

Recorrido por Hv1: un canal selectivo de protones

Victoria Pinto

Cátedra de Bioinformática, Área de Biotecnología y Biología Molecular, Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, La Plata, Argentina.

RESUMEN

A lo largo del cuatrimestre se fueron introduciendo diferentes métodos y técnicas bioinformáticas para el estudio de proteínas. Este análisis se propone estudiar de forma secuencial, estructural y funcional al canal voltaje operado en humanos, Hv1.

Para esto, se realizó una búsqueda de homología, fijando los parámetros necesarios. Luego, con alguna de estas secuencias encontradas, se hizo un alineamiento múltiple para corroborar la importancia de Asp112 en la selectividad de los protones.

También se realizó la predicción de estructura secundaria y se modeló la estructura terciaria. Se construyó, además, un árbol filogenético utilizando secuencias de proteínas homólogas de distintos mamíferos que expresan el canal. Por último, se hizo un análisis de la función del canal.

PALABRAS CLAVE: canal de protones voltaje operado; mamíferos; bioinformática

INTRODUCCIÓN

En este trabajo presento un estudio bioinformático sobre la proteína Hv1, un canal de protones voltaje operado en humanos. El gen HVCN1 que codifica para la proteína, se encuentra en el cromosoma 12 de humanos y se expresa en varios tejidos del organismo. Existen cuatro isoformas, de las cuales se trabajará con la isoforma 1, cuya longitud es de 273 aminoácidos (Uniprot ID: Q96D96).

En condiciones fisiológicas normales el pH extracelular es ligeramente alcalino (7,3-7,4); y el pH citosólico, menor que dicho valor (7,2). Dos propiedades celulares dan base a la tendencia del citosol a acidificarse: en primer lugar el potencial de membrana con el interior celular negativo respecto al extracelular, brinda una fuerza impulsora importante para la entrada de protones y la salida de bases negativamente cargadas, como el HCO_3^- . En segundo lugar, diversas reacciones metabólicas vitales para la célula, como la producción de ATP por la glicólisis en el citoplasma o la fosforilación oxidativa en la mitocondria, generan H^+ como productos de la reacción, que aumentan cuando se incrementa la actividad celular. La acumulación gradual de H^+ es compensada mediante su extrusión continua a través de la membrana plasmática por distintos mecanismos.

El canal Hv1 se activa con la despolarización de la membrana celular, mediando una corriente saliente de cargas positivas que tiende a hiperpolarizar la membrana. Además, se encuentra en su conformación abierta sólo cuando el gradiente electroquímico es favorable a la salida espontánea de protones, por lo cual se puede considerar al canal como un estricto extrusor pasivo de H^+ . Esta fuerte dependencia con el pH indica que su actividad será mayor al aumentar el pH extracelular o disminuir el pH intracelular.

Con respecto a la inactivación, evidencias experimentales indican que los canales Hv1 no presentan en esta situación una dependencia con el voltaje. Sin embargo, cuando se evocan corrientes de H^+ demasiado prolongadas (~20 seg.) Se puede observar un decaimiento de la corriente que se asemeja al proceso de

inactivación. Se vio que la misma corriente mantenida en el tiempo es la que produce un cambio en la concentración de H⁺, modificando su gradiente. Esto reduce la fuerza impulsora, produciendo una disminución de la magnitud de la corriente cada vez mayor, de forma similar a lo que ocurre cuando los canales presentan un estado inactivo.

En contraste con los típicos canales de cationes dependientes de voltaje, Hv1 contiene solo el dominio de detección de voltaje (VSD) y carece del dominio de poro central (esencial para otros canales de cationes). La secuencia de aminoácidos de Hv1 está relacionada con el VSD de otros canales de cationes. Otra diferencia es que Hv1 es un dímero, donde cada subunidad contiene tanto el sensor de voltaje como la vía de los protones. Se requieren regiones N- y C-terminales del canal para la dimerización, logrando una activación cooperativa (figura 1). Evidencias muestran que el canal Hv1 también podría ser funcional como monómero. Se postula que los dímeros se mantienen por interacciones Coiled-Coil entre los dominios intracelulares C-terminal, por lo que si se remueve o mutan los dominios C-terminal, podría obtenerse la expresión funcional del monómero (Asuaje, 2018).

Los canales Hv1 ofrecen a la célula que lo expresa, una vía de pasaje de H⁺ que permite la regulación del pH_i sin gasto de energía metabólica. En este sentido, se ha comprobado que la función del canal es muy importante en la recuperación del pH_i de diversos tipos celulares luego de ser sometidos a cargas ácidas: neutrófilos y basófilos humanos.

Además, está involucrado en enfermedades como el cáncer (Bayrhuber et al, 2019):

- La sobreexpresión de Hv1 se correlaciona significativamente con el tamaño del tumor en el cáncer de mama.

- Una isoforma más corta de Hv1 se enriquece específicamente en células de leucemia linfocítica crónica, lo que hace que Hv1 sea un objetivo farmacológico potencial.

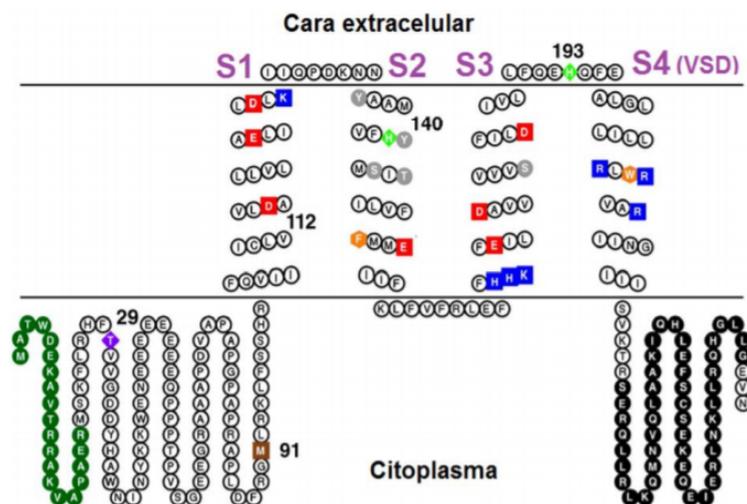


FIGURA 1. Esquema de Hv1 con sus aminoácidos. Se marcan algunos aminoácidos importantes, estudiados previamente, como el caso del Aspartato en la posición 112 (Asuaje, 2018).

MÉTODOS Y RESULTADOS

De la base de datos Uniprot se obtuvo la secuencia en formato FASTA, la cual se utilizó para la búsqueda de homólogos en BLAST. Como parámetros para asegurar homología, se aceptaron aquellos que cumplan con: un coverage del 70% (cuánto cubre la secuencia de la base de datos a la secuencia

query), un porcentaje de identidad de al menos 25% (ya que se trata de una proteína de casi 273 aminoácidos), con E-value cuyo valor máximo sea 10^{-6} . El número de hits seleccionado fue de 5000 y el word size fue modificado a 6 para una mayor sensibilidad. En la corrida del BLAST se obtuvieron 973 secuencias dentro de los parámetros mencionados anteriormente. Se encontraron principalmente en organismos eucariotas, con predominio en animales, y en mucho menor proporción en bacterias.

Adicionalmente se realizó una búsqueda por PSIBLAST con el objetivo de encontrar homólogos más remotos. Se corrió PSIBLAST con un número de hits de 500 y luego de 3 iteraciones, se obtuvieron 500 secuencias con porcentajes de identidad mayores al 30%, coverage del 70%, y E-value bajos. En especial, pudieron encontrarse homólogos remotos que en una corrida de BLAST no llegaban a encontrarse.

Del total de las secuencias homólogas encontradas usando BLAST, se seleccionaron 12, cada una perteneciente a una especie distinta y se realizó un alineamiento múltiple utilizando T-coffee (Fig.2). El aminoácido Asp112 es crucial para la selectividad de los protones. Se muestra en la figura parte del alineamiento en donde se observa que el Asp112 se encuentra 100% conservado.

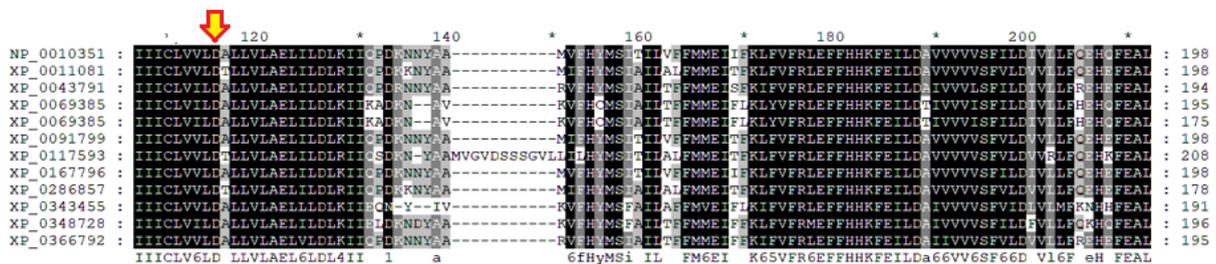


FIGURA 2. Visualización de una sección del alineamiento múltiple de doce secuencias homólogas.

Por otro lado, se realizó una búsqueda con la secuencia query en Pfam, una base de datos HMM de dominios secuenciales que pueden o no coincidir con dominios estructurales. Los resultados permitieron reconocer un dominio de transporte de iones, una región C-terminal del canal Hv1, dos regiones coiled-coil, cuatro sitios transmembrana y una zona desordenada (Fig. 3).



FIGURA 3. Esquema de los resultados obtenidos en Pfam.

El dominio ion_trans hace referencia a un dominio encontrado en canales de sodio, potasio y calcio. Esta familia tiene seis hélices transmembrana, donde las últimas dos flanquean un loop encargado de la selectividad iónica. En el caso de Hv1, no comparte la totalidad del dominio ya que sólo posee cuatro hélices transmembrana.

La familia de este dominio está formada por 3908 especies distintas, distribuidas principalmente en eucariotas y bacterias, y en menor medida en arqueobacterias (Fig. A1).

Con respecto al segundo dominio, VGPC1_C, hace referencia a la región C-terminal del canal de protones voltaje operado encontrado en eucariotas, necesario para que se dé la dimerización, ya que esta se mantiene por interacciones de hélices superenrolladas entre los dominios intracelulares C terminal.

Además, la región es esencial para la localización de la proteína en una membrana intracelular. La familia de este dominio está formada por 134 especies distintas, del dominio eucariota (Fig. A2).

A continuación, se realizó la predicción de estructura secundaria utilizando el programa Quick2D (Gabler, et al. 2020) (Fig. 4):

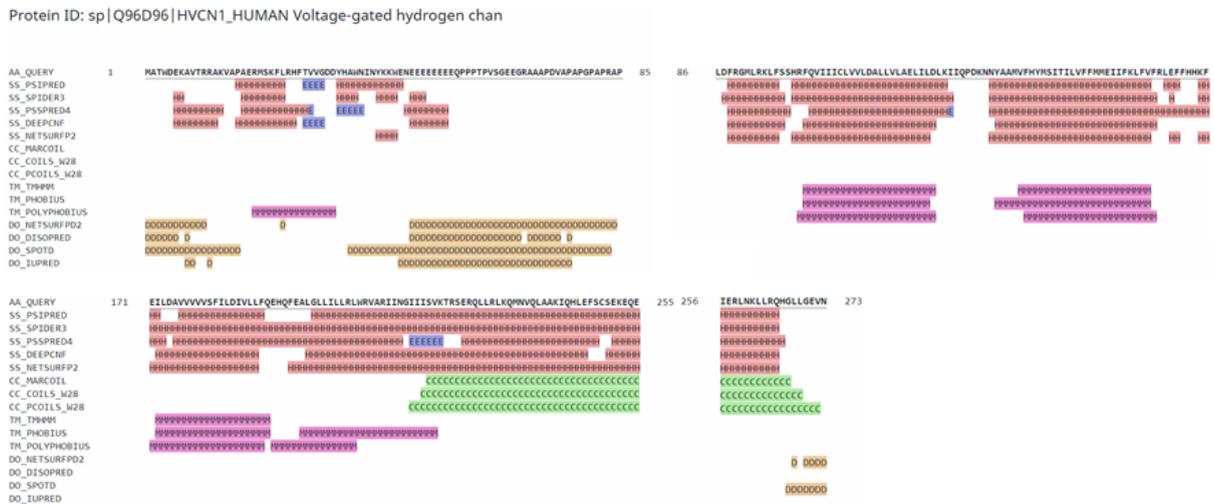


FIGURA 4. Predicción de la estructura secundaria de Hv1. Se representa en rosa las regiones α -hélice, regiones coiled-coil en verde, regiones transmembrana en violeta y las desordenadas en marrón.

Se puede ver mayormente regiones α -hélice, coiled-coil y transmembrana. Se visualiza, además, una pequeña porción desordenada.

Para profundizar en cada una de las estructuras y corroborar los resultados anteriores, se realizó una predicción con distintos servidores de regiones desordenadas, transmembrana y coiled-coil. Para la predicción de desorden se obtuvo la presencia de éste en una pequeña porción de la proteína, utilizando <https://mobidb.org/> (Fig. 5).

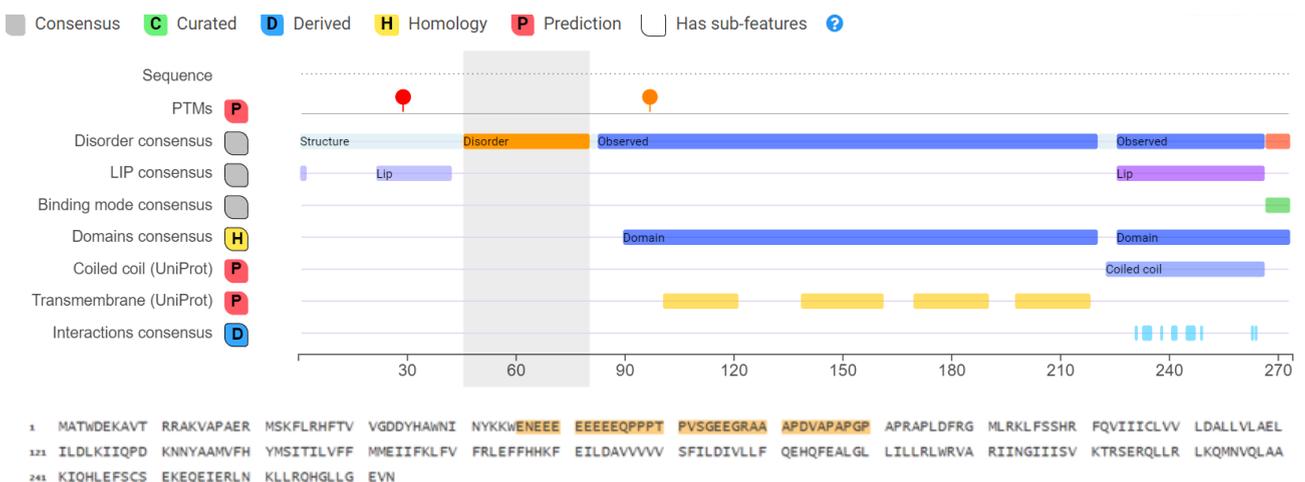


FIGURA 5. Resultado de la predicción de desorden utilizando <https://mobidb.org/>

Para la predicción de regiones transmembrana, se utilizó el programa TMHMM (Fig.A3), mediante el cual se ve una probabilidad alta (cercana a 1) de encontrar tres segmentos transmembrana, mientras que el cuarto segmento tiene una probabilidad baja, de 0,3. Sin embargo Quick2D y Phobius (Käll et al., 2004), encuentran altas posibilidades de existencia de cuatro de estas regiones (Fig.6):

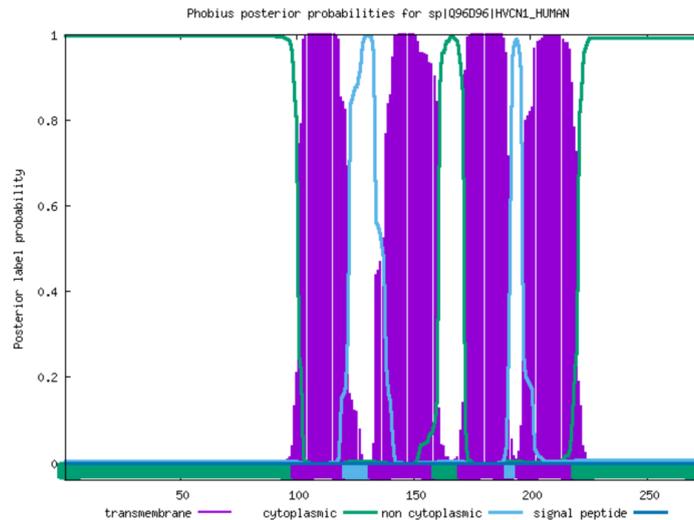


FIGURA 6. Predicción de regiones transmembrana con Phobius. Se ven cuatro regiones transmembrana: del aminoácido 101 al 122, del aminoácido 134 al 160, del aminoácido 172 al 191, y del aminoácido 197 al 220.

Por otro lado, para el análisis de regiones coiled-coil se utilizó el programa Coils (Fig.7). Predice, usando la ventana más grande, una región coiled-coil cercana al extremo C-terminal (coincidente con la predicción realizada con Quick2D). Mientras que con ventanas más chicas, existe probabilidad de encontrar además una región adicional en la zona media de la proteína.

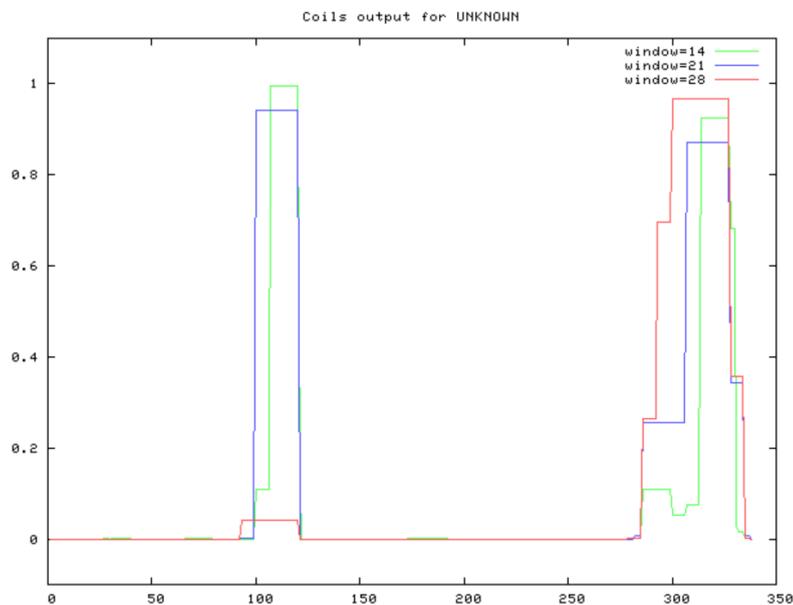


FIGURA 7. Predicción de regiones coiled-coil con Embnet.

Luego de hacer el análisis de la estructura secundaria de la proteína, se realizó un modelado de la estructura terciaria de la misma, usando la proteína homóloga en ratón como template y el programa Modeller. De los 10 modelos obtenidos, todos fueron parecidos y no muy buenos. Esto probablemente se deba a que sólo se encontró una única estructura homóloga (en una única conformación), con una resolución de 3.75 Å. Además, hubo varios residuos perdidos sin coordenadas, por lo que el programa tuvo que trabajar con poca información valiosa. Se utilizó el programa Pymol para la visualización y alineamiento de los modelos con el template (Fig.8).

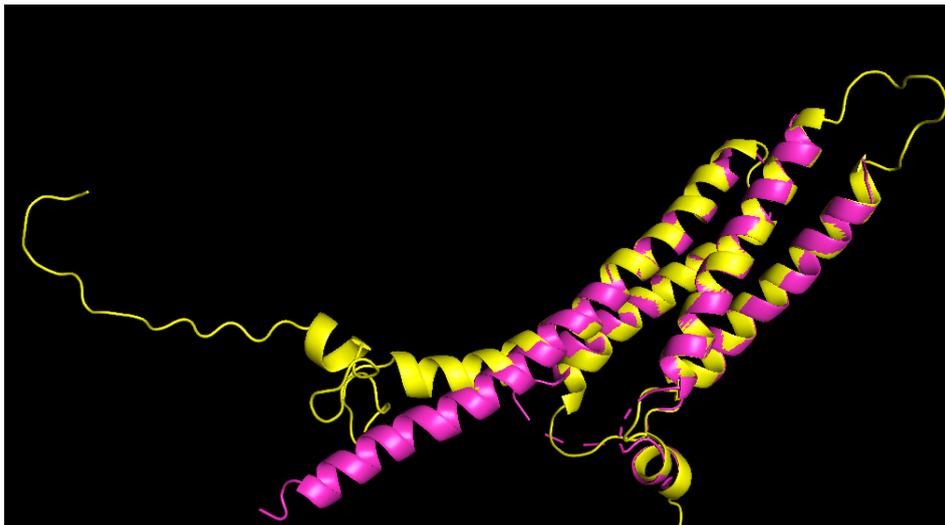


FIGURA 8. Modelado por homología. En rosa el template 3WKV, en amarillo uno de los modelos elegidos. RMSD = 0,243.

La calidad de los modelos obtenidos se analizó con Procheck y Verify (Laskowski et al., 1993; Bowie et al, 1991). El programa Verify analiza el ambiente químico en el cual se encuentra un aminoácido en una posición específica. El resultado arroja que 1,1% de los residuos tienen un score mayor a 0,2. Idealmente, al menos el 80% de los aminoácidos deberían tener ese score (Fig. A4).

Procheck, nos da información sobre el Ramachandran Plot, como forma de evaluar la estereoquímica de los aminoácidos. Es decir, los ángulos psi y phi de cada uno de los aminoácidos de la proteína. Dentro del Plot hay regiones permitidas y no permitidas, por lo que para ser considerada una estructura de buena calidad, al menos el 90% de los aminoácidos debe caer en posiciones permitidas. En este caso, hay un 88,85% dentro de la zona permitida. Lo anterior es otra verificación de que el modelado no fue bueno, tanto en términos de ambiente químico como en aspectos estereoquímicos.

A continuación se realizó un análisis filogenético. Para ello se corrió un BLAST en Uniprot, eligiendo 13 secuencias homólogas, de diferentes organismos mamíferos, ya que el canal sólo se encuentra en éstos. Se realizó un alineamiento múltiple con T-coffee y el análisis filogenético por el método de maximum likelihood utilizando IQ-TREE (incluye la elección del modelo evolutivo, siendo el mejor JTT+G4). El árbol filogenético obtenido fue editado con SNAD y visualizado mediante FigTree (Fig.9). Para el root se usó la opción de punto medio y se utilizó el mismo color para aquellos organismos que compartían el último nodo, es decir, el último evento de especiación en cada caso.

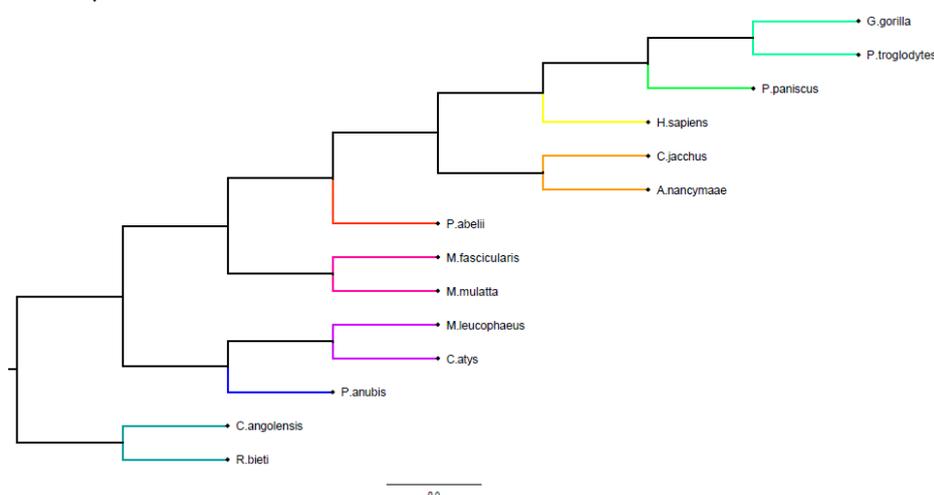


FIGURA 9. Árbol filogenético de la proteína estudiada, obtenido en FigTree

Para terminar de analizar a la proteína, se utilizaron los servidores ConSurf y Evolutionary Trace (Lua et al., 2016). Para el primero, se usó uno de los modelos estructurales obtenidos anteriormente, dando como resultado que 179 de 273 residuos tienen puntuaciones de conservación poco confiables debido a datos insuficientes en el alineamiento múltiple (Fig. 10).

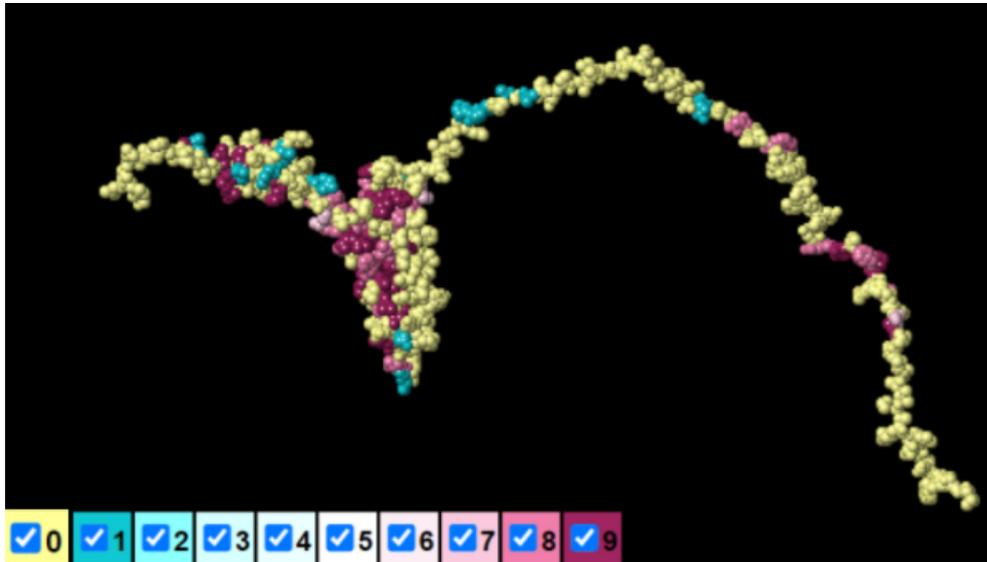


FIGURA 10. Visualización de la conservación de los residuos. El color amarillo evidencia que no hay información suficiente. Desde el celeste al púrpura la conservación aumenta.

El Evolutionary Trace calcula distancias evolutivas para ver el rango relativo de importancia estructural y funcional entre las posiciones de las secuencias alineadas de los homólogos. El rango es menor si las posiciones de la secuencia varían entre homólogos evolutivamente más cercanos y mayor si las posiciones varían entre homólogos evolutivamente distantes.

Se le pidió al programa que haga un alineamiento múltiple con secuencias homólogas seleccionadas por el mismo, arrojando un gráfico Logo y una representación de la secuencia (Fig. 11).

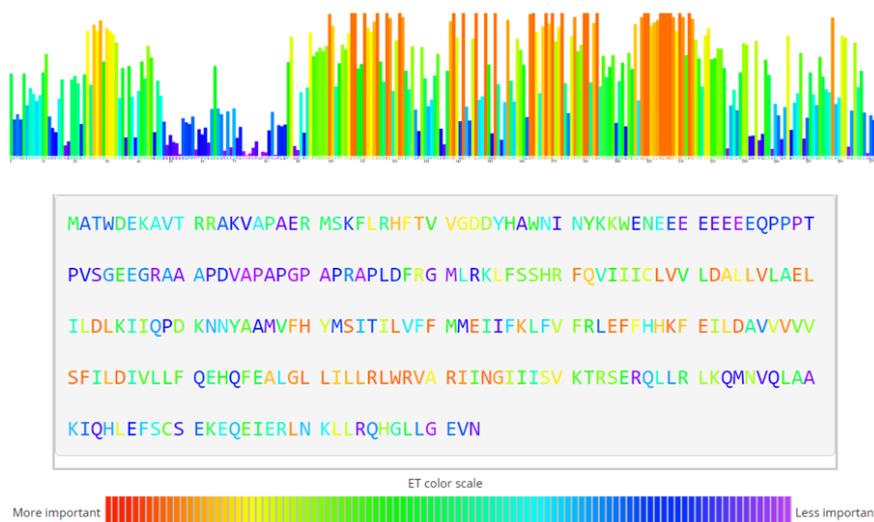


FIGURA 11. Gráfico Logo obtenido con Evolutionary Trace y secuencia en colores del canal.

Se destaca una mayor conservación de la región transmembrana del canal, que coincide con la zona de selectividad iónica. Además del Asp112, existen otros sitios destacados como: Thr29 es el sitio de fosforilación por la protein quinasa C (PKC) lo cual deriva en la potenciación del canal, His140 y His193

coordinan la unión del zinc, el cual ejerce un efecto inhibitorio. (http://sedici.unlp.edu.ar/bitstream/handle/10915/68542/Documento_completo.pdf-PDFA.pdf?sequence=1&isAllowed=y).

CONCLUSIONES Y DISCUSIÓN

Con la información disponible del canal de protones voltaje operado Hv1 se pudo realizar un estudio secuencial, estructural y funcional del mismo, mediante el uso de diferentes programas y servidores.

Se realizó un recorrido por diferentes bases de datos, extrayendo información de las mismas, incluida la secuencia del gen HVCN1 que utilizamos en la mayoría de las instancias.

Se pudo analizar la estructura secundaria del canal y aproximarse a la estructura terciaria usando el mejor homólogo encontrado.

Por otro lado, con algunos de los homólogos cercanos y remotos seleccionados, se pudo construir un árbol filogenético, y conocer, además, posiciones conservadas que podrían llegar a estar relacionadas con la función de la proteína.

BIBLIOGRAFÍA

Asuaje, A. (2018). *Rol del canal de protones operado por voltaje (HVCN1) en la homeostasis ácido-base de células tumorales T humanas y su impacto en la apoptosis (Doctoral dissertation, Universidad Nacional de La Plata)*.

Bayrhuber, M., Maslennikov, I., Kwiatkowski, W., Sobol, A., Wierschem, C., Eichmann, C., Frey, L. & Riek, R. (2019). Nuclear magnetic resonance solution structure and functional behavior of the human proton channel. *Biochemistry*, 58(39), 4017-4027.

Gabler, F., Nam, S. Z., Till, S., Mirdita, M., Steinegger, M., Söding, J., Lupas, A. N. & Alva, V. (2020). Protein sequence analysis using the MPI bioinformatics toolkit. *Current Protocols in Bioinformatics*, 72(1), e108.

TMHMM (Server v. 2.0), <http://www.cbs.dtu.dk/services/TMHMM/>

Käll, L., Krogh, A., & Sonnhammer, E. L. (2004). A combined transmembrane topology and signal peptide prediction method. *Journal of molecular biology*, 338(5), 1027-1036.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of applied crystallography*, 26(2), 283-291.

Bowie, J. U., Lüthy, R., & Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, 253(5016), 164-170.

Lua, R. C., Wilson, S. J., Konecki, D. M., Wilkins, A. D., Venner, E., Morgan, D. H., & Lichtarge, O. (2016). UET: a database of evolutionarily-predicted functional determinants of protein sequences that cluster as functional sites in protein structures. *Nucleic acids research*, 44(D1), D308-D312.