogsteen o Hoogsteen-inversa, este tipo de unión se da al unirse con una de las hebras de la doble hélice una tercera hebra para dar lugar a una triple hélice.

La formación de una triple hélice se da cuando las bases de pirimidina o de purina se unen al surco mayor de la doble hélice de ADN por medio de uniones tipo Hoogsteen con las purinas de la doble cadena Watson-Crick. El ADN de triple cadena se puede categorizar en dos tipos, figura 16:

1. Triplex-formados intermolecularmente entre los oligonucleótidos que forman triplex (TFOs) y las secuencias blanco de ADN de doble cadena.
2. Triplex-formados intramolecularmente en las regiones de homopurina/homopirimidina de ADN superenrollado, que son los elementos principales de ADN H, una de las estructuras inusuales del ADN.



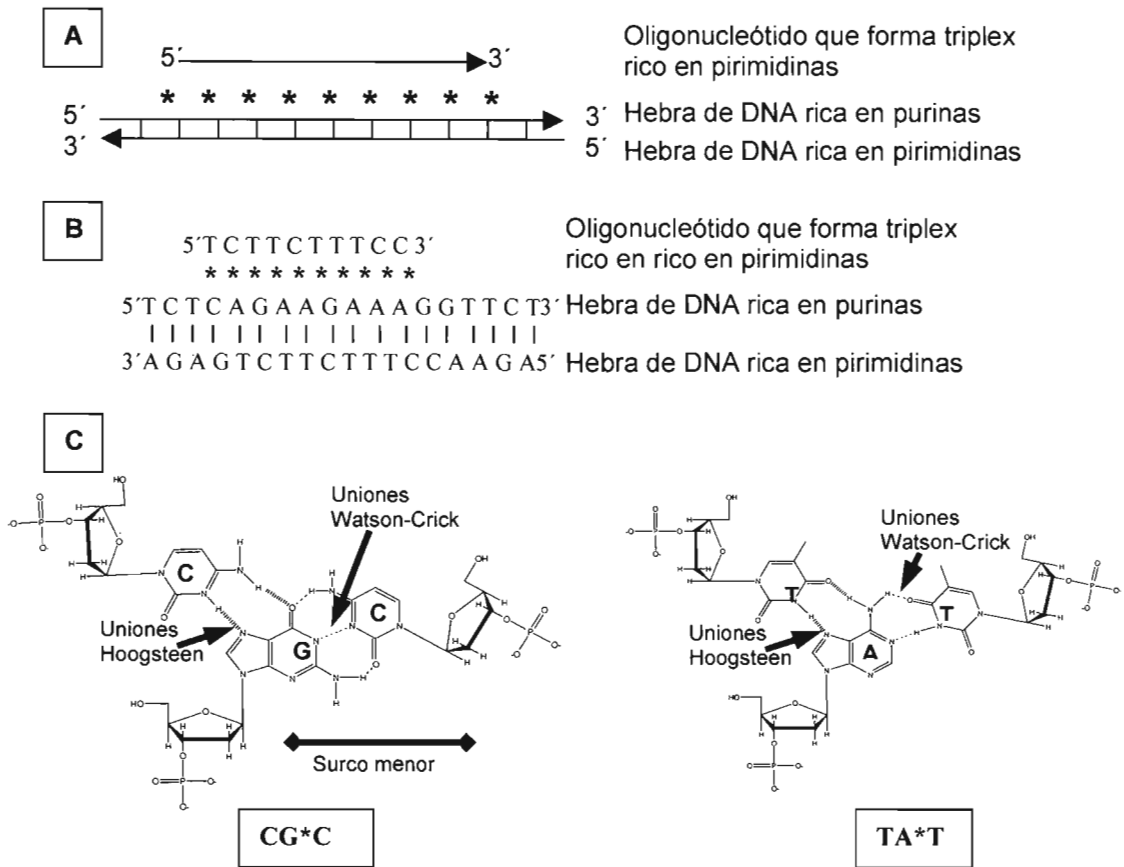
**Figura 16.** Triple hélice. A) Triplex formado entre dos moléculas de DNA de doble cadena y oligonucleótido formador de triplex. B) Triplex formado intramolecularmente.

Para describir un triplex se emplea una notación de tres letras: la primera letra corresponde a la base de la hebra rica en pirimidinas (Py) del dúplex, la segunda letra corresponde a la base complementaria en la hebra rica en purinas (Pu) del dúplex, y la tercera letra (generalmente después de un asterisco) denota la base correspondiente al oligonucleótido que forma el triplex (TFO), figura 17. El número de nucleótidos es descrito generalmente por subíndices, por ejemplo  $T_{10}A_{10}^*T_{10}$ .

Los TFO se categorizan dependiendo de su composición pirimidina o purina y su orientación con respecto al ADN de doble cadena. Por ejemplo, para un TFO de polipirimidinas unido de forma paralela a la hebra de purinas de ADN mediante uniones tipo Hoogsteen. Hay dos posibilidades canónicas (PyPu\*Py): TA\*T y CG\*C. El triplex CG\*C requiere que la citosina este protonada ( $C^+$ ), lo que generalmente se logra a bajo pH, figura 17.

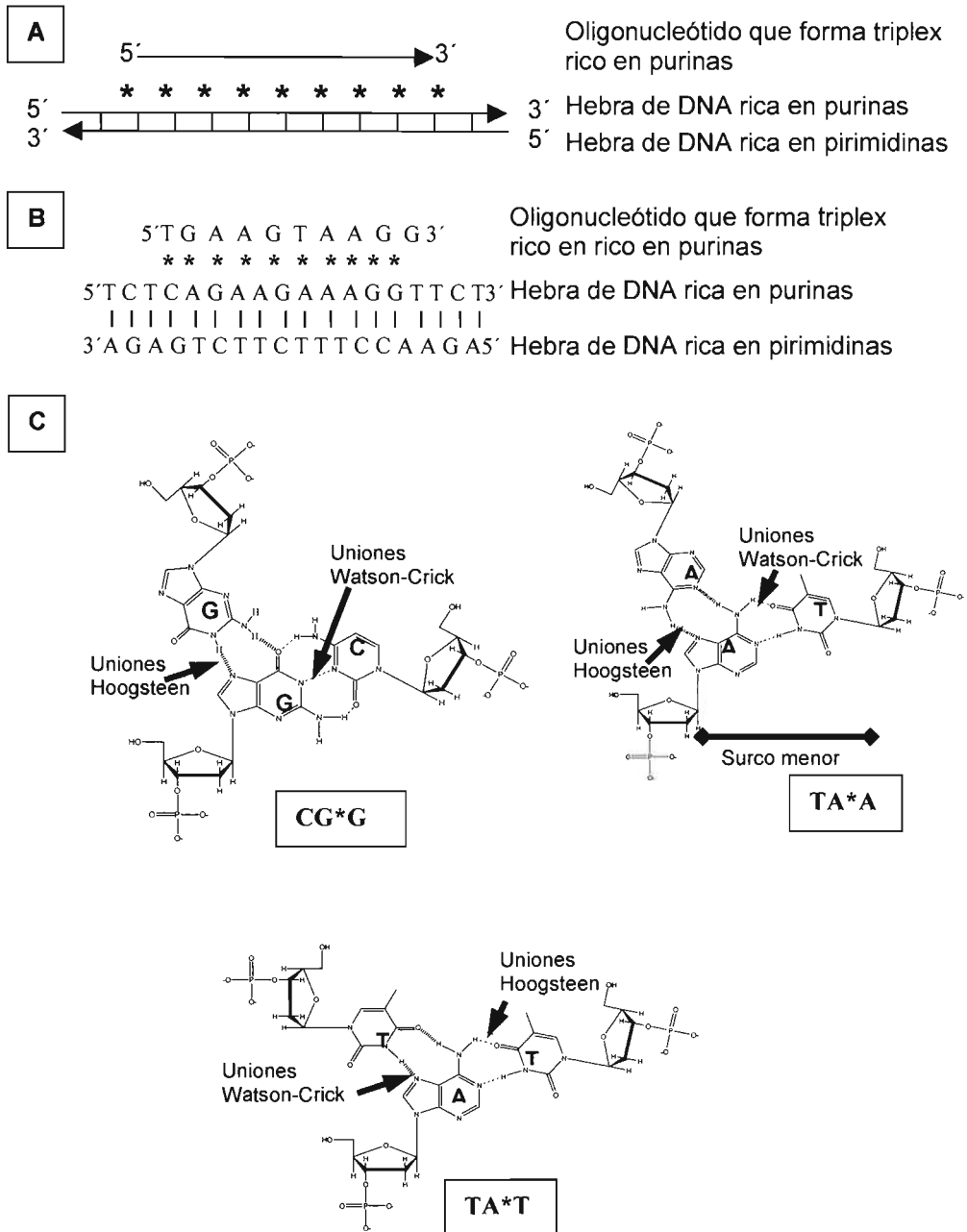
Para un TFO que se une antiparalelamente a la hebra rica en purinas de ADN mediante uniones tipo Hoogsteen inversa hay tres posibilidades canónicas (PyPu\*Pu): TA\*A, CG\*G, y TA\*T (T aparece aquí como excepción porque es una Py), figura 18.

La formación de Triplex es altamente específica; por ejemplo, 16 pirimidinas continúas reconocen y se unen únicamente a sus 16 bases complementarias en ADN genómico de  $3 \times 10^9$  bases (Strobel et al., 1991). Las polipurinas y las polipirimidinas, se estabilizan en ciertas condiciones químicas, tales como pH ( $C+G^*C$ ) y la concentración de los cationes divalentes ( $Mg^{2+}$ ,  $Ca^{2+}$  y  $Zn^{2+}$ ).



**Figura 17.** Oligonucleótidos que forman triplex, ricos en pirimidinas. A) Uniones de un TFO en orientación paralela a una hebra de polipurinas de DNA. Las flechas indican la orientación de la cadena B) Un ejemplo de oligonucleótidos que forman triplex. C) Dos triplex canónicos CG\*C y TA\*T.





**Figura 18.** Oligonucleótidos que forman triplex, ricos en purinas. A) Uniones de un TFO en orientación paralela a una hebra de polipurinas de DNA. Las flechas indican la orientación de la cadena B) Un ejemplo de oligonucleótidos que forman triplex. C) Tres triplex canónicos \*GGC, \*AAT y \*TAT.

## II. OBJETIVOS

### OBJETIVO GENERAL

Crear modelos computacionales de las vías de señalización intracelular en las que participa Ras, que faciliten la identificación de blancos terapéuticos y observar experimentalmente el efecto de su inhibición empleando oligonucleótidos que forman una triple hélice.

### OBJETIVOS ESPECÍFICOS

- a) Modelar las vías de señalización intracelular en las que participa Ras, con diferentes técnicas computacionales, que faciliten la identificación de blancos terapéuticos.
- b) Emplear oligonucleótidos que formen triple helice para inhibir a alguno de los blancos propuestos.

### III. MATERIAL Y MÉTODOS

#### III.1 Diseño de oligonucleótidos que forman una triple hélice.

Los oligonucleótidos fueron diseñados a partir de la secuencia promotora del gen que codifica para la proteína 14-3-3  $\sigma$  y del gen que codifica para la proteína 14-3-3  $\eta$ ; como control se usaron ambos genes de humano.

Se empleó el software MatInspector Public (<http://www.generegulation.com>) para realizar la búsqueda de probables sitios de unión de factores de transcripción dentro de la región promotora, así como de las cajas TATA y CAAT. La búsqueda se basó en la biblioteca Transfac 5.0. La determinación de especificidad celular de los factores de transcripción resultantes se corroboró con la base de datos de TFACTOR ([http://www.cbrc.jp/htbin/bget\\_tffactor](http://www.cbrc.jp/htbin/bget_tffactor)).

Aunque no se tiene aún un protocolo específico para determinar la secuencia y las modificaciones de las bases del TFO para dar lugar a un triplex estable, existen varios puntos que se deben cumplir:

1. Los oligonucleótidos de polipurina deben contener 65% o más de guanina para estabilizar el triplex. Una cadena de polipirimidinas rica en citosina desestabiliza el triplex (Chan y Glazer, 1997). Aunque en condiciones fisiológicas se requiere la protonación de las citosinas en la triple cadena, a pH entre 7.0 y 7.5, el triplex G\*GC es estable, por lo que se puede prescindir de la C (Sun 1991, Medhi 2002).
2. La longitud del oligonucleótido que forma triplex, debe ser entre 10 y 40 nucleótidos para disminuir la posibilidad de que la misma secuencia se encuentre en otro gen que no se quiere inhibir, cuanto más larga es la región, mayor posibilidad de formar el triplex correspondiente (Potaman, 2003).
3. Se sugiere que un oligonucleótido de polipurinas por lo menos debe contener entre 12 y 14 purinas ininterrumpidas para obtener un triplex estable (Chan y Glazer, 1997).
4. Una discontinuidad en una región de polipurinas es suficiente para causar una unión deficiente del triplex, (Potaman, 2003). Sin embargo se recomienda utilizar zonas alternadas de Polipurinas y Polipirimidinas. Para formar el triplex estable, la longitud de las porciones de Pirimidinas y de Purinas en un 5'-Pu-Py-3' debe ser por lo menos 4 bases (Frank-Kamenetskii y Mirkin 1995).
5. La inversión de algunas bases da una mayor estabilidad. Hasta la fecha, se han identificado las combinaciones siguientes para una inversión de CG, GC\*C, GC\*T y GC\*T, (Horne y Dervan, 1991). Para una inversión de TA, AT\*G (Gowers y Fox, 1999).
6. Para la mejor inhibición de un gen se recomienda que el TFO este dirigido a la región promotora para bloquear la unión de los factores de la transcripción (Chan y Glazer, 1997). En la cercanía de la región promotora se han observado secuencias de polipurinas y polipirimidinas, siendo estos sitios de alta afinidad para formar triplex que podrían tener un papel importante en la regulación de la expresión genética. Si se colocan secuencias de polipurinas/polipirimidinas en un gen reportero, y se adiciona un TFO, no se inhibe la expresión del gen; pero si se colocan antes del gen reportero se

reduce la expresión. Por lo tanto esta estrategia funciona como un regulador transcripcional en cis (Hobbs y Yoon, 1994).

7. En nuestro caso las citocinas fueron sustituidas por 2-metoxi-citosinas, para estabilizar el triplex e inhibir la acción de nucleasas (Leitner et al, 2000; Wang y Kool, 1995; Beban y Miher, 2000; Torigoe et al, 1999; Ushijima et al, 1999).

En base a los puntos mencionados se diseñaron los oligonucleótidos que se presentan en la tabla 6.

Se comprobó que los oligonucleótidos resultantes no formaran estructuras secundarias estables u homodímeros estables entre si, empleando el programa Oligo Tech. <http://www.oligoset.com>.

**Tabla 6.** Oligonucleotidos que forman triplex diseñados para inhibir la transcripción del gen que codifica para la proteína 14-3-3  $\sigma$ . Como control el oligonucleótido diseñado para inhibir la transcripción del gen que codifica para la proteína 14-3-3  $\eta$ .

Nombre	Dirigido a:	oligonucleótido
MCG-1	14-3-3 $\sigma$	5'AGGGCAGGGGAGAGGGG3'
MCG-2	14-3-3 $\sigma$	5'GGAGCAGGGGCUAGGG3'
MCG-3	14-3-3 $\sigma$	5'GGCTGAGAAGAGAACAGG3'
MCG-4	14-3-3 $\sigma$	5'GGACAAGAGAAGAGTCGG3'
MCG-5	14-3-3 $\eta$	5'TGGGCUGGGAGAGUGGC3'

```

Oligo1.dna
5'GGGAAACTGAGGCTCAGGGAAAGGAGGGG 3'
-----
SEQUENCE
1 4 7 10 13 16 19 22 25 28
5' GGG AAA CTG AGC TCT GAG GGA AAG GAG GGG 3'

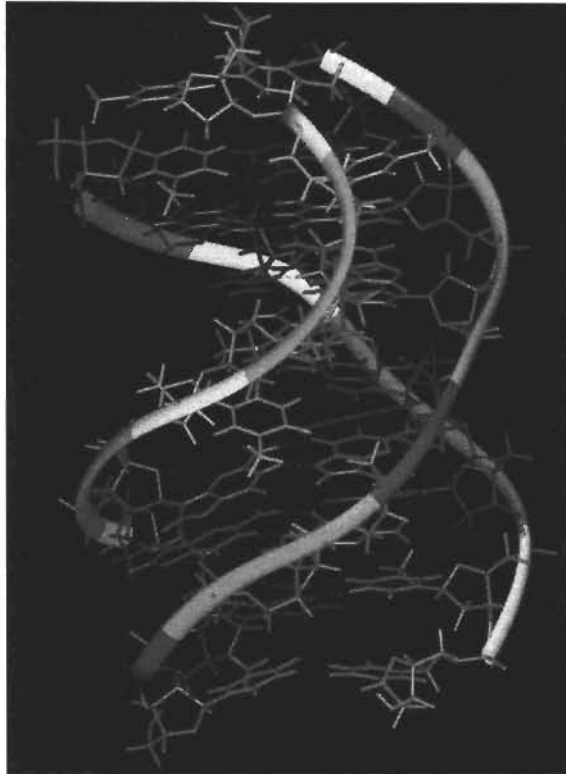
3.17 nM/OD
30.43 ug/OD
MW = 9.6k (one strand)
Primer to Target Tm (by NGC) = 83.6°C

COMPOSITION:
A 9.00 30.0%
C 2.00 6.7%
G 16.00 53.3%
T 3.00 10.0%
GC 18.00 60.0%
A+T 12.00 40.0%

STEM LOOP STRUCTURE:
4
13
TCAAAAGCC 5'
11
GAGCTCTGAGGGAAAGGAGGGG 3'
14
13
G = 2.5 kcal/mol
No Stable Secondary Structure

HOMODIMER:
No Homodimer
  
```

**Figura 19.** Visualización en pantalla del resultado que ofrece el programa OligoTech. Este programa indica si el oligonucleótido forma estructuras secundarias estables u homodímeros estables.



**Figura 20.** La imagen es una representación gráfica de un triplex. La imagen fué obtenida empleando el programa Insight de Biosym, en la Dirección general de cómputo académico de la UNAM.

Los oligonucleótidos fueron sintetizados en la Unidad de Biología Molecular del Instituto de Fisiología Celular de la UNAM. Los oligonucleótidos quiméricos con fluoresceína fueron sintetizados por la compañía "Oligos etc".

Los oligonucleótidos fueron cuantificados espectrofotométricamente determinando su absorbancia a 260 nm.

### **III.II Formación de triplex *in vitro*.**

Para determinar si los oligonucleótidos diseñados y sintetizados forman una triple cadena *in vitro* es necesario obtener una doble cadena de la secuencia blanco.

#### **1. Formación de ADN de doble cadena (dúplex) *in vitro***

Se mezclaron 30 nmoles de cada oligonucleótido complementario de cadena sencilla en un volumen final de 200  $\mu$ l en una solución compuesta por 50 mM Tris-HCl pH 7.4, 50 mM NaCl, 5 mM MgCl<sub>2</sub>. La mezcla se incubó durante 10 min a 90 °C, seguidos de 30 min a 60 °C, de 2 h a 37 °C y de 12 h a temperatura ambiente, las incubaciones se realizaron en bloques térmicos de Fisher Scientific (Catapano et al., 2000).

## 2. Formación de triplex *in vitro*.

Se incubaron 10 nmoles de ADN de doble cadena con 10 nmoles del oligonucleótido que forma la triple cadena, en un volumen final de 50  $\mu$ l a una concentración final de 200  $\mu$ M, en una solución que contiene 50 mM Tris-HCl pH 7.4, 50 mM NaCl, 5 mM MgCl<sub>2</sub>, toda la noche a 37 °C, en un bloque térmico Fisher Scientific (Catapano et al., 2000).

Con el fin de comprobar experimentalmente la formación de una triple cadena *in vitro*, se hicieron los siguientes experimentos.

## 3. Desnaturalización Térmica.

Los experimentos de desnaturalización térmica se efectuaron empleando un espectrofotómetro Hewlett Packard 8453, 845 UV-Visible con un sistema Peltier que permite el cambio de temperatura en la celda que contiene la muestra.

Se determinó el cambio en absorbancia a 260nm, incrementando la temperatura a una velocidad de 0.5 °C/min desde 10 °C hasta 90 °C. Las muestras se mantuvieron a temperatura ambiente hasta antes de iniciar el experimento. Se colocaron en una concentración final de 2.0  $\mu$ M de ADN de cadena doble o triple en 50 mM Tris-HCl pH 7.4, 50 mM NaCl, 5 mM MgCl<sub>2</sub>, en celdas de cuarzo de 0.5 cm, con tapa de teflón, para evitar la evaporación de la muestra (Alunni-Fabbroni et al., 1996).

## 4. Dicroísmo Circular.

Los experimentos se efectuaron en un espectrofotómetro para dicroísmo circular Aviv modelo 62DS, haciendo un barrido de 200 a 320 nm, empleando 2.0  $\mu$ M de ADN de cadena sencilla, doble o triple en 50 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM MgCl<sub>2</sub>, en una celda de cuarzo de paso óptico de 0.1 (capacidad máxima 200  $\mu$ l), a 37 °C. Se empleó como blanco el mismo amortiguador (Alunni-Fabbroni et al., 1996).

## 5. Movilidad en gel.

La electroforesis se realizó en geles de poliacrilamida al 15% en condiciones nativas en amortiguador 50 mM tris-boratos pH 8.0, 5 mM MgCl<sub>2</sub>. A 20  $\mu$ g de ADN de cadena sencilla, doble o triple se les adicionó 5  $\mu$ l de Ficoll 400 al 15% y azul de bromo fenol y se llevó a un volumen final de 35  $\mu$ l con amortiguador 50 mM tris-glicina pH 8.3. Cada muestra se colocó en un pozo diferente. La electroforesis se corrió a 50 V a temperatura ambiente. Cuando el colorante llegó al borde del gel, se detuvo la electroforesis y el gel se colocó dentro de una solución de bromuro de etidio al 1%, durante 2 h, pasado este tiempo se lavó durante 2 h con agua destilada y se observó la fluorescencia en un transiluminador; la imagen se obtuvo empleado una cámara digital Sony MVC-FD75 Mavica (Catapano et al., 2000).

### III.III Experimentos en líneas celulares.

Para determinar el efecto biológico del oligonucleótido diseñado, se realizaron experimentos en líneas celulares en cultivo.

Líneas celulares empleadas en este trabajo

Línea celular	Origen	Condiciones de cultivo
NIH 3T3 Línea celular empleada para verificar que ni el oligonucleótido, ni el método de transfección como tales son tóxicos para las células.	Fibroblastos de ratón	Se crecieron en medio de cultivo D-MEM (Dulbecco's Modified Eagle's medium) adicionado con 10% de suero fetal bovino y los siguientes antibióticos y antimicóticos penicilina G, sulfato de estreptomycin, anfotericina B, y se incubaron a 37 °C en una atmósfera con 5% de CO <sub>2</sub> en una incubadora Forma Scientific.
MCF-7 Línea empleada en los experimentos para determinar el efecto biológico del oligonucleótido.	Cáncer humano de mama	Se crecieron en medio de cultivo RPMI-1640 (Roswell Park Memorial Institute) adicionado con 5% de suero fetal bovino y 5% de suero de ternera y los siguientes antibióticos y antimicóticos penicilina G, sulfato de estreptomycin, anfotericina B, y se incubaron a 37 °C en una atmósfera con 5% de CO <sub>2</sub> en una incubadora Forma Scientific.

Condiciones de cultivo de las células y tratamiento con adriamicina.

Se sembraron  $2 \times 10^5$  células/ml 24 h antes de cualquier experimento y después se centrifugaban a 1,000 rpm durante 10 min, se resuspendieron en PBS (phosphate buffered saline solution, Fosfato de Sodio 10mM pH 7.4 y 0.9% de Cloruro de sodio); a 50  $\mu$ l de la suspensión celular se les adicionó el colorante vital azul de tripano al 0.4%; se contaron en una cámara de Neubauer y se diluyeron con medio fresco a la concentración requerida.

Es necesario causar daño al ADN para activar la transcripción del factor transcripcional p53, el cual activa la transcripción de la proteína 14-3-3 $\sigma$ . Para causar daño al ADN empleamos adriamicina que se intercala en el zurco mayor de ADN de doble cadena e inhibe a la Topoisomerasa II (Tewey et al., 1984; Gelvan and Samuni, 1986; Bodley et al., 1989; Tarr y Helden, 1990).

Para determinar la concentración necesaria de adriamicina, así como el tiempo de incubación, para observar disminución en el crecimiento celular sin causar la muerte de la totalidad de la población, se procedió de la siguiente manera.

Se sembraron  $5 \times 10^4$  células/ml en placas de 96 pozos y se incubaron empleando 0, 0.5, 1, 5, 10 y 15  $\mu$ M como concentración final de adriamicina, y se incubaron a 37 °C, las células se fijaron con formaldehído al 4% después de incubar con las diferentes concentraciones de adriamicina a las 0, 1, 2, 4, 8, 12, 24 y 32 h. Una vez que todas las células se fijaron dentro de los pozos se les adicionó una solución de cristal violeta al 0.5% en agua, en esta solución se mantuvieron durante 20 min, pasado este tiempo se lavó con agua abundante.

Las células teñidas se resuspendieron en ácido acético al 10%, posteriormente se determinó la absorbancia a 600 nm.

### 1. Efecto del oligonucleótido MCG3 sobre el crecimiento de la línea celular MCF7.

Para determinar si el oligonucleótido MCG 3 tenía algún efecto en el crecimiento celular, se sembraron  $5 \times 10^4$  células/ml en placas de 96 pozos y se incubaron empleando 0, 0.25, 0.5, 1, 2.5 y 5  $\mu\text{g}$  de oligonucleótido por ml de medio de cultivo. Antes de adicionar el oligonucleótido, éste se mezcló con el agente transfectante Metafectene (Biontix Laboratories GmbH) en una proporción 1/5, en 100  $\mu\text{l}$  de medio de cultivo libre de suero y antibiótico durante 15 min a temperatura ambiente; pasado este tiempo se realizaron las diluciones necesarias para obtener las diferentes concentraciones y en cada pozo se colocaron 100  $\mu\text{l}$  de la dilución correspondiente, se incubaron a 37 °C, 6 h después de haber adicionado la mezcla de transfección (oligonucleótido/ Metafectene), se le adicionó a cada pozo 100  $\mu\text{l}$  de medio con antibiótico, 10% de suero de ternera y 10% de suero fetal bovino. La concentración de oligonucleótido se calculó considerando 200  $\mu\text{l}$  de medio de cultivo. Después de 48 h de la adición de la mezcla de transfección se le agregó adriamicina a una concentración final de 5  $\mu\text{M}$  y se incubaron a 37 °C, las células se fijaron con formaldehído al 4% después de incubar con las diferentes concentraciones de oligonucleótido a las 0, 1, 2, 4, 8, 12, 16 y 24 h. Una vez que todas las células se fijaron dentro de los pozos, se les adicionó una solución de cristal violeta al 0.5% en agua, en esta solución se mantuvieron durante 20 min, pasado este tiempo se lavó con agua abundante. Las células teñidas se resuspendieron en ácido acético al 10%. Posteriormente se determinó la absorbancia.

El Metafectane permite la entrada del ADN a las células, ya que lo condensa en estructuras compactas, evita la destrucción osmótica y desestabiliza los lípidos de la membrana mediante fuerzas electrostáticas repulsivas.

El mismo tipo de experimento se realizó empleando a la línea inmortalizada de fibroblastos de ratón NIH 3T3, con el fin de determinar si el método de transfección o el oligonucleótido eran tóxicos para la célula. Se empleó esta línea celular porque el oligonucleótido fue diseñado contra la proteína 14-3-3  $\sigma$  de humano y no de ratón.

### 2. Citometría de flujo.

Para citometría de flujo se sembraron  $2 \times 10^5$  células /ml en cajas con 6 pozos y se les adicionó 1ml de medio de cultivo libre de suero y antibiótico, Metafectene 2.5  $\mu\text{l}/\text{ml}$  de medio de cultivo y oligonucleótido 1 $\mu\text{g}/\text{ml}$  de medio de cultivo.

La concentración final se calculó considerando un volumen final de 2 ml.

Las células con las diferentes soluciones, se incubaron a 37 °C durante seis horas. Pasado este tiempo se les adicionó 1ml de medio de cultivo con 10% de suero de ternera y 10% de suero fetal bovino.

Después de 48 h a una serie se adicionó y a otra no, adriamicina en una concentración final de 5  $\mu\text{M}$  y se incubó 8, 12 o 24 h.



Pasados los tiempos de incubación las células se lavaron durante un minuto con PBS y se despegaron de los pozos empleando 1 ml de tripsina/EDTA al 0.025 en PBS durante un minuto o hasta observar que todas las células estaban despegadas. La actividad de esta proteasa se detuvo agregando 1 ml de medio de cultivo con 10 % de suero. Las células se resuspendieron delicadamente y se transfirieron a un tubo para centrifuga de 15 ml, se centrifugaron a 1,000 rpm durante 5 min a 4 °C. La pastilla de células se lavó 2 veces con PBS, centrifugando a 1,000 rpm durante 5 min a 4 °C.

Después del segundo lavado, a la pastilla se le agregó 1ml de etanol al 70% para fijar las células, se mezclaron en vortex durante un minuto, y nuevamente se centrifugaron a 1,000 rpm durante 5min a 4 °C.

Finalmente se resuspendieron en 200µl de una solución de 0.5 mg/ml de yoduro de propidio en PBS con 0.1% de azida de sodio, pH 7.4, la suspensión se diluyó con 800 µl de PBS, se mantuvieron a 4 °C, hasta su análisis en un citómetro de flujo FACSCalibur de Becton Dickinson.

### 3. Hibridación tipo western.

Se sembraron  $5 \times 10^4$  células/ml en cajas de 30 mm y se les adicionó 1 ml de mezcla de transfección previamente incubada a temperatura ambiente, la mezcla contenía 1µg de oligonucleótido/ml de medio de cultivo y 2.5 µl de Metafectene/ml de medio de cultivo. El medio de cultivo no contenía suero y el cálculo se hizo considerando un volumen final de 3 ml de medio de cultivo. Las células se incubaron durante 6 h, pasado este tiempo se les adicionó medio de cultivo con 20% de suero. Después de 48 h se les adicionó adriamicina en una concentración final de 5 µM y se incubaron durante 8, 12, 16 y 24 h. De manera paralela, como control, se incubó otro grupo de cajas en las mismas condiciones, sin adicionarles oligonucleótido. Pasado el tiempo de incubación las células se colocaron en una cama de hielo, se les eliminó el medio de cultivo y se lavaron 2 veces con PBS, el cual se succionó empleando una bomba de vacío casi hasta sequedad. Posteriormente se les adicionaron 300 µl de una solución de lisis que contenía 50 mM Tris-HCl pH 8, 120 mM NaCl, 0.5% Nonidet P-40, 100 mM NaF y 200 mM NaVO<sub>3</sub> e inhibidores de proteasas: 10 µg/ml de leupeptina, 70 UI/ml de aprotinina y 0.58 mM de fluoruro de fenil-metil-sulfonilo (PMSF). Se incubaron en hielo durante 5 min. Pasado este tiempo y empleando un raspador se colectó todo el lisado y se transfirió a un tubo para microcentrifuga.

La mezcla se centrifugó a 4 °C durante 15 min a 12 000 rpm en una microcentrifuga. El sobrenadante se transfirió a otro tubo para microcentrifuga y se mantuvo a -20 °C hasta su uso.

Las proteínas contenidas en el sobrenadante se cuantificaron por el método de Bradford, empleando albúmina sérica bovina para la curva patrón de proteínas y extrapolando las absorbancias obtenidas.

Alrededor de 100 µg de proteína de cada muestra se corrió en un gel de acrilamida al 12.5 %, a 40 mV durante 4 h o hasta que el frente del colorante se encontrara en el borde del gel. Se transfirió a una membrana de celulosa y se hibridó con un anticuerpo policlonal específico para la proteína 14-3-3σ (Santa

Cruz Biothecchnology, Inc.), posteriormente se adicionó reactivo para quimiolumiscencia de Pierce y se expuso a placas radiográficas Kodak.

Una parte del gel que contenía a otras proteínas y que no se transfirió, se tiñó con azul de Coomasie, con el fin de determinar que la cantidad de proteína fue la misma en cada carril. Las proteínas en el gel se fijaron con una solución de isopropanol al 25 % y ácido acético al 10 % en agua, a continuación se adicionó una solución de azul de Coomasie y finalmente se eliminó el exceso de colorante con una solución para desteñir con metanol al 5 % y ácido acético al 7 % en agua. Todos los reactivos empleados en este trabajo fueron obtenidos de Sigma Chemical Products.

## V. RESULTADOS Y DISCUSIÓN

Por la naturaleza de este trabajo que incluye el desarrollo de modelos computacionales para modelar las vías de señalización intracelular y experimentos en células en cultivo para coprobar las predicciones de los modelos con el comportamiento real de las células, los resultados se dividen en dos partes:

-En la primera parte se hace la descripción de los modelos computacionales desarrollados y los resultados teóricos obtenidos con ellos.

-En la segunda parte se presentan los resultados experimentales obtenidos en células en cultivo. Se inhibe la transcripción de uno de los blancos terapéuticos propuesto por el sistema experto desarrollado por nosotros.

### 1. MODELOS COMPUTACIONALES.

Los modelos computacionales que se presentan en esta tesis se realizaron con el fin de comprender y analizar las vías de señalización en las que participa la proteína Ras. Elegimos diferentes estrategias considerando a las proteínas como componentes de un sistema dinámico discreto. En este tipo de sistemas cada elemento tiene un número de estados numerable y finito. Estos estados cambian de manera discreta en el tiempo. La relación entre elementos está determinada por reglas lógicas que indican como un elemento define el estado de otro u otros, (Cárdenas-García et al. 2000; González-Pérez et al. 2000 y 2001; Armas et al. 2000).

Un sistema dinámico discreto, es aquel en el que cada elemento tiene un número de estados numerable y finito 0, 1, 2, 3, etc. (activo, inactivo, fosforilado, etc.). La varoación del estado depende de su interacción con los demás elementos. Para definir el comportamiento temporal del sistema es necesario saber el estado inicial y las reglas con las cuales el estado siguiente se obtiene del estado anterior.

En la literatura los sistemas de este tipo también se llaman booleanos. El número de estados del sistema booleano (si consideramos como el estado del sistema, al conjunto de los estados de sus elementos) es finito. Por eso, algún estado del sistema debe repetirse tarde o temprano. Como los estados siguientes se determinan únicamente por los estados anteriores, eso significa que finalmente el sistema entra en un ciclo, que se llama el ciclo límite o atractor. Puede ser que el atractor sea un solo estado que en este caso es "estado de equilibrio". Después de llegar no hay nuevos estados o se repiten los mismos.

Los sistemas booleanos son una aproximación a la realidad, pero con un número de elementos suficientemente grande, procesos muy complicados se pueden modelar. Hay que notar que la computadora es un sistema booleano, por eso, estrictamente hablando, cualquier modelo computacional podría reducirse a un modelo Booleano. Si tenemos solo dos estados por elemento en el sistema de  $n$  elementos, el número total de estados es  $2^n$ . Entonces, el tiempo de modelaje generalmente crece exponencialmente con  $n$ , lo que da lugar a un sistema muy demandante en términos de tiempo de cómputo.

Para nuestro primer modelo empleamos un método que simplifica el análisis, al subdividir las vías de señalización en bloques. Con esto, el tiempo de

cálculo disminuye drásticamente. El mecanismo de subdivisión está basado en la topología abstracta que es la aplicación de los axiomas matemáticos de topología, aplicados a cualquier otro conjunto de puntos o elementos (O’Neil, 1982). La figura 21 corresponde a un diagrama de flujo del desarrollo de los modelos computacionales.

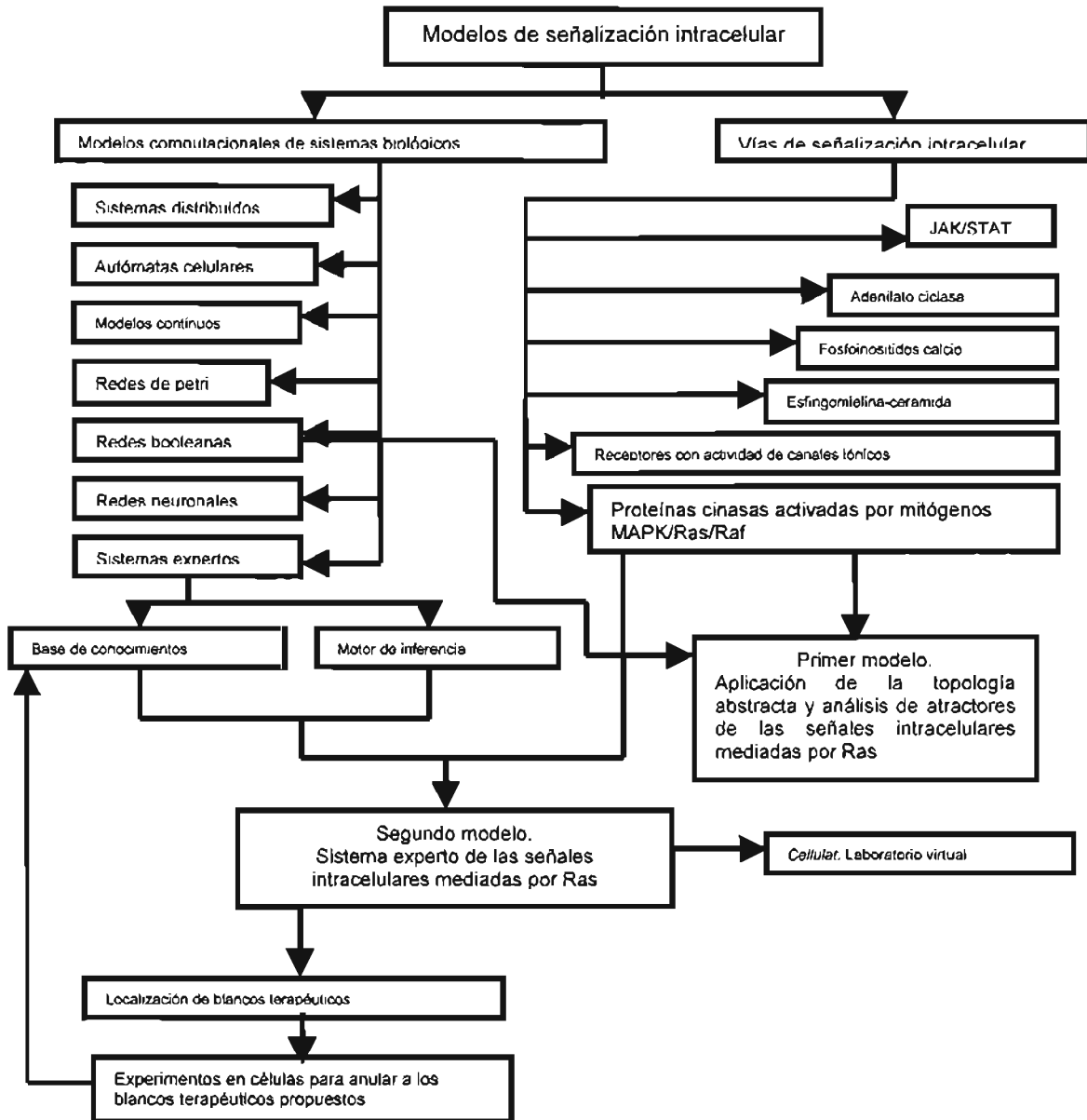


Figura 21 Representación general de los pasos que se siguieron para el desarrollo de los modelos computacionales. Ambos modelos se pueden emplear para facilitar la búsqueda de blancos terapéuticos. El sistema experto por medio de la estructura de pizarrón puede comunicarse con el laboratorio virtual *Cellular*. Los experimentos en líneas celulares relacionadas con los blancos propuestos dan lugar a información que incrementa la base de conocimientos del sistema experto.

1.1 PRIMER MODELO.

En la figura 22 se representan las vías generadas por Ras y sus proteínas efectoras Raf, RalGDS, PLC y PI3K, así como las vías que relacionan a Ras con el control del ciclo celular, consideradas por nosotros para este modelo. Los elementos que participan son proteínas, hormonas, factores de crecimiento, etc. La flecha que va de un elemento a otro indica que el primer elemento "determina" el estado del segundo elemento, por ejemplo: un elemento lleva a la activación, fosforilación o a la formación de un complejo de un segundo elemento. Si el elemento tiene más de una flecha señalando hacia él, indica que su estado está determinado por la acción cooperativa de todos los elementos conectados a él. Una flecha con la punta blanca indica inhibición. (Cárdenas-García, 2002).

Las interacciones que se observan en la figura 22 son complicadas. Si queremos considerar a estos elementos para hacer un modelo booleano simple, el número de posibles combinaciones de estados es muy grande, por lo que para su análisis lo dividimos en grupos de elementos que interactúan, que puedan ser analizados de manera más o menos independiente. La tabla 7 es una lista de los elementos que se encuentran en la figura 22.

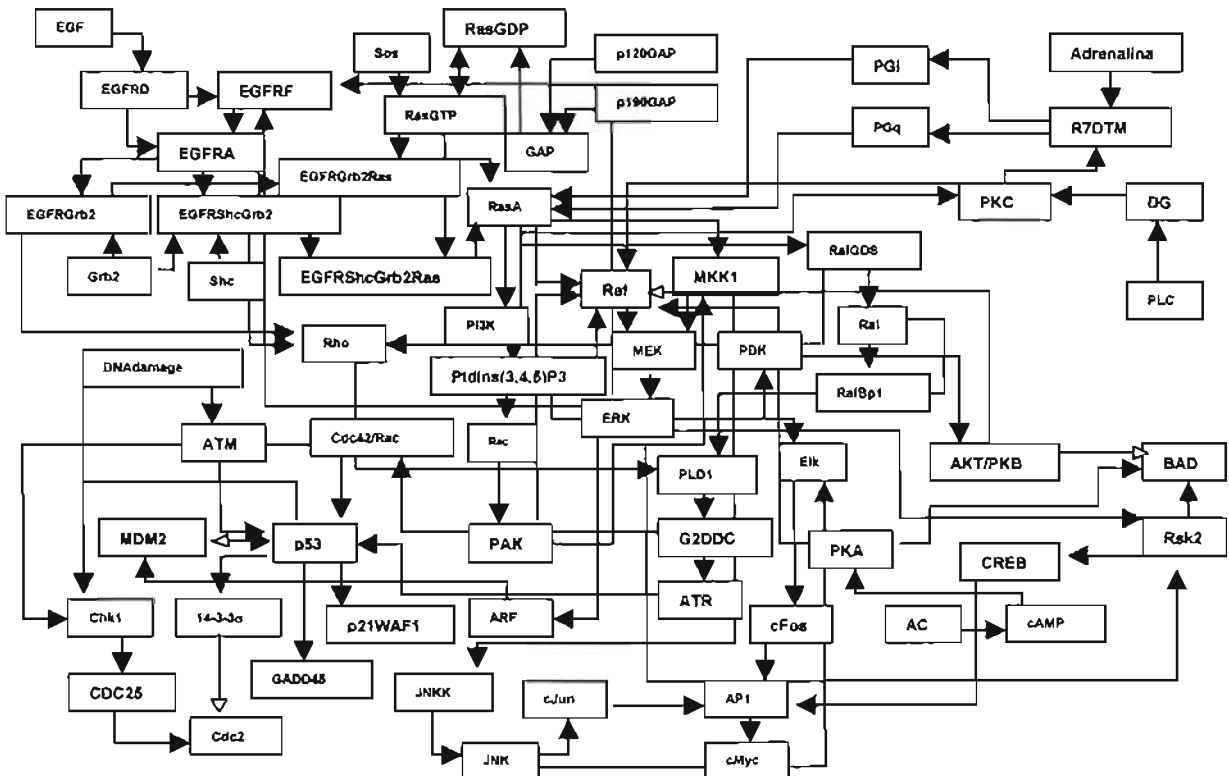


Figura 22. Interacciones entre los elementos que participan en las vías de señalización relacionadas con Ras. Las flechas indican interacciones positivas (activación, fosforilación, etc.), las flechas con punta blanca interacciones negativas (inhibición, desfosforilación, etc.).

**Tabla 7.** Elementos que participan en las vías de señalización que involucran a Ras consideradas para el primer modelo computacional. La activación o inactivación de cada elemento a considerar depende de una serie de elementos, dentro de los que se encuentran complejos proteicos.

No	Nombre del elemento	Su cambio de estado depende de:	No	Nombre del elemento	Su cambio de estado depende de:
0	EGF	EGF	33	RalBp1	Ral
1	EGFRA	EGFRD EGFRF	34	PLD1	Rho Ral
2	EGFRD	EGF	35	G2DDC	PLD1
3	EGFRF	EGFRD ERK	36	ATR	G2DDC
4	EGFRGrb2	EGFRA Grb2	37	ATM	Daño a DNA
5	EGFRShcGrb2	EGFRA Shc Grb2	38	Daño a DNA	Daño a DNA
6	EGFRGrb2Ras	EGFRGrb2 RasGTP	39	Cdc42/Rac	G2DDC
7	EGFRShcGrb2Ras	EGFRShcGrb2 RasGTP	40	p53	Cdc42/Rac ATM ATR MDM2
8	Grb2	Grb2	41	MDM2	p53 ARF
9	Shc	Shc	42	ARF	ERK
10	RasGTP	RasGDP Sos	43	p21WAF1	p53
11	Sos	Sos	44	Cdc2	CDC25 14-3-3 $\sigma$
12	RasGDP	RasGTP GAP	45	14-3-3 $\sigma$	p53
13	GAP	p120GAP p190GAP	46	GADD45	p53
14	p120GAP	p120GAP	47	CDC25	Chk1
15	p190GAP	p190GAP	48	Chk1	Cdc42/Rac daño a DNA
16	RasA	EGFRGrb2Ras EGFRShcGrb2Ras Pgi PGq	49	MEK	Raf MKK1
17	PGi	R7DTM	50	ERK	MEK
18	PGq	R7DTM	51	Elk	ERK JNK
19	R7DTM	Adrenalina PKC	52	cFos	EIK1
20	Adrenalina	Adrenalina	53	AP1	cFos cJun
21	PKC	DG PI3K	54	cMyc	AP1
22	DG	PLC	55	JNKK	MKK1
23	PLC	PLC	56	JNK	JNKK
24	RalGDS	RasA	57	cJun	JNK
25	Raf	RasA PKC PKA PAK AKT/PKB	58	PDK	PtdIns(3,4,5)P3
26	MKK1	RasA PAK	59	AKT/PKB	PDK
27	PAK	Rac	60	BAD	AKT/PKB PKA Rsk2
28	Rac	PtdIns (3,4,5)P3	61	AC	AC
29	PtdIns (3,4,5)P3	PI3K	62	cAMP	AC
30	PI3K	RasA	63	PKA	cAMP
31	Rho	EGFRGrb2 EGFRShcGrb2 RalGDS	64	Rsk2	ERK
32	Ral	RalGDS	65	CREB	Rsk2

### **Bases matemáticas.**

Matemáticamente hablando, todos los elementos involucrados en las vías de señalización constituyen un conjunto. Para analizarlo en detalle resulta necesario dividir al conjunto en subconjuntos denominados autosuficientes (SDS, por sus siglas en inglés *self-determined subsets*). Estos subconjuntos se definen como los que tienen dentro del subconjunto todos los elementos de los cuales dependen. Cualquier SDS se puede considerar de manera independiente de todos los demás elementos del sistema. Por otro lado, el SDS puede influir a los elementos externos al SDS. El mismo elemento puede ser incluido en varios SDS, entonces los diferentes SDS forman una estructura por si mismos.

Para aclarar las propiedades matemáticas de los SDS consideramos que la unión o la intersección de dos SDS también es un SDS. La prueba formal de este hecho se presenta en el artículo, (Cárdenas-García, 2002). Todos los SDS forman lo que se llama en matemáticas una topología abstracta. De hecho, la definición de una topología implica determinar cuáles de los subconjuntos del conjunto son "abiertos". Si la unión e intersección de los subconjuntos abiertos es abierta, la topología es definida (O'Neil, 1982).

La idea de esta definición es que si los elementos pertenecen al mismo subconjunto abierto están de alguna manera "cercaños". En particular, para un espacio de tres dimensiones la base de los conjuntos abiertos son esferas sin frontera alrededor de todos los puntos. Todos los conjuntos abiertos se obtienen con uniones de estos subconjuntos de base. Aunque la topología es más útil para los conjuntos infinitos, se define también para los conjuntos finitos. Identificando los SDS con los subconjuntos abiertos de una topología, vemos que la división en SDS define una topología abstracta. El siguiente paso para el análisis es tratar de simplificar la topología para obtener subconjuntos autosuficientes manejables y estructurados de una manera razonable. Para eso notamos, que algunos SDS no tienen otros SDS adentro. Estos son los más simples. Luego siguen los SDS con un SDS adentro etc. Partiendo de esta observación, se puede obtener una estructura piramidal o jerárquica. En los escalones más altos se encuentran los SDS que no tienen otros SDS adentro, el siguiente peldaño corresponderá a los que contienen un SDS, etc. Los escalones más bajos de la pirámide dependen de los escalones más altos y definen, a su vez, a los escalones aun más bajos. La estructura piramidal obtenida por nosotros se muestra en la figura 23. Esta división nos permitió visualizar en conjuntos diferentes, a los elementos relacionados con las proteínas efectoras de Ras implicadas en crecimiento celular por un lado, por otro a los elementos involucrados en la regulación del citoesqueleto, y también los elementos relacionados con el control del ciclo celular. Así, la división en conjuntos hace no sólo más fácil el análisis, sino que también proporciona una subdivisión biológicamente significativa.

Todo el procedimiento antes descrito requiere el uso de la computadora y se hace de una manera automática. Para hacer el cálculo es necesario tener una tabla en la cual se indican los elementos de los cuales depende cada uno de los elementos del sistema. La computadora calcula todos los SDS, luego asigna

los rangos a los subconjuntos en la estructura. El procedimiento se describe con mas detalle en Cárdenas-García, 2002.

Siguiendo con este primer modelo y considerando la posibilidad de que el mismo nos facilitara la detección de blancos terapéuticos potenciales, determinamos los atractores o estados finales para cada elemento. Para obtenerlos primero determinamos los estados iniciales para cada elemento (presencia, ausencia, activa, inactiva) y posteriormente aplicamos reglas lógicas de las interacciones, por ejemplo:

// regla 0, un elemento es activado por otro

// regla 1, un elemento es activado cuando otros 2 están activos

//regla 2, un elemento es activado por cualesquiera de otros 2 elementos activos, pero no activado cuando ambos están activos.

Etc.

Las reglas se aplican a todos los elementos en cada paso. Finalmente, el sistema llega a un equilibrio (no hay más cambios en los estados de los elementos después de aplicar reglas). A estos estados finales al equilibrio se llama atractores de la red, o ciclo límite. Para este análisis seleccionamos 41 proteínas, matemáticamente el número de estados iniciales posibles es de  $2^{41}$ . Eso hace imposible la búsqueda directa de todos los atractores posibles, pues la cantidad de estados iniciales y sus combinaciones es demasiado grande. Después de hacer la división jerárquica se obtuvieron 3 subgrupos. Tomamos el más pequeño. En este caso, el número de estados finales depende de que siempre esté EGF (lo cual es lógico desde el punto de vista biológico). En este caso todas aquellas proteínas que se requiere que estén presentes no importando la combinación de estados iniciales, son consideradas posibles blancos terapéuticos, ya que son esenciales para la célula. Los resultados de este análisis computacional se encuentran en mayor detalle en Cárdenas-García et al. 2000.



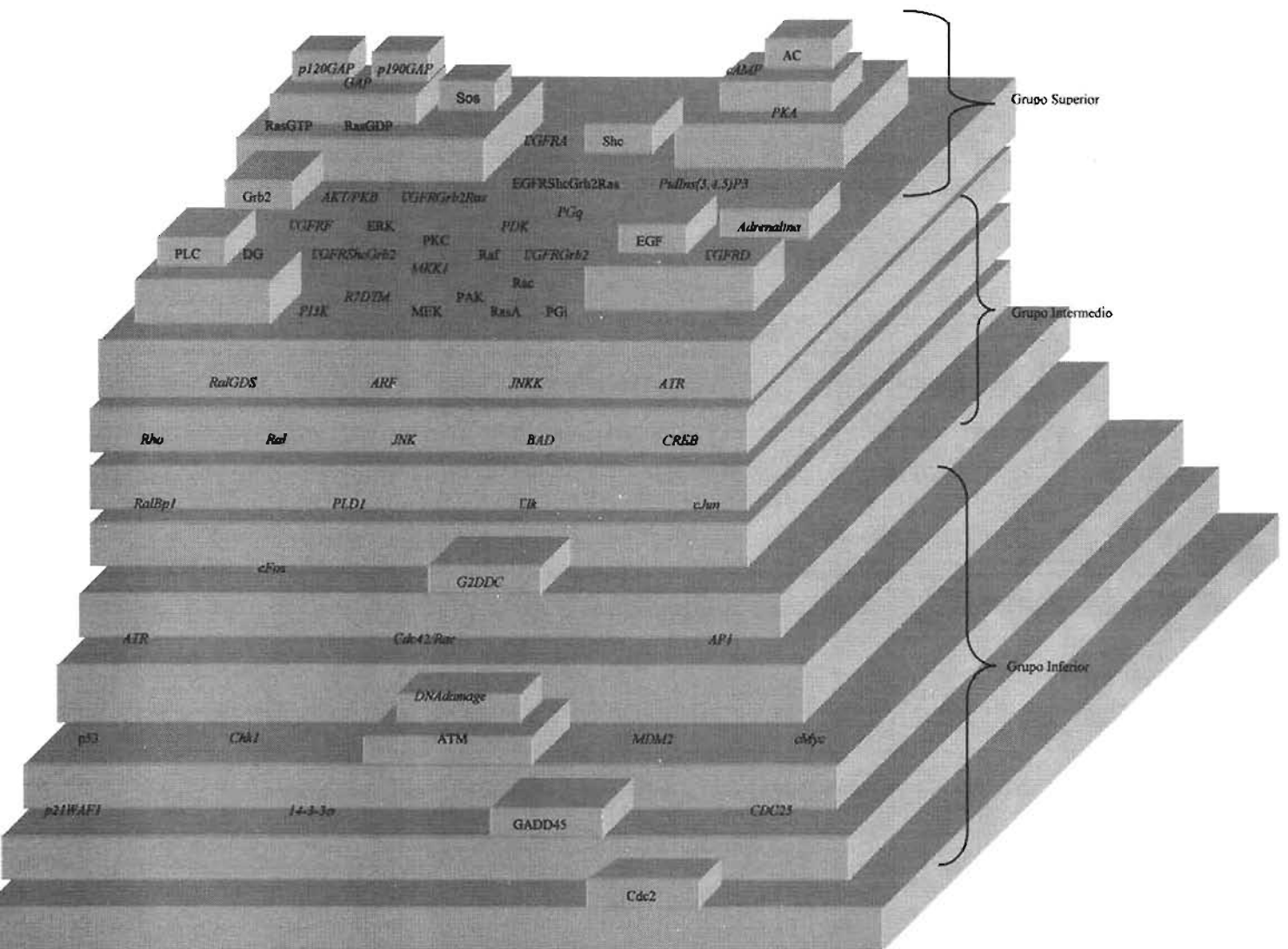


Figura 23. Estructura piramidal o jerárquica obtenida. En el grupo superior se encuentran los ligandos, receptores, cinasas, en el grupo intermedio factores de transcripción, sus monómeros y cinasas de los mismos y en el grupo inferior proteínas cuya respuesta está regulada por factores de transcripción activados por la vía de señalización.

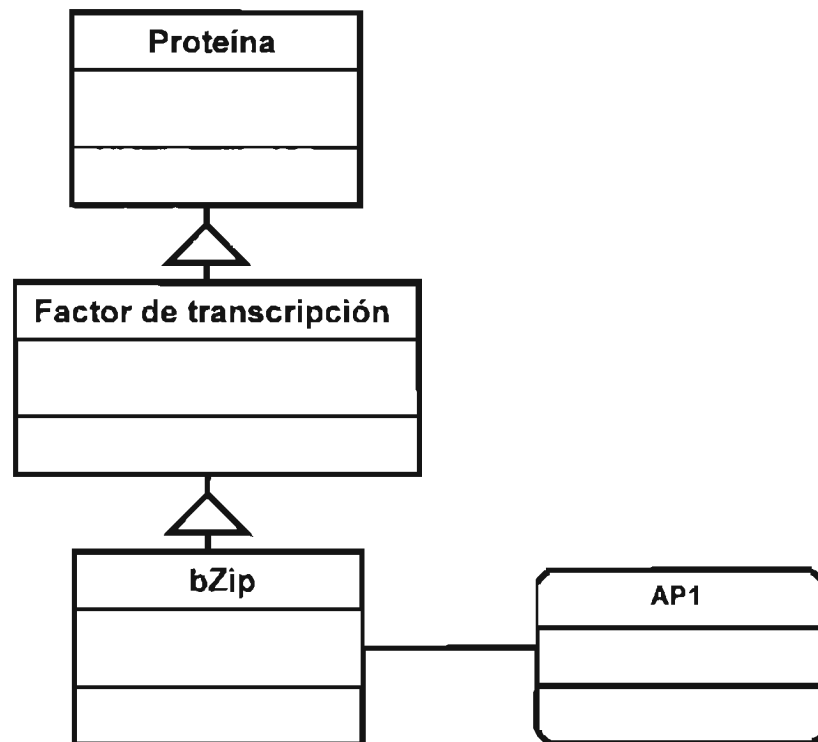
La realización de este modelo nos permitió mediante un método computacional analizar y clasificar un gran número de datos y consecuentemente sugerir posibles blancos terapéuticos; en este caso en particular los blancos terapéuticos serían aquellos que jerárquicamente se encuentran en la base de la pirámide ya que su activación se da en respuesta a un estímulo extracelular, por lo que su ausencia puede evitar que la respuesta celular se cumpla.

## 1.2. Sistema Experto.

Para el sistema experto consideramos a las vías relacionadas con 3 de las proteínas efectoras de Ras: Raf, PI3K y Ral GDS, así como las vías que relacionan a Ras con el control de ciclo celular, específicamente con p53 y ARF. También se consideraron las proteínas 14-3-3 relacionadas con estas vías.

A continuación, el factor de transcripción AP1 se presenta como ejemplo de como se almacena la información para cada uno de los elementos del sistema.

AP1 (Activator Protein 1), es un factor de transcripción tipo cierre o zipper de leucina (bZip), se considera un mensajero nuclear que regula las acciones generadas por las vías de transducción que se encienden en respuesta a factores de crecimiento, hormonas, citosinas y neurotransmisores, (Karin y Smeal, 1992). En el sistema experto, AP1 es un objeto que pertenece a la megaclass: proteínas, clase: factor de transcripción, subclase: básicos tipo cierre o zipper de leucina bZip. Como se indica en el siguiente daigram de clases de la figura 24.



**Figura 24.** Diagrama de clases que representa la clasificación de AP1. AP1 es un objeto de la clase bZip, la clase bZip pertenece a la clase factor de transcripción y la clase factor de transcripción pertenece a la clase proteína.

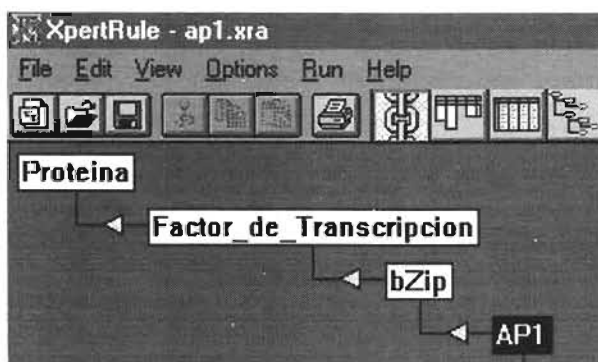


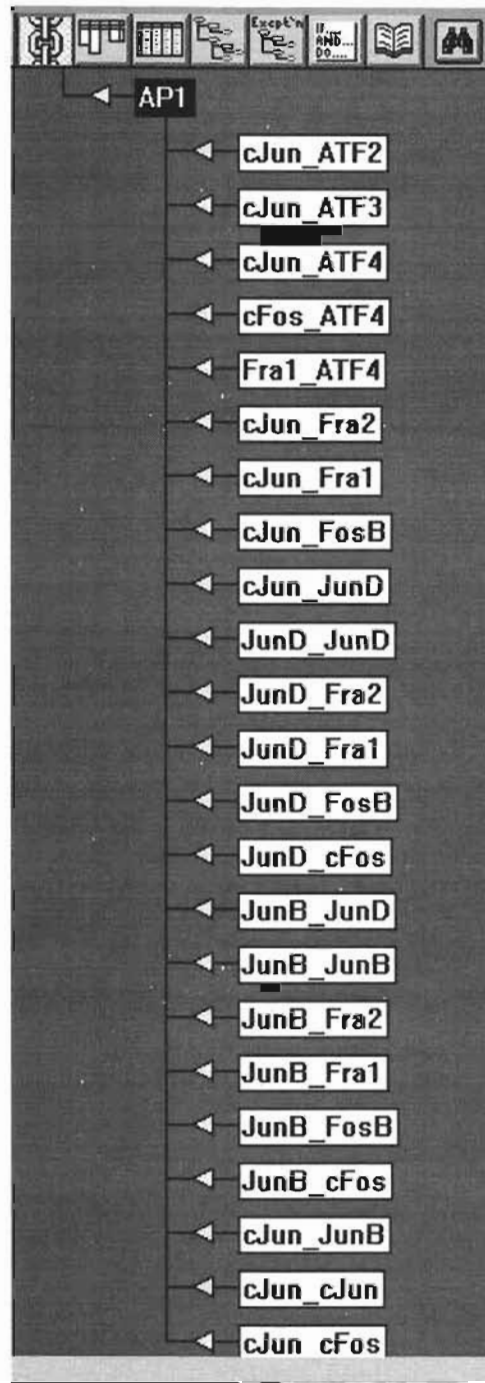
Figura 25. Esta es una imagen de la visualización en pantalla de la clasificación de AP1. AP1 es un objeto de la clase bZip, bZip es una subclase de la clase factor de transcripción y factor de transcripción es una subclase de la clase proteínas.

El factor de transcripción AP1 es una proteína dimérica que puede estar formada por proteínas de la familia Jun, proteínas de la familia Fos y proteínas ATF.

Dímeros que pueden formar a AP1:

cJun-cFos	cJun-JunB	JunD-FosB	JunB-FosB	cFos-FosB
cJun-cJun	JunB-JunB	cJun-Fra1	JunD-cFos	cFos-Fra1
cJun-JunD	JunB-JunD	JunB-Fra1	cJun-ATF4	cFos-Fra2
JunD-JunD	cJun-FosB	JunD-Fra1	cFos-ATF4	FosB-FosB
JunD-Fra2	JunB-cFos	cJun-ATF2	Fra1-ATF4	
cJun-Fra2	JunB-Fra2	cJun-ATF3	cFos-cFos	

En este caso el objeto AP1 contiene como subobjetos a cada uno de los posibles dímeros (figura 26).



**Figura 26.** Esta figura muestra la visualización en pantalla de los dímeros que pueden dar lugar a AP1. En este caso el objeto es AP1 y sus subobjetos son cada uno de los dímeros.

La actividad de AP1 es regulada por fosforilación, tanto de la proteína preexistente como de aquella sintetizada *de novo*. Las cinasas JNK, FRK y MAPK son las más importantes en dicha regulación. Fos es fosforilado en la treonina 232 por una cinasa específica la FRK, fosforilación que potencia la transcripción del gen. La regla que corresponde a este hecho es la siguiente.

	Comment	Weight	Rules	Pattern Rules
1	Estado de Fos	1.00	If FRK = Activa And FosTreonina232 = Presente Then cFos = FosforiladaTre232	

Figura 27. Esta figura muestra la visualización en pantalla de una regla de inferencia. En este caso para que cFos este fosforilado en la treonina 232 es necesario que la cinasa FRK esté activa y treonina 232 presente.

Otras formas de regulación de AP1, implícitas en la función de Jun y Fos, dependen de la estabilidad de los heterodímeros

Fos-Fos < Jun-Jun < Fos-Jun

En el sistema experto esto se programa dándole diferentes pesos a cada uno de los dímeros, dependiendo de la disponibilidad del monómero y de la estabilidad de los dímeros, como se ejemplifica en la figura 28.

	Comment	Weight	Rules
1	AP1 Presente	1.00	If cJun_cFos = Yes Then AP1 = Yes
2	AP1 Presente	0.85	If cJun_ATF2 = Yes Then AP1 = Yes
3	AP1 Presente	0.87	If cJun_ATF3 = Yes Then AP1 = Yes
4	AP1 Presente	0.88	If cJun_ATF4 = Yes Then AP1 = Yes
5	AP1 Presente	0.50	If cFos_ATF4 = Yes Then AP1 = Yes
6	AP1 Presente	0.25	If Fra1_ATF4 = Yes Then AP1 = Yes
7	AP1 Presente	0.89	If cJun_Fra2 = Yes Then AP1 = Yes
8	AP1 Presente	0.90	If cJun_Fra1 = Yes Then AP1 = Yes
9	AP1 Presente	0.91	If cJun_FosB = Yes Then AP1 = Yes
10	AP1 Presente	0.60	If cJun_JunD = Yes Then AP1 = Yes
11	AP1 Presente	0.65	If JunD_JunD = Yes Then AP1 = Yes

Figura 28. Esta figura muestra a algunas de las reglas de inferencia que determinan la presencia de AP1, en ella se observan los diferentes pesos de cada regla, estos pesos dependen de la estabilidad de los dímeros.

Una vez ejemplificado el almacenaje de la información, a continuación se muestra cómo el usuario final realiza una sesión de trabajo.

### Introducción a la sesión de trabajo.

La sesión comienza con una bienvenida al sistema experto. En esta pantalla el usuario debe seleccionar el tipo celular con que prefiere trabajar, figura 29.

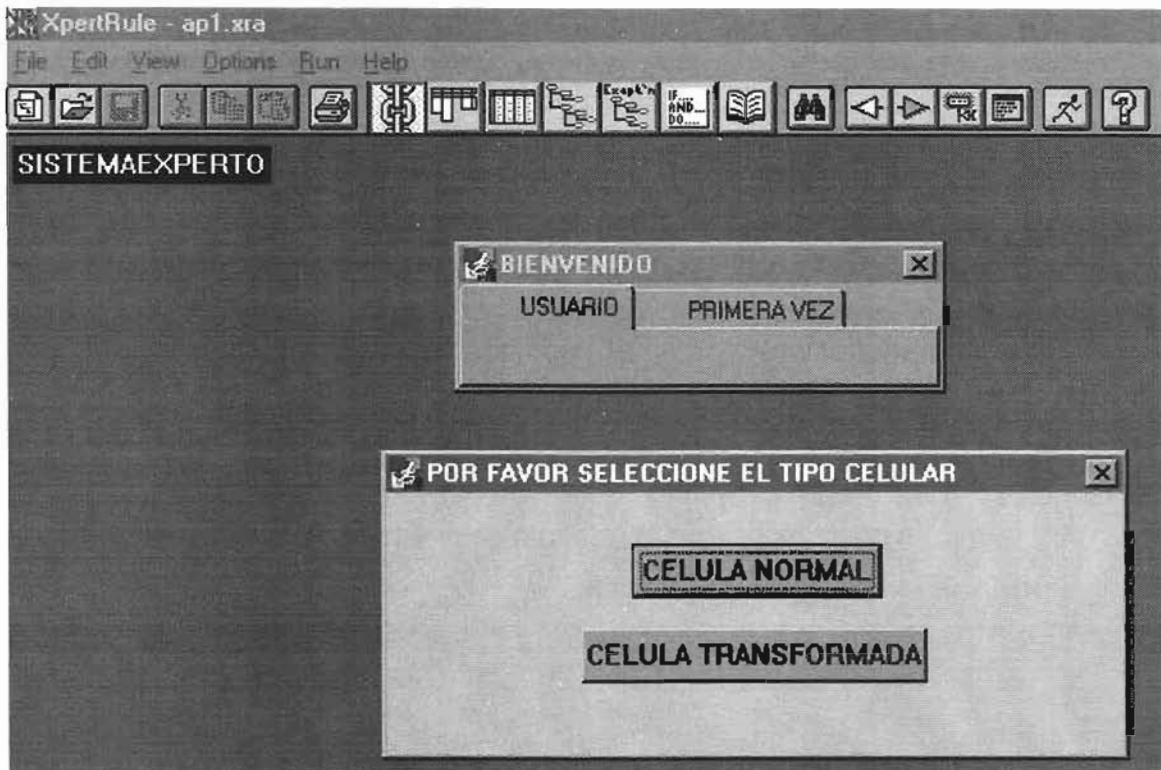


Figura 29. Esta imagen corresponde a la ventana de bienvenida en la que el usuario debe elegir el tipo celular con el que desea trabajar.

Después de esta presentación, se inicia la tarea de introducción de datos por parte del usuario, figura 30.

1. La asignación de un nombre que identifique la sesión.
2. Determinar el tipo de información con la que trabajará. Cuantitativa (si asignara valores numéricos como concentración a todos los elementos), semicuantitativa (si asignara valores numéricos como concentración a algunos de los elementos) o no cuantitativa (no se asignarán valores numéricos a ninguno de los elementos).

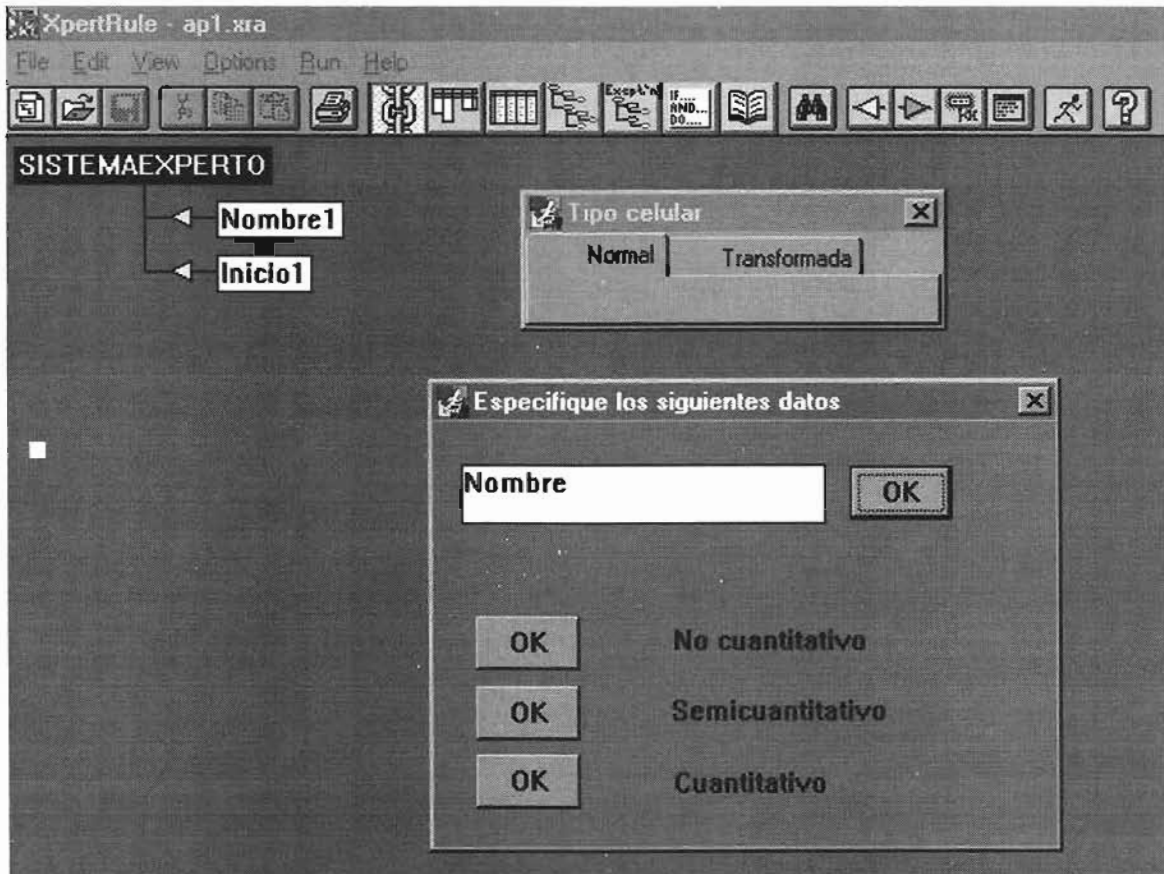


Figura 30. Esta imagen corresponde a la ventana en la que el usuario debe asignar el nombre a su sesión y determinar el tipo de sesión.

A continuación el usuario debe indicar los elementos de la base que estarán presentes en su sesión, figura 31.

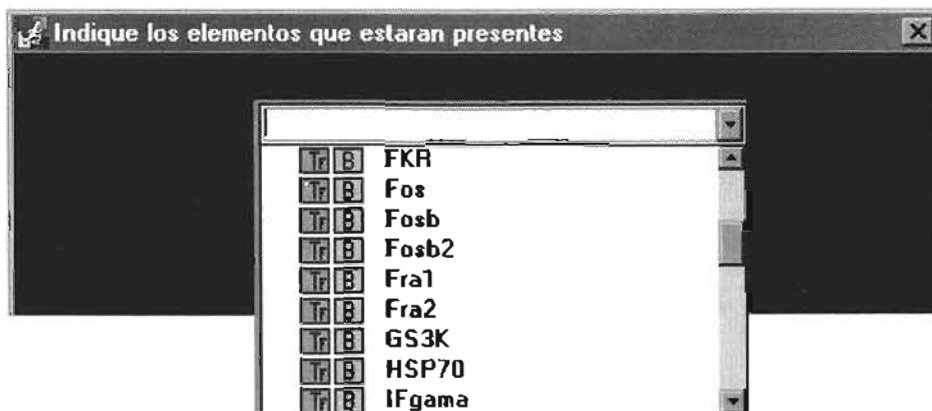


Figura 31. Esta imagen corresponde a la ventana en la que el usuario debe seleccionar el o los elementos que desea estén presentes.

A continuación se indica para cada elemento seleccionado si está o no activo, en el caso de una sección cuantitativa o semicuantitativa se debe asignar una concentración a cada elemento. Con base en la información proporcionada por

el usuario, el sistema experto determina haciendo uso del motor de inferencia si la combinación de elementos da lugar o no a una célula viable. El sistema nos indica si la célula sobrevive o muere, mostrando una ventana como la que se encuentra en la figura 32.

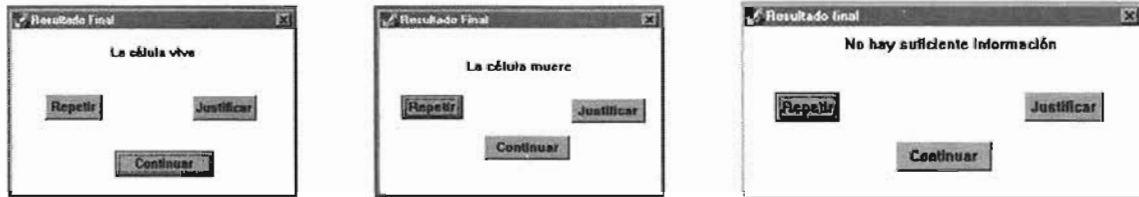


Figura 32. Esta imagen corresponde a las posibles ventanas que aparece con el resultado final, en ellas se pueden seleccionar diferentes opciones. Repetir si se quiere modificar alguno de los parámetros y repetir la operación, justificación si se quiere saber el por qué del resultado final y continuar para continuar o salir.

De manera automática, el sistema elimina todas las interacciones en las que participa el elemento eliminado o elementos eliminados, e indica el resultado final. De esta forma es posible seleccionar posibles blancos terapéuticos, ya que aquellos elementos que sean necesarios para la supervivencia celular serán posibles blancos, en el caso que la idea sea eliminar a la célula. A continuación se presentan dos ejemplos. En el primer ejemplo se eliminó a la proteína Ras y en el segundo a la proteína 14-3-3σ. Al seleccionar la opción justificación, el sistema nos indicará cuál fue la vía alterna que la célula utilizó y logró sobrevivir para el caso de Ras figura 33. Para el caso de la proteína 14-3-3 la célula no sobrevive, la justificación se presenta en la figura 34.

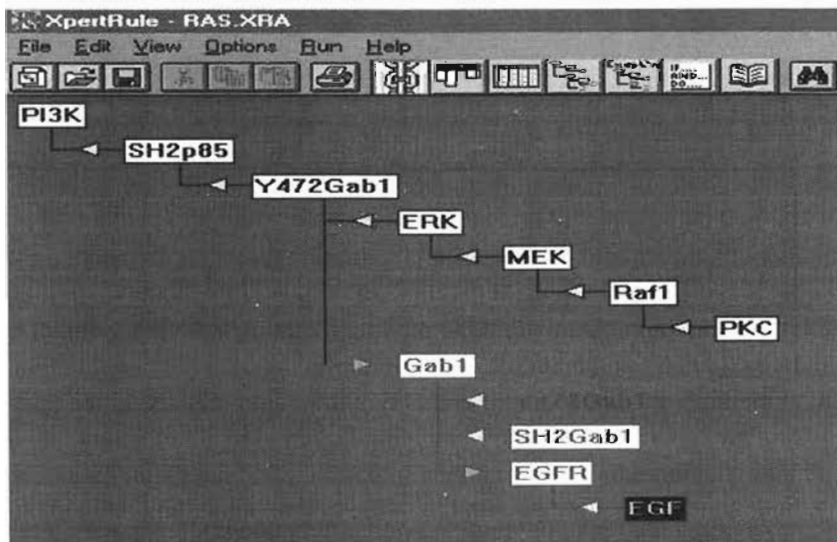
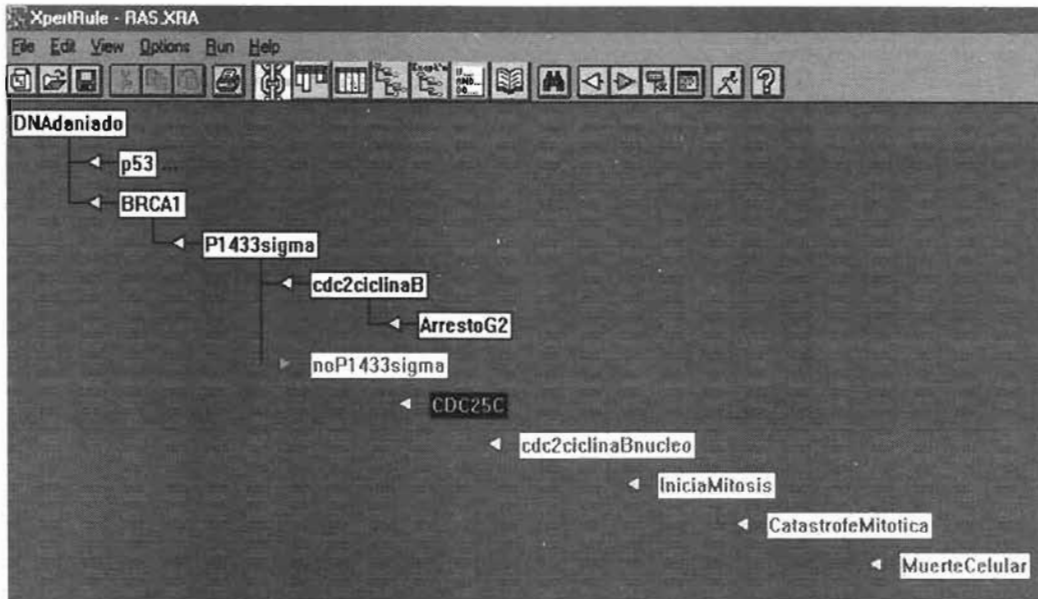


Figura 33. Esta imagen corresponde a la ventana que justifica porque la célula sigue viviendo en ausencia de Ras, ya que dos de sus proteínas efectoras Raf1 y PI3K son activadas en su ausencia.





**Figura 34.** Esta imagen corresponde a la ventana que aparece con la justificación, de porque en ausencia de la proteina 14-3-3  $\sigma$ , la célula muere. Al estar libre el complejo cdc2ciclinaB, la fosfatasa CDC25C lo desfosforila y el complejo entra a núcleo iniciandose la mitosis, lo que lleva a la célula a una catástrofe mitotica y finalmente a la muerte.

## 2.0 RESULTADOS EXPERIMENTALES.

Los resultados experimentales de esta tesis serán enviados para su publicación una vez que se tenga la patente de los oligonucleótidos utilizados.

El sistema experto desarrollado para buscar posibles blancos terapéuticos, indica que la ausencia de la proteína 14-3-3  $\sigma$  (figura 34), da lugar a la muerte celular, por lo que decidimos corroborar este resultado experimentalmente.

Para eliminar a la proteína 14-3-3  $\sigma$  empleamos oligonucleótidos que forman una triple hélice en la región promotora del gen que codifica para esta proteína, evitando de esta manera su transcripción y por ende su traducción.

Para comprobar que los TFO diseñados por nosotros realmente forman una triple cadena con su secuencia blanco realizamos experimentos que confirman la formación de triplex *in vitro*.

### 2.1 Formación de triplex *in vitro*.

#### 2.1.1 Desnaturalización térmica.

El ADN de doble cadena en solución al incrementar la temperatura se desnaturaliza, es decir la doble cadena se abre, dando lugar a un incremento en la absorbencia a 260 nm, llamado efecto hipercrómico. Al someter a una triple cadena al mismo incremento de temperatura, el incremento en la absorbencia es mayor, como se muestra en la figura 35 en una gráfica de densidad óptica (DO) contra temperatura.

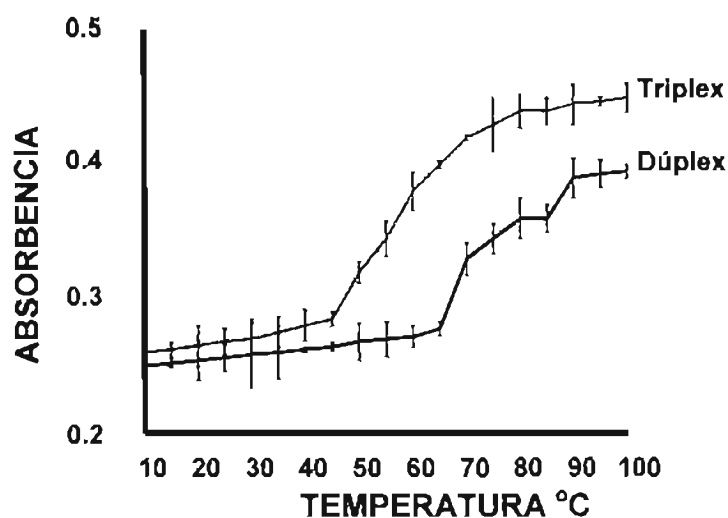
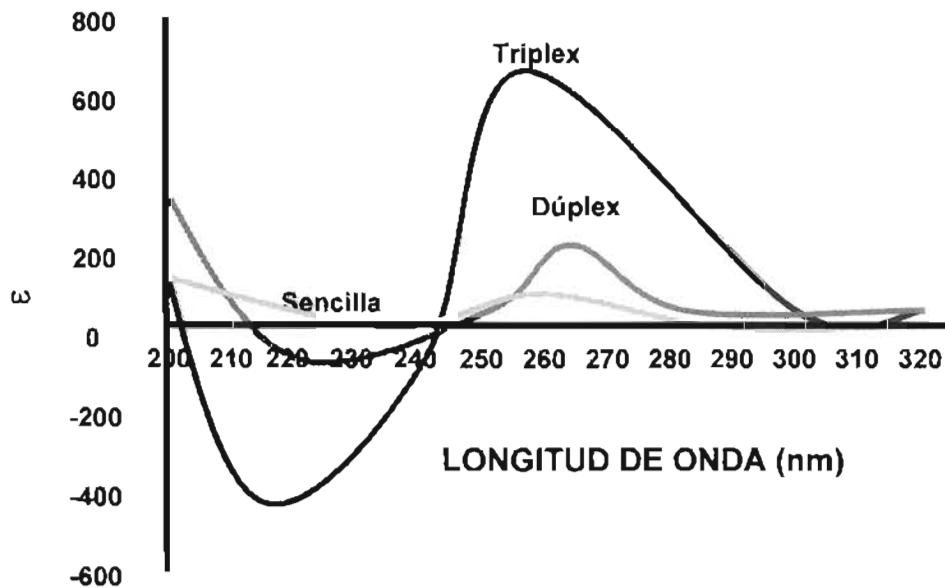


Figura 35. En esta gráfica de temperatura en °C, contra absorbencia a 260 nm podemos observar ADN de doble cadena curva inferior y de triple curva superior. A partir de los 50 grados centígrados se observa un incremento en la absorbencia de la triple cadena con respecto a la doble, estos resultados coinciden con los reportados en la literatura y demuestran la existencia de una triple cadena. Las gráficas representan un promedio de 3 experimentos, las barras indican la desviación estándar. El mismo tipo de experimento se realizó con los oligonucleótidos MCG1, MCG2, MCG3 y MCG4. En todos los casos se obtuvieron espectros similares. La curva que se muestra aquí corresponde al oligonucleótido MCG3.

### 2.1.2 Dicroísmo circular.

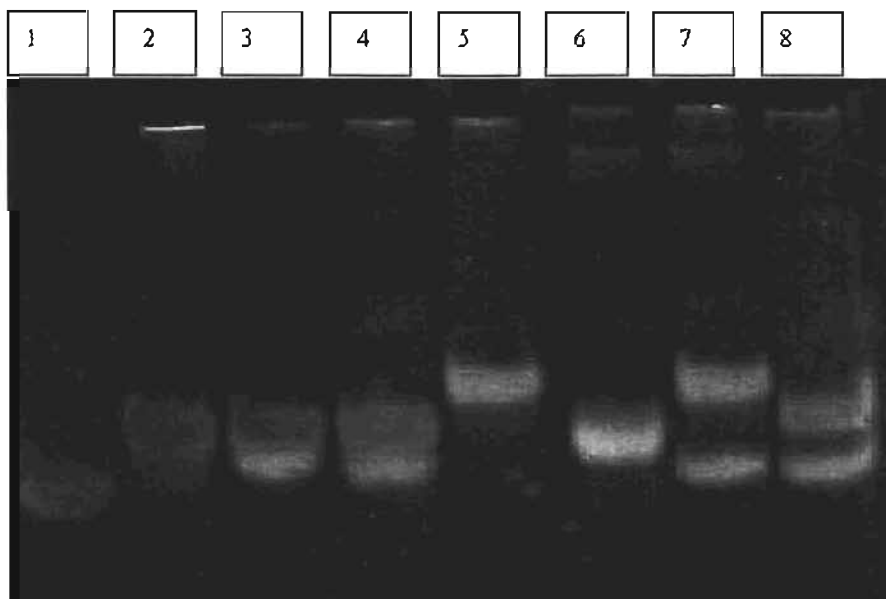
El dicroísmo circular es la diferencia de la desviación de la luz a la derecha y la desviación de la luz a la izquierda, de tal manera que una molécula entre más asimétrica sea presentará una mayor elipticidad negativa y/o positiva. Al comparar los espectros de dicroísmo circular de la cadena sencilla, el dúplex y el triplex, se observa en el triplex incrementó la elipticidad negativa y positiva, con respecto al dúplex y a la cadena sencilla, como se muestra en la figura 36. El hecho de que en la triple cadena la elipticidad sea mayor que en el dúplex y en la cadena sencilla indica una mayor asimetría de la molécula.



**Figura 36.** En esta gráfica elipticidad contra longitud de onda, se muestran los espectros de dicroísmo circular para la cadena sencilla (gris claro), el dúplex (gris) y el triplex (negro). Se observa un incremento en la elipticidad negativa y positiva de la triple cadena con respecto a la doble y de la doble con respecto a la sencilla. Los datos obtenidos coinciden con los reportados en la literatura y demuestran la presencia de una triple cadena. El mismo tipo de experimento se realizó con los oligonucleótidos MCG1, MCG2, MCG3 y MCG4. En todos los casos se obtuvieron espectros similares. Los espectros que aquí se muestran corresponden al oligonucleótido MCG3.

### 2.1.3 Movilidad en gel.

Para observar diferencias en el patrón de migración de cadena sencilla, doble o triple de ADN, se realizó una electroforesis en gel de poliacrilamida en condiciones nativas, figura 37. Se utilizó la misma cantidad de muestra en cada carril (20  $\mu\text{g}$ ). Debido a que los TFO tienen fluoresceína en su extremo 5' y que al gel de acrilamida se le adiciona el agente intercalante bromuro de etidio que también fluoresce al exponerlo a la luz ultravioleta, fue posible observar la formación de triplex, así como la diferencia en la migración de cadena sencilla, doble y triple. En la fotografía de la figura 37 se observa en el carril 5 una triple cadena más estable con respecto a las que se presentan en los carriles 3, 4 y 6. El carril 5 corresponde al triplex formado entre el oligonucleótido MCG3 y su dúplex correspondiente.



**Figura 37.** En esta imagen podemos observar en los carriles (1) una cadena sencilla, (2) una doble cadena, (3) triple cadena entre MCG1 y su dúplex correspondiente, (4) triple cadena entre MCG2 y su dúplex correspondiente, (5) triple cadena entre MCG3 y su dúplex correspondiente, (6) triple cadena entre MCG4 y su dúplex correspondiente, (7) la mezcla resultante de la incubación de los oligonucleótidos MCG3, MCG5 y el dúplex correspondiente a MCG3, (8) la mezcla resultante de la incubación de el oligonucleótido MCG5 y el dúplex correspondiente a MCG3. El color verde corresponde a la fluoresceína y el naranja al bromuro de etidio. Se confirma la formación de una triple cadena *in vitro*, por lo menos entre la doble cadena correspondiente al oligonucleótido MCG3 y el oligonucleótido MCG3 (carril 5). La triple cadena es muy inestable o se pierde durante la electroforesis en el caso de los oligonucleótidos MCG1 y MCG2 y su respectiva doble cadena correspondiente (carriles 3 y 4). El oligonucleótido MCG4 y su doble cadena correspondiente forman un triplex parcialmente estable con respecto a formado por MCG3 y su doble cadena correspondiente, carril 6.

Hasta aquí queda demostrado que al menos uno de los oligonucleótidos diseñado (MCG3 carril 5) es capaz de formar una triple cadena *in vitro*, sin embargo hay que demostrar que *in vivo* es capaz de inhibir la transcripción de su gen blanco, por lo que realizamos los siguientes experimentos en líneas celulares en cultivo.



## MCG3

```

Homo sapiens t---aggaatctggccttag-agcctgcccctgtttctctctcagccaggcatagccat
|      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Mus musculus tctttaggaatctcgctc-gcagcctgcc-----tct-----ag-ccag-cagagccat

Homo sapiens agtcaagcccagcaggttctct-caggagctgtctggggtgttgatgggtggatgacgctg
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Mus musculus aggccagcctagcaggtt-ccttcaggagctatctggg-tgttgatg----a---cgctg

Homo sapiens ctgaacaagt ttggtagctgttctaagcacaactggcttgatactgttcccacg
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Mus musculus ctgaacaagt ttggcaactgctctaagcacaacgggcttgacgctgttcccacg

Homo sapiens aggctgctaaccctctctctctctctctctcttctgctgctctgcccgggagtc-tccagtgtg
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Mus musculus a--tttgctaaccctctctctctctctctctcttggcatcctgcc-gcgateatccagctctg

Homo sapiens tgcgggggcttaaggacctctgaggaccgctgctctctgctctccaggaaatggcctgg
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Mus musculus tgcgggggcttaaggagctctgaggaccgacctctctgctctccaggaaatggcctgg

Homo sapiens ggggagccaggcaccccggcacctccacctgctaactgtggcccatctgccaccatctg
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Mus musculus ggggagccaggcaccccggcacctccacctgctaactgtggcccatctgccaccatctg

Homo sapiens tgcctacagggtctg-ccc--ccc--agcctgcccggcctgtgtgctctctag-gac-cc
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Mus musculus tgcctacagggtctgccccgctcccagcctgtctactgtgtgtctctagagccctcc

Homo sapiens c-at-ag--g-gggcaggggctggcctctttgccc--attcccg-ctccatgccggcca
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Mus musculus ccaccaacagttggcagggcctgggctcta-gtccaggaaccccactccttgctggcca

Homo sapiens gagtgtgaaaagccataacgcacgcagccatcagcacaataatgtgactctacgctgata
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Mus musculus gagtgtgaaaagccata-c-c--gcagcc-tcagcacagtactgtaacact--gctga-a

Homo sapiens tgct--ccctctc-tctccactgactt-ccccttcccg-gatttggtaggtgtcaagac
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Mus musculus c-cttgccctcccctctccct--cgtgcccctcccagatttggtaggtgtggagtc

Homo sapiens t---aggaatctggccttag-agcctgcccctccaccctcagatcaggcat
|      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Mus musculus tctttaggaatctcgctc-gcagcctgcc-----acc-----ag-tcag-cag

Homo sapiens caagcccagcaggttctct-caggagctgtctggggtgttgatgggtggat
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Mus musculus ccagcctagcaggttctctcaggagctatctggg-tgttgatg----a---tgctg

Homo sapiens ctgaacaagt ttggtagctgttctaagcacaactggcttgatactgttcccacg
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Mus musculus ctgaacaagt ttggcaactgctctaagcacaacgggcttgacgctgttcccacg

```

Figura 38. Alineamiento de la región promotora de los genes que codifican para la proteína 14-3-3  $\sigma$  de humano (*Homo sapiens*) y de ratón (*Mus musculus*). Se indican los sitios a donde van dirigidos los oligonucleótidos MCG1, MCG2, MCG3 y MCG4. Los nucleótidos sombreados corresponden a las cajas TATA y CAT.

Las células de las diferentes clonas fueron transfectadas con el oligonucleótido y el agente transfectante Metafectane. Se determinó viabilidad por densitometría con el método de cristal violeta a las 0, 24, 48 y 72 h después de la transfección, la gráfica obtenida se muestra en la figura 39. Se encontró que ni el oligonucleótido o el agente transfectante Metafectane son tóxicos para las células, ya que no observamos disminución en la viabilidad de las mismas, aún incubando con 15  $\mu$ g de oligonucleótido por mililitro de medio de cultivo durante 72 h.

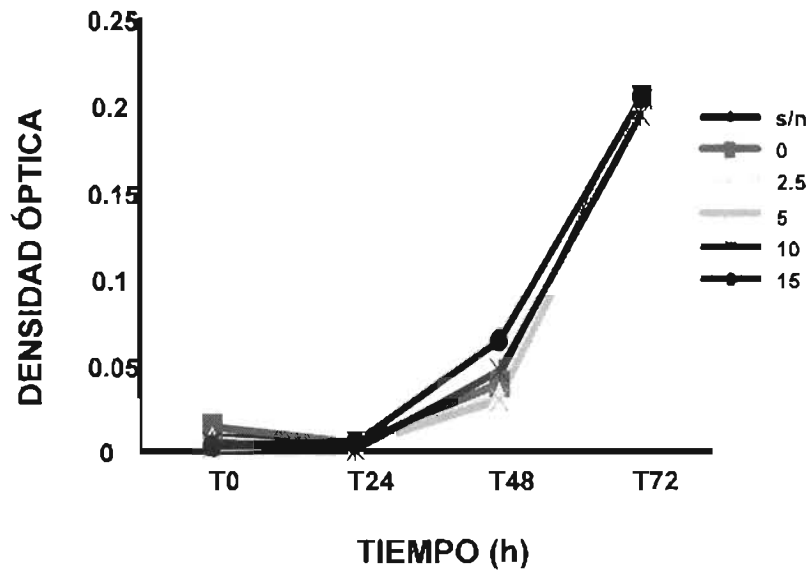
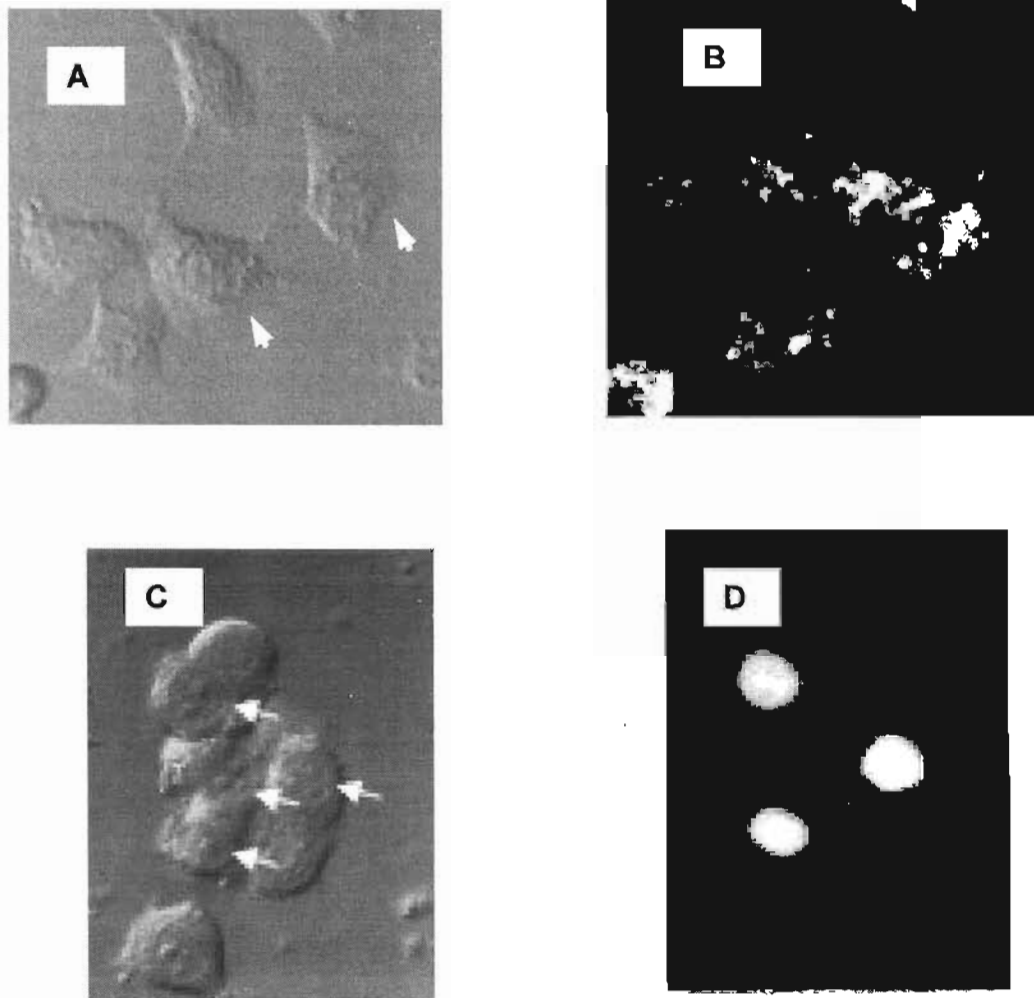


Figura 39. En esta curva de crecimiento que nos muestra densidad óptica contra tiempo, podemos observar que el crecimiento es similar con diferentes concentraciones del oligonucleótido. Los datos obtenidos fueron con la clona C10, las concentraciones de oligonucleótido son en  $\mu\text{g}$  por mililitro de medio de cultivo y metafectane en una proporción 1/5.

Una vez que se determinó que los oligonucleótidos no son tóxicos por sí mismos en células de fibroblastos de ratón, procedimos a determinar el efecto de uno de los oligonucleótidos anti-Hoogsteen (MCG 3), por presentar mayor estabilidad de triplex *in vitro*. Los experimentos se realizaron en la línea celular humana de cáncer de mama MCF 7.

### 2.2.1 Internalización del oligonucleótido.

La internalización de oligonucleótido se puede ver ya que este tiene fluoresceína en su extremo 5'. En las fotografías de la figura 40 observamos que 24 h después de la transfección, el oligonucleótido se encuentra en el citoplasma celular y a las 48 h en el núcleo. Las fotografías A y B corresponden a 24 h y las C y D a 48 h.



**Figura 40.** Fotografías de las células MCF 7 transfectadas con el oligonucleótido MCG3. "A" y "C" Imagen de las células con luz visible, 24 h y 48 h, después de la transfección, "B" y "D" Imagen de las células 24 h y 48 h, después de la transfección empleando lámpara de mercurio. Las flechas señalan a las células fluorescentes.

### **2.2.2 Efecto de la concentración del oligonucleótido MCG3 sobre el crecimiento de la línea celular MCF7 en cultivo.**

Para determinar la concentración de oligonucleótido suficiente para la observación de un cambio en el crecimiento celular, es necesario primero causar daño al ADN ya que el gen que codifica para la proteína 14-3-3 $\sigma$  se activa en respuesta a daño a ADN. Para causar daño a ADN empleamos al fármaco adriamicina. En la figura 41 se presenta la curva de crecimiento celular en respuesta a diferentes concentraciones de adriamicina. Elegimos 5  $\mu$ M ya que a esta concentración las células presentan un menor crecimiento con respecto al control pero no todas mueren, por lo que el porcentaje que sobrevive debe presentar daño a ADN.



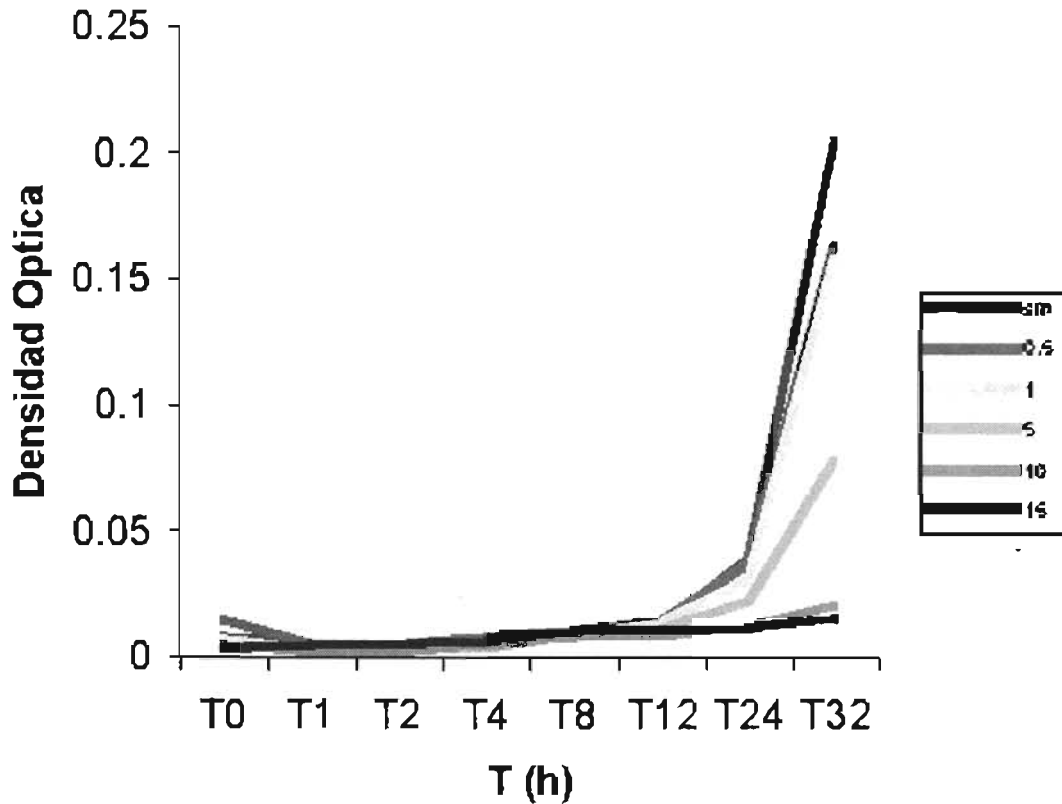
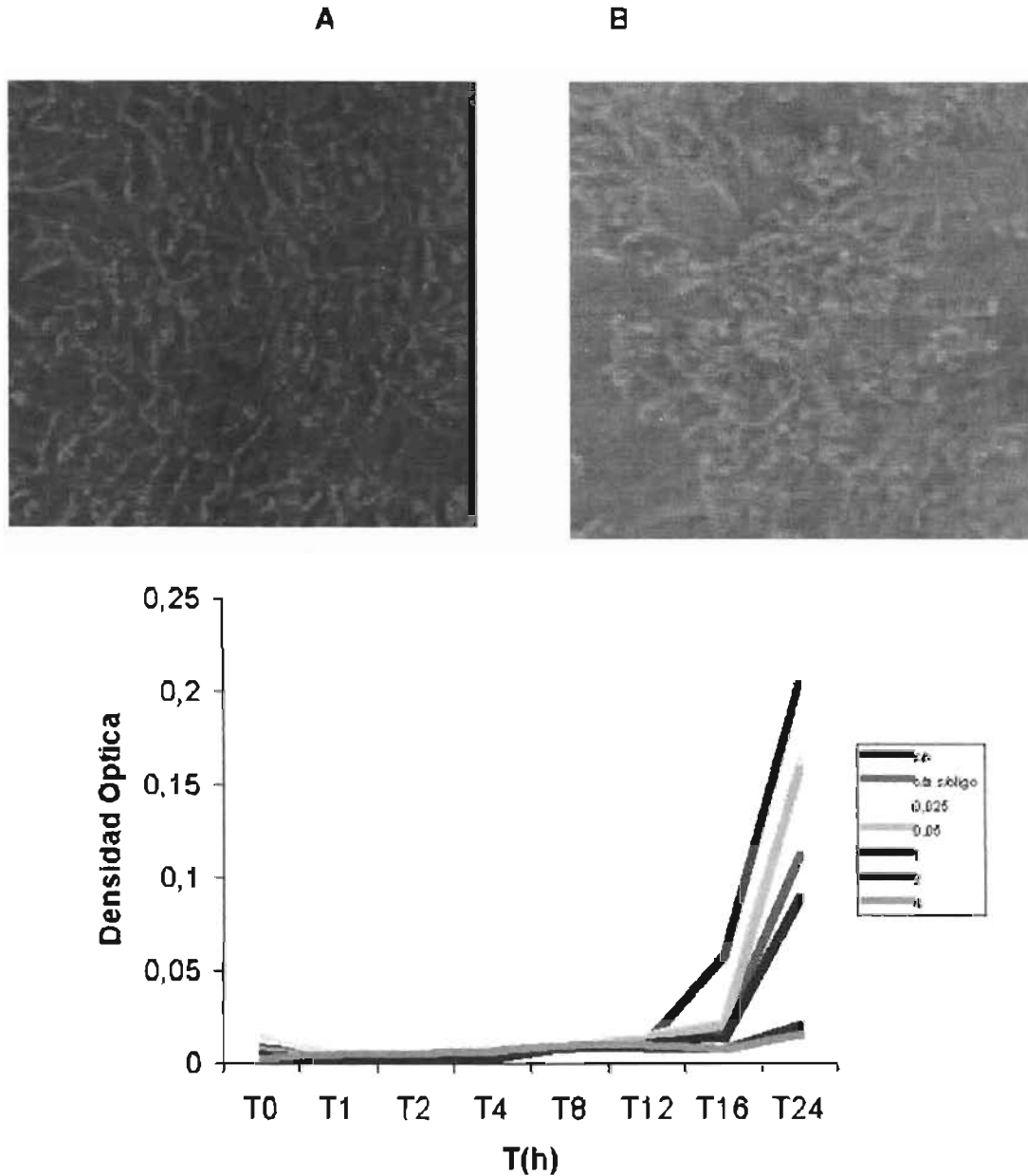


Figura 41. Curva de crecimiento empleando diferentes concentraciones  $\mu$ molares de adriamicina, a diferentes tiempos de incubación. La densidad óptica representa viabilidad celular.

Empleando 5  $\mu$ M de adriamicina después de transfectar a las células con diferentes concentraciones del oligonucleótido MCG3, se realizó otra curva de crecimiento, que se presenta en la figura 42. En esta gráfica observamos que la concentración de oligonucleótido necesaria para observar un cambio en el crecimiento celular es de 1  $\mu$ g/ml, 16 h después de adicionar adriamicina. En esta gráfica también se observa que el efecto del oligonucleótido sobre la viabilidad celular, depende de su concentración. La viabilidad celular presenta una disminución del 80 % con respecto al control al emplear la concentración de oligonucleótido de 4  $\mu$ g/ml de medio de cultivo.



**Figura 42.** Curva de crecimiento. Esta curva se realizó empleando diferentes concentraciones en  $\mu\text{g}$  de oligonucleótido por ml de medio de cultivo. Se grafica densidad óptica que corresponde a viabilidad celular, con respecto al tiempo. El oligonucleótido se adicionó 48 h antes de adicionar adriamicina en una concentración final de  $5 \mu\text{M}$ , el tiempo corresponde al tiempo de incubación con adriamicina. Se observa que la disminución en la población de células vivas depende de la concentración de oligonucleótido, ya que a mayor concentración, menor densidad óptica. En los recuadros A y B se presentan dos fotografías de las células en un microscopio óptico con un objetivo de 100X. (A) corresponde a la imagen de 1hr después de adicionarles adriamicina y (B) a 12h después de adicionarles adriamicina.

### 2.2.3 Efecto del oligonucleótido MCG3 sobre el ciclo celular de la línea celular MCF7 en cultivo.

Cuando hay daño a ADN, los niveles de p53 se incrementan, p53 activa la transcripción de 14-3-3  $\sigma$ . 14-3-3  $\sigma$  se une a cdc2 fosforilada y esto impide que la fase G2 de ciclo celular continúe. En ausencia de 14-3-3  $\sigma$  cdc2 es desfosforilada por Cdc25, cdc2 desfosforilada se une a la ciclina B y se inicia la mitosis, figura 43.

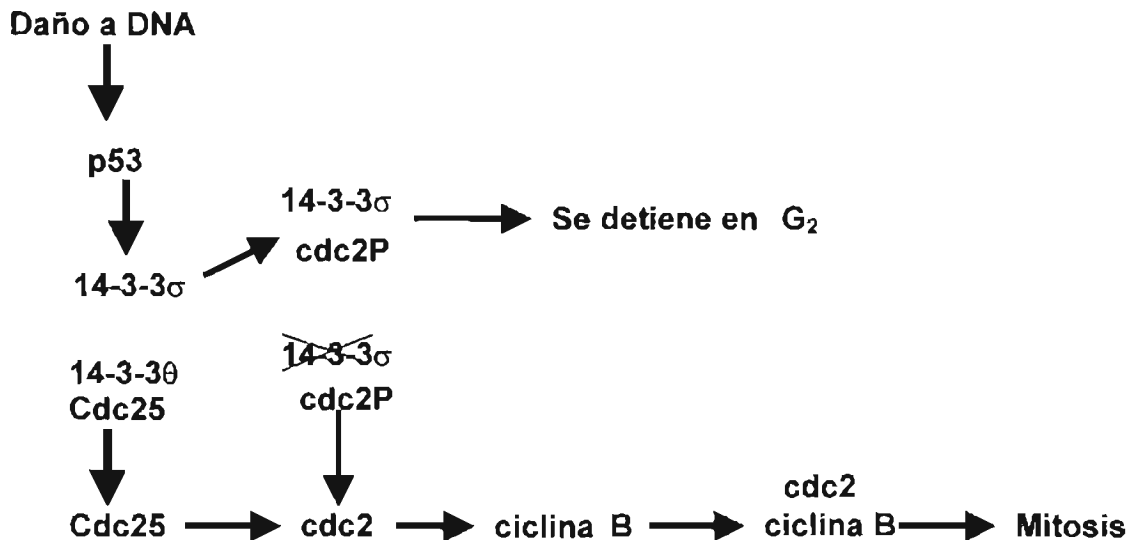
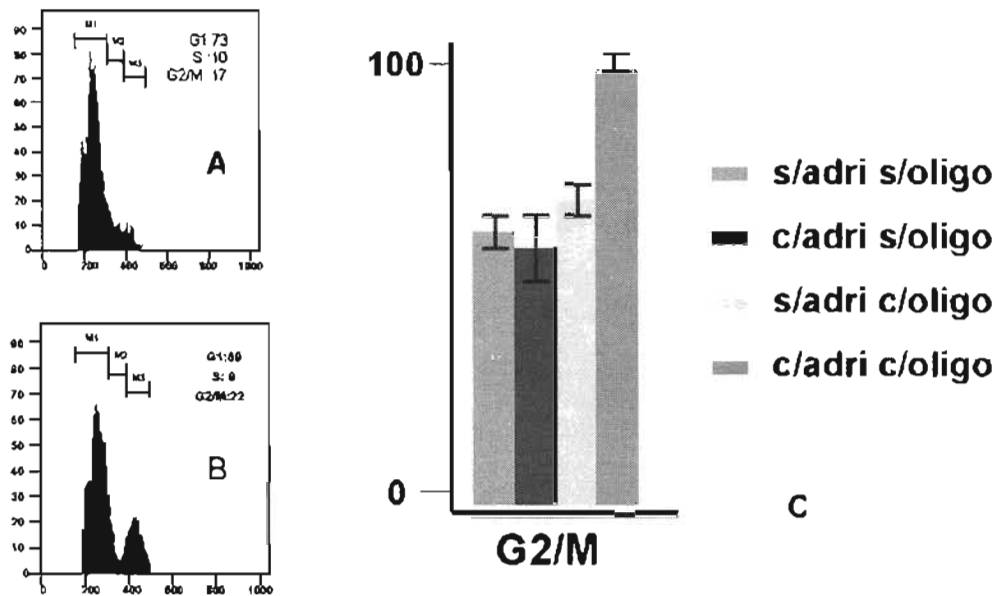


Figura 43. Esquema del papel de la proteína 14-3-3  $\sigma$  en el control del ciclo celular. Cuando hay daño en DNA, los niveles de p53 se incrementan, p53 activa la transcripción de 14-3-3  $\sigma$ . 14-3-3  $\sigma$  se une a cdc2 fosforilada y esto impide que la fase G2 de ciclo celular continúe. En ausencia de 14-3-3  $\sigma$ , cdc2 es desfosforilada por Cdc25, cdc2 desfosforilada se une a la ciclina B y se inicia la mitosis.

Con el fin de determinar si existe un incremento en la población de células en G2/M, se calculó el porcentaje de células en cada una de las fases del ciclo celular por citometría de flujo. En la figura 44 se presenta un histograma de barras del índice mitótico relativo obtenido en células tratadas o no con adriamicina y transfectadas o no con el oligonucleótido, se observa un incremento de aproximadamente el 20 % de células en mitosis, con respecto al control. En la figura 44 se presenta un recuadro que corresponde a un ejemplo de las cuentas obtenidas en el citometro de flujo. Se observa un incremento de la población en G2/M en las células transfectadas con el oligonucleótido y tratadas con adriamicina con respecto a las células control. Las población de las células en S en ambos casos es similar y hay una disminución de la población en G1 con respecto al control.

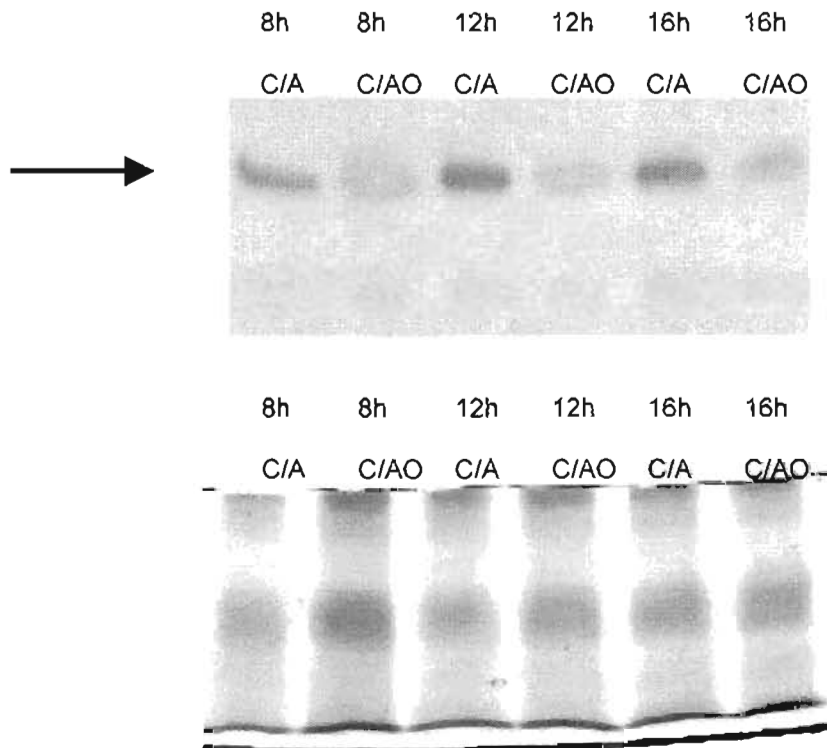


**Figura 44.** Índice mitótico relativo de células con diferentes tratamientos: primera barra células control (sin adriamicina, ni oligonucleótido), segunda barra células tratadas 12 h con adriamicina 5  $\mu$ M, no transfectadas con el oligonucleótido, tercera barra células transfectadas con 1  $\mu$ g de oligonucleótido por ml de medio de cultivo, sin tratamiento con adriamicina y cuarta barra células transfectadas con 1  $\mu$ g de oligonucleótido por ml de medio de cultivo, 48 h antes de tratar por 12 h con adriamicina 5  $\mu$ M. El 100% corresponde a las células tratadas con adriamicina y transfectadas con el oligonucleótido. En los recuadros A y B se presentan las cuentas obtenidas en el citómetro de flujo, en ambos casos la mayor parte de la población celular se encuentra en G1. (A) corresponde a las células control (sin adriamicina, ni oligonucleótido) y (B) corresponde a las células transfectadas con el oligonucleótido y tratadas con adriamicina. En B se observa un incremento de las células en G2/M y una disminución en las células en G1 con respecto a A.

#### 2.2.4 Efecto del oligonucleótido MCG3 sobre la expresión de la proteína 14-3-3 $\sigma$ en la línea celular MCF7 en cultivo.

De acuerdo con los resultados anteriores, el oligonucleótido MCG3 disminuye la viabilidad de la línea celular en cultivo MCF-7. Pero aún no hemos determinado si esta disminución se deba a la formación de triplex y disminución en la transcripción del gen que codifica para 14-3-3  $\sigma$  ~~con~~secuentemente su traducción. Para determinarlo, realizamos una hibridación tipo Western. En la figura 45 se observa una autoradiografía obtenida después de hibridar el extracto proteico de células transfectadas con 1  $\mu$ g de oligonucleótido por ml de medio de cultivo, e incubadas 48 h después, durante 8, 12 ó 16 h con adriamicina. La hibridación se hizo con un anticuerpo policlonal de cabra contra 14-3-3  $\sigma$ , un anticuerpo contra IgG de cabra, y finalmente se adicionó reactivo para quimioluminiscencia, y se expuso en la oscuridad a una placa para rayos X. En la figura 46 se señala con una flecha la banda correspondiente a la proteína 14-3-3  $\sigma$ . Se observa una disminución en la banda en las células transfectadas con el oligonucleótido con respecto al control. En un análisis densitométrico se determinó por densidad óptica relativa, que la disminución corresponde

aproximadamente al 40% a las 8h, 60% a las 12h y 50% a las 26h. En la parte de abajo de esta misma figura se presenta la fotografía del gel de acrilamida teñido con azul de Coomasie, se observa que la concentración de proteína si bien no es igual es similar en todos los casos con excepción del carril correspondiente a 8h con oligonucleótido y adriamicina. Tanto al gel como a la radiografía se les determinó la densidad óptica relativa, para ajustar la concentración.



**Figura 45.** En el panel A se presenta la fotografía de una autoradiografía obtenida después de una hibridación tipo Western, del anticuerpo contra 14-3-3  $\sigma$  y un extracto proteico de Células de la línea MCF7 en cultivo incubadas con adriamicina 5  $\mu$ M, durante 8, 12 o 16 h y con 1  $\mu$ g de oligonucleótido MCG3/ml de medio de cultivo. El panel B corresponde a una fotografía del gel de poli-acrilamida teñido con azul de Coomasie, como control que indica que la cantidad de proteína que se colocó en cada carril es similar. La densidad óptica relativa obtenida se ajusto con respecto a la densidad óptica relativa obtenida para las proteínas. C/A con adriamicina, C/AO con adriamicina y oligonucleótido.

## VI. CONCLUSIONES

Las técnicas computacionales en las últimas décadas se han empleado para representar y modelar sistemas biológicos.

En la actualidad programas computacionales son utilizados de manera rutinaria en los laboratorios de investigación de diferentes áreas biológicas, para apoyar su investigación, por ejemplo existen los programas que facilitan el análisis de secuencias, el diseño de templados para PCR, la interpretación de microarreglos, los que permiten integrar la proteómica y la genómica funcional de un sistema de estudio, etc.

En esta tesis se han aplicado técnicas para modelar las vías de señalización intracelular en las que participa la proteína Ras, estos modelos se hicieron con el fin de obtener información que puede ser útil para la identificación de posibles blancos terapéuticos.

Presentamos dos modelos computacionales. El primer modelo nos permite dividir una vía de señalización en subgrupos fácilmente analizables, también nos permite determinar estados finales de cada uno de los elementos. Esto se hace considerando las interacciones que se dan entre los elementos, por ejemplo, si un elemento activa o inactiva a otro, o si se requiere de dos elementos activos para activar o inactivar a un tercero, etc.

Los estados finales nos señalan a los elementos o proteínas indispensables en esa vía de señalización, ya que aquellos elementos que siempre estén activos, serán requeridos y nunca podrán ser reemplazados aunque la célula tome una vía alterna, en este caso esos son los posibles blancos terapéuticos.

El segundo modelo es un sistema experto que crece de manera continua de acuerdo a los reportes experimentales de la literatura. Se eligió esta técnica computacional ya que permite evaluar el estado de cada componente con respecto al tiempo, además se puede asignar un valor numérico como la concentración a cada elemento del sistema.

Para este modelo se consideran a las vías de señalización en las que se involucra Ras y tres de sus proteínas efectoras Raf1, Ral-GDS y PI3K. Este modelo nos permite eliminar a cada una de las proteínas dentro de la célula de manera virtual y nos indica si la célula vive o muere al carecer de alguna de ellas, de esta manera identificamos a la proteína 14-3-3  $\sigma$  como un posible blanco terapéutico.

Este modelo en un futuro podrá ser integrado a un laboratorio virtual, que facilite la visualización de las vías de señalización cuando una proteína disminuya, aumente, desaparezca o aparezca en un tipo celular en el que usualmente no debe encontrarse, o que nos facilite la comprensión de resultados que experimentales sean contradictorios, ya que al visualizar las vías de señalización como un todo, podemos aclarar el resultado, al localizar interacción no reportadas que surjan *in silico*.

*Cellulat* es el laboratorio virtual que estamos desarrollando e integra varias técnicas computacionales en una estructura basada en agentes autónomos (González et al. 2003).

En este trabajo empleamos oligonucleótidos que forman una triple cadena, con el ADN de doble cadena del gen que codifica para la proteína 14-3-3  $\sigma$  en

células en cultivo de la línea celular de cáncer de mama MCF7 para determinar el efecto de este en el crecimiento de las mismas. El oligonucleótido diseñado es capaz de formar una triple cadena *in vitro*, y posiblemente *in vivo*, ya que disminuye a la proteína 14-3-3 $\sigma$ , 60%. Este oligonucleótido también disminuye la viabilidad celular de la línea celular MCF7 en cultivo, esta disminución es dependiente de la concentración de oligonucleótido.

En este caso en particular nosotros proponemos que el oligonucleótido puede ser empleado como coadyuvante en el control de crecimiento celular, ya que para observar su acción biológica es necesario causar daño a ADN, en este caso empleamos adriamicina para ese efecto.

El uso de oligonucleótidos que forman una triple cadena con fines terapéuticos cada vez es más prometedor ya que después que se completó el trabajo de esta tesis, se demostró la inhibición de crecimiento tumoral al emplearlos en células de humanas de cancer de prostata (Carbone et al., 2004) en células humanas de carcinoma pancreático (Cogoi et al., 2004) y en (Re et al. 2004).

## REFERENCIAS

1. Adjei, A. A. (2001) Blocking oncogenic Ras signaling for cancer therapy. *J. Natl. Cancer Inst.* 93 (14) 1062-1074.
2. Aitken, A. (2002) Functional specificity in 14-3-3 isoform interactions through dimer formation and phosphorylation. Chromosome location of mammalian isoforms and variants. *Plant. Mol. Biol.* 50 (6) 993-1010.
3. Alunni-Fabbroni, M.; Pirulli, D.; Manzini, G. y Xodo, L.E. (1996) (A,G)-oligonucleotides form extraordinary stable triple helices with a critical R.Y sequence of the murine c-Ki-ras promoter and inhibit transcription in transfected NIH 3T3 cells. *Biochemistry* 35 (50) 16361-16369.
4. Andoh, T.; Kato, T. Jr.; Matsui, Y. y Toh-e, A. (1998) Phosphoinositide-specific phospholipase C forms a complex with 14-3-3 proteins and is involved in expression of UV resistance in fission yeast. *Mol. Gen. Genet.* 258 (1-2) 139-147.
5. Aoki, H.; Hayashi, J.; Moriyama, M.; Arakawa, Y. y Hino, O. (2000) Hepatitis C virus core protein interacts with 14-3-3 protein and activates the kinase Raf-1. *J. Virol.* 74 (4) 1736-1741.
6. Armas, M.; Armas, O.; Cárdenas, M.; Vasconcelos, J. y Lagunez, J. (2000) Modelling signal transduction with boolean paradigms. En: *Mathematics and computers in modern science* 165-168. Ed. Mastorakis, N. E. World Scientific and engineering society press.
7. Arnott y Selsing (1974) Structures for the polynucleotide complexes poly (dA) with poly (dT) and poly (dT) with poly (dA) with poly (dT). *J. Mol. Biol.* 88 (2) 509-521.
8. Beban, M. y Miller, P. S. (2000) Pyrimidine motif triplexes containing polypurine RNA or DNA with oligo 2'-O-methyl or DNA triplex forming oligonucleotides. *Biochim. Biophys Acta.* 1492 (1) 155-162.
9. Behrens, P.; Rothe, M.; Wellmann, A.; Krischler, J. y Wernert, N. (2001) The Ets-1 transcription factor is up-regulated together with MMP 1 and MMP 9 in the stroma of pre-invasive breast cancer. *J. Pathol.* 194 (1) 43-50.
10. Bhalla, U. S. e Iyengar, R. (1999) Emergent properties of networks of biological signaling pathways. *Science* 283 (5400) 381-387.
11. Blagosklonny, M.V. y Pordee, A. B. (2001) Exploiting cancer cell cycling for selective prediction of normal cells. *Cancer Res.* 61 (11) 4301-4305.
12. Bodley, A.; Liu, L. F.; Israel, M.; Seshadri, R.; Koseki, Y.; Giuliani, F. C.; Kirschenbaum, S.; Silber, R. y Potmesil, M. (1989) DNA topoisomerase II-mediated interaction of doxorubicin and daunorubicin congeners with DNA. *Cancer Res.* 49 (21) 5969-5978.
13. Bray, D. (1990) Intracellular Signalling as a parallel distributed process. *J. Theor. Biol.* 143 (2) 215-231.
14. Bray, D. y Lay, S. (1994) Computer simulated evolution of a network of cell-signaling molecules. *Biophys J.* 66 (4) 972-977.
15. Bray, D. (1995) Protein molecules as computational elements in living cells. *Nature* 376 (6538) 307-312.



16. Brighman y Fell (2000) Differential feedback regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signalling in PC12. *FEBS Lett.* 482 (3) 169-174.
17. Brunet, A.; Bonni, A.; Zigmond, M. J.; Lin, M. Z.; Juo, P.; Hu, L. S.; Anderson, M. J.; Arden, K. C.; Blenis, J. y Greenberg, M. E. (1999) Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. *Cell* 96 (6) 857-868.
18. Cacace, A. M.; Michaud, N. R.; Therrien, M.; Mathes, K.; Copeland, T.; Rubin, G. M. y Morrison, D. K. (1999) Identification of constitutive and ras-inducible phosphorylation sites of KSR: implications for 14-3-3 binding, mitogen-activated protein kinase binding, and KSR overexpression. *Mol. Cell Biol.* 19 (1) 229-240.
19. Campbell, J. K.; Gurung, R.; Romero, S.; Speed, C. J.; Andrews, R. K.; Berndt, M. C. y Mitchell, C. A. (1997) Activation of the 43 kDa inositol polyphosphate 5-phosphatase by 14-3-3 zeta. *Biochemistry* 36 (49) 15363-15370.
20. Carbone, G. M.; Napoli, S.; Valentini, A.; Cavalli, F.; Watson, D. K. y Catapano, C. V. (2004) Triplex DNA-mediated downregulation of Ets2 expression results in growth inhibition and apoptosis in human prostate cancer cells. *Nucleic Acids Res.* 32 (14) 4358-4367.
21. Cárdenas-García, M.; Lagúnez-Otero, J. y Korneev, N. (2000) Efficient Attractor Analysis Based on Self-Dependent subsets of elements-An Application to Signal Transduction Studies. *ISMB* (8) 86-92.
22. Cárdenas-García, M.; Lagúnez-Otero, J. y Korneev, N. (2002) The application of abstract topology to Ras-related signal transduction pathways. *In Silico Biology* (2) 40-49.
23. Catapano, C. V.; McGuffie, E. M.; Pacheco, D. y Carbone, G. M. (2000) Inhibition of gene expression and cell proliferation by triple helix-forming oligonucleotides directed to the *c-myc* gene. *Biochemistry* 39 (17) 5126-5138.
24. Cobellis, G.; Missero, C.; Simionati, B.; Valle, G. y Di Lauro, R. (2001) Immediate early genes induced by H-Ras in thyroid cells. *Oncogene* 20 (18) 2281-2290.
25. Cogo, S.; Quadrioglio, F. y Xodo, L.E. (2004) G-rich oligonucleotides inhibits the binding of a nuclear protein to the Ki-ras promoter and strongly reduces cell growth in human carcinoma pancreatic cells. *Biochemistry* 43 (9) 2512-2523.
26. Cohen, P. T. (2002) Protein phosphatase 1-targeted in many directions. *J. Cell Sci.* 115 (Pt 2) 241-256.
27. Craparo, A.; Freund, R. y Gustafson, T. A. (1997) 14-3-3 (epsilon) interacts with the insulin-like growth factor I receptor and insulin receptor substrate I in a phosphoserine-dependent manner. *J. Biol. Chem.* 272 (17) 11663-11669.
28. Crespo, P. y León, J. (2000) Ras proteins in the control of the cell cycle and cell differentiation. *Cell. Mol. Life Sci.* 57 (11) 1613-1636.
29. Cullere, X.; Rose, P.; Thathamangalam, U.; Chatterjee, A.; Mullane, K. P.; Pallas, D. C.; Benjamin, T. L.; Roberts, T. M. y Schaffhausen, B. S. (1998)

- Serine 257 phosphorylation regulates association of polyomavirus middle T antigen with 14-3-3 proteins. *J. Virol.* 72 (1) 558-563.
30. Chan y Glazer (1997) Triplex DNA: fundamentals, advances, and potential applications for gene therapy. *J. Mol. Med* 75 (4) 267-282.
  31. Chan, T. A.; Hermeking, H.; Lengauer, C.; Kinzler, K.W. y Vogelstein, B. (1999) 14-3-3 sigma is required to prevent mitotic catastrophe after DNA damage. *Nature*, 401 (6753) 616-620.
  32. Chan, T, A.; Hwang, P. M.; Hermeking, H.; Kinzler, K. W. y Vogelstein, B. (2000) Cooperative effects of genes controlling the G<sub>2</sub>/M checkpoint. *Genes & Dev.* 14 (13) 1584-1588.
  33. Changeux, J. P. (1993) Chemical signalling in the brain. *Scientific American* (269) 30-37.
  34. Chilton-Maneill, S.; Ho, M.; Hawkins, C.; Gassas, A.; Zielenska, M. y Baruchel, S. (2004) c-kit expression and mutational analysis in medulloblastoma. *Pediatr. Dev. Pathol* 7 (5) 493-498.
  35. D'Agnano, I.; Valentini, A.; Fornari, C.; Bucci, B.; Starace, G.; Felsani, A. y Citro, G. (2001) Myc down-regulation induces apoptosis in M14 melanoma cells by increasing p27(kip1) levels. *Oncogene* 20 (22) 2814-2825.
  36. Davy, A.; Feuerstein, C. y Robbins, S. M. (2000) Signaling within a caveolae-like membrane microdomain in human neuroblastoma cells in response to fibroblast growth factor. *J. Neurochem* 74 (2) 676-683.
  37. Doi, A.; Fujita, S.; Matsuno, H.; Nagasaki, M. y Miyano, S. (2004) Constructing biological pathway models with hybrid functional Petri nets. In *Silico Biol.* 4 (2) 0023.
  38. Donaldson, J. C.; Dempsey, P. J.; Reddy, S.; Bouton, A. H.; Coffey, R. J. y Hanks, S. K. (2000) Crk-associated substrate p130(Cas) interacts with nephrocystin and both proteins localize to cell-cell contacts of polarized epithelial cells. *Exp. Cell Res.* 256 (1) 168-178.
  39. Dougherty, M. K. y Morrison, D. K. (2004) Unlocking the code of 14-3-3. *J. Cell Sciences* 117 (Pt 10) 1875-1884.
  40. Downward, J. (1998) Lipid-regulated kinases: some common themes at last. *Science* 279 (5351) 673-674.
  41. Edwards, C. (1995) Computational models for cellular information processing systems. unpublished undergraduate thesis, University of Liverpool.
  42. Etoh, T.; Ueo, H.; Inoue, H.; Sato, K.; Utsunomiya, T.; Barnard, G. F.; Kitano, S. y Mori, M. (2001) Clinical significance of K-Ras mutations in intraoperative tumor drainage blood from patients with colorectal carcinoma. *Ann. Surg. Oncol.* 8 (5) 407-412.
  43. Fanger, G. R.; Widmann, C.; Porter, A. C.; Sather, S.; Johnson, G. L. y Vaillancourt, R. R. (1998) 14-3-3 proteins interact with specific MEK kinases. *J. Biol. Chem.* 273 (6) 3476-3483.
  44. Feldman, B. J.; Reid, T. R. y Cleary, M. L. (1997) Pim1 cooperates with E2a-Pbx1 to facilitate the progression of thymic lymphomas in transgenic mice. *Oncogene* 15 (22) 2735-2742.
  45. Fisher, M.; Malcom, G. y Paton, R. (2000) Spatio-logical processes in intracellular signalling. *BioSystems* 55 (1-3) 83-92.

46. Frank-Kamenetskii, M. D. y Mirkin, S. M. (1995) Triplex DNA structures. *Annu. Rev. Biochem.* 64: 65-95
47. Fruman, D. A.; Meyers, R. E. y Cantley L. C. (1998) Phosphoinositide kinases. *Annu. Rev. Biochem.* 67: 481-507.
48. Funato, T.; Satou, J.; Kozawa, K.; Fujimaki, S.; Miura, T. y Kaku, M. (2001) Use of c-myc antisense oligonucleotides to increase the sensitivity of human colon cancer cells to cisplatin. *Oncol Rep.* 8 (4) 807-810.
49. Furlanetto, R. W.; Dey, B. R.; Lopaczynski, W. y Nissley, S. P. (1997) 14-3-3 proteins interact with the insulin-like growth factor receptor but not the insulin receptor. *Biochem. J.* 327 ( Pt 3) 765-771.
50. García, R.; Bowman, T. L.; Niu, G.; Yu, H.; Minton, S.; Muro-Cacho, C. A.; Cox, C. E.; Falcone, R.; Fairclough, R.; Parsons, S.; Laudano, A.; Gazit, A.; Levitzki, A.; Kraker A. y Jove R. (2001) Constitutive activation of Stat3 by the Src and JAK tyrosine kinases participates in growth regulation of human breast carcinoma cells. *Oncogene* 20 (20) 2499-2513.
51. Gelvan, D. y Samuni, A. (1986) Cellular targets of adriamycin-induced damage in *Escherichia coli* *Biochem Pharmacol.* 35 (19) 3267-3275.
52. Ghomashchi, F.; Zhang, X.; Liu, L. y Gelb, M. H. (1995) Binding of prenylated and polybasic peptides to membranes: affinities and intervesicle exchange. *Biochemistry* 34 (37) 11910-11918.
53. González, P. P.; Cárdenas-García, M.; Camacho, D.; Franyuti, A.; Rosas, O. y Lagúnez-Otero, J. (2003) Cellulat: An agent-based intracellular signalling model. *BioSystems* 68 (2-3) 171-185.
54. González-Pérez, P.P.; Gershenson, C.; Cárdenas-García, M. y Lagúnez-Otero, J. (2000) Modelling intracellular signalling networks using behaviour-based systems and the blackboard architecture, En: *Mathematics and computers in modern science* 219-223. Ed. Mastorakis, N. E. World Scientific and engineering society press.
55. Gorgoulis, V. G.; Zacharatos, P.; Mariatos, G.; Liloglou, T.; Kokotas, S.; Kastrinakis, N.; Kotsinas, A.; Athanasiou, A.; Foukas, P.; Zoumpourlis, V.; Kletsas, D.; Ikonopoulou, J.; Asimacopoulos, P. J.; Kittas, C. y Field, J. K. (2001) Deregulated expression of *c-mos* in non-small cell lung carcinomas: relationship with p53 status, genomic instability, and tumor kinetics. *Cancer Res.* 61 (2) 538-549.
56. Gorre, M. E.; Mohammed, M.; Ellwood, K.; Hsu, N.; Paquette, R.; Rao, P. N. y Sawyers, C. L. (2001) Clinical Resistance to STI-571 Cancer Therapy Caused by BCR-ABL Gene Mutation or Amplification. *Science* 293 (5531) 876-880.
57. Gowers, D. M. y Fox, K. R. (1999) Towards mixed sequence recognition by triple helix formation. *Nucleic Acids Res.* 27 (7) 1569-1577.
58. Hansen, R. K.; Parra, I.; Hilsenbeck, S. G.; Himmelstein, B. y Fuqua, S. A. (2001). Hsp27-induced MMP-9 expression is influenced by the Src tyrosine protein kinase yes. *Biochem. Biophys. Res. Commun.* 282 (1) 186-193.
59. Harashima, N.; Tanaka, K.; Sasatomi, T.; Shimizu, K.; Miyagi, Y.; Yamada, A.; Tamura, M.; Yamana, H.; Itoh, K. y Shichijo, S. (2001) Recognition of the

ESTA TESIS NO SALI  
DE LA BIBLIOTECA

- Lck tyrosine kinase as a tumor antigen by cytotoxic T lymphocytes of cancer patients with distant metastases. *Eur. J. Immunol.* 31 (2) 323-332.
60. Hausser, A.; Storz, P.; Link, G.; Stoll, H.; Liu, Y. C.; Altman, A.; Pfizenmaier, K. y Johannes, F. J. (1999) Protein kinase C mu is negatively regulated by 14-3-3 signal transduction proteins. *J. Biol. Chem.* 274 (14) 9258-9264.
  61. Hayne, C.; Tzivion, G. y Luo, Z. (2000) Raf-1/MEK/MAPK pathway is necessary for the G2/M transition induced by nocodazole. *J. Biol. Chem.* 275 (41) 31876-31882.
  62. Henriksson, M. L.; Francis, M. S.; Peden, A.; Aili, M.; Stefansson, K.; Palmer, R.; Aitken, A. y Hallberg, B. (2002) A nonphosphorylated 14-3-3 binding motif on exoenzyme S that is functional in vivo. *Eur. J. Biochem.* 269 (20) 4921-4929.
  63. Hermeking, H.; Lengauer, C.; Polyak, K.; He, T. C.; Zhang, L.; Thiagalingam, S.; Kinzler, K. W. y Vogelstein, B. (1997) 14-3-3 sigma is a p53-regulated inhibitor of G2/M progression. *Mol. Cell* 1 (1) 3-11.
  64. Hesketh, R. (1994) *The oncogene handbook*. Academic Press.
  65. Hinoi, T.; Kishida, S.; Koyama, S.; Ikeda, M.; Matsuura, Y y Kikuchi, A. (1996) Post-translational modifications of Ras and Ral are important for the action of Ral GDP dissociation stimulator. *J. Biol. Chem* 271 (33) 19710-19716.
  66. Hobbs, C. A. y Yoon, K. (1994) Differential regulation of gene expression in vivo by triple helix-forming oligonucleotides as detected by a reporter enzyme. *Antisense Res. Dev.* 4 (1) 1-8.
  67. Holcombe, M. (1994) From VLSI trough machine models to cellular metabolism. *Computing with biological metaphors*, Paton, R. ed., Chapman and Hall, 11-25.
  68. Hancock, J. F.; Paterson, H. y Marshall, C. J. (1990) A polibasic domain or palmitoylation is required in addition to the CAAX motif to localize p21ras to the plasma membrane. *Cell* 63 (1) 133-139.
  69. Horne, D.A. y Dervan, P.B. (1991) Effects of an abasic site on triple helix formation characterized by affinity cleaving. *Nucleic Acids Res.* 19 (18) 4963-4965.
  70. Hu, Q.; Klippel, A.; Muslin, A. J.; Fantl, W. J. y Williams, L. T. (1995) Ras-dependent induction of cellular responses by constitutively active phosphatidylinositol 3-kinase. *Science* 268 (5207) 100-102.
  71. Hung, C. F.; Hsu, K. F.; Cheng, W. F.; Chai, C. Y.; He, L.; Ling, M. y Wu, T. C. (2001) Enhancement of DNA vaccine potency by linkage of antigen gene to a gene encoding the extracellular domain of Fms-like tyrosine kinase 3-ligand. *Cancer Res.* 61 (3) 1080-1088.
  72. Ikeda, Y.; Koizumi, N.; Kusano, T. y Sano, H. (2000) Specific binding of a 14-3-3 protein to autophosphorylated WPK4, an SNF1-related wheat protein kinase, and to WPK4-phosphorylated nitrate reductase. *J. Biol. Chem.* 275 (41) 31695-31700 y 275 (52) 41528.
  73. Improta-Brears, T.; Ghosh, S. y Bell, R. M. (1999) Mutational analysis of Raf-1 cysteine rich domain: requirement for a cluster of basic aminoacids for interaction with phosphatidylserine. *Mol. Cell Biochem.* 198 (1-2) 171-178.

74. Jans, D. A. (1994) Nuclear signaling pathways for polypeptide ligands and their membrane receptors? *FASEB* (8) 841-847.
75. Janssens, V. y Goris, J. (2001) Protein phosphatase 2A a highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochem. J.* 353 (Pt 3) 417-439.
76. Jeremic, B. (2001) Yes! Time for global strategy to improve the treatment of lung cancer is right now. *Lung Cancer* 32 (2) 209-210.
77. Karin, M. y Smeal, T. (1992) Control of transcription factors by signal transduction pathways: the beginning of the end. *Trends Bioche. Sci.* 17 (10) 418-422.
78. Karp, P. D. y Paley, S. M. (1994) Representations of metabolic knowledge: pathways. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2: 203-211.
79. Kato, S.; Masuhiro, Y.; Watanabe, M.; Kobayashi, Y.; Takeyama, K. I.; Endoh, H. y Yanagisawa, J. (2000) Molecular mechanism of a cross-talk between oestrogen and growth factor signalling pathways. *Genes Cells* 5 (8) 593-601.
80. Kauffman, S. A. (1991) Antichaos and adaptation. *Scientific American* 265 (2) 64-70.
81. Kiuru-Kuhlefelt, S.; Sarlomo-Rikala, M.; Larramendy, M. L.; Soderlund, M.; Hedman, K.; Miettinen, M. y Knuutila, S. (2000) FGF4 and INT2 oncogenes are amplified and expressed in Kaposi's sarcoma. *Modern Pathol.* 13 (4) 433-437.
82. Korsmeyer, S. J. (1999) BCL-2 gene family and the regulation of programmed cell death. *Cancer Res.* 59 (7 Suppl.) 1693s-1700s.
83. Kosako, H.; Gotoh, Y.; Matsuda, S.; Ishikawa, M. y Nishida, E. (1992) Map kinase activator is a serine/threonine/tyrosine activated by threonine phosphorylation. *EMBO J.* 11 (8) 2903-2908.
84. Koveker, G. B. (2000) Growth factors in clinical practice *Int. J. Clin. Pract.* 54 (9) 590-593.
85. Ku, N. O.; Liao, J. y Omary, M. B. (1998) Phosphorylation of human keratin 18 serine 33 regulates binding to 14-3-3 proteins. *EMBO J.* 17 (7) 1892-1906.
86. Kumagai, A.; Yakowec, P. S. y Dunphy, W. G. (1998) 14-3-3 proteins act as negative regulators of the mitotic inducer Cdc25 in *Xenopus* egg extracts. *Mol Biol Cell.* 9 (2) 345-354.
87. Kuwana, T.; Peterson, P. A. y Karlsson, L. (1998) Exit of major histocompatibility complex class II-invariant chain p35 complexes from the endoplasmic reticulum is modulated by phosphorylation. *Proc. Natl. Acad. Sci.* 95 (3) 1056-1061.
88. Lacal, J. C. y McCormick, F. (1993) *The ras Superfamily of GTPases.* CRC Press Inc, USA.
89. Lagunéz-Otero, J. (1998) The cell as an expert system. *Dagstuhl seminar report* 215, p60. ISSN 0940-1121.
90. Laronga, C.; Yang, H. Y.; Neal, C. y Lee, M. H. (2000) Association Of The Cyclin-Dependent Kinases And 14-3-3 Sigma Negatively Regulates Cell Cycle Progression *J. Biol. Chem.* 275 (30) 23106-23112.

91. Lee, H.; Akita, R. W.; Sliwkowski, M. X. y Maihle, N. J. (2001) A naturally occurring secreted human ErbB3 receptor isoform inhibits heregulin-stimulated activation of ErbB2, ErbB3, and ErbB4. *Cancer Res.* 61 (11) 4467-4473.
92. Leitner, D.; Schroder, W. y Weisz, K. (2000) Influence of sequence-dependent cytosine protonation and methylation on DNA triplex stability. *Biochemistry.* 39 (19) 5886-5892.
93. Levy, S. (1992) *Artificial life: the quest for a new creation.* Jonathan Cape, London.
94. Liu, Y. C.; Liu, Y.; Elly, C.; Yoshida, H.; Lipkowitz, S. y Altman, A. (1997) Serine phosphorylation of Cbl induced by phorbol ester enhances its association with 14-3-3 proteins in T cells via a novel serine-rich 14-3-3-binding motif. *J. Biol. Chem.* 272 (15) 9979-9985.
95. Liu, Y. C. y Altman, A. (1998) Cbl: complex formation and functional implications. *Cell Signal.* 10 (6) 377-385.
96. López-Girona, A.; Furnari, B.; Mondesert, O. y Russell, P. (1999) Nuclear localization of Cdc25 is regulated by DNA damage and a 14-3-3 protein. *Nature.* 397 (6715) 172-175.
97. McMahon, M. y Woods, D. (2001) Regulation of the p53 pathway by Ras, the plot thickens. *Biochim Biophys Acta* 1471(2) M63-M71.
98. MacNicol, M. C.; Muslin, A. J. y MacNicol, A. M. (2000) Disruption of the 14-3-3 binding site within the B-Raf kinase domain uncouples catalytic activity from PC12 cell differentiation. *J. Biol. Chem.* 275 (6) 3803-3809.
99. Madruga, J.; Briegel, K.; Diebold, S.; Boehmelt, G.; Vogel, F. y Zenke, M. (2000) Dendritic cells conditionally transformed by *v-rel* ER oncogene express lymphoid marker genes. *Immunobiology.* 202 (4) 394-407.
100. Mahyar-Roemer, M. y Roemer, K. (2001) P21 Waf1/Cip1 can protect human colon carcinoma cells against p53-dependent and p53-independent apoptosis induced by natural chemopreventive and therapeutic agents. *Oncogene* 20 (26) 3387-3398.
101. Mansour, S. J.; Matten, W. T.; Hermann, A. S.; Candia, J. M.; Rong, S. y Fukusawa, K. (1994) Transformation of mammalian cells by constitutively active MAP kinase kinase. *Science* 265 (5174) 966-970.
102. Marshall, C. J. (1994) MAP kinase kinase kinase, MAP kinase kinase and MAP kinase, *Curr. Opin. Genet. Dev.* 4 (1) 82-89.
103. Marijuan, P. C.; (1994) *Enzymes, automata, and artificial cells. Computing with Biological Metaphors,* Paton, R. (ed), Chapman and Hall, 50-68.
104. Masters, S. C.; Pederson, K. J.; Zhang, L.; Barbieri, J. T. y Fu, H. (1999) Interaction of 14-3-3 with a nonphosphorylated protein ligand, exoenzyme S of *Pseudomonas aeruginosa*. *Biochemistry* 38 (16) 5216-5221.
105. Medhi, C. (2002) The models of proton assisted and the unassisted formation of CGC base triplets. *J. Chem. Inf. Comput. Sci.* 42 (3) 598-601.
106. McCarthy, Minsky, Newell y Simon (1995) Dartmouth computer conference. [http://livinginternet.com/iii\\_ai.htm](http://livinginternet.com/iii_ai.htm).

107. McIlroy, J.; Chen, D.; Wjasow, J.; Michaeli, T. y Backer, J. M. (1997) p85-p110 phosphatidylinositol 3-kinase stimulates DNA synthesis by Ras and p70 S6 kinase-dependent pathways. *Mol. Cell. Biol.* 17 (1) 248-255.
108. McMahon, M. y Woods, D. (2001) Regulation of the p53 pathway by Ras, the plot thickens. *Bioch. Bioph. Acta* 1471 (2) M63-M71.
109. McPherson, R. A.; Harding, A.; Roy, S.; Lane, A. y Hancock, J. F. (1999) Interactions of c-Raf-1 with phosphatidylserine and 14-3-3. *Oncogene* 18 (26) 3862-3869.
110. Milligan, G. (1993) Mechanisms of multifunctional signalling by G protein-linked receptors. *TIPS* (14) 239-244.
111. Moore, B.W. y Pérez, V. J. (1967) Specific acidic proteins of the nervous system. *Physiological and Biochemical Aspects of nervous integration*. Ed FD Carlson. Prentice-Hall, 343-359.
112. Morgan, A. R. y Wells, R. D. (1968) Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences. *J. Mol. Biol.* 37 (1) 63-80.
113. Morrison, D. K. y Cutler, R. E. (1997) The complexity of Raf-1 regulation. *Curr. Opin. Cell Biol.* 9 (2) 174-179.
114. Munday, A. D.; Berndt, M. C. y Mitchell, C. A. (2000) Phosphoinositide 3-kinase forms a complex with platelet membrane glycoprotein Ib-IX-V complex and 14-3-3zeta. *Blood*. 96 (2) 577-584.
115. Muslin, A. J.; Tanner, J. W.; Allen, P. M. y Shaw, A. S. (1996) Interaction of 14-3-3 with signaling proteins is mediated by the recognition of phosphoserine. *Cell*. 84 (6) 889-897.
116. Nishizuka, Y. editor (1992) Signal transduction: crosstalk. *Trends Biochem. Sci.* 17 (10) 367-443.
117. Nurse, P. (2000) A long twentieth century of the cell cycle and beyond. *Cell* 100 (1) 71-78.
118. O'Neil, P. V. (1982) Fundamental concepts of topology. Gordon and Breach New York, 2<sup>nd</sup> edition.
119. Papp, T.; Schipper, H.; Pemsel, H.; Bastrop, R.; Muller, K. M.; Wiethage, T.; Weiss, D. G.; Dopp, E.; Schiffmann, D. y Rahman, Q. (2001) Mutational analysis of N-ras, p53, p16INK4a, p14ARF and CDK4 genes in primary human malignant mesotheliomas. *Int. J. Oncol.* 18 (2) 425-433.
120. Paton, R. C. (1993) Some computational models at the cellular level, *BioSystems*, 29 (2-3) 63-75
121. Paton, R. C. (1995) Specifying logical agents in cellular hierarchies, *Proceedings of IPCAT (Information Processing in Cell and Tissues)*, Paton, R. C., Holcombe, M. y Staniford, G. (eds) 302-317.
122. Peng, C. Y.; Graves, P. R.; Thoma, R. S.; Wu, Z.; Shaw, A. S. y Piwnicka-Worms, H. (1997) Mitotic and G2 checkpoint control: regulation of 14-3-3 protein binding by phosphorylation of Cdc25C on serine-216. *Science*. 277 (5331) 1501-1505.
123. Petosa, C.; Masters, S. C.; Bankston, L. A.; Pohl, J.; Wnag, B.; Fu, H. y Liddington, R. C. (1998) 14-3-3zeta binds a phosphorylated Raf peptide and

- an unphosphorylated peptide via its conserved amphipathic groove. *J. Biol. Chem.* 273 (26) 16305-16310.
124. Porebska, I.; Harlozinska, A. y Bojarowski, T. (2000) Expression of the tyrosine kinase activity growth factor receptors (EGFR, ERB B2, ERB B3) in colorectal adenocarcinomas and adenomas *Tumour Biol.* 21 (2) 105-115.
  125. Potaman, V. N. (2003) Applications of triple-stranded nucleic acid structures to DNA purification, detection and analysis. *Expert. Rev. Mol. Diagn.* 3 (4) 481-496.
  126. Praseuth, D.; Guieysse, A. L. y Hélène, C. (1999) Triple helix formation and the antigene strategy for sequence-specific control of gene expression *Biochim. Biophys. Acta*, 1489 (1) 181-206
  127. Pratt, M. A.; Satkunarathnam, A. y Novosad, D. M. (1998) Estrogen activates raf-1 kinase and induces expression of Egr-1 in MCF-7 breast cancer cells *Mol. Cell Biochem.* 189 (1-2) 119-25.
  128. Pritchard, L. y Dufton, M. J. (2000) Do proteins learn to evolve?. The Hopfield network as a basis for the understanding of protein evolution. *J. Theor. Biol.* 202 (1) 77-86.
  129. Qiu, W.; Zhuang, S.; von Lintig, F. C.; Boss, G. R. y Pilz, R. B. (2000) Cell type-specific regulation of B-Raf kinase by cAMP and 14-3-3 proteins. *J. Biol. Chem.* 275 (41) 31921-31929.
  130. Raff, M.C. (1992) Social controls on cell survival and cell death. *Nature* 356 (6368) 397-400.
  131. Ramocki, M. B.; White, M. A. Konieczny, S. F. y Taparowsky, E. J. (1998) A role for RaIGDS and a novel Ras effector in the Ras-mediated inhibition of skeletal myogenesis. *J. Biol. Chem.* 273 (28) 17696-17701.
  132. Re, R. N.; Cook, J. L. y Giardina, J. F. (2004) The inhibition of tumor growth by triplex-forming oligonucleotides. *Cancer Lett.* 209 (1) 51-53.
  133. Rittinger, K.; Budman, J.; Xu, J.; Volinia, S.; Cantley, L. C.; Smerdon, S J.; Gambin, S J.; y Yaffe, M. B. (1999) Structural analysis of 14-3-3 phosphopeptide complex identifies a dual role for the nuclear export signal of 14-3-3 in ligand binding. *Mol. Cell* 4 (2) 153-166.
  134. Rivera, J.; Lozano, M. L.; Corral, J.; González-Conejero, R.; Martínez, C. y Vicente, V. (2000) Platelet GP Ib/IX/V complex: physiological role. *J. Physiol. Biochem.* 56 (4) 355-365.
  135. Rodríguez-Viciano, P.; Warne, P. H.; Dhand, R.; Vanhaesebroeck, B.; Gout, I. y Fry, M. J. (1994) Phosphatidylinositol 3-kinase as a direct target for Ras. *Nature* 370 (6490) 527-532.
  136. Rosenquist, M.; Alsterfjord, M.; Larsson, C. y Sommarin, M. (2001) Data mining the Arabidopsis genome reveals fifteen 14-3-3 genes. Expression is demonstrated for two out of five novel genes. *Plant. Physiol.* 127 (1) 142-149.
  137. Sakamoto, Y.; Kitajima, Y.; Edakuni, G.; Sasatomi, E.; Mori, M.; Kitahara, K. y Miyazaki K. (2001) Expression of Trk tyrosine kinase receptor is a biologic marker for cell proliferation and perineural invasion of human pancreatic ductal adenocarcinoma. *Oncol. Rep.* 8 (3) 477-484.
  138. Senften, M.; Dilworth, S. y Ballmer-Hofer, K. (1997) Multimerization of polyomavirus middle-T antigen. *J. Virol.* 71 (9) 6990-6995.



139. Shaharabany, M.; Abramovitch, R.; Kushnir, T.; Tsarfaty, G.; Ravid-Megido, M.; Horev, J.; Ron, R.; Itzchak, Y. y Tsarfaty, I. (2001) In vivo molecular imaging of met tyrosine kinase growth factor receptor activity in normal organs and breast tumors. *Cancer. Res.* 61 (12) 4873-4878.
140. Shen, Y.; Mamers, P.; Jobling, T.; Burger, H. G. y Fuller, P. J. (1996) Absence of the previously reported G protein oncogene (*gip2*) in ovarian granulosa cell tumors. *J Clin Endocrinol Metab* 81 (11) 4159-4161.
141. Shi, Y.; Tang, D.; Deng, J. y Su, C. (1998) Detection of *gsp* oncogene in growth hormone-secreting pituitary adenomas and the study of clinical characteristics of acromegalic patients with *gsp*-positive pituitary tumors. *Chin. Med. J.* 111 (10) 891-894.
142. Shirouzu, M.; Morinaka, K.; Koyama, S.; Hu, C. D.; Hori-Tamura, N.; Okada, T.; Kariya, K.; Kataoka, T.; Kikuchi, A. y Yokoyama, S. (1998) Interactions of the amino acid residue at position 31 of the c-Ha-Ras protein with Raf-1 and RalGDS. *J. Biol. Chem.* 273 (13) 7737-7742.
143. Spaargaren, M. y Bischoff, J. R. (1994) Identification of the guanine dissociator stimulator for Ral as a putative effector loop of R-Ras, H-Ras K-Ras and rap. *Proc. Natl. Acad. Sci.* 91 (26) 11089-11093.
144. Strobel, S. A.; Doucette-Stamm, L.A.; Riba, L.; Housman, D. E. y Dervan, P. B. (1991) Site-specific cleavage of human chromosome 4 mediated by triple-helix formation. *Science.* 254 (5038) 1639-1642.
145. Sun, J. S.; Mergny, J. L.; Lavery, R.; Montenay-Garestier, T. y Hélène, C. (1991) Triple helix structures: sequence dependence, flexibility and mismatch effects. *J. Biomol. Struct. Dyn.* 9 (3) 411-424.
146. Taguchi, T.; Kiyokawa, N.; Sato, N.; Saito, M. y Fujimoto, J. (2000) Characteristic expression of Hck in human B-cell precursors. *Exp. Hematol.* 28 (1) 55-64.
147. Takai-Igarashi, T. y Kaminuma, Ts. (1998) A pathway finding system for the cell signaling network database. *J. Comp. Biol.* 5 (4) 747-754.
148. Tan, Y.; Ruan, H.; Demeter, M. R. y Comb, M. J. (1999) p90(RSK) blocks bad-mediated cell death via a protein kinase C-dependent pathway. *J. Biol. Chem.* 274 (49) 34859-34867.
149. Tarr, M. y Helden, P. D. (1990) Inhibition of transcription by adriamycin is a consequence of the loss of negative superhelicity in DNA mediated by topoisomerase II. *Mol. Cell Biochem.* 93 (2) 141-146.
150. Tewey, K. M.; Rowe, T. C.; Yang, L.; Halligan, B. D. y Liu, L. F. (1984) Adriamycin-induced DNA damage mediated by mammalian DNA topoisomerase II. *Science* 226 (4673) 466-468.
151. Thormeyer, D. y Baniahmad, A. (1999) The *v-erbA* oncogene. *Int. J. Mol. Med.* 4 (4) 351-358
152. Toker, A. y Cantley, L. C. (1997) Signalling through the lipid products of phosphoinositide-3-OH kinase. *Nature* 387 (6634) 673-676.
153. Torigoe, H.; Shimizume, R.; Sarai, A. y Shindo, H. (1999) Triplex formation of chemically modified homopyrimidine oligonucleotides: thermodynamic and kinetic studies. *Biochemistry.* 38 (44) 14653-14659.

154. Ushijima, K.; Ishibashi, T.; Yamakawa, H.; Tsukahara, S.; Takai, K.; Maruyama, T. y Takaku, H. (1999) Inhibition of restriction endonuclease cleavage by triple helix formation with RNA and 2'-O-methyl RNA oligonucleotides containing 8-oxo-adenosine in place of cytidine. *Biochemistry* 38 (20) 6570-6575.
155. Van Dam, H. y Castellazzi, M. (2001). Distinct roles of Jun : Fos and Jun : ATF dimers in oncogenesis. *Oncogene*. 20 (19) 2453-2464.
156. Van Der Hoeven, P. C.; Van Der Wal, J. C.; Ruurs, P. y Van Blitterswijk, W. J. (2000) Protein kinase C activation by acidic proteins including 14-3-3. *Biochem. J.* 347 (Pt 3) 781-785.
157. Van Der Hoeven, P. C.; Van Der Wal, J. C.; Ruurs, P.; Van Dijk, M. C. y Van Blitterswijk, J. (2000) 14-3-3 isotypes facilitate coupling of protein kinase C-zeta to Raf-1: negative regulation by 14-3-3 phosphorylation. *Biochem. J.* 345 (Pt 2) 297-306.
158. Vanhaesebroeck, B.; Leever, S. J.; Panayotou, G. y Waterfield, M. D. (1997) Phosphoinositide 3-kinases: a conserved family of signal transducers. *Trends. Biochem. Sci.* 22 (7) 267-272.
159. Wakui, H.; Wright, A. P.; Gustafsson, J. y Zilliacus, J. (1997) Interaction of the ligand-activated glucocorticoid receptor with the 14-3-3 eta protein. *J. Biol. Chem.* 272 (13) 8153-8156.
160. Wang, W. y Shakes, D. C. (1996) Molecular evolution of the 14-3-3 protein family. *J. Mol. Evol.* 43 (4) 384-398
161. Wang, B.; Yang, H.; Liu, Y. C.; Jelinek, T.; Zhang, L.; Ruoslahti, E. y Fu, H. (1999) Isolation of high-affinity peptide antagonist of 14-3-3 proteins by phage display. *Biochemistry* 38 (38) 12499-12504.
162. Weber, C. K.; Slupsky, J. R.; Kalmes, H. A. y Rapp, U. R. (2001) Active Ras induces heterodimerization of cRaf and BRaf. *Cancer Res.* 61 (9) 3595-3598.
163. Waterman, M. J.; Stavridi, E. S.; Waterman, J. L. y Halazonetis, T. D. (1998) ATM-dependent activation of p53 involves dephosphorylation and association with 14-3-3 proteins. *Nat. Genet.* 19 (2) 175-178.
164. Weber, C. K.; Slupsky, J. R.; Kalmes, H. A. y Rapp, U. R. (2001) Active Ras induces heterodimerization of cRaf and BRaf. *Cancer Res.* 61 (9) 3595-3598.
165. Wilkinson, M. G. y Millar, J. B. A. (2000) Control of the eukariotic cell cycle by MAP kinase signaling pathways. *FASEB J.* 14 (14) 2147-2157.
166. Widen, C.; Zilliacus, J.; Gustafsson, J. A. y Wikstrom, A. C. (2000) Glucocorticoid receptor interaction with 14-3-3 and Raf-1, a proposed mechanism for cross-talk of two signal transduction pathways. *J. Biol. Chem.* 275 (50) 39296-39301.
167. Wurthner, J. U.; Mukhopadhyay, A. K. y Peimann, C. J. (2000) A cellular automaton model of cellular signal transduction. *Comput. Biol. Med.* 30 (1) 1-21.
168. Xing, H.; Kornfeld, K. y Muslin, A. J. (1997) The protein kinase KSR interacts with 14-3-3 protein and Raf. *Curr. Biol.* 7 (5) 294-300.

169. Yaffe, M. B.; Winkler, K.; Volinia, S.; Caron, P. R.; Aitken, A.; Leffers, H.; Gamblin, S. J.; Smerdon, S. J. y Cantley, L. C. (1997) The structural basis for 14-3-3: phosphopeptide binding specificity. *Cell* 91 (7) 961-971.
170. Yip-Schneider, M. T.; Miao, W.; Lin, A.; Barnard, D. S.; Tzivion, G. y Marshall, M. S. Regulation of the Raf-1 kinase domain by phosphorylation and 14-3-3 association. *Biochem J.* 351 (Pt 1) 151-159.
171. Yang, J.; Winkler, K.; Yoshida, M. y Kornbluth S. (1999) Maintenance of G2 arrest in the *Xenopus* oocyte: a role for 14-3-3-mediated inhibition of Cdc25 nuclear import. *EMBO J.* 18 (8) 2174-2183.
172. Yoshida, T.; Muramatsu, H.; Muramatsu, T.; Sakamoto, H.; Katoh, O.; Sugimura, T. y Terada, M. (1988) Differential expression of two homologous and clustered oncogenes, *Hst1* and *Int-2*, during differentiation of F9 cells *Biochem. Biophys. Res. Commun.* 157 (2) 618-625.
173. Yuryev, A.; Ono, M.; Goff, S. A.; Macaluso, F. y Wennogle, L. P. (2000) Isoform-specific localization of A-RAF in mitochondria. *Mol. Cell Biol.* 20 (13) 4870-4878.
174. Zamecnik, P. C. y Stephenson, M. L. (1978) Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. *Proc. Natl Acad. Sci.* 75 (1) 280-284.
175. Zarn, J. A.; Zimmermann, S. M.; Pass, M. K.; Waibel, R. y Stahel, R. A. (1996) Association of CD24 with the kinase *c-fgr* in a small cell lung cancer cell line and with the kinase *lyn* in an erythroleukemia cell line. *Biochem Biophys Res Commun.* 225 (2) 384-391.
176. Zha, J.; Harada, H.; Osipov, K.; Jockel, J.; Waksman, G. y Korsmeyer, S. J. (1997) BH3 domain of BAD is required for heterodimerization with BCL-XL and pro-apoptotic activity. *J. Biol. Chem.* 272 (39) 24101-24104.
177. Zha, J.; Harada, H.; Yang, E.; Jockel, J. y Korsmeyer, S. J. (1996) Serine phosphorylation of death agonist BAD in response to survival factor results in binding to 14-3-3 not BCL-X(L) *Cell* 87 (4) 619-628.
178. Zhai, J.; Lin, H.; Shamim, M.; Schlaepfer, W. W. y Canete-Soler, R. (2001) Identification of a novel interaction of 14-3-3 with p190RhoGEF. *J. Biol. Chem.* 276 (44) 41318-41324.
179. Zhang, L.; Chen, J. y Fu, H. (1999) Suppression of apoptosis signal-regulating kinase 1-induced cell death by 14-3-3 proteins. *Proc. Natl. Acad. Sci.* 96 (15) 8511-8515.
180. Zhang, B. H. y Guan, K. L. (2000) Activation of B-Raf kinase requires phosphorylation of the conserved residues Thr598 and Ser601. *EMBO J.* 19 (20) 5429-5439.
181. Zilliacus, J.; Holter, E.; Wakui, H.; Tazawa, H.; Treuter, E. y Gustafsson, J. K. (2001) Regulation of glucocorticoid receptor activity by 14-3-3-dependent intracellular relocalization of the corepressor rip140. *Mol. Endocrinol.* 15 (4) 501-511.
182. Zohn, I. E.; Symons, M.; Chrzanowska-Wodnicka, M.; Westwick, J. K. y Der, C. J. (1998) *mas* oncogene signaling and transformation require the small GTP-binding protein Rac. *Mol. Cell Biol.* 18 (3) 1225-1235.

*In Silico* Biology 2, 0040 (2002); ©2002, Bioinformation Systems e.V.

# The application of abstract topology to RAS-related signal transduction pathways

Maura Cárdenas-García <sup>a,1</sup>, Jaime Lagunez Otero <sup>a,2</sup> and Nikolai A. Korneev <sup>b,3</sup>

<sup>a</sup>Instituto de Química UNAM, Circuito Exterior CP 04510, Coyoacán, México

Phone: (52) 555 6224424

Fax: (52) 555 6162203

<sup>b</sup>Instituto Nacional de Astrofísica óptica y Electrónica, Apartado Postal 51 y 216, Sta. María Tonanzintla, Puebla

Phone: (52) 222 2663100

Fax: (52) 222 2472940

E-mail: <sup>1</sup>maura@servidor.unam.mx, <sup>2</sup>lagunez@servidor.unam.mx, <sup>3</sup>korneev@inaoep.mx

Edited by E. Wingender; received April 30, 2002; revised and accepted June 02, 2002; published June 13, 2002

## Abstract

Ras is a protein related to cancer development. It is a convergence point for different signal transduction pathways that allow the cell to respond to external stimuli with different cell functions like growth, division, death, etc. In this paper, we analyze the signal pathways generated by different Ras effectors (Raf, RalGDS and PI3K), and the pathway relating Ras to the cell cycle control. We show that the interaction among different elements of these pathways induces a topologic structure in the set of elements. We discuss properties of this topology and give an algorithm to build it. The application of topological concepts makes easier the interaction analysis. Using a computational algorithm, we can create isolated, independently manageable sub-groups. Then we construct their hierarchical structure. The procedure allows us to visualize groups of elements related to the Ras effectors involved in cell growth, the elements involved in the cytoskeleton regulation, and the elements related to the cell cycle control. Thus the division in sub-groups does not only make easier the analysis, but it also provides a biologically meaningful subdivision.

**Key words:** Cell signal transduction pathways, abstract topology, Ras protein

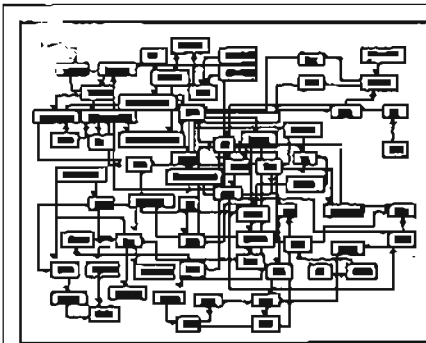
## Introduction

The cells determine their growth, division, death, etc., on the basis of extracellular signals transmitted to their interior through specific receptors located in the plasmatic membrane. This happens by means of intracellular signal transduction pathways. Some signaling pathways are deregulated in the cancer cells due to the oncogen expression, or to the over-expression or repression of control elements. Thus, the analysis of these pathways can help to understand the modifications of the cell on its way from a normal

healthy state to a transformed one and finally to cancer generation or death. In particular, we are interested in the analysis of intracellular signaling pathways in which Ras protein participates. This protein is the convergence point for different pathways related to cancer development.

In Figure 1 we present the pathways generated by Ras effectors (Raf, RaIGDS, PLC and PI3K) and the pathway connecting Ras with a cell cycle control. The participating agents are proteins, hormones, growth factors, etc. In a model we call them elements. An arrow from one element to another means that the first "determines" the state of the second one (e.g. one element leads to the activation, phosphorylation or complex formation of a second element). If the element has two or more arrows entering into it, this means that its state is determined by a cooperative action of all the elements connected to it. An open arrow indicates an inhibitory effect. [Macara, 1991; Lacal and McCormick, 1993; Feig, 1993; Lowy, 1993; Burgering and Bos, 1995; Kawasaki *et al.*, 1996; Denhardt, 1996; Khosravi-Far *et al.*, 1998; Cárdenas-García, 2002; McMahon and Woods, 2001; Crespo and León, 2000].

The interaction picture, as shown in Figure 1, is quite complicated. Even if we consider a simple Boolean-type model, the number of possible combinations of states for the model is very large. For the analysis, it is desirable to divide interacting elements into groups that can be analyzed more or less independently. In this paper we present a mathematical procedure to get such organization and we discuss the behavior of simple Boolean-type dynamic models within the framework of a methodology reported at the ISMB 2000 conference [Cárdenas-García *et al.*, 2000]. We explain this in more detail and discuss the consequences for models with attractor cycles. Finally, we apply the method to the model of interacting proteins and show that this procedure allows the introduction of a useful organization.



**Figure 1:** Interactions between elements of Ras-related signal transduction pathway. Solid arrows correspond to positive interactions (activation, phosphorylation, etc.), open arrows to negative ones (inhibition, dephosphorylation, etc).

## Ras-related signal transduction pathways

Extracellular stimuli of diverse nature, whether mediated by cell surface receptors of the tyrosine kinase, non-receptor tyrosine kinases, or G protein-coupled types, all converge in Ras, which is transiently activated upon receipt of these stimuli. After having been activated, Ras-GTP relays the signals downstream through an association with a growing number of signaling intermediaries. In mammalian cells, Ras has three well-established effectors: Raf, PI3K and RaIGDS. A growing number of Ras-interacting proteins such as MEKK1, Rin1, AF-6 and Nore-1 await further confirmation as Ras effectors as well.

Raf is involved in cell differentiation, PI3K in the cell growth and RaIGDS in both cell growth and regulation

of cytoskeletal maintenance. Ras participates in cell cycle regulation through p53. If we put together all these pathways, we obtain a complex interaction model as shown in Figure 1. [Cárdenas-García, 2002; McMahon and Woods, 2001; Crespo and León, 2000].

## Mathematical basis of the theory

Mathematically speaking, all the elements involved in signal transduction pathways constitute the *set* under study. We introduce the concept of *self-determined subsets (SDS's)* of the set.

### Definition

A subset is a Self-Determined Subset (*SDS*) if its elements depend only on elements of the same subset.

Any *SDS* is autonomous in the sense that it is entirely separate from all other system elements and it can be analyzed independently. On the other hand, a *SDS* can act as an external influence on the elements that do not belong to it. We show how one can systematically find all self-defined subsets of a system for a given type of interaction. These self-defined subsets form an abstract topology of the set. Knowing the set's topology, we can introduce hierarchic structure. We will now make some statements about *SDS's* with corresponding proofs.

### Statement 1

Intersection or union of any two *SDS's* is also a *SDS*.

The fact that the union of two *SDS's* is also self-determined is quite evident. The proof for the case of the intersection is as follows: Let us consider element A, which belongs to two *SDS's*,  $S_1$  and  $S_2$ . If A depends on the element B which belongs to  $S_1$ , but not to the intersection, then  $S_2$  cannot be self-determined since there is the element B which determines the element from  $S_2$  but does not belong to  $S_2$ . Thus, the element belonging to intersection can depend only on elements also belonging to the intersection, which proves the statement.

### Statement 2

All *SDS's* constitute a topology of the set.

We can identify self-determined subsets with open subsets. From Statement 1, any intersection or union of *SDS's* is also a *SDS*, thus this property of topology is fulfilled. [O'Neil, 1982]. The whole set is obviously self-determined, and the empty set is a *SDS* as well. Thus, all usual requirements for topology are satisfied.

Now we give an explicit way to construct the topology. For this let us define the minimal *SDS*.

### Definition

The minimal *SDS* including the given element A is the intersection of all *SDS's* that have A as their element.

In topological terms, it is the minimal neighborhood of A, which always exists for a finite set. Practically, the minimal *SDS* can be constructed by making a list that initially includes only the element A. At the second step, it includes all elements that define A. At the third step, it includes also all the elements that define elements defining A, etc. The list will stop to grow and it will be the minimal *SDS* of A. The two extreme cases are the complete set, and the element A, if it does not depend on anything except itself. It is possible that two different elements A and B can have the same minimal *SDS*.

We can consider the Figure 1 as a graph. If the direction of arrows is inverted, the task of finding the minimal **SDS** of *A* is equivalent to traversing the oriented graph starting from the vertex *A*. Two popular algorithms for the purpose are Depth-First Search (DFS) and Breadth-First Search (BFS). The complexity of both of them is known to be  $O(e+v)$ , where  $e$  is the number of edges and  $v$  is the number of vertices in the graph. Thus, the complexity of algorithm for finding all **SDS**'s is  $O(N(N+K))$  where  $N$  is the number of elements and  $K$  is the number of connecting lines.

#### Statement 3

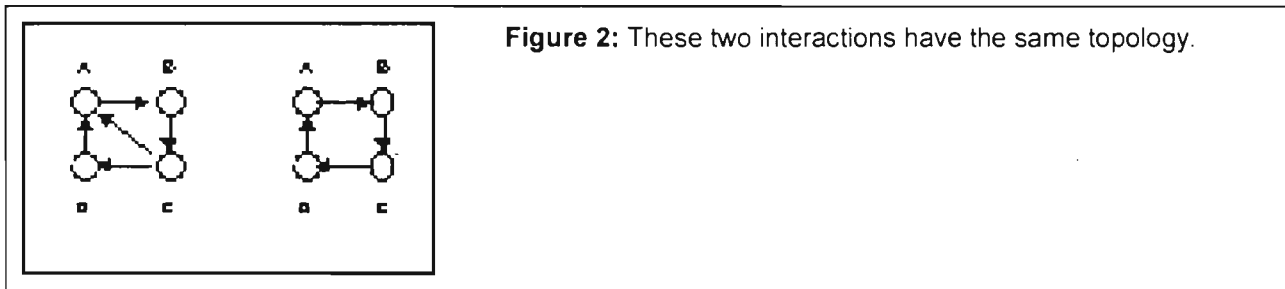
Any **SDS** can be represented as a union of minimal **SDS**'s.

Let *S* be an **SDS**. If the element *A* is included in *S*, then the minimal **SDS** of *A* also belongs to *S*. Thus, *S* is the union of minimal **SDS**'s of all its elements. (The union of **SDS**'s is also a **SDS** according to the Statement 1). From statement 3 it follows that minimal **SDS** form what is called a *base* of topology, [O'Neil, 1982]. Thus, from the practical point of view, it is sufficient to determine minimal **SDS**'s. All the other **SDS**'s are obtained as unions of minimal **SDS**'s.

#### Statement 4

It is possible to build a system of dependencies corresponding to any given topology.

The procedure for doing this is the following. For a given finite topology, we can determine a minimal open subset including the element *A* as the intersection of all open subsets including *A*. Then we can say that *A* is defined by all elements that enter in this minimal subset. Repeating the procedure for all elements, it is possible to determine all interactions. It is easy to see that generally the interaction corresponding to a given topology is not unique (Figure 2), thus the procedure gives only one of many possible interactions. Statements 2 and 4 mean that for finite systems the topological structure and interaction between system elements are directly related to one another.



**Figure 2:** These two interactions have the same topology.

### Building the hierarchical structure

Taking all **SDS**'s, as we did above, gives the most complete or richest topology. The disadvantage is that two **SDS**'s can share some elements. It can be convenient to choose a smaller topology by joining some **SDS**, so that two **SDS**'s have either zero intersection or their intersection is equal to one of them. Such topology base introduces in the system a hierarchy; and we give here a method of building it.

#### Definition

The  $0^{\text{th}}$  order **SDS** as a minimal one that has no **SDS**'s inside.

The intersection of two different  $0^{\text{th}}$  order **SDS**'s is empty (otherwise the intersection **SDS** will lie inside intersecting subsets, which contradicts the definition).

**Statement 5**

There is at least one 0<sup>th</sup> order **SDS** in a set.

**Proof:** Let us take any **SDS**. If it has no **SDS** inside, it is what we need. If it has, we repeat the search for the inner **SDS**. Finally, as the set is finite we will arrive to the 0<sup>th</sup> order **SDS** we are looking for.

The 1<sup>st</sup> order *minimal SDS* is defined as having only 0<sup>th</sup> order **SDS** inside. The intersection of two 1st order minimal subsets is the 0<sup>th</sup> order or empty. Similar to Statement 5, if the whole set is not 0<sup>th</sup> order, there will always be at least one 1<sup>st</sup> order **SDS**. Generally speaking *N<sup>th</sup> order minimal SDS* has only **SDS** up to N-1<sup>th</sup> order inside.

Now let us construct the hierarchy. We start from 0<sup>th</sup> order subsets. At the next step we add 1st order minimal subsets and join all those that have non-empty intersections. As a result, we get subsets formed of minimal ones with order less than 2 that do not intersect. At the next step we add 2<sup>nd</sup> order minimal subsets and unite them in the same way. Following this procedure, we will finally get the base having the desired property. The whole structure can be visualized as a topographic map with height levels, or a pyramid. The subset of elements within any height line is self-determined. Elements that correspond to mountain peaks belong to the 0<sup>th</sup> order subsets, and only themselves determine them. The elements with a higher order stand in a lower position.

## Consequences of the topological structure for attractors of Boolean models

Boolean systems are regarded as simple examples of dynamic systems. Initially every element can be in a finite number of discrete states. The state of the system (or subsystem) is described by the states of all its elements. If the number of elements is finite, and every element has a finite number of states, then the total number of states for the system is finite. Of course, this number can easily be very big. If we have N elements with 2 states each, it is  $2^N$ . This gives a result of  $2^{66}$  (approximately  $7.4 \times 10^{19}$ ) states for our model system, which includes 66 elements with two states each.

The evolution of the Boolean system is given by a set of rules that determine the state of the system at the next step from its state at the previous step. The final result of the development is the attractor, which can be the single equilibrium state, or the cycle (a set of states which follow each other, and the last state of the cycle leads to the first one).

The existence of the hierarchical structure imposes certain limitations on the attractor cycle. Let us consider the example of the **SDS**  $P = V + S$ , which includes **SDS** S. Note, that V is generally not a **SDS**, because the elements of V can depend both on the elements of S and on the elements of V. As S is autonomous, it defines its final states. Let us suppose that it forms a cycle with states  $S_1, S_2, \dots, S_M$ . Then a pair (V, S) defines the total state of P. At the first moment it is  $(V_1, S_1)$ . After the cycle of S is completed, the state of P will be  $(V_2, S_1)$ . Finally, after some integer number of cycles of S, the state of P will return to its initial value. So, the number of states in the internal cycle of S must be the divisor of this number in the cycle of P.

The practical rule for finding the attractors for **SDS** P is the following. First, we must find attractors for **SDS** S, which is included in P. Then, we need to trace development for initial states that include the attractor states of S and all possible states of V. For a given state  $V_1$ , we must take all possible attractor states  $S_1, S_2, \dots, S_M$ , and then we have to look for their evolution, considering that states S can form a cycle.

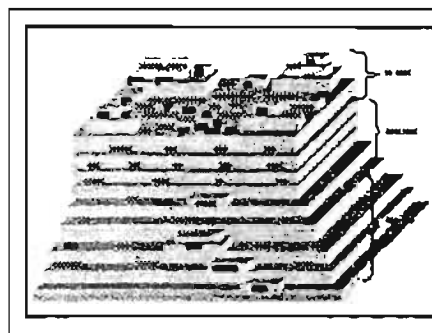


The existence of subsets can greatly simplify the search of possible attractors. Let us consider that  $P$  has  $2N$  elements, and  $S$  has  $N$  elements, with two states each. Then, the exhaustive search in  $P$  requires considering  $2^{2N}$  initial states. If we first consider the inner  $SDS$ , the search requires  $2^N$  initial states leading to presumably much smaller number  $H$  of final states. Then the number of necessary initial states is reduced to  $(H+1) 2^N$ , which represents a drastic reduction if  $N$  is big.

### Application of the topological approach to the Ras-related signal transduction pathway

We applied the algorithm described above to the system of 66 transduction pathway elements listed in Figure 1. They include proteins, hormones, growth factors, etc. We took into account pathways generated by 3 Ras effectors (Raf, RaIGDS and PI3K), and the pathway connecting Ras with a cell cycle control.

We developed a computer program, which took as the input a list containing names of elements and a list of the elements they depend on. At the output, the program gives a list of self-defined subgroups and their orders. After this the hierarchic structure can be easily built. This is shown in Figure 3.



**Figure 3:** The hierarchical structure for Ras-related signal transduction pathways.

Though the algorithm is abstract, the structure we obtained has quite clear biological meaning. In the upper part of the pyramid we find extracellular elements that activate Ras, their receptors, and two of Ras effectors Raf1 and PI3K that are involved in cell growth regulation. Some important kinases are also in this group. In the middle part, we find the elements that are related to the third Ras effector, RaIGDS and are involved in the regulation of cytoskeleton. The group also includes transcription factors. The lower part elements are involved in a cell cycle regulation. In this group there are two transcription factors for cell growth regulation. Generally, the hierarchy does not follow the direction from the extracellular elements to transcription factors. It is organized mostly by cell function regulation types.

The part of the upper group, which does not belong to lower-order  $SDS$ 's, includes 20 elements. This means, that if the Boolean model is used, all final states of this group can be calculated in a reasonable amount of time, similar to the way it was done in [Cárdenas-García *et al.*, 2000].

Thus, the algorithm does not only facilitate the division of elements in easily analyzable sub-groups, but it also allows us to visualize the elements involved in different biological functions.

### Discussion and conclusions

We suggested the method of introducing a structure in the model of many-element interactions. The

method is based on abstract topology. It permits to isolate, in a systematic way, the subsets of elements, which are self-determined, and to build the hierarchy. The computation time, in many cases, depends exponentially on the system size; isolating subsets, which can be analyzed separately, can significantly reduce the amount of calculation.

As an example of the application, we consider the cellular signal transduction pathway related to the oncogen protein Ras. The abstract procedure that we apply permits biologically meaningful division and facilitates further analysis. From the discussion after Statement 2 it follows that the computation time for finding all minimal *SDS*'s generally behaves as  $O(N^2)$ . Hierarchy determination complexity is  $O(NM^2)$ , where  $M$  is a number of orders. The calculation for 100 elements with a program written in C was taking fraction of second.

The topological analysis can be useful when the interactions between different elements of the system are known, but it is not clear how to group these elements. Of course, in some situations (for example if all elements are included in a feedback loop), the topology will be trivial. One can expect task simplification only if the topological structure is rich enough. The hierarchy we introduced is not the only possibility of uniting self-dependent subsets; other procedures of uniting subsets may be better for different situations.

## Acknowledgments

We thank Sergio Cárdenas for help with manuscript preparation, Pedro Pablo González Pérez for useful discussions, and a reviewer for pointing out the analogy with graph traversing algorithm.

## References

- Burgering, B. M. and Bos, J. B. (1995). Regulation of Ras-mediated signalling: more than one way to skin a cat. *Trends Biochem. Sci.* **20**, 18-22.
- Cárdenas-García, M., Lagunez, J. and Korneev, N. (2000). Efficient Attractor Analysis Based on Self-Dependent subsets of elements - An Application to Signal Transduction Studies. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **8**, 86-92.
- Cárdenas García, M. (2002). Search of therapeutic targets using a model of intracellular signals mediated by RAS. D. Thesis, Facultad de Química e Instituto de Química, UNAM.
- Crespo, P. and León, J. (2000). Ras proteins in the control of the cell cycle and cell differentiation. *Cell. Mol. Life Sci.* **57**, 1613-1636.
- Denhardt, D. T. (1996). Signal-transducing protein phosphorylation cascades mediated by Ras/Rho proteins in the mammalian cell: the potential for multiplex signalling. *Biochem. J.* **318**, 729-747.
- Feig, L. A. (1993). The Many Roads That Lead to Ras. *Science* **260**, 767-768.
- Kawasaki, H., Moriguchi, T., Matsuda, S., Li, H. Z., Nakamura, S., Shimohama, S., Kimura, J., Gotoh, Y. and Nishida, E. (1996). Ras-dependent and Ras-independent activation pathways for the stress-activated-protein-kinase cascade. *Eur. J. Biochem.* **241**, 315-321.
- Khosravi-Far, R., Campbell, S., Rossman, K. L. and Der, C. J. (1998). Increasing Complexity of Ras Signal Transduction: Involvement of Rho Family Proteins. *Adv. Cancer Res.* **72**, 57-107.
- Lacal, J. C. and McCormick F., eds. (1993). *The ras Superfamily of GTPases*. CRC Press, Boca Raton, Ann Arbor, London, Tokyo

- Lowy, D. R. (1993). Function and Regulation of Ras. *Ann. Rev. Biochem.* 62, 851-891.
- Macara, I. G. (1991). The Ras superfamily of molecular switches. *Cell Signalling* 3, 179-187.
- McMahon, M. and Woods, D. (2001). Regulation of the p53 pathway by Ras, the plot thickens. *Biochim. Biophys. Acta* 1471, M63-M71.
- O'Neil, P. V. (1982). *Fundamental concepts of topology*. 2nd edition, Gordon and Breach, New York.

Figure 1

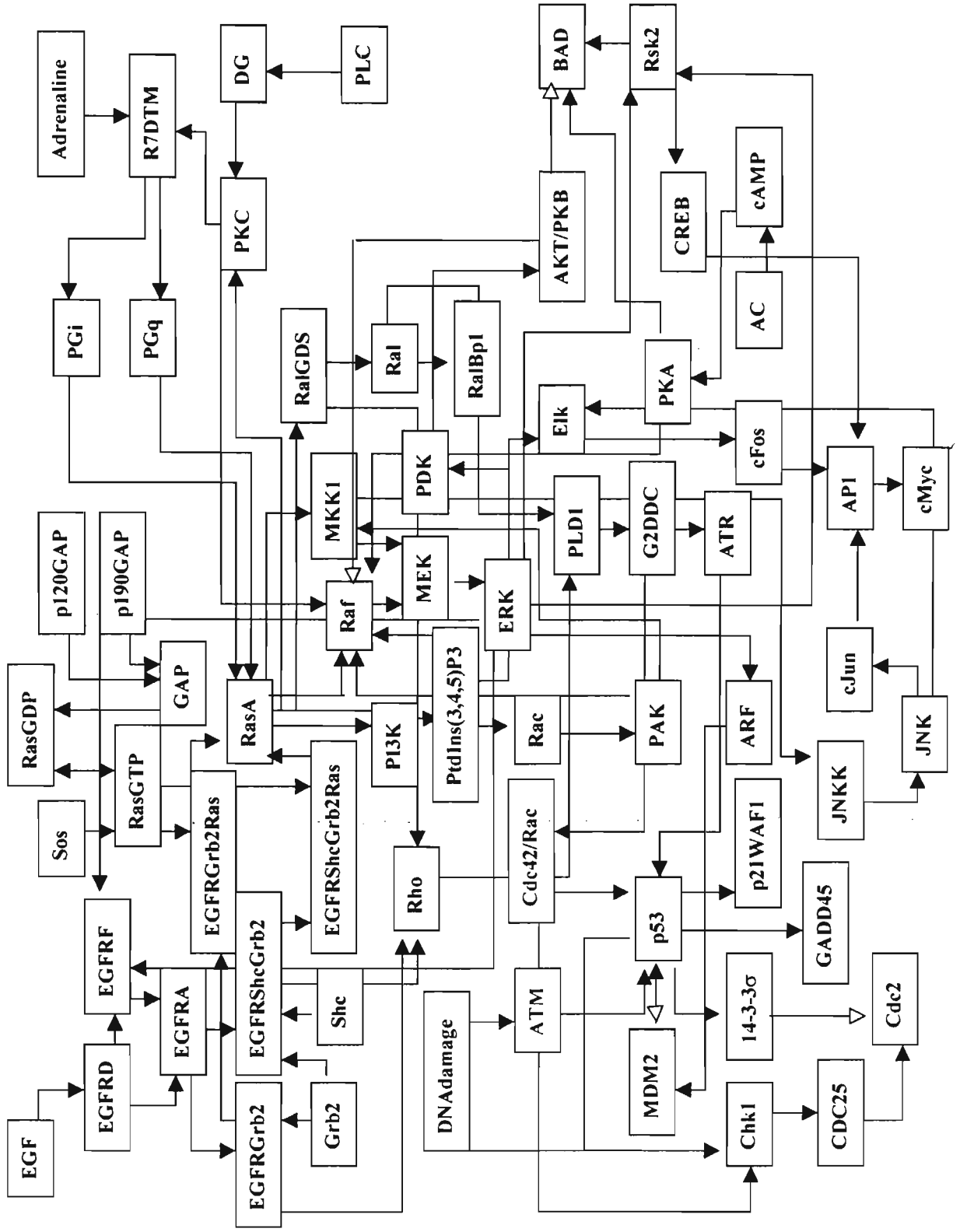


Figure 2

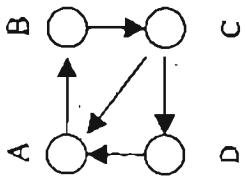
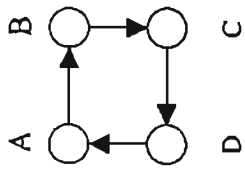
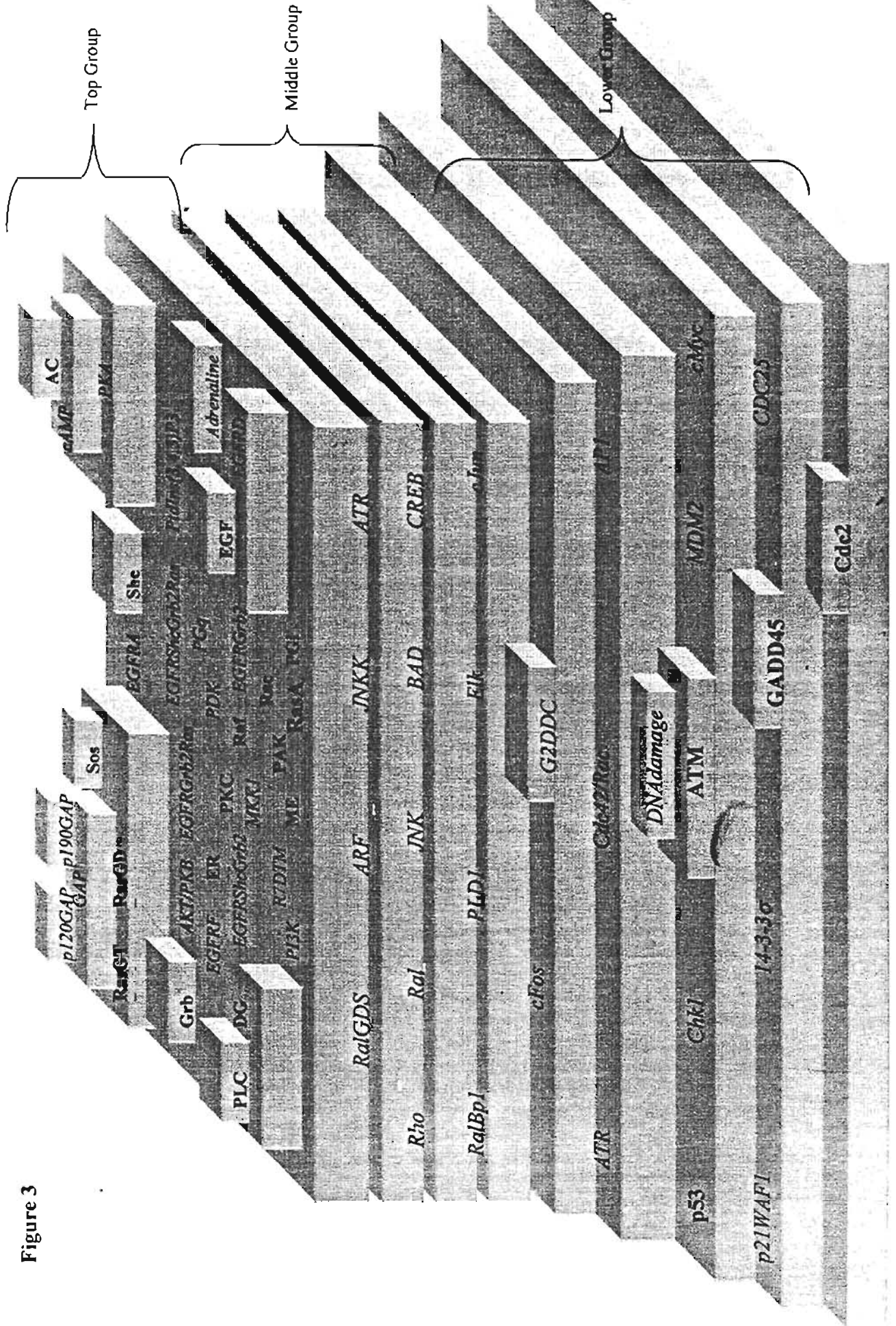


Figure 3





## Cellulat: an agent-based intracellular signalling model

Pedro Pablo González\*, Maura Cárdenas, David Camacho,  
Armando Franyuti, Octavio Rosas, Jaime Lagúnez-Otero

*Instituto de Química, Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510 Mexico, D.F., Mexico*

### Abstract

The theory of behaviour-based systems (or autonomous agents) constitutes a useful approach for the modelling of intracellular signalling networks. In this sense, a cell can be seen as an adaptive autonomous agent or as a society of such agents, where each can exhibit a particular behaviour depending on its cognitive capabilities. We present an intracellular signalling model obtained by integrating several computational techniques into an agent-based paradigm. Cellulat, the model, takes into account two essential aspects of the intracellular signalling networks: (1) cognitive capacities, which are modelled as the agent abilities to interact with the surrounding medium and (2) a spatial organisation, this last obtained using a shared data structure through which the agents communicate between them. We propose a methodology for the modelling of intracellular signalling pathway using Cellulat and we discuss the goal of a virtual laboratory based on our model and presently under development.

© 2002 Elsevier Science Ireland Ltd. All rights reserved.

*Keywords:* Intracellular signalling pathways; Autonomous agents; Blackboard architecture; Virtual laboratory

### 1. Introduction

Models of multiple interactions of proteins such as those found in signal transduction allow for the visualisation of the general network. They can permit the prediction of the effects at the cellular level of changes or perturbations to the system. The changes that are required to be studied include protein mutations or lesions to sections of the signalling pathway.

It is evident, due to the difficulty in having a complete overview of the cell machinery, that the more experimental data about cellular function we obtain, the more important this computational model becomes. The complexity of intracellular communication networks makes them difficult to model with any isolated technique. However, when integrating the most relevant features of known techniques in a single computational system, it is possible to obtain a robust model of signal transduction.

Within the computer sciences, artificial intelligence is one of the main areas used to model biological systems. This is due to the great variety of techniques and methods that support this

\* Corresponding author. Tel.: +52-5-622-4511; fax: +52-5-616-2203.

E-mail address: [ppgp@servidor.unam.mx](mailto:ppgp@servidor.unam.mx) (P.P. González).

research area, many of which are inherited from disciplines such as cognitive sciences and neuroscience, both of which have an important biological component. Among the main techniques of artificial intelligence and computer sciences used to model cellular signalling networks are artificial neural networks (Bray and Lay, 1994; Pritchard and Dufton, 2000), Boolean networks (Edwards, 1995), Petri nets (Fuss, 1987), rule-based systems (Cárdenas-García, 2001; Lagúnez-Otero, 1998), cellular automata (Edwards, 1995; Wurthner et al., 2000), and multi-agent systems (Fisher et al., 1999; Paton et al., 1995; Schwab and Pienta, 1997).

Virtual laboratories have been developed in different areas such as physics, electronics, robotics, ethology, physiology, etc. to reproduce experiments and can be extremely useful for education and training. In particular, they can help visualise emergent properties of diverse systems. One of the reasons for our interest in the analysis and understanding of the signalling pathways is to observe the generation of new properties. In principle, it is possible to observe these processes in a virtual laboratory based on the paradigm that we are presenting here. Indeed as we have done in previous work (Gershenson et al., 2000).

The theory of behaviour-based systems or autonomous agents (Brooks, 1986; Maes, 1994) constitutes a useful approach for the modelling of intracellular signalling networks (González et al., 2000b). The model permits one to take into account communication between agents via a shared data structure, in which other cellular compartments and elements of the signalling pathways can be explicitly represented. In this sense, the blackboard architecture (Nii, 1989) becomes a natural paradigm to simulate the intracellular milieu.

In this paper, we propose a robust model of intracellular signalling obtained by joining the main structural and functional characteristics of behaviour-based systems and the blackboard architecture. We consider that a cell can be seen as a society of autonomous agents, where each agent communicates with the others through the creation or modification of signals on a shared data

structure, named "blackboard". The autonomous agents model determined functional components of intracellular signalling pathways, such as signalling proteins and other mechanisms. The blackboard levels represent different cellular compartments related to the signalling pathways, whereas the different objects created on the blackboard represent signal molecules, activation or inhibition signals or others elements belonging to the intracellular medium.

Furthermore, when the autonomous agents are used in an intracellular signalling model, the cognitive capabilities of the signalling pathway can be taken into account. The use of blackboard architecture permits the modelling of a high level of spatial organisation exhibited by the intracellular signalling networks and structural characteristics of the intracellular medium.

The rest of the paper is structured as follows: Section 2 presents an overview of the information processes of signalling intracellular networks. In Section 3, some computational models for intracellular signalling are briefly reviewed; Section 4 presents the intracellular signalling model proposed. By instantiating the paradigm with the EGF receptor signalling pathway, Section 5 describes the methodology used. Section 6 is focussed on the final goal of this work: the creation of an intracellular signalling virtual laboratory.

## 2. An overview of signalling intracellular networks

In multicellular organisms, cell decisions about survival, secretion of some molecule type, growth, differentiation, senescence or death, are made on the basis of external signals. These stimuli include growth factors, hormones, cytokines, neuropeptides, ions and cell-cell adhesion signals. It is essential to integrate information from multiple sources for the cell to respond appropriately to a wide range of conditions, and thereby obtain the adaptability and consequential survival of the organism.

A specific set of cell surface receptors allows cells to perceive changes in the extracellular environment. One of the most exciting recent developments in molecular and cell biology has



been the step by step construction of protein signalling cascades that trace the path of the effects of an extracellular stimulus from the external membrane to the cell nucleus. A protein processing network mediates the transmission of extracellular signals to their intracellular targets.

In general, the external signals are transmitted to the interior of the cells through receptors activating diverse signalling pathways. They can follow a single way to generate an answer, or branch out and give rise to others pathways to do so. From this second way of signal transduction emerges an interconnected network, because pathways usually corresponding to different stimuli cross and generate alternative trajectories.

The intracellular signalling implies several molecular processes. This processing can be as simple as the direct introduction of the signal to the nucleus and the activation of the expected genes for a specific response. On the other hand, they can be very complicated and include multiple stages. For example, the receptor could activate effector proteins like membrane tyrosine kinases directly, or through some intermediaries as G proteins and adapter proteins. They activate the cytosolic protein network described below and in turn, into the nucleus a specific set of transcription factor proteins, which determine the transcription of genes involved in the specified cellular response.

Several pathways of signal transduction systems have been described for the extracellular signal membrane receptors and they could be included in to two general ones: receptor linked G proteins and receptor linked effector protein, and some specific cases include:

- Adenylate cyclase system
- Phosphoinositides–calcium system
- Mitogen activated protein kinases system (MAKPs)
- JAK/STAT activation system
- The sphingomyelin–ceramide system
- The ionic channel function receptor system

The MAPK cascade is a central signal transduction pathway that is activated by growth factors, and is known to be involved in diverse cellular functions. In particular, we modelled the EGF

activated pathway which crosstalks with the mentioned systems. The processes related to ageing and cellular proliferation constitute our main interest.

It should be mentioned that we are not modelling a specific cellular type. Once we have the functional model with all the elements involved in one pluripotential cell, it is relatively easy to extend it to cellular types.

### 3. Computational models for intracellular signalling

Computational models in signal transduction pathways have been made using different points of view. Each research group chose the approach which seemed best for it. The different models have been proposed according to the perspective that each research group has of the pathway they want to model. This perspective involves a range from the types of information processing present at cellular level, such as sequential, parallel, distributed, concurrent and emergent; to the cognitive capabilities exhibited by certain signal transduction pathway component, such as memory, learning, pattern recognition and handling fuzzy data.

In this sense, several computational approaches have been proposed to model the cellular signalling pathways, such as artificial neural networks. Table 1 summarises the main characteristics of these computational approaches, taking into account the idea behind the approach, the cognitive capabilities that can be modelled, types of present information processing, and the part of the cellular signalling to be modelled.

From each model we took some points that we considered suitable for our work. Thus, our approach sees the dynamics of the cellular signal transduction in terms of a collection of autonomous agents communicating among them through a shared data structure, where each agent is implemented as an artificial neural network, a package of production rules, a Boolean network or a molecular automata, depending of the complexity of the task carried out by the agent and the knowledge degree or cognitive capabilities required by it.

Table 1  
Summary of different computational approaches for the modelling of cellular signalling

Computational approach	Idea behind the approach	Cognitive capabilities	Types of information processing	Cellular signalling pathway modelled	References
Boolean networks	The cell can be modelled as a network of two-states components interacting between them. The state of each component in the network depends of a particular Boolean function.	Boolean logic computation	Parallel	Intracellular signalling, genetic networks	Kauffman, 1991; Edwards, 1995; Karp and Paley, 1994; Armas et al., 2000
Expert systems	The interactions (activation, phosphorylation, etc.) between signalling network components are modelled using production rules.	Knowledge-based inference and deduction	Sequential, parallel	Intracellular signalling	Lagúnez-Otero, 1998; Cárdenas-García, 2001; Takai-Igarashi and Kaminuma, 1998
Cellular automata	The interaction between cells or molecules is modelled as a matrix, where the state of an element of the matrix depends on the states of the neighbouring elements.	None	Parallel	Extracellular signalling, intracellular signalling	Marijuan, 1994; Levy, 1992; Wurthner et al., 2000
Petri nets	The cell is seen as a connected graph with two types of nodes. One type represents elements, such as signalling molecules and proteins, whereas the other type represents transitions, such as activations.	None	Sequential, concurrent	Intracellular signalling	Holcombe, 1994
Artificial neural networks (ANN)	The proteins in signalling networks are seen as artificial neurons in ANN. Like an artificial neuron, a protein receives weighted inputs, produces an output, and has an activation value.	Memory, learning, pattern recognition	Distributed, emergent	Protein networks, intracellular signalling	Bray, 1990; Bray and Lay, 1994; Bray, 1995; Pritchard and Dufton, 2000; Paton, 1993
Distributed systems (agents)	The cell is seen as a collection of agents working in parallel. The agents communicate between them through messages.	Memory, learning, pattern recognition, handling fuzzy data, adaptive action selection	Distributed, parallel, emergent	Intracellular signalling	Paton et al., 1995; Fisher et al., 1999
Continuous models	The behaviour of proteins populations depend on their physicochemical variables as concentration and affinity. The dynamic of the model is represented by this variable changes with respect to continuous time, all this is implemented through differential equations.	None	Distributed, emergent	EGF and NGF receptors linked to MAPK kinases. Ca, PKC, PKA, CaM and CaMKII systems	Brighman and Fell, 2000; Ying and Ferrell, 1996; Bhalla and Iyengar, 1999

#### 4. Cellulat: an intracellular signalling model

##### 4.1. Initial considerations

Our model considers that the cell as well as its components can act as agents with cognitive capabilities. As such, the cell perceives its external environment through its surface receptors and acts on them by generating internal messages, which will be able to affect its behaviour and ultimately the behaviour of the organism of which the cell is a part. The cell's external environment is also dynamic and complex, in which an extensive range of signals and combinations of these can exist. The cell decides how to relate the received external signals (signalling molecules) to its internal signals (activation/inactivation signals), so that their goals (cellular functions) can take place. Finally, the cell is able also to adapt to its environment. That is, it is constantly adapting its behaviour to the changes that take place in the environment and to the signals received.

We consider the main signalling pathway components as adaptive autonomous agents and require a medium through which the communication among them takes place, having the results produced by these interactions, e.g. activation and phosphorylation levels, recorded. The blackboard architecture (Nii, 1989; González and Negrete, 1997; Negrete and González, 1998) provides an ideal paradigm to satisfy the requirements mentioned before. The idea behind the blackboard architecture is to handle the problem of shared information between multiple expert agents.

The blackboard architecture is defined in terms of three basic components: (1) a set of independent modules, named knowledge sources, which contain specific knowledge about a problem domain, (2) a shared data structure, named blackboard, through which the knowledge sources communicate to each other, and (3) a control mechanism, which determines the order in which the knowledge sources will operate on the blackboard.

The knowledge sources can be built as adaptive autonomous agents when the knowledge is replaced by a behaviour repertoire and an adaptive action selection mechanism has been allowed to control the interactions that the first ones execute

on blackboard (González, 2000). In this way, these agents can model intracellular signalling pathway components such as receptors, non-catalytic proteins, and enzymes. On the other hand, the blackboard structure allows for a modelling of the intracellular medium structure through its levels. That is, different cellular structures involved in the intracellular signalling could be mapped to different blackboard levels. The blackboard structure also provides a continuous trace of all interactions occurring among the agents. This trace can be seen as a topologic map distributed between the blackboard levels and its elements reflect the different degrees of activation, inactivation, phosphorylation and desphosphorylation that characterise the intracellular signalling in a given time.

These previous considerations constitute the functional and structural essence of the intracellular signalling model described below.

##### 4.2. The model

Our proposal consists in modelling the cell as an autonomous agent, which in turn is composed by a society of these agents, where each agent communicates through a blackboard with others. The model constitutes a refinement and adaptation of an action selection mechanism structured on a blackboard architecture previously developed by us, called Internal Behaviours Network (IBeNet) (Gershenson et al., 2000; González, 2000; González et al., 2000a). Although the IBeNet was initially built to control the action selection in autonomous agents (physical robots, animats, or artificial creatures simulated on a computer), it constitutes a working environment for the bottom-up modelling of information processing systems characterised by: (1) modularity, (2) parallel, distributed and emergent processing, (3) coordination and opportunistic integration of several tasks in real time, (4) use of several abstraction or context levels for the different types of information that participate in the processing network, (5) decision making, and (6) cognitive capabilities such as adaptive action selection, memory and learning (González, 2000).

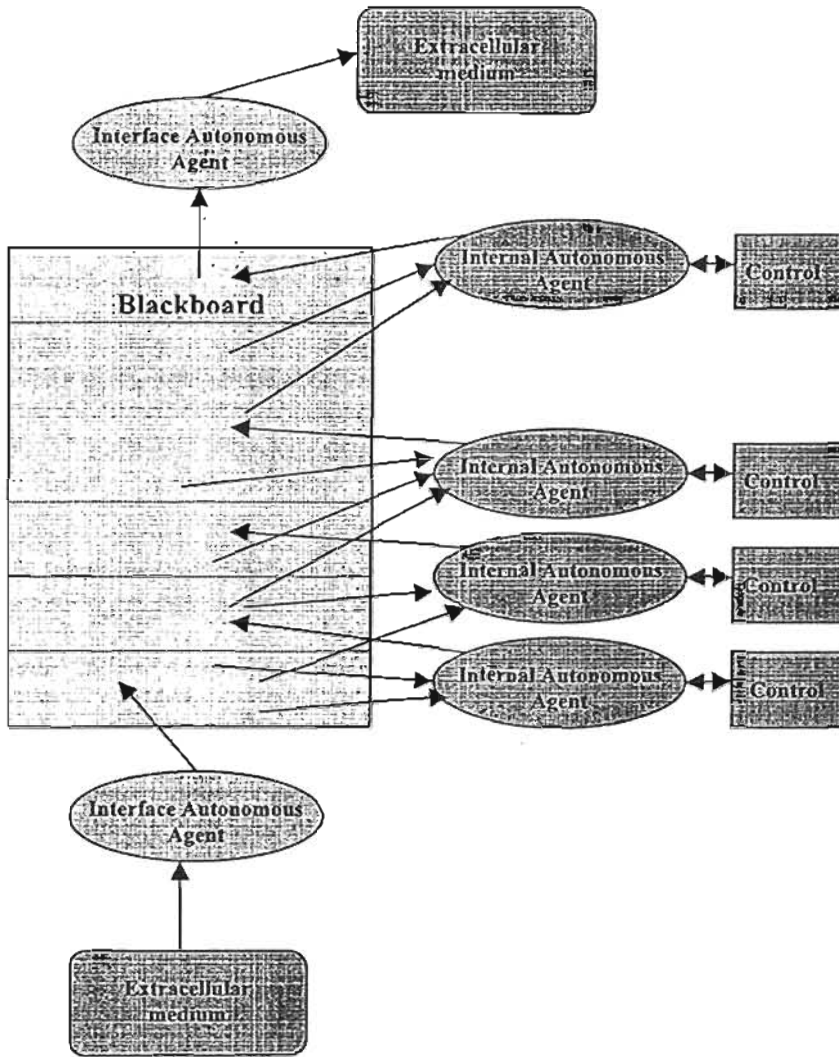


Fig. 1. Architecture of Cellulat.

As mentioned above, the intracellular signalling model has been named Cellulat (Fig. 1). Four main components define the structure:

- the blackboard
- the internal autonomous agents
- the interface autonomous agents
- the control mechanisms

Different levels in the blackboard correspond to different static and nested cellular compartments through which the signal transduction takes place. In the cell this compartmentalisation allows for the

self-organisation of structures from stochastic process, such as the creation of signalling cascades from diffusion. The compartmentalisation is characterised by the following properties: (1) interaction restrictions for the different signalling components and (2) different rates for diffusion reaction processes. The above mentioned can be obtained using the blackboard architecture in Cellulat. The solution elements recorded on the blackboard represent two main types of intracellular signals: signalling molecules and activation/inactivation signals. Both types of signals are

synthesised or created by internal autonomous agents and these, either directly or indirectly, promote the activation/inactivation of other internal autonomous agents. Other types of cellular elements or structures can be represented on the blackboard as well.

The term "internal autonomous agent" has been used to identify the proteins whose tasks deal with the creation or modification of signals on the blackboard. An internal autonomous agent obtains a signal or combination of signals from a determined blackboard level and transduces these into other signals on the same or other blackboard level. The way in which a signal is transduced depends on the cognitive capabilities of the internal autonomous agent. Internal autonomous agents model components of the intracellular signalling network such as proteins and other molecules necessary to carry out the signal transduction.

On the other hand, the function of an interface autonomous agent is to establish the communication between the blackboard and the external medium (they are similar to sensors or actuators in a behaviour-based system). The recognition of a signal by an interface agent depends both on the signal characteristics and the cognitive capabilities of the agent. In this manner, each external signal or combinations of these are recognised by different agents in different ways. Each agent, independently of its type, is composed of three main components: a set of attributes, a repertoire of behaviours and an action selection mechanism. The set of attributes and its values define the main characteristics of the agent. These characteristics reflect the state of the agent in a given time. The repertoire of behaviours involves all possible actions that the agent can execute. Each behaviour has a condition part and an action part. The condition part must be fulfilled in order for the corresponding action part to be executed. The condition part is satisfied when certain attribute values, signals recorded on the blackboard and/or external signals to blackboard are observed. The action part establishes the action that the agent must execute. The execution of an action could consist in the modification of attributes values of the agent or the others agents and/or in the

creation or modification of signals on the blackboard. The way in which both parts are linked depends on the complexity of the intracellular component modelled by the agent. For this reason, agents which model complex components could use more advanced techniques, such as neural networks, or any combination of other techniques, to link them. Agents which model less complex components could use less sophisticated but useful techniques, such as production rules or Boolean networks.

The role of the action selection mechanism is to select, at every moment in time, the most appropriate action among all possible actions that the agent could execute. This selection is done taking into account both internal and external signals. The internal signals are specified through attribute values of the agent, whereas the external signals to agent are the signals or elements recorded on the blackboard or other signals outside the blackboard. This last kind of signal corresponds to the extracellular signals. The extracellular milieu is represented in our architecture as other data structure, which is independent to blackboard. The extracellular signals are elements belonging to this data structure.

The work of both types of agents is event-directed. This is, each element registered on the blackboard (e.g. secondary messengers or activation/inactivation signals) or each extracellular signal perceived (e.g. ligands) constitutes an event, which could satisfy the condition part of certain behaviours in one or more autonomous agents.

In Figs. 2 and 3 the attributes and behaviour repertoire of two intracellular signalling components in Cellulat are shown. Fig. 2. shows an interface autonomous agent whose identity, attributes and behaviour repertoire correspond to a type I receptor. As can be seen in Fig. 2, *Free current concentration (FCC)* and *Vector of ligands (VL)* are two kinds of attributes of this agent, whereas *External-Interaction* is a behaviour. The condition part of this behaviour is fulfilled when there is a ligand in the extracellular milieu; this ligand is specified in the VL attribute of receptor agent and both attributes values FCC belonging to receptor agent and FCC belonging ligand object exceed certain thresholds. Fig. 3 represents an



Fig. 2. Receptor model in Cellulat.

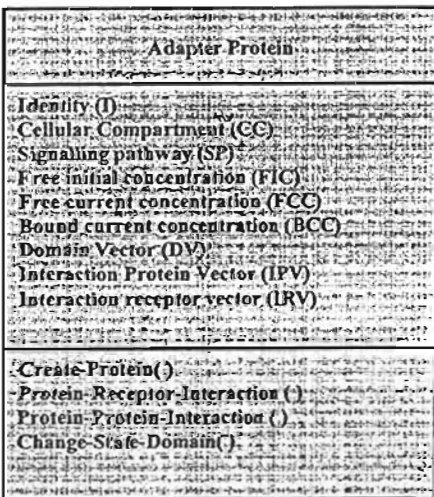


Fig. 3. Adapter protein model in Cellulat.

internal autonomous agent which models an adapter protein. As shown in Fig. 3, FCC and *Interaction receptor vector* (IRV) are two kinds of attributes of this agent, whereas *Protein-Receptor-Interaction* is a behaviour. The condition part of this behaviour is fulfilled when a receptor activa-

tion signal has been created on the blackboard and this receptor is specified in the IRV attribute of the adapter protein agent and both attributes values FCC belonging to adapter protein agent and FCC belonging to receptor agent exceed certain thresholds. In both cases, the behaviours of the agents shown in Figs. 2 and 3 have been modelled using production rules, as a first approach.

The architecture of Cellulat involves two types of control mechanisms. One of these control mechanisms is part of the agent structure and it was described above as the action selection mechanism. The other control mechanism is external to the agent and its role is to carry out the scheduling among different actions to be executed by different agents in a given time. This type of control involves a competition process, which takes place in order to select the more appropriate action to be finally executed. Our agents act concurrently and the logical consequences of their interactions generate a logical relative time sequence.

#### 4.3. Cognitive capabilities of Cellulat's agents

It is known that certain types of proteins exhibit several cognitive capabilities such as pattern recognition, memory, and handling fuzzy data (Fisher et al., 1999; Bhalla and Iyengar, 1999). The evolution, learning and emergence of properties have also been suggested in protein networks (Pritchard and Dufton, 2000). In Cellulat, both types of autonomous agents exhibit several cognitive capabilities mentioned above. These capabilities allow the autonomous agents to exhibit adaptive behaviour. These cognitive capabilities are to be supported by several artificial intelligence paradigms including the following approaches: rule-based reasoning, probabilistic reasoning, artificial neural networks, Boolean networks, fuzzy logic systems, and multidimensional logic systems.

#### 4.4. Modelling spatial organisation in Cellulat

Another important aspect to consider in the modelling of intracellular signalling networks is their spatial organisation. Recent experimental data clearly demonstrate that many intracellular

signalling networks exhibit a high level of spatial organisation. Cellular functions could depend on the spatial organisation of the cell's components. New intracellular signalling models take this into account (Fisher et al., 2000). Cellulat models the spatial organisation taking into account two organisational criteria of the blackboard architecture. The first is horizontal organisation, given by the different abstraction levels of the blackboard (cellular compartments), which allow intra level signal processing. The other is vertical organisation, given by columns crossing different blackboard levels. These columns arise as a result of the adjoining work of several internal autonomous agents that operate at a same vertical section of blackboard, which cover different levels. We have named these columns "agency columns", where an agency column corresponds to a given type of signalling pathway.

Convergence and divergence of agency columns can occur, and these processes could be related with evolution and learning of the signalling networks (González, 2000). In this way, the model proposed allows modelling of intracellular signalling pathways by taking into account their spatial organisation. This is, the two information processing levels present in Cellulat (horizontal and vertical) allow us to establish what may be called a "topology preserving map".

##### 5. Modelling EGF receptor signalling pathway: an instantiation of the paradigm

The creation of an intracellular signalling network can be seen as an incremental process of definition and refinement of the pathway elements such as receptors, enzymes and secondary messengers, which are expressed as interface autonomous agents, internal autonomous agents or objects recorded on blackboard levels in Cellulat. Applying part of the methodology previously presented by González et al. (2000b) we modelled a MAPK signalling pathway activated by the epidermal growth factor (EGF). Fig. 4 shows an interaction model of the MAPK signalling pathway. The identity of proteins belonging to MAPK signalling pathway are shown in Table 2.

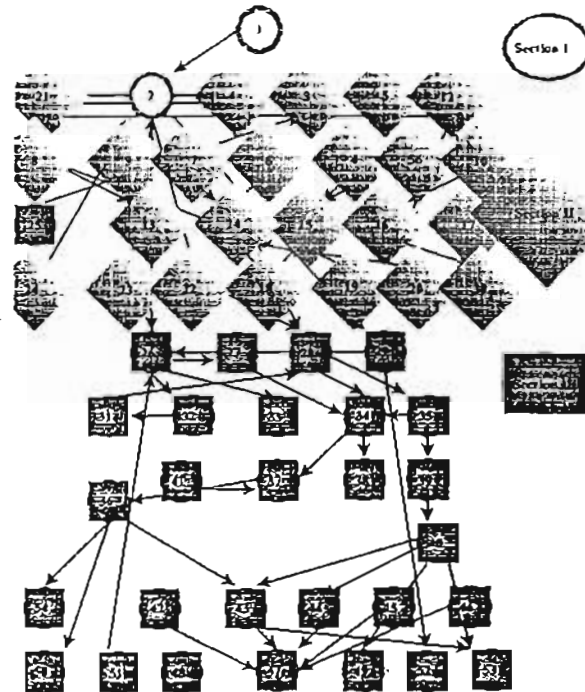


Fig. 4. MAPK signalling pathway interaction model. Numbers are individual proteins. Arrows represent the relation between two proteins. Two or more arrows mean cooperative or alternative action. Sections are indicated.

As can be seen in Fig. 4, the signalling pathway has been divided in three sections as part of the modelling methodology. The sections are the following:

- First section: from ligand to the activation of the receptor.
- Second section: from the activated receptor to the activation of cytosolic kinases.
- Third section: from the activated cytosolic kinases to the activation of transcription factors.

Here, we will present only a few steps of the previously described methodology. The EGF receptor (EGFR) signal pathway is one of the main regulators of cell proliferation, differentiation, senescence and apoptosis. The elements included in the three sections are explained below.

Polypeptides such as growth factors, differentiation factors and hormones are crucial components of the regulatory systems that coordinate develop-

Table 2  
MAPK signal transduction elements

Number	Identity	Number	Identity	Number	Identity	Number	Identity
1	EGF	15	Rap	29	JAK	43	cJun
2	EGFR	16	cdc42	30	MLK	44	JunD
3	Grb2	17	Rac	31	PKC	45	FosB
4	Shc	18	KSR	32	PDK	46	JunB
5	Gab2	19	RasGRp	33	AKT	47	API
6	Sos	20	RasGRF	34	MEK	48	IPI
7	p120	21	Src	35	MEKK	49	Elk1
8	p190	22	PLCg	36	RSK	50	CREB
9	p115	23	SHP2	37	ERK1	51	cAb1
10	cdc42 GAB	24	SHP1	38	ERK2	52	STAT
11	RapGAP	25	Pyk2	39	JNKK	53	ATF2
12	Dbl	26	PI3K	40	JNK	54	SRF
13	Rho	27	PAK	41	MKP1		
14	Ras	28	Raf	42	cFos		

The numbers correspond to Fig. 4.

ment of multicellular organisms. Many of these factors mediate their pleiotropic actions by binding to and activating the cell surface receptors with an intrinsic protein tyrosine kinase activity. It appears that ligand-induced activation of the kinase domain and its signalling potential are mediated by receptor dimerisation.

Ligand binding and the subsequent conformational alteration of the extracellular domain induce receptor dimerisation, which stabilises interaction of kinase function by molecular interaction. Receptor dimerisation permits the transmission of a conformational change from the extracellular domain to the cytoplasmic domain and it allows its interaction with the following proteins of the signalling cascade. These interactions give rise to the first section of the model. The identified components belonging to this section are the following:

- Signalling molecules: EGF
- Receptors: EGFR

Fig. 5 presents the value assignment to the attributes defining one of the previously identified components. The elements of the vector of ligands represent the ligand identity, the affinity of the receptor for the ligand and the interaction relationship between them. The elements of the vector of phosphorylation sites represent the identity of

ATTRIBUTE	VALUE
Identity	EGF
Cellular Compartment	Membrane
Signalling Pathway	
Free Initial Concentration	30.0
Free Current Concentration	30.0
Bound Current Concentration	0.0
Vector of Ligands	
Vector of Phosphorylation Sites	
Phosphorylation State	Non Phosphorylated
Activation State	Inactivated
Monomer Current Concentration	30.0
Dimer Current Concentration	0.0
Self Phosphorylated Dimer Concentration	0.0
Phosphorylated Dimer Concentration	0.0
Done	

Fig. 5. Initial values of an interface agent EGFR.

the molecule, identity of the phosphorylation protein, and the site state. When this identity is equal to the receptor identity (in this case, EGFR), it is a self-phosphorylation site.

Controlling the state of phosphorylation is an important mechanism by which signalling molecules regulate the activity of other proteins. A



common theme in molecular signalling is the kinase cascade, in which a linear series of kinases is activated by phosphorylation by a kinase. The first kinase is specifying growth factor receptor. In this phase of the signalling cascade, phosphatases that regulate the state of activation and inactivation, and adapter proteins that facilitate interaction between two proteins also participate. This gives rise to the second section of the model. The identified components belonging to the second section are the following:

- Adapter proteins: GRB2, SHC, Gab2
- Guanine interchanging proteins: Sos, p120, p190, p115, cdc42GAP, RapGAP
- G proteins: Rho, Ras, Rap, cdc42, Rac
- Enzymes: KSR, RasGRP, RasGRF, Src, PLC $\beta$ , SHP-2

Fig. 6 presents the value assignment to the attributes defining a previously identified component in second section, whereas Figs. 7–9 display the values of some attributes of this component. The elements of the interaction protein vector

ATTRIBUTE	VALUE
Identity	Grb2
Cellular Compartment	Cytosol
Signaling Pathway	
Free Initial Concentration	40.0
Free Current Concentration	40.0
Bound Current Concentration	0.0
Domain Vector	
Interaction Protein Vector	
Interaction Receptor Vector	
Done ✓	Cancel ✕

Fig. 6. Initial values of an internal agent GRB2.

Pathway
Mitogen activated protein kinase system (MAPK)
Adenylate Cyclase System
Cross-phosphatidylation system
Membrane associated protein tyrosine phosphorylation
JAK/STAT activation system
Estrogen/tyrosine Receptor system

Fig. 7. Signalling pathway vector of GRB2.

Domain Name	Complementary Domain	State
SH3A	polyproline	Free
SH2	phosphotyrosine	Free
SH3B	polyproline	Free

Fig. 8. Domain vector of GRB2.

Protein	Interaction	Affinity	IPMA	APMA
Gab2	SH3	1	1.2	3

Fig. 9. Interaction protein vector of GRB2.

represent the protein identity, the domain identity, the affinity and interaction relationship values.

The third section of the model corresponds to the activation of transcription factors by kinase proteins. The transcription factors activate the transcription of genes codifying for proteins involved, in this case, in cell growth. Some of these formed proteins belong to this signalling pathway, and others give rise to transcription factors. This is a reason why a positive or negative feedback can be observed. This third section includes the following components:

- Enzymes: Pyk2, PI3K, PAK, Raf, JAK, PAK, MLK, PKA, PKC, PDK, AKT, MEK,

MEKK, RSK, ERK1, ERK2, JNKK, JNK, SEK, S6K, MKP1

- Transcription factors: c-Fos, c-Jun, JunD, FosB, JunB, IP-1, Elk-1, CREB, c-Ab1, STAT, ATF2

The signalling components shown in Figs. 5 and 6 were created using the interface of a virtual laboratory based on Cellulab, which is briefly discussed in the next section. All the components of the MAPK signalling pathway identified in each of these three sections are shown in Table 3

Fig. 10 shows the semantics assigned to the different blackboard levels and types of agent in Cellulab for this pathway. For example, the agents of the "Adapter protein" type would be GRB2, SHC, and Gab2; and some of the agents of the "G-protein" type would be Rho, Ras, and Rap. The elemental behaviour repertoire of all these agents has been defined using production rules as a first approach. Once an agent is able to interact with its external and internal milieu in an adequate way, other more complex behaviours can be defined through an improvement of the production rule scheme and/or including other information processing models more sophisticated.

Once defined the elements belonging to each of the sections of the MAPK signalling pathway, according to our methodology, each section is tested, thereby assembling the sections to form the complete pathway.

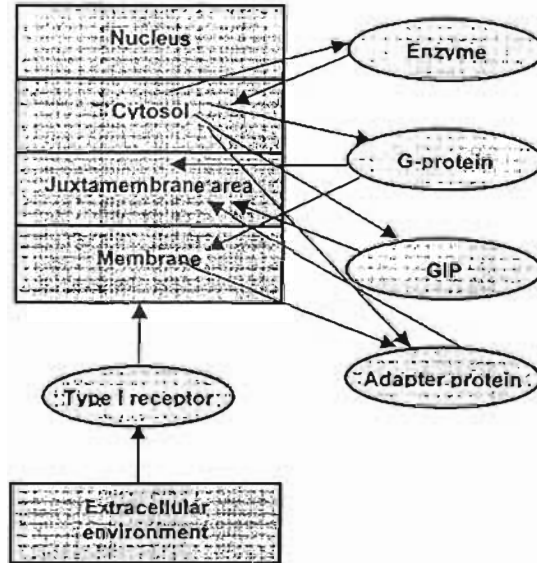


Fig. 10. Model of EGFR signalling pathway in Cellulab.

## 6. Towards a virtual laboratory of intracellular signalling

As discussed, each cell in a multicellular organism receives specific combinations of chemical signals generated by other cells or from their internal milieu. The final effect of the signals received by a cell can be translated in short term responses: secretion of some molecule, metabolism and shape changes. Long term responses: proliferation, differentiation, apoptosis and carcinogen-

Table 3  
Identified components of the MAPK signalling pathway

Section	Component type	Component identity
I	Signalling molecule	EGF
	Receptor	EGFR
II	Adapter protein	Grb2, Shc, Gab2
	Guanine interchanging protein	Sos, p120, p190, p115, cdc42GAP, RapGAP, Db1
	G protein	Rho, Ras, Rap, Cdc42, Rac
	Enzymes	KSR, RasGRP, RasGRF, Src, PLC $\gamma$ , SHP2, SHP1
III	Enzymes	Pyk2, PI3K, PAK, Raf, JAK, MLK, PKA, PKC, PDK, AKT, MEK, MEKK, RSK, ERK1, ERK2, JNKK, JNK, MKP1
	Transcription factors	cFos, cJun, JunD, FosB, JunB, AP1, IP1, Elk1, CREB, cAb1, STAT, ATF2

esis. Once the extracellular signals bind to the receptors, different signal transmission routes are activated, generating a complex information processing network.

In the future, in order to understand how to engineer such a system, we would like to see the effects of perturbations on the systems such as adding elements or taking them out as knock-outs. The expected effects would be directly on the cognitive capacity of the network and ultimately on the decisions taken by the cell in order to differentiate, proliferate or become senescent. Pathologies and natural processes can be followed in the computation of the interactions made by the components in the modelled network. The paradigm presented here is the backbone of the virtual laboratory. With the virtual laboratory we hope that etiologies and expected results of putative therapeutic strategies can be visualised. Figs. 11 and 12 show part of the virtual laboratory interface in its initial stage of development. Through component browser the different signalling components are created, modified and visualised.

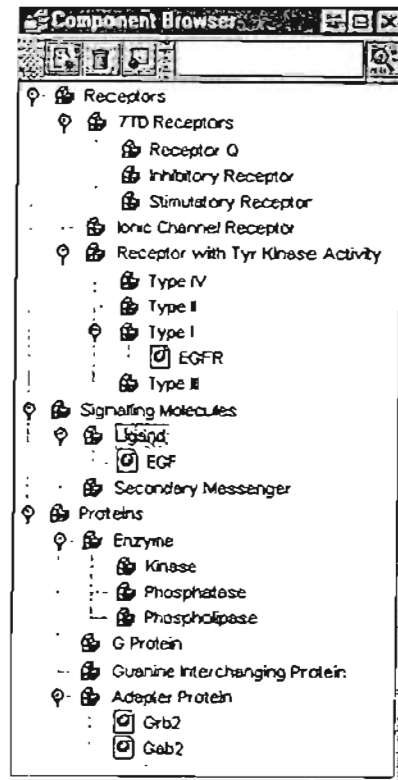


Fig. 12. Signalling component browser.

## 7. Conclusion

We have constructed an intracellular signalling model where the cognitive capabilities are coded using behaviour-based paradigms and the blackboard architecture. In Cellulat, the fusion of the agent-based approach and the blackboard architecture constitutes an initial platform to incorporate other bioinformatic paradigms, in which logics describe the different molecular mechanisms

of the cell. Among the paradigms to be incorporated are neural and Boolean networks. Recruiting these techniques, the complexity of the topology and cognitive capacities of intracellular signalling system can be studied. In our group we are currently following them in the context of pathological states, in particular, senescence and cancer.

We hope that by making the system available to the community, the requirement of resources in a given experimental laboratory will be diminished. As mentioned above, other simulators have been successfully implemented in order to facilitate the orientation of the experimental strategy. This saves time and can avoid the risk of using dangerous or delicate equipments and materials. The expected behaviour of a system that has components interacting with each other by activations and inhibitions need not be studied in an experimental or clinical setting. The interactions can be modelled "in silico" by Boolean expressions and the

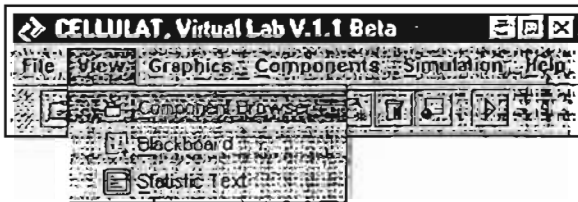


Fig. 11. Initial interface of the intracellular signalling virtual laboratory.

outcomes can be derived theoretically. We will modify components in the system in order to emulate mutations. The resulting state of the system should correspond to a pathological one if the initial conditions were a healthy context or it can be an improvement in case the modification corrects a pathological component.

It should be mentioned that with the construction of the model, we also expect to observe emerging cognitive properties of the networks. It is evident that this will occur as we have found this in other systems having the same structure, architecture, communication and types of components. Among the more interesting properties found are the adaptive action selection, non-linear association, pattern recognition, memory and the handling of fuzzy data. These results will be presented in future publications.

#### Acknowledgements

We would like to thank E. St. James for excellent discussions and proofreading. Part of this work was supported by DGAPA/UNAM.

#### References

- Armas, M., Armas, O., Cárdenas, M., Vasconcelos, J., Lagúnez-Otero, J., 2000. Signal transduction analysis using logic programming paradigms. In: Mastorakis, N. (Ed.), *Mathematics and Computers in Modern Science*. WSES Press, Greece, pp. 165–168.
- Brighman, F.A., Fell, D.A., 2000. Differential feed back regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signalling in PCL12 cells. *FEBS Letters* 482, 169–174.
- Bhalla, U.S., Iyengar, R., 1999. Emergent properties of networks of biological signaling pathways. *Science* 283, 381–387.
- Bray, D., 1990. Intracellular signalling as a parallel distributed process. *Journal of Theoretical Biology* 143, 215–231.
- Bray, D., Lay, S., 1994. Computer simulated evolution of a network of cell signalling molecules. *Biophysical Journal* 66 (4), 972–977.
- Bray, D., 1995. Protein molecules as computational elements in living cells. *Nature* 376, 307–312.
- Brooks, R.A., 1986. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation* RA-2, 14–23.
- Cárdenas-García, M., 2001. Búsqueda de blancos terapéuticos empleando un modelo de señales intracelulares mediadas por Ras. Tesis Doctoral, UNAM, México. (<http://132.248.11.4>).
- Edwards, C.F., 1995. Computational models for cellular information processing systems. unpublished undergraduate thesis, University of Liverpool.
- Fisher, M.J., Paton, R.C., Matsuno, K., 1999. Intracellular signalling proteins as 'smart' agents in parallel distributed processes. *BioSystems* 50, 159–171.
- Fisher, M.J., Malcom, G., Paton, R.C., 2000. Spatio-logical processes in intracellular signalling. *BioSystems* 55, 83–92.
- Fuss, H., 1987. Simulation of biological systems with Petri nets—introduction to modelling of distributed systems. In: Moller, D. (Ed.), *Advances in System Analysis*. Vieweg, Braunschweig, Wiesbaden, pp. 1–12.
- Gershenson, C., González, P.P., Negrete, J., 2000. Action selection properties in a software simulated agent. *Lecture Notes in Artificial Intelligence* 1793, 634–648.
- González, P.P., Negrete, J., 1997. REDSIEG: A cooperative network of expert systems with blackboard architectures. *Expert Systems* 14 (4), 180–189.
- González, P.P., 2000. Redes de Conductas Internas como Nodos-Pizarrón: Selección de acciones Aprendizaje en un Robot Reactivo. Tesis Doctoral, Instituto de Investigaciones Biomédicas/UNAM, México. (<http://132.248.11.4>).
- González, P.P., Negrete, J., Barreiro, A.J., Gershenson, C., 2000a. A model for combination of external and internal stimuli in the action selection of an autonomous agent. *Lecture Notes in Artificial Intelligence* 1793, 621–633.
- González, P.P., Gershenson, C., Cárdenas, M., Lagúnez-Otero, J., 2000b. Modelling intracellular signalling networks using behaviour-based systems and the blackboard architecture. In: Mastorakis, N. (Ed.), *Mathematics and Computers in Modern Science*. WSES Press, Greece, pp. 165–168.
- Holcombe, M., 1994. From VLSI through machine models to cellular metabolism. In: Paton, R. (Ed.), *Computing with Biological Metaphors*. Chapman and Hall, pp. 11–25.
- Karp, P.D., Palcy, S.M., 1994. Representations of metabolic knowledge: pathways. In: *Proceedings of Second International Conference on Intelligent Systems for Molecular Biology*, pp. 203–211.
- Kauffman, S.A., 1991. Antichaos and adaptation. *Scientific American* 265 (2), 64–70.
- Lagúnez-Otero, J., 1998. The cell as an expert system. *Dagstuhl Seminar Report* 215, 60.
- Levy, S., 1992. *Artificial Life: the Quest for a New Creation*. Jonathan Cape, London.
- Maes, P., 1994. Adaptive autonomous agents. *Journal of Artificial Life* 1 (1), 2.
- Marijuan, P.C., 1994. Enzymes, automata, and artificial cells. In: Paton, R. (Ed.), *Computing with Biological Metaphors*. Chapman and Hall, pp. 50–68.
- Negrete, J., González, P.P., 1998. Net of multi-agent expert systems with emergent control. *Expert Systems with Applications* 14 (1), 109–116.

- Nii, H.P., 1989. Blackboard systems. In: Barr, A., Cohen, P.R., Feigenbaum, E.A. (Eds.), *The Handbook of Artificial Intelligence*, vol. IV. Addison-Wesley Publishing Company.
- Paton, R.C., 1993. Some computational models at the cellular level. *BioSystems* 29, 63–75.
- Paton, R.C., Staniford, G. and Kendal, G., 1995. Specifying logical agents in cellular hierarchies. In: Paton, R.C., Holcombe, M., Staniford, G. (Eds.), *Proceedings of JPCAT (Information Processing in Cells and Tissues)*, pp. 302–317.
- Pritchard, L., Dufton, M.J., 2000. Do proteins learn to evolve? The Hopfield Network as a basis for the understanding of protein evolution. *Journal of Theoretical Biology* 202, 77–86.
- Schwab, E.D., Pienta, K.J., 1997. Modeling signal transduction in normal and cancer cells using complex adaptive systems. *Medical Hypotheses* 48, 111–125.
- Takai-Igarashi, T., Kaminuma, T.s., 1998. A pathway finding system for the cell signaling network database. *Journal of Computational Biology* 5 (4), 747–754.
- Wurthner, J.U., Mukhopadhyay, A.K., Peitmann, C.J., 2000. A cellular automaton model of cellular signal transduction. *Computers in Biology and Medicine* 30, 1–21.
- Ying, C., Ferrell, J., 1996. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Science, USA* 93, 10078–10083.