



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE QUÍMICA

“Estudio proteómico de los cambios de fosforilación y glicosilación en Cáncer de mama.”

TESIS

**QUE PARA OBTENER EL TÍTULO DE:
QUÍMICO FÁRMACO BIÓLOGO**

**PRESENTA:
MERINO VAQUERO JULIO CÉSAR**



MÉXICO, D.F

2010



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.

JURADO ASIGNADO:

PRESIDENTE: Profesor: DRA. MARINA GAVILANES RUIZ

VOCAL: Profesor: M. en C. LUIS TONATIHUT SÁNCHEZ LINARES

SECRETARIO: Profesor: DRA. NORA ANDREA GUTIÉRREZ NÁJERA

1er. SUPLENTE: Profesor: DRA. VANESSA REBECA MAYA AMPUDIA

2° SUPLENTE: Profesor: M. en C. RAÚL GENARO AGUILAR CABALLERO

SITIO DONDE SE DESARROLLÓ EL TEMA:

INSTITUTO NACIONAL DE MEDICINA GENÓMICA (INMEGEN). PÉRIFÉRICO SUR,
No. 4124, Torre Zafiro II, PISO 6, COLONIA JARDINES DEL PEDREGAL. DEL.
ALVARO OBREGÓN, MEXICO D.F. C.P. 01900.

ASESOR DEL TEMA: DRA. NORA ANDREA GUTIÉRREZ NÁJERA.

SUSTENTANTE: MERINO VAQUERO JULIO CÉSAR

Dedicatorias y agradecimientos.

A mi padre, por estar ahí en las buenas y en las malas, por desvivirte y aguantarme todas las tonterías que he hecho en mi vida. Gracias por regañarme y abrazarme cuando lo necesitaba. Y lo más importante, gracias por enseñarme a ser un hombre. Te amo papá.

A mí madre en paz descansa, porque a pesar de no estar aquí eres un pilar fundamental en mi vida. ¡Sí se pudo!.

A mi segunda madre, Erika. Gracias a tí pude terminar esta carrera. Gracias por aguantar mis malos humores y mis groserías, por estar ahí cuando te necesitaba. Simplemente gracias por darme una parte de tu vida y no solo te quiero, te admiro.

A mis hermanas: Iety y Fabi. Gracias por entenderme y estar a mi lado. Buena vibra a las dos y a disfrutar la vida, ya nos toca.

A la mujer que amé, amo y amaré toda mi vida. Gracias Lidia por enseñarme a vivir. Por las buenas, las malas y las muy malas. Gracias por la mejor etapa de mi vida, eres la cosa más linda que he conocido y eso no cambiará nunca. Mi pequeña vakera bebe. Gracias por ser mi vakera y más vale amar una chispa que odiar la obscuridad. ¡Te amo!, amor de mi eternidad.

A mis hermanos de vida: Juana, Oscar, Vianney y mi pequeña mamá Claus. Gracias por todo. Por los regaños, los abrazos y los hombros donde llorar y valla que los moje, por las buenas y constates borracheras. Sin ustedes mi vida no sería la misma. A la chiva, gracias por las buenas charlas y presentarme al amor de mi vida, te quiero chiva.

A los de la prepa y los Chávez: shaggy, vivian, mega, momis, mars, viry, chucho, Iram, homero, ivan y lupe. Por los buenos cuernavacasos y Tlalmanalco, porque sin ustedes mi vida sería aburrida, jajaja.

A mis amigos de la Fac: Axel, Pony, Kory, Rafa, mi valedor querido (Alicia), Gabo, Memo, Piffs y demás personitas que compartieron una buena siesta en cualquier clase. Gracias por hacer muy amena esta hermosa carrera.

A mi asesora, la Dra. Nora Gutiérrez por su confianza, porque ambos aprendimos en esta aventura llamada tesis.

A la Dra. Marina. Gracias por su paciencia y mil disculpas por lo ocurrido.

En fin, a todas esas personitas que han sido parte de mi aprendizaje. Las que me hicieron llorar, las que me hicieron reír pero sobre todo a las que dejaron un cachito de su vida en mi corazón. ¡Gracias!.

Índice.

Abreviaturas.....	6
Glosario.....	8
Resumen.....	13
Capítulo 1. Introducción.....	15
Capítulo 2. Antecedentes	19
2.1 Cáncer.....	19
2.1.1 Factores de riesgo.....	20
2.1.2 Factores Internos (genéticos).....	20
2.1.3 Factores Externos (Ambientales).....	21
2.2 Desarrollo de cáncer a nivel celular.....	22
2.2.1 Genes supresores de tumores.....	24
2.2.2 Oncogenes.....	25
2.2.3 Propagación del Cáncer (Angiogénesis y Metástasis).....	26
2.3 Clasificación del cáncer por la AJCC y UICC.....	28
2.3.1 Otros tipos de clasificación.....	30
2.3.2 Métodos para la clasificación.....	30
2.4 Terapias.....	31
2.5 Cáncer de mama.....	32
2.5.1 Factores que influyen en el desarrollo del Cáncer de mama.....	33
2.5.2 Anatomía y fisiología de la mama.....	34
2.5.3 Tipos de Cáncer de mama.....	36
2.5.4 Diagnóstico en Cáncer de mama.....	39
2.5.5 Pronóstico en Cáncer de mama.....	40
2.5.6 Terapias en Cáncer de mama.....	43
2.6 Estructura, formación, clasificación y estudio de proteínas.....	48
2.6.1 Fosforilación.....	51
2.6.2 Glicosilación.....	55
2.7 Biomarcadores.....	60
Capítulo 3. Objetivos.....	64
3.1 Objetivo general.....	64
3.2 Objetivos particulares.....	64
Capítulo 4. Hipótesis.....	64
Capítulo 5. Metodología. Protocolo de análisis de Fosfoproteínas y Glicoproteínas.....	65
5.1 Líneas celulares.....	66

5.3 Protocolo para purificación de glicoproteínas mediante kits Qiagen.....	70
5.4 Precipitación de proteínas.....	72
5.5 Cuantificación de proteínas.....	73
5.6 Preparación y tinción de geles SDS-PAGE.....	76
5.7 Análisis de geles obtenidos.....	78
5.7.1 Protocolo para la digestión de las bandas seleccionadas con tripsina.....	78
5.7.2 Análisis de fracciones de fosfoproteínas y glicoproteínas mediante LC/ESI-MS/MS.	81
Capítulo 6. Resultados.	86
6.1 Rendimiento de la extracción de fosfoproteínas y glicoproteínas de las líneas celulares MDA-MB 231, MCF7, Hs578T y Hs578BsT.	86
6.3 Comparación del patrón de bandas de las fosfoproteínas y glicoproteínas de las líneas celulares MDA-MB 231, MCF7 Y Hs578T en contra del patrón de bandas de las fosfoproteínas y glicoproteínas de la línea celular control Hs578BsT.....	91
6.3.1 Fosfoproteínas.	92
6.3.2 Glicoproteínas.	92
6.4 Resultados del análisis por LC/MS- MS/MS de las proteínas seleccionadas de los geles corridos con las fracciones obtenidas de la extracción de glicoproteínas y fosfoproteínas.	95
6.5 Selección final de posibles biomarcadores fosforiladas.	99
6.7 Selección final de posibles biomarcadores glicosilados.....	103
Capítulo 7. Discusión. Biomarcadores fosforilados y glicosilados seleccionados.	106
Capítulo 8. Conclusiones.	122
9. Referencias Bibliográficas.....	123
ANEXO I.....	129
ANEXO II.....	140
ANEXO III.....	143
ANEXO III.....	152
ANEXO IV.....	163

Abreviaturas.

ADN. Ácido Desoxirribonucleico.

AHA1. Activator of heat shock 90 kDa protein ATPase homolog 1.

AJCC. American Joint Committee for Cancer.

ATCC. American Type Culture Collection (Colección Americana de Tipos de Células).

ARN. Ácido Ribonucleico DMEM. Medio modificado de Eagle's Dulbecco's.

Ca mama. Cáncer de mama.

CDC. Centro de prevención y control de enfermedades (Center for Disease Control and Prevention).

CD44E. Proteína CD44 en su forma epitelial.

CHAP's. 3-[(3-cloramidopropil)dimetilamonio]-1-propanosulfonato.

2D. Gel en 2 dimensiones.

DCIS. Carcinoma ductal *in situ*.

DIGE. Electroforesis diferencial en geles (Difference in gel electrophoresis).

DMEM. Medio Modificado de Dulbecco's Eagle.

DTT. Ditiotreitól.

EGF. Factor de Crecimiento Epitelial.

FGF-1. Factor de Crecimiento Fibroblástico tipo 1.

FNIP1. Folliculin interacting protein 1.

HSF1. Heat shock transcription factor 1.

ESI-MS. Espectrometría de masas por ionización de electroespray (electrospray ionization mass spectrometry).

HEPES-base. Ácido 4-(2-hidroxietil)-1-piperazinaetansulfónico.

hnRNP U. Ribonucleoproteína U.

HPLC. Cromatografía líquida de alta resolución (High Pressure Liquid Chromatography).

HSP- α . Proteína de choque térmico HSP 90- α .

ILC. Carcinoma lobular infiltrante o invasivo.

IBC. Cáncer de mama inflamatorio.

IDC. Carcinoma ductal infiltrante.

IMHQ. Inmunohistoquímica.

kDa. Kilo Daltones.

LCIS. Carcinoma lobular *in situ*.

MALDI MS o TOF. Espectrometría de masas por absorción/desorción de la matriz asistida por laser (matrix-assisted laser desorption/ionization mass spectrometry).

MRM. Masectomía Radical Modificada.

MS. Espectrometría de Masas.

MS/MS. Espectrometría de Masas en Tándem (Mass Spectrometry).

NCL. Nucleolina.

NCI. Instituto Nacional de Cáncer (National Cancer Institute).

NIST. Instituto Nacional de Estándares y Tecnología (National Institute of Standards and Technology).

80 kDa. Subunidad β de la Glucosidasa II.

OMS. Organización Mundial de la Salud.

PSA. Persulfato de amonio.

pI. Punto isoeléctrico.

PTGES3. Prostaglandin E synthase 3 (cytosolic).

RER. Retículo endoplásmico rugoso.

RE. Receptores de Estrógenos.

RP. Receptores de Progesterona.

SAR. Scaffold-Attached Región.

SDS. Dodecil Sulfato de Sodio.

SI. Sistema Inmune.

SMYD3. SET and MYND domain containing 3.

TBS. Solución Tris-Base.

TERT. Telomerase reverse transcriptase.

TFA. Ácido trifluoroacético.

TMC. Tratamiento conservador de la mama.

TMN. Tumor, ganglios y metástasis.

TNF- α . Factor de Necrosis Tumoral Alfa.

TOMM34. Translocase of outer mitochondrial membrane 34.

UICC. Unión Internacional en contra del Cáncer (International Union Against Cancer).

UNC45A. UNC-45 homolog A.

VEGF. Factor de Crecimiento Endotelial Vascular.

Glosario.

***Acilación.** Proceso de agregar un grupo acilo a un compuesto. El compuesto que provee el grupo acilo es denominado el grupo acilante y esta reacción se realiza gracias a una reacción de sustitución.

***Aminoácido.** Unidad fundamental de un péptido. Una molécula orgánica la cual se compone de un grupo amino ($-\text{NH}_2$), un grupo carboxílico ($-\text{COOH}$; ácido), un carbono α y una cadena R la cual puede ser un ión hidrógeno o un grupo aromático.

***Angiogénesis.** Generación de vasos sanguíneos para proporcionar los nutrientes necesarios en el crecimiento tumoral.

***Anticuerpo.** Glucoproteínas del tipo gamma globulina. Se encuentran de forma soluble en la sangre u otros fluidos corporales de los vertebrados, actúan como receptor de los linfocitos B y son empleados por el sistema inmune para identificar y neutralizar elementos extraños (antígenos).

***Antígeno.** Molécula extraña a la cual ataca el anticuerpo (puede provenir de bacterias, hongo, virus, etc).

***BCR1.** Gen que se encuentra presente en más de 45% de los casos de Cáncer de mama hereditario y 80% en los casos de cáncer de ovario.

***BCR2.** Gen que se encuentra relacionado con formas poco comunes de Cáncer de mama heredado.

***Biomarcador.** Producto del cuerpo (metabolitos) que tal vez respondan a una interacción negativa con el medio (carcinogénesis, mutagénesis, etc.).

***Biomolécula.** Compuesto químico que se encuentra en los organismos vivos y que están formados por sustancias químicas compuestas principalmente por carbono, hidrógeno, oxígeno, nitrógeno, sulfuro y fósforo; éstas son el fundamento de la vida y cumplen funciones imprescindibles para los organismos vivos.

***Cambio postraduccionaL.** Modificaciones que sufren las cadenas polipeptídicas para adquirir su estado activo.

***Cáncer de mama.** Grupo de tumores malignos epiteliales, los cuales se desarrollan en las glándulas mamarias con la consecuente metástasis de la región de los nódulos linfáticos y órganos periféricos y finalizando con la muerte del paciente.

***Carcinogénesis.** Transformación progresiva de las células normales en células malignas, que se produce por la adquisición de autonomía de las mismas, lo que es un

reflejo de una regulación y expresión anormal de su carga génica. Como consecuencia final se induce una neoplasia.

***Células somáticas.** Células que constituyen la mayoría de las células de nuestro cuerpo. Contienen toda la información genética de un individuo, organizada en 23 pares de cromosomas, en el humano se las denomina células diploide: $2n$ cromosomas, 2×23 cromosomas.

***Ciclinas.** Familia de proteínas que regula la actividad de Cdks y el control de la progresión a través del ciclo celular.

***c-MYC.** Factor de transcripción que controla la expresión de varios genes involucrados en la división celular. La forma mutante de este gen actúa como un oncogen y se ha relacionado con varios tipos de cánceres, incluyendo el neuroblastoma.

***Cofactor.** Componente no proteico, termoestable y de bajo peso molecular, necesario para la acción de una enzima y formar una haloenzima,

***Coenzima.** Cofactores unidos débilmente a la apoenzima y actúan fundamentalmente como uno de los sustratos específicos de la reacción, que lleven a cabo.

***Cromatografía.** Técnica analítica empleada para la identificación de compuestos.

***Enlace intermolecular.** Unión que como resultado de las fuerzas de carácter electrostático que se establecen entre las moléculas, consigue mantenerlas unidas en una red, ejemplo de éstos son: enlaces de hidrógeno, fuerzas electrostáticas, fuerzas de Van der Waals, entre otros.

***Enlace intramolecular.** Enlaces que unen entre sí conjuntos de átomos para dar lugar a moléculas ó estructuras superiores a los átomos. En este tipo de enlaces se incluyen el iónico, el covalente y el metálico. En proteínas, el ejemplo más claro son los puentes disulfuro.

***ErbB2.** Factor de Crecimiento Epidermal 2, un receptor de membrana que es un receptor de tirosin cinasas que está asociado con la traducción de señales para estimular el crecimiento celular.

***Factor de riesgo.** Atributo o exposición de una persona, una población o el medio que están asociados a la probabilidad de que ocurra un evento.

***Factores internos.** Acontecimientos que ocurren dentro del individuo, es decir a nivel genético, los cuales contribuyen a la expresión del proteoma en espacios y tiempos determinados.

***Factores externos.** Acontecimientos ambientales que rodean al individuo, los cuales

contribuyen a la expresión del proteoma en espacios y tiempos determinados.

***Fase móvil.** Fase en la cromatografía que contiene el analito a analizar y la cual entra en contacto con la fase estacionaria.

***Fase estacionaria.** Matriz que posee las características para retener cierto tipo de analitos por la cual se desplaza la fase móvil.

***Fosforilación.** Adición de un grupo fosfato inorgánico (PO_4^{3-}) a cualquier otra molécula. En las proteínas, esta adición es catalizada por proteína cinasas.

***Genes supresores.** Genes encargados de la proliferación celular, por lo tanto son reguladores negativos de crecimiento y cuando no están presentes en la célula o se encuentran inactivos, las células dejan de crecer normalmente y adquieren propiedades proliferativas anormales.

***Genética.** Disciplina de la biología, la cual se encarga del estudio de la herencia y la variación de ésta en los organismos.

***Glándulas mamarias.** Órgano característico de los mamíferos, se desarrollan a partir de crestas mamarias definidas que se extienden a ambos lados del tronco en el embrión de mamífero, éstas se hallan en ambos sexos, pero en los machos dejan de desarrollarse mucho antes de la pubertad. Se emplean para el alimento de las crías, además de ser necesarias para la regulación hormonal en las hembras.

***Glicación.** Adición de una glucosa a una molécula, a nivel bioquímico ocurre cuando las moléculas de glucosa reaccionan con una proteína, uniéndose a ésta ejemplos de estas proteínas son: proteínas que transportan el colesterol, la capa protectora que rodea los nervios, la hemoglobina, los glomérulos de los riñones y con proteínas en el cristalino de los ojos.

***Grupo prostético.** Cofactores que están fuertemente unidos a la apoenzima.

***Glucosilación.** Adición de azúcares mediante una glucosidasa, este tipo de modificación se divide en dos grandes grupos: O-glicoproteínas (los grupos aceptores son Ser y Thr) y N-glicoproteínas (el residuo aceptor es Asn).

***Hidroxilación.** Adición de un grupo hidroxilo por hidrolasas a los residuos de tirosina.

***Hormonas esteroideas.** Grupo de hormonas hidrofóbicas que son derivadas del colesterol.

***Metástasis.** Capacidad de diseminación de las células neoplásicas.

***Metilación.** Adición de un grupo metilo en los amino α de los residuos de Lys, His y Arg.

***Mutaciones.** Alteración o cambio en la información genética de un organismo y que por

lo tanto produce un cambio en su expresión (proteoma). Se presenta súbita y espontáneamente. Se puede transmitir o heredar a la descendencia.

***Neoplasia.** Alteración de la proliferación y muchas veces de la diferenciación celular, que se manifiesta por la formación de una masa o tumor.

***Nitrosilación.** Modificación postraducciona en la cual se añade grupo nitroxilo a una proteína.

***Nitración.** Proceso bioquímico general que introduce un grupo nitro en un residuo de Tyr.

***Oncogenes.** Protooncogenes alterados, los cuales dirigen la producción de proteínas como ciclinas, factores de crecimiento, receptores entre otros, los mismos que estimulan la proliferación celular y son capaces de generar una multiplicación descontrolada de las células.

***Péptido.** Unión de dos o más aminoácidos mediante un enlace peptídico que une el extremo carboxilo terminal de un aminoácido y el extremo amino del otro.

***Péptido señal.** Secuencia hidrofóbica unida al amino terminal de una cadena polipeptídica que indica el lugar al cual debe dirigirse dentro y fuera de la célula.

***Poliacrilamida.** Acrilato que pertenece a un tipo de polímeros vinílicos que se deriva de monómeros acrilatos, que son ésteres que contienen grupos vinílicos, es decir, dos átomos de carbono unidos por una doble ligadura, directamente enlazados al carbono del carbonilo. Este polímero es el más empleado en la separación electroforética.

***Prenilación.** Adición de tipos específicos de lípidos (grupos preniles) al C terminal de los residuos de cisteína de la cadena polipeptídica.

***Proteína.** Macromoléculas formadas por una cadena lineal de aminoácidos, desempeñan un papel fundamental para la vida y son las biomoléculas más versátiles y diversas e imprescindibles para el crecimiento del organismo ya que realizan una enorme cantidad de funciones diferentes.

***Proteoma.** Expresión de proteínas en un lugar y momento específico por parte de un organismo.

***Proteómica.** Disciplina de la genética encargada del estudio del proteoma de los organismos.

***p53.** Factor de transcripción (codificado por el gen supresor tumoral *p53*) que detiene el ciclo celular en G1 en respuesta al daño en el ADN e induce apoptosis.

***RB1.** Gen supresor tumoral que reduce la probabilidad de que una célula en un

organismo multicelular se transforme en una célula cancerígena. Delaciones en él se encuentran relacionadas con el desorden del Retiblastoma.

***Replicación.** Síntesis de ADN a través de una cadena que sirve como molde.

***SUMOilación.** Sistema de modificación post-traducciona que consiste en la unión covalente a determinadas proteínas de un péptido similar a la ubiquitina denominado SUMO (small ubiquitin-like modifier), el cual tiene como finalidad la degradación de proteínas.

***Traducción.** Generación de una cadena polipeptídica apartir de un mRNA.

***Transcripción.** Síntesis de una molécula de RNA a partir de una cadena molde de ADN.

***Tubulina.** Proteína de citoesqueleto que se polimeriza para formar los microtúbulos.

***Ubiquitinación.** Vía de degradación de proteínas dependiente de ATP de los mamíferos en la cual se adiciona una ubiquitina al extremo C-terminal de un péptido para ser destruido por el proteasoma 26S.

***WNT7B.** Gen que codifica para la proteína Wnt-7b, la cual pertenece a una familia que se encuentra relacionada con la oncogénesis y en varios procesos de crecimiento celular. Este gen se ha observado en la progresión del cáncer gástrico, de estómago y pancreático.

Resumen.

El cáncer es una de las cuatro principales causas de defunciones en la población mundial. El Cáncer de mama es el principal tipo de neoplasias que presentan las mujeres y representa más del 13% de defunciones, equivalente a 7.9 millones. Se prevé que para el año 2030 esta cifra se incrementará en un 45%, es decir será causante de más de 12 millones de defunciones anuales. Se ha hecho hincapié en una terapia preventiva que garantice un diagnóstico temprano, ya que el pronóstico sería más favorable si la detección de esta enfermedad se hace en un estadio temprano. Desde la década de los noventa se reportaron proteínas que se encontraban relacionadas con algún tipo de cáncer y que se expresaban solo en pacientes enfermos y no así en los sanos ^{1,2}. A estas proteínas se les dio el nombre de biomarcadores, los cuales podemos definir como: “Un producto del cuerpo (metabolitos) que tal vez respondan a una interacción negativa con el medio (carcinogénesis, mutagénesis, etc.)”⁴. En años recientes esta área ha tenido un avance sobresaliente gracias al desarrollo de la genómica y sus múltiples ramas, entre ellas la proteómica, la cual nos permite el estudio del proteoma y que se define como: “La expresión de productos proteicos en un lugar y tiempo determinado”. El uso de las herramientas de la proteómica permite la separación e identificación de varias proteínas a la vez. Una de las características principales de las proteínas es que pueden sufrir una o varias modificaciones postraduccionales por medio de enlaces no covalentes y covalentes o bien mediante la unión de alguna molécula, como pueden ser azúcares (derivados de glucosa) y los grupos fosfatos. Ambas modificaciones son capaces de modular vías metabólicas de la célula, las cuales pueden contribuir a su crecimiento o a su destrucción (muerte programada), es por esto que el siguiente estudio tuvo como finalidad la identificación de fosfoproteínas y glicoproteínas para ser propuestas como posibles biomarcadores en los aspectos de diagnóstico, terapéuticos y/o de pronóstico en el Cáncer de mama. Este estudio se realizó en un modelo celular en el cual se utilizaron cuatro líneas celulares: MCF7, MDA-MB231, Hs578T y Hs578BsT (esta última como control). De éstas se obtuvieron fracciones de fosfoproteínas y glicoproteínas, las primeras, fueron aisladas mediante el uso de anticuerpos contra serina, treonina y tirosina fosforilada, y las segundas aisladas mediante la unión a una lectina con afinidad al grupo O-glico de las glicoproteínas. Una vez obtenidas las fracciones fueron separadas en geles SDS-PAGE y los geles resultantes fueron digitalizados mediante el programa

LABSCANER™, permitiendo así la realización de una proteómica comparativa. Una vez seleccionadas las bandas, fueron extraídas, digeridas y analizadas mediante LC/ESI-MS/MS y el algoritmo de búsqueda MASCOT. Los resultados fueron procesados con herramientas bioinformáticas. El resultado fue la identificación de 3 fosfoproteínas y 2 glicoproteínas. Las fosfoproteínas identificadas fueron la HSP90, la Nucleolina y la Ribonucleoproteína U (hnRNAU), las cuales se encuentran relacionadas al proceso de la transcripción. Interviniendo en la unión de diversas moléculas y en la estabilización de la sobreexpresión de oncogenes. Las tres proteínas se localizan cerca del núcleo.

Las glicoproteínas identificadas fueron la CD44E y la 80K-H. La CD44E está relacionada con el crecimiento tumoral y se localiza en la membrana, actúa mediante la unión a p53 y causa un incremento en las CD44s, las cuales se han relacionado con el desarrollo celular y la angiogénesis. La 80K-H es la subunidad β de la glucosidasa, se ha reportado que participa en la estabilización de los factores de crecimiento que se sobreexpresan como el FGF1.

Cada una de las proteínas identificadas tienen características que las hacen candidatas a ser biomarcadores. Por ello es necesaria una caracterización de vías metabólicas mediante técnicas genómicas y proteómicas como la inmunohistoquímica, la inmunocitoquímica, el western blot, la cromatografía, entre otras.

Capítulo 1. Introducción.

En las últimas décadas se han realizado numerosos esfuerzos para prevenir, detectar y curar diversas enfermedades degenerativas, ejemplo de esto es el cáncer, el cual de acuerdo a la OMS se define como:

“Un proceso de crecimiento y diseminación incontrolados de células el cual puede aparecer prácticamente en cualquier lugar del cuerpo, y que genera un tumor que suele invadir el tejido circundante y puede provocar metástasis en puntos distantes del organismo” (OMS 2010).

El cáncer es responsable de 7.9 millones de defunciones anuales, lo cual lo posiciona como la segunda causa de mortandad después de las enfermedades cardiovasculares. Se prevé que para el 2030 esta cifra se incrementara en un 45%, representando casi 12 millones de defunciones. En el 2002 se convocó a la convención contra el Cáncer que fue celebrada en París, en ella se acordó una alianza entre investigadores, profesionales de la salud, pacientes, gobiernos, industria y medios de comunicación para luchar contra esta enfermedad, dando como resultado “La Carta de París”, que entre otras cosas establece el 4 de febrero como el Día Mundial de lucha contra el Cáncer.

De acuerdo a la OMS, los cánceres de mayor incidencia para hombres son: el cáncer de pulmón y próstata, mientras que para las mujeres son el cáncer cervicouterino y el de mama, este último es considerado como la neoplasia más frecuente entre las mujeres europeas y americanas.

En México, durante los años 2004 a 2007 el Cáncer de mama fue la principal causa de ingreso hospitalario en las mujeres, desplazando al cáncer de cuello uterino en el 2006 como la primera causa de defunciones, mientras que para los varones, la leucemia es la primera causa de ingreso hospitalario y el cáncer de próstata, seguido por el de tráquea, bronquios y pulmón.

A pesar de los exhaustivos estudios realizados para el entendimiento de la epidemiología y la mejora en el diagnóstico de la enfermedad, una de las principales causas de decesos es la detección tardía, debido a diferentes factores como:

- Una ineficiente cultura de la autoexploración. Apesar de haber mejorado en los últimos años no es suficiente, debido a la pobre difusión en la población de diferentes estratos sociales. En la actualidad muchos de los proyectos de salud se han enfocado a la medicina preventiva, la cual comienza a dar frutos. Sin embargo,

no se debe olvidar que existen comunidades en todo el mundo que no poseen ni la mínima información para la generación de una cultura preventiva de salud.

- Falta de recursos del sector salud. Los países participantes en la Carta de París han destinado una gran cantidad de recursos para la prevención, diagnóstico y tratamiento del Cáncer de mama como resultado de la convención celebrada en París, desgraciadamente el número de pacientes diagnosticados con esta enfermedad va en aumento, lo cual hace insuficiente estos recursos.
- La falla en los métodos de diagnóstico, aunado al poco entendimiento de la epidemiología de los blancos terapéuticos y la etiología de los cánceres humanos. Esta situación se ha tratado de abordar con el nacimiento de nuevas técnicas pertenecientes a la genómica, proteómica, entre otras. Las cuales tienen como objetivo fundamental la localización de blancos terapéuticos para el tratamiento y la búsqueda de biomarcadores para un diagnóstico temprano y confiable.

Un biomarcador es definido como: “Un producto del cuerpo (metabolitos) que tal vez respondan a una interacción negativa con el medio (carcinogénesis, mutagénesis, etc.)”⁴. En el artículo publicado por Bidart y colaboradores¹ en el siglo XX se habla de biomoléculas presentes en pacientes con cáncer como el Antígeno específico de próstata, la α -fetoproteína y la CA 12.5 que son considerados como los primeros biomarcadores reportados, a estas moléculas se les sumaron varias más. Estas moléculas no fueron del todo confiables debido a las técnicas poco sensibles y específicas que se emplearon para su identificación y caracterización. En años recientes este problema se ha ido resolviendo gracias al desarrollo de las disciplinas como la genómica, la proteómica por citar algunas, las cuales han proporcionado diversas técnicas, permitiendo la identificación de marcadores moleculares confiables, es decir, biomoléculas que varían su expresión en el transcurso de la enfermedad, éstas pueden ser utilizadas para un diagnóstico eficaz y la identificación de diversos estadios de la enfermedad o para la generación de una terapia individualizada que amplíe la expectativa y calidad de vida del paciente.

A principios de los 90's se observa una relación existente entre pacientes diagnosticadas con Cáncer de mama y dos proteínas: BRCA1 y BRCA2, relación confirmada tiempo después por diversos estudios. Se creyó que estas proteínas se comportaban de forma

homóloga a enfermedades con herencia autosómica ^{1,2}, ya que se observó su presencia en familiares de primera línea, estos genes han sido ampliamente estudiados pero no del todo comprendidos. Hoy en día se sabe que BRCA1 se encuentra presente en más del 45% de los casos de Cáncer de mama hereditario y 80% en los casos de cáncer de ovario, mientras que BCRA2 ha sido detectada en formas poco comunes de cáncer heredado. Otro ejemplo de proteínas relacionadas con cáncer son las tirosin cinasas y factores de crecimiento, los cuales intervienen en los procesos de iniciación, progresión y metástasis mediante la señalización angiogénica y mitogénica dentro de la célula y son regulados por procesos de fosforilación y defosforilación. Los ejemplos más claros de la importancia de estas proteínas se han visto reflejados en la generación de tratamientos cuyo blanco terapéutico son proteínas pertenecientes a estas familias, por ejemplo: Bevacizumab® el cual inhibe al factor de crecimiento endotelial vascular (VEGF), y el Trastuzumab®, el cual se une a ErbB2⁵, un marcador específico que al igual que BCR-ABL⁵ participan en el crecimiento celular.

Otro tipo de moléculas relacionadas con Cáncer de mama son los O-glicanos que son estructuras complejas y diversas, las cuales participan en la conformación estructural de las proteínas evitando su agregación, se ha observado que durante el cáncer ocurren cambios en la etapa de elongación de los oligosacáridos, teniendo como consecuencia la formación de antígenos asociados al cáncer por ejemplo: Antígeno Tn, Antígeno Sialil-Tn y Antígeno T, los cuales interactúan con proteínas de membrana (Mucinas) provocando la pérdida de polaridad y protegiendo a la célula tumoral del reconocimiento mediante la vía efectora del SI⁶.

En este trabajo nos proponemos identificar proteínas a través de técnicas proteómicas que en una célula de Cáncer de mama se modifiquen por fosforilación o glicosilación respecto a las de una célula sana, y así proponerlas como proteínas que puedan ser posibles biomarcadores o marcadores moleculares de esta enfermedad. Para esto se utilizará un sistema modelo de células en cultivo de Cáncer de mama, el cual incluye 4 líneas celulares: Hs578t, MCF7, MDA-MB231 y como control la línea celular de tejido normal de mama Hs578BsT.

De estas líneas se extraerán fracciones celulares eluidas en columnas de afinidad a fosfoproteínas y glicoproteínas (Qiagen™) que se separarán mediante geles SDS-PAGE para su posterior análisis por técnicas de espectrometría de masas. Estas herramientas que aseguran la presencia de estas modificaciones y su búsqueda en la base de datos

para conocer la relación existente entre proteína-Cáncer de mama.

Mediante este estudio se pretende obtener información de las proteínas que son modificadas en el Cáncer de mama, las cuales pueden servir para un diagnóstico temprano de pronóstico o bien como blanco terapéutico, colaborando así con la lucha en contra del Cáncer de mama.

Capítulo 2. Antecedentes

2.1 Cáncer.

El cáncer es uno de los cuatro problemas de salud mundial más graves de acuerdo a la OMS, superado por las enfermedades cardiovasculares, diabetes y enfermedades crónicas respiratorias. Representa el 13% de los decesos ocurridos en el 2007 y se prevé que en años posteriores esta cifra se incremente. De acuerdo a la OMS, definimos al cáncer como: “Un proceso de crecimiento y diseminación descontrolado de células, el cual puede aparecer prácticamente en cualquier lugar del cuerpo, que genera un tumor que suele invadir el tejido circundante y puede provocar metástasis en puntos distantes del organismo” o bien “Una enfermedad genética, producto de la proliferación anómala de un clon celular”. Décadas atrás, este tipo de enfermedades era considerada como incurable y en ocasiones como un sinónimo de muerte, hoy en día este concepto ha cambiado gracias a los avances en el diagnóstico y tratamiento que han permitido ampliar notablemente la expectativa de vida, hasta un punto en el cual los pacientes pueden ser considerados como curados, es decir en un estado en el cual ha desaparecido el tumor, pero que necesita un tratamiento que garantice que no existirá una reincidencia.

El cáncer tiene como principal característica la generación de neoplasias las cuales se pueden presentar en diversas partes del cuerpo de cualquier individuo sin importar género, edad, origen étnico y estatus social, siendo las más comunes por género las presentadas en la Tabla 1. El cáncer es considerado como una enfermedad multifactorial en la cual intervienen factores externos (ambientales) e internos (genéticos) que una vez presentes desencadenan una serie de procesos los cuales culminan en el desarrollo de neoplasias (tumores). Dichos tumores pueden estar primero localizados y de no ser tratados adecuadamente ocurren una serie de procesos como la angiogénesis y metástasis, de las cuales hablaremos más adelante, que permiten la diseminación de éstos en varias partes del organismo trayendo consigo complicaciones que pueden terminar en un deceso.

Tabla 1. Tipos de cáncer más comunes clasificadas por sexo. En la tabla se observan en orden descendente de acuerdo al número de defunciones reportadas en 2007 por la OMS (World Health Organization, www.who.int).

SEXO.	TIPOS DE CÁNCER
FEMENINO.	Mama. Pulmón. Estómago. Colón-recto. Cuello uterino.
MASCULINO.	Pulmón. Estómago. Hígado. Colón-recto. Esófago. Próstata.

2.1.1 Factores de riesgo.

Un factor de riesgo es toda aquella circunstancia, cambio o agente (químico o físico) que ocasione un desequilibrio en el organismo el cual desencadena la formación de algún tipo de malestar o enfermedad que se ve reflejado en la expresión proteica del organismo. Estos se pueden dividir en dos grande grupos:

- Factores Internos.
- Factores Externos.

2.1.2 Factores Internos (genéticos).

Cuando el cáncer fue descubierto como un desorden crónico se consideró como una enfermedad de tipo mendeliano debido a que muchas familias poseían el mismo desorden en varias generaciones consecuentes, hoy en día, gracias a los adelantos en el área de la genética se sabe que el cáncer se ocasiona por alteraciones génicas las cuales juegan un papel importante en el desarrollo de la enfermedad pero no decisivo, sin embargo el hecho de presentar estas características (mutaciones) no hacen obligatorio el desarrollo de la enfermedad, debido a que se necesitan una serie de procesos para desencadenarla. Estos procesos se llevan a cabo en el medio que rodea al organismo (Factores externos). El origen del cáncer o carcinogénesis comienza con un clon celular que es sometido a un proceso de varias fases en las cuales se dan una serie de alteraciones en el ADN, que repercuten en el desarrollo de las células y dan inicio a los procesos de promoción y progresión, proceso que termina en la transformación de células normales a células neoplásicas. Estas modificaciones se dan gracias a la interacción de factores genéticos y agentes externos (Factores ambientales). Estos factores pueden ser:

físicos, químicos y/o biológicos y generalmente inciden en genes supresores y oncogenes del genoma humano (Tabla 2), los cuales son encargados de regular el desarrollo celular. Este tipo de alteraciones se lleva a cabo en el ADN, y se pueden perder bases, adquirirlas, silenciarlas o cambiarlas, los daños más comunes en el ADN son: inserción, delección, silenciamiento, mutaciones con sentido, mutaciones sin sentido, impronta génica, entre otros. Todos estos cambios son reflejados en el proteoma de la célula, lo cual trae como consecuencia la transformación de células diferenciadas a células no diferenciadas causando la pérdida de la funcionalidad del tejido⁴.

Tabla 2. Genes mutados relacionados a cáncer. La tabla nos señala algunos genes mutados, sus consecuencias y el cáncer con el cual se asocian (Ruin, 2004).

Alteración Hereditaria	Gen(es) afectados.	Consecuencias.	Cáncer asociado.
Traslocación cromosómica.	BCL y ABL	Pérdida de función mediante una fusión de ambos genes.	Leucemia mieloide crónica.
Delección Homocigótica.	NFI	Desregularización de p21RAS	Neurofibrosarcoma. Feocrocitoma.
Mutaciones Sin sentido. Amplificación génica.	APC EGFR.	Anula el gen supresor tumoral. Sobreexpresión el receptor de la tirosinasa.	Carcinoma colorrectal. Glioma.
SNP's e hipermutabilidad.	TP53. Hmsh2 y HmLh1	Inestabilidad genética, Funcionalidad dañada del guardián del genoma	Carcinoma colorrectal. Cáncer de Pulmón. Síndrome de Muir-Torre.
Microsatélites.	HRAS	Pérdidas cromosómicas(genes)	Carcinoma colorrectal.
Impronta génica	IGF2 Y H19	Efectos sobre dotación génica.	Síndrome de Beckwith-Wiedemann y tumor de Wilms.
Metilaciones.	Múltiples.	Alteración en la expresión de la cromatina.	Carcinoma colorrectal.

2.1.3 Factores Externos (Ambientales).

Los principales factores externos que incrementan el riesgo para el desarrollo de cáncer de acuerdo al nivel económico de los países, reportados por la OMS son: el consumo de tabaco, de alcohol, la ingesta insuficiente de frutas y hortalizas, las infecciones por virus de la hepatitis B, de la hepatitis C y por virus del papiloma humano para países de bajos y de medianos ingresos, mientras que en los países desarrollados el consumo de tabaco y alcohol, el exceso de peso y la obesidad ocupan este rubro, sin olvidar la exposición de agentes químicos, los cuales pueden ocasionar daños a nivel genético potencializando

con ésto el desarrollo de cáncer, incrementando el riesgo a un mayor tiempo de exposición o bien cuando las sustancias tiene un mayor potencial mutagénico.

La edad es un factor importante que incrementa la probabilidad de desarrollar cáncer, debido a la pérdida de eficacia de los procesos de reparación celular sumada a la acumulación de factores como el tiempo de exposición hacia cancerígenos, en la Figura 1 se presenta la relación existente entre la edad y la incidencia, teniendo una relación proporcional con el incremento de la edad, es decir a mayor edad mayor es el riesgo de padecer cáncer.

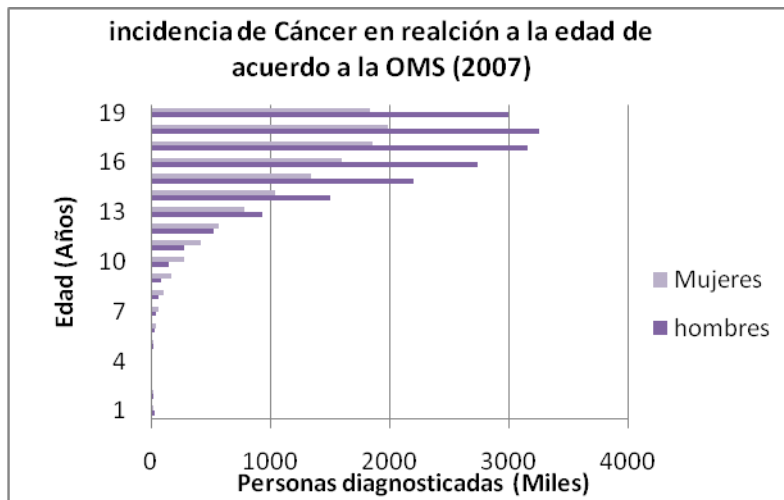


Figura 1. Desarrollo de diferentes tipos cáncer respecto a la edad. En la figura se observa la relación positiva de edad y desarrollo de cáncer en hombres y mujeres. (National Cancer Institute, www.nci.com).

2.2 Desarrollo de cáncer a nivel celular.

La carcinogénesis es un proceso multifactorial, durante el cual se da la transformación de células normales para la creación de una población clonética de células neoplásicas transformadas, que tienen como características proceder de una célula y “heredar” las anomalías de la progenitora. Estas anomalías se dan generalmente en la expresión de genes que están encargados de regular el crecimiento celular, genes represores y de diferenciación¹¹. La mayoría de estos cambios se dan en células somáticas y sólo un pequeño porcentaje corresponde a una mutación en las líneas germinales de los individuos. Ejemplo de genes mutados son: p53, RB1, BRC1, BCR2 ó INK4A entre otros, los cuales al encontrarse dañados permiten la proliferación

descontrolada de la célula a pesar de que ésta posea una gran cantidad de errores en su ADN replicado.

El proceso de carcinogénesis es similar en los diversos tipos de cáncer y se divide en tres etapas: iniciación, promoción y progresión (Figura 2), las cuales terminan en la generación de neoplasias (tumores) que son grandes masas de células no diferenciadas con alteraciones características de cada uno de los tejidos afectados. La iniciación es la etapa en la cual un daño en el ADN se hace permanente por un fallo en genes que se encargan de repararlos o por que la velocidad de duplicación no permite corregirlos. La promoción es el proceso en el cual influyen factores epigenéticos que coadyuvan en el desarrollo de células dañadas. La progresión implica la producción de cambios hereditarios más profundos los cuales aparecen a medida que la población celular se incrementa¹⁸.

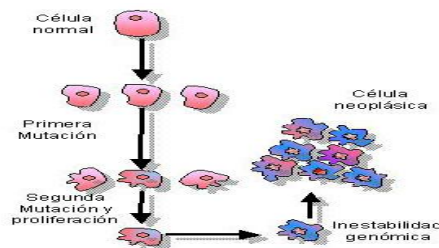


Figura 2. Etapas celulares de la carcinogénesis: iniciación, promoción y progresión. Estas etapas terminan en la generación de neoplasias (Ruin, 2004).

La iniciación requiere de factores externos y/o internos capaces de desencadenar una cascada de reacciones en las cuales se pueden involucrar tres clases de genes y sus productos proteicos; las primeras relacionadas con el crecimiento celular, por ejemplo: p53 que está encargada de garantizar que no exista daño alguno en la duplicación de ADN y así permitir el paso de la fase G1 a S en la mitosis de no ser así la célula dañada es llevada a mecanismos de apoptos. Otra clase de genes que intervienen son los oncogenes, que son protooncogenes en las células normales, los cuales han perdido heterocigosidad mediante una serie de rearreglos del genoma (Figura 3).

La progresión tumoral se genera por cambios genéticos adquiridos por células cancerosas individuales dando como resultado la aceleración de su crecimiento y la

selección secuencial de subpoblaciones celulares más agresivas a medida que se desarrolla el tumor, incrementando con esto la inestabilidad genética que se puede presentar en genes supresores de tumores, encargados de procesos apoptóticos y bloqueantes del ciclo celular, y en los oncogenes que son protooncogenes alterados, los cuales participan en la inhibición de apoptosis y favoreciendo procesos proliferativos.

2.2.1 Genes supresores de tumores.

Se encargan de la proliferación celular, son reguladores negativos del crecimiento celular por lo que cuando no están presentes en la célula o se encuentran inactivos éstas dejan de crecer normalmente y adquieren propiedades proliferativas anormales. La inactivación de estos genes se da por un evento denominado pérdida de heterocigocidad (LOH), la cual se ocasiona por mutaciones o por deleciones de ambos alelos. El ejemplo más representativo es el gen del retinoblastoma RB1, el cual se ha asociado con el retinoblastoma o bien el TP53 que se encuentra relacionado con diversos tipos de cánceres. Los genes supresores asociados al Cáncer de mama se presentan en la Tabla 3.

Tabla 3. Genes supresores asociados con Cáncer de mama. En la tabla se enlista los genes supresores asociados a Cáncer de mama y al función en la que participan (Ruin, 2002).

Genes supresores tumorales asociados con Cáncer de mama.	
Gen.	Función.
RB	Codifica para la proteína pRB, uno de los principales controles del ciclo celular.
P53	Codifica para la proteína p53, la cual puede detener la división celular e inducir a las células anormales a procesos apoptóticos.
BCRA1	Codifica para la proteína BCRA1, la cual regula el ciclo celular evitando la proliferación incontrolada, ya que se encarga de la reparación del ADN.
BCRA2	Codifica para la proteína BCRA2, la cual está involucrada en la reparación del daño cromosomal.

2.2.2 Oncogenes.

Son protooncogenes alterados que dirigen la producción de proteínas, como: ciclinas, factores de crecimiento, receptores, entre otros. Todos ellos involucrados en el crecimiento celular que mediante la estimulación de la proliferación celular son capaces de generar una multiplicación descontrolada de las células, de modo que algunos de ellos fuerzan a las células a que sintetizen de forma masiva determinados factores de crecimiento. El más representativo en el Cáncer de mama es ERBB2 que es un receptor de hereulina asociado al Factor del Crecimiento Epidérmico (EGF). La activación de estos oncogenes se puede dar por mutaciones cromosómicas o génicas (Figura 3). A nivel génico, el gen puede sufrir mutaciones que afecten los exones de modo que se altere la estructura de sus productos y con ello su actividad fisiológica o bien la mutación puede ocurrir sobre “enhancers” cercanos a la región del promotor, las cuales son regiones potenciadoras de transcripción, lo que ocasiona que el producto normal se expresa en cantidades anormalmente altas. A nivel cromosómico puede ocurrir un rearrreglo del cromosoma acercando una secuencia de ADN (exón) normalmente lejana al oncogen afectando así su tasa de expresión, la lectura del gen o produciendo supresiones o recombinaciones anormales derivadas de esta traslocación¹².

Tabla 4. Oncogenes relacionados con cáncer. La tabla nos presenta algunos oncogenes, su función y con el tipo de cáncer se relacionan (Ruin, 2002).

Oncogén	Función	Cáncer humano asociado
<i>SIS</i>	Factor de crecimiento derivado de plaquetas	
<i>ERBB1</i>	Receptor del factor de crecimiento epidérmico	
<i>FMS</i>	Receptor del CSF-1	
<i>KIT</i>	Receptor del factor de crecimiento de células madre	
<i>MET</i>	Receptor del factor de crecimiento hepático	Cáncer renal papilar
<i>HER2/NEU/ERBB2</i>	Receptor de hereulina	Cáncer de mama, de ovario, de vejiga, glioblastoma
<i>RET</i>	Receptor del factor neurotrófico	Cáncer medular de tiroides
<i>TRKA</i>	Receptor del factor de crecimiento neuronal	
<i>ERBA</i>	Receptor (nuclear) de hormona tiroidea	
<i>RAS (H, K, N)</i>	GTPasa	Cáncer de pulmón, de páncreas, de vejiga, colorrectal
<i>BCR-ABL</i>	Tirosinasa citoplásmica	Leucemia mieloide crónica
<i>RAF</i>	Tirosinasa citoplásmica	
<i>SRC</i>	Tirosinasa citoplásmica	
<i>FOS</i>	Factor de transcripción	
<i>JUN</i>	Factor de transcripción	
<i>MYC (c, L, N)</i>	Factor de transcripción	Linfoma de Burkitt, neuroblastoma, cáncer de pulmón
<i>BCL6</i>	Factor de transcripción	Linfoma de células grandes
<i>BCL2</i>	Anti-apoptosis	Cáncer de próstata, linfoma

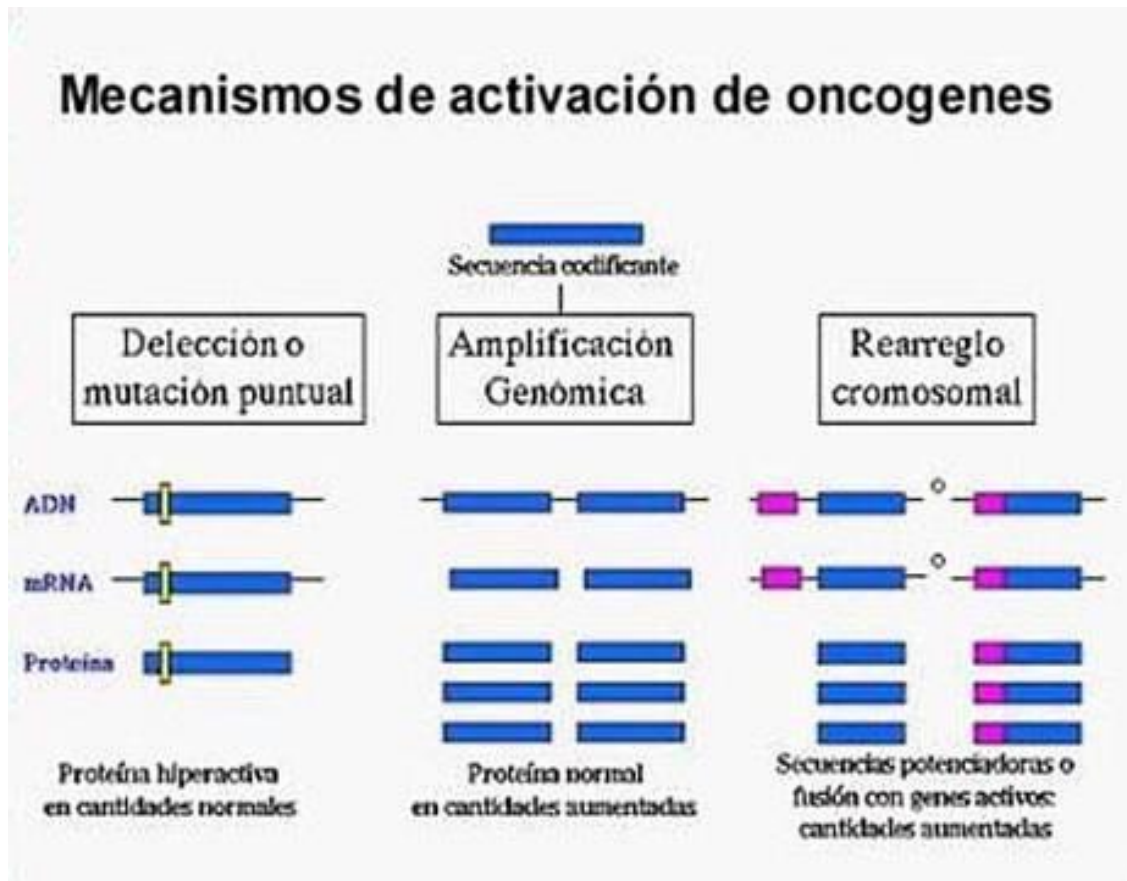


Figura 3. Mecanismos de activación de protooncogenes. La figura ejemplifica los tres tipos de modificaciones que pueden presentar los protooncogenes para activarse (Nelson y Cox, 2000).

Una vez presentes las mutaciones en los genes, el crecimiento de células cancerosas es descontrolado, el resultado es el incremento del número de mutaciones en el grupo celular, lo que hace posible la progresión o crecimiento tumoral, el cual de acuerdo al tejido que pertenece, posee características distintivas que son empleadas para su identificación y clasificación⁴, esta estatificación se revisará en los capítulos posteriores.

2.2.3 Propagación del Cáncer (Angiogénesis y Metástasis).

Durante el cáncer se incrementa notablemente la tasa de crecimiento celular por los factores ya descritos, la magnitud del incremento depende de tres características principalmente: la duración media del ciclo celular para las células en proliferación, la porción de células en crecimiento y la tasa de muerte celular. El crecimiento celular descontrolado requiere de una gran cantidad de nutrientes, motivo por el cual es necesario un proceso denominado "Angiogénesis", mismo que engloba la formación de

nuevos vasos sanguíneos que son capaces de abastecer de los nutrientes necesarios a las células tumorales. Este proceso es estimulado por varios factores de crecimiento, el principal de estos factores es el Factor de Crecimiento Endotelial Vascular o VEGF por sus iniciales en inglés (para el cual ya existe una terapia capaz de inhibirlo) que es estimulado por la hipoxia en el tumor y la producción de algunos oncogenes mutados (*RAS* y *Src*). Una vez que se generan los vasos sanguíneos necesarios para la obtención de nutrientes, el tumor es capaz de diseminarse, es decir de invadir tejidos circundantes, esta característica es indicativa de estadios más graves de cáncer y se le conoce como “Metástasis”, que es considerada como una de las últimas fases de la progresión. Ésta fase se compone de varias etapas: intravasación (salida del tumor primario), supervivencia de las células a la circulación, llegada a órganos periféricos, extravasación (salida de la circulación), migración hacia el espacio intersticial e inicio del crecimiento. Estas fases culminan con la angiogénesis de los vasos cercanos a la nueva localización del tumor agravando con esto el cáncer. La metástasis se ve influenciada por proteínas de fijación como las integrinas, cadherinas y las enzimas proteolíticas (activadores de plasminogénos, catepsinas). La expresión de éstas indica estadios graves en pacientes con cáncer, es por esto que la detección temprana es primordial para aumentar la expectativa de sobrevivida a la enfermedad de los pacientes.

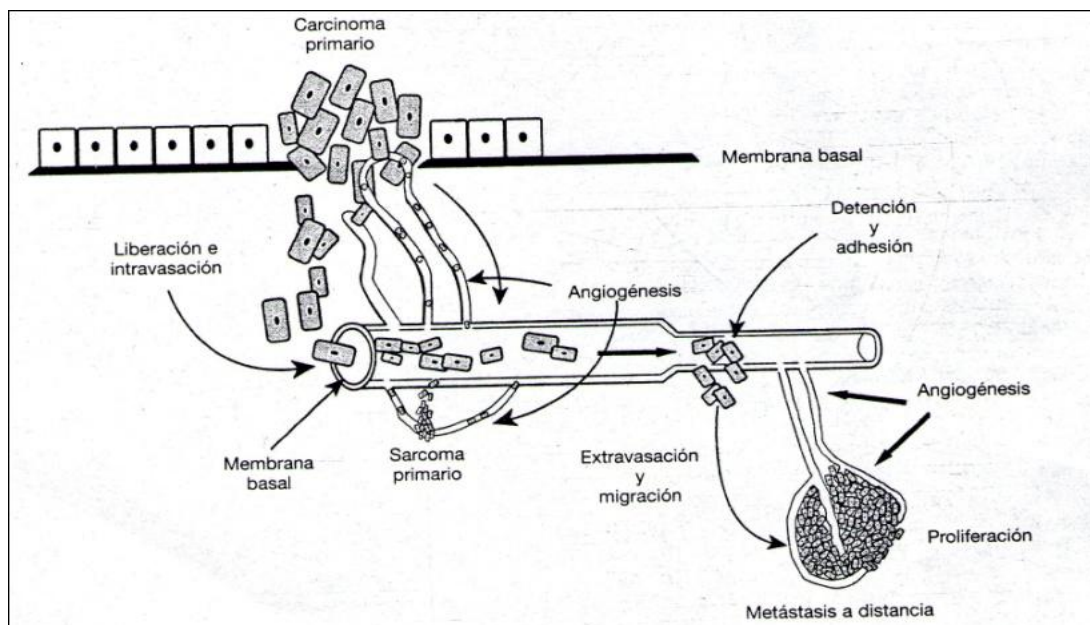


Figura 4. Proceso de diseminación del Cáncer. La figura nos presenta los procesos fundamentales para llevar a cabo la diseminación de cáncer: metástasis y angiogénesis (Ruin, 2002).

2.3 Clasificación del cáncer por la AJCC y UICC.

Una vez que se desarrolla el cáncer es necesario tener un criterio de clasificación que permita conocer la fase de desarrollo en que se encuentra, el daño que pueda causar y qué tipo de terapias son las adecuadas para su tratamiento. Esta clasificación empleada por médicos y profesionistas de la salud se basa en aspectos anatómicos e histológicos y es realizada por dos instituciones, la AJCC (American Joint Committee for Cancer) y la UICC (International Union Against Cancer) por sus siglas en inglés. De acuerdo a esto, los tumores pueden clasificarse de tres maneras distintas:

- Clasificación Histopatológica. Se basa en las características que presenta el tumor en sus distintas etapas, estas características son:
 - Histológicas. Localización del tumor (ovárico, testicular, mamario, etc.)
 - Grado de malignidad. Toma en cuenta el nivel de diferenciación del tejido y va de G1 hasta G4, siendo G1 bien diferenciado, G2 moderadamente diferenciados y G3 a G4 poco ha muy poco diferenciado, este último es el lugar en el que se encuentran generalmente los tumores.
 - Citogenética. Se basa en diferentes técnicas de la biología molecular y en la utilización de biomarcadores para definir el estadio en el cual se encuentra el desarrollo tumoral.
- Clasificación Anatómica por estadios. Este sistema es propuesto por Deniox y se basa en el sistema TNM (tumor, ganglios, metástasis) y busca la calificación del tumor por tamaño, diseminación hacia ganglios linfáticos y presencia de metástasis (Tabla 5).

Tabla 5. Clasificación TNM (tumor, ganglios y metástasis). Sistema aprobado por la la AJCC y la UICC para la clasificación de ls diferentes tipos de cánceres (National Cancer Institute, www.cancer.gov)

Característica	Clasificación	Características generales.
Tumor primario(T)	C	Sin evaluación.
	T0	No evidencia de tumor, no existe presencia de lesión, macro ni microscópica.
	Tis	Carcinoma <i>in situ</i> .
	T1	Lesiones superficiales, confinadas en el órgano afectado, móvil y Resecable.
	T2	Lesión localizada, infiltrante, comienza extensión profunda hacia tejido colindantes, pérdida de movilidad, no existe fijación.
	T3	Lesión avanzada, existe una mayor diseminación, comienza fijación entre tumor y tejido afectado, son resecables parcialmente para reducir su tamaño en algunos casos como el Cáncer de mama solo el tamaño son suficientes para clasificarse en este estadio.
	T4	Producción de fistulas (vícera) o sinus (órgano sólido), adherencia hacia otros órganos, difícil de aborar desde cualquier tratamiento terapéutico.
Región ganglionar (N).	NX	Sin evaluación.
	N0	No existe evidencia de cáncer en ganglios
	N1	Ganglios palpables, móviles hasta la primera estación de drenaje linfático. Tamaño $\leq$ 1 cm y $\leq$ 3 cm, solitario y encapsulados.
	N2	Mayores a 3 cm, múltiples, afectación a ganglios regionales, extensión extra capsular no adheridos.
	N3	Adheridos a estructuras vecinas, $\geq$ 6cm.
Metástasis distante.	MX	No valorado
	M0	No existe presencia de metástasis.
	M1	Presencia de metástasis.

- **Clasificación por etapas o estadio.** Si bien es necesaria la clasificación de la enfermedad tomando en cuenta los componentes ya descritos existen más de 40 combinaciones posibles, lo que dificulta definir el estado del paciente. Es por ésto que surge la clasificación por etapas, la cual toma la clasificación TMN como base, dando como resultado 5 grupos.

Tabla 6. Clasificación anatómica por estadios. En la tabla se exponen los 5 niveles de la calificación anatómica y alguna de sus características (National Cancer Intstute, www.cancer.gov).

Clasificación por etapas.	
Etapas o estadio	Definición
Etapas 0	Carcinoma <i>in situ</i> (cáncer inicial que está presente sólo en la capa de células en donde empezó).
Etapas I, Etapas II y Etapas III	Los números más altos indican enfermedad más extensa: tamaño mayor del tumor y/o diseminación del cáncer a ganglios linfáticos vecinos y/o a órganos adyacentes al tumor primario.
Etapas IV	El cáncer se ha diseminado a otro órgano.

2.3.1 Otros tipos de clasificación.

Una gran cantidad de cánceres se apegan a la clasificación TNM pero existen otros que se clasifican de formas diferentes, ejemplo de éstos son el cáncer de cerebro y médula espinal, linfomas, leucemias, esta última no posee clasificación para su estadificación. Es por esto que varios registros de cáncer utilizan una clasificación concisa la cual se distribuye en 5 grandes categorías principalmente:

- **Desconocido**, no se conocen datos suficientes para clasificarlos.
- **In situ**, está presente solo en el lugar donde se localizan las células progenitoras.
- **Localizado**, cáncer delimitado sólo al órgano de origen y no presenta evidencia de diseminación.
- **Regional**, cáncer diseminado a ganglios linfáticos y/o tejidos u órganos circundantes.
- **Distante**, cáncer diseminado a órganos distantes o a ganglios linfáticos distantes.

2.3.2 Métodos para la clasificación.

Para la clasificación de estos tumores y con ello la determinación de la gravedad de la enfermedad y la indicación del tratamiento adecuado, se utilizan diversos métodos de diagnóstico que nos proporcionen resultados confiables, éstos se dividen en dos grandes grupos:

- Estudios Clínicos-diagnóstico. Son métodos no invasivos, los cuales son recomendados en regiones de fácil acceso. Consisten principalmente en:
 - Exámenes físicos.
 - Métodos Radiográficos.
 - Pruebas de laboratorio.

- Estudios quirúrgico-patológicos. Son métodos invasivos, debido a que el lugar de localización del cáncer es de difícil acceso para la obtención de las muestras.

Una vez definido el estadio de la enfermedad el médico debe asignar el tratamiento adecuado con la finalidad de tener la menor afectación de la funcionalidad de los órganos o tejidos dañados y garantizar con esto una mejor calidad de vida. Si bien lo correcto es una terapia individualizada, la cual en los últimos años gracias a las herramientas de biología molecular no se ve tan lejana, la realidad es que no se practican, las terapias que se ocupan actualmente son menos agresivas hacia los pacientes y mucho más eficientes.

2.4 Terapias.

Hoy en día las terapias que se ofrecen para el tratamiento de cáncer son una combinación de varios tratamientos para elevar su eficiencia y garantizar la erradicación de la enfermedad. Estos tratamientos pueden ser:

- Radioterapias.
- Cirugías.
- Tratamiento biológico.
 - ✓ Inmunoterapia.
 - ✓ Viroterapia oncolítica.
 - ✓ Transferencia de genes.
- Quimioterapias.

La combinación de un diagnóstico oportuno y temprano mediante diversas técnicas como la autoexploración, identificación de biomarcadores (citogenéticas), extracción de biopsias, por citar algunas y una terapia adecuada al estadio en el cual se encuentre el paciente han mejorado el porcentaje de supervivencia y calidad de vida en las últimas décadas.

2.5 Cáncer de mama.

El Cáncer de mama es la neoplasia sólida más frecuente entre las mujeres europeas y latinoamericanas (Tabla 7) y se define como: “Un grupo de tumores malignos epiteliales los cuales se desarrollan en las glándulas mamarias y la consecuente metástasis de la región de los nódulos linfáticos y órganos periféricos y finalizando con la muerte del paciente”⁴.

El riesgo de padecer esta enfermedad es del 12% en mujeres, lo que significa que 1 de cada 8 mujeres es propensa a desarrollar Cáncer de mama y solo un 3.6% de esta población diagnosticada termina en defunciones es decir 1 de cada 28 pacientes. Se estimó que para este año existieran un total de 192, 370 mujeres diagnosticadas (dato estadístico no actualizado) con este tipo de cáncer, de las cuales 40,170 podrían terminar en defunciones, debido a un diagnóstico tardío o a un terapia ineficaz (National Cancer Institute). Si bien la cifras no corresponden a México éstas son un claro ejemplo del crecimiento de la población diagnosticada con esta enfermedad, es por estas cifras tan alarmantes que es necesario mejorar las técnicas de diagnóstico y aumentar la cultura de la autoexploración.

Tabla 7.- Incidencia del Cáncer de mama de acuerdo al origen étnico. En la tabla podemos apreciar que las mujeres caucásicas tienen el mayor porcentaje de incidencia y las Indias americanas el menor (National Cancer Institute, www.cancer.gov).

Incidencia por origen étnico.	
Raza/Origen étnico	Mujeres
Todas las Razas	123.8 por 100,000 mujeres.
Blancos.	127.8 por 100,000 mujeres.
Negros	117.7 por 100,000 mujeres.
Asia/Islas del pacífico.	89.5 por 100,000 mujeres.
Indios Americanos/ Nativos de Alaska	74.4 por 100,000 mujeres.
Hispanos.	88.3 por 100,000 mujeres.

La población más vulnerable a padecer esta enfermedad son las personas de más de 40 años, las cuales representan más del 85% de los casos totales diagnosticados entre el 2002-2006. El hecho de encontrar un porcentaje tan elevado en una pequeña población se asocia a la ineficacia que se va adquiriendo en mecanismos de reparación celular así como el tiempo de contacto con factores ambientales como carcinógenos físicos, químicos y biológicos (Tabla 8).

Tabla 8. Relación del desarrollo del Cáncer de mama respecto a la edad. En la tabla se observa un incremento significativo en el porcentaje de acuerdo al incremento de la edad (National Cancer Institute, www.cancer.gov).

Incidencia de Cáncer de mama 2002-2006.	
Edad.	Porcentaje de Incidencia
20 años.	0.0%
20-34 años.	1.9%
35-44 años.	10.5%
45-54 años.	22.5%
55-64 años.	23.7%
65-74 años.	19.6%
75-84 años.	16.2%
□ 85 años.	5.5%

2.5.1 Factores que influyen en el desarrollo del Cáncer de mama.

Además de la edad existen varios factores ambientales importantes que hacen que el riesgo de padecer Cáncer de mama se incremente. Este tipo de factores se dividen en dos grandes grupos:

➤ Clínicos.

- Género. Mayor incidencia en mujeres que hombres (casos raros).
- Edad.
- Antecedentes familiares de primer grado. Si el familiar de primera línea padeció cáncer en una edad temprana, el riesgo de padecer esta enfermedad se incrementa conforme avanza su edad.
- Antecedentes personales. Presencia de lesiones benignas o malignas.
- Presencia de Carcinoma lobular *in situ* (CLIS) ó hiperplasia atípica. Pacientes previamente diagnosticados con este tipo de daños son más propensas a desarrollar cáncer.

➤ Otros factores.

- Menopausia tardía.
- Edad creciente del primer parto.
- Uso de Anticonceptivos.
- Tratamiento sustitutivo hormonal (estrógenos).

Es por ésto que el Cáncer de mama es una enfermedad multifactorial, ya que no sólo se requiere de antecedentes genéticos, sino de la combinación de los factores ambientales. El riesgo se incrementa cuando una paciente cuenta con dos o más familiares afectadas, lo que fue observado por el Dr. M. C. King en la década de los 90's, dando como resultado la propuesta de los primeros dos marcadores moleculares de Cáncer de mama: *BCRA1* y *BCRA2*, ambos se identificaron en una *población* de pacientes jóvenes y hoy en día son utilizados para realizar pruebas que permitan conocer el riesgo de las pacientes cuando poseen familiares de primera línea diagnosticadas con esta enfermedad. Ambas proteínas son la razón por la cual se ha hecho un gran énfasis para localizar marcadores moleculares a los cuales se les denominan “biomarcadores” que pueden ser utilizados para el diagnóstico de pacientes asintomáticas, dando la posibilidad de detectarlo en estadios tempranos y lo cual, en combinación con una terapia adecuada, disminuirá las cifras de mujeres hospitalizadas y de las defunciones por este tipo de neoplasias.

2.5.2 Anatomía y fisiología de la mama.

Las glándulas mamarias son una característica distintiva de los mamíferos, las cuales han evolucionado con la finalidad amamantar a sus crías y con esto proporcionar los nutrientes necesarios hasta que sean capaces de alimentarse por sí mismos, además de aumentar un lazo afectivo entre cría y madre. Es por esto que en los mamíferos machos son rudimentarias. Las mamas alcanzan una madurez gracias a hormonas ováricas, la progesterona y estrógenos principalmente, las cuales se asocian al desarrollo de Cáncer de mama con la expresión de receptores a estrógenos (RE) y receptores a progesterona¹⁶ (RP), o bien con protección contra cáncer¹⁴. En una mujer adulta las mamas se localizan hacia abajo de la segunda o tercer costilla hasta la sexta o séptima (pliegue inframamario) y hacia afuera, desde el borde lateral del esternón hasta la línea axilar anterior o medioaxilar³ y consiste en 15 a 20 lóbulos de tejido glandular de tipo tubuloalveolar, los cuales están sostenidos por tejido conectivo fibroso, el tejido adiposo se encarga de rellenar los espacios entre cada lóbulo, todo esto se encuentra rodeado por tejido conectivo subcutáneo, la bolsa retromamaria es la encargada de proporcionarle movilidad³.

La mama se conforma de tres estructuras importantes en las cuales se desarrollan diversos tipos de neoplasias, estas son: pezón, aréola y glándulas mamarias. La primera revestida por un epitelio escamoso, queratinizado y estratificado, posee papilas dérmicas largas que permiten que los capilares aporten sangre cerca de la superficie, característica que proporciona la coloración, además de poseer fibras musculares lisas que permiten su erección, el pezón es innervado por la distribución dermatómica del nervio torácico T4. La aréola posee glándulas sebáceas, sudoríparas y areolares accesorias, estas últimas producen pequeñas elevaciones sobre la aréola, y las otras dos se localizan a lo largo del reborde areolar. La tercera estructura de la mama son las glándulas mamarias que se compone de 15 a 20 lóbulos irregulares de glándulas tuboalveolares, separadas por tejido conectivo que a su vez se dividen en lobulillos. Cada una de las glándulas termina en un conducto llamado conducto galactóforo. La innervación de las mamas está dada por estímulos de ramas anteriores y laterales de los nervios intercostales, cuatro a seis, provenientes de los nervios espinales. Todas éstas son innervadas por el sistema periférico autónomo simpático y son parte del componente sensorial.

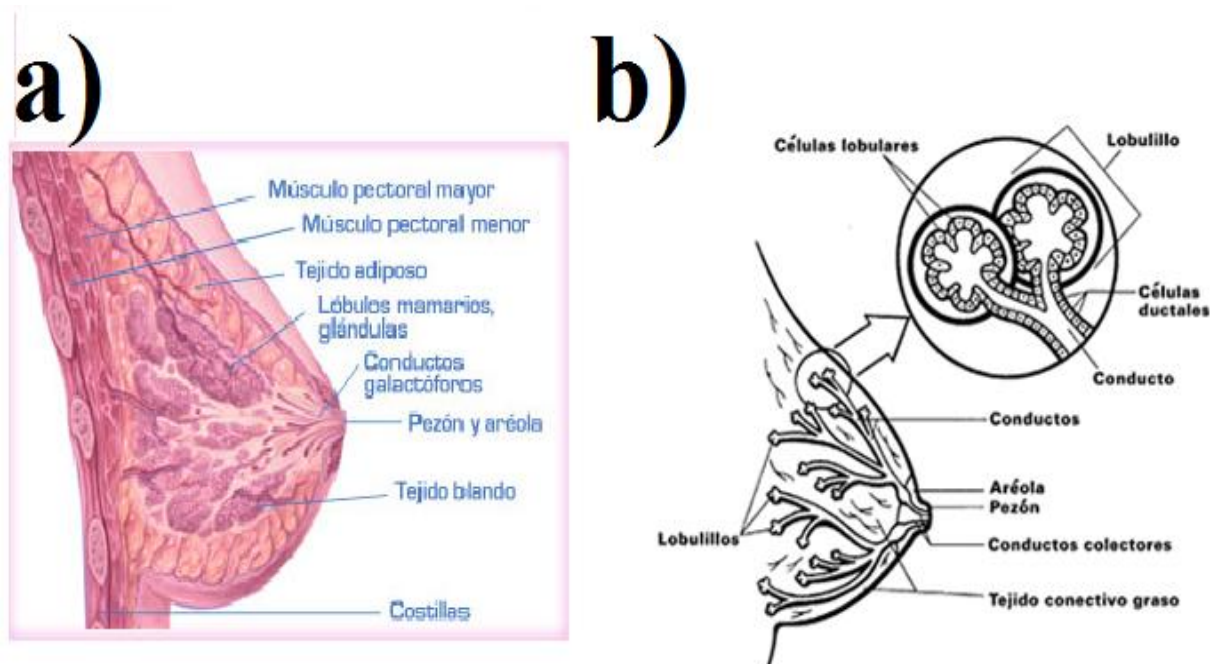


Figura 5. a) Estructuras principales de la mama: pezón, aréola y glándulas mamarias. b) Partes de la glándula mamaria, lóbulos, lobulillos y conductos (Ruin, 2002).

2.5.3 Tipos de Cáncer de mama.

Existen diversos tipos histológicos de carcinomas en la mama, un gran porcentaje de éstos se origina en las células epiteliales, las cuales revisten los ductos y lobulillos mamarios, mientras que los casos más extraños se originan en las células estromales las cuales rodean las glándulas epiteliales. Los tipos más frecuentes de cáncer en orden porcentual son:

- El carcinoma ductal *in situ* (DCIS), es el tipo más común de Cáncer de mama no invasivo por lo cual es considerado como una lesión benigna, se origina en las células de las paredes de los conductos mamarios, es localizado, es decir, solo se desarrolla en los conductos mamarios y predispone a la paciente a padecer lesiones de mayor gravedad de no ser tratado adecuadamente. Se ha observado que el 30% de las pacientes que sufrieron de esta neoplasia son propensas a desarrollarla nuevamente cuando sólo fueron tratadas con quimioterapias, este porcentaje disminuye si se realiza resección del tejido hasta en un 50%, si no presenta metástasis y por consecuencia no se ha extendido. Es curable en el 100% de los casos, el método de detección para el DCIS es la mamografía.
- El carcinoma ductal infiltrante (IDC). Esta neoplasia comienza en los ductos mamarios, es capaz atravesar la pared de éstos y así diseminarse a los tejidos cercanos mediante la metástasis. El IDC representa cerca del 80% de los casos diagnosticados con cáncer invasivo y es diagnosticado mediante la utilización de exámenes físicos como la autoexploración y pruebas de laboratorio, pruebas de imágenes como: mamografía, ultrasonidos, IMR, y biopsias. La característica principal de este tipo de cáncer es la disminución de Receptores de estrógenos (RE), Receptores de progesterona (RP) y CD44s¹² con respecto al DCIS, características empleadas para definir el estadio de la enfermedad, la terapia utilizada en este tipo de neoplasias son la radiación, quimioterapias, terapias hormonales y tratamientos dirigidos.
- El carcinoma lobular *in situ* (LCIS). Es un área que posee células con un crecimiento anormal, el cual incrementa el riesgo de padecer cáncer de tipo invasivo. Este tipo de neoplasias al igual que DCIS es considerada como una

neoplasia benigna, la cual se desarrolla en los lóbulos y glándulas productoras de leche. Esta se presenta en mujeres premenopáusicas, no ocasionan sintomatología, son Receptores de estrógenos (RE) y Receptores de progesterona (RP) positivos, por lo tanto, además de la mamografía en la cual se deben observar pequeñas calcificaciones y las biopsias, una buena alternativa para su diagnóstico son la detección de estos receptores. Es altamente recomendable revisiones anuales una vez diagnosticado, además de autoexploración y revisiones periódicas, las pacientes con LCIS generalmente no son sometidas a ninguna terapia pero se puede utilizar Tamixofen o bien Raloxifeno los cuales son una terapia hormonal que inhibe a las aromatasas que son encargadas de la síntesis de estrógenos a partir de andrógenos androstenediona o testosterona. La importancia de inhibir esta síntesis es detener la interacción receptor- estrógeno/progesterona, ya que intervienen en la síntesis de RNA y proteínas dando como resultado metástasis y neoplasias dañinas.

- El carcinoma lobular infiltrante o invasivo (ILC). Comienza en los lóbulos de las glándulas mamarias y es capaz de romper las paredes de éstos y diseminarse al tejido mamario circundante, después hacia los blancos cercanos a la mama hasta llegar a los nódulos linfáticos. Este tipo de neoplasia representa del 10 al 15% de los casos diagnosticados con cáncer invasivo, la característica principal de este tipo de cáncer es la sobreexpresión de una proteína llamada HerbB2 hasta en un 15% en pacientes con ILC. HerbB2 es una proteína de membrana unida a la superficie de receptores de tirosin cinasa e interviene en el crecimiento y la diferenciación celular¹³.

El ILC se diagnostica mediante exámenes físicos que buscan pequeñas protuberancias en la mama, además de enrojecimiento y endurecimiento; también se utilizan mamografías, biopsias, resonancia magnética y ultrasonidos. Una vez diagnosticado, el ILC puede ser tratado mediante quimioterapias basadas en taxoles, los cuales inhiben la formación de microtúbulos, deteniendo con ésto la mitosis en un lapso de 3 a 6 meses. La terapia hormonal (SERMs, inhibidores de aromatasas) es utilizada en este tipo de cáncer y busca intervenir en la diferenciación y crecimiento celular, otro tipo de terapia es la dirigida contra HerbB2 en donde se utiliza Trastuzumab© contra el complejo de la tirosin cinasa.

- El Cáncer de mama inflamatorio (IBTC). Es el tipo de cáncer menos común que se puede presentar en la mama, éste representa entre el 1-5% de los casos totales. Las características principales del IBC es su agresividad y la rapidez con la que se desarrolla, además de expresar HerB2, RE y RP, los cuales como ya se ha mencionado intervienen en el crecimiento y diferenciación celular, y que sirve como método de diagnóstico. Este tipo de cáncer suele presentar enrojecimiento e hinchazón, características que se acreditan al bloqueo de vasos linfáticos. Para su diagnóstico se utilizan: rayos X de pecho, tomografía computarizada, escáner de hueso, Test de Función de hígado y Test Herb2, RE y RP.

El IBTC se puede dividir en tres estadios de acuerdo a su nivel de diseminación en IIIB, IIIC y IV siendo el último el de mayor gravedad, el tratamiento de este tipo de cáncer se basa en la quimioterapias que varía de acuerdo a su desarrollo, por ejemplo en etapas tempranas se utilizan las antraciclinas (dexorubicin, epirubicin, etc.) combinadas con taxoles (docetaxel, paclitaxel, etc) y la terapia dirigida (Trastuzumab©), que son anticuerpos monoclonales que bloquean los receptores expresados en membrana Herb2. Si la terapia no es efectiva se utiliza radiación o bien una solución más drástica como la mastectomía.

- Enfermedad del Pezón de Paget. Es un adenocarcinoma intraepidérmico, es decir que afecta la mucosa y la dermis de las glándulas mamaria, el cual se caracteriza por la presencia de eritema y segregaciones en el área del pezón y la areola. Está enfermedad se presenta en el 5% de las pacientes diagnosticadas con CDIS. Se cree que el Pezón de Paget es debido a la migración de células dañadas de los ductos y lóbulos hacia la mucosa de las glándulas mamarias, como característica citológica principal es su elevada presencia de Herb2.¹⁶

2.5.4 Diagnóstico en Cáncer de mama.

En la actualidad se ha dado un gran énfasis en la autoexploración como terapia preventiva y diagnóstico para el Cáncer de mama, la cual es confirmada con estudios como la mastografía. De acuerdo al The American College of Surgeons, la autoexploración en la cual se localizan pequeños bultos representa cerca del 42% de un prediagnóstico, de este porcentaje el 24% es diagnosticado con cáncer y el restante con neoplasias benignas. De tenerse sospechas de presencia de bultos se recurre a los siguientes criterios para discriminar masas benignas de las malignas:

- 1) Síntomas. El Cáncer de mama suele ser asintomático y sin presencia de dolor. Neoplasias benignas (nódulos mamarios funcionales) se asocian a dolor e inflamación.
- 2) Edad. La presencia de Cáncer de mama se da principalmente en mujeres mayores de 30 años. En mujeres jóvenes, la presencia de nódulos se asocia o un fibroadenoma, el cual es una protuberancia benigna, firme o móvil.
- 3) Relación nódulo-tejido adyacente. La presencia de un nódulo móvil el cual esté libre de tejido mamario suele corresponder a un fibroadenoma.

Los criterios mencionados no pueden indicarnos por sí solos la presencia o ausencia de cáncer, pero son una guía para indicar los métodos diagnósticos más adecuados que sirvan para descartar o confirmar la presencia de cáncer, éstos pueden ser:

- Estudios Clínicos-Diagnóstico. Son métodos no invasivos los cuales son recomendados en regiones de fácil acceso y consisten principalmente en:
 - ✓ Exámenes físicos. Principalmente la autoexploración y estudios exploratorios por parte del médico. Este método nos proporciona información sobre el tamaño, lugar y diseminación del cáncer.
 - ✓ Métodos Radiográficos. Nos proporcionan imágenes internas del paciente, para el Cáncer de mama los más comunes son: mamografía la cual detecta entre el 85% y 90% de los casos, además de la mamografía existen técnicas como la ecografía, resonancia magnética nuclear (RMN). Estas pruebas nos proporcionan información acerca del sitio, tamaño y diseminación del cáncer.

- ✓ Pruebas de laboratorio. Son pruebas histológicas en diferentes fluidos corporales como sangre u orina, en donde se buscan sustancias que reflejen algún cambio metabólico que se presenta en pacientes enfermos. Ejemplos de éstos son los biomarcadores de los que hablaremos en el capítulo 7, en los cuales se ha puesto un gran énfasis en las últimas décadas, ya que presentan grandes ventajas como la facilidad de acceder a las muestras y que podrían proporcionar información concisa del estadio en el cual se encuentra el paciente.
- Estudios quirúrgico-patológicos. Son métodos invasivos debido a que el lugar de localización del cáncer es de difícil acceso para la obtención de las muestras, son necesarias las biopsias y otros métodos quirúrgicos. En este tipo de metodologías se dan reportes citológicos y patológicos como: tamaño, diseminación, tipos de células, grado (similitud entre células sanas y cancerígenas), tipo de células localizadas en fluidos corporales. Además es necesario que el médico reporte todo lo ocurrido durante la cirugía en donde describirá el tamaño, apariencia, información sobre ganglios linfáticos y órganos vecinos.

2.5.5 Pronóstico en Cáncer de mama.

La estadificación del Cáncer de mama comenzó a principios del siglo XX, con la finalidad de entenderlo y así poder asignar un tratamiento adecuado a la etapa de desarrollo, este sistema se basó en la diseminación hacia ganglios linfáticos y la metástasis a distancia y conforme se han descubierto nuevos factores, se han ido incorporando. En la actualidad los factores pronósticos utilizados para el cáncer de mama más empleados son:

- Ganglios linfáticos axilares. La positividad de los ganglios es un factor de riesgo de enfermedad sistémica interior en los pacientes. La presencia de ganglios positivos indican metástasis en la neoplasia mamaria, esto nos indica que se han afectado otros tipos de tejido.
- Tamaño del tumor. Este factor de pronóstico se combina con la afectación de ganglios y la presencia de metástasis dando como resultado 5 posibles estadios, 0, I, II, III y IV (Tabla 6), que indican gravedad de la enfermedad

en orden progresivo. Aquellas pacientes diagnosticadas en los estadios III y IV tienen un pronóstico desfavorable con respecto a las de estadios 0, I y II debido a las características y tamaño que presentan los diversos tumores.

- Sistema de gradación histológica. Basado en el patrón de crecimiento y morfología de las células cancerosas respecto a las normales y la diseminación del tumor. Utiliza dos criterios: el primero es histológico y compara a las células cancerosas con respecto a las normales del mismo tipo de tejido. El segundo es nuclear, compara tamaño y forma del núcleo de las células y porcentaje en que se encuentran activas es decir, aquellas que se encuentran en división mitótica. Para el Cáncer de mama la escala de estatificación es el grado histológico de Nottingham.

Tabla 9. Gradación histológica de acuerdo a Nottingham. Esta clasificación se divide en 5 grados la tabla nos indican las características mas representativa de cada uno (National Cancer Institute, www.cancer.gov).

Grado.	Características.
GX	No es posible asignar un grado (Grado indeterminado)
G1	Bien diferenciado (Grado bajo)
G2	Moderadamente diferenciado (Grado intermedio)
G3	Mal diferenciado (Grado alto)
G4	Indiferenciado (Grado alto)

- Edad. El Cáncer de mama se presenta generalmente en pacientes mayores a los 30 años. En pacientes que presentan la enfermedad a una edad temprana se desarrollan tipos de cáncer más agresivos con un pronóstico desfavorable en su tratamiento.
- Receptores esteroides. Principalmente RE y RP, los cuales se han utilizado como factores pronósticos y para un diagnóstico temprano. La ausencia de RE y RP en los primeros 5 años de la enfermedad se relaciona con un desarrollo menos agresivo respecto a los pacientes con RE y RP positivos; una vez transcurrido este tiempo no se ha encontrado correlación con la

agresividad de los tumores. La sola determinación de los receptores no es suficiente para asignar una terapia adecuada ni para dar un pronóstico confiable. Por lo que debe de combinarse con los demás factores pronósticos mencionados.

- **Conteo mitótico.** La elevada tasa de crecimiento celular durante los diversos tipos de cáncer trae consigo una elevada actividad mitótica en el lugar en donde se localiza la neoplasia. Un conteo mitótico elevado indica un desarrollo rápido de la neoplasia.
- **Expresión de ErbB2.** La expresión de muchos factores de crecimiento (GF) y sus receptores se ha relacionado en cánceres de mama primario. La expresión de ErbB2 trae consigo un pronóstico negativo e indica tipos de cánceres resistentes a la terapia hormonal, este gen se encuentra sobreexpresado en un 20% por amplificación genética en pacientes con cánceres en estadios graves.
- **Catepsina D.** Es una proteasa lisosomal ácida secretada por las células afectadas, la cual se ha relacionado con actividades promotoras del crecimiento y proteolíticas extracelulares, la sobreexpresión de esta proteína se vincula con pronósticos desfavorables.
- **Receptor de laminina y colagenasa tipo IV.** La detección de ambas proteínas por técnicas como la inmunohistoquímica en el tejido afectado se correlaciona con un elevado potencial metastásico en Cáncer de mama, indicando así fases finales de la neoplasia y asociándolos con un pronóstico negativo.
- **uPA.** La producción elevada de enzimas proteolíticas que favorecen la metástasis es una característica en etapas tardías de cáncer de mama dando un pronóstico negativo en pacientes que la expresan. La uPa es un activador de plasminogeno del tipo urocinasa y se ha asociado a la

remodelación tisular y migración celular.

- Angiogénesis. Es un proceso empleado por el tumor para la obtención de nutrientes asociado con el grado de malignidad. A una mayor actividad angiogénica el pronóstico es negativo, es por esto que hoy en día existen fármacos que tienen como blanco terapéutico disminuir esta actividad (Bevacizumab©).

2.5.6 Terapias en Cáncer de mama

Una vez que la paciente es diagnosticada, la expectativa y calidad de vida dependen del tratamiento y el seguimiento del mismo. Durante muchos años y aún hoy en día el tratamiento recomendado para tratar el Cáncer de mama, son los inhibidores de angiogénesis y los agentes citotóxicos, ya que ambos tienen un efecto en la disminución de la metástasis en el tejido afectado. Estudios recientes indican una relación positiva entre la disminución del riesgo de padecer Cáncer de mama con altos niveles de vitamina D₃ debido a que esta molécula y su receptor el VDR (Vitamin D receptor) han mostrado un comportamiento de reguladores negativos en el crecimiento celular ⁷.

Debido al creciente número de pacientes con esta enfermedad, en los últimos años se ha realizado un gran esfuerzo por conseguir terapias alternas que coadyuven con las ya existentes, además de garantizar la seguridad del paciente y proporcionen mejores resultados. Las terapias más utilizadas son las siguientes:

- Radioterapias. Consiste en utilizar radiación ionizante, la cual se puede aplicar de forma externa (sobre la piel del lugar afectado) o bien mediante la implantación de un dispositivo bajo la piel (intradérmico). Este tratamiento tiene como finalidad la destrucción de células dañadas y es utilizado en conjunto con la cirugía para incrementar su eficiencia. Una gran desventaja es la pobre selectividad hacia células dañadas, causando diversos efectos adversos de entre los cuales destaca la inmunosupresión.
- Cirugías. Para la mayor parte de los profesionistas de la salud, la resección de la masa tumoral o tejido afectado ya sea parcial o total, es la terapia más adecuada. Hoy en día se utilizan dos técnicas: MRM (Mastectomía Radical Modificada), la cual incluye una resección total de la mama, en ocasiones el músculo pectoral y

los niveles de ganglios axilares 1 y 2, y el TMC (Tratamiento Conservador de la mama) que consiste en la resección completa de la tumoración macroscópica, ambas cirugías van acompañadas de un tratamiento con fármacos (quimioterapias) para garantizar su eficacia. En años recientes se ha tratado de reducir la utilización de las cirugías y tomarla como última alternativa debido al gran impacto psicológico que puede llegar a causar al paciente, a pesar de ser una de las terapias más efectivas existentes hoy en día.

- **Terapia génica.** Se basa en la inserción de genes capaces de corregir daños en el ADN o de inhibir alguna vía metabólica que desencadenen la muerte de las células dañadas (Tabla 10), para esto se utilizan varias clases de vectores virales y no virales: Adenovirus, Retrovirus, Poxvirus, Virus del Herpes Simple, Lentivirus entre otros.

Tabla 10. Principales clases de terapias génicas empleadas en el tratamiento de Cáncer de mama (De Vita, 2004)

Terapia.	Objetivos.	Ejemplos.
Mutación compensada.	Corrección de defectos que generan fenotipos malignos, por ejemplos genes involucrados en el crecimiento o que inducen apoptosis.	Reemplazamiento de p53 y BCRA1 en células tumorales de Cáncer de mama.
Quimioterapia molecular o genes suicidas.	Transformación de un profármaco a una molécula capaz de destruir células cancerosas.	Gen de la cinasa timidina mediante el virus del Herpes Simple.
Oncolisis viral.	Infección de virus y con esto la introducción de genes capaces de producir lisis de las células.	No existen ejemplos aprobados.
Inmunoterapia génica.	Estimula en sistema inmunológico para que pueda reconocer antígenos expresados por el tumor y con esto lo lleve a la destrucción.	No existen ejemplos aprobados.
Terapia angiogénica de genes	Interrupción de angiogénesis con la finalidad de detener el suministro de nutrientes del tumor.	Inhibición del factor de crecimiento vascular-endotelial (VEGF).

- **Quimioterapia.** Es un tratamiento basado en una serie de fármacos aprobados por la FDA⁹, los cuales buscan la disminución de la metástasis mediante la destrucción de células neoplásicas. Esta terapia se divide en: terapia adyuvante y terapia metastásica, ambas utilizadas después de la resección de tejido dañado. La acción de los quimioterapéuticos se dirige en su totalidad a frenar la proliferación y/o el

crecimiento celular. Para esto, actúan sobre el ADN, el ARN o la división mitótica, y sólo excepcionalmente en inhibir la síntesis de proteínas.

La quimioterapia adyuvante es empleada en etapas tempranas y tumores pequeños que en ocasiones son benignos. Se emplea con la finalidad de prevenir el desarrollo de enfermedades metastásicas, las ventajas que ofrece esta terapia es la reducción de posible recurrencia y muerte de pacientes con cáncer en estadios tempranos. La quimioterapia metastásica se utiliza en pacientes con estadios avanzados y que tienen como característica principal la presencia de metástasis y angiogénesis, esta terapia es asignada mediante estudios clínico-diagnósticos y/o quirúrgico-patológicos que son capaces de indicarnos el estadio de esta enfermedad, los fármacos más empleados en ambas terapias son:

- Antibióticos antitumorales. Principalmente las antraciclinas que son fármacos obtenidos de *Streptomyces spp*, su mecanismo de acción es la intercalación en el ADN y la inhibición de la Topoisomerasa II, causando la muerte de las células cancerosas. Las más utilizada en el Cáncer de mama son: Doxorubicin, Epirubicin, Mitoxantrone y Mitomicina-C. Los principales efectos adversos son la mielosupresión y una deficiencia de neutrófilos, náuseas, vómito, diarrea, y pérdida de cabello.
- Inhibidores de síntesis de microtúbulos. Se dividen en dos grupos el primero son los alcaloides de la vinca los que son una familia de fármacos derivada de la planta del bígaro que inhiben la síntesis de microtúbulos, los representantes de este grupo son: vinorelbin, vinblastin. Cada uno ellos posee efectos secundarios distintos pero comparten dos: la mielosupresión y neurotoxicidad, ambos empleados en combinación con otros fármacos debido a que su efecto terapéutico es pobre. El segundo grupo son los taxanos, que son diterpenos obtenidos de plantas o de forma sintética, inhiben el desensamblaje del citoesqueleto gracias a la unión a los dímeros de tubulina, el resultado es el bloqueo de G2/M ya que no se desarrolla adecuadamente el huso mitótico. Este grupo además, posee propiedades antiangiogénicas originando la inhibición de la proliferación endotelial, migración y formación de túbulos, llevando a la muerte las células

tumorales. El paclitaxel y docetaxel son los dos taxanes más empleados en el tratamiento contra el Cáncer de mama, sus efectos secundarios más comunes son la mielosupresión, neuropatías periféricas y retención de líquidos corporales.

- Antimetabolitos. Son sustancias que bloquean el crecimiento celular al interferir con la síntesis del ADN en la etapa “S” del ciclo celular, simulando sustancias que participan en la síntesis e inhibiendo la producción de moléculas. Los más empleados para el tratamiento de Cáncer de mama son: 5-fluorouracil, metotrexato, capecitabina y gemcitabina.
- Agentes alquilantes y compuestos de platino. Son fármacos que añaden grupos alquilo cargados negativamente, generando sitios nucleofílicos ricos en electrones, formando así aductos los cuales alteran la estructura y función del ADN. Estos pueden operar en cualquier momento del ciclo celular, sin embargo son más eficaces durante la síntesis de ADN que es llevada a cabo en la fase S del ciclo celular, los más utilizados son: ciclofosfamida, tiotepa e ifosfamida. Existen compuestos de platino los cuales tienen el mismo mecanismo de acción y son: cisplatino y carboplatino.
- Terapia hormonal. Consiste en la administración de fármacos que bloquean la acción de receptores y sus hormonas asociadas a una disminución en su positividad de acuerdo al grado de malignidad. La terapia hormonal sólo es utilizada en pacientes que tienen receptores hormonales positivos, que representa entre un 60% al 70% de las pacientes diagnosticadas. Esta terapia posee las siguientes familias de fármacos:
 - Inhibidores de aromatasas. Durante el Cáncer de mama, la función de los ovarios disminuye y con esto la síntesis de estrógenos en sitios extragonadales y se incrementa en otros sitios a causa de la aromatización de precursores androgénicos de origen adrenal. Como principal reacción se lleva a cabo la conversión de la androstenediona o estrona y consecuentemente a estradiol, el cual estimula la producción de hormonas que incrementan el crecimiento de células tumorales, el uso de estos

inhibidores disminuye la producción de estradiol y con esto reprime el crecimiento celular, ejemplos de este tipo de inhibidores son: los inhibidores no esteroideos de triazol (anastrozol, letrozol y fadrozol) y los inhibidores esteroideos de aromatasas (formestano y exemestano).

- Antiestrógenos. Compuestos que antagonizan la acción de los estrógenos compitiendo por su receptor, debido a que los estrógenos están implicados en la proliferación y diferenciación de las células y son factores de riesgo para el desarrollo de Cáncer de mama. Los antiestrógenos se clasifican en tipo I o parciales (agonista/antagonista) que son de naturaleza no-esteroidal, presentan propiedades agonistas y antagonistas y también son denominados Moduladores Selectivos del Receptor de Estrógenos (MSRE), porque pueden actuar de manera diferente dependiendo del órgano blanco, se unen a los RE y permite su dimerización. Se unen al complejo hormona- receptor a elementos de unión a receptor de estrógenos (ERE) inhibiendo la actividad del factor de transcripción TAF-2. Los más empleados para el tratamiento de Cáncer de mama son: tamoxifen, toremifeno, clomifeno, droloxifeno, benztiofenos y raloxifeno, por otra parte los de Tipo II o antiestrógenos puros que son de naturaleza esteroideal que inhiben la acción de los estrógenos gracias a una elevada afinidad por el RE, además de inhibir en la dimerización de los mismos, lo cual no permite la unión al ERE, y como consecuencia impide a la transcripción de los genes. El ICI 182 780 (faslodex) es el único que ha sido probado en la clínica.
- Análogos de Gh-Rh. Son polipéptidos que inhiben el efecto pulsátil de la GnRH sobre el hipotálamo dando como resultado la anulación de la función ovárica o testicular los más empleados son: goselerin, leuprolin, buselerin y triptolerin.

2.6 Estructura, formación, clasificación y estudio de proteínas.

Una célula es capaz expresar hasta 5000 proteínas en un lapso de tiempo y espacio determinado, a lo cual se le conoce como proteoma. Éste se puede modificar de acuerdo a las condiciones a las que se exponga el organismo. La Proteómica se encarga de estudiar este tipo de cambios en el proteoma de los diversos organismos, además de tratar de explicar las propiedades y funciones de estas proteínas mediante diversas técnicas como la espectrometría de masas. Recordemos que una proteína es “Una macromolécula que lleva a cabo todas las actividades celulares”¹⁹, las cuales reflejan la información contenida en los genes. Estas macromoléculas se componen de varios aminoácidos unidos mediante un enlace peptídico que se forma por la unión del extremo carboxilo (-COOH) de un aminoácido y el extremo amino (NH₂) del aminoácido siguiente, partes que junto con el carbono α y una cadena R conforman un aminoácido. La última otorga las propiedades características de cada uno de ellos como: carga, polaridad, tamaño, entre otros (Figura 6).

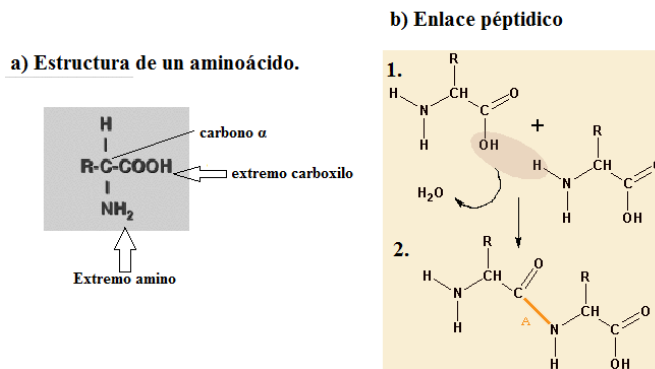


Figura 6. a) Estructura de una aminoácido, b) Formación de un enlace peptídico, 1) unión de extremo amino y carboxilo con la pérdida de una molécula de agua, 2).Formación del enlace (Nelson y Cox, 2000).

Existe un gran número de aminoácidos que pueden conformar una proteína, pero sólo 20 son necesarios para la formación de estructuras que intervienen en procesos celulares del humano, con este número de aminoácidos se pueden generar miles de proteínas estructurales o funcionales, ésto gracias a sus enlaces intramoleculares e intermoleculares. Una proteína que posea alguna función se compone generalmente de al menos 100 a 200 aminoácidos unidos por un enlace covalente (peptídico) y estabilizado por enlaces no covalentes como: puentes de hidrógeno, fuerzas de Van der Waals, fuerzas

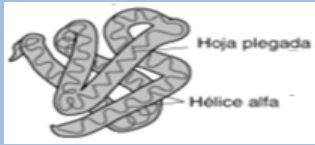
iónicas, etc. Estas interacciones nos dan como resultado 4 diferentes tipos de estructuras: primaria, secundaria, terciaria y cuaternaria, las propiedades de cada una de ellas las observamos en la Tabla 11. Se clasifican por su estructura, su composición o bien su valor nutricional para su estudio, estos grupos pueden sufrir una o más modificaciones postraduccionales, como el plegamiento, o la adición de algún grupo prostético, estos cambios en general tienen como finalidad alterar la estructura y con ello modificar su actividad biológica o bien permitirle la incorporación hacia la estructura celular de la que forman parte (matriz, pared celular, membranas, etc). La adición de algún grupo prostético en una cadena polipeptídica es una de las modificaciones postraduccionales de entre las que se destacan:

- Acilación
- Fosforilación
- Metilación
- Hidroxilación
- Glucosilación
- Sulfonilación
- Prenilación
- Glicación
- Nitrosilación
- Nitración

El enlace entre proteínas es otra modificación importante, ya que se emplean por la célula como señal para iniciar la degradación de una cadena polipeptídica cuando ya se tiene la cantidad necesaria o bien cuando la cadena tiene algún problema en su estructura, los mecanismos más conocidos son la SUMOlación y la Ubiquitinación.

Se han asociado varias de estas modificaciones a diferentes tipos de cáncer, la fosforilación y glucosilación son las dos modificaciones más importantes, ya que entre las dos son capaces de cambiar más del 60% de proteoma de un organismo^{6,7}. Muchas de las proteínas importantes para el metabolismo celular requieren de una ser fosforiladas y representa más del 12% del total del proteoma de un organismo, los ejemplos más representativos de este grupo son las cinasas y las CDks, las cuales son encargadas de regular una gran variedad de vías metabólicas. En el caso de la glicosilación, más del 50% del proteoma de un organismo es susceptible a esta modificación en particular, las más estudiadas son las proteínas de membrana. Las proteínas más representativas de este grupo son las mucinas.

Tabla 11. Niveles estructurales de las proteínas. La tabla nos presenta la clasificación de proteínas de acuerdo a su estructura y sus características principales (Nelson y Cox, 2000).

Estructura.	Propiedades y características.	
Primaria.	Esqueleto de una proteína. Cadena lineal de al menos 2 aminoácidos, unidos por enlaces peptídicos.	
Secundaria.	Estructura tridimensional de una cadena polipeptídica. Estabilizada por puentes de hidrógeno. 2 tipos : • α- hélice. Su conformación se asemeja a un espiral, 3.6 aminoácidos por cada giro, aminoácidos incapaces de formar esta estructura: Asp, Thr, Ile, Pro y Gly debido a sus propiedades físicoquímicas. • β-plegada u Hoja β. Conformación plisada es decir en forma de una hoja la cual compone por varios segmentos de perpendiculares a la lámina uno tras otro. Con una distancia de 3.5 Å se forman puentes de hidrógeno entre cada grupo carboxilo y amino de los segmentos, la hoja puede ser antiparalela o paralela dependiendo de la dirección de los segmentos. • Giro β. Formado por 4 aminoácidos, estabilizado por un puente de hidrógeno entre el carbonilo del primer aminoácido y el amino del cuarto, el tercer aminoácido suele ser glicina o prolina, ya que son capaces de cambiar de dirección por el impedimento estérico.	
Terciaria.	Formada por una o varias estructuras secundarias. Es la conformación activa de una proteína, tridimensional, estabilizada por uniones no covalentes, fuerzas de Van der Waals, puentes disulfuro, interacción con iones metálicos, etc. Existen dos tipos principalmente las proteína fibrosas (forma elongada): las cuales conforman principalmente proteínas estructurales como el colágeno, elastina, queratina, y las proteínas globulares (forma compacta).	
Cuaternaria.	Compuesta por dos o más estructuras terciarias estabilizadas por puentes disulfuro, pueden ser homodímeros (subunidades iguales) o heterodímeros (subunidades distintas).	

2.6.1 Fosforilación.

Una de las modificaciones más comunes de proteínas es la fosforilación en sus residuos de serina (Ser), treonina (Thr) y tirosina (Tyr). Esta modificación juega un papel importante en la señalización, activación, propiedades estructurales y de estabilización a nivel celular y se ha observado su importancia en varios pasos de diferentes enfermedades como en los procesos de oncogénesis, apoptosis y desórdenes inmunológicos. Alrededor de una tercera parte de los genes traducidos por los organismos eucariontes pueden sufrir de este tipo de cambios.

La fosforilación específica de los residuos ya mencionados es el mecanismo más común para la regulación de la actividad de proteínas celulares. Las cinasas catalizan la adición de una porción del grupo fosfato de un donador generalmente ATP al grupo hidroxilo de los aminoácidos mientras que la fosfatasas intervienen en la remoción del mismo haciendo este proceso reversible, característica que hace importante esta modificación en la regulación de vías metabólicas (Figura 7). Las cinasas dependen de un ión divalente como el Mg^{2+} , Mn^{2+} o Ca^{2+} y su actividad es regulada por diferentes señales intracelulares las cuales inician gracias a receptores en la membrana o señalizaciones extracelulares, desencadenando una cascada de reacciones que utiliza la célula para activar o desactivar funciones, modular interacciones electrostáticas o conformacionales. Existen dos tipos de cinasas: cinasas de tirosina y las cinasas de serina y treonina y de igual modo dos tipos de fosfatasas, la fosfatasa de tirosina y la fosfatasa de serina y treonina, las cuales se desequilibran en enfermedades crónicas.

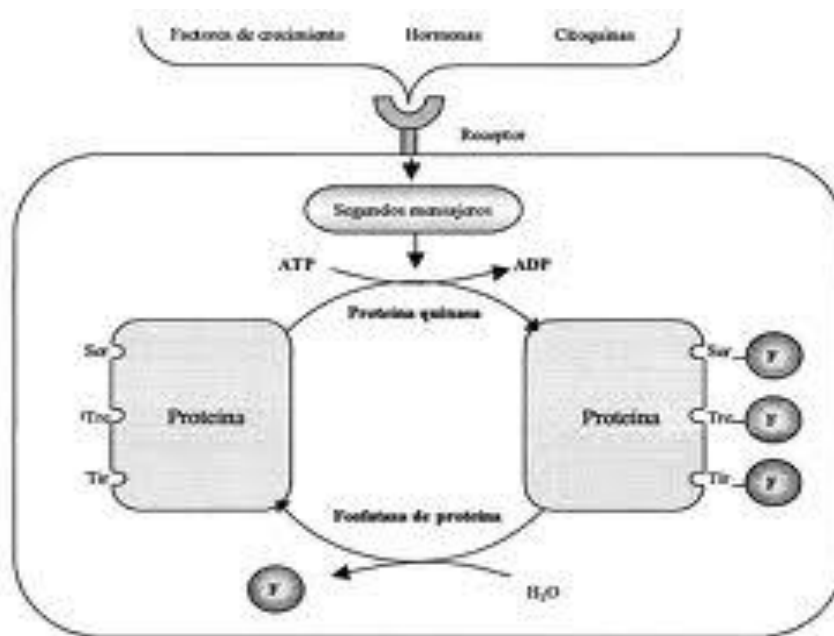


Figura 7. Mecanismo general de fosforilación y defosforilación de proteínas (Nelson y Cox, 2000).

La fosforilación en residuos de aminoácidos es clasificada en O-fosforiladas, N-fosforiladas y como acilfosfatos. Las primeras son la clase principal en términos de cuantificación y localización en el organismo. Las fosfatasa catalizan la defosforilación específica de proteínas, permitiendo así el encendido o apagado de su actividad. Desde hace años se sabe que la fosforilación se relaciona con varios procesos celulares como: señalización transmembranal, amplificación de señales intracelulares y el control del ciclo celular.

La fosforilación específica en los residuos de tirosina está relacionada a factores de crecimiento, los cuales son capaces de estimular a una familia de receptores de superficie celular con actividad intrínseca de tirosin-quinasas, como el Factor de Crecimiento Epidermal (EGF), el Factor de Crecimiento Plaquetario (PGF), ambos con potencial mutagénico, y el Factor de Crecimiento de Fibroblastos (FGF) con poder angiogénico, todos ellos relacionados con el crecimiento y diferenciación celular, motivo por el cual son blancos terapéuticos idóneos para el tratamiento de cáncer, ya que son expresados en la membrana y se pueden bloquear. El estudio de estos factores de crecimiento ha dado como resultado la generación de nuevos medicamentos que inhiben el crecimiento de la neoplasia mediante el bloqueo de la

angiogénesis y por consecuencia de la metástasis, los cuales han demostrado ser efectivos en la Fase clínica III (Herceptin y Avastin).

La fosforilación efectuada por cinasas de serina y treonina son otro grupo importante en la señalización del cáncer. Estas modificaciones tienen como característica principal su estabilidad en medios ácidos y labilidad en los medios básicos. Mientras que las fosfotirosinas son estables en medios básicos.

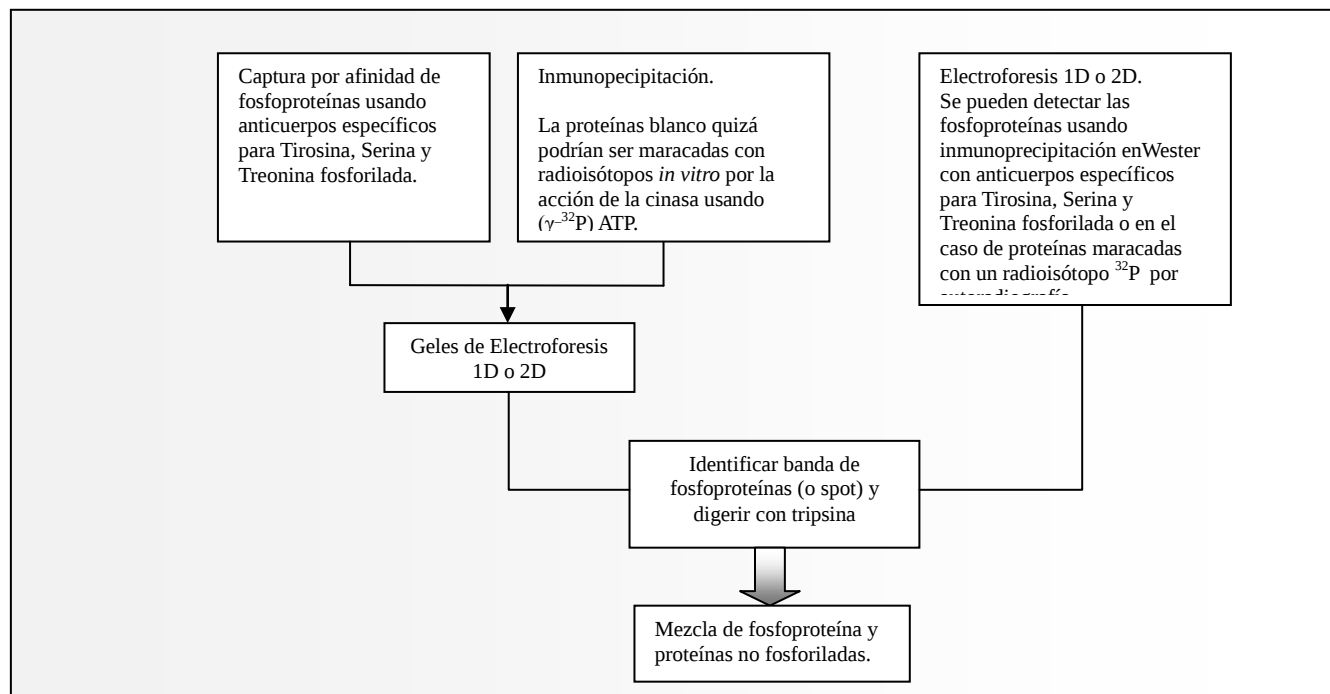
Estas características son aprovechadas por algunas técnicas empleadas para la identificación de éstas modificaciones como son la inmunodetección por Western y la inmunoprecipitación en donde se usan anticuerpos específicos de estos residuos o se identifican regiones fosforiladas y proteínas marcadas por fluorescencia. Otro de los protocolos más empleados es una previa separación mediante la utilización de cromatografía de afinidad y su posterior identificación con MS y/o MS/MS. Estas técnicas tienen una mayor aplicación en la identificación de O-fosforilaciones con peso molecular mayor a 80 kDa. Otras de las técnicas más empleadas en la identificación de esta clase de proteínas son los geles SDS-PAGE y los 2D, la primera utilizando la carga y masa de los péptidos dando como resultado una serie de bandas a lo largo de un polímero, la segunda utilizando el pI de las proteínas además de la masa. Una vez localizadas las manchas de interés, lo recomendable es la extracción, digestión e identificación por espectrometría de masas (Figura 8).

Preparación de la muestra.

Mezcla de proteínas fosforiladas y no fosforiladas.

Las proteínas metabólicas podrían ser marcadas con radioisótopos *in vivo* genralmente (^{32}P).

Enriquecimiento/separación
Fosfoproteínas.



Identificación de fosfoproteínas/
Péptidos y sitios.

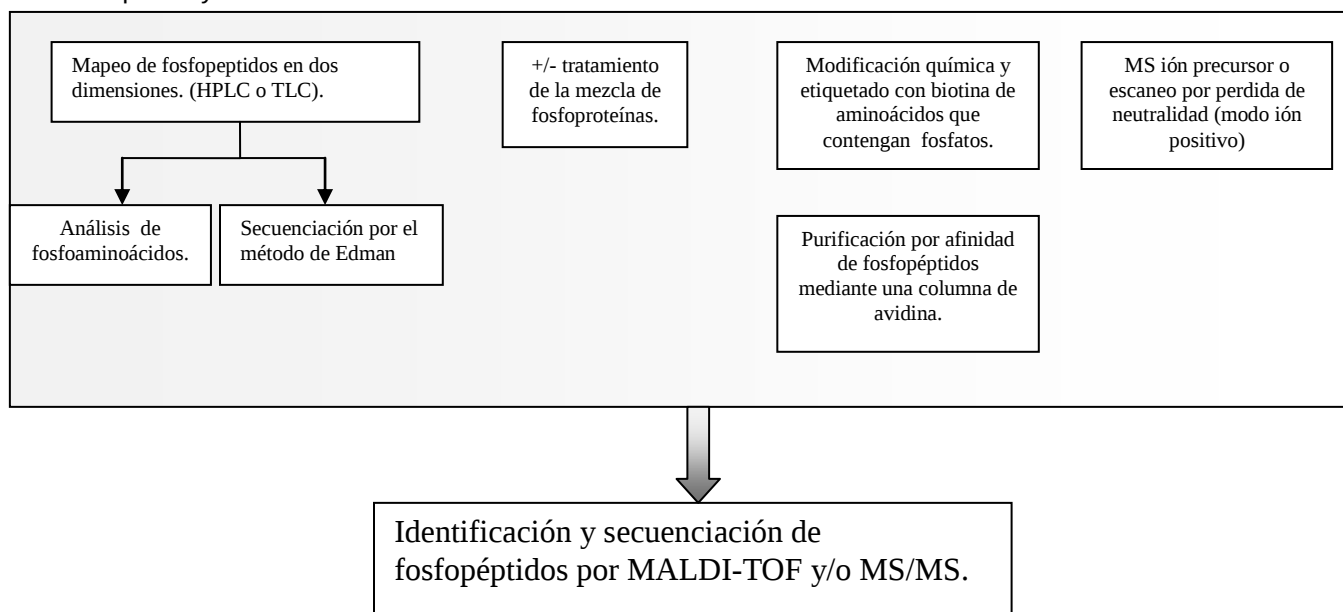


Figura 8. Técnicas empleadas para le identificación de fosfoproteínas (Cox. 2007).

2.6.2 Glicosilación.

El gran número de proteínas presentes en las células no es el único factor que contribuye en la complejidad del proteoma, un segundo factor con la misma importancia son las modificaciones postraduccionales, las cuales generalmente regulan la actividad de rutas metabólicas o la localización de las mismas.

Las glicoproteínas poseen uno o varios oligosacáridos que pueden representar del 3% hasta el 40% de la masa total de la proteína. Los residuos son de diferente composición y están unidos covalentemente, pueden encontrarse dentro y fuera de la célula, por ejemplo en la superficie de la membranal, en la matriz extracelular o bien en organelos, como el aparato de Golgi, gránulos secretores y lisosomas.

La porción de oligosacáridos presentes en glicoproteínas es menor que la presente en los proteoglicanos, sin embargo, éstas son ricas en información debido a que forman muchos sitios específicos de reconocimiento y poseen una elevada afinidad para la unión a otras proteínas. Los oligosacáridos se componen generalmente de galactosa (gal), fucosa (Fuc), N-Acetilglucosamina (GlcNAc) y acetilgalactosamina (GalcNAc), manosa (Man), ácido siálico (NANA), ácido N-acetilneuramínico y s-inositol.

Una de las modificaciones más comunes junto con la fosforilación es la glicosilación de residuos de serina, treonina, y asparagina con mono u oligosacáridos, estas glicoproteínas están constituidas por tres regiones principalmente:

- 1) Núcleo o core. En esta región se encuentran los carbohidratos más cercanos a la unión con la cadena polipeptídica.
- 2) Esqueleto. Esta región determina la longitud del oligosacárido. La unión de monosacárido-monosacárido esta mediada por enlaces glicosídicos.
- 3) Región periférica. Se encuentran antígenos de importancia biológica (Grupo sanguíneo).

Esté tipo de modificaciones juega un papel importante en procesos celulares como la adhesión, la señalización celular, la estabilización de procesos celulares, el tráfico de proteínas, la interacción de ligandos con sus receptores y la oncogénesis en varios tipos de enfermedades (Cáncer de mama, artritis reumatoide, etc.), quizá puedan ser causadas por un defecto en la glicosilación de proteínas. Se ha

observado que casi la mitad de todos los productos proteicos en eucariontes pueden ser modificados por la adición de un carbohidrato, el cual puede o no intervenir de forma importante en su función.

Las glicoproteínas se dividen en dos grandes grupos de acuerdo a la unión entre el azúcar y el aminoácido: en N- y O-glicanos, el primer grupo tiene como característica principal que la unión del oligosacárido con el átomo de nitrógeno se da en un residuo de asparagina (Asn) y se lleva a cabo en el retículo endoplásmico rugoso (RER). En el segundo grupo la unión del oligosacárido se lleva en el átomo de oxígeno de los residuos de serina (Ser) y treonina (Thr), este tipo de unión es llevada en el aparato de Golgi.

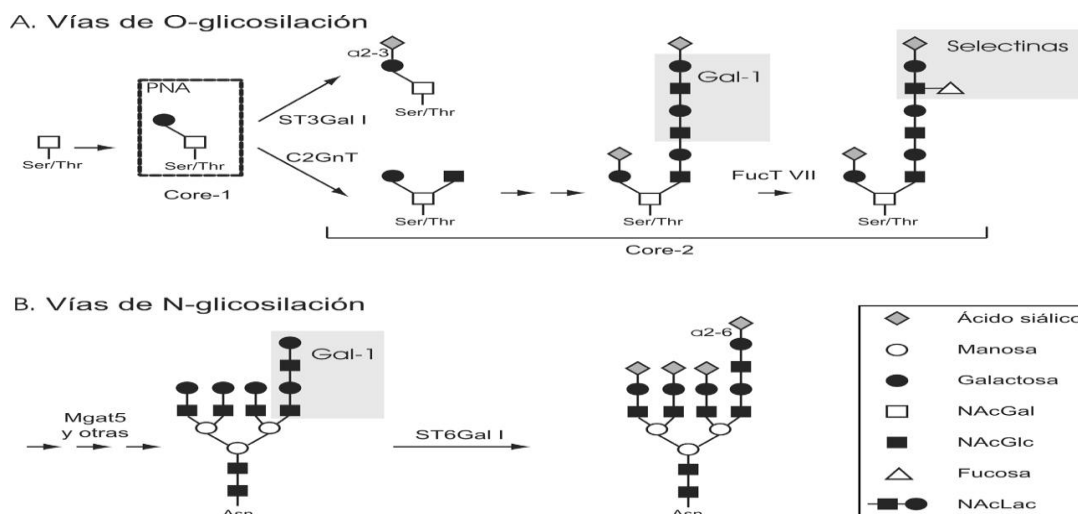


Figura 9. Tipos de glicosilación. A) O-glicosilación, B) N-glicosilación (Toscano, 2006).

La N-glicosilación puede formar moléculas de tres tipos: el primero altamente compuestas de manosa que se conforman por un proceso secuencial y que pueden ser: ricas en manosa, híbrido y complejas (Figura 9), siendo la primera clase el tipo más importante. La síntesis de los N-oligosacáridos comienza en el citosol que se encuentra cercano a la membrana del RER y comienza con la generación de dolicol pirofosforil-N-acetilglucosamina, la cual gracias a la glicosiltransferasas después de una serie de reacciones da como resultado el $\text{Man}_5\text{GlcNAc}_2\text{-PP-Dol}$, este sustrato ingresa al lumen del RER en donde ocurren

varios pasos que una vez finalizados lo llevan la cara del citosol y ahí es transformada a $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2\text{-PP-dol}$, un oligosacárido rico en manosa el cual es capaz de atacar a un residuo de asparagina de una cadena nascente.

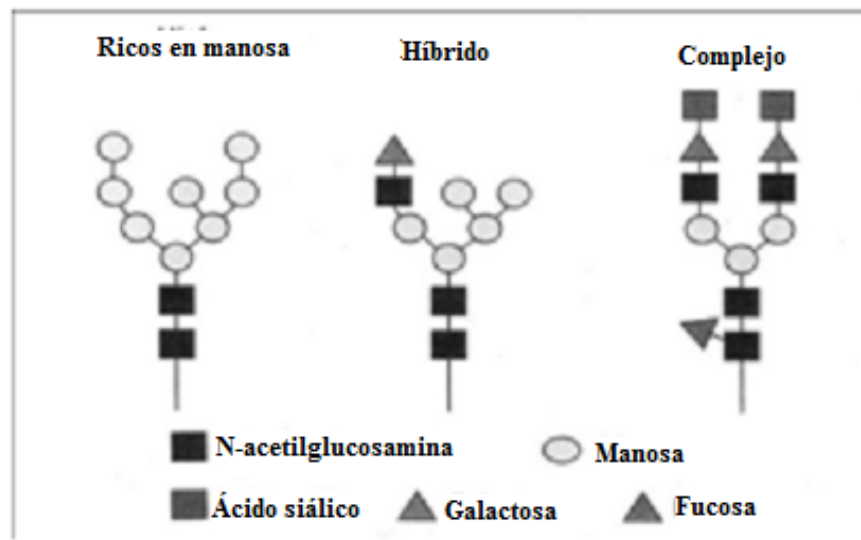


Figura 10. Tipos de oligosacáridos que se puede unir en una N-Glicosilación (Simpson, 2003).

Se ha relacionado a los O-glicanos que poseen una N-acetil-D-galactosamina (que se encuentra unida al hidroxilo de la serina o treonina), con el cáncer debido a que existe una mayor diversidad como resultado de los 8 núcleos descritos²⁰. Debido a su gran diversidad y complejidad, los O-glicanos participan en funciones diversas como la conformación de la estructura secundaria, terciaria y también participan evitando la agregación proteica.

Una de las principales etapas afectadas durante el cáncer es la de elongación de los O-glicanos. Cuando hay acortamiento de las cadenas de oligosacáridos, se exponen azúcares que generalmente se encuentran enmascarados en la superficie celular, el resultado es la formación de antígenos asociados al cáncer.

La participación en diversas funciones y la formación de antígenos hace que los O-glicanos sean candidatos para ser propuestos como biomarcadores moleculares. Algunos ejemplos de O-glicanos reportados en el Cáncer de mama son la MUC1 que es una mucina con un peso aproximado de 250 a 500 kDa, ésta se encuentra

asociada a la membrana y se ha relacionado con la metástasis. Una mucina es una glicoproteína de elevado peso molecular que se caracterizan por tener cadenas complejas de oligosacáridos unidas a proteína por enlaces O-glicosídicos.

La glicosilación es catalizada por las glicosiltransferasas, las cuales transfieren a un donador adecuado de la proteína un mono u oligosacárido. Esta reacción se realiza en el RE (retículo endoplásmico), el conjugado resultante es llevado por todo el RE hasta el aparato de Golgi y de ahí, mediante el péptido señal se dirige hacia la membrana plasmática, lugar en donde se localiza la mayoría de este tipo de proteínas para ser secretada hacia la matriz celular¹¹.

Los estudios proteómicos de glicoproteínas se realizan mediante el fraccionamiento de muestras complejas, las cuales pueden ser analizadas por geles 1D y 2D o espectrometría de masas (Figura 11). Un método muy utilizado para el fraccionamiento y enriquecimiento de glicoproteínas es la cromatografía de afinidad mediante lectinas (Tabla 12). Estas son proteínas vegetales con una elevada afinidad hacia los azúcares, capaces de unírseles de forma reversible¹⁷.

Gracias a la proteómica comparativa es posible la identificación de péptidos que posean estas modificaciones y realizar comparaciones entre el proteoma de células de tumores y las células de tejidos sano para localizar diferencias entre estas dos muestras y así proponer posibles biomarcadores.

Tabla 12. Lectinas empleadas en la separación de glicoproteínas. La tabla nos presenta las lectinas más empleadas en la separación de glicoproteínas y sus ligandos (Qiagen, 2000).

Lectina y origen.	Abreviatura.	Ligando(s).
Plantas		
Concavalina A	ConA	Man α 1-OCH ₃
Lectina 4 (<i>Griffonia simplicifolia</i>)	CS4	Lewis b (le ^b) tetrasacárido
Aglutinina de germen de trigo	WGA	Neu5Ac(α 2 $\rightarrow$ 3)Gal (β 4 $\rightarrow$ 4)Glc GlcNac(β 4 $\rightarrow$ 4)GlcNac
Ricina		Gal(β 1 $\rightarrow$ 4)Glc
Animal.		
Galectina-1		Gal(β 1 $\rightarrow$ 4)Glc
Proteína A por afinidad a manosa	MBP-A	Octa sacáridos ricos en manosa
Virales		

Hemaglutinina del virus de la influenza.	HA	Neu5Ac(α 2 $\rightarrow$ 6)Gal (β 1 $\rightarrow$ 4)Glc
Proteína 1 del virus del polio	VP1	Neu5Ac(α 2 $\rightarrow$ 3)Gal (β 1 $\rightarrow$ 4)Glc
Bacterial.		
Enterotoxina	LT	Gal
Toxina del cólera	CT	Pentasacárido GM1

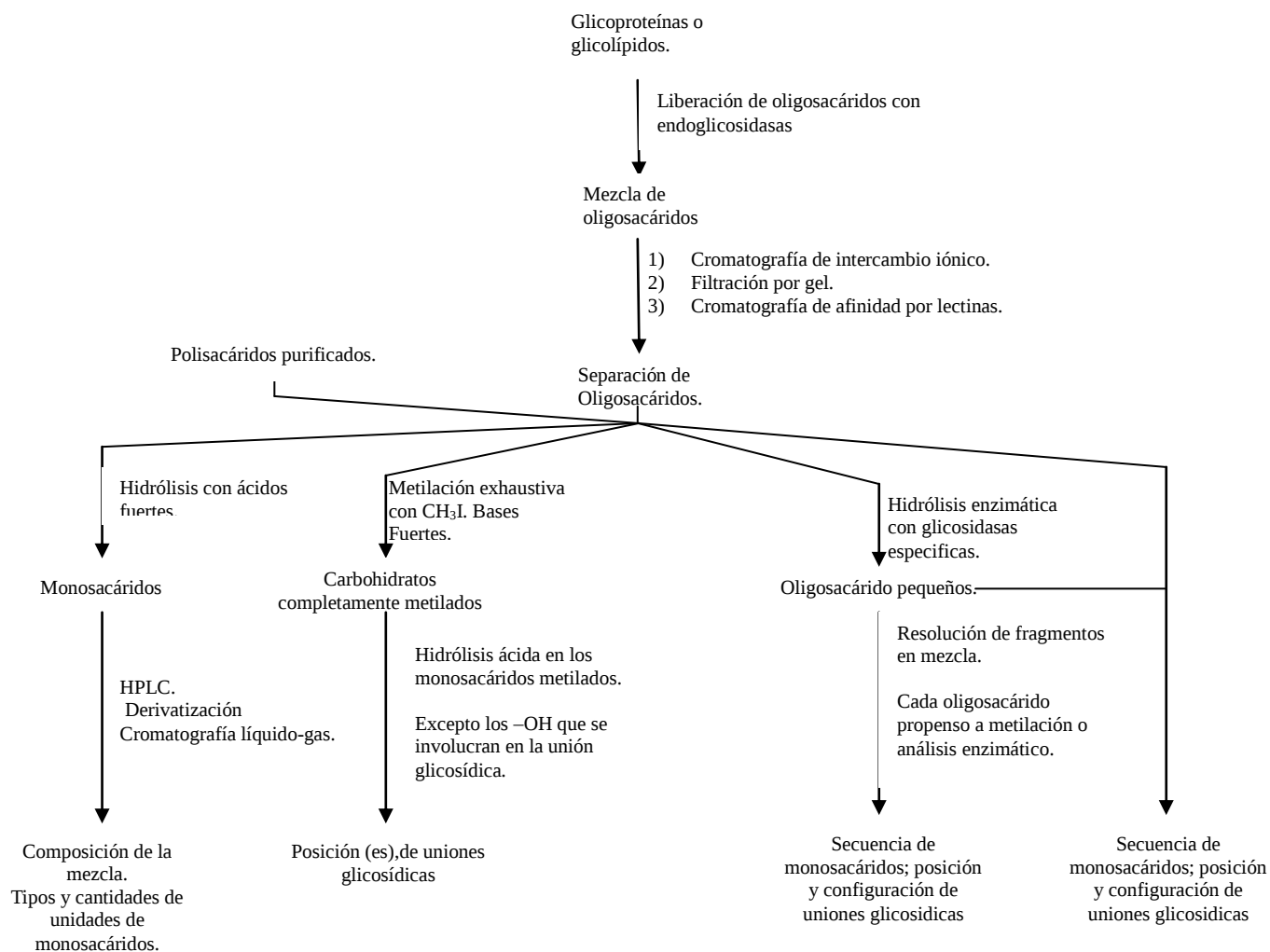


Figura 11. Métodos empleados para la identificación de glicoproteínas (Karp, 2004).

2.7 Biomarcadores.

A pesar de décadas de investigación sobre el Cáncer de mama, las tasas de mortalidad se han incrementado debido al poco conocimiento de la etiología molecular de la enfermedad, a la falta de métodos de diagnóstico sensibles que permitan la detección de esta enfermedad en estadios tempranos y al pobre entendimiento de la epidemiología de los blancos terapéuticos. Es por esto que se ha realizado un gran esfuerzo en la generación de métodos de diagnóstico que tengan como características principales una elevada sensibilidad y especificidad garantizando así un diagnóstico confiable, que la obtención de muestras sea sencilla y que además estos métodos puedan emplearse en estadios tempranos, garantizando con esto una terapia adecuada y efectiva.

Uno de los métodos de diagnóstico sensible, específico y que la obtención de las muestras para su realización es sencilla, además de haber demostrado que puede ser empleado en la detección de cáncer en estadios tempranos es el diagnóstico mediante la utilización de biomarcadores moleculares.

Un biomarcador se define de acuerdo a la Academia de Ciencias como: “Un indicador de la señalización de eventos en sistemas biológicos o muestras”, o bien como “Una biomolécula que se expresa bajo ciertas circunstancias, en cierto tiempo y lugar, la cual varía en su nivel de expresión y estructura, lo que repercute en su actividad disminuyéndola o incrementándola”.

La promoción para el estudio de biomarcadores inició en la década de los 80's gracias a dos instituciones, la NIST (National Institute of Standards and Technology) y CDC (Center for Disease Control and Prevention). Ambas instituciones concuerdan que el estudio de biomarcadores debe realizarse mediante la utilización de diversas técnicas moleculares y que la principal finalidad de los biomarcadores es la de coadyuvar con el diagnóstico y pronóstico de cáncer.

Para el estudio y caracterización de biomarcadores los investigadores necesitan satisfacer ciertas condiciones como:

- Desarrollar y promover procesos de validación de biomarcadores consistentes, eficientes y efectivos.
- Una retroalimentación entre las bases de datos que contienen información

sobre el uso y descubrimiento de los mismos.

- Repetibilidad y reproducibilidad de biomarcadores propuestos.
- Mejorar los accesos de muestras biológicas o experimentales que sostengan la investigación y desarrollo de éstos.

Las primeras biomoléculas en las cuales se observó una relación con el cáncer fueron: el antígeno carcinoembrionario (Gold y Freedman, 1975), el cual se relacionó con el cáncer colorectal hasta en un 100%, debido al diseño inadecuado en la parte experimental; el antígeno específico de próstata (Stamer, 1987) relacionado ampliamente con el cáncer de próstata, éste antígeno presenta hasta un 30% de falsos positivos, debido al pobre desarrollo de técnicas proteómicas cuando se descubrió y por último una glicoproteína de membrana denominada CA-125 la cual se relacionó con cáncer de ovario (Bast, 1981). Cada una de estas biomoléculas tenía como característica principal la falta de especificidad y selectividad debido al diseño experimental y al pobre desarrollo de las técnicas empleadas para su identificación y caracterización.

Gracias al desarrollo de la genómica y la aparición de diversas ramas como la proteómica, en los años 90's se reportan los primeros posibles biomarcadores relacionados con cáncer de mama por Bidart (Bidart, 1994), éstos fueron BRCA1 y BRCA2, los cuales han sido ampliamente estudiados y hoy en día se sabe que BRCA1 se encuentra presente en más del 45% de los casos de Cáncer de mama hereditario y 80% en los casos de cáncer de ovario, por lo que se utiliza como un biomarcador para cáncer ovárico, mientras que BRCA2 ha sido detectado en formas poco comunes de Cáncer de mama heredado.

En un esfuerzo realizado por el National Cancer Institute de evitar propuestas erróneas de biomarcadores señaló 5 fases obligatorias. Para que una biomolécula pueda ser considerada como un biomarcador debe pasar por 5 fases, las primeras dos tienen como finalidad demostrar la presencia de estas biomoléculas en líneas celulares y tejido, además de su realización la caracterización y validación de estas biomoléculas mediante técnicas de biología molecular. Las tres fases restantes buscan demostrar la repetibilidad y reproducibilidad de los biomoléculas

identificadas. Una vez que las biomoléculas pasan con éxito las 5 etapas puede ser considerada como un biomarcador. Los biomarcadores se clasifican en tres grandes grupos:

- a) Biomarcadores de diagnóstico. Estos biomarcadores son sobreexpresados en diferentes tipos de cáncer y que deben de ser sensibles y altamente específicos.
- b) Biomarcadores de pronóstico. Son biomarcadores que nos pueden indicar el estatus en el cual se encuentra la enfermedad.
- c) Biomarcadores predictivos. Son biomarcadores que sirven para el monitoreo de la respuesta hacia los tratamientos.

Las características principales que deben de cumplir los biomarcadores son las siguientes:

- 1) Precisión y reproducibilidad. Se debe demostrar en las fases de desarrollo mediante métodos estadísticos y estudios de laboratorio, la utilización de controles internos y un gran número de muestras es necesaria para demostrar estas características.
- 2) Especificidad/Sensibilidad. Un biomarcador debe ser suficientemente sensible para proporcionar resultados precisos con una cantidad muy pequeña. Una elevada sensibilidad podría evitar falsos negativos, mientras que una alta especificidad podría evitar falsos positivos en las muestras. La utilización de controles y estándares internos es necesaria para evaluar estos parámetros.
- 3) Recuperación en base de datos y análisis retrospectivos. La comparación y administración de datos son importantes en el desarrollo y validación de biomarcadores para su uso en estudios poblacionales, es por esto que deben existir bases de datos capaces de proporcionar información sobre los diversos estudios y sobre el desarrollo de un mismo biomarcador en diversos laboratorios.
- 4) Modelos experimentales y credibilidad. La relación existente entre la exposición y los efectos es ambigua, por lo cual se debe comprobar esta relación con el uso de un modelo experimental celular o animal.

- 5) Obtención de muestras. El método por el cual se obtenga la muestra analizar debe ser lo menos invasivo posible, por ésto la técnica por la cual se detecte al biomarcador debe poseer un elevado rendimiento.

Algunos ejemplos de los biomarcadores de diagnóstico identificados relacionados con Cáncer de mama son:

- El Ki67, es un antígeno nuclear presente en las fases activas del ciclo celular (G1, S, G2, M) y ausente en la fase G0.
- El PCNA, una proteína que juega un papel en el inicio de la proliferación celular incrementando la actividad de la ADN polimerasa. Se han detectado niveles elevados en las fases S-G2 y M, por lo que existe una correlación elevada entre la existencia de esta proteína con un elevado índice mitótico y una correlación negativa con la expresión de receptores de estrógenos (RE) y progesterona (RP), además de existir una correlación de la inmunexpresión de PCNA y Ki67 con la existencia de metástasis ganglionar.
- *Her-2/neu (c-erbB-2)*, un gen que codifica a una proteína homóloga para el receptor del factor de crecimiento epidérmico (EGFR). Este gen se encuentra amplificado en los carcinomas de mama y en otros tumores.
- *p53* es un oncogen cuya alteración es el cambio genético más común identificado en tumores. Su producto proteico se encuentra relacionado con la regulación del ciclo celular en el paso de G1 a S, donde es el encargado de verificar que la célula cumpla con todos los requerimientos para continuar con el ciclo celular.

Para la identificación de estos biomarcadores podemos utilizar diferentes técnicas para purificarlas, separarlas e identificarlas. Entre ellas destacan geles SDS-PAGE, 2D y la espectrometría de masas principalmente (consultar Anexo I).

Capítulo 3. Objetivos.

3.1 Objetivo general.

- Identificar posibles biomarcadores de Cáncer de mama que presenten cambios postraduccionales como fosforilación y glicosilación en su estructura.

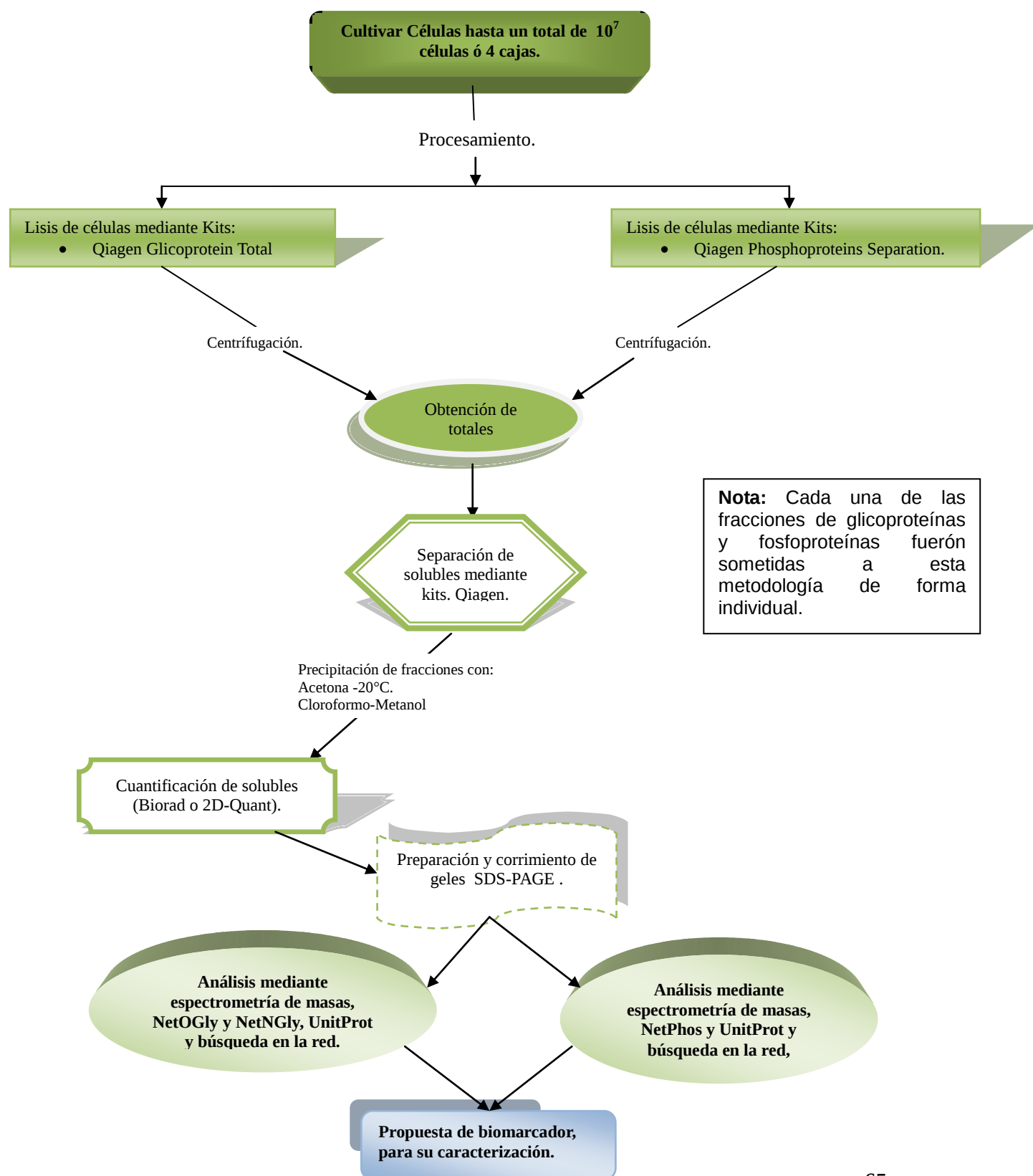
3.2 Objetivos particulares.

- Identificar la presencia de fosfoproteínas y glicoproteínas en el proteoma de líneas celulares de diversos tipos de Cáncer de mama.
- Identificar las proteínas que presentan estos cambios postraduccionales en su estructura.
- Proponer nuevos biomarcadores que colaboren con el diagnóstico de Cáncer de mama.

Capítulo 4. Hipótesis.

Existen cambios postraduccionales, los cuales funcionan como un interruptor para encender o apagar rutas metabólicas, ejemplos de éstos son la fosforilación y la glicosilación. Durante el desarrollo del cáncer muchas de estas vías son desreguladas, por lo que debe existir una diferencia en los niveles de fosforilación y glicosilación de proteínas para producir dicho desorden. Estas proteínas podrían ser propuestas como biomarcadores específicos de Cáncer de mama, los cuales podrían coadyuvar a un diagnóstico temprano de la enfermedad.

Capítulo 5. Metodología. Protocolo de análisis de Fosfoproteínas y Glicoproteínas.



5.1 Líneas celulares.

Se utilizaron 4 líneas celulares de Cáncer de mama (ATCC), de las cuales 3 provienen tejido canceroso y una de tejido sano (línea control). Algunas de sus características son:

- Línea celular MDA-MB 231. Células epiteliales de tejido mamario, adherentes, obtenidas de un derrame pleural de una paciente caucásica de 74 años de edad, diagnosticada con adenocarcinoma. Sus características principales son la expresión del factor epidermal de crecimiento (EGF) y el factor α transformador de crecimiento. Al igual que la línea MCF7 expresa el oncogen *WNT7B*, el cual codifica para algunas proteínas de señalización. Estas células fueron cultivadas en un medio Leibovitz's L-15 (ATCC No. de catálogo 30-2008, Apéndice II) a 37°C, en una atmósfera compuesta por 95% aire y 5% CO₂, hasta alcanzar una población de 1.5×10^6 células por caja petri.
- Línea celular MFC7. Células epiteliales obtenidas de un derrame pleural de una paciente caucásica de 74 años, diagnosticada con adenocarcinoma, la cual conserva varias características de tejido mamario diferenciado como la capacidad de producir estradiol mediante la estimulación de receptores citoplásmicos para estrógenos, la capacidad de la formación de estructuras celulares y la expresión del oncogen *WNT7B*. El crecimiento de estas células es inhibido por TNF- α y la secreción de proteínas de unión a Factores de Crecimiento similares a la Insulina (IGFBP's por sus siglas en inglés) BP-2; BP-4; BP-5, la secreción de éste factor y proteínas es modulada con un tratamiento contra estrógenos. Estas células se cultivaron en un medio DMEM (ATCC No. de catálogo 30-2002, Apéndice II) a 37°C, en una atmósfera compuesta por 95% aire y 5% CO₂, hasta alcanzar una población 3×10^6 células por caja petri.

- Línea celular Hs578T. Son Células epiteliales originarias de tejido periférico con carcinoma ductal infiltrate, son no tumorigénicas y fueron tomadas de una paciente caucásica con carcinoma de 74 años de edad. Mediante microscopia electrónica se han observado agregados de proteínas granulares como caseína, desmosomas y un vascularizado retículo endoplasmático liso. Al igual que la línea Hs578BsT no posee receptores para estrógenos y/o para virus endógenos, además presenta una hipotriploidia. Estas células se cultivan en un medio DMEM (ATCC No. catálogo 30-2002, Apéndice II) adicionado con insulina bovina (0.01 mg/mL) y suero fetal bovino (concentración final 10%) en una atmósfera compuesta por 95% de aire y 5% de CO₂, a 37°C, hasta alcanzar una población de 4.5×10^6 células por caja petri.
- Línea celular Hs578BsT. Esta línea celular se obtuvo de tejido periférico de la mama de una mujer caucásica de 74 años de edad. Son células fibroblásticas adherentes, en su origen quizá fueron mioepiteliales, ya que poseían microfilamentos y grupos de vesículas de pinocitosis similares a las observadas en células mioepiteliales *in vivo*, no presentan desmosomas, ni receptores de estrógenos pero si receptores a virus endógenos. Esta línea celular se propaga en un medio Hybri-Care (ATCC No. de catálogo 46-X) al cual se le adiciona EGF y suero fetal bovino hasta una concentración final del 10% a 37°C en una atmósfera compuesta por un 95% de aire y 5% de CO₂, hasta alcanzar un conteo de 3×10^6 en cada caja o bien tener más de 5 semanas en crecimiento, debido a un lento crecimiento.

Una vez alcanzado el crecimiento de cada una de las líneas celulares se toma el mismo número de cajas para la separación de fosfoproteínas y glicoproteínas y se procede a las metodologías indicadas en cada uno de los manuales de uso de los kits comerciales.

Material.

- Líneas celulares
 - MDA-MB 231.
 - MCF7.
 - Hs578T.
 - Hs578BsT.
- Qiagen Phosphoproteins Separation.
- Qiagen Glicoprotein Total.
- Tubos eppendorf de 1.5 mL.
- Microcentrífuga.
- Amortiguador TBS 1X.
- Micropipetas 5, 20, 100, 200 y 1000 μ L.

5.2 Protocolo para purificación de fosfoproteínas mediante Kits Qiagen.

- 1) Recolectar 10^7 células (4 cajas aproximadamente) en un tubo Falcon de 15 mL, centrífugar a 1500 rpm, dejar a 4°C (baño de hielo), procurando no dejar sin amortiguador las células.
- 2) Preparar dos soluciones (A y B) hasta llegar a una concentración de 0.25% de CHAPS en ambos:
 - 875 μ L de solución stock CHAPS (10%w/v) a 35 mL de amortiguador de lisis Phosphoprotein (solución A).
 - 75 μ L de solución stock CHAPS (10% w/v) a 3 mL de amortiguador de elución Phosphoprotein (solución B).
- 3) Adicionar una tableta de inhibidor de proteasa y 10 μ L de solución stock de Benzonasa a 5 mL de solución A, mezclar para así tener la solución C.
- 4) Retirar el sobrenadante de las células, adicionar 5 mL de solución A y la solución C, resuspender el pellet con pipeteo gentil.
- 5) Incubar por 30 min a 4°C, mezclar brevemente con vortex cada 10 min para lisar las células bacterianas.
- 6) Después de 30 min, centrífugar a 10000 g a 4°C por 30 min, decantar.

Cuantificar el sobrenadante,

- 7) Durante la centrifugación, preparar una columna de purificación de fosfoproteínas (Quiagen) separando la tapa superior, rompiendo el botón de cierre, y permitiendo el flujo. Adicionar 4 mL de solución A para equilibrar la columna permitiendo la salida de amortiguador.
- 8) Tomar un volumen de sobrenadante que contenga aproximadamente 2.5 mg de proteína total y ajustar su concentración a 0.1 mg/mL con la adición de solución A. Se puede obtener un volumen cercano a 25 mL.
- 9) Poner la mitad de solución obtenida en Paso 8 (12.5 mL) en el reservorio superior de la columna de purificación de fosfoproteínas (Quiagen), cuando casi todo el volumen esté dentro de la columna, adicionar la mitad restante de la solución y permitir el paso a través de la columna, recolectar la fracción y si se desea realizar un estudio de proteínas no fosforiladas.
- 10) Adicionar 6 mL de Solución A para lavar la columna, permitiendo el flujo a través de la columna, recolectar en un tubo y descartar.
- 11) Adicionar 500 µL de Solución B (amortiguador de elución) a la columna, y coleccionar las fracciones, repetir este paso 6 veces, la mayor cantidad de fosfoproteínas se obtienen en las fracciones 3 y 4.
- 12) Determinar la concentración de las fracciones.
- 13) Proceder al análisis mediante electroforesis SDS-PAGE,

Notas.

- La lisis se lleva a cabo a 4°C (baño de hielo). La purificación de proteínas se lleva a cabo a temperatura ambiente, ya que el inhibidor de proteasa previene la degradación proteolítica por lo que no es necesario temperaturas bajas además de que a 4°C disminuye el rendimiento de la purificación. No se debe utilizar amortiguador de fosfatasa para el lavado de células, porque puede interferir con la unión de fosfoproteína-resina se recomienda la utilización de TBS 1X Amortiguador pH 7.4 (Anexo II).

*El paso por la columna puede tardar aproximadamente 50 min, es necesario permitir la unión de fosfoproteínas y la columna. La purificación, elución, cuantificación pueden realizarse a temperatura ambiente (15-25°C), una vez

cuantificadas las fracciones es necesario almacenarlas a -8°C.

- De requerirse más inhibidor de proteasa se recomienda el uso de “Phosphatasa Inhibitor Cocktail” de Sigma (No. de catálogo P2850).
- NO utilizar vanadato, molibdato o flúor, ya que interfiere en el proceso de purificación.

5.3 Protocolo para purificación de glicoproteínas mediante kits Qiagen.

- 1) Colectar las células utilizando un gendarme. Centrifugar por 5 min a 450x g y lavar las células con amortiguador HEPES-based o TBS por duplicado (Anexo II). Almacenar a 4°C. Durante la centrifugación, por cada “spin colum” que vaya a utilizar, preparar:
 - Solución 1: 3 mL de el amortiguador de unión con 30 µL de solución inhibidora de proteasa (100X) y 300 µL de solución detergente.
- 2) Con pipeteo gentil, resuspender el pellet de células correspondientes a 1×10^7 células en 1 mL de Solución 1. Incubar por 15 min a 4°C. Agitar brevemente cada 5 min.
- 3) Completar la ruptura de las células utilizando una aguja y jeringa, succionar el lisado lentamente en la jeringa, expulsar de tal forma que las células se estrellen en la pared del tubo, repetir esto 11 veces. Centrifugar el lisado a 10000x g y 4°C por 20 min, decantar y colocar el sobrenadante en un baño de hielo.
- 4) Durante la centrifugación preparar las “spin columns” aflojando el tapón de la columna y colocarlo en un tubo colector de 2 mL. Centrifugar la “spin colum” por 2 min a 500 rpm.
- 5) Descartar el flujo que atravesó la columna y poner 500 µL de amortiguador de unión en la columna. Centrifugar por 2 min a 500 rpm en una microcentrífuga.
- 6) Descartar el buffer que atraviese la columna y aplicar 500 µL de la muestra obtenida en el paso 3. Incubar por un minuto y centrifugar por 2 min a 500 rpm.

- 7) Colectar y descartar las fracciones eluidas ya que no contienen fosfoproteínas
- 8) Repetir pasos 10 y 11 con toda la muestra obtenida en el paso 3.
- 9) Adicionar 750 μ L de Solución 1 para lavar la columna, centrífugar por 2 min a 500 rpm y descartar la fracción. Repetir este paso.
- 10) Durante la centrifugación preparar:
Solución 2: 1 μ L de Inhibidor de proteasas, 10 μ L de solución detergente a 100 μ L de amortiguador de elución ME.
Solución 3: 2 μ L de Inhibidor de proteasas, 20 μ L de Solución detergente a 200 μ L de amortiguador de elución SE.
- 11) Transferir la columna a un tubo de microcentrífuga limpio. Adicionar 100 μ L de Solución 2, incubar por 1 min y centrífugar a 500 rpm
- 12) Repetir el paso anterior por duplicado pero con solución 3, recolectar la fracción y determinar la concentración por el método de Lowry.
- 13) Analizar mediante la electroforesis SDS-PAGE.

Notas.

- La lisis se lleva a cabo a 4°C (baño de hielo), la purificación de proteínas es a temperatura ambiente, ya que el inhibidor previene la degradación proteolítica pues a 4°C disminuye el rendimiento de la purificación. NO se debe utilizar amortiguador de fosfatasas para el lavado de células ya que puede interferir con la unión de fosfoproteína-resina se recomienda la utilización de TBS 1X bufer pH 7.4 o HEPES 1X (Apendice II).
- NO utilizar vanadato, molibdato o flúor, ya que interfiere en el proceso de unión ConA-glicoproteína.

5.4 Precipitación de proteínas.

La precipitación tiene como finalidad la concentración, el desalado o eliminación de compuestos que pudieran intervenir en el análisis posterior de las fracciones de interés. En ésta se emplean una gran variedad de disolventes, los cuales deben de ser elegidos de acuerdo a las propiedades de las muestras en la precipitación de proteínas. Los disolventes más empleados son la acetona o bien el cloroformo en combinación con metanol.

Material biológico.

- Fracciones obtenidas.

Materiales.

- Acetona (-20°C)
- Cloroformo.
- Metanol.
- Tubos de microcentrífuga (1.5 mL)
- Microcentrífuga.
- Desecador rotatorio.
- Micropipetas 5, 20, 100, 200 y 1000 µL.

5.4.1 Precipitación de fosfoproteínas.

Precipitación con Cloroformo-Metanol de las fracciones obtenidas.

Esta técnica se emplea para concentrar las muestras proteicas y eliminar componentes que intervengan con la electroforesis SDS-PAGE.

- 1) Colocar 150 µL de la muestra en un tubo eppendorf.
- 2) Adicionar un volumen de cloroformo, 3 volúmenes de metanol y 4 de agua bidestilada.
- 3) Mezclar con vortex y centrífugar a 15000 rpm por 15 min.
- 4) Descartar con pipeta el sobrenadante.
- 5) Lavar con cuatro volúmenes de metanol.
- 6) Centrífugar a 15 000 rpm por 10 min.

- 7) Descartar el sobrenadante con micropipeta. Dejar secar el pellet a temperatura ambiente.
- 8) De acuerdo a la técnica empleada reconstituir en el amortiguador requerido DIGE (7M Urea, 2M Tiourea, 30 mM Tris, 4% CHAPS pH 8.5). Almacenar a -60°C.

5.4.2 Precipitación de glicoproteínas.

Precipitación por acetona de las fracciones obtenidas.

Esta técnica se emplea para la concentración y desalado de las muestras proteicas obtenidas para su utilización en SDS-PAGE.

- 1) Adicionar cuatro volúmenes de acetona fría (-20°C) a la fracción de la proteína e incubar por 1 h a 4°C.
- 2) Centrífugar por 10 min a 12 000g en una centrífuga a 4°C. Descartar el sobrenadante y secar a temperatura ambiente, no desecar el pellet ya que se dificulta su resuspensión.
- 3) De acuerdo a la aplicación, resuspender el pellet en el amortiguador requerido DIGE (7 M Urea, 2 M Tiourea, 30 mM Tris, 4% CHAPS pH 8.5). Almacenar a -60°C.

Una vez precipitadas las fracciones y reconstituidas en el amortiguador adecuado, se procede a la cuantificación de las mismas.

Nota. La preparación del amortiguador DIGE se encuentra en el Apéndice II.

5.5 Cuantificación de proteínas.

El ensayo empleado para la cuantificación proteica es una prueba colorimétrica que tiene como finalidad la determinación de proteína total en presencia de agentes reductores y detergentes, que está basada en el método Lowry (Lowry, 1951) el cual nos permite realizar una determinación cuantitativa de las proteínas. A la muestra se añade un reactivo que forma un complejo colorido con los residuos de tirosina de las proteínas, el color resultante es directamente proporcional a la cantidad de proteína presente, según la ley de Lambert-Beer.

El método consta de dos etapas: en la primera, los iones Cu^{2+} (medio alcalino) se unen a las proteínas formando complejos con los átomos de nitrógeno de los enlaces peptídicos. Estos complejos de Cu^{2+} -proteína son de color azul claro y provocan el desdoblamiento de la estructura tridimensional de la proteína, exponiendo así los residuos de tirosina que van a participar en la segunda etapa de la reacción. El Cu^{2+} se mantiene en solución alcalina en forma de su complejo con tartrato. La segunda etapa es la reducción (medio básico) del reactivo de Folin-Ciocalteu por los grupos fenólicos de los residuos de tirosina presentes en la mayoría de las proteínas, actuando el cobre como catalizador. El principal constituyente del reactivo de Folin-Ciocalteu es el ácido fosfomolibdotúngstico de color amarillo, que al ser reducido por los grupos fenólicos da lugar a un complejo de color azul intenso (Figura 12).

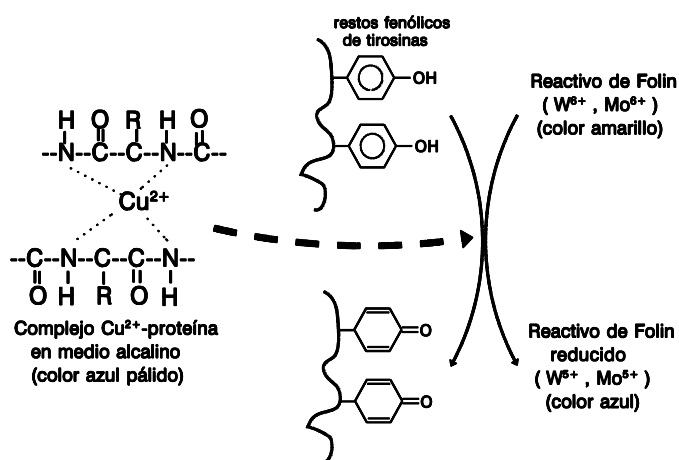


Figura 12. Reacción entre la proteína y los reactivos en la reacción de Lowry (Lowry, 1951).

Materiales.

- Fracciones concentradas (fosfoproteína y glicoproteínas)
- Kit RC-DC Protein Assay de BIO-RAD.
- Microplacas de 96 carriles.
- Multidetector©.
- Centrífuga.

Protocolo para la cuantificación de proteína empleando el Kit RC-DC Protein Assay de BIO-RAD.

1) Preparar un volumen adecuado de reactivo A':

Adicionar 20 μL de "S" por cada 1 mL de "A", cada carril requiere 25 μL de esta mezcla (estable por una semana, en caso de presentar precipitado calentar a 37° y mezclar).

2) Si existe un medio que interfería:

- Agregar 25 μL de reactivo I.
- Agregar 25 μL de reactivo II.
- Centrífugar a 15 000 g durante 10 min.
- Remover el sobrenadante.

*Si las muestras no contienen detergente omitir el paso descrito.

3) Preparar una curva estándar de 0-14 μg de proteína a partir de BSA 2 mg/mL.

4) Colocar en cada carril 1-3 μL de muestra, 25 μL de reactivo "A'" y 200 μL de reactivo "B".

5) Agitar en termomezclador e incubar por 15 min.

6) Leer absorbancia a 620 nm.

* Se utiliza el Detector de microplaca (DTX 880 Multimodetector Beckman Couter).

Se debe de realizar el ajuste de pozo, éste ajuste se realiza calculando la diferencia del valor de absorbancia de la muestras menos el valor de absorbancia del carril vacío.

5.6 Preparación y tinción de geles SDS-PAGE.

Como ya se mencionó, la electroforesis SDS-PAGE se basa en la migración de proteínas cargadas a través de un campo eléctrico que va de negativo a positivo es utilizada como método analítico y es recomendada para muestras pequeñas (no mayores a 30 µg), el porcentaje (T%) empleado en este tipo de geles va del 5 al 15% (a mayor T menor es el tamaño de poro), siendo uno de los más empleados a 12.5%, con un espesor de 1mm.

Tabla 13. Preparación de geles SDS-PAGE al 12.5% (Qiagen, 2000).

Reactivos.	Gel separador	Gel concentrador.
Acrilamida-bis-Acrilamida. (3.0%-0.8%)	3.187 mL	801.18 µL
Amortiguador Tris pH 8.8 (separador).	2.499mL	
Amortiguador Tris pH 6.6 (concentrador).		1,5 mL
SDS 10%.	99.95 µL	60 µL
H ₂ O destilada.	4.0181 mL	3.603 mL
Persulfato de Amonio (PSA)	49.97 µL	30 µL
TEMED	8.567 µL	3.529 µL
Cantidad Total.	10 ML	6 mL

Una vez polimerizados los geles se colocan en una camara electroforética con amortiguador LAEMLI (Anexo II) y se realizan los siguientes pasos:

- Se toma un volumen de la muestra previamente cuantificada equivalente a 25 µg, depositándolo en un tubo de microcentrífuga y se mezcla en una relación 1:1 con el amortiguador de Carga (Anexo II).
- Se toma el volumen obtenido de la mezcla y se deposita en cada uno de los carriles, en uno de los carriles se agregan 7.5 µL de marcador de peso molecular. Una vez cargadas las muestras, se agrega amortiguador LAEMLI hasta cerrar el circuito.
- Las condiciones de corrimiento son baño de hielo (4°C) con un voltaje de 45 a 120 Volts incrementando cada 10 min 15 Volts hasta llegar a 120 Volts, y deteniendo la generación del campo eléctrico cuando el frente de corrida se encuentre a 1 cm del borde inferior del gel.

Tinción de geles.

Los colorantes orgánicos (bromofenol, Azul de Coomassie, etc.) fueron los primeros en emplearse para la tinción de proteínas, ya que una rápida fijación, tinción y destinción son esenciales en las proteínas pequeñas para evitar difusión. Estos colorantes se emplean en muestras donde las proteínas son abundantes (mayor a 20 µg). La tinción con Azul de Coomassie puede emplearse para la determinación de proteínas cuando son abundantes, pero no para la determinación de pureza o de proteínas traza. La unión proteína-colorante es totalmente reversible en condiciones apropiadas. Existen diferentes tipos de tinción como la plata que más sensible (menor a 10 µg) y la cual produce una coloración carmelita o negra, aunque algunas proteínas tienen colores característicos como las lipoproteínas, que tienden a colorearse de azul y algunas glicoproteínas que aparecen amarillas, carmelitas o rojas.

Tinción con Azul de Coomassie.

Materiales.

- Azul de Coomassie.
- Acetona al 5%
- Agua destilada
- Escáner.

Protocolo.

- Una vez detenida la corrida se coloca el gel en Azul de Coomassie (Anexo II) por 12 o 24 h para garantizar con esto la tinción de las bandas.
- Una vez transcurrido el tiempo, se deposita el colorante en un recipiente para su almacenamiento y su posterior reuso, el gel es lavado con agua destilada hasta que el agua presente la menor cantidad de colorante.
- Se desecha por completo el agua del último lavado y se coloca el gel en ácido acético al 5% para evitar la difusión de los péptidos.
- El gel es escaneado y analizado por LABSCANER™ o GE Healthcare™.

5.7 Análisis de geles obtenidos.

Una vez escaneados los geles, se procede a un análisis comparativo entre los carriles de fosfoproteínas de cada una de las líneas celulares mencionadas. Se procede de la misma forma para las glicoproteínas, buscando en el análisis bandas que estén presentes en el control y ausentes en las demás líneas o viceversa, y/o bandas que se encuentren con una menor intensidad cuando se comparan el control y las líneas con cáncer ya que ésto puede indicar un cambio en la expresión proteica. Una vez seleccionadas las bandas, se procede a cortarla y extraerla, es depositada en un tubo de microcentrífuga de 1.5 mL y se procede a la digestión con tripsina.

5.7.1 Protocolo para la digestión de las bandas seleccionadas con tripsina.

Materiales.

- ✓ Acetonitrilo 50%- NH_4HCO_3 50 mM.
- ✓ Acetonitrilo al 100%.
- ✓ DTT 10 mM.
- ✓ Iodoacetamida 100 mM.
- ✓ NH_4HCO_3 100 mM.
- ✓ Tripsina 25 ng/ μL .
- ✓ Solución de extracción (50% (vol/vol) acetonitrilo-5%(vol/vol), ácido fórmico).
- ✓ TFA 0.1%
- ✓ 0.1%TFA-50% acetonitrilo.

Protocolo.

- 1) Cortar bandas o manchas de proteínas (aprox. 1 mm³ o 2 mm³) con la punta de una micropipeta, colocarlos en tubos eppendorf, remover el exceso del líquido.
- 2) Agregar 500 μL de una solución de acetonitrilo 50%- NH_4HCO_3 50 mM a cada uno de los cortes, incubar a 50°C durante 5 min.
- 3) Eliminar el sobrenadante y repetir el paso anteriores hasta limpiar completamente el gel de Azul de Coomassie.

- 4) Agregar 100 μL de acetonitrilo al 100% y dejar deshidratar la piezas de los geles por 5 min a temperatura ambiente (hasta que se tornen de un color blanco opaco y reduzcan su tamaño). Repetir este procedimiento dos veces.
- 5) Remover el sobrenadante y secar completamente las piezas del gel a temperatura ambiente (tubo destapado cubierto con un paño para evitar contaminación).
- 6) Añadir 30 μL de DTT 10 mM. Dejar reaccionar por 30 min a temperatura ambiente.
- 7) Remover cuidadosamente el DTT con la punta de una pipeta y descartar el sobrenadante.
- 8) Añadir 30 μL de iodoacetamida 100 mM. Dejar alquilar por 30 min a temperatura ambiente.
- 9) Transcurrido el tiempo, remover el sobrenadante con la punta de una pipeta y descartarlo.
- 10) Añadir 200 μL de acetonitrilo 100% y dejar deshidratar las piezas por 5 min, remover el sobrenadante.
- 11) Rehidratar las piezas adicionando 200 μL de NH_4HCO_3 100 mM, incubar por 10 min. Transcurrido el tiempo remover cuidadosamente el sobrenadante y descartar.
- 12) Añadir 200 μL de acetonitrilo 100% y dejar deshidratar las piezas durante 5 min. Remover y descartar el sobrenadante.
- 13) Secar completamente las piezas del gel como se indica en el paso 5 o bien en una centrífuga con vacío por 2 ó 3 min.
- 14) Añadir 15 μL (o hasta cubrir la pieza del gel) de una solución de tripsina (25 ng/mL) y dejar que se rehidraten por aproximadamente 10 min. Mezclar ocasionalmente con vortex.
- 15) Transcurrido el tiempo, remover el excedente de la solución de tripsina y añadir 20 μL de NH_4HCO_3 50 mM (volumen suficiente para cubrir la pieza del gel. Incubar a 37°C por aproximadamente 12 h.
- 16) Retirar el sobrenadante y depositarlo en tubos de microcentrífuga

etiquetados adecuadamente.

- 17) Añadir 20 μL de NH_4HCO_3 50 mM a cada una de las piezas de los geles e incubar por 10 min, mezclando ocasionalmente con vortex. Colectar el sobrenadante y depositarlo en los tubos utilizados en el paso 16.
- 18) Añadir 20 μL de solución de extracción a las piezas de los geles, incubar por 10 min, mezclando ocasionalmete con vortex. Bajar la muestra al fondo del tubo centrífugando por 30 s. Colectar el sobrenadante en el tubo utilizado en el paso 16 (extracto total). Repetir este paso por triplicado, dejar incubar por 30, 20 y 10 min respectivamente.
- 19) Reducir el volumen del extractacto total a 20 μL evaporando en una centrífuga con vacío a temperatura ambiente.
- 20) Limpiar mediante el uso de las Zip-Tips C8 (Millipore).
 - ◆ Checar el pH del extracto, que debe ser < 4 , de no serlo bajar el pH con TFA 0.1% o con ácido fórmico.
 - ◆ Adicionar la punta de la Zip-Tip con 10 μL de acetonitrilo al 100% (tomar y tirar por triplicado) y posteriormente con 10 μL de TFA 0.1% (tomar y tirar por triplicado).
 - ◆ Tomar 10 μL de la muestra (bajar y subir por 50 veces).
 - ◆ Lavar con TFA 0.1% la Zip-Tip (tomar y tirar por 3 veces).
 - ◆ Eluir con 10 μL de una solución 0.1% TFA-50% acetonitrilo en un tubo de microcentrífuga los péptidos adheridos a la Zip.
 - ◆ Ajustar el volumen de la muestra a 20 μL con ácido acético al 1% solo en caso necesario.
 - ◆ Analizar las muestras obtenidas (fosfoproteínas y glicoproteínas) mediante LC/ESI-MS/MS.

5.7.2 Análisis de fracciones de fosfoproteínas y glicoproteínas mediante LC/ESI-MS/MS.

La ionización por electrospray (ESI/LC-MS/MA) es una técnica utilizada en espectrometría de masas para producir iones (ESI-MS). Es útil en la producción de iones a partir de macromoléculas, pues supera la propensión de éstas a fragmentarse cuando se ionizan.

El funcionamiento de esta técnica es el siguiente: se introduce la muestra disuelta en un disolvente volátil por un capilar de metal pequeño y cargado. La repulsión de las cargas eléctricas hace que el líquido salga del capilar formando una nube de pequeñas gotas (10 µm) altamente cargadas, el solvente se evapora y las moléculas del analito se aproximan, repelen y estallan las gotas cuando las cargas positivas se repelen. Este proceso se repite hasta que el disolvente se elimine y solo queden iones, los cuales son analizados por el detector.

Una vez analizado por el detector se entrega una lista a través del software o algoritmo de búsqueda MASCOT, la base de datos donde se nos indica el resultado, se nos entrega una lista de posibles proteínas las cuales se deben seleccionar con los siguientes criterios:

- 1) La proteína representada debe de ser de *Homo sapiens* (Figura 13).
- 2) Se deben presentar al menos 3 secuencias péptidicas, estas deben poseer un ion score mayor a 45 (Figura 13).
- 3) El general score debe ser mayor a 200 (Figura 13).
- 4) Las proteínas deben ser de un origen distinto a la epidermis ya que esto indica contaminación de las muestras (Figura 13).
- 5) Una vez cumplidos los cuatro puntos anteriores, las secuencias de las proteínas seleccionadas se ingresan a NetNGlyc, NetOGlyc y NetPhos, todas son herramientas obtenidas de Expasy.org. La primera es un programa capaz de predecir los sitios de glicosilación en las proteínas de eucariontes, en caso de que la proteína no posea sitios

para que la modificación se realice *in vivo* es indicado por el programa, estas mismas características están presentes en la segunda herramienta. Si los programas señalan que las modificaciones no son posibles *in vivo* se descartan como posibles biomarcadores glicosilados o fosforilados según sea el caso y si las modificaciones son capaces de realizarse *in vivo* son analizadas de acuerdo al punto 6 (Figura 14).

- 6) Una vez que se ha confirmado que las proteínas seleccionadas son de *Homo sapiens*, presentan al menos 3 secuencias peptídicas con un ion score mayor al 45 de probabilidad, tienen un general score a 200 y no son de origen epidermal se utiliza la base de datos UnitProtein para obtener información de las proteínas seleccionadas. Esta base de datos nos da varias características importantes acerca de la proteína como: peso molecular, cantidad de aminoácidos que la componen, estructura tridimensional, conformación de la estructura secundaria, ubicación celular, entre otras. Además de realizarse una búsqueda de información bibliográfica de la relación existente entre la proteína propuesta y el Cáncer de mama o algún otro tipo de cáncer. Una vez obtenida esta información se decide si la proteína identificada es un posible candidato a un biomarcador obtenido por una proteómica comparativa, estas proteínas deben de ser caracterizadas y más estudiada para que pueda ser considerada como biomarcador de Cáncer de mama.

Mascot Search Results: Protein View - Windows Internet Explorer

http://www.mattscience.com/cgi/protein_view.pl?files=/data/20100203/FmpObTeR.dat&hit=g131645&db_id=1&px=1&ave_thresh=51&ignoreionscore=

Peptide Summary Report ... Peptide Summary Report ... Mascot Search Results ...

Mascot Search Results

Protein View

Match to: g131645 Score: 499
glyceraldehyde-3-phosphate dehydrogenase [Homo sapiens]
Found in search of mas56.tmp

Nominal mass (M₀): 36031; Calculated pI value: 8.26
NCBI BLAST search of [g131645](#) against nr
Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Homo sapiens](#)

Variable modifications: Deamidated (NQ), Oxidation (M), Propionamide (C)
Cleavage by Trypsin: cuts C-term side of RR unless next residue is P
Sequence Coverage: 48%

Matched peptides shown in Bold Red

1 MGKVKVGVNG FGRIGRLVTR AAFNSGRVDI VAINDPFIDL NYVMVMFQYD
51 STHGKPHSTV KAENGLKLVN GNPTIFPQR DSKIKWQDA GAETTVFSTG
101 VFTMEKAGA HLQGAQRVI ISAPSDADAP FVGVNNGKY DSKIKIINA
151 SCTINCLAPL ARYIDENFPI VGLKNTTVVA ITATQCTYD PSGLKARDGR
201 GALKQILPAS YQAAVAVRV IPELIGLQTO HAPVVPFANY SYVDLCTLE
251 KFAKIDIDIK VYQASEGL KGLGVTHQ VVSDPNST HSTTPDAGG
301 TALNDHPVKL ISWYNEFQY SNRVVDLHAG HASKK

Show predicted peptides also

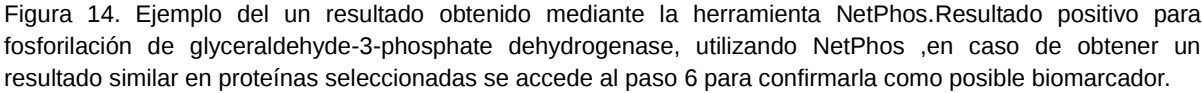
Select All Select None Search Selected Error tolerant

1. [g131645](#) Mass: 36031 Score: 499 Matches: 18 (4) Sequences: 11 (4) eMFI: 0.55
glyceraldehyde-3-phosphate dehydrogenase [Homo sapiens]
☐ Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calcd)	Delta	Miss	Score	Expect	Rank	Unique	Peptide
☒ 23	398.2390	794.4634	794.4109	0.0525	0	22	47	1		K.LTQGLR.V
☒ 26	403.2641	804.5136	804.4243	0.0894	0	52	0.059	1		K.VGVNGPGR.I
☒ 27	403.6419	805.2693	805.4083	-0.1389	0	(42)	0.51	1		K.VGVNGPGR.I + Deamidated (NQ)
☒ 35	435.7809	869.5473	869.4858	0.0615	0	37	1.3	1		K.VIPFLDQK.L
☒ 72	601.2744	1200.5342	1200.5995	-0.0653	0	61	0.0063	1		R.VVDLMAGASK.E
☒ 73	401.2558	1200.7456	1200.5995	0.1461	0	(48)	0.12	1		R.VVDLMAGASK.E
☒ 77	609.2209	1216.4273	1216.5944	-0.1671	0	(27)	18	1		R.VVDLMAGASK.E + Oxidation (M)
☒ 91	665.6894	1329.3643	1329.6421	-0.2778	1	24	26	3		R.VVDLMAGASK.E
☒ 103	706.5196	1411.8227	1410.7831	0.2397	0	50	0.061	3		R.GALKQILPASQAAK.A
☒ 120	807.3825	1612.7504	1610.8937	-0.1433	0	83	4.1e-05	1		K.LVINGNPFTIPQR.D
☒ 135	882.2732	1762.5319	1762.7951	-0.2632	0	49	0.084	1		K.LISWYNEFQYH.V
☒ 152	738.5951	2212.7634	2212.1020	0.6614	0	49	0.074	3		R.VIISAPSDADAPFVGVNNGKY
☒ 169	1298.1758	2594.3369	2594.3527	-0.0157	0	(32)	4.5	1		K.VIENPQVIGLMTTVHAIATQK.T
☒ 170	865.7995	2594.3767	2594.3527	0.0240	0	57	0.015	1		K.VIENPQVIGLMTTVHAIATQK.T
☒ 171	649.8621	2595.4191	2594.3527	1.0665	0	(49)	0.1	2		K.VIENPQVIGLMTTVHAIATQK.T
☒ 172	871.1711	2610.4913	2610.3476	0.1438	0	(42)	0.36	1		K.VIENPQVIGLMTTVHAIATQK.T + Oxidation (M)
☒ 173	653.7692	2611.0478	2610.3476	0.7002	0	(30)	6.4	1		K.VIENPQVIGLMTTVHAIATQK.T + Oxidation (M)
☒ 208	1010.1592	4036.6076	4035.8875	0.7201	0	26	11	4		K.GILGYTHQVVDSPNSDTHSSTFDAGIALNDHPVK.L

LOCUS CAA25833 335 aa linear PRI 18-APR-2005
DEFINITION glyceraldehyde-3-phosphate dehydrogenase [Homo sapiens].
ACCESSION CAA25833
VERSION CAA25833.1 GI:31645
DBSOURCE EMBL accession [X01677.1](#)
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM [Homo sapiens](#)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorhina; Catarrhini; Hominoidea; Hominidae; Homo.
REFERENCE 1
AUTHORS Hanauer, A. and Mandel, J.L.
TITLE The glyceraldehyde 3 phosphate dehydrogenase gene family: structure of a human cDNA and of an X chromosome linked pseudogene; amazing complexity of the gene family in mouse
JOURNAL EMBO J. 3 (11), 2627-2633 (1984)
PUBMED [5071958](#)
REFERENCE 2 (residues 1 to 335)
AUTHORS Arcari, P., Martinelli, R. and Salvatore, F.
TITLE The complete sequence of a full length cDNA for human liver glyceraldehyde-3-phosphate dehydrogenase: evidence for multiple mRNA species
JOURNAL Nucleic Acids Res. 12 (23), 9179-9189 (1984)
PUBMED [6096821](#)
REFERENCE 3
AUTHORS Tso, J.Y., Sun, X.H., Kao, T.H., Reese, E.S. and Wu, R.
TITLE Isolation and characterization of rat and human glyceraldehyde-3-phosphate dehydrogenase cDNAs: genomic complexity and molecular evolution of the gene
JOURNAL Nucleic Acids Res. 13 (7), 2485-2502 (1985)

Figura 13. Ejemplo del listado del programa MASCOT en la identificación de péptidos. Se observa la pagina web a la cual nos lleva el link obtenido por el análisis mediante LS/ESI-MS/MS en la que se presentan los cuatro primeros criterios para proponer un biomarcador fosforilado o glicosiado.



Capítulo 6. Resultados.

6.1 Rendimiento de la extracción de fosfoproteínas y glicoproteínas de las líneas celulares MDA-MB 231, MCF7, Hs578T y Hs578BsT.

Inicialmente necesitábamos aislar glicoproteínas y fosfoproteínas de extractos totales de proteína obtenidas de células en cultivo para su posterior separación en geles de poliacrilamida. Las células fueron cultivadas de acuerdo a la metodología descrita hasta llegar a un conteo poblacional de 10^6 células aproximadamente. Las cuales fueron procesadas de acuerdo a los protocolos de extracción. Una vez realizada la extracción se procedió a precipitar y concentrar la muestra de acuerdo a lo mencionado; es decir, en el caso de las fosfoproteínas se concentró la muestra utilizando las membranas incluidas para ultrafiltración, mientras que para las glicoproteínas se utilizó el protocolo de precipitación mediante acetona a -20°C . Ambas fracciones se reconstituyeron con un amortiguador de 10 mM Tris-Cl pH 7.4 y se cuantificó la proteína. El resultado se muestra en la Tabla 14, en ella podemos observar que la cantidad de proteína obtenida es inferior a la cantidad necesaria para los geles de 1D ó 2D, los cuales requieren de por lo menos 15 y 50 μg de proteína respectivamente. Debido a estos resultados se decidió modificar el protocolo de extracción, precipitación y concentración de las proteínas provenientes de las diferentes líneas celulares.

Tabla 14. Fracciones proteicas obtenidas de las líneas celulares MDA-MB 231, MCF7 y Hs578T por cromatografía de afinidad. Las fracciones fueron precipitadas con acetona a -20°C y reconstituidas con amortiguador Tris-Cl pH 7.4.

Línea.	MDA-MB 231	MCF7	Hs578T
Total Glicoproteínas (μg)	482.5346	285.3744	411.7798
Glicoproteínas (μg)	3.8915	6.3293	0.3221
% Glicoproteínas	0.8065	2.2179	0.0782
Total Fosfoproteínas (μg)	553.6325	411.2536	320.1923
Fosfoproteínas (μg)	nd	nd	nd
% Fosfoproteínas	0.0000	0.0000	0.0000

*nd= no se detecto

En la segunda extracción se modificó el protocolo de precipitación. El protocolo empleado fue el de cloroformo-metanol con la finalidad de eliminar las sustancias contaminantes de los amortiguadores utilizados durante la cromatografía. Una vez precipitadas, las proteínas fueron resuspendidas en un amortiguador con 7 M Urea, 2 M Tiourea, 30 mM Tris, 4% CHAPS pH 8.5, debido a que este amortiguador posee una mayor capacidad de disolución.

El resultado del rendimiento de las fracciones obtenidas se muestra en la Tabla 15, en ella se observa que las cantidades de las fracciones finales de glicoproteínas y fosfoproteínas fueron cercanas a 50 µg, cantidad suficiente para realizar una separación electroforética por duplicado empleando geles de poliacrilamida SDS-PAGE.

Tabla 15. Fracciones proteicas obtenidas de las líneas celulares MDA-MB 231, MCF7, Hs578T, Hs578BsT por cromatografía de afinidad. Las fracciones fueron precipitadas con una mezcla de cloroformo-metanol y reconstituidas en amortiguador DIGE.

Línea.	MDA-MB 231	MCF7	Hs578t	Hs578BsT
Total Glicoproteínas (µg)	1626.4344	1317.2589	1551.1831	Nd
Glicoproteínas (µg)	55.3110	50.1428	56.8852	5.8878
% Glicoproteinas	3.4008	3.8066	3.6672	Nd
Total Fosfoproteínas (µg)	1162.6068	1276.6382	1190.3846	125.4266
Fosfoproteínas (µg)	52.5534	52.0014	59.6403	5.9514
% Fosfoproteínas	4.5203	4.0733	5.0102	4.7448

*nd se detecto.

En el caso del control se obtienen rendimientos bajos, ya que el crecimiento celular utilizado fue menor al indicado en los procedimientos, pues esta línea celular tiene crecimiento muy lento. La población de células que se logró alcanzar en los cultivos era menor a 10^4 células por caja petri, sin embargo se decidió procesarlas ya que si se dejaba incubando el cultivo más tiempo, las células entraban en el proceso de senescencia. A partir de varias extracciones se obtuvo la cantidad necesaria de fosfoproteínas y glicoproteínas para realizar un gel de SDS-acrilamida para ser teñido mediante la tinción de plata, la cual tiene una mayor sensibilidad.

6.2 Electroforesis en geles de poliacrilamida SDS-PAGE de las fracciones de fosfoproteínas y glicoproteínas de las líneas celulares MDA-MB 231, MCF7, Hs578T Y Hs578BsT obtenidas por cromatografía de afinidad.

Una vez que se obtuvieron glicoproteínas y fosfoproteínas en cantidades suficientes para un gel y su réplica, se cargó cada uno de los carriles con un volumen equivalente a 25 µg de proteína de cada una de las fracciones de fosfoproteínas (Figura 16) y de glicoproteínas (Figura 17). Se corrieron las fracciones de proteínas totales, las fracciones sin fosfoproteínas y las fracciones de fosfoproteínas. Las primeras dos tienen la finalidad de monitorear que las columnas separen las proteínas adecuadamente y el carril de la fracción de fosfoproteínas es el carril del cual seleccionamos bandas. Se hizo un comparativo de bandas de las diferentes fracciones pertenecientes a la misma línea celular. En las tres fracciones se pudo observar que el patrón de bandas fue distinto y que a través de la separación cromatográfica se obtuvieron proteínas fosforiladas. Por ejemplo, en el caso de las fosfoproteínas de la línea Hs578T (carril 9), se presentó una banda con un peso molecular de aproximadamente 50 kDa. Esta banda se presentó en el extracto total y con fosfoproteínas, pero no en la fracción de la que se eliminó este tipo de proteínas por cromatografía.

En el caso de las glicoproteínas observamos un resultado similar al descrito en el gel de fosfoproteínas, por ejemplo existen bandas en los carriles de glicoproteínas de todas la líneas celulares por arriba de 250 kDa, las cuales se definieron mejor una vez que la muestra pasó por la columna de afinidad. A su vez existen bandas que no aparecieron en las fracciones de glicoproteínas, por ambas observaciones se concluye que la columna de afinidad retiene sólo cierto tipo de proteínas (glicoproteínas) y deja pasar a las que no poseen residuos de carbohidratos en su estructura.

De igual modo se cargó en otro gel un volumen equivalente a 5 µg para cada una de las fracciones en existencia del control. Este gel fue teñido con nitrato de plata y se presenta como una tinción negativa (opción incluida en el programa LABSCANNER™) para tener una mayor visualización de las bandas obtenidas. Sólo se corrieron las fracciones de fosfoproteínas, glicoproteínas y totales debido a no

contarse con las demás fracciones, ya que al concentrarse y disolverse en amortiguador no se detectó la proteína (Figura 18).

FOSFOPROTEÍNAS SDS-PAGE (KIT Qiagen).

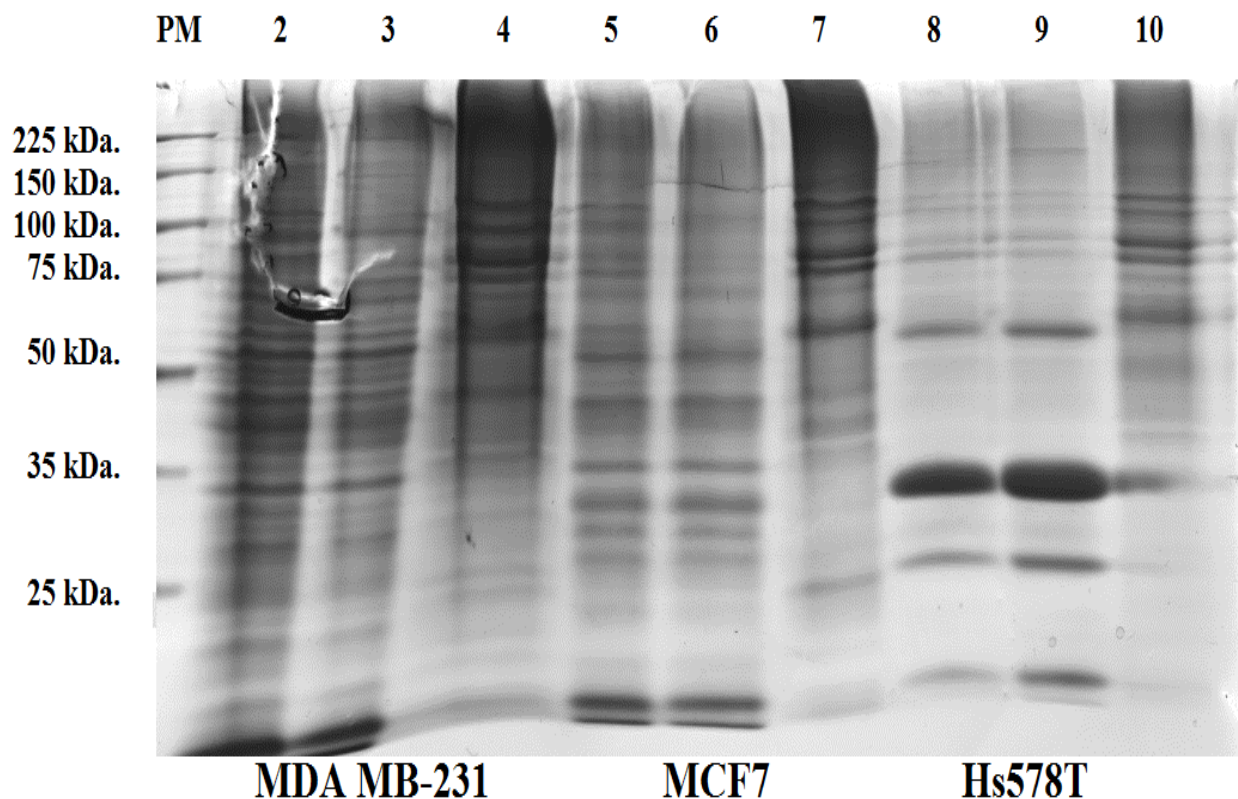


Figura 16. Gel SDS-PAGE al 12.5% de las fracciones proteicas obtenidas mediante el protocolo de extracción de fosfoproteínas de las líneas celulares MDA-MB 231, MCF7 y Hs578T. Cada carril contiene 25 µg de proteína. El gel se corrió a 120 Volts por 4 h a 4 °C, fue teñido con Azul de Coomassie y digitalizado con LABSCAN™. Los carriles 2, 5 y 8 son las fracciones totales. Los carriles 4, 7 y 10 son las fracciones sin fosfoproteínas y los carriles 3, 6 y 9 son las fracciones con fosfoproteínas de las líneas celulares.

GLICOPROTEÍNAS SDS-PAGE (KIT Qiagen)

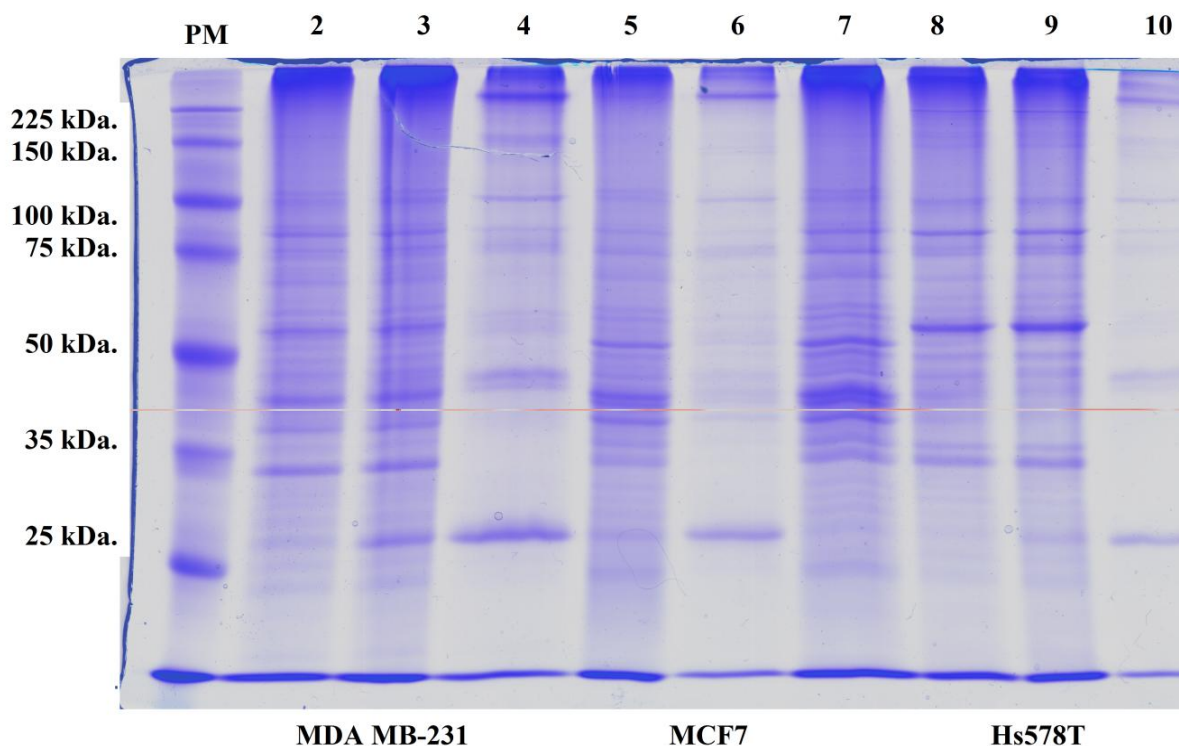


Figura 17. Gel SDS-PAGE al 12.5% de las fracciones proteicas obtenidas mediante el protocolo de extracción de glicoproteínas de las líneas celulares MDA-MB 231, MCF7 y Hs578T. Cada carril contiene 25 µg de proteína. El gel se corrió a 120 Volts por 4 h a 4 °C, fue teñido con Azul de Coomassie y digitalizado con LABSCAN™. Los carriles 3, 5 y 9 son las fracciones totales. Los carriles 2, 6 y 7 pertenecen a las fracciones sin glicoproteínas y los carriles 4, 6 y 10 son las fracciones de glicoproteínas de las líneas celulares.

CONTROL (Hs578BsT)

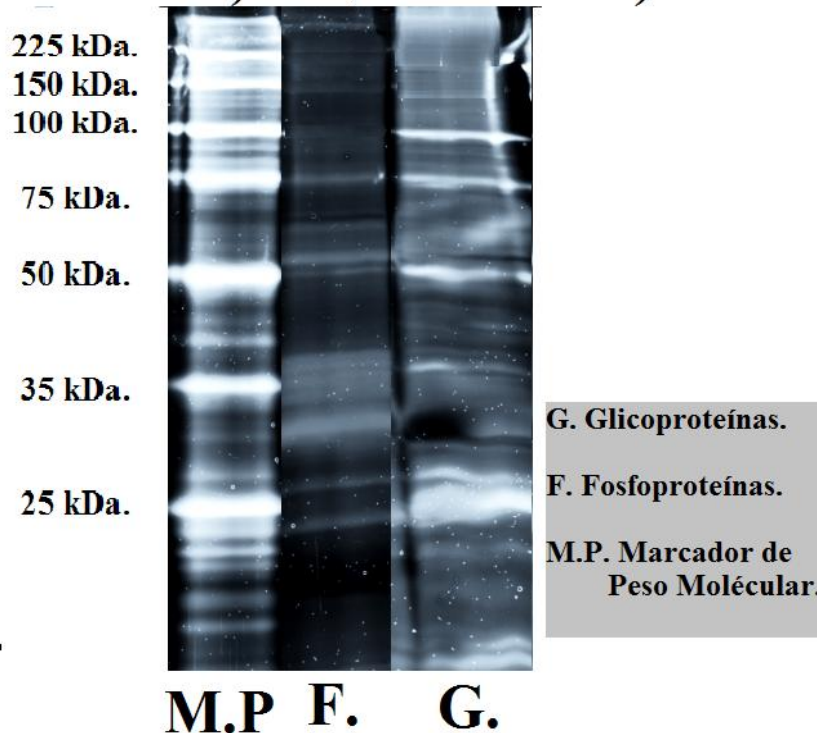


Figura 18. Gel SDS-PAGE al 12.5% de las fracciones obtenidas mediante los protocolos de extracción de glicoproteínas y fosfoproteínas de la línea celular Hs578BsT.

Los carriles del gel se cargaron con un volumen equivalente 5 μ g de las fracciones que contienen fosfoproteínas y glicoproteínas. El gel se corrió a 120 Volts por 4 h a 4 °C, fue teñido con nitrato de plata y digitalizado con LABSCAN™.

6.3 Comparación del patrón de bandas de las fosfoproteínas y glicoproteínas de las líneas celulares MDA-MB 231, MCF7 Y Hs578T en contra del patrón de bandas de las fosfoproteínas y glicoproteínas de la línea celular control Hs578BsT.

Se compararon las fracciones de fosfoproteínas y glicoproteínas de la línea celular control (Hs578BsT) con las de las líneas celulares MDA-MB 231, MCF7 Y Hs578T. Se seleccionaron únicamente las bandas que se encontraron presentes en las líneas celulares MDA-MB 231, MCF7 Y Hs578T y ausentes en la línea celular Hs578BsT (control).

6.3.1 Fosfoproteínas.

Se tomaron cada una de las imágenes correspondientes a las fracciones de fosfoproteínas y se compararon, de esta comparación se identificaron 4 bandas que no se presentaron en el control. Estas bandas tienen los siguientes pesos moleculares (Figura 19):

- 1) Cercana a los 50 kDa.
- 2) De 100 kDa
- 3) Entre los marcadores de 75 y 100 kDa.
- 4) Entre los marcadores de 100 y 150 kD.

6.3.2 Glicoproteínas.

De acuerdo a la teoría, el tipo más común de glicosilaciones *in vivo* en los organismos son las O-glicosilaciones o bien las glicosilaciones tipo mucina, las cuales tienen un peso molecular mayor a 150 kDa. Los oligosacáridos que se adhieren a las glicoproteínas afectan directamente el peso de éstas debido a su gran tamaño, es por esta razón que esperábamos obtener diferencias entre las bandas con pesos moleculares elevados (mayores de 150 kDa), pesos similares al de las mucinas, ya que son las glicoproteínas que se han relacionado con el cáncer. En los geles que se corrieron observamos bandas arriba y debajo de 150 kDa, la observación de bandas menores a 75 kDa la acreditamos, a la degradación de estas glicoproteínas originalmente de un mayor peso molecular, y que se fragmentaron por el tratamiento de obtención. La otra opción es que el carbohidrato que tienen unido no sea de peso molecular elevado. Esto lo verificaremos en los resultados de masas mediante la técnica de LS/ESI-MS/MS (Electrospray). Los pesos de las bandas señaladas son los siguientes:

- | | |
|-----------------------|---------------------|
| 1) Entre 25 y 35 kDa. | 5) Mayor a 225 kDa. |
| 2) 50 kDa. | 6) Mayor a 225 kDa. |
| 3) Entre 50 y 75 kDa. | 7) 75 kDa. |
| 4) 150 kDa. | |

FOSFOPROTEÍNAS (Qiagen)

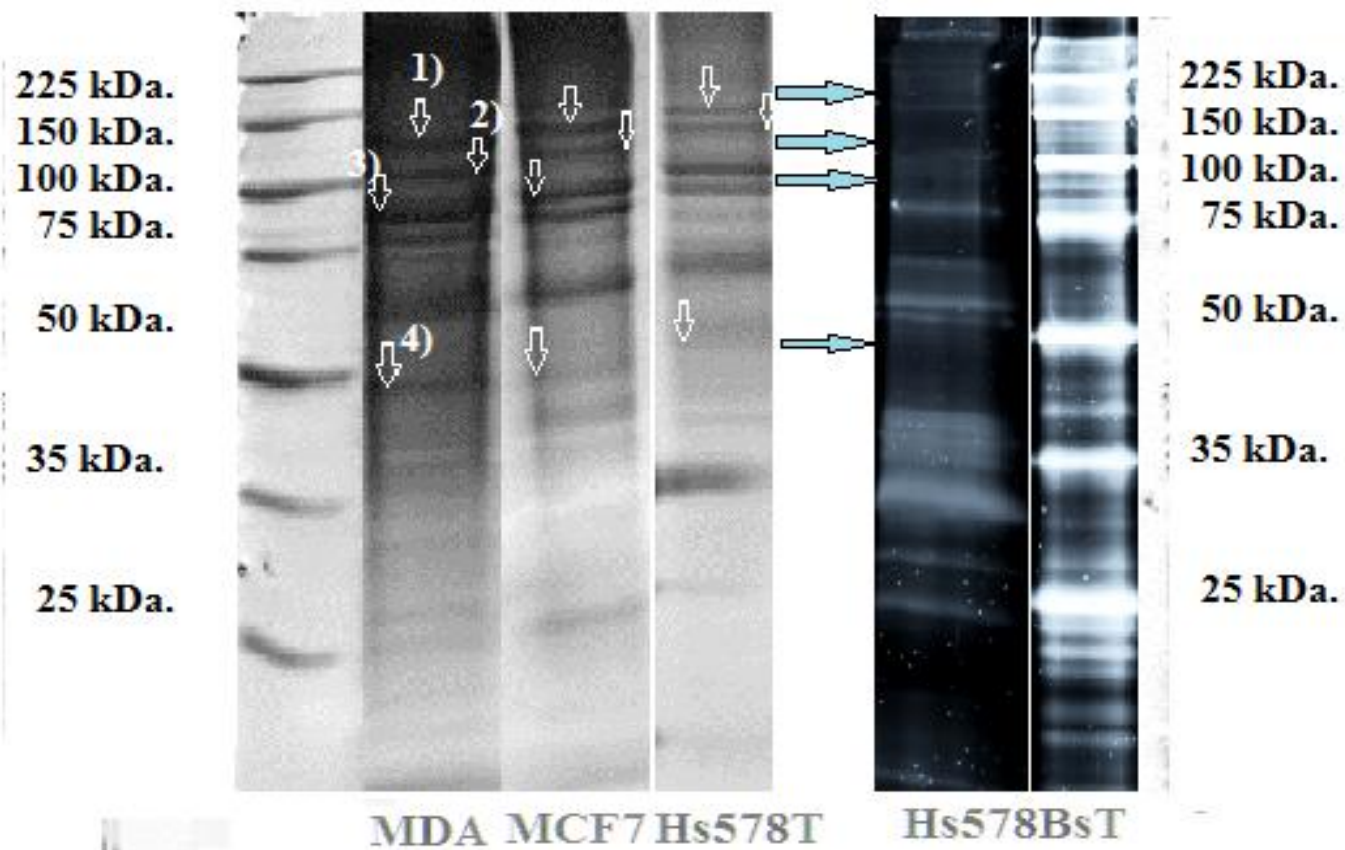


Figura 19. Bandas de fosfoproteínas de las líneas celulares MDA-MB 231, MCF7, Hs578T y Hs578BsT (línea celular control). En la figura se señalan cuatro bandas que se encuentran presentes en las líneas celulares MDA-MB 231, MCF7 y Hs578T, pero que están ausentes en la línea control Hs578B. Estas bandas fueron extraídas, digeridas y analizadas mediante LC/MS-MS/MS (Electrospray).

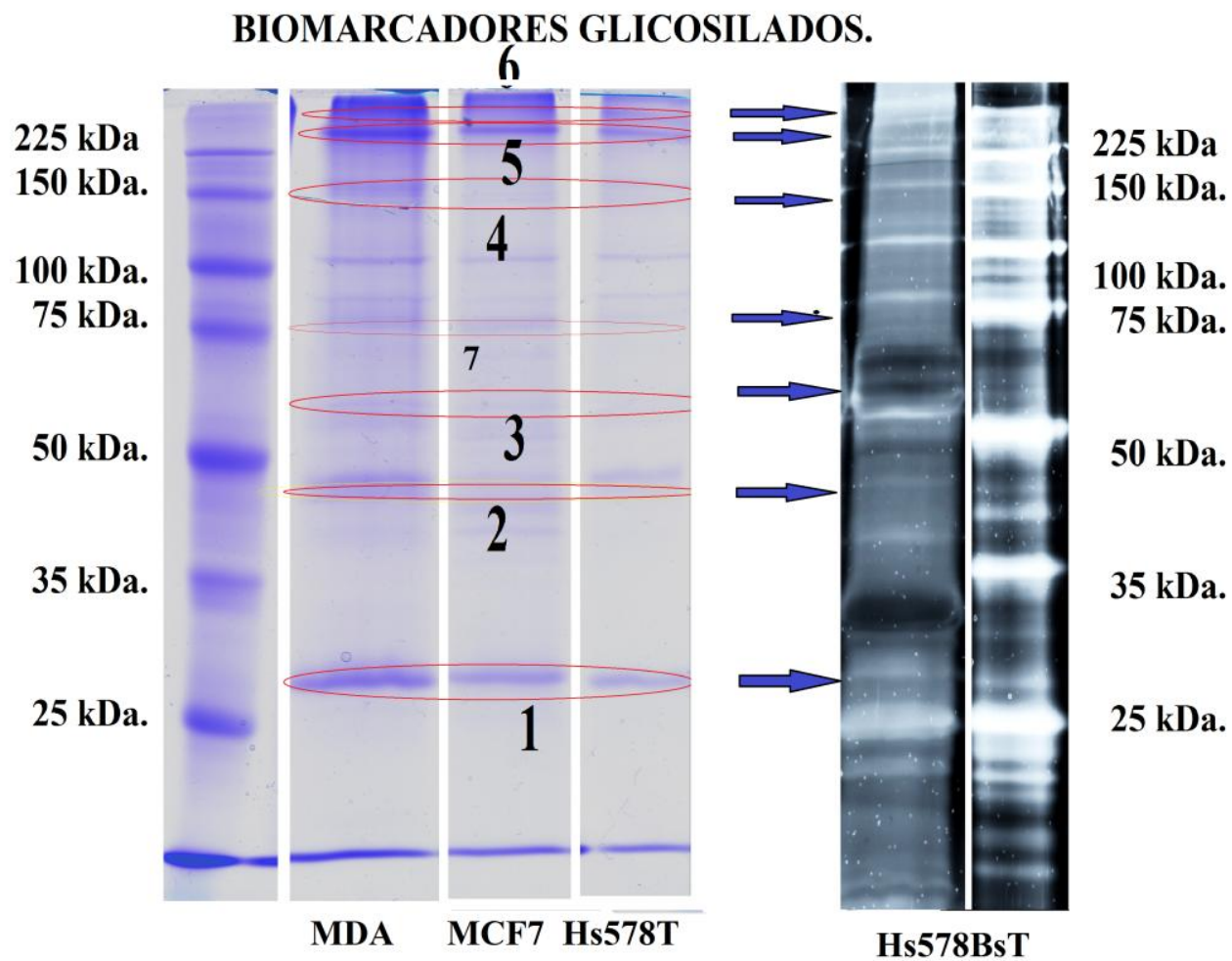


Figura 20. Bandas de glicoproteínas de líneas celulares MDA-MB 231, MCF7 Y Hs578T y Hs578BsT (línea celular control). En la figura se señalan siete bandas que se encuentran presentes en las líneas celulares MDA-MB 231, MCF7 y Hs578T, pero que están ausentes en la línea control Hs578BsT. Estas bandas fueron extraídas, digeridas y analizadas mediante LC/MS-MS/MS (Electrospray).

6.4 Resultados del análisis por LC/MS- MS/MS de las proteínas seleccionadas de los geles corridos con las fracciones obtenidas de la extracción de glicoproteínas y fosfoproteínas.

Una vez que se seleccionaron las bandas y se procesaron por LS/ESI-MS/MS, se entregó una hoja con los resultados del análisis por espectrometría de masas como la que se presenta en las Figuras 21-23 y se hace una selección de las proteínas identificadas de acuerdo a los puntos indicados en la metodología. La serie de cuadros (Figuras 21-23) corresponde a la banda número 1 de la fracción de fosfoproteínas. La Figura 23 ejemplifica una banda descartada por no posser un general mayor a 200, ni al menos 3 secuencias peptídicas con un ion score mayor a 45.

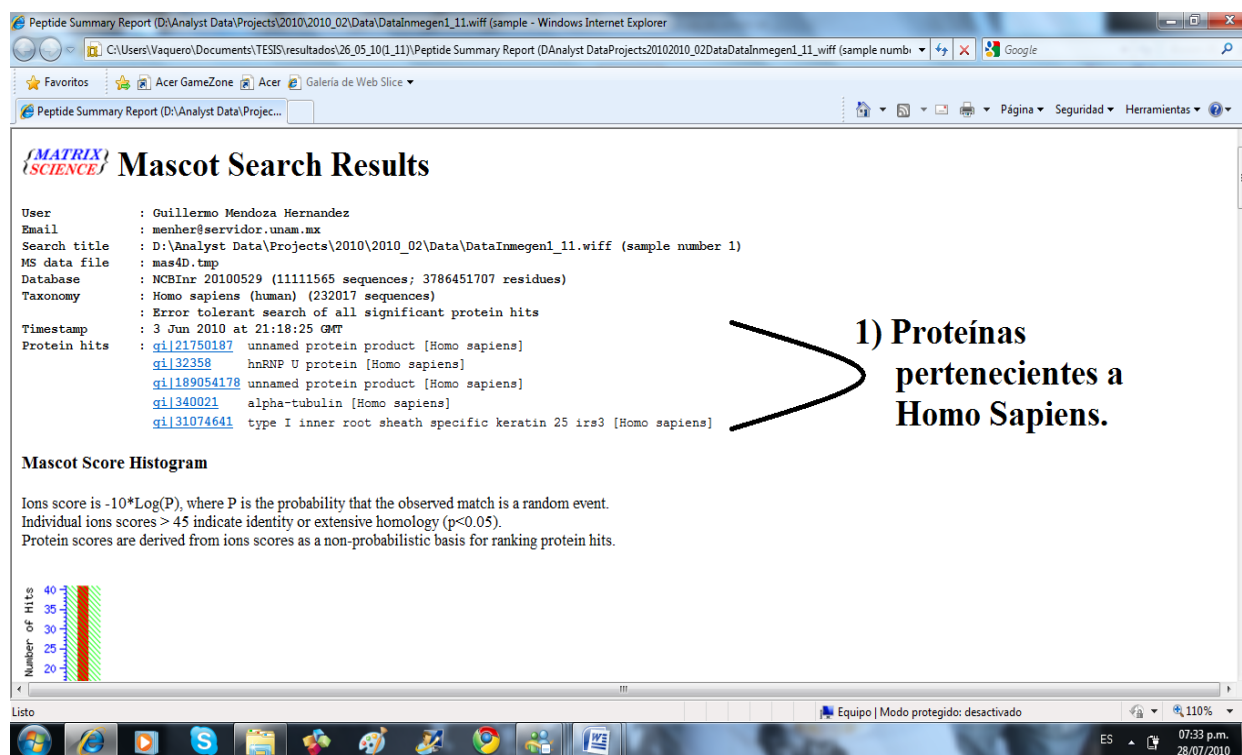


Figura 21. Hoja de resultados MASCOT del análisis por LS/ESI-MS/MS de la banda número 1 de la fracción de fosfoproteínas (1). En la Figura se muestra el resultado de la Nucleoproteína y como es que cumple con el primer punto establecido para la selección de posibles biomarcadores. Este punto se refiere a que las proteínas pertenezcan a *Homo sapiens*.



Figura 22. Hoja de resultados MASCOT del análisis por LS/ESI-MS/MS de la banda número 1 de la fracción de fosfoproteínas (2). En la figura se muestran los siguientes dos puntos necesarios para elegir un posible biomarcador. Estos son un general score mayor a 200 y 3 secuencias peptídicas con un ion score mayor a 45.

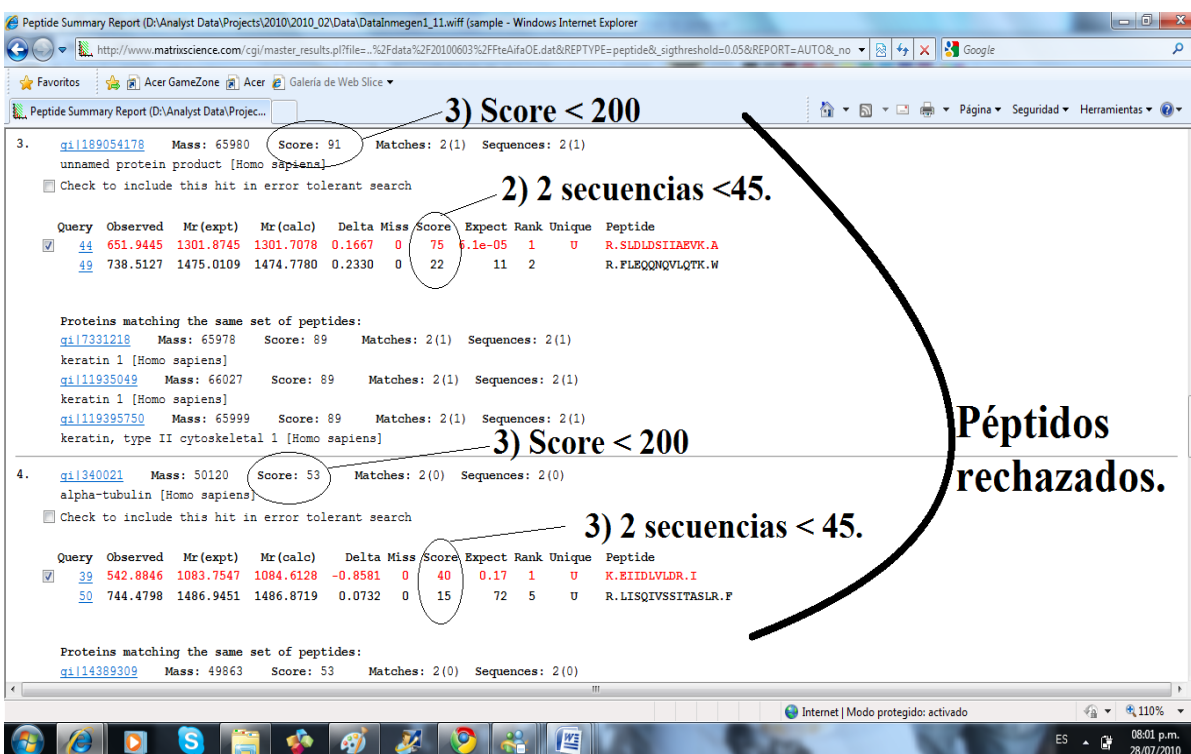


Figura 23. Hoja de resultados MASCOT de muestras identificadas como Queratina-1 y α -tubulina. En la figura se muestran resultados del análisis por LC/MS-MS/MS que fueron rechazados por no presentar un general score mayor a 200, ni poseer 3 secuencia peptídicas con un ion score mayor a 45.

Una vez seleccionada la Nucleolina (de acuerdo a los resultados de LC/MS-MS/MS), fue analizada mediante NetPhos, la cual es una herramienta diseñada para calcular la probabilidad de los residuos de aminoácidos que se pueden fosforilar (sitios de fosforilación). La herramienta entrega una página de resultados que se compone de 5 columnas: nombre de la proteína, residuo (T o S), posición, G-score, I-score, los cuales tiene un rango de 0 a 1, la columna de respuesta (Y/N). Al final hay una gráfica en donde existe una línea basal o umbral acompañada de una serie de picos correspondientes a los residuos de aminoácidos provenientes de las proteínas analizadas. Los picos que superan el umbral son los residuos capaces de fosforilarse (Figura 24). Las dos herramientas que se emplearon para la detección de sitios de glicosilación (NetOGlyc y NetNGlyc) dieron los mismos resultados.

Los puntos que utilizan las herramientas (NetPhos, NetOGlyc y NetNGlyc) para predecir si es posible o no un cambio postraducciona (fosforilación y glicosilación) son los siguientes:

- Si el G-score es >0.5 , el residuo es capaz de fosforilarse o glicosilarse. Entre más elevado el puntaje, el porcentaje de confiabilidad es mayor.
- Las treoninas con un G-score < 0.5 pero un I-score >0.5 fueron tomadas como posibles vecinos de una glicosilación o fosforilación, esta modificación no está a más de 10 residuos de aminoácidos de distancia.
- Si la secuencia no es susceptible a este cambio postraducciona, el programa lo informa.
- Al final del análisis se reporta cada uno de los scores (puntajes) totales y los mayores a 0.6 son reportados como posibles sitios de fosforilación o glicosilación en la cadena polipeptídica.

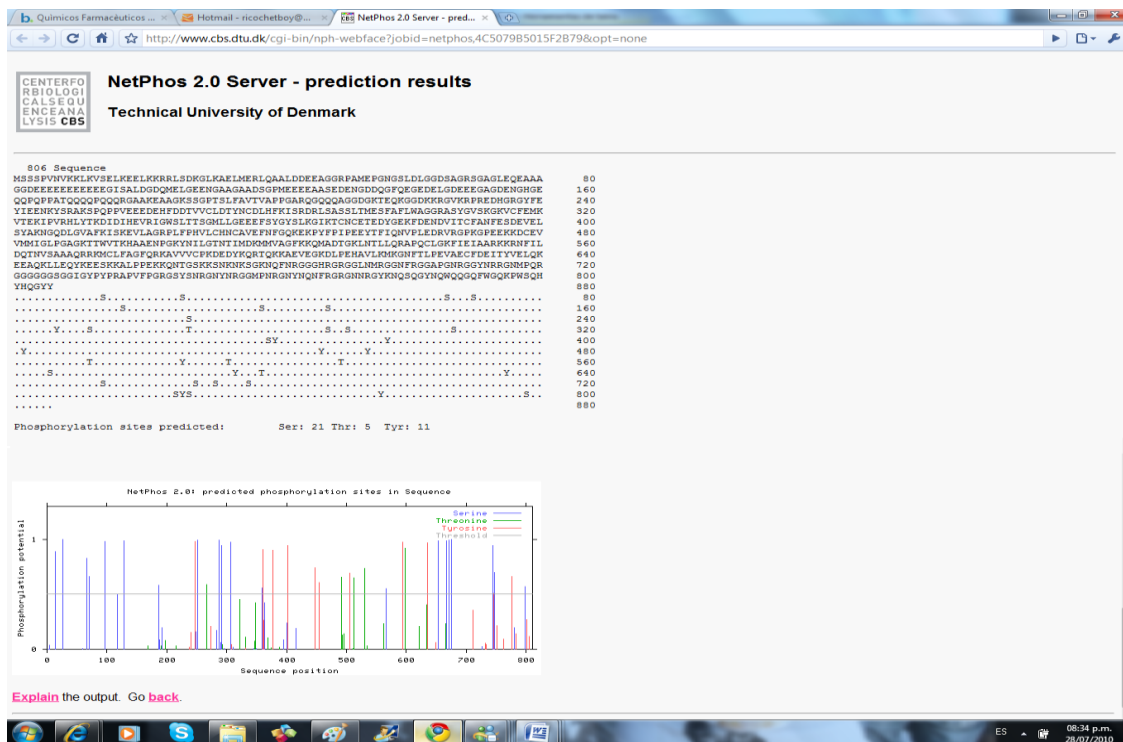


Figura 24. Hoja de resultados de NetPhos de la banda identificada como Nucleolina. En la figura observamos un resultado positivo obtenido mediante el análisis con NetPhos. Esta herramienta bioinformática nos indica el número del péptido y los péptidos que pueden sufrir una o más fosforilaciones. Los resultados se aprecian en la inferior de la figura, los picos que rebasan la línea media o umbral son los residuos de aminoácidos que pueden fosforilarse.

6.5 Selección final de posibles biomarcadores fosforilados.

En la hoja de informe que entrega la herramienta bioinformática (NetPhos) se nos presenta la cantidad de fosforilaciones que pueden ocurrir a lo largo de la secuencia peptídica, el sitio y la probabilidad de que ésta ocurra. Esto lo podemos observar en la gráfica que se nos presenta al final de la hoja de resultados; si bien este programa es capaz de señalar varios posibles lugares, es necesario aclarar que el hecho de que exista una predicción positiva, no es que la proteína necesariamente se fosfore *in vivo*. Ya que el proceso de fosforilación no sólo se lleva a cabo si existe afinidad por parte de los residuos de aminoácidos hacia un grupo fosfato; sino también por la estabilidad de la proteína, la existencia de enlaces no covalentes y covalentes, la estructura tridimensional de la proteína y a las necesidades de la célula. Para cada proteína identificada se realiza un análisis idéntico en el cual mostramos el número de aminoácidos que son susceptibles a fosforilarse y la cantidad por residuos, es decir, cuántos son serinas (S), treoninas(T) o tirosinas(Y).

La Tabla 16 nos muestra que todas las proteínas seleccionadas fueron susceptibles a fosforilarse, este resultado lo corroboramos en la base de UnitProtein. Base de datos que contiene información básica de las proteínas estudiadas. De esta base de datos tomamos las siguientes características de cada una de las proteínas seleccionadas:

- 1) El número total de aminoácidos que componen la cadena peptídica,
- 2) La ubicación celular.
- 3) Funciones principales en el metabolismo de la célula.
- 4) Las PTM (Postraductional Modifications) a las cuales son sometidas *in vivo*.

Siendo el último criterio la prueba confirmatoria de que toda la metodología ha sido empleada correctamente, es decir que la separación cromatográfica, la identificación por LC/ESI-MS/MS, y la utilización de las herramientas bioinformáticas encaminadas al análisis proteómico dieron como resultado las proteínas enlistadas en la Tabla 16.

Tabla 16. Fosfoproteínas identificadas como posibles biomarcadores mediante LC/MS-MS/MS (Electrospray) de las líneas celulares MDA-MB 231, MCF7 y Hs578T mediante proteómica comparativa.

Banda.	Peso molecular (kDa).	Proteína.	Análisis con (Expasy.org)	NetPhos.	General score de MS	Número de péptidos identificados.
1	150.	nucleolin [Homo sapiens] gi 189306	Posibles sitios de fosforilación: Ser: 28 Thr: 17 Tyr: 4		408	13(3).
		hnRNP U protein [Homo sapiens] gi 32358	Posibles sitios de fosforilación: Ser: 21 Thr: 5 Tyr: 11		370	10 (3)
2	150	heat shock protein HSP 90-beta [Homo sapiens] gi 20149594	Posibles sitios de fosforilación: Ser: 21 Thr: 8 Tyr: 11		561	17(4)
		heat shock protein HSP 90-alpha isoform 2 [Homo sapiens] gi 154146191	Posibles sitios de fosforilación: Ser: 21 Thr: 10 Tyr: 10		357	12(3)
3	100	protein disulfide-isomerase A4 precursor [Homo sapiens] gi 4758304	Posibles sitios de fosforilación: Ser: 19 Thr: 8 Tyr: 8 No se han reportado fosforilaciones <i>in vivo</i> .		1376	40(18)
4	75	protein disulfide-isomerase precursor [Homo sapiens] gi 20070125	Posibles sitios de fosforilación: Ser: 13 Thr: 6 Tyr: 4 No se han reportado fosforilaciones <i>in vivo</i> .		946	28 (7)
		Chain A, Human Protein Disulfide Isomerase, Nmr gi 159162689	Posibles sitios de fosforilación: Ser: 1 Thr: 1 Tyr: 1 No se han reportado fosforilaciones <i>in vivo</i> .		315	7(3)

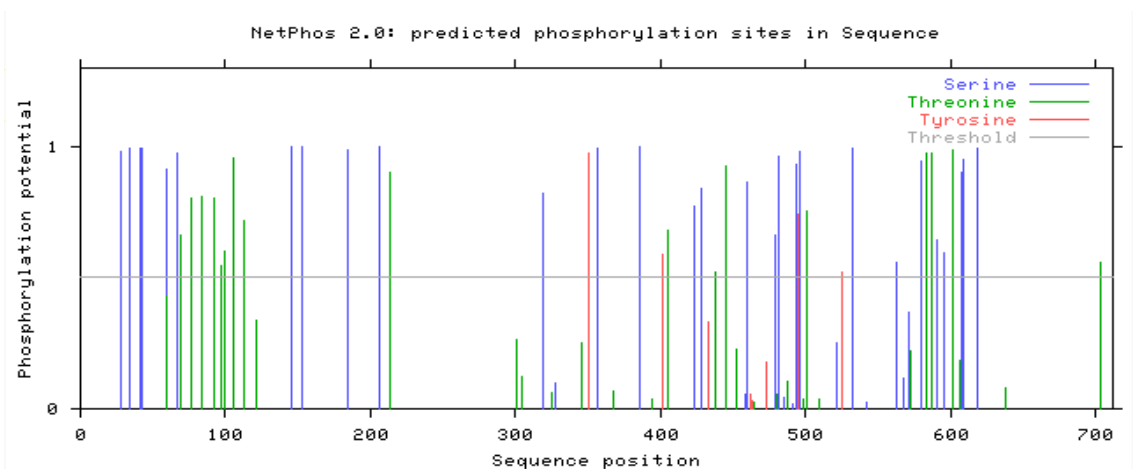


Figura 25. Sitios de fosforilación de la nucleolina (NetPhost, Expasy.org).

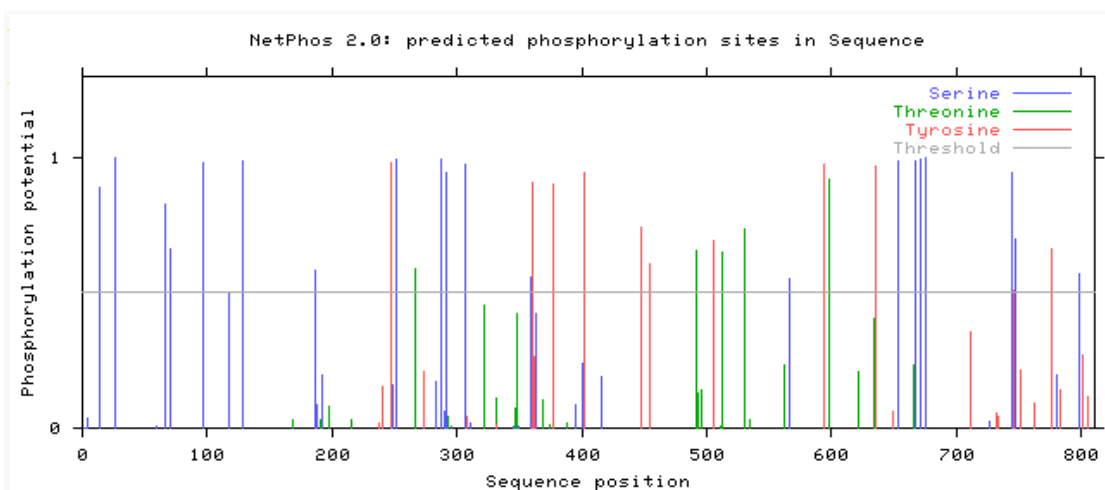


Figura 26. Sitios de fosforilación de la Ribonucleoproteína U (hnRNP U) (NetPhost, Expasy.org).

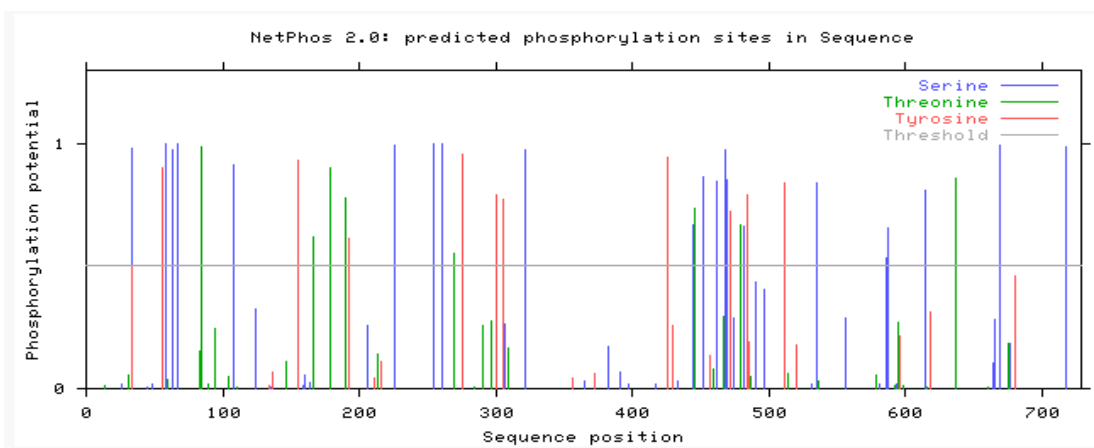


Figura 27. Sitios de fosforilación de la proteína de choque térmico HSP 90- α (NetPhost, Expasy.org).

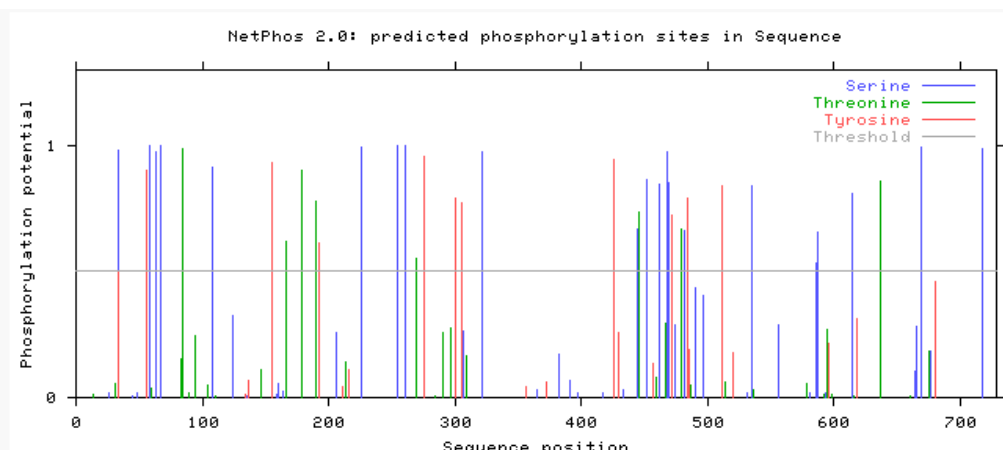


Figura 28. Sitios de fosforilación de la isoforma 2 de la proteína e choque térmico HSP 90- β (NetPhost, Expasy.org).

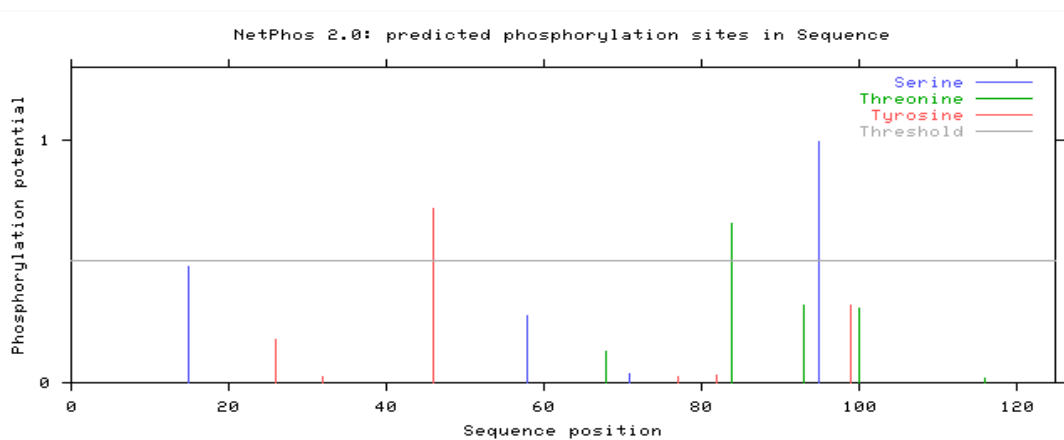


Figura 29. Sitios de fosforilación de disulfuro Isomerasa (NetPhost, Expasy.org).

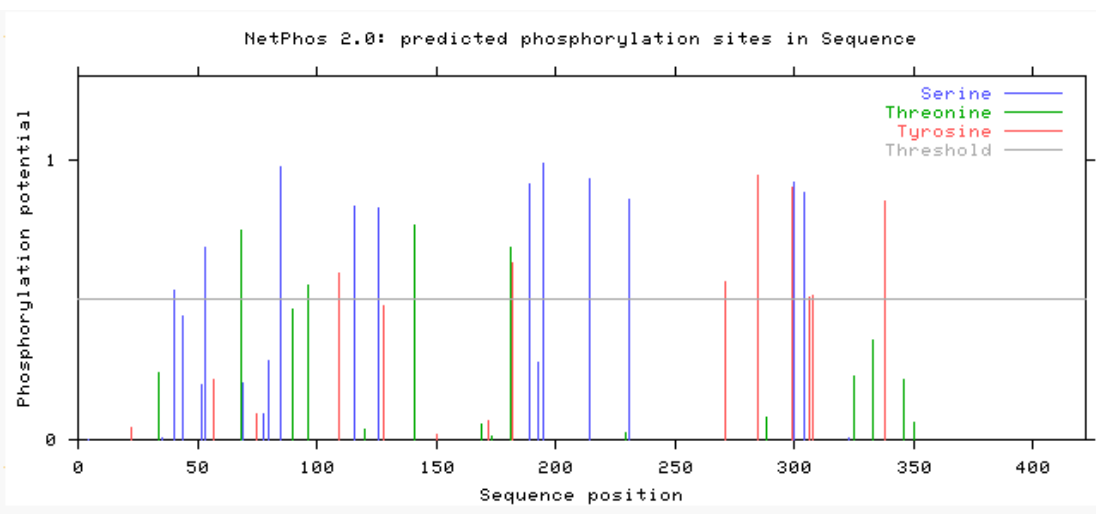


Figura 30. Sitios de fosforilación de la calreticulina (NetPhost, Expasy.org).

6.7 Selección final de posibles biomarcadores glicosilados.

Se repitió el mismo procedimiento empleado en la selección de biomarcadores fosforilados con las bandas seleccionadas de glicoproteínas. Solamente se cambió la herramienta utilizada, en lugar de utilizar Netphos se emplearon dos herramientas, la primera fue NetNGlyc, capaz de deducir *N*-glicosilaciones en secuencias peptídicas de humano y la segunda, NetOGlyC, la cual deduce las posibles *O*-glicosilaciones en proteínas de humano.

Los resultados se enlistan en la Tabla 17 y de acuerdo a la bibliografía⁷. El programa no detectó ninguna proteína capaz de modificarse mediante una *N*-glicosilación. No podemos decir con ésto que ninguna proteína enlistada sea capaz de sufrir este tipo de modificaciones, pero la probabilidad de que ocurran es mínima, razón por la cual el programa no las señala. En la Tabla 17 podemos observar que solo dos proteínas la CD44 y la proteína de 80 kDa son capaces de llevar acabo *O*-glicosilaciones en residuos de tirosina y serina. En el caso de la Subunidad α -VLA-3 y el precursor de la reticulocalbina, el programa nos señala que no existe la probabilidad de que las *O*-glicosilaciones puedan ser llevadas a cabo, debido a que no existen sitios en los cuales se puede llevar la unión residuo-carbohidrato. Por esta razón se excluyeron de la discusión de resultados.

Tabla 17. Glicoproteínas identificadas como posibles biomarcadores mediante LC/MS-MS/MS (Electrospray) de las líneas celulares MDA-MB 231, MCF7 y Hs578T mediante proteómica comparativa.

Banda.	Peso molecular.	Proteína.	N-glicosilaciones	O-glicosilaciones.	General score de MS	Número de péptidos identificados.
1	28 kDa.			Descartada		
2	46 kDa.	reticulocalbin -1 precursor [Homo sapiens] gi 4506455	<i>In vivo</i> no se presentan N-glicosilaciones.	el programa no señala posibles sitio de glicosilación	211	4 (2)
3	60 kDa.			Descartada.		
4	150 kDa.	VLA-3 subunit [Homo sapiens] gi 220141	<i>In vivo</i> no se presentan N-glicosilaciones.	<i>In vivo</i> no se presentan O-glicosilaciones.	267	3(1)
5	>225 kDa.	CD44E (epithelial form) [Homo sapiens] gi 29801	<i>In vivo</i> no se presentan N-glicosilaciones.	8 posibles sitios de O-glicosilaciones todos ellos en residuos de tirosina.	211	3(3)
		VLA-3 subunit [Homo sapiens] gi 220141	<i>In vivo</i> no se presentan N-glicosilaciones.	<i>In vivo</i> no se presentan O-glicosilaciones.	260	5(3)
6	>225 kDa.	CD44E (epithelial form) [Homo sapiens] gi 29801	<i>In vivo</i> no se presentan N-glicosilaciones.	8 posibles sitios de O-glicosilaciones todos ellos en residuos de Tirosina.	280	5(2)
7	75 kDa.	80K-H protein [Homo sapiens] gi 182855	<i>In vivo</i> no se presentan N-glicosilaciones.	10 posibles sitios de O-glicosilación, 5 en residuos de tirosina y 5 en residuos de serina.	286	6(3)

*Se descartaron las proteínas obtenidas en el análisis mediante LC/ESI-MS/MS ya que no cumplieron con los parámetros establecidos.

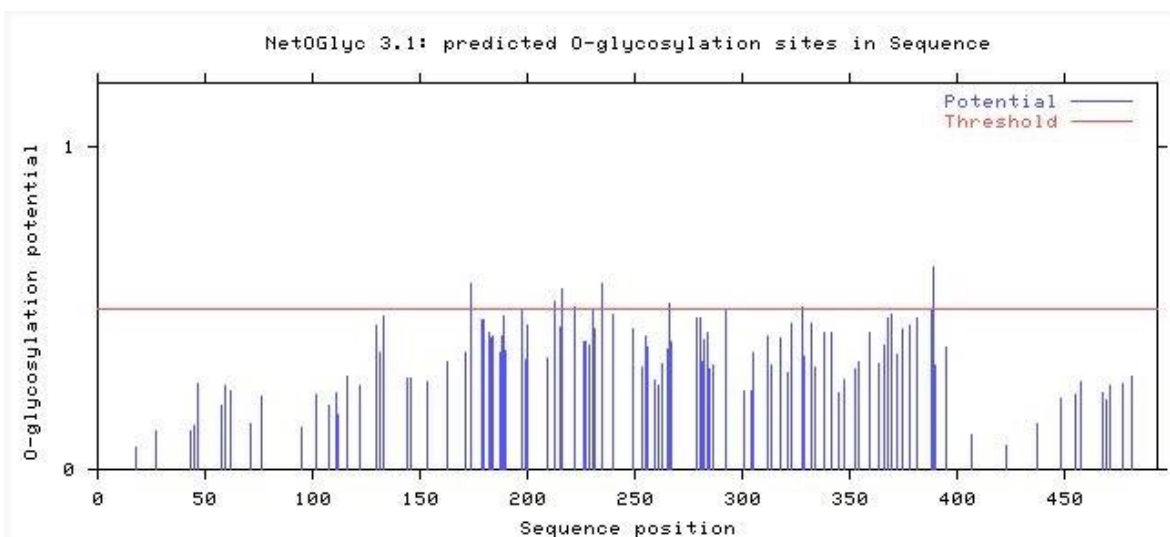


Figura 31. Posibles sitios de Glicosilación de la CD44E (NetOGly, Expasy.org).

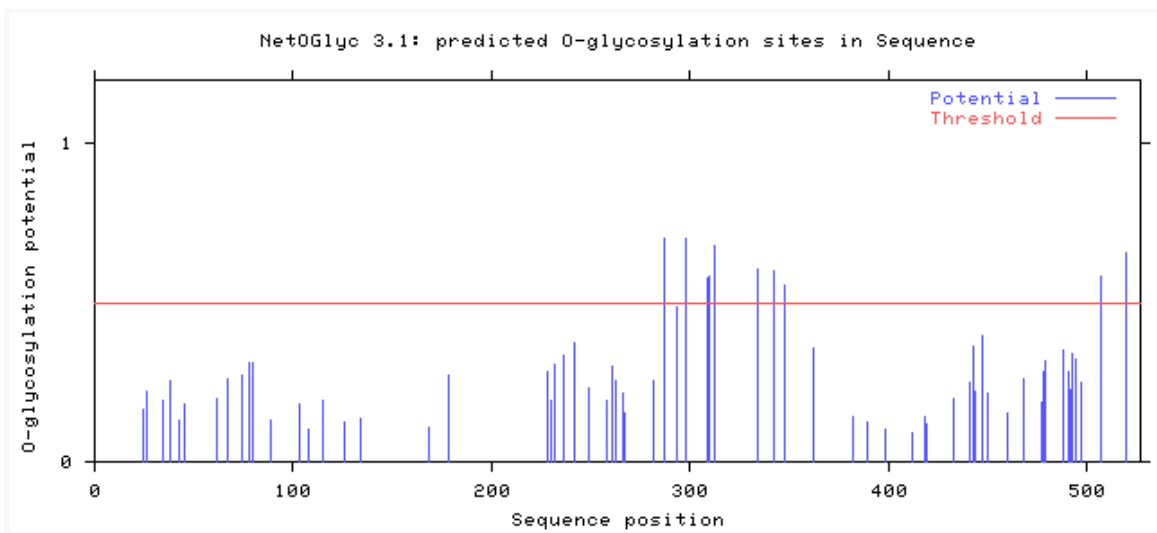


Figura 32. Posibles sitios de Glicosilación de la proteína 80K-H (NetOGly, Expasy.org).

Capítulo 7. Discusión. Biomarcadores fosforilados y glicosilados seleccionados.

En las últimas décadas, las técnicas proteómicas han tenido un gran desarrollo, pues han demostrado ser una herramienta invaluable en la identificación y caracterización de diversas biomoléculas como las proteínas. Un claro ejemplo de la importancia del desarrollo de las técnicas proteómicas son los diversos estudios en las que han sido utilizadas como herramientas principales en sus metodologías. Un elevado porcentaje de estos estudios, tienen como objetivo principal la búsqueda de diferencias en los niveles de expresión proteómico de individuos sanos respecto a individuos con alguna patología crónica, es decir, la búsqueda de biomarcadores proteicos que puedan ser utilizados para el diagnóstico y/o la terapia de estas enfermedades. Además de poder emplearse como biomarcadores de pronóstico de la enfermedad.

Uno de los errores más recurrentes en la identificación de los primeros biomarcadores fue la baja confiabilidad que presentaban las técnicas proteómicas utilizadas, la cual se adjudicó al pobre desarrollo y poco entendimiento de las técnicas empleadas. Se ha realizado un gran esfuerzo por depurarlas y optimizarlas, con la finalidad de ofrecer una mayor confiabilidad y disminuir los tiempos de análisis. La única forma en la que ésto se ha logrado, es a través de la combinación de dos o más técnicas proteómicas. Esta combinación tiene la finalidad de minimizar cualquier desventaja que pueda presentar el empleo de una sola técnica. Es por esta razón que para el desarrollo de la parte experimental de este trabajo se emplearon distintas técnicas proteómicas propuestas en las Figuras 8 y 11.

Una célula es capaz de producir hasta 5 000 proteínas en un momento y espacio determinado. Estas proteínas pueden sufrir uno o varios cambios postraduccionales, y sólo entre el 10 y 15% de estos cambios corresponden a la adición de un grupo fosfato o un oligosacárido, es decir, sólo entre un 10 y 15% son fosfoproteínas o glicoproteínas. Es por lo que, la primer técnica empleada en este trabajo, fue una separación de proteínas mediante cromatografía de afinidad para

cada uno de los grupos de interés (fosfoproteínas y glicoproteínas). La separación cromatográfica eliminó todas aquellas proteínas que no cumplieran con la característica de estar fosforiladas o glicosiladas. Ésta nos dio como resultado la reducción de más de un 95% (Tabla 15) respecto a los totales de cada una de las fracciones obtenidas para las fosfoproteínas y glicoproteínas.

Las fracciones obtenidas mediante cromatografía de afinidad fueron separadas de acuerdo a su peso molecular utilizando geles de acrilamida en una electroforesis SDS-PAGE. Los geles resultantes fueron digitalizados utilizando el programa LABSCANER [™] (Figuras 16, 17 y 18). De éstos se tomaron las fracciones que corresponden a las fosfoproteínas de las líneas celulares cancerosas (MCF7, Hs578T y MDA-MB 231) y la fracción de fosfoproteínas de la línea control Hs578BsT (Figura 19), con la finalidad de realizar un análisis comparativo del patrón de bandas (Figura 19 y 20).

En la Figura 19 podemos observar como existe una diferencia del patrón de bandas que corresponde a la fracción de fosfoproteínas de la línea celular control (Hs578BsT) respecto al patrón de bandas de las líneas celulares cancerosas (MDA-MB 231, Hs578BsT y MCF7). El resultado de este análisis comparativo fue la identificación de 4 bandas presentes en las líneas celulares cancerosas pero ausentes en la línea celular control (Figura 19). Este mismo procedimiento se realizó con las imágenes digitalizadas de los geles de las glicoproteínas (Figura 20). El resultado del análisis del patrón de bandas de las fracciones de las glicoproteínas, fue la identificación de 7 bandas que se encuentran presentes en las fracciones de glicoproteínas de las líneas celulares cancerosas pero ausentes en la fracción de la línea celular control (Hs578BsT).

Las bandas identificadas en las Figuras 19 y 20 fueron ubicadas en los geles de acrilamida. Una vez ubicadas estas bandas, se extrajeron y digirieron mediante tripsina modificada (Promega Madison) para ser analizadas por LS/ESI-MS/MS. Este análisis fue realizado con un espectrómetro de masas híbrido 3200 Q-TRAP en Tándem (Applied Biosystems/MDS Acix) equipado con un nano spray. Las muestras fueron inyectadas a una columna C₁₈ UPLC de 0.7 µm (5µm, 180 µm x 20 mm) para ser desaladas. El desalado de las proteínas optimizó la ionización de los

péptidos cuando fueron introducidos a los espectrómetros en Tándem (MS/MS). Esta optimización disminuyó el ruido en el set de picos obtenidos. Lo cual nos proporciona una mayor confiabilidad de los resultados obtenidos, pues es más sencilla la identificación de coincidencia de los fragmentos proteicos identificados experimentalmente y los fragmentos que pertenecen a las secuencias reportadas que realiza el algoritmo de búsqueda MASCOT. Una vez desaladas las muestras son introducidas a los espectrómetros en Tándem (MS/MS) para ser fragmentadas. Los fragmentos proteicos obtenidos son capturados y analizados mediante una trampa de electrones lineal Q3 (Q3 Linear Ion Trap).

La interpretación de los set de picos obtenidos por MS/MS se realizó utilizando un algoritmo de búsqueda MASCOT (Versión 1.6 b9 Matrix Science, London.UK). En esta interpretación se establecieron los siguientes parámetros de confiabilidad de concordancia entre los fragmentos proteicos de las proteínas identificadas experimentalmente y las secuencias reportadas en Sub-set de proteínas Humanas del Centro Nacional de Información Biotecnológica (<http://www.ncbi.gov>, NCBI):

- 1) De los fragmentos proteicos identificados al menos 3 deben de poseer una probabilidad de coincidencia calculada mayor a 0.05, es decir, deben de existir al menos 3 secuencia peptídicas identificadas con un ión score mayor a 45. Un ión score mayor a 45 nos indica que la probabilidad de que el fragmento identificado pertenezca a la secuencia reportada es elevada.
- 2) La probabilidad de homología o identidad debe ser mayor a 200, es decir, el general score debe ser mayor a 200. Sí el general score de los fragmentos proteicos identificados es mayor a 200 la probabilidad de que los fragmento pertenezcan a la secuencia reportada es elevada.

Las secuencias peptídicas que cumplan con estos parámetros, tienen una probabilidad elevada de coincidencia o similitud respecto a las secuencia peptidicas reportadas en las base de datos de NCBI.

A continuación se presenta un listado de las proteínas seleccionadas como posibles biomarcadores, las cuales cumplen con los parámetros de confiabilidad establecidos por el algoritmo de búsqueda MASCOT y los establecidos en la metodología.

Nucleolina (NCL).

Del análisis por LS/ESI-MS/MS de la banda número 1 de los geles de fosfoproteínas (Figura 19), empleando un algoritmo de búsqueda MASCOT; se identificaron un total de 13 fragmentos que pertenecen a la secuencia proteica reportada en la base de datos NCBI como Nucleolina (NCL) del *Homo sapiens*. Los 13 fragmentos identificados representan el 15% del total de la secuencia proteica de la Nucleolina (NCL).

El general score de los fragmentos fue de 408, lo que significa, que la probabilidad de homología de éstos es elevada, respecto a los fragmentos que componen la secuencia de la Nucleolina (NCL). El general score mínimo establecido de coincidencia entre las proteínas de la base de datos NCBI y la obtenida experimentalmente es 200.

De los 13 fragmentos proteicos identificados, 4 poseen un ion score mayor a 45, es decir, un porcentaje de probabilidad de coincidencia mayor a 0.05. Cumpliendo así con los parámetros de confiabilidad establecidos por el algoritmo de búsqueda MASCOT.

La nucleolina es una fosfoproteína nucleolar presente en los organismos eucariontes, la cual se encuentra relacionada con la síntesis y maduración de los ribosomas. Esta proteína se localiza en las regiones fibrilares más densas dentro del núcleo. El gen que codifica para esta proteína se compone de 14 exones y 13 intrones, los cuales representan aproximadamente 11 Kb, el producto proteico de esta secuencia es un polipéptido de 710 aminoácidos. La nucleolina contiene 13 hojas β -plegadas, 4 α -hélices (Figura 33) y es susceptible a 3 tipos de modificaciones postraduccionales: acetilación, metilación y fosforilación. Esta última puede ocurrir en 49 sitios posibles (Figura 25), pero existen reportes en los cuales se ha observado que esta proteína se puede encontrar glicosilada cuando se encuentra en la membrana celular²⁶.

La nucleolina tiene la capacidad de unirse al ADN y ARN, sus ubicaciones principales dentro de la célula son: el núcleo, nucléolo y el citoplasma.

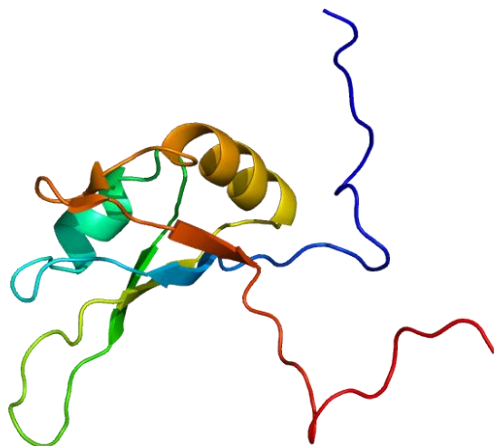


Figura 33. Estructura 3D de la nucleolina (Protein Data Bank, <http://www.pdb.org/pdb/home/home.do>).

La nucleolina se encuentra relacionada ampliamente con el crecimiento celular en células eucariontes, se le ha encontrado asociación con la cromatina intranucleolar y partículas preribosomales. Esta proteína induce o colabora con la descondensación de la cromatina debido a una interacción con la histona H1. Se ha reportado que juega un papel en la transcripción del pre-rARN y el ensamblaje del ribosoma, y en el proceso de la elongación transcripcional. La unión al RNA lo hace con el extremo 5'-UUAGGG-3' en el ADN con 5'-TTAGGG-3'. Gracias a técnicas inmunohistoquímicas se le ha relacionado con la expresión de BRCA1, proteína que se encuentra relacionada con el daño de ADN. Aquí se observó que BRCA1 co-localizaba con la nucleolina fosforilada, teniendo una relación de co-expresión durante las fases G1 y S del ciclo celular, y una tasa de incremento en el paso de estas fases, lo cual repercute en una mayor proliferación celular. Ésto es importante, ya que el paso de G1 a S es crítico para seguir adelante en la duplicación celular y quizá uno de los más estrictos pues de aquí las células que tuvieron problemas en la replicación de su material genético son llevadas a apoptosis, impidiendo así una acumulación de fallas en el ADN. Esta correlación se pierde cuando ingresan a la fase G2 y M, en donde se rompe la interacción; este comportamiento se observó al realizar un experimento en el cual se integraron pequeños RNA de Interferencia con la finalidad de disminuir la expresión de la nucleolina; al ocurrir esto, la cantidad de BCRA1 disminuyó y con ésto, el crecimiento celular descontrolado²⁷, confirmando así que la presencia de este

complejo coadyuva al incremento de la tasa de proliferación.

Existen reportes que señalan que la nucleolina se expresa de forma simultánea con los RE (receptor de estrógenos) en células cancerosas. Ésto se pudo observar mediante inmunofluorescencia y el uso de la citometría de escáner (LSC) ²⁸. En los resultados obtenidos de este experimento se observó que la nucleolina en las fases G2 y M se encuentra de forma libre en los mismos niveles tanto en células RE positivas como en células RE negativas, pero en la fase S en las células RE negativas aumenta esta forma, mientras que en las RE positivas disminuye, lo que nos indica que hay relación entre la unión de la nucleolina con moléculas que aumentan la transcripción de BCRA1 y ER ²⁹, la expresión de los RE está relacionada con los carcinomas más agresivos durante los primeros 5 años del padecimiento y después se pierde esta correlación. Además es uno de los primeros biomarcadores empleados para el diagnóstico de Cáncer de mama y hoy en día es utilizado también como biomarcador pronóstico. Las vías en las cuales participa la nucleolina en estos procesos son desconocidas, pero se sabe que son llevadas a cabo en el núcleo y el citoplasma.

La nucleolina es transportada específicamente a la superficie de las células endoteliales en proliferación donde su función principal esta pobremente definida, aunque se sabe que participa en la angiogénesis. Una vez que se une con su ligando es activada mediante fosforilación, este comportamiento se observó por la unión de receptores con antiNLC, lo que ocasionó que se internalizaran vasos sanguíneos en la célula causando un aporte mayor de oxígeno, desencadenando la apoptosis de la célula. Esto sugiere a esta proteína como un blanco para terapia antiangiogénica³⁰.

La NCL es una candidata ideal para caracterizar las vías en las cuales interviene y con esto ver su relación con otras proteínas que participan en el mismo mecanismo de acción en la enfermedad. Para poder ser propuesta como un posible biomarcador pronóstico, diagnóstico o terapéutico.

Ribonucleoproteína U (hnRNP U).

Del análisis por LS/ESI-MS/MS de la banda número 1 de los geles de fosfoproteínas (Figura 19), empleando el algoritmo de búsqueda MASCOT; se

identificaron 10 fragmentos proteicos que pertenecen a la secuencia proteica reportada en la base de datos NCBI como la Ribonucleoproteína (hnRNP U) del *Homo sapiens*. Los fragmentos identificados representan el 13% total de la secuencia reportada como la Ribonucleoproteína (hnRNP U).

El general score es de 370 puntos, es decir 170 puntos arriba del mínimo establecido (200) por el algoritmo de calculo MASCOT. Un general score mayor a 200, significa que la probabilidad de coincidencia de los fragmentos proteicos identificados experimentalmente es elevada respecto a los fragmentos que componen la secuencia de la Ribonucleoproteína U (hnRNP U).

De los 10 fragmentos proteicos identificados, 4 poseen un ion score mayor a 45, es decir, tienen una probabilidad de coincidencia mayor a 0.05, y de acuerdo a los parámetros establecidos por el algoritmo de búsqueda MASCOT, nos indican una probabilidad de coincidencia elevada entre éstos y los fragmentos que componen la secuencia proteica de la Ribonucleoproteína U (hnRNP U).

La Ribonucleoproteína U (hnRNP U) pertenece a una subfamilia de ribonucleoproteínas (hnRNPs), las cuales forman complejos con ARN nuclear heterogéneo (hnRNP's), que es un ARN de alto peso molecular el cual no ha sido procesado, es decir el ARN recién transcrito. La proteína esta compuesta de 825 aminoácidos, con un peso aproximado de 120 kDa, se localiza en el núcleo, citoplasma y en superficie membranal, las modificaciones postraduccionales a las que es susceptible son: la fosforilación, la cual se puede dar en 37 residuos diferentes (Figura 26) y la metilación en la Arg-733 y Arg-739. Esta ribonucleoproteína es componente del complejo CRD, el cual promueve la estabilización del ARNm de c-MYC. Con la unión al pre-ARNm, la hnRNP U adquiere una elevada afinidad por SAR (scaffold-attached región) de ADN y es capaz de unirse a las cadenas dobles y sencillas de ADN y ARN dependientes de ATP. La hnRNP U participa en la apoptosis. Una vez iniciada la apoptosis la unión hnRNP- ADN se desestabiliza, pero la unión con el ARN no se ve alterada.

La isoforma U de la ribonucleoproteína es abundante en el núcleo, ésta se ha relacionado con la unión de la RNA Polimerasa II al ARNm coadyuvando a la transcripción³⁰. Esta isoforma se ha identificado como un sustrato para una cinasa

que interactúa con el ADN fosforilándose en dos sitios, contribuye a la reparación del ADN³¹ y participa en la organización de la heterocromatina³³. Si bien no existen reportes en los cuales se haya asociado a esta proteína con algún tipo de cáncer, todos los procesos en los cuales ha sido reportada tienen una íntima relación con el proceso de la transcripción, ya que interviene en el procesamiento del pre-ARNm al ARNm. Esto hace pensar que si existe una hiperfosforilación o bien una producción elevada de esta proteína podría existir un incremento en la actividad de la transcripción, lo cual se reflejaría en una mayor producción proteica y con esto la célula tendría mayor material para llevar a cabo la formación de nuevas células en un lapso muy corto, contribuyendo directamente a la tasa de proliferación en un tejido determinado. Al igual que la nucleolina, su acceso para ser analizada es complicado ya que se encuentra en el núcleo.

Proteína de choque térmico HSP 90- α isoforma 2 y la proteína de choque térmico HSP 90- β .

Ambas proteínas fueron identificadas mediante un análisis por LS/ESI-MS/MS de la banda número 2 de los geles de fosfoproteínas (Figura 19). Para la búsqueda de homología entre los fragmentos de las secuencias proteicas reportadas en la base de datos NCBI y los fragmentos proteicos identificados experimentalmente se utilizó el algoritmo de búsqueda MASCOT. El resultado de este análisis fue la identificación de 10 fragmentos proteicos que pertenecen a la secuencia proteica de HSP 90- α isoforma 2. Estos fragmentos representan un 16% de la secuencia proteica reportada y tienen un general score de 357, superando el límite mínimo establecido por el algoritmo de búsqueda MASCOT.

De los fragmentos proteicos identificados, 3 poseen un ion score mayor a 45, es decir, poseen una probabilidad de coincidencia mayor a 0.05, lo que nos indica una homología elevada entre los fragmentos proteicos identificados experimentalmente y los que componen la secuencia proteica reportada de la HSP 90- α isoforma 2.

En la misma banda se identificaron 17 fragmentos proteicos que pertenecen a la HSP 90- β . Éstos representan el 21% de la secuencia proteica reportada como la HSP 90- β .

Los 17 fragmentos proteicos identificados poseen un general score de 531, por lo que, su homología respecto a los que componen la secuencia de la HSP 90- β es elevada.

De los fragmentos proteicos identificados, 4 poseen un ion score mayor a 0.05, es decir, poseen una probabilidad de coincidencia mayor a 0.05. Cumpliendo así con los dos parámetros de confiabilidad establecido por el algoritmo de búsqueda MASCOT.

La HSP 90- α isoforma 2 esta compuesta por 732 aminoácidos con un peso aproximado de 86 kDa. Interviene en el proceso contra el estrés celular, está compuesta por 10 α -hélices, 8 hojas β -plegada y un giro (Figura 34), su ubicación principal es el citoplasma, es propensa a dos cambios postraduccionales: la acetilación y la fosforilación la cual puede llevarse a cabo en 40 diferentes residuos (Figura 26.). Esta proteína puede interaccionar con AHSA1, FNIP1, HSF1, SMYD3 y TOMM34, TERT y PTGES3/P23, unión que es estabilizada por TERT.

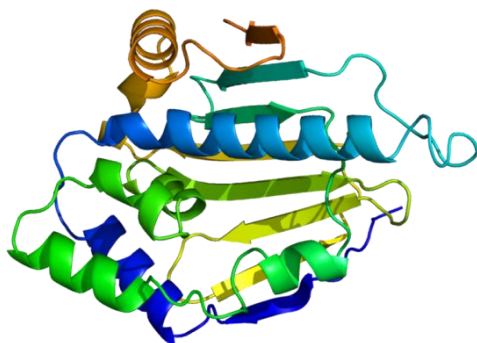


Figura 34. Estructura 3D de la Proteína de choque térmico α (Protein Data Bank, <http://www.pdb.org/pdb/home/home.do>).

La HSP 90- β es una proteína de 724 aminoácidos con un peso aproximado de 84 kDa. Está compuesta por 9 α -hélices, 8 β -plegadas y 3 giros (Figura 35). Al igual que la 90- α , se encuentra en el citoplasma y es susceptible a la acetilación y fosforilación en 41 residuos de aminoácidos (Figura 28). Esta proteína tiene actividad de cinasa y actúa como chaperona molecular. Al igual que la 90- α , interacciona con TP53/p53 y UNC45A.

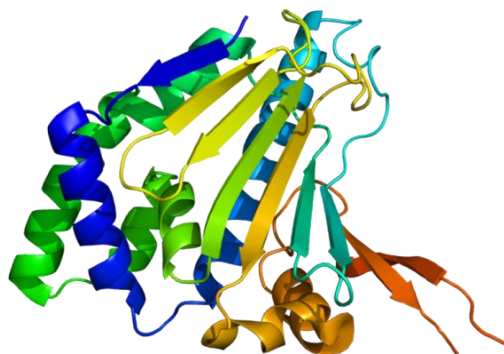


Figura 35. Proteína de choque térmico β (Protein Data Bank, <http://www.pdb.org/pdb/home/home.do>).

La proteína de choque térmico HSP 90 y sus isoformas son proteínas chaperonas las cuales se incrementan en carcinomas ductales de mama³⁶. Esta sobre-expresión lleva a la transcripción de un mayor número de proteínas, estas proteínas tienen diversas funciones y además incrementan la velocidad de crecimiento, ya que forman distintos complejos con oncogenes. Ambas proteínas de choque térmico están relacionadas con el plegamiento no covalente y con el desensamblaje y ensamblaje de polipéptidos o moléculas de ARN, incluyendo los procesos de oligomerización y también en el plegamiento y reensamblaje de proteínas y ARN desnaturalizados por estrés, por tanto son fundamentales en la respuesta al estrés. Se han relacionado con el cáncer y se han visto incrementos en los niveles de la concentración celular de HSP27 y HSP70 en el Cáncer de mama. Estos cambios dan como resultado el bloqueo de mecanismos como la apoptosis y las vías de senescencia replicativa, las cuales desencadenan la muerte de la célula. En estos procesos, la HSP90 juega un papel facilitando la transformación de estas células por la estabilización de las oncoproteínas sobre-expresadas y mutadas en tumores de Cáncer de mama, razón por la cual es un blanco perfecto para el tratamiento del carcinoma de mama y otros tipos de cáncer³⁴. También se ha demostrado relación entre la proteína p53 y la Hsp90 α mediante la formación de un complejo en células de Cáncer de mama. Este complejo, una vez formado, coadyuva a la disminución de la actividad transcripcional de p53. Este complejo provoca una deficiencia de p53 y como consecuencia se origina un crecimiento descontrolado de las células. Este comportamiento se observó en un estudio realizado por Abdur Rehman y

colaboradores³⁵ en el cual se tratan a líneas celulares con PRIMA-1, que es un compuesto que actúa dependiendo de p53 y tiene la capacidad de inducir a la apoptosis gracias a la restauración de la transcripción de p53. En el experimento se observó un incremento en la cantidad de HSP90α libre en las líneas no tratadas con PRIMA-1, las cuales corresponden a las células con un crecimiento descontrolado, mientras que en los controles, que consistieron en líneas celulares con una mutación en p53, tratadas con PRIMA-1, se detectaron niveles bajos de la HSP90α libre, confirmando que la formación del complejo HSP90α-p53 interviene en el desarrollo celular descontrolado de células cancerosas aún cuando no se conocen las vías en las que interviene este complejo³⁵. Ambas isoformas se localizan en citoplasma y no son secretadas por la célula.

Disulfuro isomerasa.

En el análisis de la banda número 4 de los geles de fosfoproteínas (Figura 19) por LS/ESI-MS/MS y utilizando un algoritmo de búsqueda MASCOT; se identificaron 7 fragmentos proteicos pertenecientes a la disulfuro isomerasa. De los cuales, 3 poseen un ion score mayor a 45, es decir, tienen una probabilidad de coincidencia mayor a 0.05.

Los 7 fragmentos proteicos identificados poseen un general score de 315 y representan el 50% de la secuencia reportada en la base de datos de NCBI como la disulfuro isomerasa. El general score nos indica que los fragmentos proteicos identificados experimentalmente poseen una homología elevada a los fragmentos que componen la secuencia proteica reportada como la disulfuro isomerasa.

En la Tabla 16 podemos observar la selección de la proteína disulfuro isomerasa.

La proteína disulfuro isomerasa es una cinasa que se ha asociado con una función anticancerígena por una ruta no conocida. La ausencia de ésta afecta en la terapia contra el cáncer⁴³. Se ha relacionado su subexpresión con tipos de cáncer agresivos en estudios realizados mediante inmunohistoquímica (INHQ). En la Figura 29 se pueden observar los posibles sitios de fosforilación que puede presentar esta proteína, sin embargo no se ha observado que esta proteína pueda sufrir de fosforilaciones *in vivo* razón por la cual es excluida en este estudio de ser

un posible biomarcador fosforilado.

CD44E.

En el análisis de la banda número 6 de los geles de glicoproteínas (Figura 20) por LS/ESI-MS/MS, empleando un algoritmo de búsqueda MASCOT; se identificaron 5 fragmentos proteicos que pertenecen a la secuencia de la proteína reportada en la base de datos NCBI como la CD44E. Éstos representan un 35% de la secuencia proteica total reportada y tienen un general score de 280, lo que significa, que tienen una probabilidad elevada de pertenecer a la secuencia proteica reportada como CD44E.

De los 5 fragmentos identificados, 3 poseen un ion score mayor a 45, es decir, la probabilidad calculada de coincidencia es mayor a 0.05. Por lo tanto, la homología de los fragmentos proteicos identificados experimentalmente es elevada respecto a los fragmentos que componen la secuencia proteica de la CD44E.

La CD44E está codificada por una secuencia de 5.74 kpb, la cual origina una proteína de la que sólo se ha identificado una fracción de 742 aminoácidos. Debido a su procesamiento (splicing), se pueden encontrar diversas isoformas de las cuales su peso molecular oscila entre 85 a 230 kDa (Figura 36). Ya que sólo su primer y quinto exónes están conservados, se han encontrado en cáncer 17 de sus variables, para el caso de la CD44E o CD44, forma epitelial, los últimos 3 exones están conservados. Esta glicoproteína se localiza en la superficie celular, y tiene la capacidad de O-glicosilarse hasta en 8 sitios. Está involucrada en la interacción célula-célula, la adhesión celular y la migración, es afin al ácido hialurónico, además de tener la posibilidad de interactuar con otros ligandos como la osteopontina, la colágena y las MMPs (metaloproteinasas de matriz celular).

La CD44 participa en una gran variedad de funciones celulares como la activación, recirculación y reciclaje de linfocitos, hematopoyesis y en la metástasis de los tumores. Esta proteína funciona como un transmisor de bioseñales para contribuir con la metástasis de las células. Ésto se observó de manera fortuita en un experimento en el cual se cultivaron células MCF7, las cuales produjeron una menor cantidad de CD44 endógeno dando origen a la línea MCF7-B5, misma que

tiene un mayor potencial para proliferar, adherirse, invadir y migrar *in vitro*, que son propiedades importantes para el desarrollo de cáncer³⁷. Esta proteína posee splicing alternativos, razón por la cual se han observado distintas isoformas.



Figura 36. Estructura 3D de la proteína CD44E (Protein Data Bank, <http://www.pdb.org/pdb/home/home.do>).

Una de las características más sobresalientes de la isoforma E, es que su tallo exhibe un receptor de cinasa o actividad de fosfatasa, el cual se localiza fuera de la matriz extracelular. La CD44 se puede comunicar a través de varios mecanismos:

- 1) Señales directas de CD44 gracias a la unión con el ácido hialurónico, el cual activa a ciertas cinasas que se encargan de potencializar la asociación de CD44 con complejos de actina del citoesqueleto iniciando así señales que terminan en la potencialización de la migración celular.
- 2) CD44 sirve como un correceptor el cual se une a otros receptores de señalización, como los miembros de la familia de ErbB (Factores de crecimiento), los cuales están relacionados con el crecimiento celular en el cáncer, facilitando con esto la asociación de mediadores intracelulares de transducción de señales.
- 3) Funcionando como una proteína de acoplamiento de proteínas pericelulares, como las MMPs o proteínas citoplásmicas como las cinasas.
- 4) Intervención de forma directa por parte de CD44, la cual ocurre en dos pasos: el primero secretando la parte del dominio transmembranal con el consecuente desprendimiento del dominio intracelular y su traslocación nuclear funcionando como un factor de transcripción de genes.

Muchos estudios relacionan a CD44 con la metástasis en Cáncer de mama pero se ha observado una variante de CD44 que disminuye, mientras la otra aumenta en carcinomas mamarios. De acuerdo a lo publicado por Allal Ouhtit y colaboradores³⁷, la expresión de CD44s (Bajo peso molecular) se incrementa causando una disminución en los niveles de CD44E, lo cual hace a esta glicoproteína membranal una candidata ideal para ser un marcador de diagnóstico junto con su isoforma S. Existen estudios en los cuales se ha demostrado que la eficacia de Herceptin® se potencializa contra el Cáncer de mama si las células expresan deficientemente a CD44³⁸, confirmando así la implicación en el crecimiento celular y cómo entrelaza esta proteína diferentes señales biológicas. Se ha relacionado a CD44 con la generación de vasos sanguíneos para el crecimiento celular (angiogénesis) ya que pruebas en ratones demuestran que la expresión deficiente de esta proteína afecta directamente la formación del epitelio y por lo tanto la generación de nuevos vasos sanguíneos, confirmando así la participación de CD44 en la metástasis y angiogénesis. Si bien existe una disminución de CD44E endógeno y un incremento de CD44s, el primero podría emplearse para la generación de vasos, mientras que CD44s se expresa en la célula para compensar la falta de CD44E.

Proteína de 80 kDa (subunidad β de la Glucosidasa II).

Del análisis de la banda número 7 de los geles de glicoproteínas (Figura 20) por LS/ESI-MS/MS, empleando el algoritmo de búsqueda MASCOT; se identificaron 6 fragmentos proteicos que pertenecen a la secuencia proteica reportada en la base de datos NCBI como la subunidad β de la Glucosidasa II.

Estos fragmentos proteicos poseen un general score de 286 y representan el 29% de la secuencia proteica reportada. De los fragmentos proteicos identificados, 4 presentaron un ion score mayor a 45, por lo que su homología es elevada respecto a los fragmentos que componen la secuencia proteica reportada como la subunidad β de la Glucosidasa II.

La subunidad β de la Glucosidasa II (PRKCSH) se compone de 528 aminoácidos. Esta proteína se puede someter a varias modificaciones postraduccionales como la glicosilación, la cual se puede llevar a cabo en 10 residuos a lo largo de la cadena (Figura 32). Junto con el heterodímero de la subunidad α catalítica (GANAB),

conforma la Glucosidasa II, la cual es una proteína que interviene en el metabolismo de los glicanos y los *N*-glicanos. Su localización celular es en el retículo endoplásmico. Cada una de las subunidades juega un papel importante; la subunidad α es la subunidad catalítica de esta proteína y la subunidad β se encarga de proporcionar estabilidad al dominio catalítico. Se ha observado que defectos en esta proteína son causa de enfermedad poliquística en el hígado. Esta enfermedad es un desorden autosómico dominante y se caracteriza por la presencia de quistes de origen epitelial, al igual que quistes hepáticos, razón por la que tiene una actividad importante durante el desarrollo de carcinomas hepáticos.

La Glucosidasa II se ha asociado al desarrollo de carcinomas mamarios gracias a la asociación que existe con la estimulación de factores de crecimiento, principalmente con el factor de crecimiento fibroblástico 1 (FGF-1), el cual se incrementa en el cáncer. La 80K-H o subunidad β de la Glucosidasa II es la encargada de estabilizar el incremento de este factor por un mecanismo que aún se desconoce, sólo se sabe que esta proteína es un sustrato de la proteína cinasa C (PKC), y que en cáncer interviene incrementándose en los procesos en los cuales está involucrada la PKC muy cerca del núcleo. Esta relación se observó gracias a estudios⁴¹ en los cuales se encontró una relación proporcional entre el FGF-1, la proteína 80K-H y el grado de tumor. Esta relación se observó al inducir la traslocación de la proteína 80K-H al núcleo, lo cual dio como resultado un mayor crecimiento celular y una mayor capacidad de invasión, todo gracias a la formación del complejo FGF-1/80K-H. Si bien no se sabe con certeza la capacidad de esta proteína para ir al núcleo y formar el complejo mencionado, se sabe que la cantidad de glicosilación se incrementa en las proteínas membranales y en enzimas como la calnexina y la calreticulina, esta última es importante en la regulación de los niveles de Ca^{2+} en la célula que interfiere directamente con la proliferación celular. Esta proteína ha sido identificada en trabajos anteriores mediante técnicas proteómicas (geles 2D y espectrometría), si bien hemos notado que existe un incremento en los niveles de calreticulina durante el Cáncer de mama en un modelo celular, esta calreticulina para ser activa necesita glicosilarse, razón por la cual se necesitaría un elevado nivel de Glucosidasa II o bien algún cambio que potencializara su actividad. Por ello es una

candidata idónea para un posible biomarcador de diagnóstico, el cual se basaría en una diferencia de expresión.

Las dos glicoproteínas CD44 y 80K-H son proteínas que se encuentran relacionadas con el crecimiento celular, ya sea interactuando con diversos factores o bien como sustratos estabilizadores de otras proteínas. Esto les da una gran ventaja sobre las fosfoproteínas reportadas, debido a que éstas se encuentran asociadas con proteínas membranales o bien alguna de sus subunidades se localizan ahí, por lo cual son candidatas ideales para posibles biomarcadores de diagnóstico o bien como blancos terapéuticos. En el caso de la 80K-H, a pesar de no conocerse la o las rutas metabólicas en las que participa, se sabe que se encuentra íntimamente relacionada con la capacidad invasiva y de crecimiento de los tumores. Es interesante ver como se ha reportado una disminución de la isoforma E y un incremento en las isoformas “s” o de bajo peso molecular, lo cual pone de manifiesto la actividad de cada una de ellas. Además, las formas fosforiladas podrían correlacionar también inversamente en su expresión implicándolas en la enfermedad.

Las herramientas proteómicas que se han ido incorporando a lo largo de los años han contribuido a la identificación de nuevos biomarcadores y blancos terapéuticos como los que se proponen en este trabajo. Estos se encuentran en la primera etapa para ser propuestos como biomarcadores de cáncer, sin embargo es alentador pensar que en un futuro podrán colaborar a reducir el elevado número de muertes por el Cáncer de mama.

Capítulo 8. Conclusiones.

De un estudio de proteómica comparativa, en el cual se utilizó cromatografía de afinidad y electroforesis SDS-PAGE para la separación de fosfoproteínas y glicoproteínas. Y para la identificación de éstas se empleo un análisis por LS/ESI-MS/MS (Electrospray) con el algoritmo de búsqueda MASCOT. De este estudio se obtuvieron 3 proteínas que son propuestas como posibles biomarcadores fosforilados.

- ✓ Nucleolina. Involucrada con la transcripción en la célula y la expresión de BRCA1, colaborando con el desarrollo celular.
- ✓ Ribonucleoproteína U (hnRNAU). La cual estabiliza el ARN incrementando la cantidad de proteína transcrita.
- ✓ HSP-90. Encargada de mantener estables a las proteínas en condiciones de estrés y relacionada con la estabilización de oncogenes.

En la fracción de las proteínas glicosiladas se obtuvieron 2 proteínas, las cuales son propuestas como posibles biomarcadores.

- ✓ CD44E. Un receptor de membrana relacionado con la angiogénesis y el crecimiento celular.
- ✓ 80K-H. La subunidad β de la glucosidasa II que se ha asociado con la estabilización del FGF (Factor de Crecimiento Fibroblástico) y con la estabilización de la Glucosidasa II, la cual participa en la glicosilación de las proteínas.

Perspectiva.

Estas 5 proteínas tendrán que ser validadas y sometidas a los criterios que establece el Instituto Nacional de Cáncer (NCI), para poder ser propuestas como biomarcadores que coadyuven en un diagnóstico temprano de Cáncer de mama y con esto poder contribuir a la disminución en la tasa de defunciones y pacientes afectadas con este tipo de neoplasias, mejorando su pronóstico y calidad de vida.

9. Referencias Bibliográficas.

1. Bidart J, Thuillier F, Augereau C, Chalas J, Daver A, Jacob N, Labrousse Y, Voitor H. Kinetics of serum tumor marker concentrations and usefulness in clinical monitoring. Clin. Chem. 1999; 45(10):1695-1707.
2. Serova O, Mazoyer S, Puget N, Dubois V, Tonin P, Shugart Y, Goldgar D, Narod SA, Lynch HT, Lenoir GM. Mutations in BRCA1 and BRCA2 in breast cancer families: are there more breast cancer-susceptibility genes?. Am. J. Hum. Gene. 1997; 60(5):486-495.
3. Huquet E, McMahon J. Differential expression of human Wnt genes 2, 3, 4, and 7B in human breast cell lines and normal and disease states of human breast tissue. Canc. Res. 1994; 54(10):2615-2621.
4. Copeland III M. The Breast; Comprehensive Management of Benign and Malignant Disorders. 3° Edition. Souders Elsevier. New York (USA) 2004.1628 p.
5. Ruck K. Encyclopedic reference of genomics and proteomics in molecular medicine. 2° Edition. Springer Dgante. New Jersey (USA). 2006. 2112 p.
6. Yoon Pin Lim. Mining the tumor Phosphoproteome for Cancer Marker. Clin. Cancer Res. 2005; 11(3):3163-3170.
7. Gallegos I, Coutiño R, Martínez G, Hernández P. Marcadores glicosilados en Cáncer de mama. REB. 2008; 27 (2):52-59.
8. Winer E, Gralow J, Diller L, Gralow J, Ozols R, Bajorin D, Cheson B, Sandler H, Winer E, Bonner J, Demetri G, Curran W, Ganz P, Kramer B, et. al. Clinical cancer advances 2008: Major research advances in cancer treatment, prevention, and screening—A report from the American Society of Clinical

Oncology. J. Clin Oncol. 2009; 27(5): 812-826.

9. Ruin P. Oncología Clínica: Un enfoque multidisciplinario para Médicos y estudiantes. Diork Servicios Integrales de Edición (traductor). 8ª Edición. Elvieser Science. New York (USA). 2002.937 p.
- 10.Dalkic E. Cancer-Drug Associations: A Complex System. 3° Edition. Plos One. London (UK) 2010. 554 p.
- 11.Cross D. Gene Therapy for Cancer Treatment: Past, Present and Future, Clin. Med. Res. 2006; 4(3): 218–227.
- 12.Aragues R, Sander C, Baldo Oliva. Predicting cancer involvement of genes from heterogeneous data. BMC Biol. 2008; 9(172):1-18.
- 13.Kim H. Expression of MT-1 MMP, MMP2, MMP9 and TIMP2 mRNAs in ductal carcinoma *in situ* and invasive ductal carcinoma of the breast. J. Med. 2006; 47(3):333–342.
- 14.Lanfranchi A. Normal breast physiology: The reasons hormonal contraceptives and induced abortion increase breast-cancer risk. Clin. Med. 2009; 76(3):236–249.
- 15.Meibodi NT, Ghoyunlu VM, Javidi Z, Nahidi Y. Clinicopathologic evaluation of mammary Paget's disease. Indian J Dermatol. 2008; 53(1):21–23.
- 16.Ramachandra S, Machin L, Ashley S, Monaghan P, Gusterson BA. Immunohistochemical distribution of c-erB-2 in *in-situ* breast carcinoma: a detailed morphological analysis. J. Pathol. 1990; 16(17):7-14.

17. Hamdan M, Dominic M, Nibbering D, et al. Cancer biomarkers: Analytics techniques for discovery. *J. Biol. Med.* 2007; 80(3):137–138.
18. Theodore L, De Vita V. *Cancer Principles and Practice of Oncology*. 8^o Edition. Souders Elsevier. New York (USA). 2004. 2035p.
19. Qiagen. Qproteome™ Glycoprotein Fractionation Handbook, 2^o Edition. Qiagen Sample and Assay. Darmstadt (Germany). 2007. 32p.
20. Donegan W, Spratt J. *Cancer of Breast*. 5^o Edition. Souders Elsevier. New Jersey (USA). 2002. 1050p.
21. Redel G. *Encyclopedic Dictionary of Genetics, Genomics and Proteomics*. 2^o Edition. Springer. London (UK). 2006. 129p.
22. Hamdan M, Righetti P. *Proteomics today: protein Assessment and Biomarkers Using Mass Spectrometry, 2D Electrophoresis, and Microarray Technology*. 1^o Edition. John Wiley and Sons. New Jersey (USA). 2005. 426p.
23. Wilson S, Suk W. *Biomarkers of Environmentally Associated Disease: Technologies, Concepts and Perspectives*. 1^o Edition. John Wiley and Sons. New Jersey (USA). 2002. 582p.
24. Nelson D, Cox M. *Lehninger Biochemistry Principles*. 3^o Edición. Worth Publishers. New York (USA). 2005. 1119p.
25. Gerald Karp. *Biología Celular y Molecular: Conceptos y Experimentos*. 3^o Edición. Mc-Graw Hill Interamericana. 2000. 899p.

26. Przybyło M, Pocheć E, Drabik A, Silberring J, Kremser M, Schadendorf D, Laidler P, Lityńska A. The new face of nucleolin in human melanoma. *Cancer Immunol. Immun.* 2009; 58(9):1471-80.
27. Tulchin N, Chambon M, Juan G, Dikman S, Strauchen J, Ornstein L, Billack B, Woods NT, Monteiro AN. BRCA1 protein and nucleolin colocalize in breast carcinoma tissue and cancer cell lines. *Am J. Pathol.* 2010; 176(3):1203-1214.
28. Masiuk M, Urasinska E, Domagala W. Simultaneous measurement of nucleolin and estrogen receptor in breast cancer cells by laser scanning cytometry. *Anticancer. Res.* 2004; 24(2B):963-966.
29. Masiuk M, Urasinska E, Domagala W. Intranuclear nucleolin distribution during cell cycle progression in human invasive ductal breast carcinomas in relation to estrogen receptor status. *Anticancer Res.* 2007; 27(6B): 3957-3962.
30. Fogal V, Sugahara K, Ruoslahti E, Christian S. Cell surface nucleolin antagonist causes endothelial cell apoptosis and normalization of tumor vasculature. *Angiogenesis.* 2009; 12(1):91-100.
31. Kiledjian M, Dreyfuss G. Primary structure and binding activity of the hnRNPU protein: binding RNA through RGG box. *The EMBO J.* 1992; 11(7):2655 – 2664.
32. Berglund F, Clarke P. hnRNP-U is a specific ADN-dependent protein kinase substrate phosphorylated in response to ADN double-strand breaks. *Biochem. Biophys. Res. Commun.* 2009; 381(1):159-164.

33. Ameyar-Zazoua M, Souidi M, Fritsch L, Robin P, Thomas A, Hamiche A, Percipalle P, Ait-Si-Ali S, Harel-Bellan A. Physical and functional interaction between heterochromatin protein 1alpha and the RNA-binding protein heterogeneous nuclear ribonucleoprotein U. *J. Biol Chem.* 2009; 284(41):27974-9.
34. Calderwood S. Heat shock proteins in breast cancer progression - A suitable case for treatment?. *Int J Hyperthermia.* 2010; 26 (7):681-685.
35. Rehman A, Chahal MS, Tang X, Bruce JE, Pommier Y, Daoud SS. Proteomic identification of heat shock protein 90 as a candidate target for p53 mutation reactivation by PRIMA-1 in breast cancer cells. *Breast Cancer Res.* 2005; 7(5):765–774.
36. Zagouri F, Sergentanis T, Nonni A, Papadimitriou C, Michalopoulos N, Domeyer P, Theodoropoulos G, Lazaris A, Patsouris E, Zografos E, Pazaiti A, Zografos G. Hsp90 in the continuum of breast ductal carcinogenesis: Evaluation in precursors, preinvasive and ductal carcinoma lesions. *BMC Cancer.* 2008; 10(353):1-9.
37. Ouhtit A, Zakaria Y, Elmageed A, Abdraboh M, Lioe T, Raj M. *In vivo* Evidence for the Role of CD44s in Promoting Breast Cancer Metástasis to the Liver. *Am. J. Pathol.* 2007; 171(6):2033-2039.
38. Cavaliere C, Corvigno S, Galgani M, Limite G, Nardone A, Veneziani BM. Combined inhibitory effect of formestane and herceptin on a subpopulation of CD44+/CD24low breast cancer cells. *Cancer. Sci.* 2010; 10(7):1661–1669.
39. Cao G, Savani RC, Fehrenbach M, Lyons C, Zhang L, Coukos G, Delisser HM.. Involvement of Endothelial CD44 during *in vivo* Angiogenesis. *Am. J. Pathol.* 2006; 169(1):325-336.

40. MeTappel A. Lysosomal enzymes and initiation of breast cancer. *Med. Hypo.* 2005; 64(2):288-289.
41. Forough R, Lindner L, Partridge C, Jones B, Guy G, Clark G. Elevated 80K-H protein in breast cancer: a role for FGF-1 stimulation of 80K-H. *Int. J. Biol. Markers.* 2003; 18(2):89-98.
42. Sánchez E, Rísquez R, Chasseraud M, Ahidouch A, Ortiz C, Ouadid-Ahidouch H, García JM. Synthesis of N-, S-, and C-glycoside castanospermine analogues with selective neutral alpha-glucosidase inhibitory activity as antitumour agents. *Chem. Commun.* 2010; 46(29):5328-5330.
43. Huang S, Tomich J, Wu H, Jongs A, Holcenbergo J. Human deoxycytidine kinase. Sequence of cADN clones and analysis of expression in cell lines with and without enzyme activity. *J. Biol. Chem.* 1989; 264(6):14762-14768.
44. Qiagen. *Phosphoprotein Handbook. Qiagen Sample and Assay Second Edition.* Darmstadt (Germany). 2007. 27p.
45. Farmacología Básica: <http://farmacologiabasica.blogspot.com> (Octubre, 2010)
46. National Institute of Cancer: <http://www.cancer.gov/> (Mayo, 2010).
47. World Heart Organization: <http://www.who.int/en/> (Abril, 2010).
48. Protein Data Base: <http://www.pdb.org/pdb/home/home.do> (Julio, 2010).

ANEXO I.

Técnicas para la identificación de proteínas.

- **Cromatografía.** Si bien se ha tenido un gran desarrollo en diversas técnicas espectrométricas que permiten identificar cambios postraduccionales de miles de proteínas o identificarlas a partir de uno de sus fragmentos, la separación de éstas en grupos que posean características semejantes como el peso molecular o un cambio postraducciona, como la fosforilación y glicosilación por mencionar algunos, reducen el tiempo de análisis considerablemente, debido a la disminución en el grupo de análisis.

Existen diferentes tipos de cromatografías, las cuales se dividen en dos grandes grupos: el primero es el de adsorción y ejemplos de este grupo son la cromatografía de intercambio iónico, la de fase inversa, la de fase normal entre otras. El segundo grupo es el de no adsorción algunos ejemplos son la filtración en gel y la exclusión molecular.

Una de las cromatografías más empleadas por el elevado porcentaje de selectividad hacia los sustratos es la cromatografía de afinidad. En esta cromatografía la proteína a purificar posee afinidad específica por un ligando como su cofactor, un anticuerpo o un ión metálico. Lo que ocasiona una interacción específica entre moléculas. Las columnas cromatográficas se componen por una matriz o fase sólida en la cual se encuentran anticuerpos o moléculas afines hacia ciertos grupos químicos (fosfatos, carbohidratos), por la cual pasa la fase móvil, que contiene una mezcla de diversas proteínas originando una interacción péptido-matriz mediante enlaces débiles de los tipos de Van der Waals, enlaces de hidrógeno, hidrofobicidad o interacciones iónicas. Estos enlaces se pueden debilitar mediante un amortiguador que posea características distintas a las de la fase móvil inicial, generalmente un pH distinto. En el caso de la cromatografía de afinidad se utiliza una solución que contiene ligandos libres. Las proteínas eluidas se pueden analizar por geles 1D y 2D o bien mediante diferentes variantes de espectrometría de masas.

En la separación de fosfoproteínas, las columnas de afinidad más utilizadas son aquellas que contienen en la matriz anticuerpos contra treonina, tirosina y serina fosforilada. Para la separación de glicoproteínas las columnas más empleadas son las que contienen lectinas de origen vegetal por ejemplo la Concavalina A (Con A).

➤ Cromatografía líquida de alta resolución (HPLC). Es una herramienta utilizada para la identificación de proteínas que tiene como característica principal la utilización de bombas que deben: ser resistente a la corrosión, alcanzar presiones estables de hasta 6000 psi, estar libre de pulsaciones, ser capaces de generar intervalos de caudales de flujo que vallan de 0.1 mL/min hasta 10 mL/min, permitir un control y reproducibilidad de elución del solvente. Estas características permiten que la velocidad a la cual corre la fase móvil sobre la estacionaria disminuya y así el tiempo de contacto entre fases sea mayor, obteniendo una mayor resolución. Existen variantes de esta técnica debido a la utilización de diferentes matrices:

- Cromatografía de adsorción. En donde la fase sólida es un sólido generalmente sílice (silica) o en ocasiones alúmina.
- Cromatografía de reparto. En donde se unen covalentemente compuestos químicos a un soporte de sílica. Ésta se subdivide en fase normal y fase reversa, en la primera se utilizan como fase sólida compuestos polares (agua, trietilenglicol) y en la fase móvil se utilizan compuestos no polares, por lo tanto las moléculas de este tipo eluyen primero. Para la fase reversa se utilizan matrices no-polares (hidrocarburos alifáticos) y como fase móvil sustancias polares por lo cual moléculas con estas características eluyen primero.
- Cromatografía Iónica. En ésta se utilizan columnas rellenas con resinas de intercambio iónico y se emplean para la separación e identificación de moléculas que se comportan como iones.
- Cromatografía de exclusión molecular. La fase estacionaria de este tipo de cromatografía se compone por partículas poliméricas o de

sílice, las cuales contienen una red uniforme de poros que llevan a cabo un fraccionamiento de acuerdo al tamaño molecular, eluyendo primero las moléculas de mayor tamaño y por último las de menor tamaño.

- **Electroforesis en SDS-PAGE.** Esta técnica de separación se basa en la migración de proteínas cargadas a través de un campo eléctrico que va de negativo a positivo. Este método de purificación no es comúnmente utilizado para muestras grandes, ya que existen técnicas más simples que no afectan tanto la estructura de las proteínas a analizar ni su función. Por eso, este tipo de electroforesis se utiliza como método analítico, es decir, en el caso de estructuras proteómicas, para proteómica cualitativa y de identificación. Es recomendada para muestras pequeñas (no mayores a 30 µg), en las cuales se tenga la finalidad de identificar la presencia o ausencia de alguna proteína o bien para una proteómica comparativa, en la cual se haya utilizado una purificación en condiciones fuertes.

Esta electroforesis se basa en la masa molecular de los péptidos, generalmente se busca cargar a las proteínas negativamente mediante el uso de un amortiguador con un pH más elevado que el pI de cualquiera de las proteína a separar, para que éstos sean capaces de migrar hacia el polo positivo. La separación se realiza en geles hechos de un polímero reticulado llamado poliacrilamida, la cual actúa como un tamiz molecular, retardando la migración de las proteínas aproximadamente en proporción a su relación carga-masa. Esta migración puede verse afectada por la estructura cuaternaria de la proteína, por lo cual se emplean detergentes encargados de romper cualquier tipo de enlace (inter e intracatenario) y garantizando que la estructura de la proteína no sea un factor que altere la migración, el detergente más utilizado es el Dodecil Sulfato de Sodio (SDS), el cual además de hacer que las proteínas pierdan su estructura nativa por la unión a las zonas apolares del péptido, otorga un carga negativa adicional proporcional al tamaño de los péptidos ya que por que cada dos

aminoácidos existe una molécula de SDS. El resultado de este método es la obtención de bandas a lo largo del polímero (Figura 37) pueden visualizarse gracias a una tinción con Azul de Coomassie o con nitrato de plata. Las bandas teñidas pueden ser comparadas entre sí o con algún grupo control, el empleo de marcadores de pesos moleculares facilita esta operación.

Una de las grandes desventajas que presenta esta técnica es que en una banda se pueden encontrar varias proteínas que poseen el mismo peso molecular pero diferente pI , para solucionar este problema surge la electroforesis en 2 dimensiones (2D).

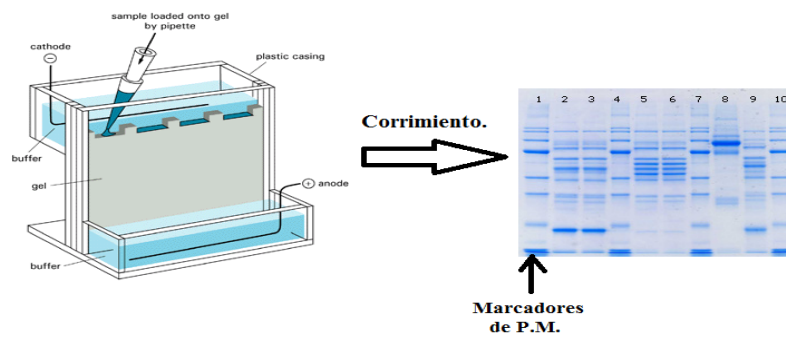


Figura 37. Electroforesis SDS-PAGE. La figura nos muestra la dirección en la cual se encuentra el campo eléctrico y un gel típico resultante de esta técnica teñido con Azul de Coomassie (Nelson y Cox, 2000).

- **Electroforesis 2D.** Su importancia radica en la capacidad que posee para separar miles de proteínas en un solo análisis, lo que hace que tenga un papel importante en la identificación de proteínas expresadas en diferentes enfermedades.

La electroforesis 2D (dos dimensiones) nace de la combinación de dos técnicas; la primera denominada “separación horizontal”, que se basa en el enfoque isoeléctrico, el cual separa proteínas de acuerdo a su pI . En este paso se utilizan bandas de acrilamidas de cierta longitud (5, 10 y 18 cm) y rangos de pH que pueden cubrir sólo un intervalo o bien toda la escala (1-14). El fundamento de esta técnica, es el establecimiento de un gradiente

de pH en la tira de acrilamida permitiendo que una mezcla de ácidos y bases orgánicas de bajo peso molecular se distribuyan en un campo eléctrico generado a través del gel. Cuando una muestra de proteínas es aplicada, cada proteína migrará hasta llegar al pH que coincida con su pI, resolviendo así la desventaja de los geles 1D que es, tener proteínas con un mismo peso molecular pero diferente pI. La segunda etapa consiste en la separación por el uso de la electroforesis SDS-PAGE o “separación vertical”, la cual hace una separación por peso molecular. El resultado de estas etapas es la obtención de un gel en el cual se identifican manchas.

Una vez que se han separado las proteínas por estos métodos, se puede obtener información de cada una de las manchas en bases de datos (Expasy.org). También se puede seleccionar las manchas de importancia y analizarse por espectrometría de masas. Existen variantes de esta técnica entre la que destaca el DIGE (Difference in gel electrophoresis). En esta variante se emplean fluoróforos (CyDye), que puede marcar distintas muestras, permitiendo así la generación de una mezcla de varias muestras y realizando una comparación entre éstas, gracias al tinte con los fluoróforos.

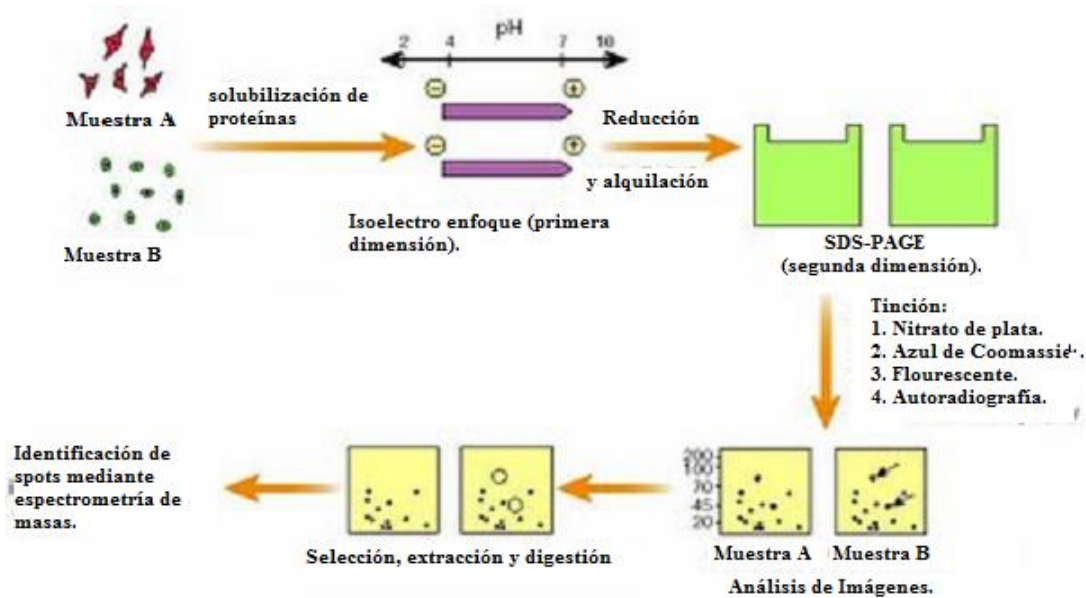


Figura 38. Análisis por Geles 2D (Farmacología básica, <http://farmacologiabasica.blogspot.com>).

- Chips de proteínas o microarreglos. De éstos se pueden obtener varia información de proteínas que pueden estar alteradas en ciertas enfermedades. Esta información incluye dos rubros, el primero, la determinación de los niveles de expresión y el segundo la determinación de interacciones selectivas con otras biomoléculas. Eexisten matrices de fase inversa que tienen la capacidad de inmovilizar un gran repertorio de proteínas. Estos microarreglos pueden obtenerse de forma comercial (previamente diseñados) o diseñarse de acuerdo a su uso. La aplicación clínica más utilizada para este tipo de microarreglos es la identificación de proteínas capaces de producir anticuerpos en enfermedades autoinmunes, por ejemplo el lupus eritematoso o la artritis reumatoide. Las matrices más empleadas para este tipo de arreglos son de alta densidad (> 100 elementos/matriz), en las cuales se colocan los anticuerpos de las proteínas que se pretenden identificar y la muestra debe ser etiquetada directamente con un fluoróforo o un hapteno o bien un elemento que permita la develación del resultado. Posteriormente las proteínas son cuantificadas mediante un detector. Existen además de la técnica descrita, diversas variantes como las siguientes:

- ✓ Microarreglos microspot ELISA y anticuerpos. Son utilizados para trazar el perfil cuantitativo de la expresión de proteínas en cultivos celulares o de muestras clínicas. Las matrices son de baja densidad (9 a 100 elementos/matriz). En este tipo de microarreglos los anticuerpos son conocidos y se pretenden capturar antígenos de las muestras analizadas. Para la detección del antígeno, éste debe ser marcado directamente con un fluoróforo, o un segundo ligante, generando así un ensayo similar a una ELISA.
- ✓ Microarreglos de captura de un solo anticuerpo. Se utilizan para la obtención de un perfil cualitativo. Este ensayo consiste en múltiples anticuerpos conocidos, que se encuentran fijos a una superficie sólida. Los microarreglos detectan la presencia de antígenos en una muestra. Una vez empleado el microarreglo se realiza una

comparación entre muestras de organismos enfermos y un control o muestra normal. En este tipo de microarreglos, los anticuerpos son unidos a una matriz a la cual se le agrega la muestra y se procede a la adición de un hapteno, capaz de indicar si la reacción antígeno-anticuerpo se ha llevado a cabo.

- ✓ Microarreglos de antígenos fase inversa. La aplicación de los arreglos de antígenos es la detección de auto-anticuerpos en enfermedades autoinmune. Se utilizan matrices de baja densidad y se utiliza suero o plasma como muestras. Este microarreglo empleado para el tratamiento de varias muestras con la finalidad de detectar un pequeño número de antígenos (1 a 3). La detección se realiza con anticuerpos relacionados con un fluorocromo o reactivo de detección de color.
- Espectrometría de masas. Esta técnica es la más utilizada para la identificación de secuencias de polipéptidos cortos; se ha convertido en una herramienta invaluable para la identificación rápida de proteínas desconocidas. En la espectrometría de masas primero el analito es ionizado con vacío, con la finalidad de cargarlo y que así pueda ser introducido a un campo eléctrico o magnético; los caminos que puede tomar a través del campo son una función de la relación masa-carga-radio característica de cada uno de ellos, lo que permite la deducción del tamaño con gran precisión. Durante mucho tiempo esta técnica no se utilizó en macromoléculas debido a las condiciones extremas a las que someten, fue hasta finales de la década de los 80's en donde surgieron variables menos destructivas con las que pudieron realizarse un análisis en macromoléculas, de estas técnicas destacan:
 - Tándem MS o MS/MS. Es una herramienta para la obtención de secuencias polipeptídicas. Su funcionamiento es el siguiente: la solución que contienen las proteínas desconocidas, es tratada con proteasas o reactivos químicos que hidrolizan y generan péptidos

pequeños. Esta mezcla es inyectada en un dispositivo integrado por dos espectrómetros de masas acoplados por un sistema de vacío (Figura 39); en el primer espectrómetro se ordenan los fragmentos ionizados y fragmentados de manera que sólo unos cuantos pasen hacia la cámara de vacío donde se fragmentan aún más y acceden al segundo espectrómetro. Esta operación se hace con la menor cantidad de agua posible. En el segundo espectrómetro la medida de todos los péptidos se encuentra en función de la relación a m/z ; los fragmentos generan uno o varios conjuntos de picos, los cuales representan todos los fragmentos obtenidos por el primer espectrómetro y la cámara de vacío. El tamaño de estas partículas depende de los puntos de rupturas en la proteína, o de la posición del carboxilo o amino terminal. Esta técnica es capaz de identificar con un alto porcentaje de certeza a cada uno de los aminoácidos en una cadena peptídica, excepto en los residuos de leucina e isoleucina.

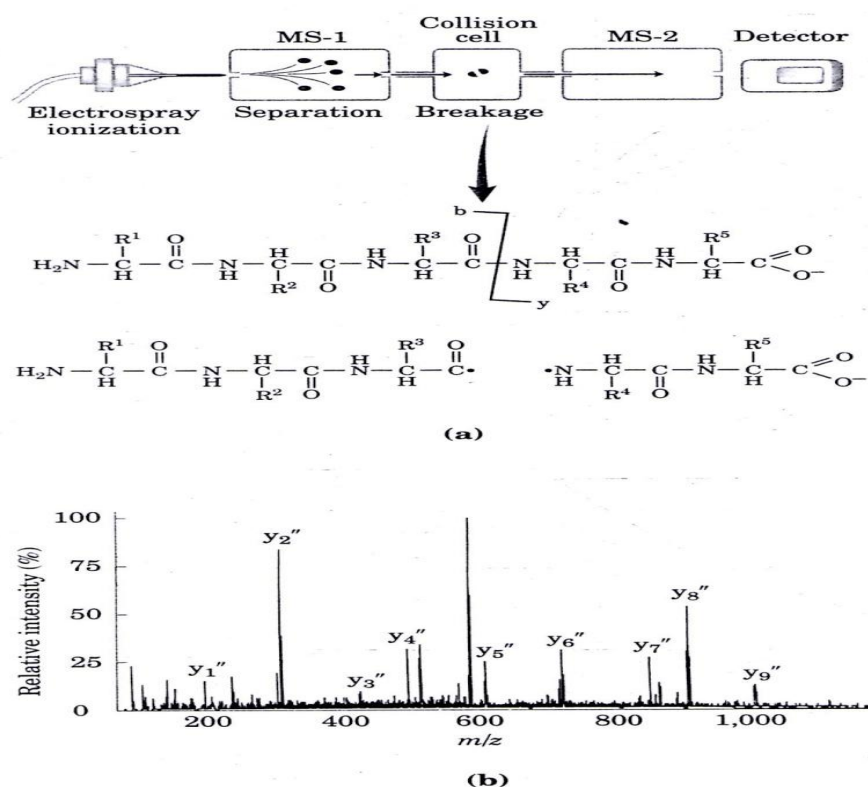


Figura 39. Técnica MS/MS. Esta técnica es una de las más utilizadas en la identificación de proteínas (Nelson y Cox, 2005).

- MALDI MS o TOF (matrix-assisted laser desorption/ionization mass spectrometry). Descrita por primera vez en 1988 por Tanaka, Karas y Hillen Kamp, hoy en día es una de las técnicas de ionización no volátiles más empleadas. En esta técnica la principal modificación es que las proteínas son colocadas en una matriz “light-absorbing” con una concentración cercana a 10 mg/L (Tabla 20); las proteínas deben haber sido previamente disueltas en un solvente adecuado con una concentración de picomoles en microLs. Tanto la matriz como la muestra son colocadas en una plataforma que es introducida al espectrómetro de masas para ser expuesta a pulsos de láser. La finalidad de estos pulsos es la de ionizar a la matriz y a la muestra. La naturaleza de los pulsos es hecha por analizadores TOF o bien analizadores de masas (MS). Una vez ionizada son desorbidas por la matriz e introducidas al sistema de vacío y llevadas a un detector (Figura 40), un punto fundamental en la realización de esta técnica es la elección de la matriz, ya que existen reportes de matrices que generan un mayor desempeño en ciertas muestra (Tabla 18).

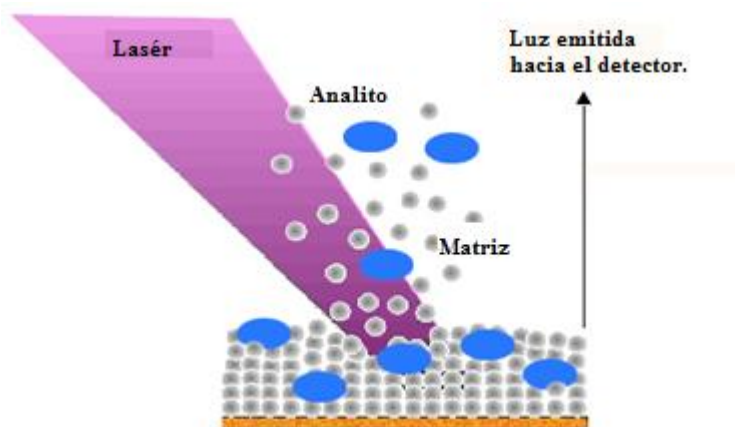


Figura 40. Técnica MALDI-TOF (Biopedia Org, <http://biopedia.org>).

Tabla 18. Matrices empleadas para la identificación de proteínas mediante MALDI MS. Las matrices que se enlistan en la tabla son las más comúnmente empleadas para la realización de MALDI MS (Zhu y Karas, 1995).

Matriz	Aplicación.
Ácido 2,5-Dihydroxybenzoico (DHB)	Péptidos, proteínas, lípidos y oligosacáridos.
Ácido 3,5-Dimetoxi-4-hidroxicinámico (ácido sinapínico)	Péptidos, proteínas y glicoproteínas.
Ácido α -Ciano-4-hidroxicinámico (CHCA)	Péptidos, proteínas, lípidos y oligonucleótidos.

- ESI-MS (electrospray ionization mass spectrometry). Es la segunda técnica más empleada para la identificación de péptidos descrita por Dole (Dole, 1968). La técnica del electrospray consiste fundamentalmente en que las cadenas polipeptídicas en solución son directamente forzadas a que la solución que las contiene cambie al estado gaseoso en una cámara de vacío y de ahí son trasladadas a un detector.

En 1990 se realiza una variante denominada nanoESI (Wilm y Mann, 1996), en la cual se utilizan flujos lentos de la solución (20-40 nL/min) que pasan a través de una aguja cargada que se mantiene con un elevado potencial, dispersando la solución en una fina nube de microgotas. El solvente que rodea las macromoléculas es evaporado rápidamente y el resultado de esto son iones de las macromoléculas cargados y así son introducidos a una fase gaseosa no destructiva; los protones adicionados en el proceso en el cual el analito pasa por la aguja le dan una carga adicional a la macromolécula y la relación m/z de la molécula puede ser analizada en la cámara de vacío por un detector de espectrometría de masa o MS (Figura 41).

El resultado de cada una de estas técnicas se puede entregar en diferentes formatos, de acuerdo a la base de datos empleadas para su identificación (Figura 41).

Si bien existe una gran combinación de técnicas que son utilizadas para la identificación de cambios postraduccionales, las descritas son las más empleadas para la identificación de proteínas que puedan servir con fines de posibles biomarcadores de diagnóstico, pronóstico o terapéuticos.

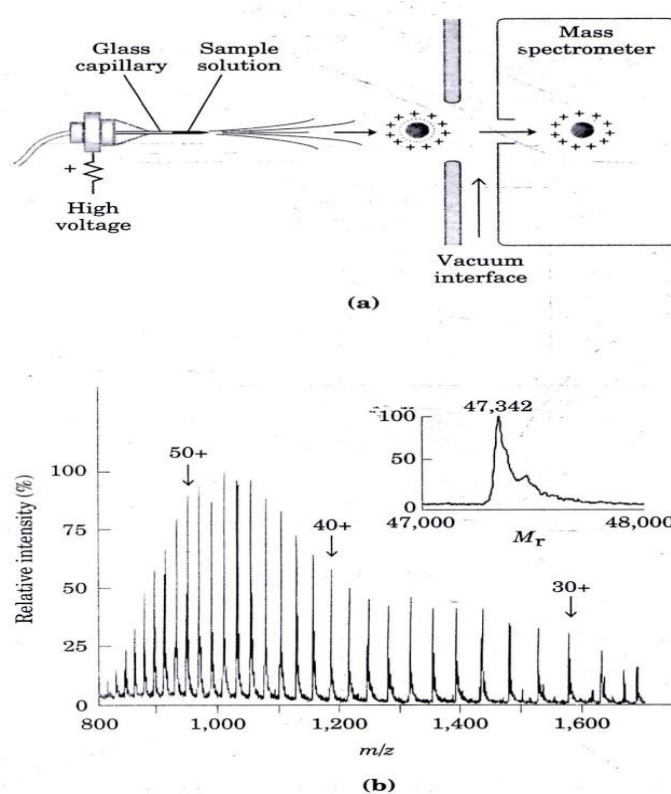


Figura 42. Técnica ESI-MS o electrospray. Esta técnica es empleada para la identificación de proteínas (Nelson y Cox, 2005).

ANEXO II.

Preparación de medios de cultivo y amortiguadores empleados en la identificación de biomarcadores fosforilados y glicosilados.

1) Medios de cultivo

- a. Formulación del Medio modificado de Eagle's Dulbecco's o DMEM(ATCC, No. de catálogo 30-2002), el cual fue utilizado en el cultivo de las líneas celulares MCF7, Hs578T y Hs578BsT.

Tabla 19. Componentes de 1L de medio de cultivo DMEM (ATCC, <http://www.atcc.org/>).

Componentes.	Cocentración (g/L)
Sales Inorgánicas.	
CaCl ₂ (anidro)	0.20000
Fe(NO ₃) ₃ ·9H ₂ O	0.00010
MgSO ₄ (anidro)	0.09770
KCl	0.40000
NaHCO ₃	1.50000
NaCl	6.40000
NaH ₂ PO ₄ ·H ₂ O	0.12500
Amino ácidos	
L-Arginina·HCl	0.08400
L-Cistina	0.06260
L-Glutamina	0.58400
Glicina	0.03000
L-Histidina·HCl·H ₂ O	0.04200
L-Isoleucina	0.10500
L-Lisina·HCl	0.14600
L-Metionina	0.03000
L-fenilalalina 0.06600	0.06600
L-Serina 0.04200	0.04200
L-Treonina 0.09500	0.09500
L-Triptofano 0.01600	0.01600
L-Tirosina·2Na·2H ₂ O	0.10379
L-Valina	0.09400

- b. Formulación de Medio de Leibovitz's (ATCC, No. de catálogo 30-2008), el cual fue empleado en el cultivo de la línea celular MDA-MB231

Tabla 20. Componentes del medio de Leibovitz's (ATCC, <http://www.atcc.org/>).

Componentes.	Cocentración (g/L)
Sales Inorgánicas.	
CaCl ₂ (anidro)	0.14000
MgCl ₂ ·6H ₂ O 0.20000	MgCl ₂ ·6H ₂ O 0.20000
MgSO ₄ (anidro)	0.09767
KCl	0.40000
KH ₂ PO ₄	0.06000
NaCl	8.00000
NaH ₂ PO ₄	0.19000
Amino ácidos	
L-Arginina(base)	0.50000
L-Alanina	0.22500
L-Asparagina·H ₂ O	0.50000
Glicina	0.20000
L-Cisteina (base)	0.12000
L-Glutamina	0.30000
L-Histidina (base)	0.25000
L-Isoleucina	0.12500
L-Leucina	0.12500
L-Serina 0.04200	0.20000
L-Treonina 0.09500	0.30000
L-Triptofano 0.01600	0.02000
L-Tirosina·2Na·2H ₂ O	0.43000
L-Valina	0.10000
L-Lisina·HCl	0.09370
L-Metionina	0.07500
L-fenilalanina	0.12500
Vitaminas.	
Cloruro de Colina	0.00100
Riboflavina-5-PO ₄ PO ₄ ·Na·2H ₂ O	0.00010
Mio-Inositol	0.00200
Nicotinamida	0.00100
Ácido D-Pantotenico	0.00100
Acido Fólico	0.00100
Tiamina·PO ₄ ·Cl·2H ₂ O	0.00100
D-Galactosa	0.90000
Rojo de Fenol, Sal de Sodio	0.01000
Piruvato de Sodio	0.55000

2) Preparación de Amortiguadores.

a) Buffer HEPES (1 L):

10mM HEPES 2.38=HEPES (M.M 238.31 g/mol).

150 mM NaCl 8.77=NaCl (M.M 58.44 g/mol).

Disolver HEPES y NaCl en 900 mL de agua destilada y ajustar pH a 7.4 usando NaOH.

Ajustar el volumen a 1 L.

b) Buffer 1xTBS (1 L):

10mM Tris-Cl=1.21g de Tris base (P.M 121.14 g/mol).

150 mM NaCl=8.77 g de NaCl (M.M 58.44 g/mol).

Disolver Tris-Cl y NaCl en 900 mL de agua destilada y ajustar pH a 7.4 usando HCl.

Ajustar el volumen a 1L.

c) Amortiguador para Electroforesis SDS LaemmLi. (10X).

Compuesto.	Concentración.	Masa o Volumen.
Tris base (P.M 121.1)	250 mM	303 g
Glicina (P.M 75.07)	1.92 M	1441 g
SDS (P.M 288.38)	1% (w/v)	100 g
Agua bi-distilada	-	10L

- El pH de esta solución puede no está ajustado.

d) Solución de Azul de Bromofenol.

Compuesto.	Concentración.	Masa o Volumen.
Azul de Bromofenol.	1%	100 mg
Tris-base	50 mM	60 mg
Agua bi-distilada	---	10mL

e) Solución SDS 10%.

Compuesto.	Concentración.	Masa o Volumen.
SDS (P.M 288.38)	10% (w/v)	5.0 g
Agua bi-distilada	---	50mL

- Filtrar la solución a través de un filtro de 0.45 µm. Almacenar a temperatura ambiente.

f) Persulfato de amonio (PSA) al 10%.

Compuesto.	Concentración.	Masa o Volumen.
Persulfato de Amonio (P.M 228.20)	10% (w/v)	1.0g
Agua bi-distilada	---	10mL

Si se presentan cristales, preparar nueva solución.

g) Azul de Coomassie al 5% G-250.

Compuesto.	Concentración.	Masa o Volumen.
Azul de Coomassie G-250	5%	0.5g
Agua bi-distilada	---	a 10 mL

- Agitar por algunos minutos hasta que se disperse el Aul de Coomassie. El colorante no se disuelve por completo.

h) DIGE.- 7M Urea, 2M Tiourea, 30 mM Tris, 4% CHAPS pH 8.5

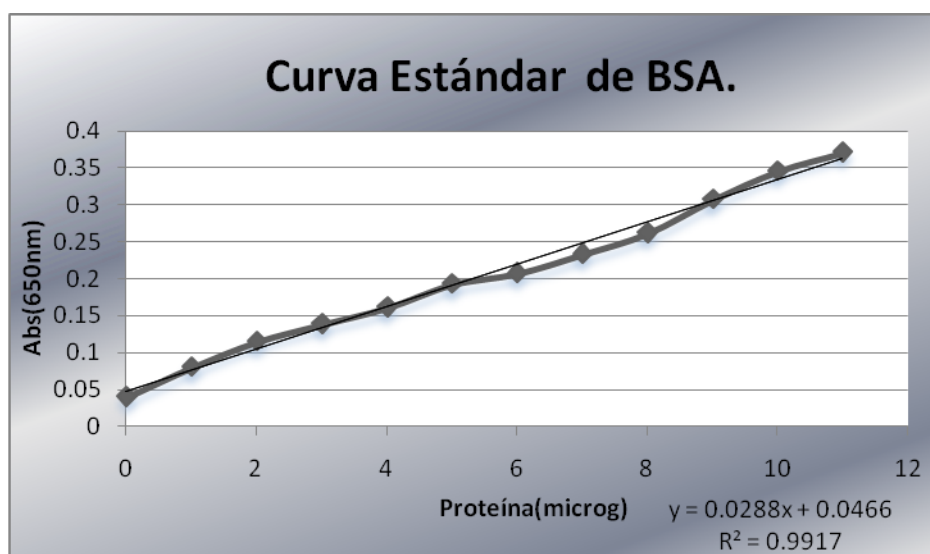
ANEXO III.

Tablas y cálculos para la cuantificación de las fracciones de fosfoproteínas y glicoproteínas obtenidas mediante cromatografía de afinidad.

A. Primera separación.

Tabla 21. Valores de absorbancia para el cálculo de la curva estándar de BSA.

CURVA ESTÁNDAR												
Proteína(mg)	0	1	2	3	4	5	6	7	8	9	10	11
Abs sin prot.	0.0496	0.0488	0.0485	0.0503	0.0487	0.0482	0.0468	0.0484	0.0486	0.0466	0.047	0.0418
Abs con prot.	0.0896	0.1294	0.1637	0.1892	0.2102	0.2412	0.2546	0.2822	0.3108	0.3541	0.3931	0.4131
Absorbancia.	0.04	0.0806	0.1152	0.1389	0.1615	0.193	0.2078	0.2338	0.2622	0.3075	0.3461	0.3713



Gráfica 1. Curva estándar de BSA (0µg-11µg). Absorbancia vs Concentración de Proteína.

GLICOPROTEÍNAS.

Tabla 22. Cuantificación de las fracciones de glicoproteínas obtenidas por cromatografía de afinidad de la línea celular MDA-MB 231. Fracciones fueron precipitadas con Acetona (-20) y reconstituidas en Buffer tris pH 7.4.

Muestra	MDA total.		MDA s/Glico		MDA GLICO	
Volumen cargado (μL).	3	3	3	3	3	3
Sin proteína.	0.0471	0.0475	0.0418	0.0425	0.0463	0.0476
Con proteína.	0.6617	0.6514	0.6603	0.6556	0.5764	0.5969
Diferencia	0.1644	0.1537	0.1683	0.1629	0.0799	0.0991
Proteína (mg).	2.4268	3.3636	4.0622	3.8038	-0.1675	0.7512
Concentración (mg/mL).	0.8089	1.1212	1.3541	1.2679	-0.0558	0.2504
Promedio.	0.9651		1.3110		0.0973	
Volumen total obtenido (μL).	1000		1000		300	
Volumen original de muestra (μL).	100		500		300	
Volumen de resuspensión (μL).	50		150		40	
Factor de disolución.	10		2		1	
μg resuspendidos.	48.2535		196.6507		3.8915	
Total (μg).	482.5346		393.3014		3.8915	

Tabla 23. Cuantificación de las fracciones de glicoproteínas obtenidas por cromatografía de afinidad de la línea celular MCF7. Fracciones fueron precipitadas con Acetona (-20) y reconstituidas en Buffer tris pH 7.4.

Muestra	Total MCF7		MCF7 s/Glico		MCF7 GLICO	
Volumen cargado (μL).	4	4	4	4	4	4
Sin proteína.	0.0477	0.0458	0.0479	0.0479	0.0487	0.0465
Con proteína.	1.0955	1.1087	1.0909	1.0112	1.0482	1.0177
Diferencia	0.139	0.1541	0.1342	0.0545	0.0907	0.0624
Proteína (mg).	2.0914	2.4746	1.9695	-0.0533	0.8655	0.1472
Concentración (mg/mL).	0.5228	0.6187	0.4924	-0.0133	0.2164	0.0368
Promedio.	0.5707		0.2395		0.1266	
Volumen total obtenido (μL).	1000		1000		300	
Volumen original de muestra (μL).	100		500		300	
Volumen de resuspensión (μL).	50		150		50	
Factor de disolución.	10		2		1	
μg resuspendidos.	28.5374		35.9296		6.3293	
Total (μg).	285.3744		71.8591		6.3293	

Tabla 24. Cuantificación de las fracciones de glicoproteínas obtenidas por cromatografía de afinidad de la línea celular Hs578T. Fracciones fueron precipitadas con Acetona (-20) y reconstituidas en Buffer tris pH 7.4.

Muestra	Hs578 total.		Hs578T s/GLICO		Hs578t glico	
Volumen cargado (μL).	4	4	4	4	4	4
Sin proteína.	0.0471	0.0481	0.0428	0.0415	0.0487	0.048
Con proteína.	0.6472	0.6525	0.6885	0.6737	0.6012	0.5929
Diferencia	0.1499	0.1542	0.1955	0.182	0.0525	0.0947
Proteína (mg).	3.2058	3.3827	5.0823	4.5267	-0.8025	0.9342
Concentración (mg/mL).	0.8014	0.8457	1.2706	1.1317	-0.2006	0.2335
Promedio.	0.8235		1.2011		0.0164	
Volumen total obtenido (μL).	1000		1000		300	
Volumen original de muestra (μL).	100		500		300	
Volumen de resuspensión (μL).	50		150		20	
Factor de disolución.	10		2		1	
μg resuspendidos.	41.1780		180.1698		0.3292	
Total (μg).	411.7798		360.3395		0.3292	

FOSFOPROTEÍNAS.

Tabla 25. Cuantificación de las fracciones de fosfoproteínas obtenidas por cromatografía de afinidad de la línea celular MDA-MB-231. Fracciones fueron precipitadas con Acetona (-20) y reconstituidas en Buffer tris pH 7.4.

Muestra	TOTALES MDA		S/FOSFO MDA		FOSFO MDA	
Volumen cargado (μL).	2	2	2	2	2	2
Sin proteína.	0.0483	0.0452	0.0499	0.0441	0.0479	0.0458
Con proteína.	0.7276	0.7303	0.7074	0.6907	0.6106	0.6008
Diferencia	0.0939	0.0997	0.0721	0.0612	-0.0227	-0.0304
Proteína (mg).	1.3937	1.5590	0.7726	0.4621	-1.9282	-2.1476
Concentración (mg/mL).	0.6969	0.7795	0.3863	0.2311	-0.9641	-1.0738
Promedio.	0.7382		0.3087		-1.0189	
Volumen total obtenido (μL).	5000		10000		2500	
Volumen original de muestra (μL).	500		500		2500	
Volumen de resuspensión (μL).	75		75		25	
Factor de disolución.	10		20		1	
μg resuspendidos.	55.3632		23.1517			
Total (μg).	553.6325		463.0342			

Tabla 26. Cuantificación de las fracciones de fosfoproteínas obtenidas por cromatografía de afinidad de la línea celular MCF7. Fracciones fueron precipitadas con Acetona (-20) y reconstituidas en Buffer tris pH 7.4.

Muestra	TOTALES MCF7		S/FOSFO MCF7		FOSFO MCF7	
Volumen cargado (μL).	2	2	2	2	2	2
Sin proteína.	0.0413	0.0426	0.0418	0.0417	0.0487	0.0465
Con proteína.	0.7171	0.7053	0.6862	0.6929	0.5765	0.5853
Diferencia	0.0904	0.0773	0.0590	0.0658	-0.0576	-0.0466
Proteína (mg).	1.2940	0.9208	0.3994	0.5932	-2.9225	-2.6091
Concentración (mg/mL).	0.6470	0.4604	0.1997	0.2966	-1.4613	-1.3046
Promedio.	0.5537		0.2481		-1.3829	
Volumen total obtenido (μL).	5000		10000		2500	
Volumen original de muestra (μL).	500		500		2500	
Volumen de resuspensión (μL).	100		70		25	
Factor de disolución.	10		20		1	
μg resuspendidos.	55.3704		17.3704			
Total (μg).	553.7037		347.4074			

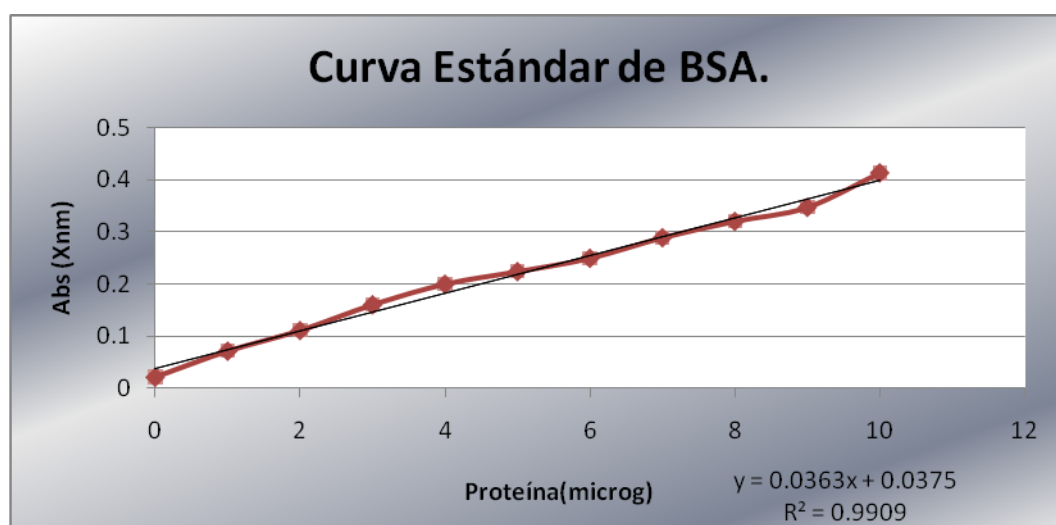
Tabla 27. Cuantificación de las fracciones de fosfoproteínas obtenidas por cromatografía de afinidad de la línea celular Hs578T. Fracciones fueron precipitadas con Acetona (-20) y reconstituidas en Buffer tris pH 7.4.

Muestra	TOTALES Hs578T		S/FOSFO Hs578T		FOSFO Hs578T	
Volumen cargado (μL).	2	2	2	2	2	2
Sin proteína.	0.0408	0.0436	0.0443	0.0415	0.0487	0.048
Con proteína.	0.7066	0.6985	0.6836	0.6775	0.6112	0.6219
Diferencia	0.0804	0.0695	0.0539	0.0506	-0.0229	-0.0115
Proteína (mg).	1.0091	0.6986	0.2541	0.1601	-1.9339	-1.6091
Concentración (mg/mL).	0.5046	0.3493	0.1271	0.0801	-0.9670	-0.8046
Promedio.	0.4269		0.1036		-0.8858	
Volumen total obtenido (μL).	5000		10000		2500	
Volumen original de muestra (μL).	500		500		2500	
Volumen de resuspensión (μL).	75		70		25	
Factor de disolución.	10		20		1	
μg resuspendidos.	32.0192		7.2493			
Total (μg).	320.1923		144.9858			

B. Segunda separación.

Tabla 28. Valores de absorbancia para calcular la curva estándar de BSA.

CURVA ESTÁNDAR											
Proteína(μg)	0	1	2	3	4	5	6	7	8	9	10
Abs. sin prot.	0.041	0.0425	0.0448	0.0423	0.0415	0.0424	0.0421	0.0418	0.0407	0.0416	0.0418
Abs. Con prot.	0.0628	0.1142	0.156	0.2029	0.2415	0.2665	0.2917	0.3309	0.361	0.3887	0.455
Absorbancia.	0.0218	0.0717	0.1112	0.1606	0.2	0.2241	0.2496	0.2891	0.3203	0.3471	0.4132



Gráfica 2. Curva estándar de BSA (0μg-10μg). Absorbancia vs. Concentración de Proteína.

GLICOPROTEÍNAS.

Tabla 29. Cuantificación de las fracciones de glicoproteínas obtenidas por cromatografía de afinidad de la línea celular MDA-MB 231. Fracciones fueron precipitadas con cloroformo-metanol (-20) y reconstituidas en Buffer DIGE.

Muestra	MDA total.		MDA s/Glico		MDA GLICO	
Volumen cargado (μL).	3	3	3	3	3	3
Sin proteína.	0.0403	0.0422	0.0410	0.0411	0.0414	0.0421
Con proteína.	0.7619	0.7558	0.7803	0.7956	0.6588	0.6653
Diferencia	0.2714	0.2634	0.2891	0.3043	0.1672	0.1730
Proteína (mg).	6.9038	8.6124	9.8421	10.5694	4.0096	4.2871
Concentración (mg/mL).	2.3013	2.8708	3.2807	3.5231	1.3365	1.4290
Promedio.	2.5860		3.4019		1.3828	
Volumen total obtenido (μL).	1000		1000		300	
Volumen original de muestra (μL).	100		500		300	
Volumen de resuspensión (μL).	50		150		40	
Factor de disolución.	10		2		1	
μg resuspendidos.	129.3017		510.2871		55.3110	
Total (μg).	1293.0172		1020.5742		55.3110	

Tabla 30. Cuantificación de las fracciones de glicoproteínas obtenidas por cromatografía de afinidad de la línea celular MCF7. Fracciones fueron precipitadas con cloroformo-metanol y reconstituidas en Buffer DIGE.

Muestra	Total MCF7		MCF7 s/Glico		MCF7 GLICO	
Volumen cargado (μL).	4	4	4	4	4	4
Sin proteína.	0.0402	0.0414	0.041	0.0413	0.0416	0.0414
Con proteína.	1.4255	1.4173	1.5969	1.6112	1.1412	1.1887
Diferencia	0.4765	0.4671	0.6471	0.6611	0.1908	0.2385
Proteína (mg).	10.6574	10.4188	14.9873	15.3426	3.4061	4.6168
Concentración (mg/mL).	2.6643	2.6047	3.7468	3.8357	0.8515	1.1542
Promedio.	2.6345		3.7912		1.0029	
Volumen total obtenido (μL).	1000		1000		300	
Volumen original de muestra (μL).	100		500		300	
Volumen de resuspensión (μL).	50		150		50	
Factor de disolución.	10		2		1	
μg resuspendidos.	131.7259		568.6865		50.1428	
Total (μg).	1317.2589		1137.3731		50.1428	

Tabla 31. Cuantificación de las fracciones de glicoproteínas obtenidas por cromatografía de afinidad de la línea celular Hs578T. Fracciones fueron precipitadas con cloroformo-metanol y reconstituidas en Buffer DIGE.

Muestra	Hs578 total.		Hs578T s/GLICO		Hs578t glico	
Volumen cargado (µL).	4	4	4	4	4	4
Sin proteína.	0.0415	0.0421	0.0428	0.0415	0.041	0.0424
Con proteína.	0.8634	0.8677	0.9785	1.0037	0.8661	0.8593
Diferencia	0.3717	0.3754	0.4855	0.512	0.3251	0.3667
Proteína (mg).	12.3333	12.4856	17.0165	18.1070	10.4156	12.1276
Concentración (mg/mL).	3.0833	3.1214	4.2541	4.5267	2.6039	3.0319
Promedio.	3.102366255		4.390432099		2.817901235	
Volumen total obtenido (µL).	1000		1000		300	
Volumen original de muestra (µL).	100		500		300	
Volumen de resuspensión (µL).	50		150		20	
Factor de disolución.	10		2		1	
µg resuspendidos.	155.1183		658.5648		56.3580	
Total (µg).	1551.1831		1317.1296		56.3580	

FOSFOPROTEÍNAS.

Tabla 32. Cuantificación de las fracciones de glicoproteínas obtenidas por cromatografía de afinidad de la línea celular MDA. Fracciones fueron precipitadas con cloroformo-metanol y reconstituidas en Buffer DIGE.

Muestra	TOTALES MDA		S/FOSFO MDA		FOSFO MDA	
Volumen cargado (µL).	2	2	2	2	2	2
Sin proteína.	0.0463	0.0432	0.0409	0.0421	0.0414	0.0416
Con proteína.	0.7876	0.7803	0.7174	0.7207	0.8011	0.8378
Diferencia	0.1559	0.1517	0.0911	0.0932	0.1743	0.2108
Proteína (mg).	3.1601	3.0405	1.3140	1.3738	3.6843	4.7242
Concentración (mg/mL).	1.5801	1.5202	0.6570	0.6869	1.8422	2.3621
Promedio.	1.5501		0.6719		2.1021	
Volumen total obtenido (µL).	5000		10000		2500	
Volumen original de muestra (µL).	500		500		2500	
Volumen de resuspensión (µL).	75		75		25	
Factor de disolución.	10		20		1	
µg resuspendidos.	116.2607		50.3953		52.5534	
Total (µg).	1162.6068		1007.9060		52.5534	

Tabla 33. Cuantificación de las fracciones de glicoproteínas obtenidas por cromatografía de afinidad de la línea celular MCF7. Fracciones fueron precipitadas con cloroformo-metanol y reconstituidas en Buffer DIGE.

Muestra	TOTALES MCF7		S/FOSFO MCF7		FOSFO MCF7	
Volumen cargado (μL).	2	2	2	2	2	2
Sin proteína.	0.0409	0.0405	0.0425	0.0421	0.0417	0.0411
Con proteína.	0.7688	0.7526	0.7245	0.7322	0.8118	0.8238
Diferencia	0.1425	0.1267	0.0966	0.1047	0.1847	0.1973
Proteína (mg).	2.7783	2.3282	1.4707	1.7014	3.9806	4.3396
Concentración (mg/mL).	1.3892	1.1641	0.7353	0.8507	1.9903	2.1698
Promedio.	1.2766		0.7930		2.0801	
Volumen total obtenido (μL).	5000		10000		2500	
Volumen original de muestra (μL).	500		500		2500	
Volumen de resuspensión (μL).	100		70		25	
Factor de disolución.	10		20		1	
μg resuspendidos.	127.6638		55.5114		52.0014	
Total (μg).	1276.6382		1110.2279		52.0014	

Tabla 34. Cuantificación de las fracciones de glicoproteínas obtenidas por cromatografía de afinidad de la línea celular Hs578T. Fracciones fueron precipitadas con cloroformo-metanol y reconstituidas en Buffer DIGE.

Muestra	TOTALES Hs578T		S/FOSFO Hs578T		FOSFO Hs578T	
Volumen cargado (μL).	2	2	2	2	2	2
Sin proteína.	0.0406	0.0409	0.0413	0.0425	0.0416	0.0420
Con proteína.	0.7766	0.7885	0.7236	0.7305	0.8335	0.8458
Diferencia	0.1506	0.1622	0.0969	0.1026	0.2065	0.2184
Proteína (mg).	3.0091	3.3396	1.4792	1.6416	4.6017	4.9407
Concentración (mg/mL).	1.5046	1.6698	0.7396	0.8208	2.3009	2.4704
Promedio.	1.5872		0.7802		2.3856	
Volumen total obtenido (μL).	5000		10000		2500	
Volumen original de muestra (μL).	500		500		2500	
Volumen de resuspensión (μL).	75		70		25	
Factor de disolución.	10		20		1	
μg resuspendidos.	119.0385		54.6140		59.6403	
Total (μg).	1190.3846		1092.2792		59.6403	

Tabla 35. Cuantificación de las fracciones de glicoproteínas obtenidas por cromatografía de afinidad de la línea celular Hs578BsT. Fracciones fueron precipitadas con cloroformo-metanol y reconstituidas en Buffer DIGE.

Muestra	TOTALES Hs578BsT		S/FOSFO Hs578BsT		FOSFO Hs578BsT	
Volumen cargado (μL).	2	2	2	2	2	2
Sin proteína.	0.0401	0.0414	0.0416	0.0414	0.0410	0.0414
Con proteína.	0.6567	0.6523	0.6892	0.6998	0.6522	0.6611
Diferencia	0.1166	0.1109	0.1476	0.1584	0.1112	0.1197
Proteína (mg).	0.5154	0.3208	1.5734	1.9420	0.3311	0.6212
Concentración (mg/mL).	0.2577	0.1604	0.7867	0.9710	0.1655	0.3106
Promedio.	0.2090		0.8788		0.2381	
Volumen total obtenido (μL).	2000		2000		2500	
Volumen original de muestra (μL).	250		2000		2500	
Volumen de resuspensión (μL).	75		75		25	
Factor de disolución.	8		1		1	
μg resuspendidos.	15.6783		65.9130		5.9514	
Total (μg).	125.4266		65.9130		5.9514	

ANEXO III.

Resultados de LC/MS-MS/MS de las fosfoproteínas y glicoproteínas seleccionadas como posibles biomarcadores.

1) Nucleolina.

Mascot Search Results

Protein View

Match to: **gi|189306** Score: **408**
nucleolin [Homo sapiens]
Found in search of mas4D.tmp

Nominal mass (M₀): **76298**; Calculated pI value: **4.59**
NCBI BLAST search of **gi|189306** against nr
Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Homo sapiens](#)

Variable modifications: Deamidated (NQ), Oxidation (M), Propionamide (C)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **15%**

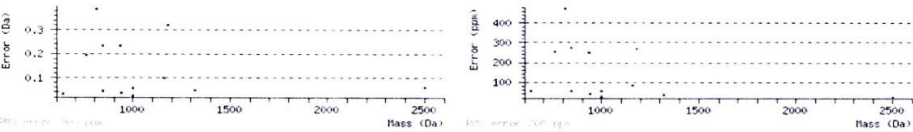
Matched peptides shown in Bold Red

1 MVKLAKAGKN QGDPKMAPF PKEVEEDSED EEMSEDEEDD SSGEEVVIQ
51 KKGKAAATS AKKVVSPTK KVAVATPAKK AAVTPGKAA ATPAKKTVP
101 AKAVTTPGKK GATPGKALVA TPGKGAAP AKGAKNGKNA KKEDSDEED
151 DDSEDEEDD EDEDEDEEI EPAAMKAAA APASEDEDD DDEDEDDDD
201 DEEDDSEEA **METTPAKGKK** AAKVVPVAK NVAEDEDEE DDEDEDDDD
251 EDEDDDDDD DEEEEEEEE EPVKEAPGKR KKEMAKQAA PEAKKQKVEG
301 TEPTTAFNLF VGNLNFNKA PELKTGISDV FAKNDLAVVD VRIGMTRKFG
351 YVDFESAEKL EKALELTGLK VFGNEIKLEK PKGDSRKER DARTLLAKNL
401 PYKVTQDELK EVFEDAAEIR LVSKDGGKSG FAXIEFKTEA DAEKTFEEKQ
451 GTEIDGRSIS LYYTGERGQN QDYRGKNST WSGESKTLVL SNLSYSATEE
501 TLQEVFEKAT FIKVPQNQNG KSKGYAFIEF ASFDEAKEAL NSCNKRRIEG
551 RAIRLELQGP RGSFNARSQP SKTLFVKGLS EDTTEETLKE SFDGKVRARI
601 VTDRETGSSK GFGFVDFNSE EDAKEAMEDG EIDGNKVTL WAKPKGGGFF
651 GRRGGRRGG GRRGGRRGG GFGGRRGRGG FGRGGFRGG RGGGGDHKFP
701 GKKTKE

Show predicted peptides also

Sort Peptides By ☒ Residue Number ☐ Increasing Mass ☐ Decreasing Mass

Start - End	Observed	Mr (expt)	Mr (calc)	Delta	Miss Sequence
72 - 79	378.8306	755.6467	755.4541	0.1926	0 K.VAVATPAK.K (Ions score 56)
325 - 333	469.3693	936.7240	936.4917	0.2323	0 K.TGISDVFAK.N (Ions score 36)
334 - 342	1000.5669	999.5597	999.5349	0.0248	0 K.NDLAVVDVR.I (Ions score 18)
334 - 342	500.8037	999.55928	999.5349	0.0579	0 K.NDLAVVDVR.I (Ions score 31)
363 - 370	844.5604	843.5531	843.5065	0.0466	0 K.ALELTGLK.V (Ions score 19)
363 - 370	422.8763	843.7381	843.5065	0.2316	0 K.ALELTGLK.V (Ions score 39)
399 - 403	634.3905	633.3832	633.3486	0.0347	0 K.NLTPYK.V (Ions score 20)
411 - 420	589.9475	1177.8805	1177.5615	0.3190	0 K.EVFEDAAEIR.L (Ions score 58)
430 - 437	940.5509	939.5437	939.5065	0.0371	0 K.GIATIEFK.T (Ions score 24)
458 - 467	580.8452	1159.6759	1159.5761	0.0998	0 R.SISLYTTGK.G (Ions score 52)
487 - 508	834.4468	2500.3184	2500.2584	0.0600	0 K.TLVLSNLSYSATEETLQVFEK.A (Ions score 49)
555 - 561	406.9274	811.8403	811.4552	0.3851	0 R.LELQGP.R (Ions score 41)
578 - 589	661.8453	1321.6760	1321.6249	0.0511	0 K.GLSEDTTEETLK.E (Ions score 20)



LOCUS AAA59954 707 aa linear PRI 07-JAN-1995
DEFINITION nucleolin [Homo sapiens].
ACCESSION AAA59954
VERSION AAA59954.1 GI:189306
DBSOURCE locus HUMNUCLEO accession [M60858.1](#)
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM [Homo sapiens](#)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
Catarrhini; Hominidae; Homo.

```
/coded_by="join(M60858.1:1181..1198,M60858.1:2159..2275,
M60858.1:3439..3916,M60858.1:4587..4784,
M60858.1:4889..4975,M60858.1:5160..5301,
M60858.1:6307..6431,M60858.1:7037..7160,
M60858.1:7620..7777,M60858.1:8292..8415,
M60858.1:8652..8785,M60858.1:9279..9405,
M60858.1:9792..10006,M60858.1:10140..10216)"
/db_xref="GDB:G00-125-908"
```

Mascot: <http://www.matrixscience.com/>

2) Ribonucleoproteína U (HnRNP).

Mascot Search Results: Protein View

Page 1 of 2

Mascot Search Results

Protein View

Match to: [gi|32358](#) Score: 370
hnRNP U protein [Homo sapiens]
Found in search of mas4D.tmp

Nominal mass (M_r): 88890; Calculated pI value: 5.60
NCBI BLAST search of [gi|32358](#) against nr
Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Homo sapiens](#)

Variable modifications: Deamidated (NQ), Oxidation (M), Propionamide (C)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: 13%

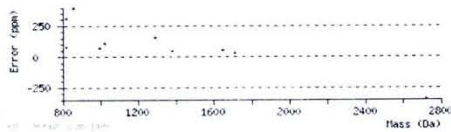
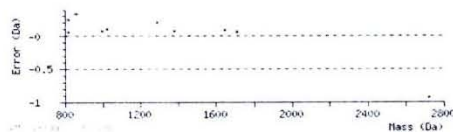
Matched peptides shown in Bold Red

```
1 YSSSPVNVKK LKYSLEKEEL KKRRLSDKGL KAEIMERLQA ALDDEEAGGR
51 PAMEPONGSL DLGGDSAGRS GAGLEQEAFA GGDEFEFEFE EEEEGISALD
101 GDQMELGREN GAAGAADSGP MEFEAAASED ENGDDQGFQE GEDELGDREE
151 GAGDENGHGE QQPQPATQQ QPQQQQRGAA KEAAGKSSGP TSLFAVTAP
201 PGARQQQQQA GGDGKTEQKG GKKRGRVRRP REDHGRGYFE YIEENKYSRA
251 KSPQPPVEEE DEHFDDTVVC LDTYNCDLHF KISRDLAS SLTMSFAFL
301 WAGGRASYGV SKGKVCPEMK VTEKIPVRHL YTKDIDIHEV RIGWSLTTSF
351 MLLGEEEFYSY GYSLKGIKTC NCETEDYGEK FENNVITCF ANFESDEVEL
401 SYAKNQDQLG VAFKISKEVL AGRPLEPHVL CHNCAVEFNF GQKEKPYFPI
451 PEYTFITQNV FLEDRVRGPK GPEEKKDCLE VMNGLPGAG KTWVYTHAA
501 ENPGKYNILG TWNTMDKQNV AGFKKQADT GKINTLLQRA PGCLGKFEI
551 AAKKKRNFIL DQTNVSAQA RKKKCLFAGF QRKAVVCPK DEDYKQIQK
601 KAEVEGKDLF EHAVLKMKN FTLPEVAECF DEITYVELQK EEAQKLEQY
651 KEESKALPP EKKQNTGSKK SNKNSGKQK FNAGGGHGR GGLNMRGGNF
701 RGGAGPNRGG YNRRGNPQR GGGGGSGGI GYPYPRAPVF PGRGYSNRG
751 NYNRGMPNR GNYNQFRGR GNNRGYKNS QGYNQWQQQ FGGQKPSQH
801 YHGGYY
```

Show predicted peptides also

Sort Peptides By ☒ Residue Number ☐ Increasing Mass ☐ Decreasing Mass

Start - End	Observed	Mr (expt)	Mr (calc)	Delta	Miss	Sequence
187 - 204	857.9875	1713.9604	1713.9050	0.0554	0	K.SSGPTSLFAVTAPPGAR Q (ions score 69)
237 - 246	646.3953	1290.7761	1290.5768	0.1993	0	R.GYFEYIEENK.Y (ions score 26)
334 - 341	498.7952	995.5758	995.5036	0.0722	0	K.DIDIHEVR.I (ions score 37)
444 - 465	908.8046	2723.3921	2724.3323	-0.9402	0	K.EKPYFPIPEEYTFIQNVPLEDR.V Deamidated (NQ) (ions score 33)
506 - 517	691.8870	1381.7595	1381.6911	0.0684	0	K.YNLTGTINTMK.M (ions score 23)
533 - 539	429.4341	856.8536	856.5130	0.3406	0	K.LNTLLQR.A (ions score 50)
547 - 553	819.5367	818.5295	818.4650	0.0644	0	K.FIEIAAR.K (ions score 31)
547 - 553	410.3668	818.7190	818.4650	0.2540	0	K.FIEIAAR.K (ions score 31)
557 - 571	824.4743	1646.9340	1646.8376	0.0964	0	R.NFILDQTNVSAQAQR.R (ions score 28)
608 - 616	511.3428	1020.6711	1020.5604	0.1107	0	K.DLPERAVLR.M (ions score 74)



LOCUS CAA46472 806 aa linear PRI 18-APR-2005
DEFINITION hnRNP U protein [Homo sapiens].
ACCESSION CAA46472
VERSION CAA46472.1 GI:32358
DBSOURCE embl accession [X65488.1](#)
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM [Homo sapiens](#)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
Catarrhini; Hominidae; Homo.
1 (residues 1 to 806)
REFERENCE
AUTHORS Kiledjian, M. and Dreyfuss, G.
TITLE Primary structure and binding activity of the hnRNP U protein:
binding RNA through RGG box
JOURNAL EMBO J. 11 (7), 2655-2664 (1992)
PUBMED [1628675](#)
REFERENCE 2 (residues 1 to 806)
AUTHORS Kiledjian, M.
TITLE Direct Submission

http://www.matrixscience.com/cgi/protein_view.pl?file=../data/20100603/FteAifaOE.dat... 07/09/2010

1992) M. Kiledjian, Howard Hughes Med Institute
 University of Pennsylvania, 422 Curie Boulevard, Philadelphia,
 PA 19104-6148, USA

FEATURES Location/Qualifiers
source 1..806
 /organism="Homo sapiens"
 /strain="Hela D98"
 /db_xref="taxon:9606"
 /clone="pGem-U21.1"
 /clone_lib="lambda GT11"
Protein 1..806
 /product="hnRNP U protein"
Region 328..443
 /region_name="SPRY"
 /note="SPRY domain; cl02614"
 /db_xref="CDD:155014"
Region 480..630
 /region_name="NK"
 /note="Nucleoside/nucleotide kinase (NK) is a protein
 superfamily consisting of multiple families of enzymes
 that share structural similarity and are functionally
 related to the catalysis of the reversible phosphate group
 transfer from nucleoside...; cl11962"
 /db_xref="CDD:125376"
CDS 1..806
 /gene="U21.1"
 /coded_by="X65488.1:42..2462"
 /db_xref="GDB:347585"
 /db_xref="GOA:Q00839"
 /db_xref="HGNC:5048"
 /db_xref="PDB:1ZBJ"
 /db_xref="UniProtKB/Swiss-Prot:Q00835"

Mascot: <http://www.matrixscience.com/>

http://www.matrixscience.com/cgi/protein_view.pl?file=../data/20100603/FteAifaOE.dat... 07/09/2010

3) Subunidad β de la proteína de choque térmico (HsP 90- β).

Mascot Search Results: Protein View

Page 1 of 4

MASCOT Search Results

Protein View

Match to: [gi|20149594](#) Score: **561**
heat shock protein HSP 90-beta [Homo sapiens]
Found in search of mas52.tmp

Nominal mass (M₀): **83212**; Calculated pI value: **4.97**
NCBI BLAST search of [gi|20149594](#) against nr
Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Homo sapiens](#)
Links to retrieve other entries containing this sequence from NCBI Entrez:
[gi|17865718](#) from [Homo sapiens](#)
[gi|1386786](#) from [Homo sapiens](#)
[gi|13436257](#) from [Homo sapiens](#)
[gi|15215418](#) from [Homo sapiens](#)
[gi|15680260](#) from [Homo sapiens](#)
[gi|16876955](#) from [Homo sapiens](#)
[gi|14404550](#) from [Homo sapiens](#)
[gi|16744975](#) from [Homo sapiens](#)
[gi|15670416](#) from [Homo sapiens](#)
[gi|1674451](#) from [Homo sapiens](#)
[gi|11962466](#) from [Homo sapiens](#)
[gi|11962466](#) from [Homo sapiens](#)
[gi|11962466](#) from [Homo sapiens](#)
[gi|123991523](#) from [synthetic construct](#)
[gi|123999414](#) from [synthetic construct](#)
[gi|168277956](#) from [synthetic construct](#)
[gi|118905381](#) from [Homo sapiens](#)

Variable modifications: Deamidated (NQ), Oxidation (M), Propionamide (C)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **21%**

Matched peptides shown in **Bold Red**

```
1 MPREVVHGEV ERTFAFQAE IAQLMSLIIN TFYSNKEIFL RELISNASDA
51 LDKIRYESLT DPSKLDGKE LKIDIIHPNQ ERTLTIVDTG IGMTKADLIN
101 NLGTIAKSGT KAFMEALQAG ADISMIGQFG VGFYSAYLVA EKVVVITKHN
151 DDEQYAWESS AGGSFTVRAD HGEPIGRGTG VILHLKEDQT EYLEERRVKE
201 VVKHNSQFQG YPITLYLEKE REKEISDDEA EEKGEKEEE DKDDEEKPKI
251 EDVGSDEEDD SGKDKKKKTK KIKEKYIDQE ELNKTPIWT RNPDIITQEE
301 YGEFYKSLTN DWEDHLAVKH FSVEGQLEFR ALLFIPRRAP FDLFENKKKK
351 NNKILYVRRV FIMDSCELI PEYLNFRGV VDSDELPLNI SREMIQQSKI
401 LKVIKKNIVK KCELPSELA EDKENYKKFY EAFSKNLKLG IHEDSTNRRR
451 LSELRLYHTS QSGDEMTSLS EYVSRMKETO KSIYYITGES KEQVANSAPV
501 ERVRKRGFEV VYMTEPIDY CVQQLKEFDG KSLVSVTKEG LELPEDEEEK
551 KMEESKARF ENLCKLMKEI LDKKVERVTE SNRLVSPCC IVTSYGVWTA
601 NMRIMKAQA LRDNSTMGYM MAKKHLEINP DHPIVETLRQ KAEAKNDKA
651 VKDLVVLIFE TALLSSGSL EDPQTHSNRI YRMKLGIGI DEDEVAASEP
701 NAAVPDEIPP LEGDEASRM REVD
```

Show predicted peptides also

Sort Peptides By ☒ Residue Number ☐ Increasing Mass ☐ Decreasing Mass

Start - End	Observed	Mr (expt)	Mr (calc)	Delta	Miss	Sequence
37 - 41	677.4770	576.4697	576.3908	0.0789	0	K.EIPLR.E (ions score 31)
42 - 53	638.3880	1274.7615	1274.6354	0.1261	0	R.ELISNASDALDK.I (ions score 78)
56 - 64	520.3491	1038.6836	1038.4869	0.1967	0	R.YESLTDPSK.L (ions score 20)
96 - 107	621.9072	1241.7998	1241.6979	0.1019	0	K.ADLINNLGTIAK.S (ions score 44)
143 - 148	658.4870	657.4798	657.4425	0.0373	0	K.VVITK.H (ions score 33)
181 - 186	361.9183	721.8221	721.4850	0.3371	0	K.VILHLK.E (ions score 39)
187 - 196	656.3643	1310.7141	1310.5626	0.1515	0	K.EDQTEYLEER.R (ions score 34)
205 - 219	905.0201	1808.0256	1808.9349	-0.9093	0	K.HSQFIGYPITLYLEK.E Deamidated (NQ) (ions score 17)
276 - 284	576.3857	1150.7568	1150.5506	0.2062	0	K.YIDQEELNK.T (ions score 25)
285 - 291	451.3274	900.6401	900.5181	0.1220	0	K.TKPIWTR.N (ions score 24)
331 - 337	415.3940	828.7734	828.5221	0.2513	0	R.ALLFIPR.R (ions score 47)
379 - 392	757.4304	1512.8463	1512.7784	0.0679	0	R.GVDSDELPLNISR.E (ions score 63)
429 - 435	446.3435	890.6724	890.4174	0.2550	0	K.FYEAFSK.N (ions score 42)
439 - 448	571.3334	1140.6523	1140.5523	0.1000	0	K.LGIHEDSTNR.R (ions score 14)
492 - 502	625.3797	1248.7449	1248.6098	0.1350	0	K.EQVANSAPVER.V (ions score 38)
532 - 538	733.5063	732.4990	732.4382	0.0608	0	K.SLVSVTK.E (ions score 40)
575 - 583	523.4138	1044.8130	1044.5927	0.2203	1	K.VKVTISNR.L (ions score 12)

http://www.matrixscience.com/cgi/protein_view.pl?file=../data/20100603/FteAifsEO.dat... 07/09/2010

DEFINITION heat shock protein Hsp 90-beta [Homo sapiens].
 ACCESSION NP_031381
 VERSION NP_031381.2 GI:20149594
 DBSOURCE REFSEQ: accession [NM_007355.2](#)
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM [Homo sapiens](#)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
 Catarrhini; Hominidae; Homo.
 REFERENCE 1 (residues 1 to 724)
 AUTHORS Zurawska,A., Urbanski,J., Matullene,J., Baraniak,J., Klejman,M.P.,
 Flilpek,S., Matulis,D. and Bieganowski,P.
 TITLE Mutations that increase both Hsp90 ATPase activity in vitro and
 Hsp90 drug resistance in vivo
 JOURNAL Biochim. Biophys. Acta 1803 (5), 575-583 (2010)
 PUBMED 20226816
 REMARK GeneRIF: findings present novel Hsp90 mutants that render cells
 resistant to Hsp90 inhibitors; show that the resistance depends on
 the increased ATPase turnover due to enhanced interaction with Aha1
 REFERENCE 2 (residues 1 to 724)
 AUTHORS Chadli,A., Felts,S.J., Wang,Q., Sullivan,W.F., Botuyan,M.V.,
 Fauq,A., Ramirez-Alvarado,M. and Mer,G.
 TITLE Celestrol inhibits Hsp90 chaperoning of steroid receptors by
 inducing fibrillization of the Co-chaperone p23
 JOURNAL J. Biol. Chem. 285 (6), 4224-4231 (2010)
 PUBMED 19963133
 REMARK GeneRIF: Data show that the small molecule celestrol inhibits the
 Hsp90 chaperoning machinery by inactivating the co-chaperone p23,
 resulting in a more selective destabilization of steroid receptors.
 REFERENCE 3 (residues 1 to 724)
 AUTHORS Zhang,T., Li,Y., Yu,Y., Zou,F., Jiang,Y. and Sun,D.
 TITLE Characterization of celestrol to inhibit hsp90 and cdc37
 interaction
 JOURNAL J. Biol. Chem. 284 (51), 35381-35389 (2009)
 PUBMED 19556714
 REMARK GeneRIF: celestrol may represent a new class of Hsp90 inhibitor by
 modifying Hsp90 C terminus to allosterically regulate its chaperone
 activity and disrupt Hsp90-Cdc37 complex.
 REFERENCE 4 (residues 1 to 724)
 AUTHORS Hosgood,H.D. III, Zhang,L., Shen,M., Berndt,S.I., Vermeulen,R.,
 Li,G., Yin,S., Yeager,M., Yuenger,J., Rothman,N., Chanock,S.,
 Smith,M. and Lan,Q.
 TITLE Association between genetic variants in VEGF, ERCC3 and
 occupational benzene haematotoxicity
 JOURNAL Occup Environ Med 66 (12), 848-853 (2009)
 PUBMED 19773279
 REMARK GeneRIF: Observational study of gene-disease association. (HUGO
 Navigator)
 REFERENCE 5 (residues 1 to 724)
 AUTHORS van de Sluis,B., Groot,A.J., Vermeulen,J., van der Wall,E., van
 Diest,P.J., Wijnenga,C., Klomp,L.W. and Voorts,M.
 TITLE COMMD1 Promotes pVHL and O2-independent Proteolysis of HIF-1alpha
 via HSP90/70
 JOURNAL PLoS ONE 4 (10), E7332 (2009)
 PUBMED 19802386
 REMARK GeneRIF: COMMD1 has a role in conjunction with HSP90beta/HSP70 in
 the ubiquitin and O(2)-independent regulation of HIF-1alpha
 Publication Status: Online-Only
 REFERENCE 6 (sites)
 AUTHORS Rikova,K., Guo,A., Zeng,Q., Possemato,A., Yin,J., Haack,H.,
 Nardone,J., Lee,K., Reeves,C., Li,Y., Hu,Y., Tan,Z., Stokes,X.,
 Sullivan,L., Mitchell,J., Wetzel,R., MacNeill,J., Ren,J.M.,
 Yuan,J., Bakalarski,C.E., Villen,J., Kornhauser,J.M., Smith,B.,
 Li,D., Zhou,X., Gygi,S.P., Gu,T.L., Polakiewicz,R.D., Rush,J. and
 Comb,M.J.
 TITLE Global survey of phosphotyrosine signaling identifies oncogenic
 kinases in lung cancer
 JOURNAL Cell 131 (6), 1190-1203 (2007)
 PUBMED 18083107
 REFERENCE 7 (sites)

http://www.matrixscience.com/cgi/protein_view.pl?file=../data/20100603/FteAifsEO.dat... 07/09/2010

AUTHORS Sugiyama,N., Masuda,T., Shinoda,K., Nakamura,A., Tomita,M. and Ishihama,Y.
 TITLE Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications
 JOURNAL Mol. Cell Proteomics 6 (6), 1103-1109 (2007)
 PUBMED 17322306
 REFERENCE 8 (sites)

AUTHORS Matsuo,K., Ballif,B.A., Smogorzewska,A., McDonald,E.R. III, Hurov,K.E., Luo,J., Bakalarski,C.E., Zhao,Z., Solimini,N., Lerenthal,Y., Shih,Y., Gygi,S.P. and Elledge,S.J.
 TITLE ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage
 JOURNAL Science 316 (5828), 1160-1166 (2007)
 PUBMED 17675192
 REFERENCE 9 (sites)

AUTHORS Nousiainen,M., Sillje,H.H., Sauer,G., Nigg,E.A. and Korner,R.
 TITLE Phosphoproteome analysis of the human mitotic spindle
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 103 (14), 5391-5396 (2006)
 PUBMED 16565220
 REFERENCE 10 (sites)

AUTHORS Rush,J., Moritz,A., Lee,K.A., Guo,A., Goss,V.L., Spek,E.J., Zhang,H., Zha,X.M., Polakiewicz,R.D. and Comb,M.J.
 TITLE Immunoaffinity profiling of tyrosine phosphorylation in cancer cells
 JOURNAL Nat. Biotechnol. 23 (1), 94-101 (2005)
 PUBMED 15597455
 REFERENCE 11 (sites)

AUTHORS Beausoleil,S.A., Jedrychowski,M., Schwartz,D., Elias,J.E., Villen,J., Li,J., Cohn,M.A., Cantley,L.C. and Gygi,S.P.
 TITLE Large-scale characterization of HeLa cell nuclear phosphoproteins
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (33), 12130-12135 (2004)
 PUBMED 15307935
 REFERENCE 12 (residues 1 to 724)

AUTHORS Tang,P.Z., Gannon,M.J., Andrew,A. and Miller,D.
 TITLE Evidence for oestrogenic regulation of heat shock protein expression in human endometrium and steroid-responsive cell lines
 JOURNAL Eur. J. Endocrinol. 133 (5), 598-605 (1995)
 PUBMED 7581591
 REFERENCE 13 (residues 1 to 724)

AUTHORS Rebbe,N.F., Hickman,W.S., Ley,T.J., Stafford,D.W. and Hickman,S.
 TITLE Nucleotide sequence and regulation of a human 90-kDa heat shock protein gene
 JOURNAL J. Biol. Chem. 264 (25), 15006-15011 (1989)
 PUBMED 2768045
 REFERENCE 14 (residues 1 to 724)

AUTHORS Lees-Miller,S.P. and Anderson,C.W.
 TITLE Two human 90-kDa heat shock proteins are phosphorylated in vivo at conserved serines that are phosphorylated in vitro by casein kinase II
 JOURNAL J. Biol. Chem. 264 (5), 2431-2437 (1989)
 PUBMED 2495519
 REFERENCE 15 (residues 1 to 724)

AUTHORS Hoffmann,T. and Hovemann,B.
 TITLE Heat-shock proteins, Hsp84 and Hsp86, of mice and men: two related genes encode formerly identified tumour-specific transplantation antigens
 JOURNAL Gene 74 (2), 491-501 (1988)
 PUBMED 2463676
 REFERENCE 16 (residues 1 to 724)

AUTHORS Rebbe,N.F., Ware,J., Bertina,R.M., Modrich,P. and Stafford,D.W.
 TITLE Nucleotide sequence of a cDNA for a member of the human 90-kDa heat-shock protein family
 JOURNAL Gene 53 (2-3), 235-245 (1987)
 PUBMED 2301534

COMMENT PROVISIONAL REFSEQ: This record has not yet been subject to final NCBI review. The reference sequence was derived from [BC004975.1](#). On Apr 15, 2002 this sequence version replaced gi:1680307.

Summary: HSP90 proteins are highly conserved molecular chaperones that have key roles in signal transduction, protein folding, protein degradation, and morphologic evolution. HSP90 proteins normally associate with other cochaperones and play important roles in folding newly synthesized proteins or stabilizing and refolding denatured proteins after stress. There are 2 major cytosolic HSP90 proteins, HSP90AA1 (MIM 140571), an inducible form, and HSP90AB1, a constitutive form. Other HSP90 proteins are found in endoplasmic reticulum (HSP90B1; MIM 191175) and mitochondria (TRAP1; MIM 606219) (Chen et al., 2005 [PubMed 16269234]).[supplied by OMIM].

Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Entrez Gene record to access additional publications.

FEATURES
 source Location/Qualifiers
 1..724
 /organism="Homo sapiens"

```

/db_xref="taxon:9606"
/chromosome="6"
/map="6p12"
Protein 1..724
/product="heat shock protein HSP 90-beta"
/note="heat shock 90kD protein 1, beta; heat shock protein
beta; heat shock 90kDa protein 1, beta; HSP84; HSP 84; HSP
90; heat shock 84 kDa"
/region 35..188
/region_name="HATPase_c"
/note="Histidine kinase-like ATPases; This family includes
several ATP-binding proteins for example: histidine
kinase, DNA gyrase B, topoisomerases, heat shock protein
HSP90, phytochrome-like ATPases and DNA mismatch repair
proteins; cl00075"
/region 191..>219
/region_name="HSP90"
/note="Hsp90 protein; pfam00183"
Site /db_xref="CDD:109449"
226
/site_type="phosphorylation"
/experiment="experimental evidence, no additional details
recorded"
Site /citation=[2]
255
/site_type="phosphorylation"
/experiment="experimental evidence, no additional details
recorded"
Site /citation=[2]
255
/site_type="phosphorylation"
/experiment="experimental evidence, no additional details
recorded"
Site /citation=[1]
261
/site_type="phosphorylation"
/experiment="experimental evidence, no additional details
recorded"
Region /citation=[11]
276..724
/region_name="HSP90"
/note="Hsp90 protein; pfam00183"
Site /db_xref="CDD:109449"
297
/site_type="phosphorylation"
/experiment="experimental evidence, no additional details
recorded"
Site /citation=[8]
484
/site_type="phosphorylation"
/experiment="experimental evidence, no additional details
recorded"
Site /citation=[10]
484
/site_type="phosphorylation"
/experiment="experimental evidence, no additional details
recorded"
Site /citation=[6]
1..724
/cls
/gene="HSP90AB1"
/gene_synonym="D6S182; FLJ26984; HSP90-BETA; HSP90B;
HSPC2; HSPCB"
/coded_by="NM_007355.2:85..2259"
/db_xref="CCDS:CCDS4909.1"
/db_xref="GeneID:3346"
/db_xref="HGNC:5255"
/db_xref="HPRD:06763"
/db_xref="MIM:140512"

```

Mascot: <http://www.matrixscience.com/>

4) Subunidad α de la isoforma II de la proteína de choque térmico (HSP 90- α II)

Mascot Search Results: Protein View

Page 1 of 3

Mascot Search Results

Protein View

Match to: [gi|154146191](#) Score: 357

heat shock protein HSP 90-alpha isoform 2 [Homo sapiens]

Found in search of mas52.tmp

Nominal mass (M_r): 84607; Calculated pI value: 4.94

NCBI BLAST search of [gi|154146191](#) against nr

Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Homo sapiens](#)

Links to retrieve other entries containing this sequence from NCBI Entrez:

[gi|92090606](#) from [Homo sapiens](#)

[gi|703087](#) from [Homo sapiens](#)

[gi|61656605](#) from [Homo sapiens](#)

[gi|111306539](#) from [Homo sapiens](#)

[gi|119602171](#) from [Homo sapiens](#)

[gi|119602172](#) from [Homo sapiens](#)

[gi|15825666](#) from [Homo sapiens](#)

[gi|158256646](#) from [Homo sapiens](#)

Variable modifications: Deamidated (NQ), Oxidation (M), Propionamide (C)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: 16%

Matched peptides shown in Bold Red

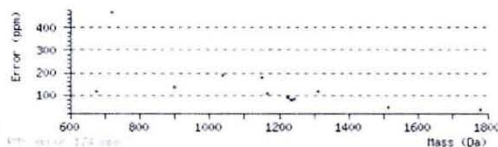
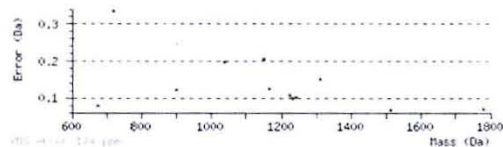
```
1 MPEETQTQDQ PMEEEEVETF AFQAEIAQLM SLIINTFYSN KEIFLRELIS
51 NSSDALDKIR YESLTDPSKL DSGKELHINL IPNKQDRTLT IVDTGIGMTK
101 ADLINNLGTI AKSGTKAFME ALQAGADISM IGQFGVGFYS AYLVAEKTV
151 ITKHNDDEQY AWESSAGGSF TVRTDTGEPM GRGTVKVLHL KEDQTEYLEE
201 RRIKEIVKHI SQFIGYPITL FVEKERDKEV SDDEAEKED KEEKEKEEEK
251 ESEDKPEIED VGSDEEEEEK DGDKKKKKKI KEKYIDQEEL NKTPIWTRN
301 PDDITNEEYG EFKYSLTNDW EDHLAVKHFS VEGQLEFRAL LFPVRRAPFD
351 LFNRRKKNNN IKLYVRRVFI MDNCEELIPE YLNFIRGVVD SEDLPLNISR
401 EMLQQSKILK VIRKNLVKKC LELFTELAED KENYKKFYEQ FSKNIKLGHI
451 EDSQNRKKLS ELLRYTTSAS GDEMVSLEKY CTRMKENQKH IYYITGETKD
501 QVANSAFVER LRKHGLEVIY MIEPIDEYCV QQLKEFEGKT LSVVTKEGLE
551 LPDEDEEKKK QEEKKTKFEN LCKIMKDILE KKVEKVVVSN RLVTSPCCIV
601 TSTYGTWANN ERIMKAQALR DNSTMGYMAA KKHLEINPDH SIETLRQKA
651 EADKNKSVK DLVILLYETA LLSSGFSLED PQTHANRIYR MIKLGIGIDE
701 DDPTADDTSA AVTEEMPPL E GDDDTSRMEE VD
```

Show predicted peptides also

Sort Peptides By

Residue Number Increasing Mass Decreasing Mass

Start	End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence
42	46	677.4770	676.4697	676.3908	0.0789	0	R.EIFLR.E (Ions score 21)
61	69	520.3491	1038.6836	1038.4869	0.1967	0	R.YESLTDPSK.L (Ions score 20)
101	112	621.9072	1241.7998	1241.6979	0.1019	0	R.ADLINNLGTIAK.S (Ions score 44)
186	191	361.9183	721.8221	721.4850	0.3371	0	K.VILHLK.E (Ions score 39)
192	201	656.3643	1310.7141	1310.5626	0.1515	0	K.EDQTEYLEER.R (Ions score 34)
210	224	890.0118	1778.0091	1777.9403	0.0688	0	K.HSQFIGYPITLFEVK.E (Ions score 13)
284	292	576.3857	1150.7568	1150.5506	0.2062	0	K.YIDQEELNK.T (Ions score 25)
293	299	451.3274	900.6401	900.5181	0.1220	0	R.TAPIWTR.N (Ions score 24)
387	400	757.4304	1512.8463	1512.7784	0.0679	0	R.GVVDSDELPLNISR.E (Ions score 63)
447	456	584.8515	1167.6885	1167.5632	0.1253	0	K.LGIHEDSQNR.K (Ions score 32)
490	499	612.8705	1223.7265	1223.6186	0.1079	0	R.HIYYITGETR.D (Ions score 25)
500	510	618.3536	1234.6927	1234.5942	0.0985	0	R.DQVANSAFVER.L (Ions score 32)



LOCUS NP_005339 732 aa linear PRI 29-AUG-2010
DEFINITION heat shock protein HSP 90-alpha isoform 2 [Homo sapiens].
ACCESSION NP_005339

http://www.matrixscience.com/cgi/protein_view.pl?file=../data/20100603/FteAifsEO.dat... 07/09/2010

VERSION NF_005339.3 GI:154146191
 DBSOURCE REFSEQ: accession [NM_005348.3](#)
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM [Homo sapiens](#)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
 Catarrhini; Hominidae; Homo.
 REFERENCE 1 (residues 1 to 732)
 AUTHORS Qin, Z., DeFee, M., Isaacs, J.S. and Parsons, C.
 TITLE Extracellular Hsp90 serves as a co-factor for MAPK activation and
 latent viral gene expression during de novo infection by KSHV
 JOURNAL Virology 403 (1), 92-102 (2010)
 PUBMED [20491233](#)
 REMARK GeneRIF: The authors show that extracellular Hsp90 inhibition
 suppresses human herpesvirus 8 gene expression during de novo
 infection, and that this effect is mediated largely through the
 inhibition of mitogen-activated protein kinase activation.
 REFERENCE 2 (residues 1 to 732)
 AUTHORS Cruickshanks, N., Shervington, L., Patel, R., Munje, C., Thakkar, D. and
 Shervington, A.
 TITLE Can hsp90alpha-targeted siRNA combined with TMZ be a future therapy
 for glioma?
 JOURNAL Cancer Invest. 28 (6), 608-614 (2010)
 PUBMED [20210521](#)
 REMARK GeneRIF: Hsp90alpha-targeted siRNA combined with temozolomide
 enhanced chemosensitivity of glioma cells.
 REFERENCE 3 (residues 1 to 732)
 AUTHORS Olson, J.E., Wang, X., Pankratz, V.S., Fredericksen, Z.S., Vachon, C.M.,
 Vierkant, R.A., Cerhan, J.R. and Couch, F.J.
 TITLE Centrosome-related genes, genetic variation, and risk of breast
 cancer
 JOURNAL Breast Cancer Res. Treat. (2010) In press
 PUBMED [20508983](#)
 REMARK GeneRIF: Observational study of gene-disease association. (HuGE
 Navigator)
 Publication Status: Available-Online prior to print
 REFERENCE 4 (residues 1 to 732)
 AUTHORS Johnston, M., Geoffroy, M.C., Sobala, A., Hay, R. and Hutvagner, G.
 TITLE HSP90 protein stabilizes unloaded argonaute complexes and
 microscopic P-bodies in human cells
 JOURNAL Mol. Biol. Cell 21 (9), 1462-1469 (2010)
 PUBMED [20237157](#)
 REMARK GeneRIF: Data suggests that HSP90 protein chaperones Argonautes
 before binding RNA and may facilitate efficient loading of small
 RNA.
 REFERENCE 5 (residues 1 to 732)
 AUTHORS McCready, J., Sims, J.D., Chan, D. and Jay, D.G.
 TITLE Secretion of extracellular hsp90alpha via exosomes increases cancer
 cell motility: a role for plasminogen activation
 JOURNAL BMC Cancer 10, 294 (2010)
 PUBMED [20553606](#)
 REMARK GeneRIF: Hsp90alpha is released by invasive cancer cells via
 exosomes and activates plasmin, a second protease that acts in
 cancer cell invasion.
 Publication Status: Online-Only
 REFERENCE 6 (residues 1 to 732)
 AUTHORS Shakhovich, R., Shue, G. and Kohtz, D.S.
 TITLE Conformational activation of a basic helix-loop-helix protein
 (MyoD1) by the C-terminal region of murine HSP90 (HSP84)
 JOURNAL Mol. Cell. Biol. 12 (11), 5059-5068 (1992)
 PUBMED [1406681](#)
 REFERENCE 7 (residues 1 to 732)
 AUTHORS Nemoto, T., Ohara-Nemoto, Y. and Ota, M.
 TITLE Association of the 90-kDa heat shock protein does not affect the
 ligand-binding ability of androgen receptor
 JOURNAL J. Steroid Biochem. Mol. Biol. 42 (8), 803-812 (1992)
 PUBMED [1525041](#)
 REFERENCE 8 (residues 1 to 732)
 AUTHORS Hutchison, K.A., Brott, B.K., De Leon, J.H., Perdew, G.H., Jove, R. and
 Pratt, W.B.
 TITLE Reconstitution of the multiprotein complex of pp60src, hsp90, and
 p50 in a cell-free system
 JOURNAL J. Biol. Chem. 267 (5), 2902-2908 (1992)
 PUBMED [1310678](#)
 REFERENCE 9 (residues 1 to 732)
 AUTHORS Ozawa, K., Murakami, Y., Eki, T., Soeda, E. and Yokoyama, K.
 TITLE Mapping of the gene family for human heat-shock protein 90 alpha to
 chromosomes 1, 4, 11, and 14
 JOURNAL Genomics 12 (2), 214-220 (1992)
 PUBMED [1740332](#)

REFERENCE 10 (residues 1 to 732)
 AUTHORS Yamazaki,M., Tashiro,H., Yokoyama,K. and Soeda,E.
 TITLE Molecular cloning of cDNA encoding a human heat-shock protein whose expression is induced by adenovirus type 12 E1A in HeLa cells
 JOURNAL Agric. Biol. Chem. 54 (12), 3163-3170 (1990)
 PUBMED [1368637](#)
 COMMENT VALIDATED [REFSEQ](#): This record has undergone validation or preliminary review. The reference sequence was derived from [DC303876.1](#), [AJ890083.1](#) and [AI250920.1](#).
 On Jul 28, 2007 this sequence version replaced gi:[40254816](#).
 Summary: HSP90 proteins are highly conserved molecular chaperones that have key roles in signal transduction, protein folding, protein degradation, and morphologic evolution. HSP90 proteins normally associate with other cochaperones and play important roles in folding newly synthesized proteins or stabilizing and refolding denatured proteins after stress. There are 2 major cytosolic HSP90 proteins, HSP90AA1, an inducible form, and HSP90AB1 (MIM 140572), a constitutive form. Other HSP90 proteins are found in endoplasmic reticulum (HSP90B1; MIM 191175) and mitochondria (TRAP1; MIM 606219) (Chen et al., 2005 [PubMed 16269234]).[supplied by OMIM].
 Transcript Variant: This variant (2) differs in the 5' UTR and coding sequence compared to variant 1. The resulting isoform (2) is shorter at the N-terminus compared to isoform 1.
 Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Entrez Gene record to access additional publications.

FEATURES
 source
 1..732
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="14"
 /map="14q32.33"
 Protein
 1..732
 /product="heat shock protein HSP 90-alpha isoform 2"
 /note="heat shock 90kD protein 1, alpha; heat shock 90kDa protein 1, alpha; heat shock 90kD protein 1, alpha-like 4; heat shock 90kD protein, alpha-like 4; HSP 86; heat shock 86 kDa; renal carcinoma antigen NY-REN-38; heat shock protein HSP 90-alpha"
 /calculated_mol_wt=84529
 Region
 40..193
 /region_name="HATPase_c"
 /note="Histidine kinase-like ATPases; This family includes several ATP-binding proteins for example: histidine kinase, DNA gyrase B, topoisomerases, heat shock protein HSP90, phytochrome-like ATPases and DNA mismatch repair proteins; cl00075"
 /db_xref="CDD:173995"
 Region
 196..>226
 /region_name="HSP90"
 /note="Hsp90 protein; pfam00183"
 /db_xref="CDD:109249"
 Region
 290..732
 /region_name="HSP90"
 /note="Hsp90 protein; pfam00183"
 /db_xref="CDD:109249"
 CDS
 1..732
 /gene="HSP90AA1"
 /gene_synonym="FLJ31884; HSP86; Hsp89; HSP89A; Hsp90; HSP90A; HSP90N; HSPC1; HSPCA; HSPCAL1; HSPCAL4; HSPN; LAP2"
 /coded_by="NM_005348.3:191..2389"
 /note="isoform 2 is encoded by transcript variant 2"
 /db_xref="CCDS:CCDS9967.1"
 /db_xref="GeneID:3320"
 /db_xref="HGNC:5253"
 /db_xref="HPRD:00777"
 /db_xref="MIM:140571"

Mascot: <http://www.matrixscience.com/>

ANEXO IV.

Resultados obtenidos del análisis con las herramientas NetPhos y NetOGly (Expasy.org) de las fosfoproteínas y glicoproteínas seleccionada como posibles biomarcadores de Cáncer de mama. NetPhos.

1) Nucleolina.

NetPhos 2.0 Server - prediction results

Page 1 of 3



NetPhos 2.0 Server - prediction results

Technical University of Denmark

```
707 Sequence
MVKLAKAGKNQGDPKMAPPPKEVEEDSEDEEMSEDEEDSSGEEVVIPOKKGKKAATSAKKVVVSPTKKVAVATPAKK 80
AAVTPGKKAAATPAKKTVTPTAKAVTTPGKKGATPGKALVATPGKKGAAPAKGAKNGKNAKKEDSDEEDDDSEDEEDD 160
EDEDEDEDEIEPAAMKAAAAAPASEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 240
DDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 320
PELKTGISDVFANDLAVDVRIGMTRKFGYVDFESAEDLEKALELTGLKVFNEIKLEKPKGKDSKKERDARTLLAKNL 400
PYKVTQDELKEVFEDAAEIRLVSKDGKSGIAYIEFKTEADAETFEKQGTEDGRSISLYYTGEKQNDYRGGKNST 480
WSGESKTLVLNLSYSATEETLQEVFEKATFIKVPQNGKSKGYAFIEFASFEDAKEALNSCNKREIEGRAIRLELQGP 560
RGSPNARSQPSKTLFVKGLSEDTTEETLKESFDGVSVRARIIVTDRETGSSKGFVDFNSEEDAKEAMEDGEIDGNKVTLD 640
WAKPKGEGGFGGRGGGFGGRGGGFGGRGGGFGGRGGGFGGRGGGFGGRGGGFGGRGGGFGGRGGGFGGRGGGFGGRGGG 720
.....S.....S.....SS.....S.....S.T.....T.....
..T.....T.....T.T.....T.....T.....S.....S.....
.....S.....T.....
.....S.....
.....Y.....S.....S.....
..Y..T.....S.....S.....T.....T.....S.....S.....
..S.....SYS.....T.....Y.....S.....
..S.....S..T..T..S..S.....T.....SS.....S.....
.....T.....
Phosphorylation sites predicted:Ser: 28 Thr: 17 Tyr: 4
```

Serine predictions

Name	Pos	Context	Score	Pred
Sequence	28	VEEDSEDEE	0.983	*S*
Sequence	34	DEEDSEDEE	0.994	*S*
Sequence	41	EEDSSGEE	0.996	*S*
Sequence	42	EDDSGEEV	0.996	*S*
Sequence	60	AAATSARKV	0.914	*S*
Sequence	67	KVVVSPTKK	0.975	*S*
Sequence	145	KKEDSDEE	0.998	*S*
Sequence	153	EDDSSEDE	0.998	*S*
Sequence	184	AAPASEDED	0.988	*S*
Sequence	206	EEDSDEEA	0.997	*S*
Sequence	319	NFNKSAPEL	0.821	*S*
Sequence	328	KTGISDVFA	0.101	.
Sequence	356	VDFESAEDL	0.994	*S*
Sequence	386	KGKDSKKER	0.998	*S*
Sequence	423	IRLVSKDGK	0.775	*S*
Sequence	428	KDGKSGIA	0.838	*S*
Sequence	458	IDGRSISLY	0.054	.
Sequence	460	GRSISLYYT	0.866	*S*
Sequence	479	GGKNSTWSG	0.665	*S*
Sequence	482	NSTWSGESK	0.965	*S*
Sequence	485	WSGESKTLV	0.042	.
Sequence	491	TLVLNLSY	0.019	.
Sequence	494	LSNLSYSAT	0.931	*S*
Sequence	496	NLSYSATEE	0.980	*S*
Sequence	522	QNGKSKGYA	0.254	.

<http://www.cbs.dtu.dk/cgi-bin/nph-webface?jobid=netphos,4C844637023321BB&opt=no...> 06/09/2010

Sequence	532	IEFASFEDA	0.993	*S*
Sequence	542	EALNSCNKR	0.022	.
Sequence	563	GPRGSPNAR	0.556	*S*
Sequence	568	PNARSQPSK	0.119	.
Sequence	571	RSQPSKTLF	0.370	.
Sequence	580	VKGLSEDTT	0.947	*S*
Sequence	591	TLKESFDGS	0.646	*S*
Sequence	595	SFDGSVRAR	0.595	*S*
Sequence	608	RETGSSKGF	0.901	*S*
Sequence	609	ETGSSKGF	0.950	*S*
Sequence	619	VDFNSEEDA	0.996	*S*

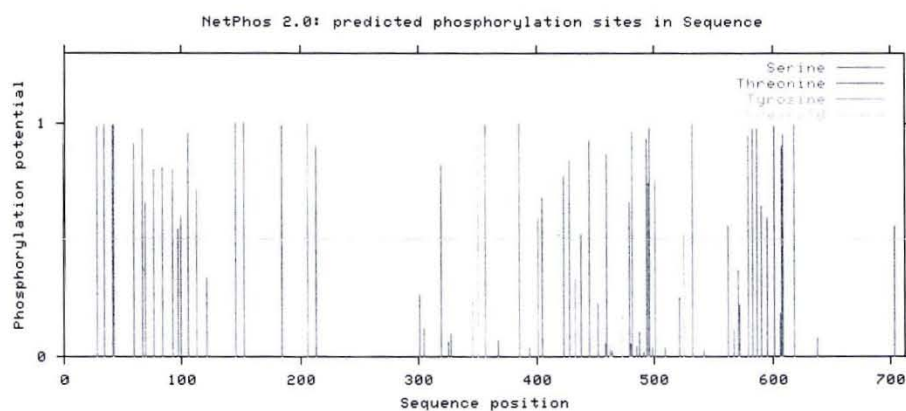
Threonine predictions

Name	Pos	Context	Score	Pred
v				
Sequence	59	KAAATSACK	0.431	.
Sequence	69	VVSPTKKVA	0.660	*T*
Sequence	76	VAVATPAKK	0.801	*T*
Sequence	84	KAAVTPGKK	0.810	*T*
Sequence	92	KAAATPAKK	0.801	*T*
Sequence	97	PAKKTVTPA	0.547	*T*
Sequence	99	KKTVTAKA	0.603	*T*
Sequence	105	AKAVTPGK	0.045	.
Sequence	106	KAVTPGKK	0.959	*T*
Sequence	113	KKGATPGKA	0.715	*T*
Sequence	121	ALVATPGKK	0.336	.
Sequence	213	EAMETTPAK	0.161	.
Sequence	214	AMETTPAKG	0.900	*T*
Sequence	301	KVEGTEPTT	0.262	.
Sequence	304	GTEPTTAFN	0.038	.
Sequence	305	TEPTTAFNL	0.121	.
Sequence	325	PELKTGISD	0.063	.
Sequence	346	RIGMTRKFG	0.251	.
Sequence	367	ALELTGLKV	0.066	.
Sequence	394	RDARTLLAK	0.034	.
Sequence	405	PYKVTQDEL	0.683	*T*
Sequence	438	IEFKTEADA	0.522	*T*
Sequence	445	DAEKTFEK	0.929	*T*
Sequence	452	EKQGTEDG	0.227	.
Sequence	464	SLYYTGKKG	0.027	.
Sequence	480	GKNSTWSGE	0.056	.
Sequence	487	GESKTLVLS	0.102	.
Sequence	498	SYSATEETL	0.034	.
Sequence	501	ATEETLQEV	0.757	*T*
Sequence	510	FEKATFIKV	0.034	.
Sequence	573	QPSKTLFVK	0.223	.
Sequence	583	LSEDTTEET	0.975	*T*
Sequence	584	SEDTEETL	0.235	.
Sequence	587	TTEETLKES	0.978	*T*
Sequence	602	ARIVTDRET	0.987	*T*
Sequence	606	TDRETGSSK	0.183	.
Sequence	638	GNKVTLDWA	0.081	.
Sequence	704	QGKKTKE-	0.560	*T*

Tyrosine predictions

Name	Pos	Context	Score	Pred
v				
Sequence	351	RKFGYVDFE	0.977	*Y*
Sequence	402	KNLPYKVTQ	0.586	*Y*

Sequence	433	KGIAYIEFK	0.331	.
Sequence	462	SISLYYTGE	0.053	.
Sequence	463	ISLYYTGEK	0.029	.
Sequence	473	QNQDYRGK	0.177	.
Sequence	495	SNLSYSATE	0.739	*Y*
Sequence	525	KSKGYAFIE	0.522	*Y*



[Explain the output.](#) [Go back.](#)

2) Ribonucleoproteína U (HnRNP).

NetPhos 2.0 Server - prediction results

Page 1 of 3



NetPhos 2.0 Server - prediction results

Technical University of Denmark

806 Sequence
MSSSPVNVKKLVSELKEELKKRRSLDKGLKAEELMERLQAALDDEEAGGRPAMEPGNGSLDLGGDSAGRSAGAGLEQEAAA 80
GGDEEEEEEEEEEGISALDGDQMEELGEENGAAGAADSGPMEEEAASEDENGDDQGFQEGEDELGDDEEGAGDENGHGE 160
QQQPATQQQQPQQRGAAKEAAGKSSGPTSLFAVTVAPPGARQQQQAGGDGKTEQKGGDKKRGVKKRPREDHGRGYFE 240
YIEENKYSRAKSPQPPVEEEDHFDFTVCLDTYNCDLHFKISRDRLSASSLTMSFAFLWAGGRASYGVSKGKVCFEMK 320
VTEKIPVRHLYTKDIDIHEVRIGWSLTSGMLLGEEFSGYSLGKIKTCNCETEDYGEKFDENDVITCFANFESDEVEL 400
SYAKNGQDLGVAFKISKEVLAGRPLFPHVLCHNCAVEFNFGQKEKPYFPIPEEYTFIQNVPLEDRVRGPKGPEEKDCEV 480
VMMIGLPGAGKTTWVTKHAAENPGKYNILGTNTIMDKMMVAGFKQMDTGKLTLLQRAPOCLGKFIEIAARKKRNFIL 560
DQTNVSAQAQRKMLCLFAGFQRKAVVCPKDEDYKQRTQKKAIEVEGKDLPEHAVLKMKGNTLPEVAECFDEITYVELQK 640
EEAQKLEQYKEESKKALPPEKKQNTGSKKSNKNSGKNQFNRRGGHRRGGLNMRGNGFRGGAPGNRRGGYNRRGNMPQR 720
GGGGGSGGIGYPYPRAPVFPGRGSYSNRGNYNRGGMPNRGNYNQNRGRGNRRGYKNQSQGYNQWQQQFWGQKFWQSH 800
YHQGY 880
.....S.....S.....S...S..... 80
.....S.....S.....S.....S..... 160
.....S.....S.....S.....S..... 240
.....Y.....S.....T.....S...S.....S..... 320
.....SY.....Y.....Y..... 400
.....Y.....Y.....Y.....Y..... 480
.....T.....Y.....T.....T.....T..... 560
.....S.....Y.....T.....Y..... 640
.....S.....S...S...S..... 720
.....SYS.....Y.....S... 800
..... 880

Phosphorylation sites predicted: Ser: 21 Thr: 5 Tyr: 11

Serine predictions

Name	Pos	Context	Score	Pred
		v		
Sequence	2	---MSSSPV	0.002	.
Sequence	3	--MSSSPVN	0.003	.
Sequence	4	-MSSSPVNV	0.034	.
Sequence	14	KLKVELKE	0.891	*S*
Sequence	26	KRRSLDKGL	0.998	*S*
Sequence	59	PGNGSLDLG	0.009	.
Sequence	66	LGGDSAGRS	0.827	*S*
Sequence	70	SAGRSAGAGL	0.660	*S*
Sequence	97	EEGISALDG	0.984	*S*
Sequence	118	GAADSGPME	0.501	*S*
Sequence	128	EEAASEDEN	0.990	*S*
Sequence	187	AAGKSSGPT	0.581	*S*
Sequence	188	AGKSSGPTS	0.087	.
Sequence	192	SGPTSLFAV	0.194	.
Sequence	248	ENKYSRAKS	0.158	.
Sequence	252	SRAKSPQPP	0.995	*S*
Sequence	283	HFKISRDRL	0.172	.
Sequence	288	RDRLSASSL	0.991	*S*
Sequence	290	RLSASSLTM	0.064	.
Sequence	291	LSASSLTM	0.947	*S*
Sequence	296	LTMESFAFL	0.004	.

<http://www.cbs.dtu.dk/cgi-bin/nph-webface?jobid=netphos,4C84468A03FC3FE3&opt=n...> 07/09/2010

Sequence	307	GGRASYGVVS	0.976	*S*
Sequence	311	SYGVSKGKV	0.017	.
Sequence	345	RIGWSLTTS	0.006	.
Sequence	349	SLTTSGMML	0.004	.
Sequence	359	EEEFSSYGYS	0.558	*S*
Sequence	363	SYGYSLKGI	0.425	.
Sequence	395	ANFESDEVE	0.084	.
Sequence	401	EVELSYAKN	0.238	.
Sequence	416	AFKISKEVL	0.189	.
Sequence	566	QTNVSAAAQ	0.551	*S*
Sequence	654	YKEESKKAL	0.985	*S*
Sequence	668	QNTGSKKSN	0.987	*S*
Sequence	671	GSKKSNKNK	0.992	*S*
Sequence	676	NKNKSGKNQ	0.997	*S*
Sequence	727	GGGGSGGIG	0.022	.
Sequence	745	PGRGSYSNR	0.945	*S*
Sequence	747	RGSYSNRGN	0.696	*S*
Sequence	780	YKNQSQGYN	0.196	.
Sequence	798	QKPWSQHYH	0.568	*S*

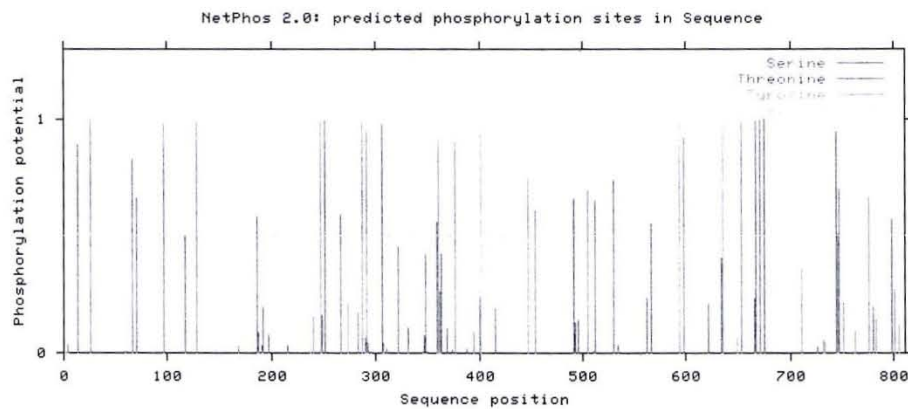
Threonine predictions

Name	Pos	Context	Score	Pred
v				
Sequence	168	QPPATQQQQ	0.031	.
Sequence	191	SSGPTSLFA	0.028	.
Sequence	197	LFAVTVAPP	0.080	.
Sequence	216	GDGKTEQKG	0.032	.
Sequence	267	HFDDEVVCL	0.587	*T*
Sequence	273	VCLDTYNCD	0.031	.
Sequence	293	ASSLTMESF	0.044	.
Sequence	322	EMKVTEKIP	0.452	.
Sequence	332	RHLYTKDID	0.111	.
Sequence	347	GWSLTTSGM	0.076	.
Sequence	348	WSLTTSGML	0.426	.
Sequence	369	KGIKTCNCE	0.105	.
Sequence	374	CNCETEDYG	0.015	.
Sequence	388	NDVITCFAN	0.018	.
Sequence	455	PEEYTFIQN	0.056	.
Sequence	492	GAGKTTWVT	0.658	*T*
Sequence	493	AGKTTWVTK	0.131	.
Sequence	496	TTWVTKHAA	0.139	.
Sequence	511	NILGTNTIM	0.005	.
Sequence	513	LGTNTIMDK	0.648	*T*
Sequence	530	QMADTGKLN	0.738	*T*
Sequence	535	GKLNTLLQR	0.029	.
Sequence	563	ILDQTNVSA	0.230	.
Sequence	598	YKQRTQKKA	0.920	*T*
Sequence	622	KGNFTLPEV	0.206	.
Sequence	634	FDEITYVEL	0.403	.
Sequence	666	KKQNTGSKK	0.231	.

Tyrosine predictions

Name	Pos	Context	Score	Pred
v				
Sequence	238	HGRGYFEYI	0.018	.
Sequence	241	GYFEYIEEN	0.152	.
Sequence	247	EENKYSRAK	0.981	*Y*
Sequence	274	CLDTYNCDL	0.208	.
Sequence	308	GRASYGVSK	0.044	.

Sequence	331	VRHLYTKDI	0.014	.
Sequence	360	EEFSYGYSL	0.907	*Y*
Sequence	362	FSYGYSLKG	0.265	.
Sequence	377	ETEDYGEKF	0.903	*Y*
Sequence	402	VELSYAKNG	0.947	*Y*
Sequence	447	KEKPYFFIP	0.739	*Y*
Sequence	454	IPEEYTFIQ	0.610	*Y*
Sequence	506	NPGKYNILG	0.695	*Y*
Sequence	594	KDEDYKQRT	0.977	*Y*
Sequence	635	DEITYVELQ	0.967	*Y*
Sequence	650	LLEQYKEES	0.062	.
Sequence	711	NRGGYNRRG	0.353	.
Sequence	732	GGIGYPYPR	0.053	.
Sequence	734	IGYPYPRAP	0.041	.
Sequence	746	GRGSYSNRG	0.512	*Y*
Sequence	752	NRGNYNRRG	0.212	.
Sequence	763	NRGNYNQNF	0.090	.
Sequence	776	NNRGYKNQS	0.663	*Y*
Sequence	783	QSQGYNQWQ	0.144	.
Sequence	801	WSQHYHQGY	0.271	.
Sequence	805	YHQGY---	0.081	.
Sequence	806	HQGY---	0.115	.



[Explain the output.](#) [Go back.](#)

3) Subunidad β de la proteína de choque térmico (HSP 90- β).

NetPhos 2.0 Server - prediction results

Page 1 of 3



NetPhos 2.0 Server - prediction results

Technical University of Denmark

724 Sequence

```

MPEEVHHGEEVETFAFQAEIAQLMSLIINTFYNSKEIFLRELISNASDALDKIRYESLTDPSKLD SGKELKIDIIPNPQ      80
ERTLTLDVTGIGMTKADLINNLGTIAKSGTKAFMEALQAGADISMIGQFGVGFYSAYLVAEKVVVITKHNDEQYAWESS      160
AGGSFTVRADHGEPGRGTVILHLKEDQTEYLEERRVKEVVKHQSQFIGYPITLYLEKEREKEISDDEAEKEKEKEE      240
DKDDEEKPKIEDVGSDEEDSGDKKKKKTKIKEKYIDQELNKTPIWTRNPDITQEEYGEFYKSLTNDWEDHLAVKH      320
FSVEGQLEFRALLFIPRRAPFDLFENKKKKNNIKLYVRRVFIMDSCDELIPEYLNFIIRGVVDSDELPLNISREMLQOSKI      400
LKVIRKNIIVKKCLELFSLEAEDKENYKKFYEAFSKNLGLIHEDSTNRRRLSELLRYHTSQSGDEMTSLSEYVSRMKETQ      480
KSIYYITGESKEQVANSFAVERVRKRGFEVVMTEPIDEYCVQQLKEFDGKSLVSVTKGLELPEDEEEKKKMEESKAKF      560
ENLCKLMKEILDKKVEKVTISNRLVSSPCCIVTSTYGTANMERIMKAQALRDNSTMGYMAKKHLEINPDHPIVETLRQ      640
KAADKNDKAVKDLVLLFETALLSSGFSLEDPQTHSNRIYRMIKLGLGIDEDEVAAEFPNAAVPDEIPPLEGDEGDASRM      720
EEVD                                                                                      800
.....YS.....Y.S...S...S.....
...T.....S.....Y.....Y.....
...T.....T.....T.Y.....S.....
.....S.....S.....T.....Y.....Y.....
.S.....ST.....S.....S.....S.Y.....T.....
.S.Y.....Y.....S.....
.....SS.....S.....T.....
.....S.....S.....
....

```

Phosphorylation sites predicted: Ser: 21 Thr: 8 Tyr: 11

Serine predictions

Name	Pos	Context	Score	Pred
Sequence	26	AQLMSLIIN	0.016	.
Sequence	34	NTFYNSKEI	0.982	*S*
Sequence	45	RELISNASD	0.006	.
Sequence	48	ISNASDALD	0.016	.
Sequence	58	IRYESLTD	0.997	*S*
Sequence	63	LTDPSKLDS	0.976	*S*
Sequence	67	SKLDSGKEL	0.998	*S*
Sequence	108	TIKSGTKA	0.916	*S*
Sequence	124	GADISMIGQ	0.328	.
Sequence	135	VGFY SAYLV	0.004	.
Sequence	159	YAWESSAGG	0.011	.
Sequence	160	AWESSAGGS	0.057	.
Sequence	164	SAGGSFTVR	0.024	.
Sequence	206	VKKHSQFIG	0.255	.
Sequence	226	EKEISDDEA	0.996	*S*
Sequence	255	EDVGSDEED	0.997	*S*
Sequence	261	EEDDSGDK	0.998	*S*
Sequence	307	EFYKSLTND	0.265	.
Sequence	322	VKHFSVEGQ	0.972	*S*
Sequence	365	FIMDSCDEL	0.033	.
Sequence	383	GVVDSDEL	0.172	.
Sequence	391	PLNISREML	0.066	.
Sequence	398	MLQQSKILK	0.018	.

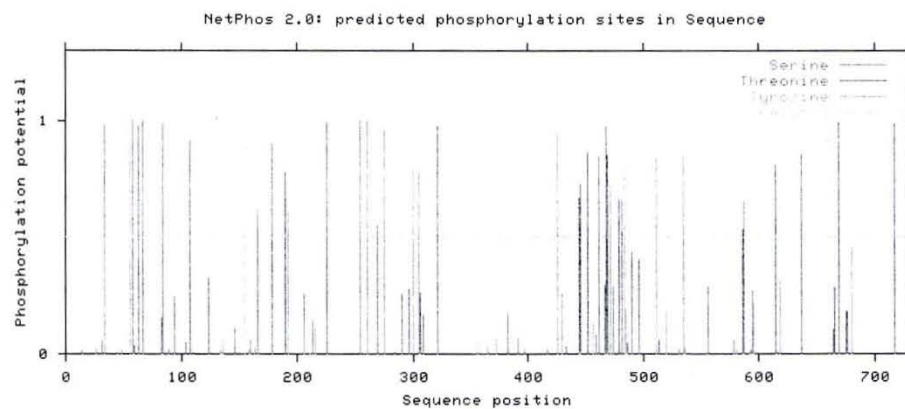
<http://www.cbs.dtu.dk/cgi-bin/nph-webface?jobid=netphos,4C8446CC03FD03D2&opt=n...> 07/09/2010

Sequence	417	LELFSELAE	0.021	.
Sequence	434	YEAFSKNLK	0.033	.
Sequence	445	IHEDSTNRR	0.666	*S*
Sequence	452	RRRLSELLR	0.863	*S*
Sequence	460	RYHTSQSGD	0.039	.
Sequence	462	HTSQSGDEM	0.844	*S*
Sequence	468	DEMTSLSEY	0.976	*S*
Sequence	470	MTSLSEYVS	0.852	*S*
Sequence	474	SEYVSRMKE	0.291	.
Sequence	482	ETQKSIYYI	0.664	*S*
Sequence	490	ITGESKEQV	0.433	.
Sequence	497	QVANSAFVE	0.407	.
Sequence	532	FDGKSLVSV	0.018	.
Sequence	535	KSLVSVTKE	0.841	*S*
Sequence	556	KMEESKAKF	0.291	.
Sequence	581	KVTISNRLV	0.019	.
Sequence	586	NRLVSSPCC	0.535	*S*
Sequence	587	RLVSSPCCI	0.656	*S*
Sequence	594	CIVTSTYGW	0.016	.
Sequence	615	LRDNSTMGY	0.812	*S*
Sequence	665	TALLSSGFS	0.106	.
Sequence	666	ALLSSGFSL	0.280	.
Sequence	669	SSGFSLEDP	0.995	*S*
Sequence	677	PQTHSNRIY	0.184	.
Sequence	718	DEDASRMEE	0.990	*S*

Threonine predictions

Name	Pos	Context	Score	Pred
v				
Sequence	14	EEVETFAFQ	0.011	.
Sequence	31	LIINTFYSN	0.056	.
Sequence	60	YESLTDPSK	0.038	.
Sequence	83	PQERTLTLV	0.155	.
Sequence	85	ERTLTLVDT	0.985	*T*
Sequence	89	TLVDTGIGM	0.021	.
Sequence	94	GIGMTKADL	0.244	.
Sequence	104	NNLGTIAKS	0.052	.
Sequence	110	AKSGTKAFM	0.008	.
Sequence	147	VVITKHND	0.108	.
Sequence	166	GSFTVRAD	0.621	*T*
Sequence	179	IGRGTKVIL	0.902	*T*
Sequence	190	KEDQTEYLE	0.776	*T*
Sequence	214	GPITLYLE	0.144	.
Sequence	269	KKKKTKKIK	0.550	*T*
Sequence	285	ELNKTKPIW	0.008	.
Sequence	290	KPIWTRNPD	0.258	.
Sequence	297	PDDITQEEY	0.276	.
Sequence	309	YKSLTNDWE	0.166	.
Sequence	446	HEDSTNRRR	0.738	*T*
Sequence	459	LRHTSQSG	0.078	.
Sequence	467	GDEMTSLSE	0.297	.
Sequence	479	RMKETQKSI	0.670	*T*
Sequence	487	IYYITGESK	0.051	.
Sequence	514	VVYMTPEID	0.060	.
Sequence	537	LVSVTKEGL	0.029	.
Sequence	579	VEKVTISNR	0.053	.
Sequence	593	CCIVTSTYG	0.010	.
Sequence	595	IVTSTYGWT	0.271	.
Sequence	599	TYGWTANME	0.015	.
Sequence	616	RDNSTMGYM	0.005	.
Sequence	637	PIVETLRQK	0.861	*T*

Tyrosine predictions				
Name	Pos	Context	Score	Pred
Sequence	33	INTFYSNKE	0.501	*Y*
Sequence	56	DKIRYESLT	0.903	*Y*
Sequence	134	GVGFYSAYL	0.012	.
Sequence	137	FYSAYLVAE	0.068	.
Sequence	155	DDEQYAWES	0.933	*Y*
Sequence	192	DQTEYLEER	0.615	*Y*
Sequence	211	QFIGYPITL	0.040	.
Sequence	216	PITLYLEKE	0.112	.
Sequence	276	IKEKYIDQE	0.957	*Y*
Sequence	301	TQEEYGEFY	0.793	*Y*
Sequence	305	YGEFYKSLT	0.772	*Y*
Sequence	356	NIKLYVRRV	0.040	.
Sequence	373	LIPEYLNFI	0.062	.
Sequence	426	DKENYKKFY	0.945	*Y*
Sequence	430	YKKFYEAFS	0.258	.
Sequence	457	ELLRYHTSQ	0.135	.
Sequence	472	SLSEYVSRM	0.724	*Y*
Sequence	484	QKSIYYITG	0.790	*Y*
Sequence	485	KSIYYITGE	0.188	.
Sequence	512	FEVVMTEP	0.840	*Y*
Sequence	520	PIDEYCVQQ	0.177	.
Sequence	596	VTSTYGWTA	0.215	.
Sequence	619	STMGYMMAK	0.315	.
Sequence	681	SNRIYRMIK	0.461	.



[Explain the output.](#) [Go back.](#)

<http://www.cbs.dtu.dk/cgi-bin/nph-webface?jobid=netphos,4C8446CC03FD03D2&opt=n...> 07/09/2010

4) Subunidad β de la isoforma II de la proteína de choque térmico (HSP 90- α II).

NetPhos 2.0 Server - prediction results

Page 1 of 4



NetPhos 2.0 Server - prediction results

Technical University of Denmark

781 Sequence
MPEETQTQDQPMEEVEVTFQAEIAQLMSLIINTFYNSKEIFLRELISNSSDALDKIRYESLTDPSKLDGSGKELHINL 80
IPNKQDRILTIIVDTGIGMTKADLINNLGTIAKSGTKAFMEALQAGADISMIGQFGVGFYSAYLVAEKVTITKHNDDEQY 160
AWESSAGGSFTVRTDTGEPMGRGTVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEKEKED 240
KEEKEKEEKESEDKPEIEDVGSDEEEKKDGDGKKKKKKIKEYIDQEELNKTPIWTRNPDDITNEEYGEFYKSLTNDW 320
EDHLAVKHFSVEGQLEFRALLFVPRRAPDLFENRKKKNNIKLYVRRVFIMDNCEELIPEYLNFRGVVDSDDLPLNISR 400
EMLQSQSKILKIRKNLVKKCLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYTTSASGDEMVSLEKDY 480
CTRMKENQKHIYYITGETKDQVANSFVERLRKHGLEVIYMIPIDEYCVQQLKEFEGKTLVSVTKEGLELPEDEEEKKK 560
QEEKTKFENLCKIMKDILEKKVEKVVSNRLVTSPPCIVTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDH 640
SIETLRQKAEADKNDKSVKDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTADDTSAAVTEEMPPLE 720
GDDTSTRMEEVDGIREFNPHEATSHXCKPRXTEINHSPALPHAISXFXRMHMXMSAPIENS 800
.....T.T.....Y.S.....S..... 80
.....T.....S.....Y 160
.....T.....T.....T.....Y.....S.....S..... 240
.....S.....S.....Y.....Y.....Y..... 320
.....S..... 400
.....Y.....S.....S.....S..... 480
.....Y.....T.....S.....Y..... 560
.....T.....S.....S.....S..... 640
S.....T.....S.....S.....S..... 720
.....TS.....S.....S..... 800

Phosphorylation sites predicted: Ser: 23 Thr: 10 Tyr: 10

Serine predictions

Name	Pos	Context	Score	Pred
v				
Sequence	31	AQLMSLIIN	0.016	.
Sequence	39	NTFYNSKEI	0.982	*S*
Sequence	50	RELISNSSD	0.029	.
Sequence	52	LISNSSDAL	0.157	.
Sequence	53	ISNSSDALD	0.025	.
Sequence	63	IRYESLTDP	0.997	*S*
Sequence	68	LTDPKLD	0.976	*S*
Sequence	72	SKLDGSGKEL	0.997	*S*
Sequence	113	TIKSGTKA	0.916	*S*
Sequence	129	GADISMIGQ	0.328	.
Sequence	140	VGIFYSAYLV	0.004	.
Sequence	164	YAWESSAGG	0.011	.
Sequence	165	AWESSAGGS	0.057	.
Sequence	169	SAGGSFTVR	0.026	.
Sequence	211	VKKHSQFIG	0.837	*S*
Sequence	231	DKEVSDDEA	0.997	*S*
Sequence	252	EEKESEDKP	0.987	*S*
Sequence	263	EDVGSDEEE	0.997	*S*
Sequence	315	EFYKSLTND	0.265	.
Sequence	330	VKHFSVEGQ	0.972	*S*
Sequence	391	GVVDSDEL	0.172	.
Sequence	399	PLNISREML	0.066	.
Sequence	406	MLQSQSKILK	0.018	.

<http://www.cbs.dtu.dk/cgi-bin/nph-webface?jobid=netphos,4C84471703EB1B4D&opt=n...> 07/09/2010

Sequence	442	YEQFSKNIK	0.020	.
Sequence	453	IHEDSQNRK	0.701	*S*
Sequence	460	RKKLSELLR	0.015	.
Sequence	468	RYYSASGD	0.059	.
Sequence	470	YTSASGDEM	0.954	*S*
Sequence	476	DEMVSLKDY	0.996	*S*
Sequence	505	QVANSAFVE	0.574	*S*
Sequence	543	KTLVSVTKE	0.463	.
Sequence	589	KVVVSNRLV	0.036	.
Sequence	595	RLVTSPCCI	0.707	*S*
Sequence	602	CIVTSTYGW	0.016	.
Sequence	623	LRDNSTMGY	0.812	*S*
Sequence	641	NPDHSIIET	0.997	*S*
Sequence	658	KNDKSVKDL	0.998	*S*
Sequence	673	TALLSSGFS	0.106	.
Sequence	674	ALLSSGFSL	0.280	.
Sequence	677	SSGFSLBDP	0.995	*S*
Sequence	709	ADDTSAAVT	0.864	*S*
Sequence	726	DDDTSRMEE	0.993	*S*
Sequence	744	HEATSHXCK	0.918	*S*
Sequence	757	EINHSPALP	0.718	*S*
Sequence	765	PHAISXFXR	0.007	.
Sequence	775	HXXMSAPIE	0.004	.
Sequence	781	PIENS----	0.006	.

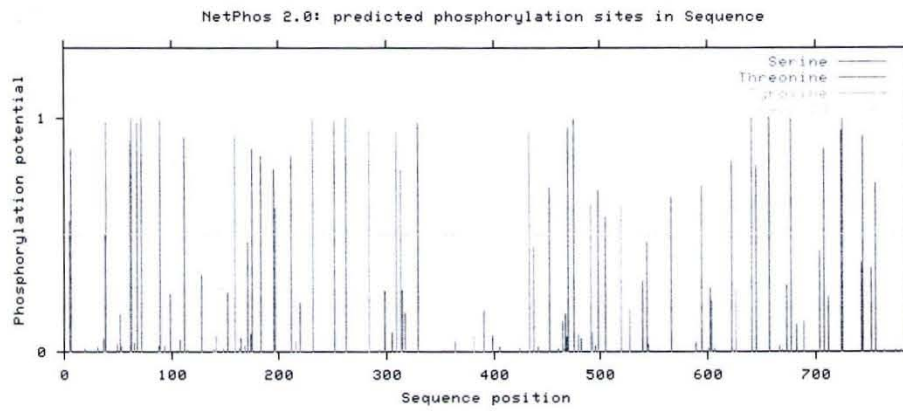
Threonine predictions

Name	Pos	Context	Score	Pred
v				
Sequence	5	MPEETQTQD	0.561	*T*
Sequence	7	EETQTQDQP	0.866	*T*
Sequence	19	EEVETFAFQ	0.011	.
Sequence	36	LIINTFYSN	0.056	.
Sequence	65	YESLTDPSK	0.038	.
Sequence	88	KQDRTLTIV	0.022	.
Sequence	90	DRTLTIIVDT	0.991	*T*
Sequence	94	TIVDTGIGM	0.026	.
Sequence	99	GIGMTKADL	0.244	.
Sequence	109	NNLGTIAKS	0.052	.
Sequence	115	AKSGTKAFM	0.008	.
Sequence	149	AEKVTVITK	0.011	.
Sequence	152	VTVITKHND	0.251	.
Sequence	171	GGSFVTRTD	0.467	.
Sequence	174	FTVRTDTGE	0.074	.
Sequence	176	VRTDTGEPM	0.867	*T*
Sequence	184	MGRGTKVIL	0.831	*T*
Sequence	195	KEDQTEYLE	0.776	*T*
Sequence	219	GYPITLFVE	0.210	.
Sequence	293	ELNKTKEPW	0.008	.
Sequence	298	KPIWTRNPD	0.258	.
Sequence	305	PDDITNEEY	0.078	.
Sequence	317	YKSLTNDWE	0.166	.
Sequence	425	LELFTLAE	0.015	.
Sequence	467	LRYYSASG	0.162	.
Sequence	482	KDYCTRMKE	0.056	.
Sequence	495	IYYITGETK	0.025	.
Sequence	498	ITGETKDQV	0.686	*T*
Sequence	540	FEGKTLVSV	0.298	.
Sequence	545	LVSVTKEGL	0.029	.
Sequence	566	EKKTKFEN	0.657	*T*
Sequence	594	NRLVTSPCC	0.092	.
Sequence	601	CCIVTSTYG	0.010	.

Sequence	603	IVTSTYGWT	0.271	.
Sequence	607	TYGWTANME	0.015	.
Sequence	624	RDNSTMGYM	0.005	.
Sequence	645	SIETLRQK	0.789	*T*
Sequence	669	LLYETALLS	0.008	.
Sequence	683	EDPQTHANR	0.115	.
Sequence	704	EDDPTADD	0.431	.
Sequence	708	TADDTSAAV	0.403	.
Sequence	713	SAAVTEEMP	0.234	.
Sequence	725	GDDDTSRME	0.944	*T*
Sequence	743	PHEATSHXC	0.380	.
Sequence	752	KPRXTEINH	0.355	.

Tyrosine predictions

Name	Pos	Context	Score	Pred
v				
Sequence	38	INTFYSNKE	0.501	*Y*
Sequence	61	DKIRYESLT	0.903	*Y*
Sequence	139	GVGFYSAYL	0.012	.
Sequence	142	FYSAYLVAE	0.068	.
Sequence	160	DDEQYAWES	0.933	*Y*
Sequence	197	DQTEYLEER	0.615	*Y*
Sequence	216	QFIGYPITL	0.040	.
Sequence	284	IKEKYIDQE	0.957	*Y*
Sequence	309	TNEEYGEFY	0.939	*Y*
Sequence	313	YGEFYKSLT	0.772	*Y*
Sequence	364	NIKLYVRRV	0.040	.
Sequence	381	LIPEYLNFI	0.062	.
Sequence	434	DKENYKKFY	0.945	*Y*
Sequence	438	YKKFYEQFS	0.445	.
Sequence	465	ELLRYYTSA	0.127	.
Sequence	466	LLRYYTSAS	0.014	.
Sequence	480	SLKDYCTRM	0.074	.
Sequence	492	QKHIYYITG	0.632	*Y*
Sequence	493	KHIYYITGE	0.077	.
Sequence	520	LEVIYMIEP	0.620	*Y*
Sequence	528	PIDEYCVQQ	0.177	.
Sequence	604	VTSTYGWTA	0.215	.
Sequence	627	STMGYMAAK	0.257	.
Sequence	667	VILLYETAL	0.026	.
Sequence	689	ANRIYRMIK	0.130	.



[Explain the output.](#) [Go back.](#)

5) CD44 E.

NetNGlyc 1.0 Server - prediction results

Page 1 of 2



NetNGlyc 1.0 Server - prediction results

Technical University of Denmark

Asn-Xaa-Ser/Thr sequons in the sequence output below are highlighted in blue.
Asparagines predicted to be N-glycosylated are highlighted in red.

Output for 'Sequence'

Name: Sequence Length: 493

```

MDKFWHAAWGLCLVPLSLAQIDLNITCRFAGVFHVEKNGRYSISRTEAADLCKAFNSTLPTMAQMEKALSIGFETCRYG      80
FIEGHVVIPIRIHPNSICAANNTGVYILTYNTSQYDTCFNASAPPEEDCTSVTDLPNAFDGPITITIVNRDGTTRYVQKGE      160
YRTNPEDIYPSNPTDDDDVSSGSSSERSTSGGYIFYTFSTVHPIPEDSPWITDSTDRIPTNMDSSHSTTLQPTANPNT      240
GLVEDLDRGTPLSMTTQQSNSQSFSTSHGLEEDKDHPPTSTLTSSNRNDVTGGRDPNHSEGSTHLEGYTSHYPHTKE      320
SRTFIPVTSKATGSFGVTAVTGDNSNVNRSLSGDQDTFHPSGGSHTHGSESDGSHSGSQEGGANTTSGPIRTPQIPE      400
WLIILASLLALILAVCIAVNSRRRCGQKKLVINSNGAVEDRKPSGLNGEASKSQEMVHLVNKESSETPDQFMTADE      480
TRNLQNVDKIGV
.....N.....n.....N.....
.....n.....n.....N.....N.....n.....n.....
.....N.....n.....n.....n.....n.....
.....n.....n.....N.....N.....
.....n.....n.....n.....n.....
.....n.....
.....

```

(Threshold=0.5)

SeqName	Position	Potential	Jury agreement	N-Glyc result
Sequence	25 NITC	0.7530	(9/9)	+++ SEQUON ASN-XAA-SER/THR.
Sequence	39 NGRY	0.6472	(7/9)	+
Sequence	57 NSTL	0.7594	(9/9)	+++ SEQUON ASN-XAA-SER/THR.
Sequence	94 NSIC	0.6463	(8/9)	+
Sequence	100 NNTG	0.4908	(6/9)	- SEQUON ASN-XAA-SER/THR.
Sequence	101 NTGV	0.6187	(8/9)	+
Sequence	110 NTSQ	0.6913	(9/9)	++ SEQUON ASN-XAA-SER/THR.
Sequence	120 NASA	0.5967	(6/9)	+
Sequence	137 NAFD	0.6459	(9/9)	++ SEQUON ASN-XAA-SER/THR.
Sequence	149 NRDG	0.6422	(7/9)	+
Sequence	164 NPED	0.7314	(9/9)	++ WARNING: PRO-X1.
Sequence	172 NPTD	0.6689	(9/9)	++ WARNING: PRO-X1. SEQUON ASN-XAA-SER/THR.
Sequence	223 NMDT	0.6997	(9/9)	++
Sequence	237 NPNT	0.4517	(6/9)	-
Sequence	239 NTGL	0.7009	(9/9)	++
Sequence	260 NSQS	0.6551	(9/9)	++
Sequence	287 NRND	0.5933	(8/9)	+
Sequence	289 NDVT	0.3702	(8/9)	-
Sequence	299 NHSE	0.3889	(8/9)	- SEQUON ASN-XAA-SER/THR.
Sequence	346 NSNV	0.5609	(7/9)	+
Sequence	348 NVNR	0.6819	(9/9)	++
Sequence	350 NRSI	0.6722	(9/9)	++ SEQUON ASN-XAA-SER/THR.
Sequence	387 NTTS	0.6313	(8/9)	+
Sequence	422 NSRR	0.6748	(9/9)	++ SEQUON ASN-XAA-SER/THR.

<http://www.cbs.dtu.dk/cgi-bin/nph-webface?jobid=netNglyc,4C844A1E027C45A7&opt=...> 07/09/2010

Sequence	436	NSGN	0.4734	(7/9)	-
Sequence	439	NGAV	0.5226	(6/9)	+
Sequence	451	NGEA	0.6176	(8/9)	+
Sequence	465	NKES	0.6771	(8/9)	+
Sequence	483	NLQN	0.5161	(3/9)	+
Sequence	486	NVDM	0.4709	(6/9)	-

×

[Graphics in PostScript](#)

[Explain the output.](#) [Go back.](#)



NetOGlyc 3.1 Server - prediction results

Technical University of Denmark

Name: Sequence Length: 493
MDKFWHAAWGLCLVPLSLAQIDLNITCRFAGVFHVEKNGRYSISRTEADLCKAFNSTLPTMAQMEKALSIGFETCRYG
FIEGHVVIPRIHPNSICAANNTGVYILTYNTSQYDTCFNASAPPEEDCTSVTDLPNAFDGPITITIVNRDGRYVQKGE
YRTNPEDIYPSNPTDDDVSSGSSSSRSSTSGGYIFYTFSTVHPIPEDDSPWITDSTDRIPRTNMDSSHSTTLQPTANPNT
GLVEDLDRTPLSMTTQQSNSQSFSSTHEGLEEDKDHPPTSTLTSSNRNDVTGGRDPNHSEGSTHLEGYTSHYPHTKE
SRTFIPVTSAKTGSFVTAVTVGDSNSNVNRLSGDQDTFHPSGGSHTHGSESDGSHSGSQEGGANTTSGPITPQIPE
WLILASLLALALILAVCIAVNSRRRCGQKKLVINSNGGAVEDRKPSGLNGEASKSQEMVHLVNKESSETPDQFMTADE
TRNLQNVDKIGV

.....
.....T.....T.....T.....T.....
.....T.....T.....
.....T.....T.....
.....

Name	S/T	Pos	G-score	I-score	Y/N	Comment
Sequence	S	18	0.071	0.028	.	-
Sequence	T	27	0.119	0.053	.	-
Sequence	S	43	0.121	0.040	.	-
Sequence	S	45	0.137	0.028	.	-
Sequence	T	47	0.266	0.199	.	-
Sequence	S	58	0.201	0.041	.	-
Sequence	T	59	0.259	0.054	.	-
Sequence	T	62	0.241	0.312	.	-
Sequence	S	71	0.141	0.082	.	-
Sequence	T	76	0.225	0.034	.	-
Sequence	S	95	0.133	0.021	.	-
Sequence	T	102	0.232	0.025	.	-
Sequence	T	108	0.196	0.058	.	-
Sequence	T	111	0.239	0.057	.	-
Sequence	S	112	0.167	0.018	.	-
Sequence	T	116	0.290	0.053	.	-
Sequence	S	122	0.263	0.109	.	-
Sequence	T	130	0.444	0.118	.	-
Sequence	S	131	0.360	0.111	.	-
Sequence	T	133	0.474	0.298	.	-
Sequence	T	144	0.282	0.079	.	-
Sequence	T	146	0.281	0.078	.	-
Sequence	T	153	0.270	0.032	.	-
Sequence	T	163	0.334	0.084	.	-
Sequence	S	171	0.365	0.462	.	-
Sequence	T	174	0.577	0.233	T	-
Sequence	S	179	0.463	0.098	.	-
Sequence	S	180	0.464	0.120	.	-
Sequence	S	182	0.426	0.019	.	-
Sequence	S	183	0.410	0.020	.	-
Sequence	S	184	0.414	0.034	.	-

<http://www.cbs.dtu.dk/cgi-bin/nph-webface?jobid=netOglyc.4C844AJA03EC7B44&opt=...> 07/09/2010

Sequence
LEFTALS 199
0.009

NetPhos 2.0 Server - prediction results

Sequence	S	187	0.362	0.023	.	-
Sequence	S	188	0.412	0.089	.	-
Sequence	T	189	0.476	0.068	.	-
Sequence	S	190	0.367	0.060	.	-
Sequence	T	197	0.494	0.055	.	-
Sequence	S	199	0.340	0.023	.	-
Sequence	T	200	0.450	0.159	.	-
Sequence	S	209	0.347	0.031	.	-
Sequence	T	213	0.520	0.091	T	-
Sequence	S	215	0.444	0.021	.	-
Sequence	T	216	0.561	0.114	T	-
Sequence	T	222	0.502	0.020	T	-
Sequence	S	226	0.397	0.025	.	-
Sequence	S	227	0.394	0.024	.	-
Sequence	S	229	0.387	0.057	.	-
Sequence	T	230	0.492	0.019	.	-
Sequence	T	231	0.439	0.032	.	-
Sequence	T	235	0.438	0.577	T	-
Sequence	T	240	0.482	0.044	.	-
Sequence	T	249	0.434	0.075	.	-
Sequence	S	253	0.314	0.051	.	-
Sequence	T	255	0.412	0.145	.	-
Sequence	T	256	0.382	0.063	.	-
Sequence	S	259	0.279	0.019	.	-
Sequence	S	261	0.258	0.034	.	-
Sequence	S	263	0.326	0.029	.	-
Sequence	S	265	0.373	0.044	.	-
Sequence	T	266	0.516	0.057	T	-
Sequence	S	267	0.398	0.034	.	-
Sequence	T	279	0.469	0.444	.	-
Sequence	T	280	0.472	0.061	.	-
Sequence	S	281	0.336	0.033	.	-
Sequence	T	282	0.401	0.062	.	-
Sequence	T	284	0.426	0.054	.	-
Sequence	S	285	0.309	0.078	.	-
Sequence	S	286	0.324	0.034	.	-
Sequence	T	292	0.500	0.057	.	-
Sequence	S	301	0.245	0.078	.	-
Sequence	S	304	0.246	0.031	.	-
Sequence	T	305	0.360	0.064	.	-
Sequence	T	312	0.413	0.052	.	-
Sequence	S	313	0.325	0.059	.	-
Sequence	T	318	0.407	0.042	.	-
Sequence	S	321	0.298	0.031	.	-
Sequence	T	323	0.454	0.036	.	-
Sequence	T	328	0.504	0.088	T	-
Sequence	S	329	0.353	0.093	.	-
Sequence	T	332	0.453	0.064	.	-
Sequence	S	334	0.318	0.068	.	-
Sequence	T	338	0.423	0.051	.	-
Sequence	T	341	0.423	0.209	.	-
Sequence	S	345	0.236	0.021	.	-
Sequence	S	347	0.278	0.018	.	-
Sequence	S	352	0.311	0.040	.	-
Sequence	S	354	0.336	0.066	.	-
Sequence	T	359	0.426	0.051	.	-
Sequence	S	363	0.330	0.099	.	-
Sequence	S	366	0.384	0.023	.	-
Sequence	T	368	0.468	0.030	.	-
Sequence	T	369	0.481	0.081	.	-

Sequence	S	372	0.358	0.037	.	-
Sequence	S	374	0.433	0.033	.	-
Sequence	S	378	0.445	0.052	.	-
Sequence	S	381	0.469	0.036	.	-
Sequence	T	388	0.498	0.036	.	-
Sequence	T	389	0.485	0.629	T	-
Sequence	S	390	0.320	0.062	.	-
Sequence	T	395	0.380	0.098	.	-
Sequence	S	407	0.107	0.033	.	-
Sequence	S	423	0.075	0.028	.	-
Sequence	S	437	0.143	0.048	.	-
Sequence	S	448	0.221	0.036	.	-
Sequence	S	455	0.233	0.031	.	-
Sequence	S	457	0.271	0.056	.	-
Sequence	S	468	0.236	0.034	.	-
Sequence	S	469	0.215	0.079	.	-
Sequence	T	471	0.258	0.043	.	-
Sequence	T	477	0.265	0.089	.	-
Sequence	T	481	0.291	0.071	.	-

x

[Graphics in PostScript](#)

[Explain the output.](#) [Go back.](#)

NetNGlyc 1.0 Server - prediction results

70-KDA



Technical University of Denmark

Output for 'Sequence'

[illegible]

```
(Threshold=0.5)
```

SeqName	Position	Potential	Jury agreement	N-Glyc result
Sequence	27 NHHF	0.5579	(6/9)	+
Sequence	52 NDDY	0.5307	(5/9)	+
Sequence	72 NGSF	0.6278	(8/9)	+
Sequence	79 NTGY	0.6249	(7/9)	+
Sequence	90 NRVN	0.7770	(9/9)	+++
Sequence	93 NDGV	0.5789	(8/9)	+
Sequence	107 NSGV	0.7637	(9/9)	+++
Sequence	114 NTCK	0.7566	(9/9)	+++
Sequence	374 NKFE	0.6254	(5/9)	+
Sequence	392 NLEQ	0.6448	(9/9)	++
Sequence	404 NGEF	0.5196	(6/9)	+
Sequence	420 NEYY	0.6436	(9/9)	++
Sequence	475 NRST	0.5451	(4/9)	+

X

Graphics in PostScript

Explain the output. Go back.

<http://www.cbs.dtu.dk/cgi-bin/nph-webface?jobid=netNglyc,4C844AE803FA3AA4&opt...> 07/09/2010

80-KOA



NetOGlyc 3.1 Server - prediction results

Technical University of Denmark

Name: Sequence Length: 527

MLLPLLLLLPMCWAVEVKRPRGVSLTNHHFYDESKPFTCLDGSATIPFDQVNDYDCCKDGSDEPGTAACPNGSFHCTNT
GYKPLYIPSNRVNDGVCDCCDGTDEYNSGVICENTCKEKGRKERESLQQMAEVTREGFRLKKILIEDWKKAREEKQKKLI
ELQAGKKSLEDQVEMLRVTKEEAKEPEREAKQHQKLWEEQLAAAKAQEQELAADAFKELDDMDGTVSVELQTHPEL
DTDGDGALSEAEQAALLSGDTQTDATSFYDRVWAAIRDKYRSEALPTDLPAPAPDLTEPKKEQPPVPSSPTEEEEEEEE
EEEEEEEEEEEEEDSEEAPPPLSPQPASPAEEDKMPPYDEQTQAFIDAAQEARNKFEEAERSLKDMEESIRNLEQEISFD
FGPNGEFAYLYSQCYELTTNEYVYRLCPFKLVSQKPKLGGSPTSLGTWGSWIGPDHDKFSAMKYEQGTGCWQGPNRSTTV
RLLCGETMTVSTTEPSRCEYLMELMTPAACPEPPPEAPTEDDHDEL

.....
.....
.....T.....T.....SS.T.....
.....S.....S.....S.....
.....T.....T.....

Name	S/T	Pos	G-score	I-score	Y/N	Comment
Sequence	S	24	0.167	0.029	.	-
Sequence	T	26	0.223	0.106	.	-
Sequence	S	34	0.193	0.029	.	-
Sequence	T	38	0.256	0.049	.	-
Sequence	S	43	0.133	0.035	.	-
Sequence	T	45	0.184	0.101	.	-
Sequence	S	62	0.196	0.057	.	-
Sequence	T	67	0.263	0.221	.	-
Sequence	S	74	0.272	0.054	.	-
Sequence	T	78	0.314	0.029	.	-
Sequence	T	80	0.311	0.019	.	-
Sequence	S	89	0.129	0.059	.	-
Sequence	T	103	0.183	0.033	.	-
Sequence	S	108	0.103	0.057	.	-
Sequence	T	115	0.195	0.020	.	-
Sequence	S	126	0.125	0.020	.	-
Sequence	T	134	0.137	0.127	.	-
Sequence	S	168	0.108	0.025	.	-
Sequence	T	178	0.270	0.046	.	-
Sequence	T	228	0.282	0.050	.	-
Sequence	S	230	0.193	0.035	.	-
Sequence	T	232	0.303	0.226	.	-
Sequence	T	236	0.336	0.052	.	-
Sequence	T	242	0.375	0.080	.	-
Sequence	S	249	0.232	0.054	.	-
Sequence	S	258	0.192	0.078	.	-
Sequence	T	261	0.298	0.020	.	-
Sequence	T	263	0.253	0.069	.	-
Sequence	T	266	0.214	0.030	.	-
Sequence	S	267	0.154	0.018	.	-
Sequence	S	282	0.252	0.078	.	-

<http://www.cbs.dtu.dk/cgi-bin/nph-webface?jobid=netOglyc,4C844AF803FAEB61&opt=...> 07/09/2010

80 kDa

Sequence	T	287	0.477	0.704	T	-
Sequence	S	293	0.485	0.405	.	-
Sequence	T	298	0.703	0.173	T	-
Sequence	S	309	0.575	0.478	S	-
Sequence	S	310	0.585	0.107	S	-
Sequence	T	312	0.677	0.405	T	-
Sequence	S	334	0.607	0.051	S	-
Sequence	S	342	0.601	0.322	S	-
Sequence	S	348	0.555	0.049	S	-
Sequence	T	362	0.355	0.087	.	-
Sequence	S	382	0.140	0.023	.	-
Sequence	S	389	0.126	0.033	.	-
Sequence	S	398	0.101	0.068	.	-
Sequence	S	412	0.092	0.054	.	-
Sequence	T	418	0.140	0.031	.	-
Sequence	T	419	0.120	0.104	.	-
Sequence	S	433	0.200	0.083	.	-
Sequence	S	441	0.246	0.027	.	-
Sequence	T	443	0.362	0.159	.	-
Sequence	S	444	0.222	0.053	.	-
Sequence	T	447	0.398	0.064	.	-
Sequence	S	450	0.216	0.030	.	-
Sequence	S	460	0.156	0.030	.	-
Sequence	T	468	0.259	0.034	.	-
Sequence	S	477	0.187	0.038	.	-
Sequence	T	478	0.280	0.020	.	-
Sequence	T	479	0.315	0.058	.	-
Sequence	T	488	0.351	0.032	.	-
Sequence	T	491	0.283	0.288	.	-
Sequence	S	492	0.229	0.022	.	-
Sequence	T	493	0.341	0.220	.	-
Sequence	T	494	0.321	0.311	.	-
Sequence	S	497	0.247	0.018	.	-
Sequence	T	507	0.583	0.287	T	-
Sequence	T	520	0.658	0.555	T	-

x

[Graphics in PostScript](#)[Explain the output.](#) [Go back.](#)<http://www.cbs.dtu.dk/cgi-bin/nph-webface?jobid=netOglyc,4C844AF803FAEB61&opt=...> 07/09/2010