

18



**UNIVERSIDAD NACIONAL
AUTONOMA DE MEXICO**
FACULTAD DE QUIMICA

**TRABAJO MONOGRAFICO DE ACTUALIZACION
BIOLOGIA MOLECULAR *IN SILICO*
MODELOS PARA LA TRANSDUCCION DE
SEÑALES CELULARES**

289616

Que para obtener el título de
QUIMICO FARMACEUTICO BILOGO
presenta

DAVID CAMACHO TRUJILLO



México, D. F.

**EXAMENES PROFESIONALES
FACULTAD DE QUIMICA**

2001



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

**UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO
FACULTAD DE QUIMICA**

**TRABAJO MONOGRÁFICO DE ACTUALIZACION
BIOLOGIA MOLECULAR *IN SILICO*
MODELOS PARA LA TRANSDUCCION DE SEÑALES CELULARES**

**QUE PARA OBTENER EL TITULO DE
QUIMICO FARMACEUTICO BIOLOGO**

PRESENTA

DAVID CAMACHO TRUJILLO

JURADO ASIGNADO:

Presidente	Homero Hernández Montes
Vocal	Rodolfo Pastelín Palacios
Secretario	Jaime Lagúnez Otero
1er. Suplente	Rogelio Rodríguez Sotres
2o. Suplente	Mireya Rodríguez Penagos

**Sitio donde
se desarrollo el tema:**

Instituto de Química UNAM.

Asesor del tema:

Dr. Jaime Lagúnez Otero.

Sustentante:

David Camacho Trujillo.

Agradecimientos

A mis padres y para a aquellos que a través de mi vida me han transmitido conocimiento sin mezquindad, como el Prof. José Luis, La "chapis", Arturo y Jaime.

A las personas que me han apoyado en diferentes sentidos y momentos; Lendy, die Klunkert und Krebs familien, Elvira Serio, Roberto Amador, Jaime.

A todos mis amigos; "Jorgito", Sammy, Villlobos, Laila, Cesar, Silke, François...

Y

A quien corresponda...

Dedicatoria.

A Tere y Arturito, porque aquí esta su tiempo, su cariño...parte de su vida expresada a través de mi.

Al recuerdo de Bettina, la más extraordinaria persona que he conocido y amado.

"als das kind... kind war..."

Indice:

- 1 **INTRODUCCION**
 - 1.1 PLANTEAMIENTO DEL TEMA
 - 1.2 OBJETIVO
 - 1.3 ENFOQUE
- 2 **INFORMACION GENERAL**
 - 2.1 **EL PROBLEMA BIOQUIMICO**
 - 2.1.1 ORIGEN Y EVOLUCION DE LA TRANSDUCCION DE SEÑALES
 - 2.1.2 NATURALEZA DE LA TRANSDUCCION DE SEÑALES
 - 2.1.3 MODELOS GENERALES DE LA TRANSDUCCION DE SEÑALES
 - 2.1.3.1 PROTEINAS G LIGADAS A RECEPTORES
 - 2.1.3.2 ENZIMAS LIGADAS A RECEPTORES
 - 2.1.4 ESTRUCTURAS EN EL SISTEMA
 - 2.1.4.1 LINEALES vs. NO LINEALES
 - 2.1.4.2 DOMINIOS SH₂ SH₃
 - 2.1.4.3 AGREGADOS MULTIPROTEICOS
 - 2.1.5 COMPUTO INDIVIDUAL, LA COMPLEJA ACTIVACION DE RAF-1
 - 2.1.6 COHERENCIA EN EL SISTEMA
 - 2.2 **EL PROBLEMA *IN SILICO***
 - 2.2.1 INTRODUCCION
 - 2.2.2 COMPUTACION NUMERICA
 - 2.2.3 MODELOS Y PARADIGMAS DE LA BIOINFORMATICA
 - 2.2.3.1 INTELIGENCIA ARTIFICIAL
 - 2.2.3.2 SISTEMAS EXPERTOS
 - 2.2.3.3 REDES NEURONALES
 - 2.2.3.4 ALGORITMOS GENETICOS
 - 2.2.3.5 AUTOMATAS CELULARES
 - 2.2.3.6 REDES DE PETRI
 - 2.2.3.7 AGENTES INTELIGENTES
 - 2.2.4 TEORIA DE LA COMPLEJIDAD
- 3 **DISCUSION**
 - 3.1 REDELIMITACION DE LA TRANSDUCCION DE SEÑALES?
 - 3.2 ANALISIS DE LAS METODOLOGIAS.
 - 3.2.1 LAS APROXIMACIONES TRADICIONALES
 - 3.2.2 SOBRE EL ENFOQUE DE LA BIOINFORMATICA
 - 3.2.3 PROPUESTAS DE CONCEPTUALIZACION.
 - 3.2.3.1 CARACTERÍSTICAS DE LOS PROCESOS BIOQUÍMICOS PARA EL DISEÑO DE MODELOS DE TRANSDUCCION DE SEÑALES.
 - 3.2.4 MECANISMOS BIOQUÍMICOS DE CONTROL
 - 3.2.4.1 MECANISMOS BIOQUÍMICOS DE DIVERGENCIA Y CONVERGENCIA
 - 3.2.4.2 MECANISMOS ESPECÍFICOS DE DIVERGENCIA
 - 3.2.4.3 MECANISMOS ESPECÍFICOS DE INTEGRACIÓN
 - 3.2.4.4 MECANISMOS BIOQUÍMICOS DE CONTROL
 - 3.2.5 OBJETOS DE ANÁLISIS
- 4 **CONCLUSIONES**
 - 4.1 PROPUESTA SOBRE LA PLATAFORMA COMPUTACIONAL
 - 4.2 PROYECCION A FUTURO
- 5 **BIBLIOGRAFIA**

1 Introducción

1.1 PLANTEAMIENTO DEL TEMA.

En el umbral del siglo XXI la gran cantidad de información generada por investigaciones diversas representa una oportunidad de integración conceptual en áreas como la inmunología, neurociencias, materiales inteligentes, etc. Pero particularmente esta integración es una necesidad en áreas como la biología molecular y la computación, las cuales tienen un crecimiento muy importante. Esto se hace evidente con los sonados resultados de proyectos como el del "genoma humano" (Human Genome Project) y la explosión en la dependencia de las computadoras para casi todas las áreas del que hacer humano.

El impresionante desarrollo de la biología molecular tiene su origen en avances tecnológicos como la utilización de la técnica PCR (polymerase chain reaction), los nuevos secuenciadores genéticos y tecnologías como los microarreglos para lectura de la expresión genética a gran escala ("microarrays"). En si mismos, estos avances tecnológicos no representan casi ningún avance conceptual para esta área, sin embargo generan una gran cantidad de información en un lapso de tiempo hasta hace poco impensable. De forma tal, para analizar e integrar toda esta información a una estructura conceptual más amplia, ha surgido la necesidad de auxiliarse de herramientas informáticas especializadas, capaces de manejar grandes cantidades de datos en corto tiempo, flexibles, modulares, expansibles y perfeccionables.

Más aún, la modelación computacional y no solo el análisis de estos volúmenes de información de los procesos de la biología molecular, se sitúa en esta coyuntura de la Biología, como un área estratégica. Casos particularmente promisorios son la simulación de la comunicación ínter e intracelular y la regulación de la expresión genética, dado que estos procesos al nivel de un organismo multicelular, pueden verse como el procesamiento de información necesario para

coordinar sus funciones superiores. Análogamente, se prestan al estudio a través de modelos informáticos, las patologías que se desprenden de los procesos bioquímicos mencionados y a mediano plazo, el sugerir posibles terapias.

El presente trabajo se circunscribe al tiempo de colaboración en Instituto de Química de la UNAM con un proyecto teórico llamado GENIA, enfocado al estudio y análisis teórico, de patologías de origen genético como el cáncer y enfermedades propias del envejecimiento. Particularmente este estudio se enfoca al denominado modulo de redes proteicas, cuyo objetivo es la simulación de las "rutas bioquímicas" relacionadas con el proceso conocido como transducción de señales celulares.

Si bien se conocía la importancia del proceso de la transducción de señales intracelulares, en los últimos años a cobrado gran auge su estudio ya que hasta ahora existen las metodologías experimentales que permiten a laboratorios al rededor del mundo arrojar basta información sobre el tema. Gracias a estos trabajos que han contribuido en el esclarecimiento de cada uno de los múltiples elementos que lo componen y sus diferentes características y niveles a los que estos actúan, es que se ha vislumbrado un sistema altamente complejo que en ocasiones arroja información aparentemente contradictoria. Por lo tanto se ha hecho patente la necesidad de integrar la información disponible, para que puedan emerger características generales si es que estas existen, o por lo menos se traiga a la luz nuevas formas de inquirir acerca de aquellos espacios vacíos que sobre el sistema existen.

Para que un organismo superior multicelular, pueda mantener su integridad y comportarse como una entidad viva, es necesario que todas sus células se coordinen para desempeñar todas las funciones que sean requeridas por ese todo, de esta forma se deben coordinar a partir de las primeras divisiones celulares posteriores a la concepción de dicho organismo, para llevar a cavo su proliferación, diferenciación, migración, maduración, especialización etc¹. Todos estos procesos requieren de una comunicación muy extensa entre todas las células, la cual debe llevarse a cabo de una forma coherente y muy precisa, tanto en tiempo como en espacio.

Todo esto es posible gracias al procesamiento individual que realiza cada célula de acuerdo con su tipo celular, su ubicación, sus estadíos tanto de su desarrollo particular como del tejido u órgano al que pertenece así como un gran número de otros factores contextuales. Ese procesamiento interno que realiza cada célula, generalmente, desde que recibe una señal externa vía receptores hasta que ésta llega a su núcleo, donde tendrá lugar la elaboración de la respuesta fenotípica a dicha señal, es a lo que se ha llamado el proceso de **transducción de señales**.

Actualmente, el sistema de la transducción de señales celulares lo intentan abordar diferentes grupos de bioinformáticos alrededor del mundo^{7,15,20,21,22,23,25,31,33,34,42,47,52,54,58}, ya que resulta atractiva la analogía del procesamiento de la información que realiza cada célula y las unidades de procesamiento en las ciencias de la computación, sin que hasta la fecha exista consenso de la metodología capaz de abordar dicho sistema complejo que realizan las células. Por lo que resultaría útil el sugerir lineamientos para el modelaje que se busca en esta área de estudio de bioinformáticos. Y que en un futuro sus desarrollos permitiesen realizar simulaciones confiables de este proceso, tal que se pudiesen hacer ensayos *"in silico"*

Así, este tipo de proyectos se podrán utilizar con un gran número de aplicaciones, como podría ser el reducir los altos costos y largos tiempos necesarios por los laboratorios experimentales para realizar y probar el diseño de nuevos fármacos. Esto último es uno de los objetivos de realizar la **Biología Molecular In Silico**

1.2 OBJETIVOS.

Lograr conceptualizar las bases bioquímicas de que depende el sistema de la Transducción de Señales Celulares. Analizar las prestaciones y limitaciones de las diferentes herramientas con que cuenta la bioinformática actual. Proponer una síntesis con el mínimo de parámetros que sean capaces de capturar el sistema bioquímico, así como la elección de una metodología adecuada para implementar un modelo teórico del sistema de transducción de señales celulares.

1.3 ENFOQUE.

Como ya se mencionó, existen diversos grupos de investigación alrededor del mundo tratando de abordar el sistema de transducciones de señales celulares desde el punto de vista de la bioinformática. Sin embargo dichos grupos están integrados mayoritariamente por investigadores del área de las ciencias de la computación más que por bioquímicos. Esto probablemente se deba a que el sistema de estudio es muy amplio y con alta complejidad, y tradicionalmente la bioquímica y sus modelos teóricos han buscado descomponer estos sistemas en fenómenos lineales, ya que no existían las herramientas teóricas para estudiarlos en forma integral.

En base a los desarrollos y planteamientos en la biología de los años cuarenta, han surgido muchas aportaciones a la ciencia de la computación provenientes del análisis de los sistemas biológicos, (los trabajos de Alan Turing y Von Newman) como el desarrollo de las llamadas redes neuronales, los algoritmos genéticos, los autómatas celulares, las redes de "Petri", etc. Así, por razones históricas son investigadores de las ciencias de la computación los que se han dado a la tarea de tratar abordar e integrar estos sistemas de alta complejidad.

Sin embargo, en el presente trabajo se busca tener un enfoque bioquímico en el análisis y planteamiento de ambos paradigmas, tanto el biológico como el computacional, ya que se consideró la probabilidad, de que se haya venido incurriendo en pérdida de información y complejidad al buscar modelar los sistemas biológicos originales con sobre simplificaciones de los sistemas biológicos originales, como se discutirá mas adelante.



2 Información general.

2.1 .EL PROBLEMA BIOQUIMICO

2.1.1 ORIGEN Y EVOLUCIÓN DE LA TRANSDUCCIÓN DE SEÑALES.

Los primeros organismos unicelulares con la organización equivalente a una bacteria aparecieron sobre la tierra 3.5 billones de años atrás y solo después de 2.5 billones de años de evolución aparecieron los primeros organismos multicelulares¹. Existen intentos por explicar este largo periodo, pero se desconoce la trayectoria que siguió el proceso evolutivo. Sin embargo existen requerimientos generales, como la necesidad de desarrollar mecanismos de comunicación para un grupo de células, indispensable para coordinar funciones y generar un organismo multicelular.

Las primeras células debieron tener intercambio de materia y energía con su entorno (termodinámicamente, sistemas abiertos) y es lógico suponer que este intercambio les significara mantenimiento, crecimiento y posibilidades de proliferación. La materia y energía que regresaban al entorno podrían ser vistas como material de desecho y energía que inevitablemente se disipaba. Sin embargo, en algún punto de la evolución este intercambio *adquirió mayor complejidad*, cuando esos sistemas unicelulares primitivos obtuvieron beneficios por la secreción de moléculas hacia su entorno. Un ejemplo de esta retroalimentación positiva se presenta actualmente en micro-ecosistemas de bacterias (unicelulares), donde algunas secretan moléculas (microcinas) a su entorno para limitar el crecimiento de otras especies y de esta forma aventajar en la competencia por los recursos del medio¹.

Otra evidencia de *comunicación* entre especies unicelulares la encontramos en las levaduras *Saccharomyces cerevisiae*, eucariotes actuales muy similares a los existentes millones de años atrás¹. En

este caso, cada célula lleva una vida independiente, pero cuando alguna alcanza cierto desarrollo, *emite un mensaje* para inducir un apareamiento sexual. Esto es, secreta un péptido denominado factor de apareamiento, que tiene como blanco otra levadura, pero que tenga un carácter opuesto respecto del apareamiento. De tal forma, cuando se da la unión ligando receptor, la célula aceptora detiene su proliferación asexual, *procesa esta información* y se prepara para la conjugación, (fenómeno previo a la reproducción donde se realiza intercambio genético entre ambas células).

Presumiblemente¹, sistemas análogos de transducción de señales evolucionaron al punto de proveerle a una comunidad de células ciertas ventajas, dando paso a los organismos multicelulares. Estas ventajas comunitarias han repercutido en células *fenotípicamente* más simples debido a especialización. Sin embargo los sistemas de transducción de señales se han transmitido íntegros en el *genotipo* de todas las células de los organismos multicelulares y por razones evolutivas, presentan similitudes estructurales y funcionales tanto para una especie, como entre organismos de especies diferentes.

2.1.2 NATURALEZA DE LA TRANSDUCCION DE LAS SEÑALES.

Las células en animales superiores se comunican utilizando cientos^{1,17,32} de tipos de *señales bioquímicas*; Proteínas, hormonas, derivados de ácidos grasos, e incluso gases en solución como el óxido nítrico y monóxido de carbono e incluso iones como Ca^{2+} , Mg^{2+} . Las cuales contienen gran cantidad de información codificada en su estructura tridimensional y su distribución topológica electrostática^{26,35,36}. La mayoría de estas moléculas son secretadas desde la célula emisora vía exocitosis, y pueden actuar como mediadores locales (señales paracrinas, autócrinas y de sinapsis) o actuar a distancia utilizando el sistema circulatorio, en este caso las células emisoras se llaman endocrinas, y las moléculas secretadas hormonas.

Estas señales pueden actuar en muchos millones de combinaciones posibles^{1,44}. Y una célula dada en un organismo multicelular, debe responder a esta gigantesca cantidad de estímulos en forma selectiva, primero de acuerdo con el conjunto de *receptores* que posee (generalmente proteínas membranales) a través de los cuales es capaz de detectar a un sub conjunto de todas las señales incidentes. Y en segundo lugar de acuerdo con la *maquinaria intracelular* que posee y que la *programa* para integrar e interpretar la información que recibe de su entorno, respondiendo a un conjunto de señales diferenciándose, a otro conjunto proliferando y a otro llevando a cabo una serie de funciones específicas, como la secreción de algún metabolito.

En los animales superiores, la mayoría de las células están programadas para depender de un conjunto particular de señales para su sobrevivencia⁷. Esto es, cada grupo de células que conforman un tejido están restringidas a ciertos "ambientes" dentro del organismo superior, donde deben recibir diferentes señales para su supervivencia y cuando alguna(s) células son privadas de ese conjunto de señales, activan un programa suicida llamado muerte celular programada o apoptosis, para evitar mayores daños en el organismo superior.

Tanto a nivel de señales (ligandos) y receptores como al de los elementos que componen la maquinaria interna de procesamiento, (generalmente proteínas cinasas) las posibilidades de interacción están dictadas por la naturaleza fisicoquímica de estos componentes. Debe existir complementariedad estructural entre las superficies de contacto para poderse acoplar, así como complementariedad electrostática. Ambas complementariedades repercuten en la afinidad y especificidad de la interacción^{22,27,35}. Evolutivamente, se han preferenciado aquellas interacciones que tienden a proveer de ventajas al organismo multicelular, interacciones que por sus consecuencias se les llama *específicas*^{5,49}.

Así mismo, algunas moléculas pueden formar estructuras multimoleculares especializadas^{41,45}, que están "diseñadas" para intervenir en el proceso de la transducción de señales, como lo serían las proteínas receptoras transmembranales y sus ligandos. O bien el conjunto de holoenzima, apoenzima, cofactor, grupo prostético, complejo requerido para la activación de una enzima. Y en general todos los complejos multimoleculares que "adquieren" propiedades nuevas, como la *aditividad*^{38,53,54} que representa la creación de alguna nueva superficie de contacto que permite interactuar con un tercer elemento, el cual no es capaz de interaccionar con ninguno de los dos componentes iniciales por separado.

En realidad el sistema de transducción de señales tiene muchísimos grados de libertad. Cabría imaginar las innumerables posibilidades de dinámica intra molecular (cambios conformacionales proteicos, transposiciones electrónicas etc.) e inter molecular (formación de complejos multiproteicos por ejemplo) para todas las moléculas. Sin embargo, la transferencia de información se realiza preferencialmente a través de alteraciones químicas como la fosforilación de residuos de aminoácidos expuestos en la superficie de las proteínas^{22,31}. Estas fosforilaciones incrementan la complementariedad tanto topológica como electrostática de las superficies de interacción entre dos proteínas. Analíticamente, se incrementa la *afinidad* entre estos elementos del sistema de transducción de señales. Lo que en última instancia lleva a la **activación** de las proteínas involucradas en la transmisión y procesamiento de la señal que generalmente son enzimas alostéricas.

Adicionalmente existen restricciones que confieren especificidad al sistema de transducción de señales. Por ejemplo, para que una colisión entre dos enzimas sea *productiva* se debe cumplir en el espacio - tiempo, con cierta orientación y velocidad al momento de la misma.^{40,11} De echo, para explicar las velocidades de reacción, se ha propuesto³⁶ que a través de las fosforilaciones se presenta una inducción para la correcta

orientación entre proteínas interactuantes, con lo cual se incrementa la eficiencia de las colisiones. Por lo que entra en juego un balance probabilístico, función de concentraciones, radios moleculares, inducción de orientación, coeficientes de difusión etc. que finalmente dictan la complejísima cinética del sistema de transducción de señales celulares.

Recapitulando, de la breve visión evolutiva y de lo anterior, se desprenden características esenciales de los sistemas de transducción de señales. Existe una *señal* que conlleva un *mensaje* a través de un *código* fisicoquímico. Esto implica debe existir un *receptor* para la señal en las *células blanco* y un *sistema para interpretar y procesar* tanto el *mensaje codificado* que contiene la señal, como para monitorear e integrar el *estado actual* de la célula para elaborar una *respuesta* adecuada al mensaje recibido.

Así mismo, se puede ver y asociar la naturaleza del sistema de transducción de señales celulares, como una serie de procesos estocásticos (mediados por difusión) de reconocimiento (complementariedad), interacción (formación de complejos) y transformación (fosforilaciones) de las moléculas (señales) involucradas, las cuales contienen y transmiten gran cantidad de información (mensaje) en su estructura tridimensional y topología electrostática (código). Y están acoplados de tal forma, que globalmente se obtiene una gran precisión. Por lo tanto, ahora procede describir los modelos bioquímicos que, en forma general se propone, describen los sistemas celulares para recibir y procesar información bioquímica.

2.1.3 MODELOS GENERALES DE LA TRANSDUCCION DE SEÑALES.

Lejos de pretender presentar toda la información existente sobre los sistemas de transducción de señales celulares, se presentaran brevemente los modelos bioquímicos generales que permitan comprender la

conceptualización anterior y los análisis posteriores. En el caso de señales solubles en medio acuoso, el punto inicial del sistema de la transducción de señales son los receptores a nivel de la membrana celular. En estos casos, se han propuesto^{1,17,32} dos modelos generales del sistema de la transducción de señales; uno de proteínas G ligadas a los receptores y otro de enzimas ligadas a los receptores.

2.1.3.1 PROTEINAS G LIGADAS A RECEPTORES

En el modelo de las proteínas G ligadas a los receptores, una molécula extracelular o ligando, se une a un receptor que se encuentra embebido en la membrana celular. Generalmente estos receptores pertenecen a los llamados serpentininas de siete pasos transmembranales, ya que estructuralmente son una cadena polipeptídica que atraviesa la membrana celular en siete ocasiones, por lo que exhiben asas intracelulares y extracelulares. Es en estas asas extracelulares que se lleva a cabo la unión con el ligando, lo cual provoca una serie de cambios conformacionales que le permiten interaccionar con una proteína intermediaria o transductora llamada proteína G, a través de su región intracelular.

Esta proteína G consiste de tres subunidades; α, β, γ y se encuentra constitutivamente unida al dominio intracelular del receptor. Su nombre se debe a su capacidad de unirse a nucleótidos de guanina, y en su conformación inactiva, guanosin di fosfato (GDP) se encuentra unido a la subunidad α del trímero. Existen diferentes tipos de subunidades, pero todas las $G\alpha$ hidrolizan GTP y lo hacen en un tiempo característico cada una. De tal forma, los cambios conformacionales del receptor inducen cambios en la subunidad α del trímero disminuyendo su afinidad por el GDP ocasionando su liberación. Como en el citoplasma la concentración del GTP es mucho mayor que la del GDP, rápidamente se une GTP al sitio de unión con nucleótidos vacante, de la subunidad $G\alpha$. Esta unión $G\alpha$ -GTP ocasiona la disociación del trímero en

subunidades $G\beta\gamma$ y $G\alpha$ -GTP, y de esta subunidad $G\alpha$ -GTP del receptor. De esta forma la subunidad $G\alpha$ -GTP libre, se asocia a otra enzima, genéricamente llamada efectora, y la activa.

La llamada enzima efectora usualmente está localizada en la membrana y su activación cambia la concentración intracelular de algún soluto de bajo peso molecular genéricamente llamado *segundo mensajero*. Esta actividad persiste hasta que la subunidad $G\alpha$ hidroliza al GTP a GDP, lo cual como ya se mencionó se realiza a un tiempo característico. Así, llegado el momento la subunidad $G\alpha$ -GDP se disocia de la enzima efectora cesando la actividad de la misma y se reasocia con algún dímero $G\beta\gamma$ libre, regenerando un trímero inactivo. Este trímero inactivo es capaz de reasociarse con algún receptor transmembranal tipo serpiente de siete pasos, restableciendo las condiciones iniciales y cerrando un ciclo.

El sistema de la adenilato ciclasa, es una ruta muy común que utiliza este modelo general^{17,32}. En ese caso la adenilato ciclasa es la enzima efectora que es activada por $G\alpha$ -GTP. Como la adenilato ciclasa cataliza la conversión de ATP-Mg a AMPc, este último actúa como un segundo mensajero. Esto es, el AMPc transmite el mensaje del ligando extracelular (el primer mensajero), al interior de la célula.

El AMPc activa a la proteína cinasa A (PKA). En su estado inactivo esta enzima está constituida por dos subunidades reguladoras que mantienen inactivas a dos subunidades catalíticas. Para que las subunidades catalíticas sean liberadas es necesario que se unan dos moléculas de cAMP a cada subunidad reguladora. De tal forma, estas subunidades de la PKA liberadas (ver fig. 1), fosforilan otras cinasas citoplasmáticas o membranales, que cambian sus actividades y se observan respuestas de corto plazo en el comportamiento de la célula. Adicionalmente a las respuestas de corto plazo, la activación de la PKA puede mediar efectos crónicos cuando una fracción del total de subunidades catalíticas activadas entra en el núcleo y fosforilan

factores de transcripción. Esto sucede en la activación de la proteína de unión a CRE (elemento de respuesta para AMPc en ADN) conocida como CREB por sus siglas en inglés, por esta vía, se da la síntesis de nuevas proteínas^{4,32}.

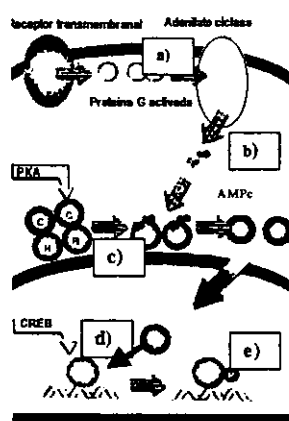


Fig.1) En esta figura se puede apreciar la secuencia; a) desde el receptor activado se libera y disocia una proteína G activada. b) Una enzima efectora, la adenilato ciclasa, es activada por la subunidad G_{α} -GTP, con la cual se incrementa la concentración de AMPc en el citoplasma. c) El AMPc se une a las subunidades reguladoras (R) de la PKA liberando las subunidades catalíticas (C). D) Una fracción de las subunidades C activadas entra al núcleo y activa vía fosforilación a CREB. e) CREB activado se une a su secuencia blanco CRE, en todos los genes que la contengan, activando su transcripción. (modificada "GENES VI" 6ª Edition Oxford University Press)

Otra ruta que utiliza el sistema de las proteínas G acopladas a la señalización, es la del fosfatidil inositol^{1,17}. Análogo a lo descrito para el modelo general, un ligando se une a un receptor tipo serpentina de siete pasos transmembranales y activa a una proteína G. La subunidad liberada G_{α} activa a una enzima efectora Fosfolipasa C beta ($PLC\beta$) a nivel de membrana. Esta enzima cataliza el rompimiento de un fosfolípido de membrana, el fosfatidilinositol 4,5 bifosfato, hacia inositol trifosfato (IP_3) y diacil glicerol; de forma que, esta ruta da como resultado la producción de dos mensajeros intracelulares. El IP_3 por ser soluble actúa liberando calcio de sus contenedores (en retículo endoplásmico), lo que permite la activación de algunas

enzimas y de otros procesos celulares que requieren de calcio. El diacil glicerol, junto con el fosfolípido membranar fosfatidilserina, así como un incremento de calcio intracelular, activan a la enzima proteína cinasa C (PKC), que es dependiente de ambos. Similar a PKA, PKC tiene efectos tanto crónicos como agudos. La PKC fosforila tanto proteínas membranales como citoplasmáticas para llevar a cabo cambios agudos en la actividad celular, y puede también fosforilar factores de transcripción para producir cambios crónicos en las proteínas expresadas en la célula. Aunque ambas PKC y PKA son capaces de fosforilar residuos de serina y de treonina (serin/treonin cinasas), las secuencias de aminoácidos que reconocen en sus substratos difieren. Por lo que PKA y PKC reconocen y fosforilan diferentes series de substratos que sin embargo algunas veces se traslapan ¹⁷.

2.1.3.2 ENZIMAS LIGADAS A RECEPTORES.

El modelo de las enzimas ligadas a los receptores es el segundo modelo^{1,17,32} general de los receptores membranales. En este modelo: el receptor, la región transmembranar, y la enzima efectora son dominios separados de la misma proteína. Existen muchas variaciones de este modelo y en muchos casos la unión del ligando con el receptor causa la dimerización del mismo, como en los receptores para factor de crecimiento epidérmico (EGFR) y de adhesión plaquetaria (PDGFR). En otros, como el receptor de insulina, los receptores se encuentran dimerizados aún en ausencia del ligando. En ambas situaciones, la unión del ligando con el receptor activa a la región catalítica (efectora) del receptor localizada en la cara intracelular de la membrana.

En otras variantes la enzima efectora es una proteína separada de la proteína extracelular de unión al ligando, pero ambas se encuentran fuertemente asociadas en la región intracelular. Los cambios

conformacionales que se producen por la unión del ligando-receptor son llevados a cabo en ausencia de proteínas de unión-transducción, tales como las proteínas G descritas anteriormente. Estos cambios activan a la proteína efectora. Y como el efector se encuentra inmovilizado en la cara interna de la membrana celular, es posible que se desarrollen cascadas muy elaboradas como las de las MAPKs (mitogen activated protein kinases) para llevar mensajes desde la superficie celular hasta el núcleo^{4,65}.

La ruta más estudiada que utiliza este modelo de señalización, de enzima unida al receptor, es la de los receptores relacionados con las proteínas cinasas de tirosina. La unión del ligando con la región extracelular del receptor activa la enzima (o su equivalente en el receptor mismo) efectora cinasa de tirosina, en la cara intracelular de la membrana. Esta región efectora con actividad de cinasa de tirosina, se autofosforila y fosforila otras proteínas sustrato en un residuo de tirosina. Como existe un dominio SH₂ (su nombre proviene de homología con la proteína Src "proto-oncogene tyrosine-protein kinase src") en múltiples proteínas citoplasmáticas como la cinasa de misma tirosina Src, las proteínas de intercambio Grb2 (Growth Factor Receptor-Bound protein 2) y SOS (Son of Sevenless protein homolog), la proteína activadora de la acción cinasa de Ras (Ras-GAP), la cinasa del fosfatidil inositol (KPI₃), fosfolipasa C gamma (PLC γ) etc, y está especializado en la interacción con tirosina fosforilada, por lo que el receptor autofosforilado sirve como un sitio de unión para proteínas que contengan este dominio SH₂. Esta interacción da como resultado la fosforilación de diversas proteínas que contienen el dominio SH₂, a través de la actividad catalítica del receptor, como cinasa de tirosina.

Adicionalmente, algunas proteínas que contienen este dominio también son activadas por el receptor sin ser fosforiladas, esto se logra a través de cambios conformacionales debidos a la interacción con el residuo

tirosina fosforilado del receptor. Un ejemplo es la proteína adaptadora Grb2, donde los cambios conformacionales que se dan en Grb2 por la interacción de sus dominios SH₂ con el receptor, facilitan su interacción con otra proteína de intercambio; SOS. SOS pasa la señal de activación a la oncoproteína monomérica; Ras (ver fig.2). Esta es una proteína débilmente asociada a la membrana que presenta actividad de GTPasa análoga a la de las subunidades α de las proteínas G. La activación de Ras causa que libere GDP y se una a GTP, y se mantenga activa en tanto que no actúe hidrolizando el GTP a GDP.

La proteína de intercambio SOS también sirve como un activador de la función hidrolítica de proteínas G (GAP, GTPase-activating protein), tal que SOS no solo actúa activando a Ras transmitiéndole la señal de activación, sino que también actúa acelerando la inactivación de Ras estimulando su acción de hidrólisis de GTP. Ras activado, a su vez activa a la oncoproteína Raf-1, una serin/treonin cinasa que también es conocida como MAPKKK. Raf-1 fosforila y activa a MEK conocida como MAPKK, la cual es una cinasa con una actividad dual (serin/ treonin) que activa a la cinasa ERK conocida como MAPK a través de la fosforilación tanto en residuos de treonina como de tirosina. Esta última cinasa perteneciente a la familia de las MAPKs, es una serin /treonin cinasa que puede fosforilar proteínas citoplasmáticas como la fosfolipasa A₂ (PLA₂) u otras proteínas como S6 (S6 es principal sustrato de las cinasas en los ribosomas eucariontes), así como factores de transcripción como ATF₂ (activating transcription factor 2) y c-Fos (su nombre se deriva del virus para Osteosarcoma descubierto por Finkel). De tal forma que esta cascada de enzimas transcribe un mensaje a nivel de la unión de un ligando extracelular en un cambio en la actividad biológica intracelular.

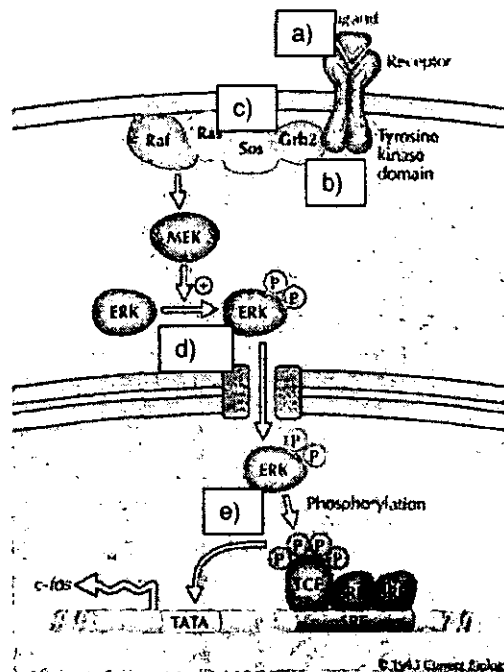


Fig. 2) En esta figura se puede apreciar la activación secuencial empleada por la ruta de la oncoproteína Ras y las cinasas conocidas como MAPKs (cinasas activadas en mitosis). Donde A) un ligando extracelular se une a un receptor B) el receptor se dimeriza, y se autofosforila en su región intracelular, causando la unión con proteínas con el dominio SH2 como la proteína adaptadora Grb2 activándola. C) Grb2 activa a la oncoproteína Ras vía la proteína de intercambio SOS. D) se da una cascada de activaciones de las proteínas cinasas conocidas como MAPKs, Ras activa a Raf, Raf activa a Mek, y Mek activa a Erk. E) La cinasa Erk se introduce al núcleo donde activa vía fosforilación, a factores de transcripción que modifican la expresión de ciertos genes. (modificada de la referencia [12])

La ruta de transducción de señales de JAK-STAT (Janus kinase/ signal transducers and activators of transcription) presenta una variación al modelo de la enzima ligado a receptor, en esta ruta la proteína receptor y la enzima efectora no son la misma proteína. El nombre de cinasas JAK se debe a que en cada molécula se encuentran dos dominios catalíticos. En esta familia de cinasas (JAK 1,2,3, etc) cada una de ellas se encuentra fuertemente asociada a nivel de membrana con algún receptor específico para alguna

citocina. La unión del ligando con el receptor activa la función de cinasa de tirosina de la enzima JAK que se encuentra fuertemente asociada al receptor. Las proteínas STAT se asocian con la JAK autofosforilada a través de sus dominios SH₂ y pasan a ser fosforiladas en residuos de tirosina. Las proteínas STAT fosforiladas en tirosinas se disocian de las JAK, y forman dímeros con otras STAT vía la interacción de sus dominios SH₂ y sus fosfotirosinas. Los dímeros de las STAT se traslocan al núcleo donde actúan como factores de transcripción. Por su parte las proteínas Jak activadas pueden adicionalmente interactuar con la proteína de intercambio Shc, por lo tanto activar la vía de las MAPK's cinasas.

La ruta de la guanilato ciclasa (enzima que cataliza la conversión de GTP a GMPc) también utiliza el modelo general de la enzima ligada al receptor. En este sistema la unión del ligando con la porción extracelular del receptor, inicia la actividad de la enzima efectora, guanilato ciclasa, en la cara citoplasmica de la membrana. La guanilato ciclasa convierte GTP a GMPc y este puede activar enzimas dependientes de él mismo, como las cinasa y fosfodiesterasa dependientes de GMPc. Como sea, el sistema de la guanilato ciclasa introduce otro concepto en la transducción de señales. Adicionalmente a la forma asociada al receptor antes descrita, la guanilato ciclasa existe como una forma citoplasmática soluble, y es capaz de transducir algún mensaje proveniente de moléculas gaseosas disueltas, NO y CO. Ambas moléculas gaseosas son liposolubles, de tal forma que difunden a través de la membrana celular de las células blanco, se unen al grupo hemo de la guanilato ciclasa, y la activan. Y así como la forma activada vía unión al receptor membranal, la guanilato ciclasa soluble convierte GTP en GMPc.

2.1.4 ESTRUCTURAS EN EL SISTEMA

2.1.4.1 LINEALES v.s. NO LINEALES.

Existen muchas rutas de señalización en cualquier célula, y en varios casos comparten una estructura funcional similar. ^{1,17,32} Estas rutas tienen un proceso que se lleva a cabo en múltiples etapas; a) un

receptor transmembranal asociado a un complejo multiprotéico en la cara interna de la membrana que; b) a su vez se conecta a una cascada de protein cinasas que por medio de fosforilaciones sucesivas de los elementos en la parte superior de la cascada hacia elementos en la parte inferior, se va transmitiendo la señal; c) esto finaliza con la activación de un conjunto determinado de factores de transcripción encargados de comenzar con lo será la respuesta fenotípica de dicha célula.

Lo antes descrito es válido para rutas aisladas y sirve para establecer los patrones básicos del proceso de señalización. Sin embargo se sabe que en la célula existen muchos factores que complican la relativa linealidad de estos modelos. Se ha observado que un ligando frecuentemente tiene diferentes efectos en diferentes células blanco y la diferencia en los receptores no es la única explicación para este comportamiento ya que en otros casos, a pesar de contar con el mismo receptor, *una misma molécula tiene efectos disímiles* en diferentes células⁵². Lo cual refleja diferencias en la *maquinaria* interna a la cual los receptores están acoplados. Tal es el caso del neurotransmisor acetilcolina, la cual induce la contracción en las células de músculo esquelético pero decrece la velocidad e intensidad de las contracciones en las células de músculo cardiaco^{1,52}.

También existe el caso donde diferentes moléculas (hormonas) que se unen a sus respectivos receptores, que están acoplados a la *misma maquinaria de procesamiento*, en el mismo tipo celular provocan respuestas muy diferentes, llegando a ser incluso opuestas. Un caso de respuestas disímiles es el que presentan las células PC12, (neuronas de ratón) que cuando son estimuladas por EGF (Epidermal Growth Factor) se induce su *proliferación*, sin embargo cuando se une el ligando NGF (Neurite Growth Factor) a su receptor se induce su *diferenciación*⁵².

Como los mencionados anteriormente, existen varios de estudios experimentales que muestran que la respuesta de una línea celular en ciertas circunstancias es diferente para otra línea celular, o puede ser diferente para la misma célula bajo condiciones distintas.^{4,6,10,12,13,14,18,24,30,48,55} De tal forma, se ha establecido que el procesamiento de la señal (o señales) que realizan las células en su interior para emitir una respuesta, es especializado para cada línea celular, altamente regulado, seguro (a prueba de errores), y adaptable frente a diferentes ambientes celulares. Por lo anterior, resulta muy difícil hacer correlaciones estímulo - respuesta sencillas, y no debe verse al sistema no solo como una simple serie de activaciones concatenadas.

Al respecto de lo que sucede en la región citoplasmática de diversas líneas celulares eucarionticas, las protein cinasas encargadas de la transducción de la señal, generalmente son serin-treonin- cinasas³², presentan convergencias (ver fig. 3). Como en el caso de MEK activada tanto por la protein cinasa Raf como por la cinasa MEKK5 en diferentes organismos como *Schizosaccharomyces pombe* y *cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Xenopus*, etc ¹⁰. (Cabe notar que estas estructuras son evolutivamente conservadas entre diferentes especies.) Formalmente Raf y MEKK5 proveen de funciones análogas en rutas paralelas^{4,10}. Así mismo, en la región citoplasmática se presentan divergencias (ver fig. 3), un caso es el de la cinasa ERK, que una vez activada es capaz de entrar al núcleo e interactuar para activar a los factores de transcripción como c-Myc y Elk-1.¹⁰

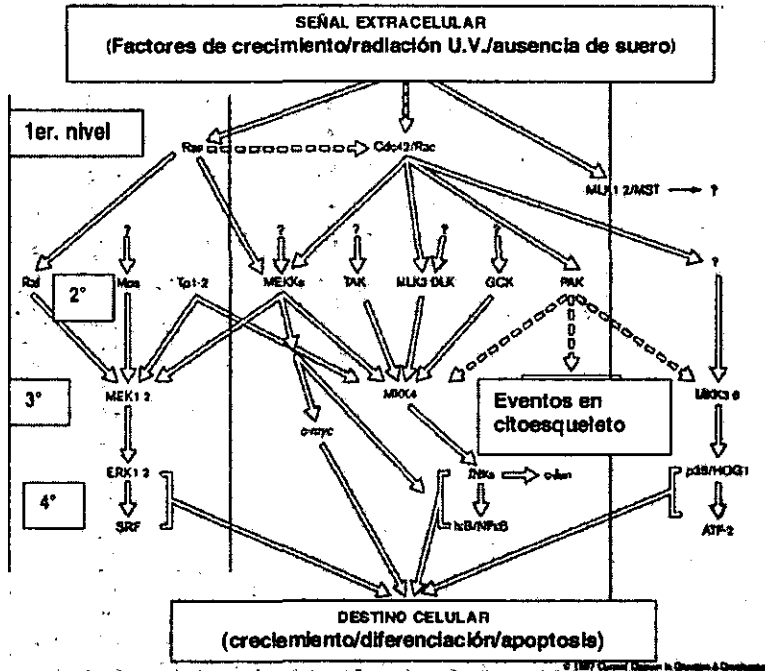


Fig. 3) En esta figura se aprecian verticalmente secuencias de activación de protein cinasas en cuatro pasos, y horizontalmente se observa que existen elementos funcionalmente equivalentes a uno de esos pasos en diferentes rutas. Adicionalmente, se observa como existen entrecruzamientos entre las diferentes rutas, lo que da como resultado una Red Protéica de activación. (modificada de [10])

Surge así un panorama de rutas paralelas de transducción de la señal con una estructura secuencial similar donde existe una corresponden funcional entre elementos de una y elementos de otra, y donde existen múltiples interacciones y entrecruzamientos entre las mismas.

Por lo tanto, más que analizar casos particulares de transducción de señales en diferentes líneas celulares, se requiere encontrar una respuesta lógica para el comportamiento general de cualquier sistema celular, que se entiende como un sistema complejo *No lineal*. Esto es, en un sistema biunívoco a un estímulo corresponde solo una respuesta, pero en este caso al dominio le corresponden más de un elemento en el

codominio. Donde el procesamiento de la información se realiza en paralelo vía diversos mecanismos ^{31,33,34,52}.

Estos entrecruzamientos entre supuestas "rutas diferentes" se deben en gran medida a que muchas proteínas del sistema de la Transducción de Señales presentan en su estructura tridimensional los mismos dominios proteicos, como los mencionados SH2 y SH3, por lo que en principio, cualquier proteína que contenga un dominio SH₂ es capaz de interactuar con cualquier otra con una fosfotirosina expuesta en su superficie. Así mismo se ha propuesto que estos dominios son los "bloques de construcción" de las proteínas involucradas en la transducción de señales celulares y que permiten con pocos elementos realizar combinaciones, (elementos de una ruta con otra) para generar una amplia gama de respuestas celulares.⁴⁵ Un ejemplo representativo de esta conceptualización modular de las proteínas involucradas en la transducción de señales se desprende de la proteína Grb2 la cual tiene en su región central un dominio SH₂ flanqueado por dos dominios SH₃.

2.1.4.2 DOMINIOS SH₂,SH₃

Los dominios SH₂ son módulos proteicos que reconocen secuencias cortas de aminoácidos con alguna fosfotirosina intermedia flanqueada por 3 a 5 residuos de aminoácidos hacia cada extremo, siendo el carboxilo terminal generalmente más largo. Los cambios en esta secuencia son los que confieren diferentes afinidades entre dominios SH₂ y sus ligandos; proteínas citoplasmáticas y región citoplasmática de receptores. Los dominios SH₂ se unen a sus fosfopéptidos con una afinidad relativamente alta (K_d=10-100nM) y con una afinidad 1000 veces menor a fosfopéptidos de secuencias aleatorias, pero básicamente no presentan afinidad por los mismos péptidos si no están fosforilados. Estructuralmente, la región de unión con los fosfopéptidos es una cuenca doble que presenta una división conformada por residuos básicos, en donde se une la fosfotirosina. (ver fig. 4).

Esta cavidad no muestra una secuencia muy estable solo existe un aminoácido invariable, una arginina que forma puentes de hidrógeno con dos oxígenos del grupo fosfato de la fosfotirosina. La segunda cuenca de unión es aún más variable y es la que permite un reconocimiento específico de los aminoácidos de la región carboxilo terminal que siguen a la fosfotirosina.

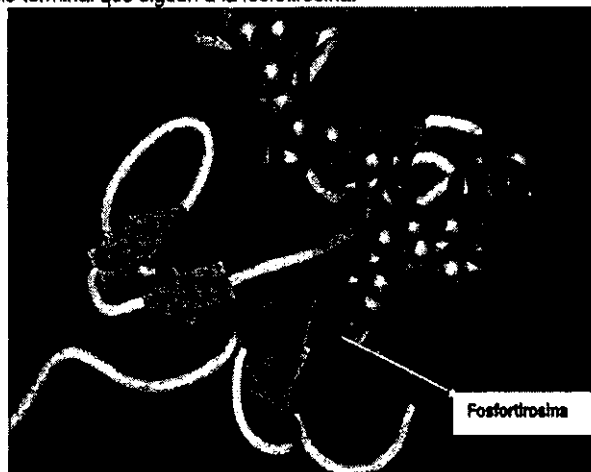


Fig. 4) En esta imagen se aprecia la estructura típica de un dominio SH₂ (representada esquemáticamente), donde se pueden diferenciar dos regiones alfa (tubos rojos) con una región Beta plegada intermedio (flechas en azul), en esta estructura se puede ver claramente que se forman dos "cuenclas" donde encaja la secuencia de reconocimiento (representa con modelos de átomos en tres dimensiones), que son los aminoácidos que flanquean a una fosfotirosina.

Esta organización bipartita del sitio de unión de los dominios SH₂ confiere a la fosforilación del residuo de tirosina la función de interruptor para la formación del complejo protéico, diferenciando entre el estado activo y el estado inactivo, por otro lado dependiendo del "contexto" de esa fosfotirosina dictar que dominio SH₂ y por ende que proteína señalizadora ha de unirse con el mismo. Así, el reconocimiento de estos residuos de fosfotirosina por las proteínas con dominios SH₂ puede afectarlas en diferentes formas, incluyendo la estimulación directa de su actividad enzimática, la relocalización dentro de la célula o bien su respectiva fosforilación en un residuo de tirosina.

Existen otro tipo de dominios denominados SH3 que aunque menos frecuente que el dominio SH₂ también se encuentran en muchas proteínas involucradas en la señalización, algunos ejemplos tomados (solo hasta la mitad de la letra C) de la base de datos PROSITE(<http://expasy.cbr.nrc.ca/cgi-bin/nicesite.pl?PS50002>) son : ABL1 y 2 (proto-oncogene tyrosine-protein kinase abl cinasa ubicua, contienen un dominio un domino SH₂ y uno SH₃), AMPH (amphiphysin, proteína asociada a membrana en neuronas, espermatozoides y células endocrinas contiene un dominio SH₃), ANM2 (protein arginine n-methyltransferase 2 metilasa ubicua contiene un dominio SH₃), BLK (tyrosine-protein kinase blk cinasa involucrada en la transducción de señales en linfocitos T, contienen un dominio un domino SH₂ y uno SH₃), BTK (tyrosine-protein kinase btk cinasa indispensable en la ontogenia de linfocitos B, contiene un domino SH₂ y uno SH₃) CCB2 y CCB4 (dihydropyridine-sensitive l-type, calcium channel beta-2 subunit, subunidades ubicuas de la enzima sensible a di-hidropirimidina, contienen un dominio SH₃), CRKL (crk-like protein cinasa ubicua, contienen un dominio un domino SH₂ y uno SH₃), CRK I y II (proto-oncogene c-crk, CRK I es más transformante que CRK II, contienen dos dominios SH₃ y uno SH₂, ambas se unen a muchas proteínas fosforiladas en aminoácidos de tirosina que se unen a GRB2. El dominio SH₃ cercano a la región carboxilo terminal funciona como un modulador negativo de la actividad transformante, mientras que el dominio cercano a la región amino terminal funciona como un modulador positivo) etc.

Adicionalmente el dominio SH₃ se encuentran en componentes del citoesqueleto, como la α -espectrina y la miosina-1 las cuales están fuertemente relacionadas con el control de la morfología celular. Por lo tanto estos dominios están involucrados en la relocalización intracelular de los elementos que los contienen, por ejemplo en los dominios SH₃ de las proteínas citoplasmáticas Grb2 y PLC- γ 1 dirigen a estas proteínas hacia la membrana citoplasmática interna y hacia las fibras de actina respectivamente.⁴⁵

Las regiones de unión de las proteínas con los dominios SH₃ consisten en péptidos de aproximadamente 10 aminoácidos con mayor contenido de prolínas, que interactúan con constantes de disociación de entre 5 y 10 μM. Estos péptidos que interactúan con los dominios SH₃ adoptan una conformación de hélice tipo II, ricos en prolínas con tres aminoácidos por vuelta. Así mismo, estos péptidos ligando presentan en total tres vueltas, dos interactúan con el dominio SH₃ y la tercera estabiliza la hélice PPII (polyproline II). El núcleo de estos ligandos tiende a presentar siete aminoácidos conteniendo la secuencia consenso X-P-p-X-P, donde X tiende a ser un radical alifático y las dos prolínas (P) conservadas son cruciales para mantener una alta afinidad.

Esos péptidos son pseudo simétricos y tienen la capacidad de interactuar en ambas direcciones, (carboxilo terminal – amino terminal; amino terminal – carboxilo terminal) por ende cada dominio SH₃ tiene distintas preferencias por diferentes péptidos. La especificidad parece provenir de las interacciones entre los residuos no prolínicos del ligando y dos asas variables del dominio SH₃ que están flanqueando a la principal superficie hidrofóbica del mismo dominio. De esta forma, la capacidad de los dominios SH₃ de interactuar en ambas direcciones con sus ligandos amplía considerablemente el rango de ligandos potenciales, y tiene complicaciones biológicas fascinantes; la orientación del ligando va a determinar la organización espacial del complejo resultante, el cual será crítico para el proceso de la transducción de señales⁴⁵.

2.1.4.3 AGREGADOS MUTIPROTÉICOS.

Los modelos multi-pasos parecieran ser complicados e innecesarios para activar a las proteínas G correspondientes. En el modelo de enzimas ligadas a receptores (antes mencionado) de activación para Ras (ver fig. 5) las dos interacciones críticas son las mediadas por los dominios SH₂ de los ligandos con el

receptor y la mediada por el dominio SH₃ con el efector SOS. El efecto neto es el gran incremento local en la concentración de SOS en la vecindad de su substrato, Ras. Esta secuencia de eventos para activar a Ras pareciera ser un método rudimentario, sería mucho más eficiente que el receptor una vez activado, directamente activase a Ras o que SOS tuviese su propio dominio SH₂ y de esta forma se eliminarían intermediarios, en este caso Grb2.

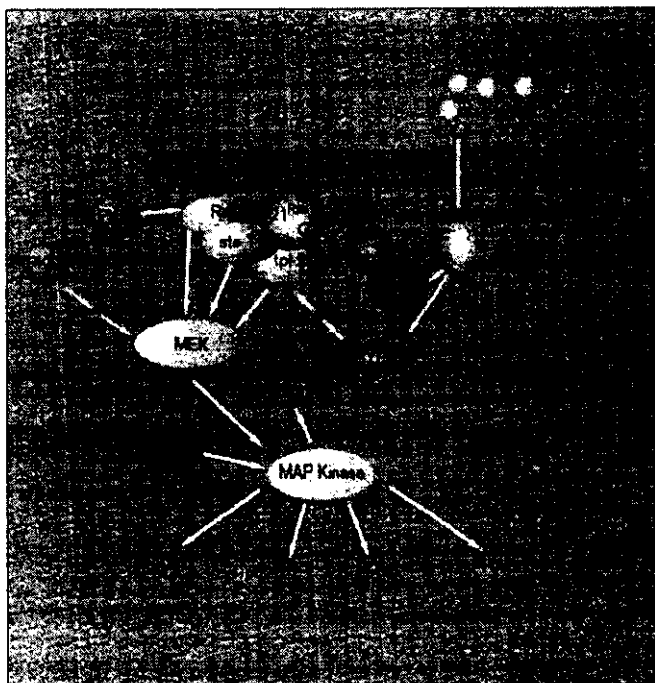


Figura 5). En esta figura se muestra: 1*) Una de las secuencias de eventos posibles para la activación de la ruta de las MAPKs cinasas vía la oncoproteína Ras, a) un receptor activado se encuentra autofosforilado en su región interna en un aminoácido tirosina, b) la proteína Shc interactúa con el receptor vía su dominio SH₂ y esta la fosforila en un aminoácido tirosina c) la proteína adaptadora Grb2 a través de su dominio SH₂ interactúa con la fosfotirosina de la proteína Shc y a través de su dominio SH₃ interactúa con la proteína de intercambio SOS, d) SOS interactúa con la oncoproteína Ras activándola. 2*) Así mismo, aquí se puede apreciar la formación de complejos multiprotéicos a través de combinaciones de proteínas con dominios SH₂ y SH₃. 3*) Adicionalmente se puede apreciar una retroalimentación en la activación de la proteína cinasa MEK por la también cinasa MAPK.

Sin embargo como ya se explicó anteriormente, es un error el considerar al complejo Receptor (EGF o PDGF)-adaptador (Grb2)-proteína de intercambio(SOS)-proteína G(Ras) como una ruta aislada y lineal. El agregado multiprotéico que surge (emerge, desde el punto de vista de teoría de la complejidad, ver más adelante) provee de una ventaja significativa al sistema de la transducción de señales; posibles combinaciones protéicas que generan puntos de ramificación de la señal o integración de estímulos. Así, como primer punto para que se den combinaciones protéicas, está la región interna del receptor que contiene múltiples sitios autofosforilados³². Los sitios de fosfotirosinas de los receptores pueden interactuar y aglomerar a diversas proteínas con el dominio SH₂, pero debido a impedimentos estéricos éstas van a competir por la unión de acuerdo con sus afinidades, concentración local. De tal forma que es posible que se den diferentes combinaciones de agregados multiprotéicos

Desde el punto de vista de la eficiencia a nivel genético, uno de los mecanismos empleados para la transducción de señales es la posibilidad de combinaciones protéicas, lo cual explica que el número de productos de los genes involucrados en la transducción no sea ilimitado, ya que no existe un gen para cada señal en cada célula. El cúmulo de información actual sugiere que muchos tipos de señales extracelulares son transducidas por un número relativamente reducido de enzimas, incluidas las tirosin cinasas , proteínas G y serin treonin cinasas, y que la especificidad de la señal surge entre otros mecanismos, del ensamblado de los complejos multiproteicos que involucran a estas proteínas.

Llegamos al punto en que más que una serie de rutas aisladas, existe (como ya se discutió) una *red de proteínas* (ver fig. 4), que puede presentar *retroalimentaciones* (ver figura 5), *orden temporal* de eventos tanto para el ensamblado de complejos multiprotéicos como para la activación secuencial de cinasas citoplásmicas, *estructura espacial* a nivel de topología celular (localizaciones en membrana vía receptores, o en citoesqueleto vía dominios SH3) así como *combinación* de elementos vía formación de diversos

agregados multiprotéicos. Adicionalmente, hasta ahora poco se sabe de las *cascadas de inactivación* que sin embargo son parte de la dinámica de estas redes que en conjunto dan una red proteica muy compleja.

2.1.5 COMPUTO INDIVIDUAL, LA COMPLEJA ACTIVACIÓN DE Raf-1

Lo hasta ahora descrito son diversos mecanismos involucrados en la transducción celular que se pueden estudiar a un nivel sub-celular están, sin embargo a nivel molecular también se presentan algunos de los mismos. Un ejemplo de las implicaciones de la formación de los agregados multiprotéicos y orden temporal a nivel molecular, es la compleja activación alostérica de la oncoproteína Raf-1^{19,37}. Esta enzima cuenta con tres dominios; CR1 que contiene dos sitios de unión secuencial con Ras (RBD y CRD), CR2 y el último CR3 es su dominio catalítico como cinasa. Raf-1 en su estado inactivo se encuentra en el citoplasma asociada a las proteínas HP (Heat Shock P90 y 50) vía su dominio CR3, y a la proteína adaptadora 14-3-3 a través del residuo fosforilado Ser 259 entre los dominios CR2 y CR1, lo cual estabiliza la conformación inactiva de Raf-1. En realidad 14-3-3 es toda una familia de proteínas que se unen a un gran número de otras proteínas través de la secuencia consenso RXSXS*XP donde S* es un residuo de serina fosforilado, presentan una distribución específica para cada tejido, y son capaces de dimerizarse pudiendo formar oligómeros, por ejemplo de Raf-1 a nivel de membrana.

Una vez que la oncoproteína Ras es activada (ver fig. 6a), esta recluta al complejo Raf-1, HSP90 y 50, 14-3-3 a la membrana por medio de una interacción con el dominio RBD (Ras Binding Domain) de Raf-1 (ver figura 6b), liberándose. La separación de la proteína adaptadora 14-3-3 del dominio CRD permite que este dominio forme una segunda superficie de interacción con Ras, (fig. 6c).

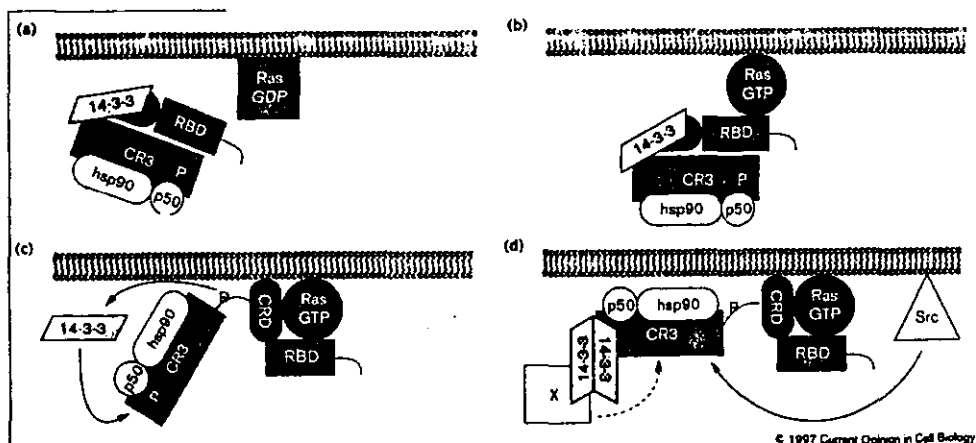


Fig. 6) Esta nueva interacción induce un cambio conformacional en Raf-1 que expone su dominio catalítico CR3, lo cual permite que la proteína libre 14-3-3 se reasocie con Raf-1 a través de la Ser-P 629 en CR3 (fig. 6d). Esta asociación estabiliza la conformación activa de Raf-1. Para completar la activación de la conformación abierta (activa) de Raf-1 se llevan a cabo otras fosforilaciones en las tirosinas 340 y 341 por miembros de la familia de las cinasas Src. (tomada de la referencia [37])

En realidad la activación de Raf-1 es mucho más compleja ya que cuenta con dieciséis sitios de fosforilación, tanto para su activación como por ejemplo en S497, S499 vía PKC, o en Y340 y Y341 vía la familia SRC y Jak-2 etc. como para inactivación en S43 vía PKA. Otros sitios son necesarios para su acoplamiento con las proteínas adaptadoras de la familia 14-3-3 como son S-259 y S-621, dos autofosforilables T 269 y T 268 constitutivos, y otros de los que se tiene poca información^{19,37}.

Lo anterior es solo un ejemplo de algunos detalles necesarios en la activación de una de las proteínas cinasas involucradas en la transducción de señales, que sirve para ilustrar otro aspecto de la complejidad del sistema real. Pero quizás lo más importante sea el analizar las posibles implicaciones que de este tipo de mecanismos se derivan y deben ser en alguna medida contemplados, aún que no los representen, por algún modelo que pretenda captar el comportamiento global del sistema

2.1.6 COHERENCIA EN EL SISTEMA.

Adicionalmente a lo descrito para las combinaciones protéicas a través de la repetición de dominios, estudios *in vitro* (<http://www.kinetekpharm.com/>) comprueban que existen cientos de combinaciones e interacción entre los diversos elementos del sistema de transducción de señales celulares, lo que parcialmente podría explicar que al activar una *misma ruta* bajo *diversas circunstancias* se puedan obtener respuestas diferentes (sensibilidad a las condiciones iniciales, lo cual es una característica de los sistemas no lineales). Por lo tanto para modelar el sistema de transducción de señales celulares es necesario encontrar los mecanismos de control celular que permiten a esta complicada red de proteínas realizar con precisión la transducción de señales, y comprender como un error puede producir una diferenciación errónea y generar un cancer^{10,24}. Las respuestas que se han propuesto hasta ahora solo son parciales, pero es posible hacer algunas acotaciones y mencionar los posibles mecanismos de control celular sobre los sistemas de transducción de señales celulares.

En primer lugar es necesario definir los estados iniciales de que parten los sistemas celulares antes de la transducción de señales. Es evidente que en una línea celular no existen todas las proteínas de todas las rutas de señalización, (de hecho como ya se mencionó, existen protein cinasas tejido-específicas). Así, una línea celular puede describirse en función o quedar definida por la existencia o ausencia de diferentes receptores membranales, la existencia o ausencia de ciertas protein cinasas, etc. Conceptualizando se diría que existen **patrones** de las proteínas que definen una línea celular.

Sin embargo, como estos patrones proteicos no son constantes en el transcurso del ciclo de vida de cualquier célula, es conveniente ir más lejos al momento de definir una línea celular. Así lo que

biológicamente es una línea celular, debe interpretarse como el conjunto posibles *estados* que el *sistema* celular puede ocupar durante su ciclo de vida ⁵⁴(matemáticamente se definiría como el "atractor" celular, ver más adelante). Esto es, el conjunto de los parámetros (patrones protéicos) que definen al sistema celular en un momento dado, representa un estado celular, y el siguiente estado celular quedará definido por el conjunto de esos parámetros para otro instante del ciclo celular. Por lo tanto, el conjunto de estados que el sistema celular puede ocupar durante su ciclo de vida, define una línea celular.

En segundo lugar (aunque funcionalmente es el último) es necesario definir las posibles respuestas o "salidas" del meta-sistema (la célula) frente a un estímulo o conjunto de estímulos, que pueden separarse en: a) Respuestas de corto plazo, como la producción y secreción de algún metabolito (como hormonas, interleucinas, anticuerpos, leche etc.), permanecer aquiescentes, cambiar su forma etc²⁸. B) Respuestas de largo plazo, como la proliferación celular; diferenciación celular, inducción para la apoptosis, y la diferenciación a hacia células neoplásicas. ^{10,24,28} Tal que las células pasan de un estado (o conjunto de estados que conforman un fenotipo) a otro, a través de un proceso no lineal muy sensible a las condiciones iniciales tanto internas como externas.

Una vez acotados los estados iniciales y las posibles respuestas, cabe restringir los grados de libertad del procesamiento de información que realizan las células o transducción de señales. Como ya se mencionó, existen *focos de emisión* de la señal a través de ciertos elementos, como podrían ser los receptores membranales, pero también lo elementos de la familia Ras, que gracias a las *caveolinas* (proteínas de la cara interna de la membrana que forman una estructura conocida como caveolae donde, mediante

interacciones débiles pero específicas, se concentran elementos de la familia de las oncoproteínas Ras) aumentan su concentración (**focalizaciones**) en las inmediaciones de los receptores, amplificando y acelerando la emisión de la señal.

Otra forma de modulación o restricción de las posibles combinaciones proteicas para el procesamiento de la señal, es la dependencia de algunos elementos para con determinadas proteínas **adaptadoras** para poder interactuar con sus proteínas blanco. Un ejemplo de esto es lo descrito en el modelo de activación de Raf-1, respecto de las proteínas Chaperonas (principalmente de la familia de las Heat Shock Protein X) o la familia de proteínas adaptadoras 14-3-3. Aquí es posible incluir a las proteínas que interactúan a través de sus dominios SH₃ con el citoesqueleto u otras proteínas "*de anclaje*" que se unen a él, y que generan **restricciones en su localización** dando como consecuencia restricciones para las posibles combinaciones proteicas antes discutidas así como el efecto de focalización sobre la concentración antes descrito.

Un parámetro muy importante para explicar diferentes respuestas del sistema celular, es el **nivel de activación** de las proteínas cinasas involucradas en la transducción de la señal (la concentración de las formas activas). Se ha propuesto que existen "**umbrales de activación**", a partir de los cuales una proteína (léase población de esas proteínas) activada puede interactuar, en forma efectiva (persistente), con otra para activarla.⁵⁸ En forma similar, se propone que existen umbrales de activación para la interacción de una proteína (a) con otras dos (b) y (c), esto es, la proteína en cuestión (a) primero interactuaría con la proteína por la que tuviese mayor afinidad (b), de saturar a esta (población) proteína y aún quedar cierta cantidad de proteínas (a) activadas, podrían interactuar con la tercera (c) proteína, por la que tendría menor afinidad. Generando así la divergencia de la señal (antes mencionada) que procesa una célula³².

Sin embargo aún restan preguntas más precisas sobre los factores o características fundamentales en que recae la modulación de la transmisión de la señal, como nuevas rutas por "descubrir", rutas de inactivación³⁹, diferenciación entre el posible "ruido" por interacciones inespecíficas entre diferentes elementos y parte del procesamiento de la información a niveles no detectados etc. Por que la estructura y dinámica se ha conservado para diferentes especies a través de la evolución. Por lo tanto el estudio de este complejo sistema mediante modelos teóricos podría contribuir en el esclarecimiento de estas preguntas, de tal forma que enseguida se presentan de algunos de los modelos desarrollados en computadoras en este sentido.

2.EL PROBLEMA IN SILICO

2.2.1 INTRODUCCION

La secuencia; conceptualización, modelado y simulación, para estudiar e intentar pronosticar procesos de la biología molecular mediante programas computacionales, implica integrar dos áreas del conocimiento con un crecimiento muy importante. Existen múltiples trabajos enfocados a lograr el paso de conceptualización ^{2,20,21,23,25,33,34,42,43,49}. Y se ha generado una amplia gama de propuestas de modelos para realizar simulaciones bioquímicas en computadoras ^{7,8,9,11,15,16,22}. Así el éxito en utilización o creación de nuevas herramientas computacionales para modelar un fenómeno particular, depende en gran medida de una clara visión de los objetivos que se buscan. Por lo tanto son necesarios una correcta conceptualización tanto del fenómeno bioquímico particular, como de las prestaciones, características de diseño y limitaciones de las diferentes herramientas computacionales existentes, o al menos las asequibles.

Durante los años en que he venido trabajando y colaborando con el proyecto GENIA del Instituto de Química de la UNAM, me ha sido claro que diversos grupos de bioinformáticos, en diferentes latitudes, fácilmente son presas de simplificaciones de los sistemas biológicos, y también durante estos años me ha sido evidente las coincidencias en razonamiento y conclusiones con algunos grupos. Por lo que considero que este paso de conceptualización y elección de modelo para realizar las simulaciones sigue siendo vigente y ha justificado múltiples publicaciones.

Como el trabajo presente está enfocado al modulo de redes proteicas y en el capítulo anterior se analizaron las características conceptuales de este sistema bioquímico, toca ahora el turno al análisis de los programas computacionales disponibles. Haciendo énfasis en los objetivos de largo plazo que se pretenden en el proyecto global GENIA, para implementar de la forma más adecuada, el modelaje del

sistema bioquímico con el sistema computacional existente. De esta forma se procederá a presentar unas brevísimas descripciones de las diferentes herramientas y propuestas computacionales. En algunos casos, se mencionarán los trabajos relacionados que se han realizado bajo cada enfoque para representar al sistema de transducción de señales o alguna parte del mismo.

2.2.2 COMPUTACION NUMERICA.

Son programas que emplean las aproximaciones numéricas tradicionales, auxiliadas por la automatización y velocidad de cálculo de las computadoras. Formalmente, se dice que son algoritmos (secuencia en el tiempo de una serie de procesos) implementados en la computadora de los modelos matemáticos de un área particular. Cuentan con la ventaja de que estas "aplicaciones" son específicas para el área en cuestión, lo cual reditua en un ahorro considerable de tiempo. Así, existen programas como: Gepasi 3.0, de Pedro Mendes y Chemical Kinetic Simulator de IBM, que han sido desarrollados bajo este enfoque. Son del dominio público y se pueden obtener fácilmente a través de la red mundial Internet.

Estos programas, una vez planteadas las ecuaciones de las reacciones que intervienen en el modelo, así como las características particulares de cada reacción como constantes de afinidad, la concentración de cada uno de los reactantes y el tipo de cinética a que obedece cada una de estas reacciones, crean los sistemas de ecuaciones diferenciales que describen el sistema que uno indicó. Luego proceden a su solución por diferentes aproximaciones de cálculo según, sea el caso, y como resultado arrojan las matrices de los datos correspondientes a las soluciones numéricas o sus respectivas gráficas. Sin embargo, al parecer nadie ha utilizado específicamente estas plataformas para realizar investigación del área en cuestión.

Un ejemplo de esta aproximación, quizás el más destacable hasta la fecha, lo constituye el trabajo de Ying⁵⁸, quienes en 1996 a través de ecuaciones diferenciales, (19 en total para un conjunto de 10 reacciones enzimáticas) crean un modelo para una ruta lineal simplificada de las MAPK cinasas. Su modelo se desarrolla para lo que se conoce en la química analítica, como química en solución homogénea y emplea los modelos matemáticos tradicionales, esto la cinética de las reacciones enzimáticas siguen el modelo de Michalis Menten. Para lograr implementar este modelo presuponen rangos para las concentraciones de algunas enzimas y los valores de las constantes de disociación desconocidos. En su modelo se estudian las consecuencias de lograr una completa activación solo después de una doble fosforilación, lo cual es una buena consideración ya que generalmente las cinasas son bifosforilables en aminoácidos separados por uno o dos aminoácidos. Con este modelo encuentran evidencia un comportamiento *global ultrasensible* como consecuencia de que la transducción se lleve a cabo en múltiple pasos. Esto es, a través de estas múltiples etapas de activación de enzimas cooperativas (alostéricas), el comportamiento global del sistema es de tipo *interruptor*, de todo o nada a pesar de que el estímulo sea gradual, Presentándose un umbral (concentración de "disparo") para el estímulo. Ellos concluyen que este comportamiento tipo interruptor podría representar una ventaja a la célula si el sistema fuese capaz de diferenciar ruido de estímulos reales. Sin embargo, el modelo implementado aún es una simplificación, ya que con una sola fosforilación se alcanza cierta activación.

En general, como principal limitante de estos modelos se encuentra su dependencia de gran cantidad de datos que permitan crear modelos deterministas. Adicionalmente, para el caso particular de la transducción de señales estos modelos resultan insuficientes para cubrir aspectos muy importantes como la heterogeneidad espacial. Esto se debe a que los modelos matemáticos en que se fundamentan estas soluciones se basan principalmente en reacciones en solo una fase y homogénea (química en soluciones

homogéneas), y como ya se aclaró en el capítulo anterior, la formación de focalizaciones por fenómenos de adsorción-reacción, restricciones espaciales y formación de agregados multiprotéicos juegan un papel importante en el comportamiento del sistema real.

Aunque bajo esta aproximación es posible realizar estudios **cualitativos** serios, resulta evidente que falta mucha información para realizar modelos deterministas. Adicionalmente los sistemas de dinámica no lineal requieren una alta especialización en el área para obtener modelos con importantes limitantes, por lo que se ha propuesto la búsqueda de alternativas. Es en este punto que se procede con la presentación y análisis de aproximaciones que han surgido en el área de la bioinformática para el problema bioquímico planteado, para posteriormente dar paso al análisis global de los mismos en la siguiente sección del presente trabajo.

2.2.3 PARADIGMAS y MODELOS EN LA BIOINFORMATICA

2.2.3.1 Inteligencia Artificial

En su forma más general la inteligencia artificial busca desarrollar técnicas para poder resolver problemas que, normalmente, requieren habilidades humanas para solucionarlos: razonamiento, adaptabilidad, intuición, aprendizaje, etc. Entre los problemas típicos están el procesamiento del lenguaje natural, el reconocimiento de patrones visuales, la solución de juegos, el diagnóstico de situaciones y la planeación de actividades. Como este tipo de problemas no pueden ser resueltos de manera algorítmica, esto es, con una definición precisa de cada paso que lleve a la solución, se ha venido proponiendo la inteligencia artificial como una opción. En general las herramientas provenientes de este nuevo campo de estudio, se han desarrollado dentro de la disciplina conocida como bioinformática, algunas de ellas han sido consideradas como promisorias para el modelado y el estudio del sistema de la transducción de señales.

2.2.3.2 SISTEMAS EXPERTOS

Un sistema experto (S.E.) es un programa computacional capaz de resolver problemas reales que requieren de lógica, procesamiento de conocimiento y tomas de decisión. Estos sistemas también pueden clasificarse en categorías, consultar, analizar y diagnosticar. Los sistemas expertos son actualmente muy utilizados en áreas específicas que formalmente requerirían de un experto humano. Utilizan un razonamiento deductivo e inductivo para dar solución a problemas que frecuentemente no están estructurados y que son imposibles de resolver por técnicas convencionales de computación. Dos de sus principales características son: que representan el conocimiento de manera simbólica y los datos se encuentran ubicados de manera explícita dentro de la arquitectura del sistema.

La plataforma computacional del grupo GENIA de Instituto de Química de la UNAM es un Sistema Experto "orientado a objetos", por lo que se trata de una plataforma (Nexpert Object) capaz de manejar tanto una base de conocimiento (diferente a las bases de datos en que el conocimiento ya está ubicado dentro de un marco de referencias cruzadas) en base a dichos objetos, como una base de reglas de inferencia (If - Then). Adicionalmente cuenta con la característica de realizar su análisis en cualquier sentido de las reglas de inferencia, por lo que presenta ventajas adicionales respecto de los Sistemas Expertos tradicionales. Es sobre esta plataforma que el grupo del proyecto GENIA ha venido desarrollando tanto la base de conocimiento de los elementos involucrados en la transducción de señales como la base de reglas de inferencia que interrelaciona a dichos objetos.

* Dentro la programación orientada a objetos (OOP por sus siglas en inglés), se dice que un objeto está definido por tres principios básicos: a) La encapsulación, esto es, los objetos poseen variables para cada instancia (objeto) y funciones ("métodos" en OOPs) no públicos sino particulares. b) La herencia, cada "clase" de objetos (solo clase en OOPs) hereda todas las funciones y métodos de su clase madre (superclase en OOPs). c) El "polimorfismo", donde múltiples instancias (objetos) de una clase comparten el mismo comportamiento (por tener los mismos métodos) pero cada uno tiene su propio estado (valores de sus variables).

2.2.3.3 REDES NEURONALES.

Son sistemas que interconectan gran cantidad de procesadores de datos sencillos, de tal modo que cada dato que entra al sistema es correlacionado con un dato de salida. Su arquitectura y funcionamiento están inspirados en el sistema nervioso de los animales. Cada procesador, análogo a una neurona, recibe gran cantidad de datos desde otros procesadores y para cada entrada ejecuta su propio cómputo de acuerdo con las funciones predeterminadas que le fueron asignadas. Para emitir una respuesta (salida) ejecuta una sumatoria de todos los cómputos individuales realizados (suma ponderada) y si la sumatoria de esos procesos es superior a un valor umbral predeterminado, dicho procesador emite una salida única, que a su vez incide en otros procesadores ^{51,56}.

Los modelos actuales de las redes neuronales son del tipo llamados de multicapas (ver fig. 7), los cuales consisten en una capa de "neuronas" de entrada, una(s) capa(s) de "neuronas" intermedias y una capa de salida .

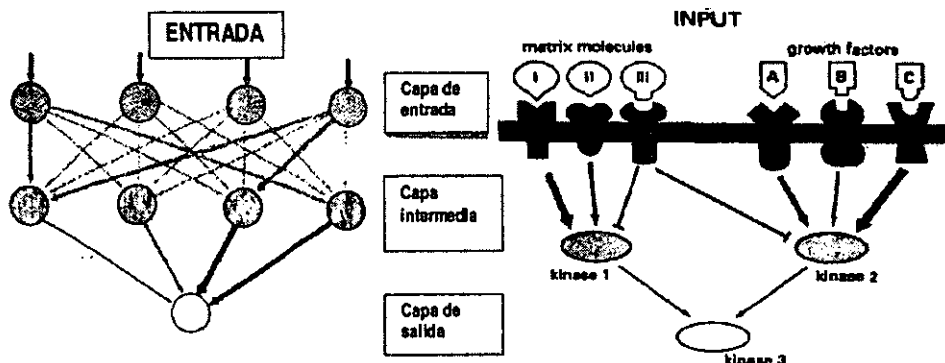


Fig. 7) En el lado izquierdo de esta figura (modificada de [1]) se puede apreciar el modelo general de una red neuronal artificial y en el lado derecho una correlación con su contraparte funcional en el sistema de transducción de señales celulares.

En la entrada de la conexión al procesador ("sinápsis"), existe un valor numérico configurable (valor de ponderación de la sinápsis) que determina la importancia de esa conexión. Gracias a que este valor es

móvil, la red puede ser "entrenada" a partir de unos cuantos ejemplos, para correlacionar datos (de entrada con salida) de forma adecuada. Esto es, la red aprende por propagación inversa, (proceso no iterativo, donde la respuesta del sistema se compara con un estándar y dependiendo de si el valor de la respuesta es superior o inferior al blanco, se ajusta *el procesamiento* del sistema *sin* alterar el valor de la *entrada*) donde las salidas son comparadas con un blanco y los valores móviles son ajustados, de tal forma tendiendo a aproximarse al blanco de salida. Este proceso de aprendizaje continúa hasta que la red logra reproducir el blanco de salida.

Las redes neuronales se aplican donde hay que emitir cierta variedad de respuestas habiendo aprendido (automáticamente) a partir de algunos ejemplos. Sus aplicaciones son muy variadas: compresión de datos (imágenes, audio, etc.), control de calidad de productos, optimización de parámetros, síntesis de voz, reconocimiento de imágenes, clasificación de objetos, eliminación del ruido en señales, etc. A diferencia del sistema experto, el conocimiento de una red neuronal está representado numéricamente (no a través de objetos) y, está distribuido entre todos los elementos que conforman la red, de este modo extirpar una neurona no afecta el rendimiento del sistema (robustez a través de la redundancia*), pero no es posible "ver" que es lo que "sabe" una neurona o un grupo de ellas.

Quizás son las redes neuronales la plataforma computacional por la que más investigadores han apostado en el área de la bioinformática, destacando los trabajos de Bray^{1,7,8,9}, donde capta en cierta medida el comportamiento global del sistema, correlacionando estímulos - respuesta no lineales. Pero sin realizar analogías directas entre los elementos del sistema y los procesos distribuidos en paralelo (PDP's, muy

* En teoría de sistemas, se dice que la **robustez** de un sistema es su capacidad para comportarse normalmente (emitir las respuestas correctas) a pesar de las **perturbaciones** que inciden en él.

Así mismo, se dice que la **redundancia** se presenta en aquellas partes del sistema que tienen una función pero no son indispensables. Esto es, aún que sean extraídas, el sistema puede continuar emitiendo respuestas correctas.

similares a las redes neuronales) en que se basan sus modelos, ni simulando las trayectorias posibles por las que debiese atravesar el sistema real.

También con modelos basados en PDP's encontramos el trabajo de Schamel y Dick⁵², que plantean la conveniencia de utilizar procesos distribuidos en paralelo para explicar la especificidad de los factores de crecimiento. En su trabajo establecen que son muy sutiles las diferencias respecto de los estímulos recibidos por las células PC12 (como ya se describió en el capítulo anterior, son neuronas de ratón) vía EGF (epidermal grow factor, factor de crecimiento epidérmico) o NGF (neurite grow factor, factor de crecimiento neuronal) ya que sus respectivos receptores están acoplados a los mismos elementos corriente abajo. Sin embargo estos estímulos muy relacionados, dan lugar a respuestas muy diferentes, como en todo sistema no lineal, lo cual es posible establecerlo con un modelo de PDP's donde patrones de estímulos muy relacionados dan lugar a diferentes respuestas.

Por último, el trabajo realizado por Iyengar y Upinder en 1999 representa quizás el trabajo más serio y extenso realizado bajo esta aproximación⁴⁷. Donde inicialmente modelan por separado cuatro rutas de señalización (PLC γ -PKC, Ras-Raf-MAPK, Ca⁺⁺/CaM-PKA y Ca⁺⁺/CaM-CaMKII, ver el capítulo anterior para estas rutas) con importantes diferencias respecto de los trabajos (con ecuaciones diferenciales) de Ying y Ferrell⁵⁶, ya que en cada una de sus rutas contemplan circuitos de retroalimentación tanto positiva como negativa, y sus curvas no son solo dosis-respuesta, sino que el estímulo esta compuesto por la relación dosis – periodo (intervalo de tiempo para completar un ciclo), de tal forma que el comportamiento de encendido-apagado tipo interruptor se logra con simulaciones incluyendo el tiempo como un parámetro del sistema. Y posteriormente ellos ensayan la integración de estas cuatro rutas de señalización hasta crear así una red de rutas de señalización (fig. 8).

"población" para confrontar cada arreglo como solución a un problema concreto, estos arreglos son representados como secuencias o "cromosomas" de unos y ceros. Del conjunto inicial se seleccionan los arreglos (soluciones) que tiendan a resolver el problema de la mejor manera para ser los "padres" de la siguiente generación (población) de soluciones. La reproducción de los "padres" se efectúa; intercambiando cromosomas de los padres (entrecruzamiento), permitiendo cierto margen de mutación (cambios aleatorios en la secuencia binaria) y cierto margen de inversiones.

Todo el proceso se repite muchas veces por lo que, en cada generación las soluciones tienden a ser mejores que en la anterior. Tras muchas iteraciones se obtiene una muy buena solución para resolver el problema. Un algoritmo genético es, en esencia, un método numérico de solución de problemas, por lo que es muy útil en optimización de parámetros. Por lo que encontramos estos algoritmos genéticos asistiendo a todo tipo de modelos, incluso aquellos de carácter determinista con base en ecuaciones diferenciales, como el mencionado modelo GEPASI de Pedro Mendes, para la optimización de sus parámetros de simulación ó rastreo de los mismos a fin de que el modelo se ajuste a cierto comportamiento global.

2.2.3.5 AUTÓMATA CELULAR

Esta técnica no esta directamente relacionada con la inteligencia artificial sino con las matemáticas discretas. Un autómata celular es básicamente un programa que simula un ambiente (una retícula) donde habitan ciertos entes, llamados "células"; cada uno tiene una posición única dentro del ambiente y cierto numero de estados (ej. vivo, muerto, hambriento, etc.). El estado de cada ente es afectado por el estado de los entes circunvecinos, en algunos modelos por los cuatro adyacentes y en otros por los ocho circunvecinos, de tal forma la retícula se actualiza en cada instante. El aspecto mas destacado del autómata celular es el patrón visual que se forma en el ambiente tras varias generaciones..

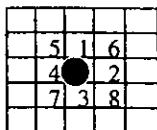


Fig. 9) En esta figura se ilustra una retícula que representa el ambiente donde existen las llamadas células, (aquí se representa como un punto negro.)

Utilizando una variante, en la que cada "célula" se mueve de manera aleatoria y puede colisionar con otras

Wolfram en 1984⁴³ estableció un modelo a partir del cual observó la formación de ciertos patrones en los estados finales:

1. Desaparecen con el transcurso del tiempo.
2. Crecen hasta un tamaño finito y se estabilizan o entran en ciclos.
3. Crecen indefinidamente.
4. Crecen y se contraen irregularmente.

Y estableció que, independientemente de las condiciones iniciales y los conjuntos de "reglas" de interacción, se pueden establecer cuatro "formas limitantes características":

1. Estructura espacial (topología) homogéneo.
2. Secuencias estados simples o estructuras periódicas
3. Comportamiento caótico aperiódico (sin que se presente ningún ciclo).
4. Estructuras locales complicadas (sin un patrón topológico definido), algunas de las cuales se propagan (como una metástasis).

Por lo que propuso que cualquier modelo basado en esta técnica presentaría estas características y que podrían ser útiles para simular sistemas biológicos reales, como la formación de patrones pigmentarios en la piel de los animales, e incluso que el proceso de transducción de señales podría ser abordado por esta técnica. Implementando esta técnica, donde los autómatas celulares representan células de un tejido Edwards en 1995⁴³ simuló la propagación de señales a través de un tejido, con cierto éxito.

2.2.3.6 REDES DE PETRI

Las redes de Petri es una forma sencilla de modelar el flujo de información y control en sistemas concurrentes y asincrónicos (Peter 1977 y Resig 1985)^{44,45}. Donde existen dos tipos de nodos: circulares (o *sitios*) y barras (o *transiciones*) los cuales son conectados en forma biunívoca a través de ciertas reglas

que establecen el lugar específico tanto de los *sitios* como de las *transiciones*, pero no pueden conectarse aleatoriamente (ver fig. 10). La ejecución es controlada por "símbolos", esto es, una transición se puede *iniciar* solo cuando hay algún *símbolo* (señal), en los "sitios de entrada". Después del disparo los símbolos son removidos de los sitios de entrada, y adicionalmente se adhiere un símbolo a cada sitio de salida.

A partir de 1994 Holcombe (Clark, L. A. "Preliminary Review of Computational Models Applicable to Cell Signalling" 1997 <http://www.csc.liv.ac.uk/~laurence/research/prereview.html>) ha venido trabajando con estas redes de Petri para modelar el "control concurrente del procesamiento enzimático". En sus modelos los símbolos representan la disponibilidad enzimática (un análogo de la concentración, y la activación), por lo que continuó desarrollando (junto con Bell a partir de 1995) un modelo alternativo llamado "maquina-X", a través una red de estados finitos de autómatas interconectados, donde los operadores de las reglas de interconexión son funciones de datos llamados "tipo-X". Estos datos tipo-X son el conjunto de datos comprendido por el total de las entradas, las salidas y los utilizados durante el proceso.

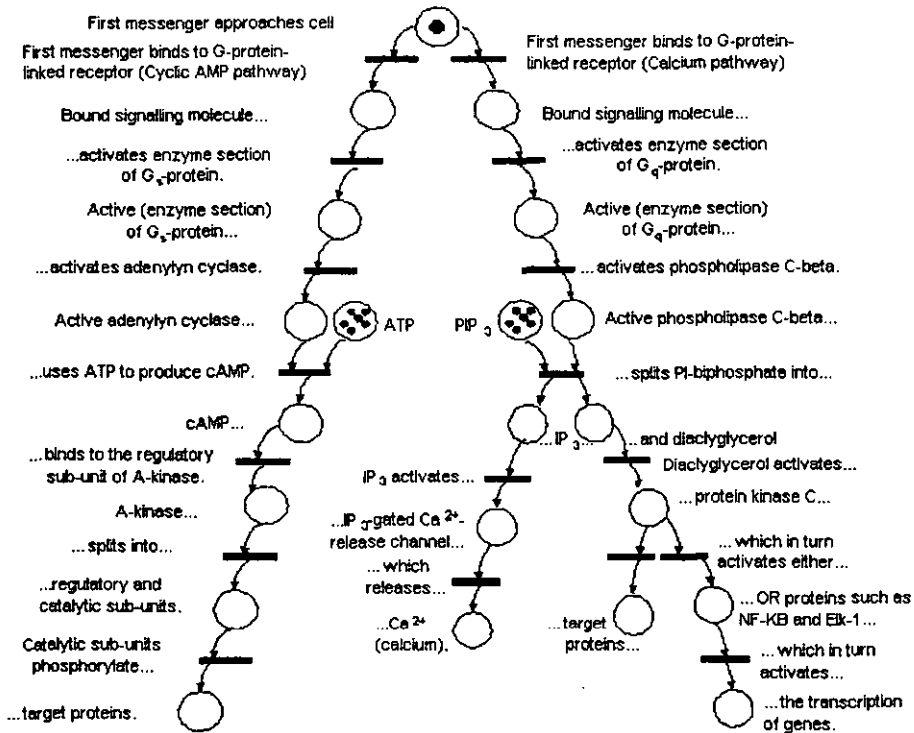


Figura 10.) Esta figura, tomada de Clark. L A "Preliminary Review of Computational Models Applicable to Cell Signalling" 1997 <http://www.csc.liv.ac.uk/~laurence/research/prereview.html> se muestra una Red de Petri que se ha desarrollado para intentar incorporar tanto las rutas de señalización del calcio como del cAMP en un modelo basado en la aproximación de "Primer mensajero". En esta figura el mensajero extracelular no está definido, por lo que el símbolo solo puede seguir una de ambas rutas, lo cual define el tipo de "primer mensajero". Aquí se nota que pueden existir muchos símbolos más en los sitios para el ATP y el PIP₃. Cabe mencionar que toda la funcionalidad de la red es lo que se observa en esta imagen, y se considera una simplificación a un sistema tan complejo con el que se describió en el capítulo anterior.

Un modelo de la máquina-X está siendo desarrollado para simular una célula con dos entradas, un procesador y sitios de salida, con la intención de incrementar detalles (por ejemplo a nivel de interacciones enzima sustrato) en el modelo global con la inserción posterior de otro modelo de menor nivel de abstracción. Por lo tanto se requiere establecer jerarquías entre los submodelos, trabajando con

cada uno a cierto *nivel de abstracción* del sistema real³⁴. Esta aproximación se ha propuesto pueda ser útil para desarrollar un modelo del sistema de transducción de señales.

2.2.3.7 AGENTES INTELIGENTES

Es una técnica emergente* para abordar los problemas que las técnicas clásicas de la inteligencia artificial no habían podido solucionar. Los agentes, más que sistemas o modelos, son una *metodología de diseño e implementación de sistemas complejos* con comportamiento discreto. Esto es, el tiempo no transcurre analógicamente, sino para intervalos discretos. En su concepción original, el agente es considerado como un ente (es como un pequeño programa autónomo) que puede ínter actuar con su medio recibiendo y enviando información, que recibe ciertas encomiendas y aplica sus habilidades para llevarlas a cabo. Por lo tanto, formalmente se dice que es un *proceso* autónomo, ya que el agente o agentes no existen sino hasta el momento de la ejecución de un programa que los contiene, esto es importante por que su comportamiento no es estático como el de un programa, sino que puede modificarse para adaptarse óptimamente a su entorno de acuerdo con su encomienda. Así, bajo estos principios se aplica una o varias técnicas de solución de problemas para construirlo. Estos agentes se han aplicado en el control de sondas espaciales, para análisis de imágenes, para búsqueda de datos en Internet, en robots autónomos, etc.

Un trabajo teórico que aplica esta metodología enfocada al sistema de transducción de señales es el realizado por Schwab y Pienta⁵³, quienes en 1996 realizan una conceptualización del cáncer como un sistema adaptable complejo, y en 1997⁵⁴ aportan las características de los mencionados sistemas, correlacionándolos con el sistema de la transducción de señales, proponiendo las conveniencias de modelar el sistema bioquímico utilizando sistemas distribuidos *multiagentes*, a los cuales definen como los

* En la teoría de la complejidad, una propiedad emergente es aquella que se sitúa en un nivel de funcionalidad superior a los componentes de un sistema. Esto es, es una característica global del sistema que surge de las interacciones de los elementos del sistema. (ver más adelante)

componentes activos de los sistemas complejos y adaptables. Adicionalmente sugieren el uso de otras metodologías de solución como los citados algoritmos genéticos para auxiliar a su modelo a evolucionar y lograr una mejor funcionalidad. Concluyendo que el sistema de la transducción de señales a nivel unicelular pudiese servir como un subsistema para modelar el proceso de carcinogénesis.

Análogos a los trabajos de Bray con las redes neuronales están los trabajos de Paton en 1993⁴³, bajo el enfoque de sistemas multiagentes, en donde parte de la premisa de que la célula puede ser vista como una red de componentes funcionando en paralelo. De esa forma, el procesamiento de la información (durante la transducción de las señales) estaría distribuido entre cierto número de *agentes* que pueden trabajar independientemente pero con posibilidad de comunicarse entre sí. Por lo tanto es posible modelar cada "nodo" en la red de transducción de señales como un agente local individual, con la habilidad de realizar su respectiva tarea y comunicarse con otros *agentes*.

Recientemente Paton y Fisher (1999)⁴² amplían su propuesta al considerar al sistema de transducción de señales como un proceso estocástico pero auxiliado por efectos de localización de ciertos elementos acorde con las últimas evidencias provenientes del área de investigación experimental. De tal suerte que consideran a las proteínas involucradas, como entidades fuertemente dependientes de su contexto, y consideran que los trabajos de Bray con redes neuronales se encontrarían con serias dificultades para incorporar este nuevo factor de heterogeneidad espacial (de hecho esto también es complicado para modelarse vía ecuaciones diferenciales), por lo que insiste en la conveniencia de modelar al sistema de manera global vía agentes computacionales. Aclarando que, si bien esta metodología es más adecuada, no necesariamente cada "agente" debe representar una proteína como tal, sino más bien reproducir su comportamiento. Por todo lo anterior y por coincidir con muchos planteamientos de esta aproximación basada en agentes se consideró conveniente el incluir una breve introducción a las bases teóricas que dan

sustento a esta nueva aproximación, que más que estar relacionada con una técnica de inteligencia artificial en particular, se relaciona con los planteamientos de los sistemas complejos y la vida artificial.

2.2.4 TEORIA DE LA COMPLEJIDAD.

En los últimos años ha surgido una nueva metodología para el estudiar y modelar de sistemas que tengan un elevado número de grados de libertad (no lineales). Esta aproximación es inductiva (opuesta a lo tradicional), y la finalidad del modelo generado por ella, es demostrar que el sistema se comporta en forma lógica³⁸. En esta nueva aproximación, los componentes individuales del sistema son "reducidos" (del concepto reduccionista) hasta sus características esenciales y solo son preservadas aquellas características comunes a todos los componentes. Estas características son parámetros con valores asignados sobre la base de que tan necesarios son para el éxito de un componente individual dentro del contexto de su ambiente. Pero adicionalmente, se incluyen mecanismos de modelado que permiten a los componentes individuales, que puedan experimenten con el perfil de sus características y sean capaces de adaptarse a su entorno. Para lograr esto, se deben establecer reglas simples acerca de las interacciones entre los elementos del sistema, y estas reglas deben ser diseñadas para permitirles a los elementos del modelo, flexibilidad al interactuar con su entorno. Esta flexibilidad es la que hace posibles respuestas no lineales³⁹.

El rigor matemático (formalismo) que se desprende de lo anterior (teoría de la interacción), a demostrado ser adecuado para implementar y aproximarse inductivamente al modelado de sistemas no lineales empleando lo que se conoce como la teoría de la complejidad. En la base de la teoría de la complejidad se encuentra la ciencia del caos y de hecho, a dos de los principios de la teoría del caos se les atribuye los fundamentos para el desarrollo de la teoría de la complejidad. El primero es el uso de los ya mencionados

modelos no lineales. La no-linealidad es responsable de la "sensibilidad" de los sistemas complejos a las condiciones iniciales, también conocida como el "efecto mariposa".

El segundo aspecto que el formalismo de la teoría del caos hereda a la teoría de la complejidad es el "atractor extraño" o simplemente el "atractor"^{38,54}. Esto es, en un modelo lineal, dado un conjunto de condiciones iniciales, tienen su correspondencia (respuesta) en un estado discreto, o un arreglo de números. En cambio, los sistemas no lineales son muy sensibles a las condiciones iniciales, esto es, que una alteración pequeña en los valores de los parámetros iniciales de un proceso dado, puede llevar a resultados muy lejanos respecto del correspondiente al conjunto inicial. De tal forma que los sistemas no lineales no ocupan estados discretos, sino que producen un rango de posibles estados que el sistema puede ocupar. Por ejemplo en un ecosistema estos estados pueden representar diferentes nichos estrechamente relacionados.

Sin embargo la no-linealidad y los "atractores" no son suficientes para modelar sistemas complejos^{38,54}. Los modelos caóticos son solamente "reactivos" (reaccionan a su entorno) mientras que los sistemas complejos son "proactivos" (son capaces de modificarse a sí mismos y a su entorno como respuesta a sus necesidades internas). Esto es, los sistemas complejos poseen mecanismos intrínsecos que les permiten modificarse a sí mismos en respuesta a cambios de su entorno. Adicionalmente a la reacción por el cambio de condiciones los componentes individuales de los sistemas complejos son capaces de alterar sus atributos de tal forma que les permitan sacar provecho de su entorno. (Por ejemplo a través de mutaciones) Esta adaptación se da en la forma de interacciones maleables entre componentes de un

* En teoría del caos se conoce a un atractor como la región espacial hacia donde tiende un sistema. Un análogo químico puede encontrarse en los orbitales moleculares, que son regiones espaciales donde se dice, existe una mayor densidad probabilística de encontrar a un electrón. Esta analogía es útil para imaginar las características de un atractor tridimensional, sin embargo existen atractores en dos dimensiones e incluso puntuales, pero la principal diferencia está en que en un sistema, no importa desde que punto se parte, (solo depende de la cuenca de atracción) el sistema tiende a encontrarse dentro del atractor.

sistema. Estas interacciones pueden llevar a **propiedades globales** del sistema. En otras palabras, él todo es mas que la suma de las partes. Esta característica es conocida como *complejidad emergente* y los sistemas que tienen este comportamiento son conocidos como **SISTEMAS ADAPTATIVOS COMPLEJOS**^{38,54}.

Adicional a la complejidad emergente, todos los sistemas adaptativos complejos están compuestos por siete características básicas. Cuatro de ellas son *propiedades*: agregación, no-linealidad, flujo, y diversidad y las otras tres son *mecanismos*: etiquetamiento, modelos internos y bloques de construcción.

La *Agregación* se refiere a que una vez realizada la agrupación y categorización (en el paso de etiquetamiento) de componentes del sistema a fin de simplificar el modelo, la complejidad emergente regresa al modelo debido a la "agregación" (propiedades que surgen del agregado), de estos componentes simples^{38,54}.

La *no-linealidad* resulta de los diferentes niveles de interacción entre los componentes, de tal forma que una acción pequeña de un componente puede llevar a una respuesta global del sistema. (en teoría de sistemas u señales en particular, surge de la interacción de señales que independientes para dar una respuesta global no lineal).

El *Flujo* se refiere a la transferencia de información y recursos entre los componentes del sistema, lo cual tiene dos efectos, la amplificación que se presenta cuando un recurso es pasado a lo largo de una cadena (esto debe resultar familiar para el caso de la transducción de señales celulares a través de múltiples pasos) y en el cual se produce un cambio en cada elemento de la misma, y un efecto de retroalimentación resultado de la presencia de ciclos a lo largo del sistema.

La *Diversidad* es una propiedad progresiva que emerge cuando los componentes van llenando los "nichos" disponibles del sistema, de manera tal que se pueden presentar *nuevos tipos de interacción* entre los

mismos componentes. Esta propiedad es bien conocida por la genética de poblaciones, y en esa área del conocimiento, la todas las fuentes de diversidad (apareamiento aleatorio, flujo génico) tienen su origen en la mutación.

El *Etiquetamiento* son los distintivos por medio de los cuales los componentes individuales de un sistema, pueden ser reconocidos y reconocer los datos de su entorno. Estos datos son empleados junto con las reglas básicas que rigen a los componentes para construir los *modelos internos*, (pueden verse como modelos más pequeños, que se pueden observar a un nivel de descripción intermedio entre el modelo global y los componentes del mismo o "agentes inteligentes" antes descritos) sobre los cuales los componentes del sistema basan sus acciones e interacciones. Así, las estructuras básicas que un sistema utiliza para construir sus componentes individuales son referidas como los *bloques de construcción*. Estos bloques de construcción son utilizados para construir los *agentes inteligentes basados en reglas* (o simplemente "agentes"). De tal forma que los agentes son los componentes **activos** de los sistemas adaptativos complejos, y están compuestos por un conjunto de reglas "estimulo-respuesta" que define sus acciones. Individualmente cada regla es simple y son las interacciones entre reglas las que llevan a un comportamiento global complejo emergente.

Concluyendo, la teoría de la complejidad puede verse como una metodología alternativa para el modelaje de sistemas complejos y es posible y conveniente implementarla con los llamados agentes inteligentes distribuidos en paralelo, aunque esto último no es excluyente del uso de otras técnicas. Por lo que, por ejemplo, es posible utilizar algoritmos genéticos para lograr la mencionada *adaptación de los agentes*, a través de la optimización de los parámetros de los agentes, y que así estos sean capaces de adaptarse a su entorno. Análogo a esta metodología de algoritmos genéticos, pueden utilizarse otras herramientas tanto de la bioinformática como de la ya mencionada computación numérica, para incrementar la

3 Discusión.

3.1 REDELIMITACION DE LA TRANSDUCCION DE SEÑALES?

Desde los primeros modelos bioquímicos del sistema de la transducción de señales, donde se le concebía como rutas independientes de activaciones en cadena, hasta la nueva visión de una red protéica con dinámica no lineal, el concepto de la transducción de señales ha venido modificandose para tratar de explicar los datos experimentales, que en muchas ocasiones parecían contraponerse. Sin embargo aún esta visión de red protéica podría quedarse corta para explicar las mencionadas respuestas de largo plazo. Es probable que esta visión de la transducción de señales, restringida a eventos en membrana celular y citoplasma sea histórica, pero no necesariamente adecuada. Por lo tanto cabe preguntar si es necesaria una redefinición de la transducción de señales celulares, donde se considere la interacción de los genes entre si mismos y para con la denominada red protéica como parte del procesamiento de la información, con miras a definir un modelo de representación del mismo sistema de transducción de señales celulares.

En todos los trabajos analizados en este estudio, existe consenso respecto del concepto de **estímulo**: aquellas moléculas que interaccionan específicamente con un receptor celular, generalmente situado en la membrana celular, y cuya interacción desencadena una respuesta por parte de la célula. Si bien esto puede presentar múltiples variantes como las mencionadas en la introducción, no es indispensable su inclusión en el modelo que se desarrolle acerca de la transducción de señales ya que los estímulos no son parte del sistema a modelar, sino lo que desencadena el funcionar del modelo. Esto es, los estímulos son los datos que deberán correlacionarse con las respuestas que emita el modelo.

El sistema a modelar, es el **procesamiento de la información**, tanto de la señal extracelular como del estado celular. Se ha sugerido que este procesamiento de la información se lleva a cabo en espacio y tiempo en una red proteica y que esta podría ser dividida en tres regiones espaciales (ver fig. 12) principales: a) Membranal, con particularidades propias de mecanismos de flujo a través de la misma (sistema dinámico en tres dimensiones) y de difusión a lo largo de ella (sistema dinámico en dos dimensiones), b) Citoplasmática. Donde tiene lugar la química en solución, pero con las complicaciones espaciales mencionadas que impiden hablar de una solución homogénea, dadas las posibilidades de compartimentalización y focalización antes mencionadas. Y c) nuclear, contemplándose solo el flujo de las proteínas activadas que selectivamente se introducen para activar genes específicos, sin analizar las dependencias para la activación entre genes.

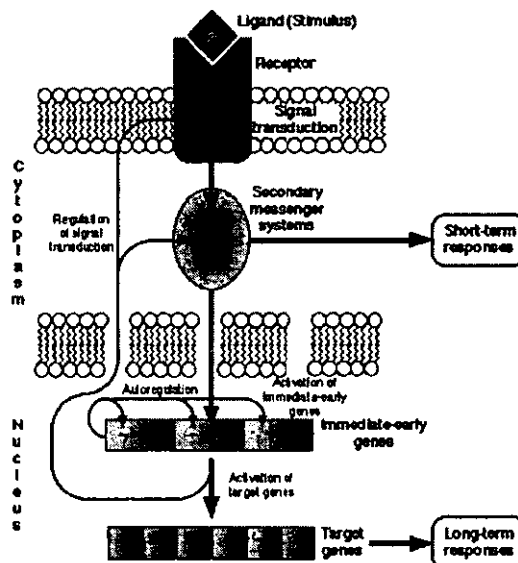


Fig. 12) En esta imagen, se muestran la división en las posibles regiones celulares de procesamiento, membrana citoplasmática y núcleo. Así como lo que se propone como respuestas de corto y largo plazo. Adicionalmente se observa en este esquema

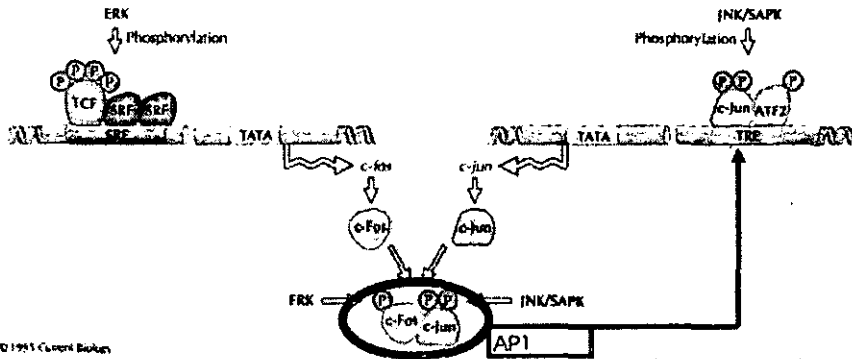
propuesto por Curran y Morgan en 1987 (en la ref. [28]), que a nivel de núcleo se llevan a cabo una serie de retroalimentaciones tanto para el mismo nivel genético como hacia la expresión de proteínas citosólicas. (tomada de la referencia [28])

Sin embargo la red proteica descrita termina, en la mayoría de los casos, en factores de transcripción que son proteínas que intervienen en el núcleo para la activación de determinados genes a través de la interacción con sus regiones reguladoras conocidas como promotores^{46,57}. Estos factores de transcripción son específicos en el reconocimiento de secuencias de ADN conocidas como elementos de respuesta (response elements) dentro de las regiones promotoras, y como estas secuencias las pueden contener un conjunto de genes, un factor de transcripción puede activar a un solo gen o a un conjunto de genes.

Esta activación de los genes por los factores de transcripción a través del reconocimiento de los elementos de respuesta, nos llevaría a considerar la "susceptibilidad" de un gen para ser activado o no, como sus condiciones de empaquetamiento en la cromatina, su grado de metilación y de acetilación etc, lo cual se discutirá más adelante. Pero un punto, quizás más relevante al respecto de los factores de transcripción, es lo desconcertante que resulta que algunos de ellos tienen como secuencias de reconocimiento a *enhancers*, los cuales promueven la transcripción de los genes que existen en su vecindad, en el espacio tridimensional del núcleo, por lo que pueden activar inespecíficamente a genes en *cis* en *trans*, corriente arriba o corriente abajo, lo cual aparentemente es muy inespecífico a menos que se considere el estado de la cromatina (ver más adelante)⁶⁰.

Un ejemplo muy importante, por ser donde convergen las rutas de las MAPK's cinasas, es el factor de transcripción dimérico AP1(ver fig. 13). Este dímero reconoce específicamente a una

secuencia llamada TRE, pero esta secuencia va a favorecer inespecíficamente, a todos aquellos genes que se encuentren corriente arriba o corriente abajo de la misma²⁹.



© 1995 Current Biology

Fig. 13) En esta imagen se aprecia (en el lado superior derecho) que existe un heterodímero AP1 formado por las proteínas c-Jun y ATF2 y cuyo producto de activación va a ser la misma proteína c-Jun. Mientras tanto, se aprecia (en el lado superior izquierdo), como una señal proveniente de las MAPKs cinasas (la cinasa ERK activada) activa la producción del factor de transcripción c-Fos. Por último, se observa (al centro) como el factor de transcripción dímérico AP1 también puede estar compuesto por las oncoproteínas c-Jun u c-Fos, y que este nuevo dímero va a realizar una retroalimentación positiva en la activación, y para la producción de más oncoproteína c-Jun. (modificada de [12])

Así, en el caso del dímero AP1, se sabe que en su forma basal existe como un heterodímero formado por el factor de transcripción ATF2 y la oncoproteína c-Jun (ver fig. 13 parte superior derecha). Pero toda vez que como respuesta a la actividad de las MAPKs cinasas se activan los factores de transcripción (y también oncoproteínas) c-Jun/c-Fos y se forma el heterodímero AP1, existe una competencia entre los dos tipos de AP1 por su secuencia de reconocimiento TRE²⁹. Adicionalmente se sabe que uno de los genes blanco de este *enhancer* (TRE) es el que codifica para el factor de transcripción c-Jun, por lo que se presenta una retroalimentación entre la última etapa de la red proteica mencionada y lo que comenzaríamos a vislumbrar como una red genética.

Lo anterior indicaría que existe un punto de convergencia entre las redes genética y proteica, y que es allí donde tiene lugar una perturbación del estado basal (léase, línea celular) de la expresión de genes. Esta perturbación bioquímicamente se realiza en pasos, primero a través del desplazamiento del heterodímero AP1 ATF2/c-Jun por el heterodímero AP1 c-Jun/c-Fos en la secuencia TRE. Acto seguido y vía retroalimentación positiva, se logra una amplificación a nivel de red genética de la señal, incrementando los niveles nucleares de los productos de genes inmediatos como c-Jun y por tanto del dímero AP1. Es a partir de esta amplificación que podrá llevarse a cabo la diversificación de la señal para la activación de otros genes que también son susceptibles de ser activados por la secuencia del *enhancer* TRE.

Por lo descrito para el caso del ubicuo factor de transcripción AP1, parece que la delimitación de la transducción hasta los factores de transcripción sin síntesis *de novo* es meramente arbitraria y no necesariamente conveniente para representar el comportamiento global del sistema. Por lo tanto parte del procesamiento de la información se encuentra delegada en lo que sería la llamada red genética. Análoga a la red proteica antes analizada, esta red genética deberá contar con un estado basal en función de la línea celular en cuestión y de su estado dentro del ciclo celular, como los puntos de anclaje con la matriz nuclear, MAR (matrix adhesion region), sus conformación como heterocromatina, proporción de acetilación y metilación etc. Todo lo cual, en conjunto se viene proponiendo como el estado del *nucleotipo*.

3.2 ANALISIS DE LAS METODOLOGIAS DE SOLUCION.

Existen dos áreas que intentan abordar el complejo comportamiento del sistema de transducción de señales celulares; A) La bioquímica tradicional, entendida como cinética enzimática. B) la

nueva ciencia de la bioinformática, integrada por múltiples metodologías tomadas de la inteligencia y vida artificial. La diferencia básica entre los primeros y los segundos radica en el nivel de información al que se aborda el sistema original y los alcances de los mismos. Mientras que los primeros son de carácter determinista, los segundos son de naturaleza fenomenológica.

En ambos casos el desarrollo de los modelos teóricos ha tenido que esperar, y no solo porque la estructura matemática para abordarlos es relativamente nueva (sistemas de ecuaciones diferenciales en el caso de los sistemas tradicionales y los paradigmas de la bioinformática en el segundo), sino porque no existían los datos e información suficientes para elaborar algún modelo ni tampoco existían los medios para corroborar los posibles modelos. Así es hasta ahora que el cúmulo de información arrojada por las nuevas tecnologías aplicadas en laboratorios al rededor del mundo, que ambos enfoques intentan integrar este tipo de sistemas no lineales, ahora asistidos por el poder y velocidad de cálculo de las computadoras, que se han constituido en el laboratorio ideal para desarrollar la biología molecular *in silico*.

3.2.1 LAS APROXIMACIONES TRADICIONALES

Bajo la aproximación de la bioquímica tradicional, el problema de la transducción de señales quizás podría reducirse a una cinética enzimática de extrema complejidad. Sin embargo, se tendrían que integrar diversas áreas más allá de la cinética en soluciones homogéneas. Por ejemplo, para el caso de la activación de la oncoproteína hidrolizadora de GTP Ras, el efecto focalizador en membrana por las estructuras conocidas como *caveolae*, requiere de un tratamiento especial. Adicionalmente, la difusión a lo largo de la membrana interna citoplasmática para la activación de SOS y el reclutamiento a este complejo de Raf-1, representan una cinética

de reacciones parecidas a las que se llevan a cabo en sistemas de reacción-difusión en dos dimensiones.

Por otro lado, en la región citosólica podría pensarse que si aplica la química en solución, pero como ya se describió se presentan efectos de focalización debido a la interacción de algunos elementos con el citoesqueleto. Así mismo, fuera del rango de la química en soluciones homogéneas, existe el flujo de ciertos elementos a través de membranas, como el paso de la versión activa de la proteína cinasa ERK hacia el núcleo. Por todo lo anterior, se tendrían que integrar sistemas numéricos especializados para cada área, acoplarlos y finalmente optimizar sus parámetros para conseguir una buena correlación para el modelo del sistema original.

Lo anterior además de requerir expertos en diferentes áreas, presenta como principal limitante el requerir de todos los parámetros que intervienen en los diferentes fenómenos fisicoquímicos descritos y que son indispensables para establecer un modelo determinista, lo cual actualmente no es posible. Esto es, para establecer un modelo cinético es necesario conocer todas las especies químicas que intervienen en el sistema tanto al nivel de reactantes como de productos y para cada uno de los pasos. Lo anterior representa un serio problema para la transducción de señales, ya que existen muchas de interacciones potenciales, tanto específicas como inespecíficas no comprobadas, pero pueden dar lugar a la formación de complejos multiprotéicos hasta ahora solo intuidos.

Por lo tanto el supuesto "ruido estocástico" que se genera a través de las interacciones hasta ahora consideradas como inespecíficas, bien pudiese contener un significado desconocido y ser un objeto de estudio en sí mismo. De omitir estas interacciones inespecíficas se pierden dos

características importantes del sistema, la primera referente a la química en general, que es "constructiva" (genera nuevas estructuras a partir de reactantes), y la segunda al sistema de transducción de señales, que es de naturaleza combinatoria debido a la existencia de los mencionados dominios ubicuos, que generan estructuras modulares tanto a nivel de una unidad proteica como de complejos multiproteicos. Por lo tanto existiría un importante sesgo de información si se omiten estas características desde el origen del modelo mismo.

Todo lo anterior sugiere que el modelo que pretenda integrar estos diferentes niveles de interacción del sistema de la transducción de señales, deberá ser un modelo capaz de auto organizarse, evolucionar y adaptarse para lograr una buena correlación con el sistema real. Esto es, en la naturaleza para introducir control (auto organización) a un sistema estocástico existen dos mecanismos conocidos; a) la retroalimentación que asegura llevar los parámetros del proceso dentro de un rango determinado y b) la robustez que adquiere el sistema frente a variaciones externas a través de la redundancia de diferentes elementos que presentan traslape de funciones. Sin embargo para los otros puntos no existen áreas tradicionales en bioquímica para ellos, por lo que se consideran importantes las aproximaciones de la inteligencia y vida artificial para resolver el planteamiento de un modelo para la transducción de señales celulares.

3.2.2 SOBRE EL ENFOQUE DE LA BIOINFORMATICA

Una impresión inicial es que la bioinformática se podría ver fuertemente robustecida incorporando más fundamentos teóricos a su estructura, provenientes de la teoría de la información. Aportaciones tales, como los límites, alcances y correlaciones con los fenómenos biológicos a que deben su origen técnicas como los llamados algoritmos genéticos, redes neuronales, redes de Petri, autómatas celulares etc.

Por otro lado, un planteamiento inicial que suele hacerse (y quizás primer error conceptual) en los modelos de la bioinformática se basa simplemente en que "el proceso de la Transducción de Señales que realiza una célula rememora claramente una unidad de procesamiento." Cuando se pierde la realidad histórica de que estas unidades de procesamiento surgieron de *sobre simplificaciones* de alguna característica funcional, que se ha observado tiene éxito en los sistemas biológicos.

En cualquier caso, un paso inicial importante es el planteamiento y conceptualización del sistema original para realizar algún modelo. De tal forma que un paradigma inicial es definir cual es el mínimo de procesos bioquímicos involucrados en la transducción de señales capaces de describir ese manejo de la información por el sistema celular. Un segundo paso es extraer las características básicas de las que dependen cada uno de esos procesos. De tal forma que cada uno se pueda caracterizar de forma dinámica e inter conectada con los demás. Con la mayor simplicidad posible pero captando el comportamiento global del sistema.

Como el presente trabajo está enfocado con la aproximación de la bioinformática se consideró importante realizar como parte de las conclusiones del mismo, una propuesta que pudiese servir de base para algún trabajo relacionado con el modelaje del sistema de transducción de señales. Por lo tanto esta propuesta está realizada pensando en el tipo de herramientas con que cuenta la bioinformática. Así las sub divisiones y categorizaciones que pudiesen sonar arbitrarias desde el punto de vista bioquímico, encuentran su razón de ser en la operatividad misma de estas herramientas.

3.2.3 PROPUESTAS DE CONCEPTUALIZACION.

Resulta importante relacionar esta sección con lo descrito en el capítulo de la Información General para la naturaleza del sistema, donde se resaltó que se trata de una serie de procesos estocásticos acoplados en forma tal que se obtiene una alta precisión en las respuestas. Y es debido a esa naturaleza estocástica que el sistema puede obtener variabilidad de respuestas (no linealidad), adaptarse, absorber cambios externos sin alterar su funcionamiento (robustez), etc. Pero como ya se mencionó, en el sistema de la transducción de señales celulares deben existir mecanismos de regulación y control para llevar a cabo su finalidad de emitir respuestas precisas.

3.2.3.1 CARACTERÍSTICAS DE LOS PROCESOS BIOQUÍMICOS PARA EL DISEÑO DE MODELOS DE TRASDUCCION DE SEÑALES.

1) **Patrones iniciales** en la existencia de diferentes elementos involucrados en la Transducción de la señal. Se propone que existen tales patrones ya que se observa que una gran cantidad de elementos que intervienen en la transducción de la señal resultarían redundantes de no existir estos patrones. Por ejemplo, para el modelo de las proteínas G, existen por lo menos 16 subunidades $G\alpha$, 5 $G\beta$ y 11 $G\gamma$ y existen restricciones para las posibles combinaciones entre las subunidades $\alpha\beta\gamma$ por lo que en realidad algunos dímeros $G\beta\gamma$ no existen en condiciones fisiológicas y no todas las subunidades $G\alpha$ pueden interaccionar con todos los dímeros $G\beta\gamma$. Así lo que sucede es que este sistema de las proteínas G provee de flexibilidad, para que diferentes receptores puedan activar diferentes **patrones** de complejos $\alpha\beta\gamma$ y a su vez, las diferentes subunidades $G\alpha$ así como $G\beta\gamma$ pueden activar diferentes enzimas efectoras.

2) **Patrones de combinación**, de las posibles señales de entrada. Queda claro que el reconocimiento de las señales externas está en función del tipo de receptores con que se cuenta, sin embargo, para un instante dado existe una combinación de señales extra celulares que inciden sobre una célula. Por lo tanto, es muy probable que existan o sean condición para

provocar una respuesta particular, un conjunto de señales. De las cuales podrían haber elementos con un nivel de concentración que se podría considerar "basal" (un ejemplo sería iones de Ca^{2+}), pero que se requiera de un incremento en los mismos para inducir una respuesta particular.

3) **Transmisión secuencial** de la Transducción de Señales. Esta estructura lineal, es útil para representar con un número particular de pasos rutas diferentes como una aproximación inicial de diseño. Esto a demostrado en sí mismo generar nueva información, como en el caso analizado y ya discutido del trabajo de Iyengar y Bhalla. Donde la ultrasensitividad surge de esta estructura secuencial y no solamente del hecho de que cada elemento de estas cascadas son enzimas alostéricas.

4) **Estructura modular**, esta propiedad debe ser tomada en cuenta por cualquier modelo ya que la formación de estructuras por combinación de módulos orgánicos es una característica ubicua y muy importante. Como en los casos para las interacciones entre secuencias polipeptídicas cortas (subdominios) y los dominios SH₂ de una cinasa con otra. Así como las interacciones de los dominios SH₃ PDZ, WW, y PH, que dan lugar a la formación de los complejos proteicos como AP1, (compuesto por homodímeros Jun-Jun. O heterodímeros Jun-ATFz (basal); o Jun-Fos). O los complejos multiprotéicos a nivel de membrana Ras-Raf-1-HSP90-HSPX-14_3_3 + activadora de Raf-1: SRC ó PKC o PKA .

3.2.4 MECANISMOS BIOQUÍMICOS DE CONTROL

3.2.4.1 MECANISMOS BIOQUÍMICOS DE DIVERGENCIA Y CONVERGENCIA

Combinación entre diferentes elementos que permiten la convergencia y divergencia debido a la homología entre estructuras químicas compartidas por los diferentes elementos (dominios,

regiones, secuencias) de tal forma que tienen lugar interacciones específicas (con algún significado) entre diferentes elementos de las llamadas diferentes rutas. Esta característica combinatoria está en función de la afinidad.

Afinidad la cual va a determinar las posibles interacciones entre los elementos involucrados en la Transducción de la Señal. Tanto al nivel de combinaciones entre los mismos, como la secuencia en que se lleven a cabo los eventos, ya que esta afinidad es a su vez función del estado del objeto en cuestión, esto es, de su alosterismo en el caso de las proteínas cinasas alostéricas, o el estado de los nuevos objetos creados.

3.2.4.2 MECANISMOS ESPECÍFICOS DE DIVERGENCIA.

Vía división de elementos involucrados en la Transducción de la Señal, como la que sucede en los complejos multiprotéicos tal es el caso de la liberación del factor de intercambio SOS, o la división de las proteínas G en sus subunidades $G\alpha$ y $G\beta\gamma$ las cuales pueden tener cada una función específica. O bien el caso de zimógenos donde destacan PKA y MEKK

Amplificación de la señal. vía incremento en las concentraciones de las versiones activas de los elementos de las cascadas de activación, por regla general en cada paso parece incrementarse en un orden de magnitud la intensidad de la señal respecto de la concentración de las versiones activas de las proteínas cinasas.

Umrales de concentración de las versiones activas o activadas de los diferentes elementos involucrados en la Transducción de la Señal. Estos umbrales de concentración tienen una

dinámica intrínseca y dan lugar a bifurcaciones entre rutas paralelas. Estos umbrales de concentración suelen ser función del periodo en que se mantiene activado el miembro anterior de una cascada de activación.

3.2.4.3 MECANISMOS ESPECÍFICOS DE INTEGRACIÓN

Vía adición, partiendo de la premisa de que los elementos involucrados en la Transducción de la Señal están compuestos por subestructuras modulares como los dominios proteicos. La adición de complejos multiprotéicos podría entenderse como un mecanismo de integración de señales de origen diverso que inciden en un momento dado sobre una célula.

Computo individual de la señal por cada uno de los elementos involucrados en la transducción de señales, vía diferentes estados de fosforilación. Con esto se hace referencia a la modulación de la activación de cada elemento a escala proteica, alosterismo expresado en los diferentes estados en que puede existir cada una de las proteínas. Un caso bien documentado es la compleja activación de la oncoproteína Raf-1. Finalmente esta característica incide sobre la afinidad que presentan unos elementos respecto de otros

3.2.4.4 MECANISMOS BIOQUÍMICOS DE CONTROL.

Retroalimentación positiva y negativa. En los sistemas bioquímicos suele presentarse la retroalimentación para asegurarse que el proceso tienda a llegar (vía amplificación) o mantenerse dentro de un rango específico de concentraciones (teoría de control de sistemas). Existen casos particulares en que lo que se pretende es que exista una cota superior o máximo de un parámetro y en tales casos por ejemplo, se da la retroalimentación inhibitoria por producto. En otros casos lo

que se desea es que exista una cota inferior o un mínimo de tal forma que se da la retroalimentación positiva por producto para amplificar la señal a dicho (Como en el caso de la transcripción del factor de transcripción C-Fos, favorecida por la misma oncoproteína)

Redundancia vista como un mecanismo que asegura el correcto funcionamiento del sistema a pesar de perturbaciones externas (robustez vía redundancia). De esta forma como a quedado claro existen múltiples elementos con funciones análogas, tanto por que algunos pertenecen a una misma familia de proteínas, como por que otros presentan los mismos dominios que generan cruces de interacción entre diferentes rutas de señalización.

Modulación de la frecuencia, en función del número de pasos de la señal a través de diferentes elementos de transducción, esto es en forma interesante se observa que está perfectamente conservado a través de la evolución un número de etapas en que se lleva a cabo la activación en forma de cascadas. En este sentido es importante hacer notar que existen caminos más cortos para algunas de estas cascadas, por lo que es interesante analizar cual es el efecto obtenido bajo esta alternativa. Un caso típico es la activación de c-Jun tanto por la vía de Ras como por la vía corta con la proteína TPA.

Modulación del periodo definido como el tiempo característico que debe tener cada elemento en un estado determinado, lo cual tiene dos repercusiones: A) esto influye sobre la amplificación global de la señal, ya que se puede incrementar la concentración a partir de alguno de los pasos. B) Puede influir sobre la secuencia de los eventos, el sistema no es heurístico sino que se trata de una secuencia de eventos lógicos, y si esto llegase a fallar puede tener graves repercusiones en el mismo (caso de oncoproteínas resistentes a la inactivación, caso más típico Ras) En el sistema

bioquímico, estos intervalos de tiempo son el resultado de las características bioquímicas de los elementos involucrados.

Depuración de ruido. Este mecanismo se sabe que existe, y está representado en la desfosforilación inespecífica, a través de las proteínas fosfatasa. Ya que si bien existe desfosforilación específica por las proteínas fosfatasa esta no es la considerada en este punto si no dentro de los mecanismos de control por retroalimentación negativa. Este mecanismo está siempre presente, y en equilibrio con como las interacciones inespecíficas, producen activaciones de diversos elementos a muy bajas concentraciones. Esto es, la actividad de las fosfatasa desfosforila todo en general, limitando el "nivel de activación basal" de diversos elementos a un equilibrio de muy bajas concentraciones, a menos que se sobre pase un umbral de actividad, eliminando así el "ruido" debido a las mencionadas activaciones inespecíficas.

3.2.4.5 OBJETOS DE ANÁLISIS.

Patrones de salida. En este punto, si el modelo esta establecido para una red proteica, el punto final son los factores de transcripción a nivel nuclear. De tal forma que deberán monitorearse cuales de esos factores de transcripción se encuentran activados y en que proporción. En lo anterior debe tenerse especial cuidado en diferenciar las isoformas con que cuentan los mismos. El caso más notable es el del dímero AP1, el cual se encuentra en niveles basales constituido por la oncoproteína c-Jun y la proteína ATF2, unido a su secuencia de reconocimiento TRE, con la afinidad que le confiere dicha combinación. Sin embargo este heterodímero constitutivo ha de ser desplazado por la formación de las versiones activas (fosforiladas) de los homodímeros AP1 Jun-Jun o bien por los heterodímeros AP1 Jun-Fos. Para los cuales varía en tan solo una base la secuencia de reconocimiento TRE. Por supuesto aquí es posible y deberá darse un

entrecruzamiento en el reconocimiento, quedando la respuesta como función de la competencia que hagan estos diferentes factores por dichas secuencias en función de sus afinidades y sus concentraciones.

4 Conclusiones

El sistema de la transducción de señales como se ha visto tanto en la información general como en el análisis de las propuestas tradicionales, es un sistema muy extenso. El cual incluye muchos fenómenos que obedecen a diferentes leyes fisicoquímicas, que para el caso de la bioinformática quizás sería válido llamarlas diferentes lógicas. Cualquiera que sea de las dos aproximaciones por la que se decida tratar de abordar el sistema, este es muy basto y requiere para sus diferentes procesos diferentes mecanismos de representación. Por lo tanto se requerirá de la integración de diferentes métodos para realizar un modelo que logre una buena reproducción de las características globales del sistema bioquímico real.

Como resulta claro en la discusión, no es posible realizar un modelo para el sistema de la Transducción de Señales utilizando tan solo un mecanismo de calculo o simulación. Por lo tanto se concluye insuficiente el tratar de englobar un sistema tan vasto, en un solo mecanismo tomado de la biología, como los antes analizados. Quizás en la solución o planteamiento de modelos para un sistema complejo como la Transducción de señales, la bioinformática salga otra vez beneficiada de la observación de los sistemas que han tenido éxito en los sistemas biológicos a lo largo de la evolución y se encuentre un nuevo paradigma.

**ESTA TESIS NO SALE
DE LA BIBLIOTECA**

4.1 PROPUESTA SOBRE LA PLATAFORMA COMPUTACIONAL

Como se viene proponiendo, ninguna metodología aislada sería capaz de capturar el complejo funcionamiento del sistema global, y cubrir con lo propuesto en los puntos 3.2.3 al 3.2.4.5 Donde se propone que cualquier modelo deberá ser un híbrido entre ambas aproximaciones. Pero sobre todo se coincide con la visión de la teoría de la complejidad para abordar el problema. La cual

esencialmente es una forma diferente de abordar los sistemas complejos, donde a través de sistemas multiagentes es posible utilizar y combinar las diferentes herramientas tanto de la computación numérica como de la bioinformática.

Adicionalmente, debe tomarse en cuenta que todavía no existe ningún estándar para el modelaje de este tipo de sistemas complejos. Donde los resultados de cualquier modelo que se desarrolle deberán poderse corroborar por cualquier otro grupo de investigación. Lo cual lleva a considerar que, o bien se utiliza una plataforma de dominio público o se deberá proveer de cierta accesibilidad remota para cualquiera que desee corroborar simulaciones realizadas.

Un buen ejemplo que combina todo lo anterior es la plataforma GNU (de dominio público) de sistemas multiagentes enfocada a vida artificial, llamada SWARM. Desarrollada en el Instituto Santa Fe de Nuevo México (E.U.). Sobre la cual se pueden implementar modelos complejos utilizando en forma concurrente, diferentes simuladores de comportamiento discreto, de sistemas reales. Como la difusión Browniana (a través de ciertas "bibliotecas" de archivos) que podría ser la base para un modelo de reacción-difusión, que normalmente aborda la computación numérica tradicional. Adicionalmente esta plataforma permite utilizar cuando sean requeridas, otras herramientas de la bioinformática tales como las redes neuronales algoritmos genéticos y autómatas celulares. Por lo que se concluye es una buena alternativa, tanto computacional como conceptual, para modelar el proceso conocido como transducción de señales celulares.

4.2 PROYECCION A FUTURO.

La nueva disciplina de la bioinformática resulta altamente promisoría tanto en el mediano como en el corto plazo, en las diferentes áreas del conocimiento y la tecnología en que incide. Cabe notar que se encuentra en una fase de pleno desarrollo y que básicamente no existen limitaciones de tipo económico para su desarrollo, ya que el laboratorio *in silico*, (estaciones de trabajo) y sus "reactivos" (información extraída de bases de datos, la mayoría de las veces gratuita a través de internet) son poco onerosos relativamente. Por lo que los esfuerzos que se están realizando a nivel nacional, pueden resultar muy importantes para evitar un rezago en el desarrollo de lo que algunos consideramos como el futuro de la investigación (*in silico*) farmacéutica: La bioinformática.

5 Bibliografía

1. Alberts B., Bray D., Watson J., Lewis J, Martin R (1994) "Molecular Biology of the Cell" 3^o Edición. Garland. Publishing, Inc. Pags. 721-785
2. Arkin Adam. (1999) "Genetic Regulation at the Nanomolar Scale" Trends in Genetics. 15:2 65-69.
3. Bhatia P, Wang Z, Friedberg E (1996) "DNA repair and transcription." Current Opinion in Genetics & Development 1996, 6: 146-150.
4. Bing S. Karin M. (1996) "Mitogen-activated protein kinase cascades and regulation of gene expression." Current Opinion in Immunology 1996 8: 402-411
5. Bolin J.T. "All in the family: Structural and evolutionary relationships among three modular proteins with diverse functions and variable assembly" (1998) Protein Science. 7:1661-1670.
6. Bourne H. R., Williams L. T., (1994) "Cell regulation." Current Opinion in Cell Biology 1994, 6: 161-162.
7. Bray,D. (1990) "Intracellular signalling as a parallel distributed process". Journal of Theoretical Biology 143, 215-231.
8. Bray,D. (1995) "Protein molecules as computational elements in living cells". Nature, 376, 307-312.
9. Bray,D. Lay,S. (1994) "Computer simulated evolution of a network of cell signalling molecules". Biophysical Journal, 66, 4, 972-977.
10. Broach J. R., Levine A. (1997) "Oncogenes and cell proliferation." Current Opinion in Genetics & Development 7: 1-6
11. Buettner M.H., Riley M.R. "Monte Carlo Simulation of Diffusion and Reaction in Two-Dimensional Cell Structures". (1995) Biophysical Journal. 68:1716-1726.
12. Cahill M.A., Janknecht R., Nordheim A., (1996) "Signalling pathways:Jack of all cascades." Current Biology 1996, 6: 16-19
13. Cantley L. C., (1997) "Cell regulation" Current Opinion in Cell Biology 9: 131-133
14. Cobb M.H. (1996) "MAP Kinase Signaling Pathways" Promega Notes Magazine 59: 37
15. Davidson E.H., Chiou-Hwa Y. (1998) "Genomic Cis-Regulatory Logic: Experimental and Computational Analysis of Sea Urchin Gene". Science 279: 1896-1902.
16. Decroly O., Goldbeter, A (1982) "Birhythmicity, chaos, and other patterns of temporal self-organization in a multiply regulated biochemical system." Proceedings of the Natural Academy of Science. USA 79: 6917-6921.
17. Eyster Kathleen M. (1998) "Introduction to Signal Transduction" Biochemical Pharmacology, 55: 1927-1938.
18. Fanger G. R, Gerwins P, Widmann C, Jarpe M B, Johnson G. L, (1997) "MEKKs, GCKs, MLKs, PAKs, TAKs, and Tpls: upstream regulators of the c-Jun amino-terminal kinases?" Current Opinion in Genetics & Development 1997 7: 67-74.
19. Force T. (1994) "Enzymatic Characteristics of the c-Raf-1 protein kinase" Proceedings of the Natural Academy of Science. USA 91, 1270-1274.
20. Hasson O. (1997) "Towards a General theory of Biological Signaling" Journal of Theoretical Biology 185, 139-156

21. Heinrich R., Schuster S. (1998) "The modeling of metabolic systems. Structure, control and optimality" *Biosystems* 47: 61-77.
22. Helms J., McCammon A (1997) "Kinase conformations: A computational study of the effect of ligand binding" *Protein Science* (1997) 6: 2336-2343.
23. Hofestädt Ralf, Julio Collado (1999) "Modeling and simulation of gene regulation and metabolic pathways" *Biosystems* 49 79-82.
24. Hunter T. (1997) "Oncoproteins Networks" *Cell*. 88: 333-346.
25. Igamberdiev A. U. (1998) "Time reflectivity and information processing in living systems: a sketch for unified information paradigm in biology" *Biosystems*. 46 95-101.
26. Janin J. "Quantifying Biological Specificity: The Statistical Mechanics of Molecular Recognition" (1996) *Proteins: Structure Functions and Genetics*. 25:438-445.
27. Jones S., Thorton J. "Principles of protein-protein interactions".(1996) *Proceedings of the Natural Academy of Science. USA*. 93: 13-20 .
28. Junius K. (1994) "Cloning, expression, and spectroscopic studies of the Jun leucine zipper domain" *European Journal of Biochemistry* 219:3 877-86
29. Karin M. (1997) "AP-1 Function and regulation" *Current Opinion in Cell Biology* 9: 240-246
30. Katz M. E., McCormick F. (1996) "Signal transduction from multiple Ras effectors." *Current Opinion in Genetics & Development* 1997, 7: 75-79.
31. Kholodenko N. B., (1997) "Quantification of information transfer via cellular signal transduction pathways" *Federation of European Biochemical Societies Letters* 4141 430-434.
32. Lewin B. (1998) "GENES VI" 6ª Edición Oxford University Press. Pags. 1053-1145
33. Liberman E.A. (1998) "Biological information and laws of nature" *Biosystems*: 46 103-106.
34. Marijuan P.C. (1995) "Enzymes, artificial cells and the nature of biological information" *Biosystems*: 35 167-170.
35. Mc.Coy A.J. "Electrostatic Complementarity at Protein/Protein Surfaces". (1997) *Journal of Molecular Biology*. 268:570-584.
36. McCammon J.A. "Theory of biomolecular recognition" . (1998) *Current Opinion on Structural Biology*. 8:245-249.
37. Morrison D. Cutler E R. (1997) " The complexity of Raf-1 regulation" *Current Opinion in Cell Biology* 9: 174-179.
38. Nadel L. Stain Daniel L. (1995) "1993 Lectures in complex Systems" Addison Wesley Pags. 143-230.
39. Neel B., Tonks N. (1997) "Protein tyrosine phosphatases in signal transduction" *Current Opinion in Cell Biology* 9: 193-204.
40. Northrup S, Erickson H.P. "Kinetics of protein-protein-association explained by Brownian dynamics computer simulaton" (1992) *Proceedings of the Natural Academy of Science. USA* 89:3338-3342.
41. Parsons J. T., Parsons J. S. (1997) " Src family protein tyrosine kinases: cooperating with growth factor and adhesion signaling pathways" *Current Opinion in Cell Biology* 9: 187-192.
42. Paton R.C. (1999) "Intracellular signaling proteins as "smart" agents in parallel distributed process" *Biosystems* 50:159-171.
43. Paton Ray C. (1993) "Some computational models at the cellular level" *Biosystems* 29:2-3 63-75
44. Pawson T. (1995) " Protein modules and signalling networks" *Nature* 373: 573-579.
45. Pawson T., Schott J.D. (1997) "Signaling Trough Scaffold, Anchoring, and Adaptor Proteins".

Science 278: 2075-2080.

46. Ranish J.A., Steven H. (1996) "Transcription: basal factors and activation." *Current Opinion in Genetics & Development* 1996, 6: 151-158.
47. Ravi i, Upinder B. S. "Emergent properties of Networks of Biological Signaling Pathways" (1999) *Science* 283 15 January 381-387
48. Robinson M.J., Cobb M.H. (1997) "Mitogen-activated protein kinase pathways" *Current Opinion in Cell Biology* 9: 180-186.
49. Root-Bernstain R., Dillon F. "Molecular Complementarity I: the Complementarity Theory of the Origin and Evolution of Life" (1997) *Journal of Theoretical Biology*. 188:447-479.
50. Saltiel A., Sawer T. (1996) "Targeting Signal Transduction in the discovery of antiproliferative Drugs" *Chemistry and Biology* 3: 887-893
51. Schalkoff R.J. (1990) "Artificial Intelligence: an Engineering Approach" Mc-Graw-Hill, Inc. Pags. 1-33
52. Schamel W.W.A, Dick T.P. (1996) "Signal Transduction: Specificity of Growth Factors Explained by Parallel Distributed Processing" *Medical Hypothesis* 47: 249-255.
53. Schwab E.D., Pienta K.J. (1996) "Cancer as a Complex Adaptative System" *Medical Hypotheses*. 47: 235-241.
54. Schwab E.D., Pienta K.J. (1997) "Modeling signal transduction in normal and cancer cells using complex adaptative systems". *Medical Hypotheses*. 48: 111-123.
55. Treisman R. (1996) "Regulation of transcription by MAP kinase cascades." *Current Opinion in Cell Biology* 1996 8: 205-215.
56. Vasconcelos J. (1998) "Paradigmas de la inteligencia artificial". *Comunicación personal*.
57. Vermeulen W., Hoeijmakers J, Egly J.M (1996) "TFIIH: a key component in multiple DNA transactions." *Current Opinion in Genetics & Development* 1996, 6: 26-33.
58. Ying C., Ferrell J., (1996) "Ultrasensitivity in the in the mitogen-activated protein kinase cascade". *Proceedings of the Natural Academy of Science. U.S.A.* 93: 10078-10083.