

David Jiménez Morales

Estudio Bioinformático sobre la presencia de proteínas glucolíticas en la superficie celular de *Saccharomyces cerevisiae*

Departamento de Microbiología II
Facultad de Farmacia
Universidad Complutense de Madrid
Tutora: Dra. Concha Gil

INTRODUCCIÓN

En los últimos años, el grupo de investigación dirigido por la Dra. Concha Gil, a raíz de una serie de datos experimentales acumulados, ha centrado una de sus líneas de investigación en la presencia de ciertas enzimas glucolíticas –típicamente citoplasmáticas- en la superficie celular de *Saccharomyces cerevisiae* y *Candida albicans*. La proteína glucolítica seleccionada como modelo para tratar de identificar la posible vías y función en esta localización es la Enolasa (Eno2p).

Nuestro objetivo consiste en tratar de contribuir a despejar las incógnitas aportando datos que puedan obtenerse a través de métodos bioinformáticos, desde descripciones fisico-químicas secuencias, de estructurales, análisis bibliográfico..

Justificación experimental

La pared celular de levaduras es una estructura rígida que protege a la célula de cambios osmóticos. También es un filtro que permite el paso de ciertas sustancias y excluye el paso a otras. No se trata de una estructura inerte si no que está sujeta a cambios en función de múltiples factores. La composición de la misma es conocida, pero los mecanismos de ensamblaje y los responsables del mismo no están bien definidos. En ese sentido se analizaron mediante electroforesis bidimensional las proteínas que son secretadas en protoplastos de levadura (células a las que se le ha extraído la pared celular) {Pitarch, Pardo, et al. 1999 2 /id}.

Entre las proteínas que se identificaron aparecen la 2-Fosfoglicerato Dehidratasa ó Enolasa (Eno2p), Fructosa-1-6-Bifosfato Aldolasa (Fba1), Gliceraldehido-3-Fosfato Deshidrogenasa (TDH3-2), entre otras.

Como describiremos más adelante, son numerosas las citas que hablan de la presencia de ciertas enzimas glicolíticas en la superficie celular no solo de *Saccharomyces* o *Candida*, si no en numerosos organismos. Pero existe controversia en la explicación de la presencia de estas proteínas típicamente citoplasmáticas en la superficie celular de levadura, como consecuencia de lisis celular o por otros posibles mecanismos desconocidos {Eroles, Sentandreu, et al. 1997 3 /id}

Proteínas de fusión Eno2p-Suc2 y Fba1p-Suc2

Mediante el diseño de las proteínas de fusión Eno2p y Fba1p con la Invertasa (Suc2p) muestran que estas dos típicas proteínas citoplasmáticas son capaces de alcanzar el exterior celular {Pitarch, Pardo, et al. 1999 2 /id}. Suc2p es una hidrolasa (rompe la sacarosa para formar glucosa y fructosa) que presenta una forma intracelular y otra forma que es secretada por la vía clásica, es decir, a través del Retículo Endoplasmático-Aparato de Golgi (RE-Golgi). Para que una proteína cualquiera sea secretada de esta forma, es necesario que disponga en su extremo N-terminal un pequeño *péptido señal* y normalmente también suele modificarse post-traduccionamente mediante glicosilación. Suc2p en su forma extracelular presenta estas características. Mediante la obtención de un mutante al que se le ha delecionado *suc2*, se consigue que la levadura sea incapaz de crecer en medios con sacarosa como única fuente de carbono. La fusión de Eno2p y Fba1p con Suc2p sin su péptido señal y la posterior traslocación en la cepa mutante permitía comprobar si esta levadura es capaz de crecer en sacarosa.

Estudio experimental sobre la presencia de Eno2p en la superficie celular

Posteriores estudios han confirmado la presencia de la Enolasa en la pared celular de *Candida* {Pitarch, Sanchez, et al. 2002 1 /id}.

Con el fin de tratar de identificar los mecanismos de secreción implicados en la localización extracelular de ciertas proteínas glucolíticas, se ha centralizado el análisis en el estudio de Eno2p, realizando los siguientes experimentos:

-Determinación de la región de Eno2p involucrada en la posible exportación de Suc2:

Tras la fusión de diferentes fragmentos de Eno2p a la invertasa extracelular (Figura 1), se llegó a la conclusión de que los primeros 169 aminoácidos del extremo N-terminal eran suficientes para el crecimiento de la levadura en sacarosa.

-Medida de la actividad enzimática de Eno2p en la superficie celular de levadura.

A partir de la adicción del sustrato de la enzima (2-Fosfo-D-glicerato) se midió la cantidad de producto resultante (Fosfoenolpiruvato). Mientras que en extractos celulares lisados el resultado fue de un 19%, en células completas fue de 0,7%. La actividad del buffer control blanco era de un 0,5%.

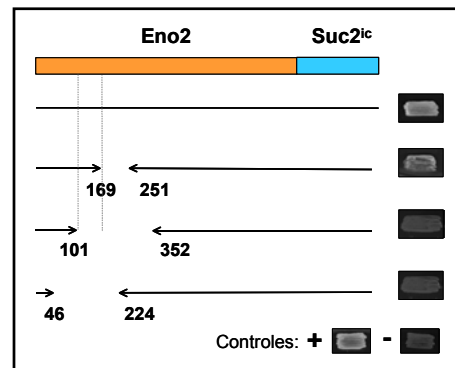


Figura 1. Representación esquemática de cada uno de los cortes efectuados con el fin de estudiar la posible región proteica mínima necesaria para la exportación.

-Citometría de flujo (FACS) empleando anticuerpos anti-enolasa humana (N300) y FITC.

No hubo detección de enolasa en la superficie celular mediante ésta técnica.

-Relación entre los niveles de expresión y la exportación de la proteína de fusión Eno2p-Suc2i

Mediante el empleo de un promotor regulable se controló la expresión de la proteína de fusión viendo que a bajos niveles de expresión hay crecimiento en sacarosa.

-Estudio del posible papel de las proteínas que forman complejos proteicos con Eno2p en la localización extracelular.

Se estudio el papel que podría ejercer las proteínas Lst8p, Spf1p y Apl6p, puesto que la delección de estos genes no provocó la ausencia de enolasa en el exterior celular. Aclarar que el mutante para Lst8p es parcial puesto que se trata de un gen esencial.

-Comparación de la Enolasa presente en citoplasma, Lipid Rafts y pared celular.

En los resultados obtenidos mediante Western-Blot (empleando como anticuerpo frente a enolasa humana N300), la enolasa fue identificada en fracción de Lipid Rafts y no en pared celular.

-Revisión de la sensibilidad de N300 frente a diferentes fragmentos de Eno2p fusionados a Suc2p.

Vías de Secreción descritas en células eucariotas

Para poder descifrar la forma mediante la cual la enolasa (y demás proteínas glucolíticas relacionadas) alcanza la superficie celular, es conveniente tener en cuenta cuáles son los mecanismos actualmente descritos de secreción para **proteínas que se encuentran solubles en citoplasma**. Básicamente pueden agruparse en dos:

- Secreción clásica dependiente del sistema RE-Golgi

Utilizan esta vía las proteínas solubles que contienen un su extremo amino terminal (N-terminal) una pequeña secuencia de unos 20 aminoácidos que dirige la translocación a través del Retículo Endoplasmático (RE) {Walter, Gilmore, et al. 1984 52 /id}. A través del transporte de vesículas y vía Golgi, finalmente son liberadas al espacio extracelular, mediante la fusión de las vesículas derivadas de Golgi con la membrana plasmática. Modificaciones post-traduccionales asociadas a esta vía es la Glicosilación. Tanto la Eno2p como el resto de glucolíticas citadas carecen de éste péptido señal.

- Secreción no clásica

En los últimos 11 años se han ido acumulando datos {Nickel 2003 4 /id} sobre mecanismos de secreción en los que no media el sistema RE-Golgi. Se dice que una proteína utiliza un mecanismo alternativo al clásico en células eucarióticas cuando:

- Carecen del péptido señal convencional.
- La exclusión de estas proteínas de los clásicos orgánulos de secreción tipo RE y el Golgi, combinado con la pérdida con las modificaciones post-traduccionales dependientes de RE-Golgi como es la N-glicosilación.
- Resistencia de estos procesos de exportación al Brefeldin A, un clásico inhibidor de la secreción dependiente de RE-Golgi.

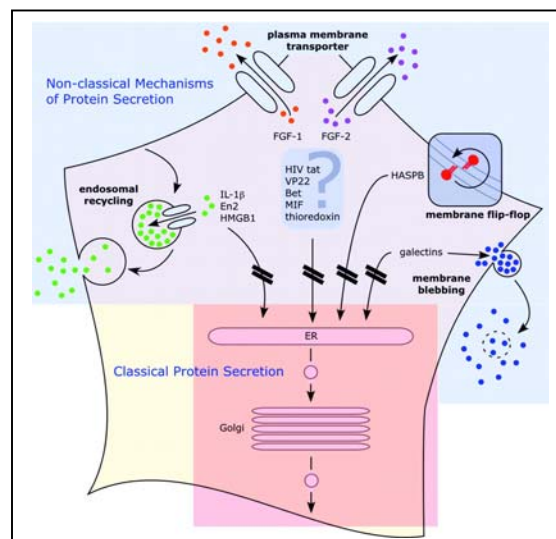


Figura 2. Resumen de los diferentes mecanismos de secreción no clásica hasta la fecha descritos

Los mecanismos de secreción no clásica hasta la fecha identificados pueden agruparse de la siguiente manera (Figura 2):

- **Exocitosis lisosomal:** podríamos citar a diferentes proteínas en principio no relacionadas como son una citokina, la interleukina-1 (IL-1 β), la homeoproteína *Engrailed* (En2) o el factor intranuclear HMGB1. Todas ellas serían translocadas al interior de ciertas vesículas intracelulares no dependientes de la maquinaria RE-Golgi, que tras la fusión con la membrana, liberarían a las mencionadas proteínas al exterior celular. La translocación al interior vesicular sería mediada por unos transportadores de membrana de tipo ABC. No obstante se ha sugerido la existencia de diferentes *poblaciones* de vesículas endolisosómicas en base a que cada proteína es secretada en respuesta a un estímulo concreto.
- **Transportadores de la Membrana Plasmática** que mediaran en la secreción de ciertas proteínas como es el caso de los Factores de Crecimiento de Fibroblastos 1 y 2 (FGF-1y FGF-2). La exportación de FGF-1 implicaría la formación del complejo FGF-1, p40-Syt1 y S100A13, seguido de la translocación directa a través de la membrana plasmática (se desconoce la maquinaria empleada). En el caso de FGF-2, los α/β heterodímeros que constituyen las Na⁺/K⁺-ATPasas funcionales en términos de transportadores de iones, serían capaces también de formar complejos funcionales que catalizaran la exportación de una manera independiente del potencial de membrana.
- **Translocación directa a través de membrana,** como sería el caso de la proteína de *Leishmania* HASPB. Para ello sería requisito previo la acetilación, palmitoilación y miristoilación de la misma. Se ha postulado que HASPB sería transferida del citoplasma a la cara externa de la membrana del Golgi, desde donde sería transportada a la cara interna de la membrana plasmática por la vía convencional. Se desconoce cómo esta proteína alcanzaría la cara externa de la membrana.
- **Formación de exosomas de membrana,** como es el caso de las galactinas. Éstas se acumularían tras determinadas regiones de la membrana plasmática a lo que seguiría un mecanismo de exportación que implicaría la formación de pequeñas vesículas de membrana, llamadas *exosomas*. En este proceso parecen estar implicadas la cadenas laterales de oligosacáridos de β -galactosa N- y O- glicosilada, junto con galactosas que contienen glucolípidos como GM₁.
- **Otras proteínas secretadas,** como es el caso de HIV-Tat, VP22, Bet, las tioredoxinas, etc, de las cuales se ha demostrado que no son secretadas por la vías clásica pero aún se desconoce los mecanismos empleados.

Ante la evidencia de los nuevos mecanismos de secreción alternativos a la vía clásica que han ido apareciendo en los últimos años, es posible plantear la posibilidad de que la Eno2p y relacionadas alcancen la superficie celular mediante un mecanismo de secreción no clásica y que así justificara, por ejemplo, la ausencia de péptido señal

Modificaciones Post-traduccionales Proteicas Asociadas a las Vías de Secreción

El principal mecanismo que dispone una célula para *marcar* una proteína que va a ser secretada es a través de las modificaciones post-traduccionales. A continuación resumiremos aquellas modificaciones que han sido relacionadas con algún tipo de mecanismo de secreción, lo cual a su vez tratará de orientar las predicciones bioinformáticas.

N- y O- Glicosilación

Muchas proteínas en eucariotas presentan cadenas de oligosacáridos unidas covalentemente a las cadenas laterales de asparragina mediante enlaces N-glicosídicos o bien a las cadenas laterales de serina o treonina mediante enlaces O-glicosídicos.

La glicosilación puede estar involucrada {Helenius & Aebi 2001 53 /id} en (i) funciones intercelulares relacionados con procesos de información, como las interacciones y señalización célula-célula, así como en (ii) ciertos papeles intracelulares como en el control de la secreción proteica a través de la vía clásica (elegir el camino a través del cual tiene que salir una determinada proteína).

Una amplia mayoría de glicoproteínas en eucariotas menores (como la levadura) son proteínas de la pared celular.

Modificaciones Lipídicas

En eucariotas, cuatro principales formas de modificaciones lipídicas han sido reconocidas:

1.- Miristoilación {Farazi, Waksman, et al. 2001 12 /id}

Es la unión covalente del **miristato**, un ácido graso saturado de 14 carbonos, a la Glicina N-terminal, de forma co-traduccional y previa eliminación del residuo de Metionina inicial (mediante metionilaminopeptidasas). Existe un caso {Zha, Weiler, et al. 2000 54 /id} de miristoilación post-traduccional y sería tras la rotura proteolítica de la proteína, lo cual revelaría un sitio de miristoilación oculto.

La N-miristoilación promueve interacciones débiles y reversibles entre proteína-membrana y proteína-proteína. Típicamente, el miristato actúa junto con otros mecanismos en la regulación del marcaje proteico y la función proteica. La forma puede ser a través de los llamados "*myristoyl-electrostatic switches*", es decir, interacciones con la membrana debido a la carga positiva del miristato. También podría ser como consecuencia del llamado "*myristoyl-conformational switch*": ya sea a través de la unión de ciertos ligandos o bien como consecuencia de cortes proteolíticos en la proteína miristoilada, el resultado sería que se provocarían cambios conformacionales que ocultaran o expusieran la cadena acil.

2.- Palmitoilación {Bijlmakers & Marsh 2003 11 /id}

Es la más abundante y aparentemente más versátil de las modificaciones Lipídicas. Es la adición de **ácido palmítico**, el cual es esterificado en el tiol libre de la cisteína (otros ácidos grasos pueden ser empleados, tanto saturados como insaturados).

Los mecanismos involucrados en la palmitoilación son poco conocidos y las enzimas relevantes están en su mayoría sin caracterizar. Además, el amplio rango de substratos, la ausencia de un claro motivo consenso, junto con el hecho de que ocurre en varias localizaciones, todo ello contribuye al oscuro conocimiento de ésta modificación.

Las funciones asociadas a la Palmitoilación también es de amplio rango: localización intracelular; efectos en la endocitosis, etc. Además, la corta vida media de la Palmitoilación (mucho menor que la de la proteína en si) es importante en los procesos de señalización y regulación.

3.- La modificación de las proteínas de membrana mediante el glicofosfatidil inositol.

Acetilación

Consiste en la adicción de un grupo acetilo en el aminoácido del extremo N-terminal. Este tipo de modificación se produce con mucha frecuencia y sus significados pueden ser variados. En ciertos casos de secreción no convencional, como en el de las galactinas, la acetilación favoreció la exportación {Mehul & Hughes 1997 55 /id}. Así mismo, la exportación de HASPB de *Leishmania* es dependiente de la acetilación {Denny, Gokool, et al. 2000 5 /id}.

Fosforilación

Es la adición de al menos un grupo fosfato a los residuos de serina, treonina o tirosina (menos frecuente es también la fosforilación del ácido aspártico y la histidina). La fosforilación es quizás la modificación post-traducciona más habitual, con un significado biológico muy amplio.

Respecto a su posible papel en la regulación de la exportación de proteínas, se ha visto que la fosforilación inhibe la secreción no clásica de la homeoproteína Engrailed 2, por tanto, es una modificación post-traducciona a tener en cuenta {Maizel, Tassetto, et al. 2002 49 /id}.

Ubiquitinación

La **ubiquitina** es un polipéptido de 76 aminoácidos altamente conservado y expresado en todas las células eucariotas, el cual es enganchado al residuo de lisina específico dentro de una proteína.

Dos de las mejores funciones caracterizadas sobre la ubiquitinación proteica es la marcación de las proteínas para su posterior degradación y la promoción de la internalización de ciertas proteínas de membrana por la vía endocítica hacia su degradación en la vacuola. Pero también juega un papel en la regulación y clasificación intracelular de las proteínas de membrana {Horak 2003 48 /id}.

Como veremos más adelante, la proteína Eno2p ha sido aislada en la fracción de membrana de levadura como consecuencia de una mutación en la maquinaria de degradación.

Proteínas Multifuncionales

Además de la manera en que la Eno2p alcanzan la superficie celular, la siguiente pregunta que debemos plantearnos es ¿tendría en la superficie celular la misma función glucolítica que en el citoplasma?

Son muchos los casos recogidos {Jeffery 1999 15 /id}, {Copley 2003 16 /id}, {Jeffery 2003 13 /id} en los cuales una misma proteína desempeña dos o más funciones. En la multifunción pueden influir múltiples variables como cambios en la localización celular, modificaciones post-traduccionales, estado oligomérico, expresión en un nuevo tipo celular, cambios en la concentración del ligando, sustrato, cofactor o producto...

Son varias las proteínas glucolíticas *pluriempleadas* que analizamos a continuación.

Una Mirada a las Enzimas Glucolíticas

Una vez analizadas las posibles vías de secreción y las modificaciones post-traduccionales relacionadas, el siguiente paso es analizar si existen enzimas glucolíticas identificadas en una localización no citoplasmática y/o con una función extra.

Son muchos los estudios que hacen referencia a la presencia de determinadas enzimas glucolíticas en la pared celular de levaduras, ya sea mediante la identificación por Espectrometría de Masas tras la separación mediante geles bidimensionales SDS/PAGE {Pardo, Ward, et al. 2000 59 /id}, {Molina, Gil, et al. 2000 60 /id} o mediante la unión de anticuerpos específicos frente a enzimas glucolíticas en el exterior de células completas {Alloush, Lopez-Ribot, et al. 1997 61 /id}, {Delgado, O'Connor, et al. 2001 62 /id}, {Edwards, Braley, et al. 1999 63 /id}, {Gil-Navarro, Gil, et al. 1997 64 /id}, estudiados de una manera global.

De forma concreta:

Fosfoglucosa Isomerasa (PGI)

Enzima glucolítica que cataliza en el citoplasma la interconversión de la Glucosa 6-P en Fructosa 6-P. En células de mamífero es secretada al exterior celular por varios tipos celulares y trabaja como *factor de motilidad autocrino*, *neuroleukina* y como *mediador de maduración y diferenciación*.

Fosfoglicerato mutasa (PGAM 1 o GPM1)

Como enzima glucolítica, su actividad catalítica es:

$2\text{-fosfo-D-glicerato} + 2,3\text{-difosfoglicerato} = 3\text{-fosfo-D-glicerato} + 2,3\text{-difosfoglicerato}$.

Por otro lado es secretada al exterior celular donde actúa como *reductora de puentes-disulfuro* en la Plasmina (ahora no encuentro la referencia)

Además se ha podido comprobar que cantidades significativas están presentes en la pared celular de la levadura {Motshwene, Brandt, et al. 2003 18 /id}.

Piruvato Kinasa (PYK)

Es la última enzima de la glucólisis y se encarga del paso de fosfoenolpiruvato a piruvato. En cuanto a datos que sean de interés con respecto a los objetivos de este estudio, cabe destacar tres cosas:

- a.- La piruvato kinasa es fosforilada por la protein-kinasa A y esta modificación supone un incremento de la actividad glucolítica {Portela, Howell, et al. 2002 19 /id}
- b.- Es exportada al núcleo, quizás a través de una vía independiente de Ran (una forma no-clásica de transporte al núcleo) {Yen, Roberts, et al. 2001 21 /id}.
- c.- La piruvato kinasa es transportada al núcleo debido a que es un factor necesario durante las etapas tempranas del ciclo celular, concretamente, en G₁ temprano (START A) {Grandori, Popolo, et al. 1990 20 /id}

Añadir que recientemente {Min-Oo, Fortin, et al. 2003 22 /id} se ha visto como una mutación en la piruvato kinasa protege frente a infección por *Leishmania* en ratones, y entre las causas se ha planteado que sea como consecuencia de una alteración de la membrana del eritrocito.

Gliceraldehido 3 Fosfato Dehidrogenasa (TDH3)

Esta enzima da inicio a la *segunda fase* dentro de la glucólisis. Forma un tetrámero que cataliza la siguiente conversión:



Sobre la TDH vamos a comentar tres datos que nos son de interés:

- Como monómero ha sido localizada en el núcleo celular donde presenta actividad Uracil DNA glicosilasa (UDG) {Meyer-Siegler, Mauro, et al. 1991 68 /id}.
- Se encuentra en la superficie del estafilococo como proteína receptora de transferrina {Modun, Morrissey, et al. 2000 67 /id}.
- En *Streptococcus pneumoniae* es una proteína de unión de plasminógeno de la superficie {Bergmann, Rohde, et al. 2004 73 /id}.
- Es una de las proteínas de la superficie celular de *Candida albicans* que une plasminógeno humano {Crowe, Sievwright, et al. 2003 66 /id}.
- Experimentos parecidos a los efectuados por nuestro grupo con la Eno2p fueron realizados construyendo proteínas de fusión con Suc2 intracelular y la Tdh3 de *Candida albicans* en levadura. La conclusión en este trabajo es que los 166 aminoácidos del extremo N-terminal son suficientes para la exportación {Delgado, Gil, et al. 2003 23 /id}.

Fosfoglicerato kinasa

Cataliza la interconversión del 1-3 Bifosfoglicerato en 3-Fosfoglicerato.

Se ha encontrado en la pared celular de *Candida albicans* {Alloush, Lopez-Ribot, et al. 1997 70 /id} donde además se le ha visto unir plasminógeno humano {Crowe, Sievwright, et al. 2003 66 /id}.

Fructosa Bifosfato Aldolasa (FBA)

Cataliza el final de la *primera fase* de la glucólisis (fructosa 1,6 bifosfato a glicerona fosfato y gliceraldehido 3-fosfato). Además de unir plasminógeno humano en la superficie celular de *Candida albicans* {Crowe, Sievwright, et al. 2003 66 /id} juega un papel importante en las infecciones de *Streptococcus pyogenes* como proteína de unión a fibronectina {Terao, Kawabata, et al. 2001 69 /id}.

La enolasa

Capítulo aparte merece esta proteína puesto que ha sido escogida por este laboratorio como enzima glucolítica (cataliza la conversión del 2-Fosfoglicerato en Fosfoenolpiruvato con la liberación de una molécula de agua) de referencia en el abordaje de la parte experimental.

Ha sido encontrada en la pared celular de *Saccharomyces cerevisiae* {Edwards, Braley, et al. 1999 72 /id} así como en *Candida albicans*, donde se ha visto que es capaz de unir plasminógeno, lo que potencia la invasión de *Candida* en las células endoteliales microvasculares del cerebro humano {Jong, Chen, et al. 2003 65 /id}.

Además de en *Candida*, también se ha visto que une plasminógeno en la superficie celular de *Streptococcus pyogenes* y *S.pneumoniae* {Pancholi & Fischetti 1998 25 /id}, {Bergmann, Rohde, et al. 2001 74 /id}, {Derbise, Song, et al. 2004 24 /id}; en patógeno oportunista *Pneumocystis carinii* {Fox & Smulian 2001 71 /id}; y también se ha detectado su expresión en

la superficie de células del tejido conectivo de humanos con ciertos desórdenes {Moscato, Pratesi, et al. 2000 28 /id}.

En cuanto a otras localizaciones extracelulares cabe citar las evidencias de localización nuclear en *Toxoplasma gondii* en fases activas de replicación del parásito.

Características de la Enolasa

Saccharomyces cerevisiae tiene dos enolasas diferentes codificadas por dos genes distintos que presentan un 94% de homología de secuencia nucleotídica: *eno1* cuya información se encuentra

almacenada en el cromosoma VII de la levadura y por otro lado está el gen *eno2*, situado en el cromosoma VIII.

La proteína en la que centraremos nuestro estudio es la Eno2p debido a que en la inmensa mayoría de los trabajos que hemos venido refiriendo la enolasa encontrada en la superficie celular es la Eno2p.

El gen de la enolasa

El gen de la enolasa (*eno2* YHR174W | ENO2 | P00925) es uno de los 3631 genes conocidos en levadura. Con 1314 nt de longitud, está situado en el cromosoma VIII de *Saccharomyces cerevisiae*, entre las coordenadas **451325 a 452638 (Figura 3)**. Con un porcentaje contenido de C-G del 47.6%(B1). Se trata de un gen **esencial**, es decir, la delección impide el crecimiento de la levadura. El producto del gen *eno2* se encarga de catalizar la interconversión del **2-fosfo-D-glicerato** en **fosfoenolpiruvato + H₂O**

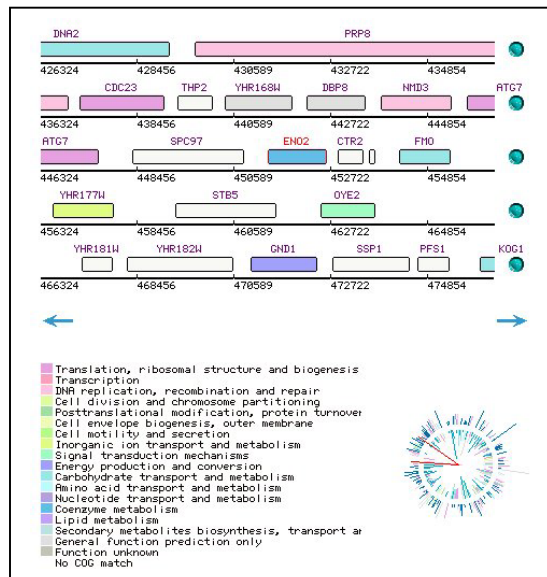


Figura 3. Representación gráfica del cromosoma VIII de *S. cerevisiae* entre 426324..476323 extraída de COG en el NCBI.

La secuencia proteica: estructura primaria

La estructura primaria es la secuencia de los aminoácidos de la enolasa.

Con una secuencia de 436 aminoácidos (sin contar la metionina inicial), alcanza un peso molecular de 46.782,9 Da. El punto Isoeléctrico teórico es de 5,67. El número y porcentaje de cada aminoácido de la secuencia es el que viene en la **tabla 1**.

El número total de aminoácidos con carga negativa (Asp + Glu) es de 59, mientras que los de carga positiva (Arg + Lys) son 49.

La vida media teórica estimada {Backhaus, Zehe, et al. 2004 9 /id} de la proteína en levadura es de 20 horas.

El índice de inestabilidad, es decir, la predicción de estabilidad de una proteína en vivo, es de 29,32 {Guruprasad, Reddy, et al. 1990 76 /id}

El índice alifático, que podría ser considerado como un factor positivo para el incremento de la termoestabilidad de proteínas globulares {Ikai 1980 77 /id} es de 88,46.

Aa - Nº	Porcentaje
Ala (A) 58	13.30%
Leu (L) 38	8.70%
Gly (G) 35	8.00%
Lys (K) 35	8.00%
Val (V) 35	8.00%
Asp (D) 30	6.90%
Ser (S) 30	6.90%
Glu (E) 28	6.40%
Asn (N) 21	4.80%
Ile (I) 20	4.60%
Thr (T) 20	4.60%
Phe (F) 15	3.40%
Arg (R) 14	3.20%
Pro (P) 13	3.00%
His (H) 10	2.30%
Tyr (Y) 10	2.30%
Gln (Q) 9	2.10%
Met (M) 9	2.10%
Trp (W) 5	1.10%
Cys (C) 1	0.20%
Asx (B) 0	0.00%
Glx (Z) 0	0.00%
Xaa (X) 0	0.00%

Tabla 1

Las propiedades de los residuos de la enolasa, es decir, su hidrofobicidad, polaridad, carga, etc quedan representados gráficamente en la **figura 4**.

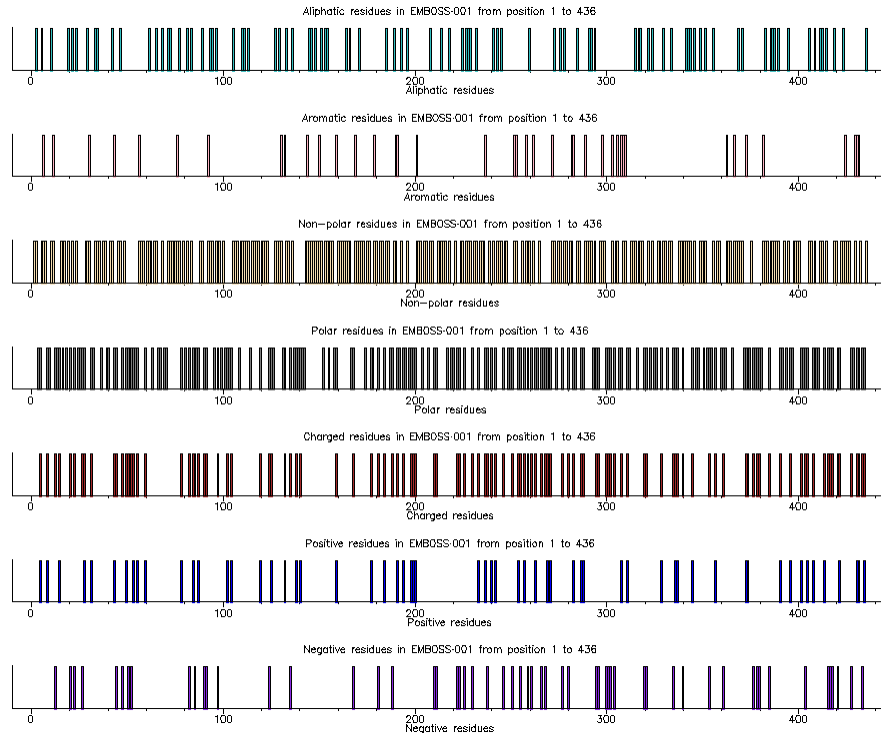
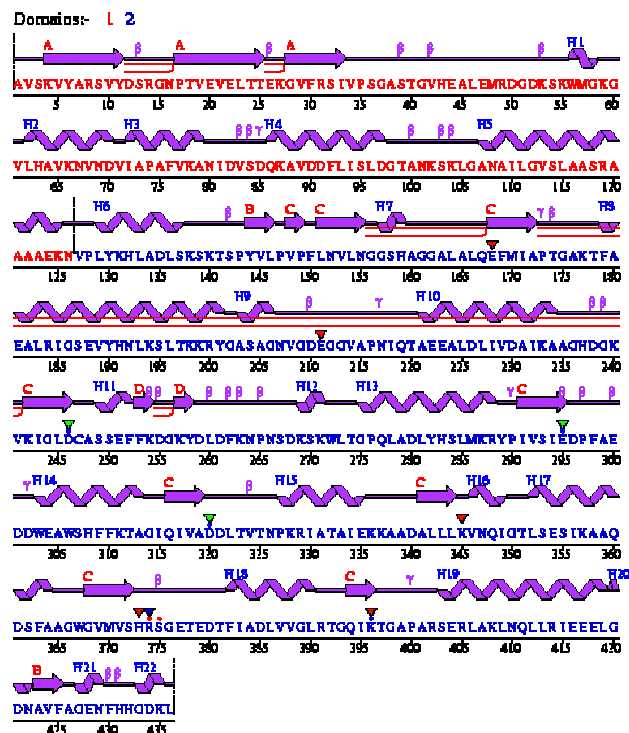


Figura 4. Representación gráfica generada por pepinfo (B2). Se representan a lo largo de la secuencia aminoacídica la distribución de residuos alifáticos, aromáticos, no-polares, polares, con carga, con carga positiva y con carga negativa.

Estructura secundaria

Los aminoácidos, mientras están siendo sintetizados y en función de la capacidad de giro del tipo de enlace, van adquiriendo una disposición espacial estable: la estructura secundaria. La enolasa de *S. cerevisiae* se encuentra cristalizada y a partir de la información contenida en dos de los archivos 4ENL.pdb y 2ONE.pdb almacenados en *The Protein Data Bank* (y que más adelante comentaremos más ampliamente), la enolasa se compone de 16 hélices alfa, 14 hélices beta y 31 regiones de giro (figura 5).



La estructura terciaria y cuaternaria de la Enolasa

Los 436 aminoácidos que constituyen la secuencia de la enolasa se disponen estructuralmente tal y como se representa en la **figura 6a**. Son numerosas las estructuras almacenadas en *The Protein Data Bank* pertenecientes a la enolasa de *S. cerevisiae* (3ENL, 4ENL, 5ENL, 6ENL, 7ENL, 1NEL, 1ELS, 1EBG, 1EBH, 1ONE, 2ONE, 1L8P).



Figura 6. a. Monómero de Eno2p (4ENL.pdb)

b. Dímero de Eno2p (2ONE.pdb)

Del análisis estructural y de la información procedente de las bases de clasificación de estructuras CATH (*Protein Structure Classification*) y SCOP (*Structural Classification of Proteins*), la estructura de la enolasa se divide en dos dominios principales (**figura 7**), que describimos a continuación:



Figura 7. Dominios de la Eno2p. En cian el C_terminal y en naranja el N_terminal.

Dominio Enolasa_N-terminal (aminoácidos 1 a 126)



Figura 8. Dominio N_terminal. En amarillo las hélices *beta* y en verde las *alfa*.

Dominio de la clase **alfa + beta proteína**, con una **arquitectura** denominada de “**2-Layer Sandwich**” (**figura 8**). El **plegamiento** es *alfa*(3) + *beta*(3): comenzando por el extremo N-terminal existen tres hojas *beta* dispuestas de forma *paralela*, tras las cuales nos continúan las hélices *alfa*.

Dominio Enolasa_C-terminal (aminoácidos 146-436)

Esta región de la proteína pertenece a la clase **Alpha-Beta**, con una arquitectura en **barril**, concretamente de topología **TIM-Barrel**. Es la más común de todas las estructuras terciarias conocidas (aproximadamente el 10% de todas las enzimas conocidas la adoptan) {Farber & Petsko 1990 78 /id}, incluido en levadura, dónde es aabrumadoramente la 'top-fold' en el proteoma {Jansen & Gerstein 2000 40 /id}. Consiste en 8 hélices *beta* dispuestas paralela y concéntricamente, como si de un barril se tratara (**figura 9**). Este grupo de hojas *beta* son hidrofóbicas, alrededor de las cuales se disponen 8 hélices *alfa*, que se van a encargar de proteger este *núcleo* central hidrofóbico (de hecho, la cara de las hélices que contacta con el *núcleo* es también hidrofóbica).

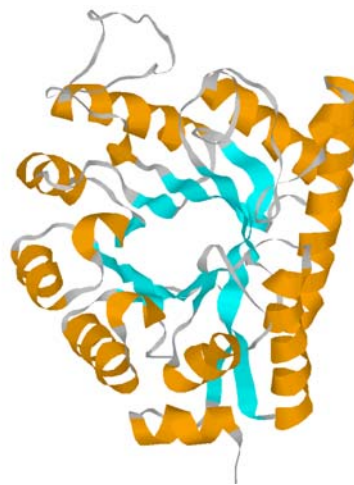


Figura 9. Dominio C-terminal. En cian las *b*-hélices. En naranja las *alfa*-hélices. las hélices *beta* y en verde las *alfa*.

A pesar de la gran cantidad de proteínas que presentan este tipo de estructura, es muy diverso el rango de funciones que son capaces de realizar. De

hecho se ha visto como la sustitución de un aminoácido puede provocar promiscuidad funcional {Schmidt, Mundorff, et al. 2003 39 /id}. Muchos son los trabajos que señalan a estos dominios para emplearlos en el diseño proteico {Vega, Lorentzen, et al. 2003 37 /id}, {Gerlt & Raushel 2003 38 /id}.



Figura 10. Detalle de los loops catalíticos. Rojo: Val153-Phe169. Azul: Ser250-Gln270

Con respecto a la **estructura cuaternaria**, dos monómeros {Zhang, Brewer, et al. 1997 30 /id} de la enolasa se asocian para constituir un homodímero que constituye una unidad cristalográfica asimétrica (**figura 6b**). Las dos subunidades son similares, aunque su superposición resulta en una pequeña distancia entre los C_{α} de 0.91Å, siendo la excepción los *bucles* catalíticos (Val153-Phe169 y Ser250-Gln270), que es de 4 Å (**figura 10**) {Zhang, Brewer, et al. 1997 30 /id}.

La enolasa requiere como cofactor al Magnesio para poder estabilizarse y llevar a cabo su reacción catalítica. Los aminoácidos que forman parte del **centro activo** de la enolasa (es decir, gozan directamente de algún papel catalítico), son His159, Glu168, Glu211, Asp246, Glu295, Asp320 y Lys345, todos ellos en el dominio C-terminal (**figura 11**). Por parte del dominio N-terminal forma parte del centro activo la Ser39, cuya mutación altera las propiedades catalíticas de la Enolasa {Brewer, Glover, et al. 1998 33 /id}, {Poyner, Larsen, et al. 2002 31 /id}. Aunque también hay que apuntar que existen otros elementos que forman parte de la reacción catalítica, como es el caso de una molécula de agua {Lebioda & Stec 1991 79 /id}, lo que conllevó a sugerir que la His373 está involucrada en el proceso catalítico {Brewer, Glover, et al. 1997 34 /id}.

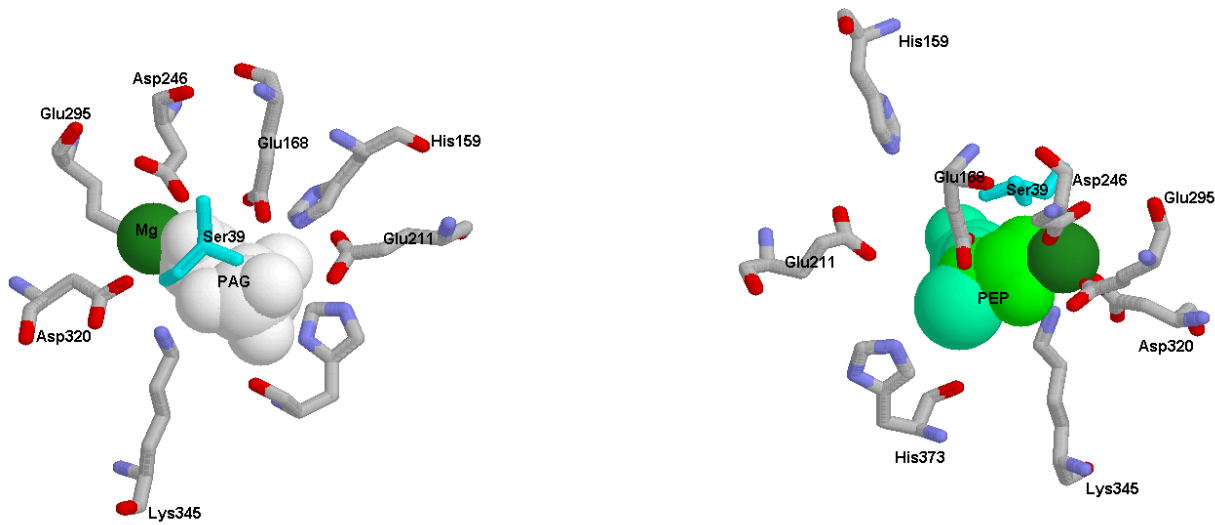


Figura 11. Representación de los aminoácidos del Centro Activo y su posición en cada uno de las dos monómeros que forman el homodímero de Eno2p y visualizado con **Rasmol**. Todos los aminoácidos han sido representados como *sticks* y coloreados en *CPK* (excepto la Ser39 en cian). El sustrato PGA (2-fosfo-D-glicerato) coloreado en blanco y el producto PEP (fosfoenolpiruvato) en verde, ambos visualizados como *Spacefill*

Modificaciones Post-Traduccionales de Eno2p

La enolasa es una proteína en la que han sido descritas varias modificaciones post-traduccionales (**figura 12**)

- Fosforilación

Los aminoácidos en los que se encuentra unido covalentemente un grupo fosfato son, por un lado la Thr378, Thr323 ó Thr325 y Ser249 ó Ser250 (no se puede asegurar el sitio exacto de la modificación puesto que en ambos casos, mediante espectrometría de masas era imposible distinguir cual de las dos treoninas y serinas era la modificada, ya que se encontraban en el mismo péptido). La identificación de la enolasa fosforilada fue mediante una aproximación sistemática {Zhou, Watts, et al. 2001 45 /id}. Por otro lado {Ficarro, McClelland, et al. 2002 46 /id} encontraron fosforilada la Ser9 también como consecuencia de una aproximación sistemática.

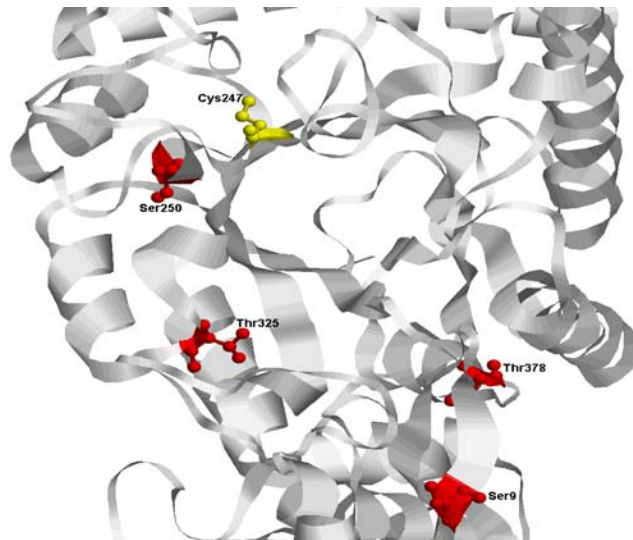


Figura 12. Eno2p representada como *ribbons*. Los aminoácidos que se fosforilan han sido representados como *ball & stick* en rojo. En amarillo la cys247

- Ubiquitinación

La adicción de ubiquitina (polipéptido de 76 aminoácidos) a un residuo de lisina específico {Hitchcock, Auld, et al. 2003 44 /id}. En este trabajo se encontró la Eno2p ubiquitinada junto con otras 210 proteínas, todas ellas aisladas en la fracción de membrana de *S.cerevisiae*, como consecuencia de una mutación en la maquinaria de degradación asociada a membrana (ERAD). Estas proteínas fueron señaladas como potenciales sustratos de esta maquinaria.

- Tiolación

Adición covalente de un grupo tilo (-SH) a la cisteína. En el caso de la enolasa es Cys247. La Eno2p se tiolizó, junto con otras enzimas glucolíticas (Tdh3, Fba1) en *S. cerevisiae* en respuesta a estrés oxidativo. La tiolación inhibe la actividad de la Eno2p y también su expresión genética (igual en Tdh3) {Shenton & Grant 2003 51 /id}.

Descripción de Proteínas Glucolíticas de interés

Piruvato quinasa (PYK1)

500 residuos que alcanzan un peso molecular de 54.533 Da. Cataliza el último paso de la glucólisis catalizando la siguiente interconversión:

ATP + piruvato = ADP + fosfoenolpiruvato.

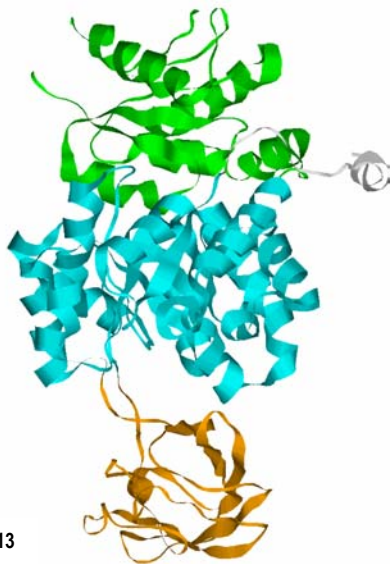


Figura 13

La estructura de la proteína en *S. cerevisiae* ha sido determinada en varias ocasiones. Nosotros elegimos 1A3X.pdb. Esta interesante proteína cuenta con tres dominios:

Dominio Central TIM BARREL (19-87, 190-355)

De las mismas características que el dominio C-terminal de la enolasa (clase *alfa-beta*, y arquitectura de barril). En la figura 13 está coloreado en cián. Entre los aminoácidos 87 a 190 sufre una "interrupción" en su estructura para formar el siguiente dominio:

Dominio Piruvato quinasa (88-189)

Estructura de clase **principalmente beta**. Se encuentra coloreado en naranja en la figura 13-

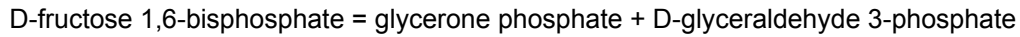
Dominio Fosfotransferasa C-terminal (356-500)

Es un dominio **fosfotransferasa** de la clase *alfa-beta*, arquitectura "**3-Layer Sandwich**" y topología **Pliege de Rossmann**.

Modificaciones Post-traduccionales: fosforilación de la Thr184 y Ser192 {Zhou, Watts, et al. 2001 45 /id}

Fructosa Bifosfato Aldolasa (FBA1 ó Aldolasa)

Proteína de 358 aa y 39.489 Da. Cataliza la siguiente interconversión:



La estructura de la proteína corresponde a *S. cerevisiae* pero utilizaremos como referencia, con la entrada 1ZEN.pdb (figura 14)

Dominio TIM-Barrel (1-287)

Presenta este dominio en la parte amino terminal de su secuencia, con las características y a vistas en la piruvato quinasa y en la enolasa.

Región PGA1_C-terminal (288-356)

Dos alfa-hélices del extremo aminoterminal que sobresalen del TIM-barrel del comienzo de la secuencia.

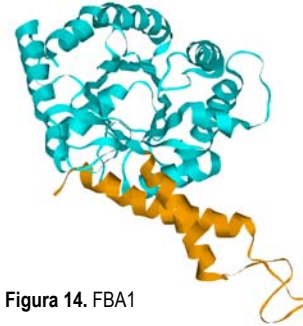


Figura 14. FBA1

Gliceraldehído 3-fosfato dehidrogenasa (TDH3 - GAPDH)

Proteína de 331 aminoácidos (aa) y peso molecular de 35.715 Da que cataliza la siguiente reacción:

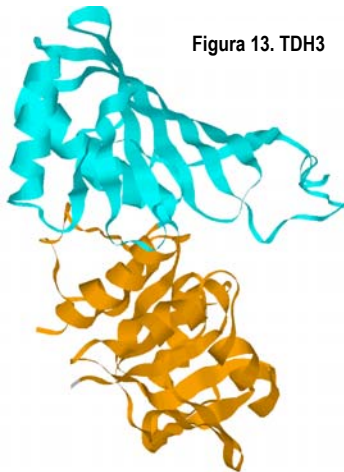
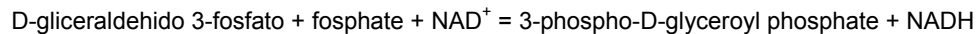


Figura 13. TDH3

La estructura de la proteína fue determinada aunque para la TDH3 de *Escherichia coli* y se encuentra almacenada en *The Protein Data Bank* con el código 1GAD {Duee, Olivier-Deyris, et al. 1996 80 /id}. Consta de los siguientes dominios (figura14):

GAPDH_N-terminal (1-147 + 314-330)

Es un dominio **Rossmann-like de unión a NAD(P)** de la clase *alfa-beta*, arquitectura "**3-Layer Sandwich**" y topología *Pliegue de Rossmann* (según CATH). En la figura 13 está coloreado en naranja.

GAPDH_C-terminal (148-313)

Dominio **holo-D-gliceraldehído-3-fosfato-dehidrogenasa**, de la clase *alfa-beta*, pero de arquitectura "**2-Layer Sándwich**". Coloreado en cian en la figura 14.

Las **modificaciones post-traduccionales** descritas para esta proteína son la **S-tiolación** de las cisteínas en respuesta a estrés oxidativo {Grant, Quinn, et al. 1999 81 /id}, {Shenton & Grant 2003 51 /id}; 4 sitios de **fosforilación**: Ser145; Thr150 ó Thr151; Thr208; Ser318 ó Thr319 {Zhou, Watts, et al. 2001 45 /id}.

3 - Fosfoglicerato kinasa (PGK1)

415aa y 44.607 Da. Cataliza la siguiente interconversión:



En cuanto a los dominios que podemos encontrar en esta proteína, basándonos en el archivo 3PGK, que pertenece a *E.coli* (**figura 14**):

Fosfotransferasa N-terminal (1-184)

Es un dominio ***fosfotransferasa (carboxil acceptor)*** que al igual que TDH3, es de la clase *alfa-beta*, arquitectura “**3-Layer Sandwich**” y topología *Pliegue de Rossmann* (según CATH).

Fosfotransferasa C-terminal (185-415)

Es igual que su extremo n-terminal pero de mayor tamaño.

Las **modificaciones post-traduccionales** que presenta esta proteína son:

- **Fosforilación:** en los residuos thr92, thr248, ser317, ser113, ser412 {Zhou, Watts, et al. 2001 45 /id}
- **Acetilación** del extremo N-terminal {Polevoda, Norbeck, et al. 1999 82 /id}



Figura 14. Pgk1.

ANÁLISIS DE SECUENCIA

Mediante el análisis de homologías tanto mediante BLAST proteína-proteína (blastp) así como PHI- and PSI-BLAST {Altschul, Madden, et al. 1997 90 /id}, los resultados no fueron indicativos.

Se puede destacar como homologías remotas con fosfopiruvatos hidratatas, muconatos racemasa, cloromuconato cicloisomerasa, muconate lactonizing enzyme, (estos muy bajos) ACT domain, muconate cicloisomerase,

Análisis mediante MAST-MEME

Con este sistema pretendemos identificar algún tipo de motivo (región altamente conservada) entre la enolasa y grupos de proteínas que hemos recopilado con las que pudiera tener alguna característica en común (tabla 2).

Interacción	Mb-Fracción		SC-MEM	SC-Levadura	HSP70	Miristoiladas	Palmitoiladas	NCP
>RN15_YEAST	>Q99325	>YJK0_YEAST	CA1463	CA1463	>SSA1	>ARF1_YEAST	>AKR1_YEAST	>IL1B_HUMAN
>MED7_YEAST	>APGB_YEAST	>MK11_YEAST	P25997	6.79470	>SSA2	>ARF2_YEAST	>GBA1_YEAST	>FGF1_HUMAN
>CCA1_YEAST	>APGD_YEAST	>AGT1_YEAST	O13430	P43076	>SSB1	>ARL1_YEAST	>PPOX_YEAST	>FGF2_HUMAN
>YNM7_YEAST	>YM34_YEAST	>MA3T_YEAST	O94039	P46598	>SSB2	>CALB_YEAST	>RAS1_YEAST	>LEG3_HUMAN
>PWP1_YEAST	>Q12142	>YJG6_YEAST	P41797	U25718		>NCS1_YEAST	>RAS2_YEAST	>TAT_HV112
>ADB6_YEAST	>Q06604	>YG2K_YEAST	CA4474	P41797		>GBA1_YEAST	>SNC1_YEAST	>TAT_HV1A2
>Q06148	>CCC1_YEAST	>SLG1_YEAST	6.5819	P87020		>PPZ1_YEAST	>SNC2_YEAST	>TAT_HV1B1
>YNA6_YEAST	>YKE6_YEAST	>glj14270689	CA4474	6.4386		>PPZ2_YEAST	>GBG_YEAST	>TAT_HV1B5
>YU1_YEAST	>D884_YEAST	>VPS1_YEAST	P46614	6.2580		>PRS4_YEAST	>TBA1_YEAST	>TAT_HV1BN
>ATC6_YEAST	>EM70_YEAST	>APG6_YEAST	6.1349	6.7003		>SIP2_YEAST	>VAC8_YEAST	>TAT_HV1BR
	>FET5_YEAST	>VP34_YEAST	P16017	P16017		>VAC8_YEAST	>CK12_YEAST	>Q9U1F9
	>GDI1_YEAST	>Q05919	O93827	Q9P842		>VP15_YEAST	>YPT1_YEAST	>HMEN_DROME
	>HOR7_YEAST	>VPS8_YEAST	P43067	P46273				>HME2_MOUSE
	>YIA2_YEAST	>WSC2_YEAST	6.1494	6.893				>HMG1_MOUSE
	>YNK6_YEAST	>WSC3_YEAST	P16017	6.4310				>TAT2_YEAST
	>Q12271	>WSC4_YEAST	Q92211	Q92211				>MDLB_PSEPU
	>IPT1_YEAST	>YBM4_YEAST	AAB96910	P16017				
	>KIN1_YEAST	>YCE9_YEAST	6.4711	Q9Y7F0				
	>KIN2_YEAST	>YCY2_YEAST	Q9P835					
	>KTR3_YEAST	>Q12025						
	>KTR6_YEAST	>Q04562						
	>LCB1_YEAST	>YEQ6_YEAST						
	>LCB2_YEAST	>YJN2_YEAST						
	>Q12246	>YN15_YEAST						
	>Q06147	>YN95_YEAST						
	>Q08989							

Estos son los resultados más significativos:

- 1.- Eno2p y proteínas implicadas en secreción no clásica.
- 2.- Motivos en común con las HSP70
- 3.- Palmitoiladas y Miristoiladas
- 4.- Fracción de Membrana

HERRAMIENTAS DE PREDICCIÓN

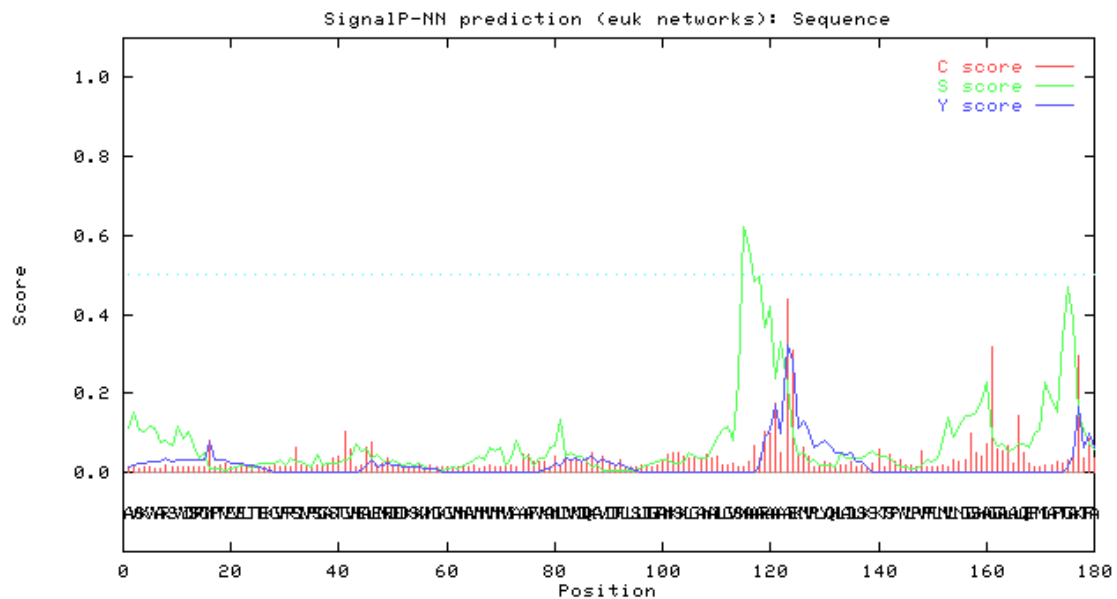
Péptido señal

The SignalP V2.0.b2 World Wide Web server predicts the presence and location of signal peptide cleavage.

Parámetros por defecto. Con solo los primeros 180 aminoácidos tal y como recomiendan....
{Nielsen, Engelbrecht, et al. 1997 83 /id}

Los resultados son:

```
>Sequence          length = 180
# Measure Position Value Cutoff signal peptide?
max. C 123      0.441 0.33 YES
max. Y 123      0.322 0.32 YES
max. S 115      0.622 0.82 NO
mean S 1-122    0.068 0.47 NO
# Most likely cleavage site between pos. 122 and 123: AAA-AE
```



Modificaciones Post-traduccionales

- Prediction of potential Myristoylation mediante el programa **NMT MyristoylCoA:Protein N-Myristoyltransferase**
(<http://mendel.imp.univie.ac.at/sat/myristate/index.html>)

No hay sitios de miristoilación

- **big-PI Predictor**

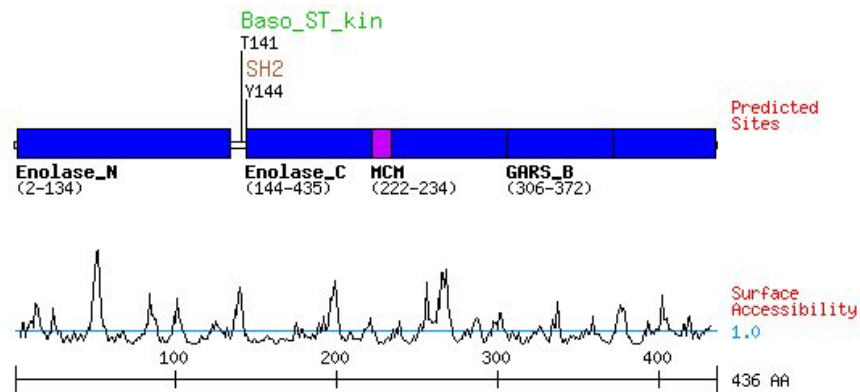
Resultados de Prediction of potential C-terminal GPI-modification sites

- **Glycosylphosphatidylinositol (GPI)**

Resultado negativo. Referencia: {Eisenhaber, Bork, et al. 1998 84 /id}

- **Motif Scan**, que se caracteriza por: Scansite identifies short protein sequence motifs that are recognized by modular signaling domains, phosphorylated by protein Ser/Thr- or Tyr-kinases or mediate specific interactions with protein or phospholipid ligands

Con una estringencia alta, los resultados han sido:



Src homology 2 group (SH2)

PLCg C-terminal SH2

Gene Card PLCG1 (1-phosphatidyl-D-myo-inositol-4,5-bisphosphate

Site Score Percentile

Sequence SA

Y144 0.1126 0.010 %

SKSKTSPYVLPVPFL 0.704

Basophilic serine/threonine kinase group (Baso_ST_kin)

PKC mu

Gene Card PRKCM (Protein kinase C, mu type (EC 2.7.1.-) (nPKC-mu)

(Protein kinase D).

Site Score Percentile

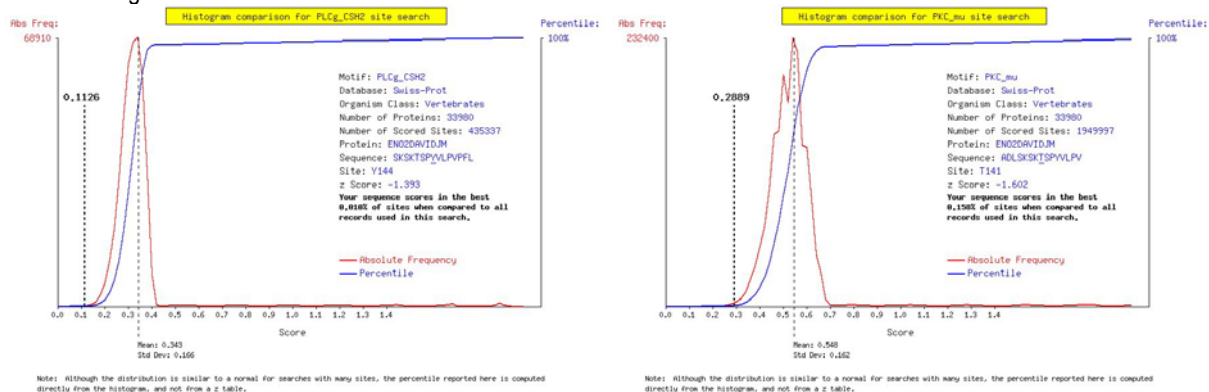
Sequence SA

T141 0.2889 0.158 %

ADLSKSKTSPYVLPV 2.878

Citación: {Obenauer, Cantley, et al. 2003 85 /id}

Las imágenes del score son estas:



El problema que son motivos de vertebrados.

- NetPhos 2.0 Server:

The NetPhos 2.0 server produces neural network predictions for serine, threonine and tyrosine phosphorylation sites in eukaryotic proteins ([Blom, Gammeltoft, et al. 1999 86 /id] Phosphorylation sites predicted: Ser: 12 Thr: 6 Tyr: 1

Serine predictions

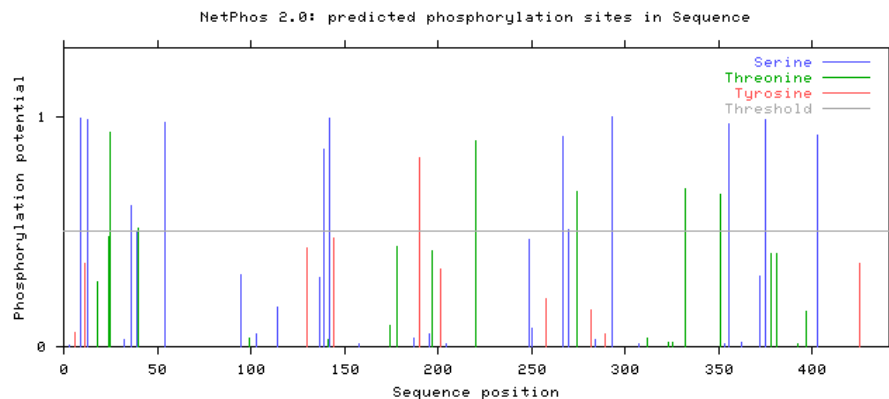
Name	Pos	Context	Score	Pred
Sequence	9	VYARSVYDS	0.995	*S*
Sequence	13	SVYDSRGNP	0.987	*S*
Sequence	36	SIVPSGAST	0.613	*S*
Sequence	54	DEDKSKWVG	0.976	*S*
Sequence	139	DLSKSKTSP	0.861	*S*
Sequence	142	KSKTSPYVL	0.995	*S*
Sequence	267	KNPESDKSK	0.914	*S*
Sequence	270	ESDKSKWLT	0.510	*S*
Sequence	293	YPIVSIEDP	0.997	*S*
Sequence	355	TLSESIKAA	0.968	*S*
Sequence	375	VSHRSGETE	0.985	*S*
Sequence	403	APARSERLA	0.918	*S*

Threonine predictions

Name	Pos	Context	Score	Pred
Sequence	25	VELTTEKGV	0.931	*T*
Sequence	40	SGASTGVHE	0.514	*T*
Sequence	220	PNIQTAEEA	0.894	*T*
Sequence	274	SKWLTGVEL	0.673	*T*
Sequence	332	ARIATAIEK	0.685	*T*
Sequence	351	NQIGTLES	0.665	*T*

Tyrosine predictions

Name	Pos	Context	Score	Pred
Sequence	190	GSEVYHNLK	0.821	*Y*

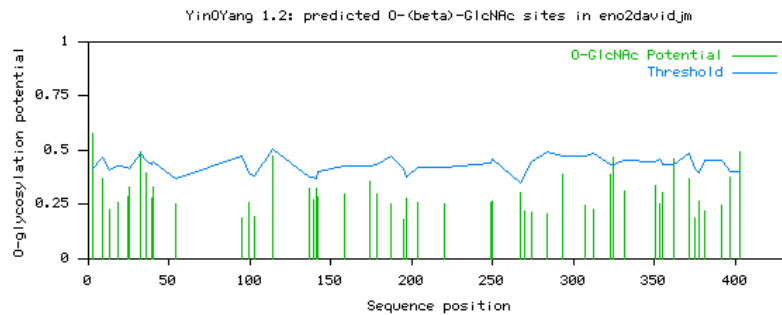


- YinOYang 1.2

Prediction Server

The YinOYang WWW server produces neural network predictions for O- β -GlcNAc attachment sites in eukaryotic protein sequences. This server can also use NetPhos, to mark possible phosphorylated sites and hence identify "Yin-Yang" sites.

SeqName	Residue	O-GlcNAc result	Potential	Thresh. (1)	Thresh. (2)
eno2davidjm	3 S	++	0.5748	0.4149	0.5096
eno2davidjm	32 S	+	0.4930	0.4861	0.6056
eno2davidjm	325 T	+	0.4698	0.4358	0.5379
eno2davidjm	362 S	+	0.4578	0.4361	0.5382
eno2davidjm	403 S	++	0.4960	0.4010	0.4909



- NetNGlyc 1.0 Prediction Results

Predice dos posibles sitios de N-glicosilación. No tiene péptido señal

SeqName	Position	Potential	Jury agreement	NGlyc result
Sequence	16 NPTV	0.7989	(9/9)	+++ WARNING: PRO-X1.
Sequence	101 NKSK	0.6727	(9/9)	++

- big-PI Predictor

None potential GPI-modification site was found.

SCAN PROSITE

> ASN_GLYCOSYLATION

Sitios de N-glicosilación

101 - 104 NKSK

> MYRISTYL

Sitios de Miristoilación

28 - 33 GVfrSI

37 - 42 GAsTV

41 - 46 GVheAL

60 - 65 GVmnAV

112 - 117 GVsmAA

156 - 161 GGshAG

157 - 162 GShaGG

161 - 166 GGalAL

202 - 207 GAsaGN

213 - 218 GVapNI

244 - 249 GLdcAS

275 - 280 GVelAD

368 - 373 GVmvSH

389 - 394 GLrtGQ

393 - 398 GQikTG

> PKC_PHOSPHO_SITE

Sitios de fosforilación de la Protein quinasa C

25 - 27 TeK

197 - 199 TkK

267 - 269 SdK

355 - 357 SiK

372 - 374 ShR

403 - 405 SeR

> CK2_PHOSPHO_SITE

Sitios de fosforilación de la Casein quinase II

9 - 12 SvyD

178 - 181 TfaE

220 - 223 TaeE

274 - 277 TgvE

293 - 296 SieD

332 - 335 TaiE

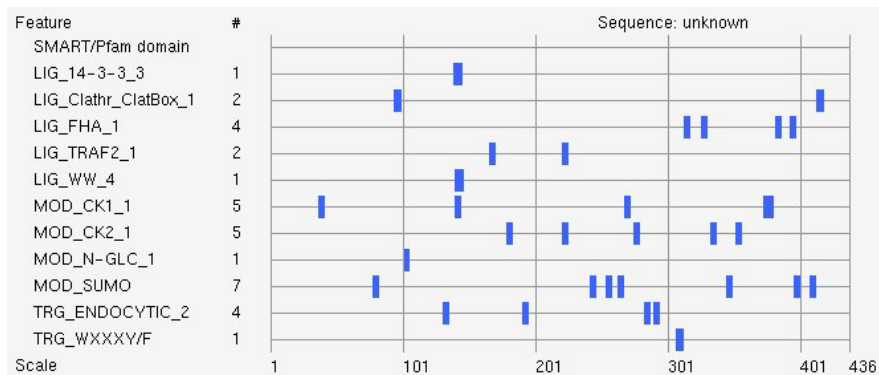
351 - 354 TlsE

- PREDICCIÓN DE SITIOS FUNCIONALES EN PROTEINAS EUCARIOTAS

The ELM server

Fuente de predicción de sitios funcionales de proteínas eucariotas. La identificación se lleva a cabo a través de patrones (expresiones regulares). {Puntervoll, Linding, et al. 2003 87 /id}

Sin especificar la localización, para *S. cerevisiae*, los resultados para Eno2p son:



ELM-NAME	Fragmento	Posición	ELM-Descripción	Compartimento Celular
LIG_14-3-3_3	KSKTSP	138-143	Consensus derived from natural interactors which do not exactly match the mode1 and mode2 ligands	nucleus, cytoplasm, plasma membrane
LIG_Clathr_ClathBox_1	LLSLD-LLRIE	93-97 412-416	Clathrin box motif found on cargo adaptor proteins, it interacts with the beta propeller structure located at the N-terminus of Clathrin heavy chain.	cytoplasm, Golgi apparatus, cytoskeleton, clathrin-coated endocytic vesicle, Golgi trans-face
LIG_FHA_1	TAGI- TNPA- TFIA-TGQI	312-315, 325-328 381-384, 392-395	FHA domain interaction motif 1, threonine phosphorylation is required	nucleus, cytoplasm, plasma membrane

LIG_WW_4	SKTSPY	139-144	Class IV WW domains interaction motif; phosphorylation-dependent interaction.	nucleus,cytoplasm
MOD_CK1_1	SGAS, SKTS SDKS, SHRS SGET	36-39, 139-142 267-270, 372-375 375-378	CK1 phosphorylation site	nucleus,cytoplasm
MOD_CK2_1	TFAE, TAEE TGVE, TAIE TLSE	178-181, 220-223 274-277, 332-335 351-354	CK2 phosphorylation site	nucleus,cytoplasm
MOD_N-GLC_1	NKS	101-103	Generic motif for N-glycosylation. Shakin-Eshleman et al. showed that Trp, Asp, and Glu are uncommon before the Ser/Thr position. Efficient glycosylation usually occurs when ~60 residues or more separate the glycosylation acceptor site from the C-terminus	extracellular, Golgi apparatus, endoplasmic reticulum
MOD_SUMO	VKAN, VKIG FKDG, FKNP LKVN, IKTG AKLN	77-80, 241-244 253-256, 262-265 344-347, 395-398 407-410	Motif recognised for modification by SUMO-1	nucleus,PML body
TRG_ENDOCYTIC_2	YQHL, YHNL YHSL, YPIV	130-133, 190-193 282-285, 289-292	Tyrosine-based sorting signal responsible for the interaction with mu subunit of AP (Adaptor Protein) complex	cytoplasm,plasma membrane, clathrin-coated endocytic vesicle

HERRAMIENTAS DE PREDICCIÓN DE LOCALIZACIÓN

PSORT

Para secuencias eucariotas {Horton & Nakai 1997 89 /id}

k = 9/23

65.2 %: cytoplasmic
13.0 %: mitochondrial
13.0 %: nuclear
4.3 %: plasma membrane
4.3 %: peroxisomal

TargetP v1.01 predicts

Localización subcelular

Predice la localización subcelular en proteínas eucariotas, asignándola en función de péptido de señal mitocondrial (mTP) péptido señal de la vía de secreción (SP) u otras localizaciones {Emanuelsson, Nielsen, et al. 2000 88 /id}

#	Name	Length	mTP	SP	other	Loc.	RC
#-----	Sequence	436	0.318	0.040	0.723	_	3
#-----							


OTRAS PREDICCIONES


SUMOPLOT™ (<http://www.abgent.com/sumoplot.html>)

Puede ayudar a explicar pesos moleculares más elevados de los esperados en geles SDS debido a la unión de proteínas SUMO (*Small Ubiquitin-related Modifier* de 11kDa) a determinados puntos de la secuencia.


```

1  AVSKVYARSV YDSRGNPTVE VELTTEKGVF RSIVPSGAST GVHEALEMRD
51  EDKSKWMGKG VMNAVNNVNN VIAAAFVKAN LDVKDQKAVD DFLLSLDGTG
101 NKSCLGANAI LGVSMMAARA AAAEKNVPLY QHLADLSKSK TSPYVLPVPF
151 LNVLNGGSHA GGALALQEFM IAPTGAKTFA EAMRIGSEVY HNLKSLTKKR
201 YGASAGNVGD EGGVAPNIQT AEEALDLIVD AIKAAGHDGK VKIGLDCASS
251 EFFKDGKYDL DFKNPESDKS KWLTVGVELAD MYHSLMKRYP IVSIEDPFAE
301 DDWEAWSHFF KTAGIQIVAD DLTVINPARI ATAIEKKAAD ALLLKVNQIG
351 TLSESIKAAQ DSFAANWGVN VSHRSGETED TFIADLVVGL RTGQIKTGAP
401 ARSERLAKLN QLLRIEEELG DKAVYAGENF HHGDKL
    
```

 Motifs with high probability

 Motifs with low probability

No.	Pos.	Group	Score
1	K396	LRTGQ IKTG APARS	0.7722
2	K242	GHDGK VKIG LDCAS	0.7556
3	K263	KYDLD FKNP ESDKS	0.7444

No.	Pos.	Group	Score
4	K254	ASSEF FKDG KYDLD	0.6778
5	K257	EFFKD GKYD LDFKN	0.6722
6	K177	IAPTG AKTF AEAMR	0.4444

Análisis Tridimensional

Estudio de la superficie celular y sus características. Para ello nos centramos en las proteínas glucolíticas descritas anteriormente. Para extraer los mapas superficiales empleamos el visualizador *Dino*, *MSMS* (*msms.i86Linux2.2.5.5.staticgcc*) con el que obtenemos los datos de superficie, un script en perl que convierte archivos pdb en xyzr (pdb2xyzr.pl), todo ello bajo entorno Linux.

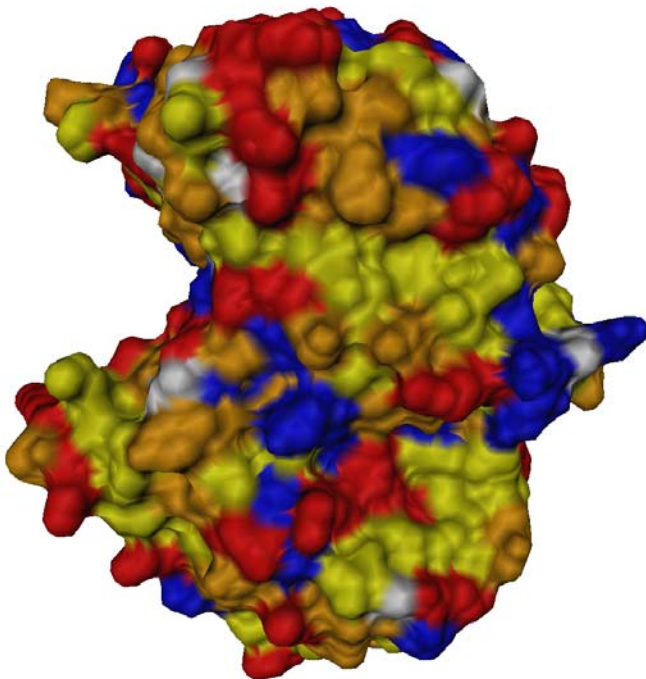
La gama de colores empleada fue:

- Rojo: aminoácidos básicos (carga positiva)
- Amarillo: aminoácidos hidrofóbicos
- Naranja: aminoácidos polares no cargados
- Azul: aminoácidos ácidos (carga negativa)

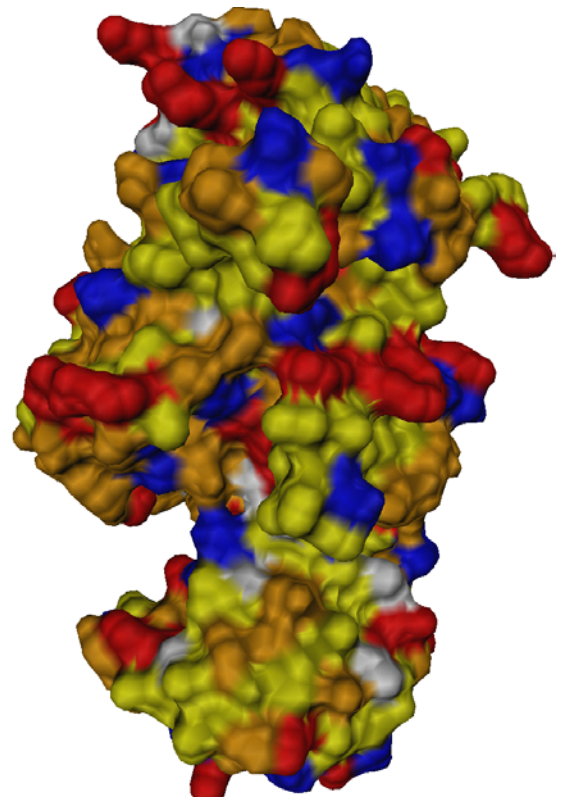
Estas fueron las figuras obtenidas

Vistas globales

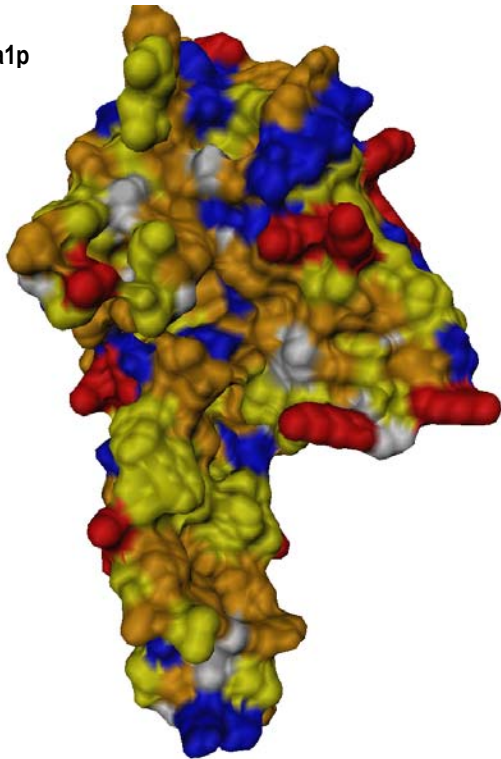
Eno2p



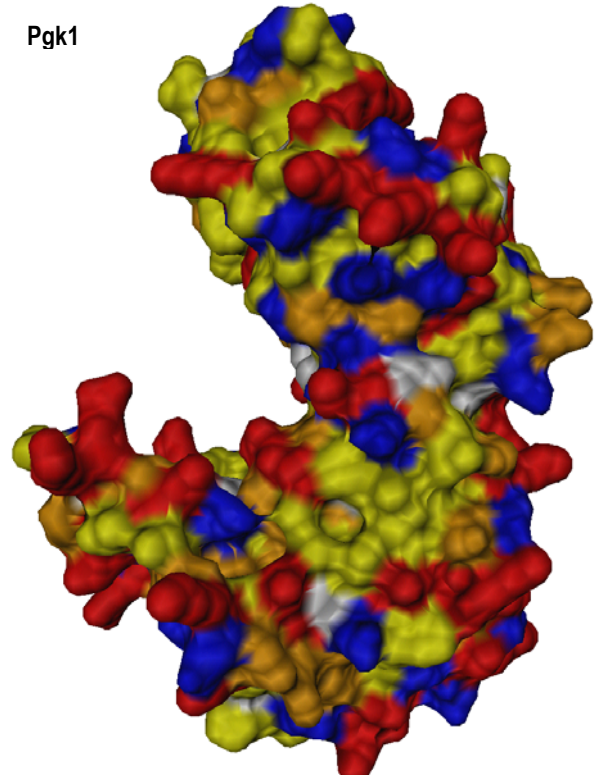
Pyk1p



Fba1p



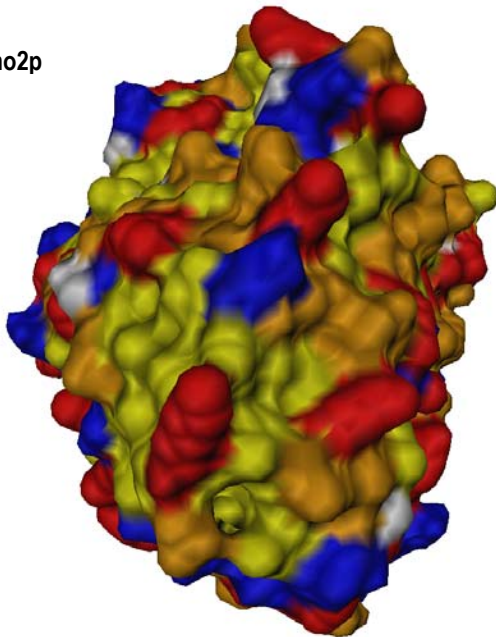
Pgk1



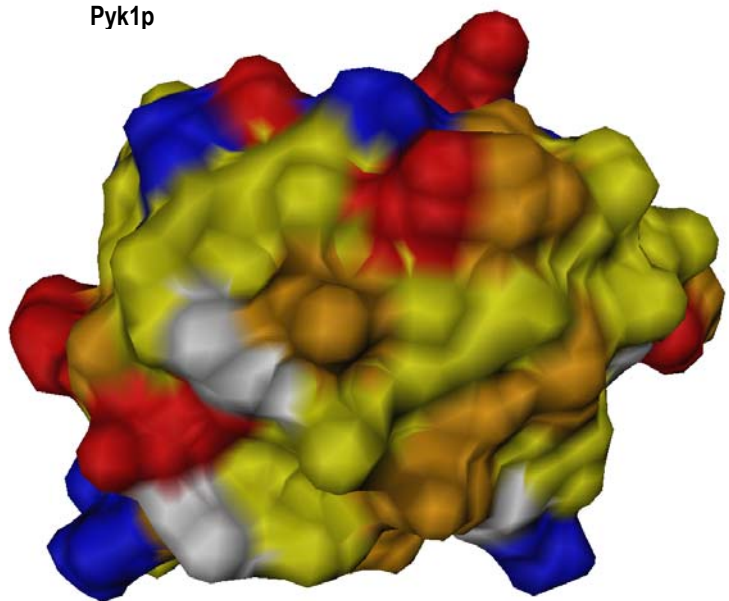
|||||Falta la Tdh3, por diosssssssss!!!!!!!!!!!!

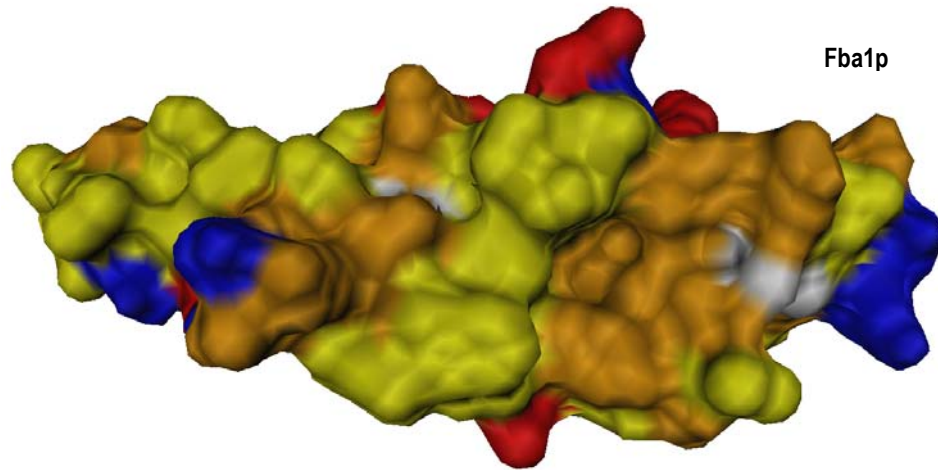
Vistas de dominios más hidrofóbicos, con más detalle

Eno2p

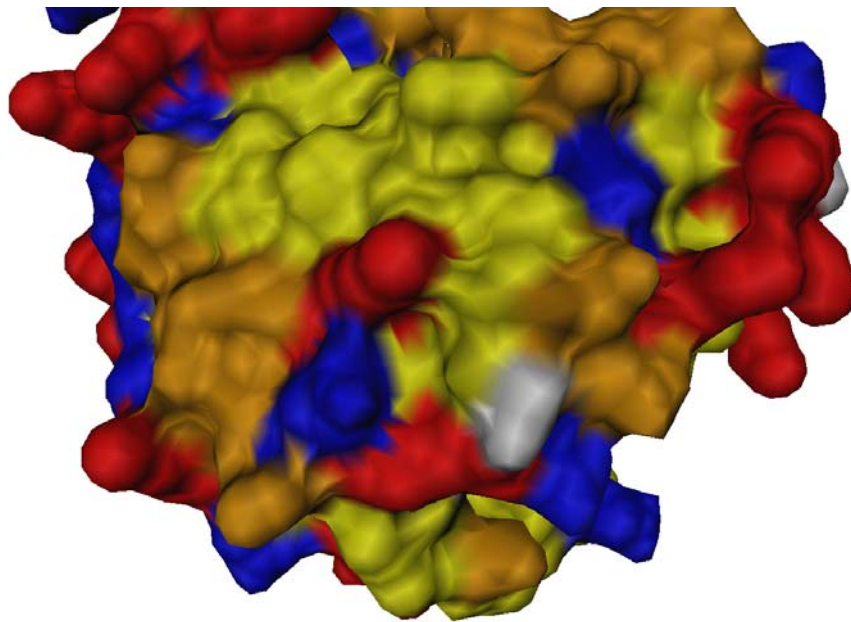


Pyk1p



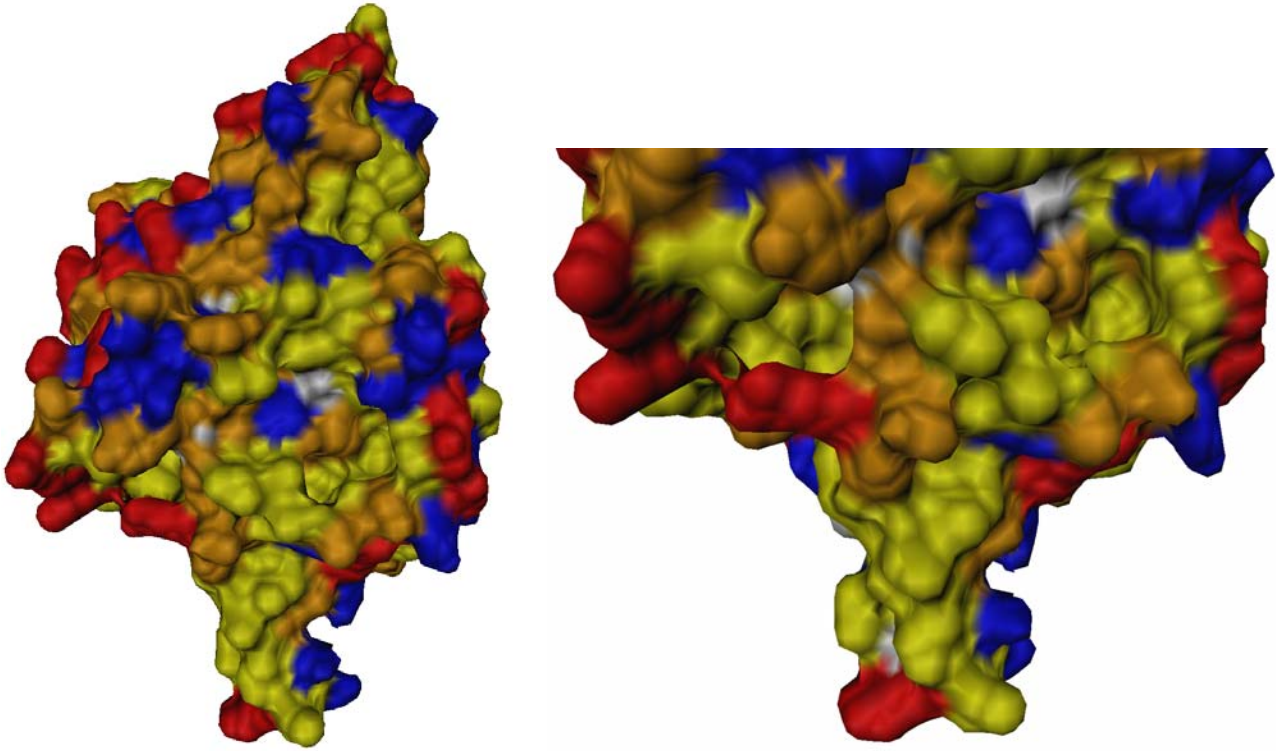


Fba1p



Pgc1p

|||||||la TDH3 de cerca!!!!!!!!!!!!

Ejemplo al azar de proteína asociada a membrana: mandelato deshidrogenasa (de cerca y de lejos)

Concha, ya está casi todo. Falta un par de cosillas. Mañana escribo la discusión. No te confundas, no aseguraré nada. Ni muchos menos voy a comparar con la mandelato deshidrogenasa. Pero tenía mucha curiosidad por saber como eran estas proteínas y fíjate tú que seleccioné la que tiene un dominio TIM-BARREL y está asociada a membrana (la parte ampliada, que es una sustitución).

En fin, el próximo capítulo, mañana.

davidjm