



INFORME FINAL DE INVESTIGACIÓN

“ANÁLISIS FILOGENÉTICO Y DE DOCKING MOLECULAR DE LA PROTEÍNA CONGLUTINA γ DE *Lupinus albus* Y SU INTERACCIÓN CON EL RECEPTOR DE INSULINA HUMANA”

LIMA–PERÚ

2016

ANÁLISIS FILOGENÉTICO Y DE DOCKING MOLECULAR DE LA PROTEÍNA CONGLUTINA γ DE *Lupinus albus* Y SU INTERACCIÓN CON EL RECEPTOR DE INSULINA HUMANA

RESUMEN

Objetivo: Realizar un análisis de las relaciones filogenéticas de la conglutina gamma de *L. albus* con otras proteínas vegetales y evaluar mediante docking molecular su interacción con moléculas relacionadas al control de la glicemia. **Materiales y métodos:** Se evaluaron las secuencias de aminoácidos del GenBank mediante BLAST, del cual se obtuvieron similitudes del 20% a 80%. Se usó la aplicación ModelGenerator_v_85 para encontrar el mejor modelo de sustitución y el servidor PHYML para generar un árbol de máxima verosimilitud. Se modeló la proteína en su forma monomérica con Phyre2. Se descargaron del PDB estructuras de insulina humana y relacionadas, se refinaron con 3DRefine y se prepararon con Chimera 1.10.2 para *docking*. Los análisis de *docking* se realizaron en ClusPro y se compararon las energías mínimas de los 3 mejores modelos elegidos por CONSRANK. **Resultados:** Las secuencias de la proteína con el 20-80% de aminoácidos conservados con la conglutina γ fueron obtenidas de GenBank. Los modelos de la substitución y la predicción de la estructura tridimensional fueron obtenidos usando ModelGenerator y Phyre2, respectivamente; usando la forma monómero de conglutina γ . El árbol filogenético de máxima verosimilitud fue construida usando PHYML, lo que demuestra que la secuencia de aminoácidos de la conglutina γ se encuentra agrupado más cercanamente a las globulinas básicas 7S de *Glicine max* y *Morus notabilis*. Para el análisis de docking molecular de la conglutina γ , se descargaron los archivos PDB de la insulina humana y las diferentes moléculas relacionadas (insulina unida al factor de crecimiento I y II, receptor humano de la insulina), refinado con 3DRefine y preparado usando Chimera. El acoplamiento molecular fue realizado en ClusPro y la energía mínima de los mejores modelos seleccionados por CONSRANK fue comparada. Esta simulación demuestra los valores de una predicción de unión altamente notable solamente con el receptor humano de la insulina. **Conclusiones:** La conglutina γ de *Lupinus albus* es una proteína con una baja conservación de aminoácidos fuera del grupo de los Lupinos, su forma monomérica presenta una alta probabilidad de la interacción con el receptor de insulina humana. Esto es una evidencia adicional para su acción de la insulino-mimético que refuerza su uso en el cotratamiento con metformina. **Palabras clave:** modelamiento, conglutina, hiperglucemia, docking molecular, máxima verosimilitud.

ABSTRACT

Objective: To perform and reports a phylogenetic study of conglutin γ from *L. albus* and evaluates its interaction with known molecules that controls glycemia using molecular docking. **Material and methods:** We assessed the amino acid sequences of GenBank by BLAST, of which similarities were obtained from 20% to 80%. The ModelGenerator_v_85 application was used to find the best replacement model and the PHYML server to generate a maximum likelihood tree. The protein was modeled in its monomeric form with Phyre2. Were downloaded from the PDB human insulin and related structures, refined with 3DRefine and prepared with Chimera 1.10.2 to docking. The analyses were carried out in ClusPro and the minimum energies of the 3 best models chosen by CONSRANK were compared. **Results:** Protein sequences with 20-80% amino acid conservation with conglutin γ were obtained from GenBank. Substitution models and tridimensional structure prediction were obtained using ModelGenerator and Phyre2, using monomeric form of conglutin γ . Maximum likelihood phylogeny was constructed using PHYML, which shows that amino acid sequence of conglutin γ is grouped with 7S basic globulins from *Glicine max* and *Morus notabilis*. For molecular docking analysis with conglutin γ , PDB files of human insulin and different related molecules (insulin-like growth factor I and II, human insulin receptor) were downloaded, refined with 3DRefine and prepared using Chimera. Molecular docking was performed on ClusPro and minimum energy of the best models selected by CONSRANK was compared. This simulation shows a highly notable binding prediction value only with human insulin receptor. **Conclusions:** Conglutin γ from *Lupinus albus* is a protein with low conservation of amino acids outside the lupin group, and its monomeric form shows a high probability of interaction with human insulin receptor. This is an extra evidence for its insulin-mimetic action which re-inforce its use in co-treatment with metformin.

Keywords: Modeling, Conglutin, hyperglycemia, molecular docking, maximum likelihood.

INTRODUCCIÓN

Actualmente una de las estrategias fundamentales para las terapias antidiabéticas son los cambios de hábitos alimenticios y entre ellas el consumo de las leguminosas juegan un papel importante en el tratamiento de la diabetes (1,2). Estas terapias dietéticas se basan en que el consumo de tres o más porciones de legumbres reducen entre un 20% a 35% el riesgo de padecer diabetes (3). El consumo de alimentos ricos en almidón, has demostrado que reducen la elevación de los niveles de glucosa postprandial (4).

El garbanzo es una de las legumbres que disminuye el índice glucémico y la respuesta insulinémica, produciendo un mejor control de la diabetes tipo 2, de manera que se dé de forma menos severa (2). El modo de acción de esta proteína es probable que se deba a la activación de algunas vías de señalización celular de la insulina, ejerciendo de esta manera un efecto mimético al de esta hormona (7).

En estos últimos años muchas especies del genero *Lupinus* han demostrado el potencial como agentes hipoglucémicos, estas conclusiones se obtuvieron en base al consumo de derivados de cereales, como el pan y la pasta, que se suplementaban con harina de semillas de altramuz (5). La responsable de estos efectos beneficiosos es una proteína minoritaria, llamada conglutina gamma, el cual procedía del altramuz blanco (*L. albus*) (6).

Estudios in vivo e in vitro han demostrados que tras su administración oral, se reducen de forma dosis-dependiente los niveles de glucosa e insulina en sangre (7,9). La administración oral de conglutina γ en ratas con elevados niveles de glucosa, produce el efecto de una reducción estadísticamente significativa de la glucemia respecto a la de la metformina, fármaco comúnmente utilizado para reducir los niveles de la glucosa. Estos hallazgos son la primera evidencia molecular del posible uso de una proteína de leguminosa en el control de la glucemia (9).

La conglutina gamma es una proteína con actividad antihiperlicemiante en ratas y cultivos celulares y es adecuado para el tratamiento de la diabetes tipo II. Se une a la insulina humana a ciertos pHs, pero no hay información sobre su interacción con otros objetivos médicos importantes.

Por ello, el presente trabajo reporta un estudio filogenético de la conglutina γ de *L. albus* con otras proteínas vegetales y evalúa su interacción con moléculas conocidas que controlan la glucemia utilizando acoplamiento molecular; así como de esta manera fundamenta su accionar insulino-mimético, lo que refuerza su utilización como tratamiento complementario a la metformina. Además, sería relevante evaluar la interacción de la globulina 7S de otras especies con las que no se ha incursionado en esta área, como las moras.

MATERIALES Y MÉTODOS

Evaluación de las secuencias de aminoácidos

Se evaluaron las secuencias primarias de aminoácidos correspondiente a la conglutina γ del *Lupinus albus* del GenBank mediante BLAST, del cual se obtuvieron similitudes del 20% a 80%. Dichas secuencias fueron utilizadas para el posterior análisis bioinformático, filogenético y obtención de la información estructural de dicha proteína.

Obtención de un árbol de máxima verosimilitud

Para determinar la relación filogenética, se usó la aplicación ModelGenerator_v_85 (10) con la finalidad de encontrar el mejor modelo de sustitución y el servidor PHYML (11) para generar un árbol de máxima verosimilitud.

Modelamiento por threading de la conglutina γ

Como la homología era remota en la *conglutina* γ , se logró modelar la proteína en su forma monomérica con Phyre2 (12).

Análisis de docking molecular

Se descargaron estructuras de insulina humana y relacionadas del PDB, se refinaron los modelos con 3DRefine (13) y se prepararon para *docking* con Chimera 1.10.2. (14) Los análisis de *docking* se realizaron en ClusPro (15) y se compararon las energías mínimas de los 3 mejores modelos elegidos por CONSRANK (16).

RESULTADOS

El análisis de las secuencias primarias de las proteínas con el 20-80% de aminoácidos conservados con la conglutina γ de *Lupinus albus* fueron obtenidas de GenBank. Los modelos de la substitución y la predicción de la estructura tridimensional fueron obtenidos usando ModelGenerator y Phyre2, respectivamente; usando la forma monómero de conglutina γ (Fig. 1).

Con respecto al árbol filogenético de máxima verosimilitud fue construida usando PHYML, lo que demuestra que la secuencia de aminoácidos de la conglutina γ se encuentra agrupado más cercanamente a las globulinas básicas 7S de *Glicine max* y *Morus notabilis* (Fig. 2).

Finalmente para el análisis de docking molecular de la conglutina γ , se descargaron los archivos PDB de la insulina humana y las diferentes moléculas relacionadas (insulina unida al factor de crecimiento I y II, receptor humano de la insulina), refinado con 3DRefine y preparado usando Chimera. El docking molecular fue realizado en ClusPro (Fig.3) y la energía mínima de los mejores modelos seleccionados por CONSRANK fue comparada (Fig. 4). Esta simulación demuestra los valores de una predicción de unión altamente notable solamente con el receptor humano de la insulina. Más detalles del docking molecular (Fig.5).

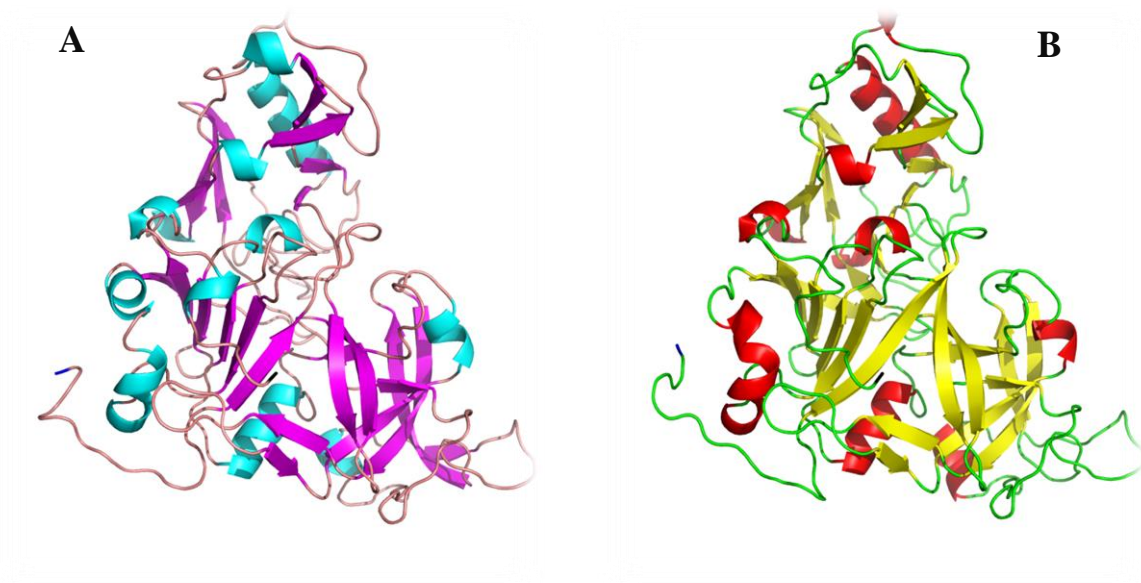


Fig. 1. Modelado tridimensional por threading del monómero de conglutina γ de *Lupinus albus* generada por Phyre2. A) Modelo (M1) de la secuencia CAC16394.1 (GenBank). B) Modelo (M2) de la secuencia CAC17729 (GenBank).² Estos modelos fueron descargados, se refinaron sus estructuras con 3D Refine y se prepararon para docking con Chimera.

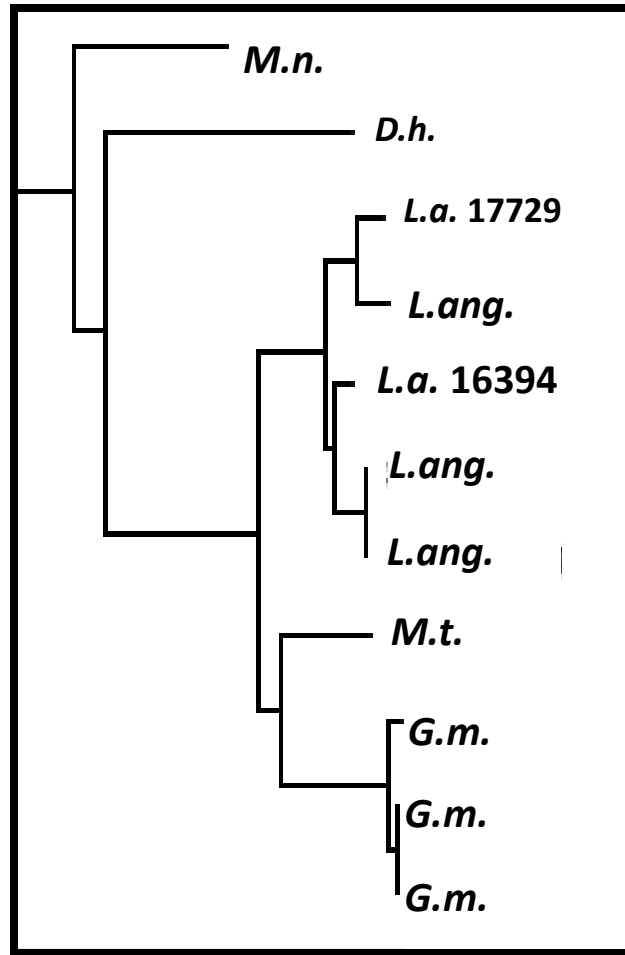


Fig. 2. Árbol de la filogenia ML de *Lupinus albus*. (*L.a.*), *Lupinus angustifolius* (*L.ang.*), *Morus notabilis* (*M.n.*), *Dorcoceras hygrometricum* (*D.h.*), *Medicago trunculata* (*M. t*) y *Glycine max* (*G.m.*). Las secuencias CAC17729 y CAC16394 de *L.ang* difieren respecto a las demás y se caracteriza en un grupo aparte.

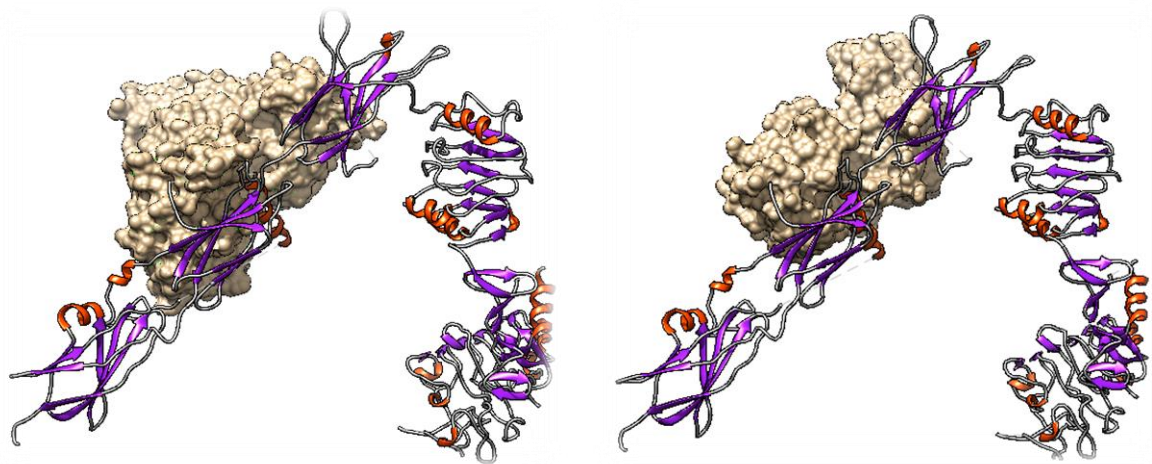


Fig. 3.

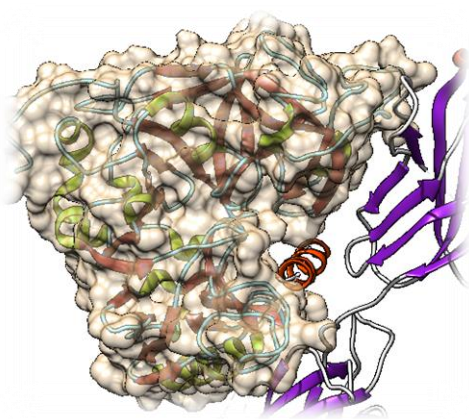
Interacción de M1 y M2 con IR. (Mejores resultados de docking seleccionados por CONStrank)

	PDB	Res(Å)	vs. 16394	vs. 17729
Insulin	1BEN	1.4	-755.6	-724.5
IGF 1	1GZR	2	-673.3	-852.9
IGF 2	3KR3	2.2	-776.3	-792.7
IR	4ZXB	3.3	-959.6	-992
SA	5IJF	2.65	-784.4	-845

Fig. 4.

Resolución y valor medio de la energía más baja de los tres mejores modelos seleccionados por CONStrank de los ensayos de docking realizados en Cluspro. La albúmina de suero fue utilizada como control negativo. Solamente con el receptor de la insulina el resultado era considerablemente más bajo.

M1



M2

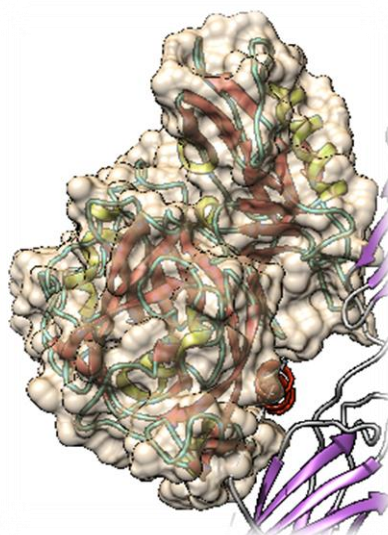


Fig. 5.

Detalle de la zona de contacto en las mejores simulaciones de docking de los modelos y la estructura del ectodominio IR. (M1 izquierda, M2 derecha) Las estructuras secundarias entre Phe 475 y Arg 634 de IR pueden estar implicadas en la interacción

DISCUSIÓN

La conglutina γ de *Lupinus albus* es una proteína con una secuencia poco conservada fuera de los lupinos y en su forma monomérica presenta una alta probabilidad de interacción con el receptor de insulina humano, lo que fundamenta su accionar insulino-mimético (6). Se observa que las secuencias de conglutina se agrupan más cercanamente a las globulinas básicas 7S de *Glicine max* (soya) y otras pocas especies. La secuencia más basal en el árbol fue la del inhibidor de xilanasas de *Triticum aestivum* (trigo). El *docking* molecular con las distintas estructuras (insulina, factor de crecimiento insulínico I y II, receptor de insulina y albúmina sérica humana, como control negativo) mostró un valor notablemente mayor en los scores sólo para el receptor de insulina humano (7, 8). En cuanto a la filogenia de máxima verosimilitud que las secuencias de conglutina se agrupan cercanamente a las globulinas básicas 7S de *Glicine max* (soya) y *Morus notabilis*, siendo el Inhibidor de xilanasas del trigo la proteína más basal. El *docking* molecular con las distintas estructuras (Insulina, Factor de Crecimiento Insulínico I y II, Receptor de Insulina Humano y Albúmina Sérica Humana, como control negativo) mostró un valor de unión notablemente mayor solo para el receptor de insulina humano (9).

CONCLUSIONES

La conglutina γ de *Lupinus albus* es una proteína con una baja conservación de aminoácidos fuera del grupo de los Lupinos, y en su forma de monómero presenta una alta probabilidad de la interacción con el receptor de insulina humana. Esto es una evidencia adicional para su acción de la insulino-mimético que refuerza su uso en el co-tratamiento con metformina.

LITERATURA CITADA

- 1.- Campos-Vega, R., Loarca-Piña, G. y Oomah, B. D. (2010). Minor components of pulses and their potential impact on human health. *Food Research International*, 43 (2), pp. 461-482. <http://doi.org/10.1016/j.foodres.2009.09.004>.
- 2.- LOVATI, M.R. et al. *British Journal of Nutrition*. 2012; vol. 107, no 01, p. 67-73.
- 3.- Kirpitch, A. R. y Maryniuk, M. D. (2011). The 3 R's of glycemic index: recommendations, research, and the real world. *Clinical Diabetes*, 29, pp. 155-159.
- 4.- Venn, B. J. y Mann, J. I. (2004). Cereal grains, legumes and diabetes. *European Journal of Clinical Nutrition*, 58 (11), pp. 1443-1461.
- 5.- Capraro, J., Magni, C., Scarafoni, A., Caramanico, R., Rossi, F., Morlacchini, M. y Duranti, M. Pasta supplemented with isolated lupin protein fractions reduces body weight gain and food intake of rats and decreases plasma glucose concentration upon glucose overload trial. *Food y Function*.2014;5 (2), pp. 375-380.
- 6.- Delgado-Andrade, C., Olías, R., Jiménez-López, J. C., & Clemente, A. (2016). Aspectos de las legumbres nutricionales y beneficiosos para la salud humana. *arbor*, 192(779), 313.
- 7.- Bertoglio, J. C., Calvo, M. A., Hancke, J. L., Burgos, R. A., Riva, A., Morazzoni, P. y Duranti, M. Hypoglycemic effect of lupin seed γ -conglutin in experimental animals and healthy human subjects. *Fitoterapia*.2011;82(7),pp.933-938.
- 8.- Terruzzi, I., Senesi, P., Magni, C., Montesano, A., Scarafoni, A., Luzi, L. y Duranti, M. (2011). Insulin-mimetic action of conglutin- γ , a lupin seed protein, in mouse myoblasts. *Nutrition, Metabolism and Cardiovascular Diseases*, 21 (3), pp. 197-205. <http://doi.org/10.1016/j.numecd.2009.09.004>.
- 9.- MAGNI, C. et al. *The Journal of nutritional biochemistry*. 2004; vol. 15, no 11, p. 646-650.
- 10.- Thomas M Keane, Christopher J Creevey , Melissa M Pentony, Thomas J Naughton and James O McInerney. Assessment of methods for amino acid matrix selection and their

use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evolutionary Biology*.2016; 6:29

11.- Guindon, S., Lethiec, F., Duroux, P., & Gascuel, O. PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic acids research*.2005; 33(suppl_2), W557-W559.

12.- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols*.2015; 10(6), 845-858.

13.- Bhattacharya, D., Nowotny, J., Cao, R., Cheng, J. 3Drefine: an interactive web server for efficient protein structure refinement. *Nucleic Acids Research*. 2016.Web Server Issue. doi: 10.1093/nar/gkw336.

14.- Chen, J. E., Huang, C. C., & Ferrin, T. E. RRDistMaps: a UCSF Chimera tool for viewing and comparing protein distance maps. *Bioinformatics*.2014; 31(9), 1484-1486.

15.- Comeau, S. R., Gatchell, D. W., Vajda, S., & Camacho, C. J. ClusPro: a fully automated algorithm for protein–protein docking. *Nucleic acids research*.2004; 32(suppl_2), W96-W99.

16.- E Chermak, A Petta, L Serra, A Vangone, V Scarano, L Cavallo, and R Oliva, CONSRANK: a server for the analysis, comparison and ranking of docking models based on inter-residue contacts. *Bioinformatics*. 2015; in press