



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 311 667**

(51) Int. Cl.:
G01N 33/68 (2006.01)
G01N 33/50 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Número de solicitud europea: **03027613 .3**
(96) Fecha de presentación : **02.12.2003**
(97) Número de publicación de la solicitud: **1426768**
(97) Fecha de publicación de la solicitud: **09.06.2004**

(54) Título: **Método para la identificación de agentes para el tratamiento de ataques, trastornos neurológicos, endocrinopatías y trastornos hormonales.**

(30) Prioridad: **03.12.2002 US 308163**

(45) Fecha de publicación de la mención BOPI:
16.02.2009

(45) Fecha de la publicación del folleto de la patente:
16.02.2009

(73) Titular/es: **UCB Pharma S.A.**
Allée de la Recherche 60
1070 Brussels, BE

(72) Inventor/es: **Lynch, Berkley;**
Nocka, Karl y
Fuks, Bruno

(74) Agente: **Elzaburu Márquez, Alberto**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Método para la identificación de agentes para el tratamiento de ataques, trastornos neurológicos, endocrinopatías y trastornos hormonales.

La presente invención se dirige en general al campo del descubrimiento de fármacos en trastornos neurológicos, endocrinopatías y enfermedades hormonales.

Los trastornos neurológicos afligen a un número sustancial de individuos y presentan un creciente reto económico para los sistemas sanitarios puesto que se conoce poco con respecto a sus causas, su diagnóstico es con frecuencia subjetivo y muchos carecen de tratamiento efectivo. En general, la actividad cerebral está fundamentalmente determinada por la capacidad de las neuronas para comunicarse en las sinapsis. Compuestos químicos neurotransmisores específicos están empaquetados en neuronas presinápticas en el interior de vesículas sinápticas las cuales se fusionan con la membrana presináptica para liberar *cuantos* del compuesto químico neurotransmisor que atraviesan la ranura sináptica para activar el tipo correspondiente de receptor residente en la membrana postsináptica. Entre estos tipos de receptores están los receptores neuronales del glutamato (GluR), los receptores del ácido γ -aminobutírico (GABAR), los receptores nicotínicos de la acetilcolina, los receptores de la serotonina, los receptores de la dopamina, y semejantes. Muchos trastornos neurológicos son resultado de la conducción impropia de corrientes eléctricas a través de las sinapsis en varios tejidos cerebrales. En la epilepsia las corrientes errantes, que se piensa están asociadas con una función impropia de las sinapsis, provocan varios niveles de ataques. Asimismo, en varias enfermedades psiquiátricas, trastornos del movimiento y trastornos neurodegenerativos las corrientes de conducción se vuelven aberrantes, desorganizadas o reducidas, provocando de este modo la condición de enfermedad. Por consiguiente, los defectos en las funciones de las vesículas sinápticas tendrán un efecto adverso sobre la neurotransmisión en general y el control de la liberación de neurotransmisores en particular.

Los ataques, incluyendo los ataques epilépticos, son consecuencia de una perturbación focal o generalizada de la función cortical, la cual puede deberse a varios trastornos cerebrales o sistémicos que incluyen, por ejemplo, el edema cerebral, la hipoxia cerebral, el trauma cerebral, infecciones del sistema nervioso central (SNC), defectos del cerebro congénitos o desarrollados, lesiones que expanden el cerebro, hiperpirexia, perturbaciones metabólicas y el uso de fármacos convulsivos o tóxicos. La epilepsia se diagnostica sólo cuando los ataques se repiten a intervalos esporádicos y durante el curso de años (o indefinidamente).

Etiológicamente, la epilepsia se clasifica como sintomática o idiopática con manifestaciones de ataques que caen en tres categorías generales: 1) tónico-clónico generalizado, 2) ausencia o pequeño mal, y 3) parcial complejo. La clasificación sintomática indica que existe una causa probable y que se puede probar un curso específico de terapia para eliminar esa causa, mientras que idiopática indica que no puede encontrarse ninguna causa obvia y puede estar unida a factores genéticos sin explicar. De las categorías de ataques, la mayoría de las personas tienen sólo un tipo de ataque, mientras que aproximadamente 30% tienen dos o más tipos.

El riesgo de desarrollar epilepsia es 1% desde el nacimiento a los 20 años de edad, y 3% a la edad de 75 años. En general, la epilepsia idiopática comienza entre los 2 y 14 años de edad. Los ataques antes de los 2 años son usualmente provocados por defectos en el desarrollo, lesiones al nacer o una enfermedad metabólica. Los que comienzan después de los 25 años de edad pueden ser secundarios al trauma cerebral, los tumores o las enfermedades cerebrovasculares, pero el 50% son de etiología desconocida.

Debido a las muchas interrelaciones que existen entre los sistemas nervioso y endocrino, los defectos en la función de la vesícula sináptica también pueden impactar en la función endocrinológica. Por ejemplo, al menos dos glándulas secretan sus hormonas sólo en respuesta a la liberación de los neurotransmisores apropiados - la médula adrenal y la glándula pituitaria posterior. Tras la secreción, las hormonas son transportadas en la sangre para provocar acciones fisiológicas en tejidos diana distantes en el cuerpo. Obviamente, las endocrinopatías que implican hiper o hiposecreción tienen consecuencias patológicas. Ejemplo de estas consecuencias son el gigantismo y el enanismo debidos, respectivamente, a la hiper o a la hiposecreción de la hormona del crecimiento.

El (LEV; ucb L059; (S)- α -etil-oxo-pirrolidina acetamida), el enantiómero (S) del antólogo etilo del piracetam, fue sintetizado durante un programa químico dirigido a identificar un fármaco nootrópico de segunda generación. Los resultados *in vivo* han demostrado una potente capacidad inesperada del LEV para suprimir ataques en el ratón audiogénico susceptible, mientras que el piracetam fue sólo débilmente activo. Aunque el LEV es una molécula no relacionada con los fármacos antiepilépticos establecidos (Margineanu *et al.*, en *Antiepileptic Drugs*; 5ª edición, pp. 419-427. Lippincott, Filadelfia (2002)), ensayos clínicos extensivos han probado que la terapia adjunta con LEV (KEPPRA, UCB S.A., Braine-l'Alleud, Bélgica) es tanto efectiva como bien tolerada en el control contumaz de los ataques parciales en adultos.

Los ensayos de enlace con LEV, realizados con membranas brutas de cerebro de rata, revelan la existencia de un sitio de enlace específico reversible, saturable y estereoselectivo. Los resultados obtenidos en membranas de hipocampo de rata sugieren que LEV marca a una única clase de sitios de enlace con modesta afinidad y con una alta capacidad enlazante. Este sitio de enlace se identifica como sitio de enlace de levetiracetam (LBS). Resultados similares se han obtenido en otras regiones de cerebro (corteza, cerebelo y cuerpo estriado). El enantiómero (R) de levetiracetam, ucb L060, muestra aproximadamente 1000 veces menos afinidad por estos sitios. El enlace de LEV parece que está con-

finado a membranas en el sistema nervioso central ya que estudios con radiomarcadores no pudieron detectar ningún enlace específico en una gama de tejidos periféricos que incluían el corazón, los riñones, el páncreas, los adrenales, los pulmones y el hígado.

5 Sin embargo, esto podría ser debido a la baja densidad del LBS en estos tejidos comparada con el sistema nervioso central y realmente el enlace específico no se produce en células PC 12, una línea celular adrenal periféricamente derivada. Los fármacos antiepilépticos más comúnmente usados carbamazepina, fenitoína, valproato, fenobarbital y clonazepam así como el agente convulsivo L-butilbiciclofosforotionato (TBPS), el picrotoxin y la bicuculina no desplazan el enlace del LEV. Sin embargo, la etosuximida, el pentobarbital, el pentilenotetrazol y la bemegrida compitieron con LEV con valores de PKi comparables a las concentraciones activas de los fármacos observadas *in vivo*. Los
10 compuestos estructuralmente relacionados, incluyendo el piracetam y el aniracetam, también desplazaron el enlace del LEV. También se ensayaron compuestos análogos del levetiracetam respecto a su actividad anticonvulsiva en el modelo de ratón audiogénico de la epilepsia. Se observó una correlación muy buena ($r^2 = 0,84$) entre la afinidad y la actividad anticonvulsiva (Noyer *et al.*, Euro. J. Pharmacol. 286:137-146 (1995)). Este alto grado de correlación es
15 un fuerte soporte para una relación causal entre el enlace al LBS y la actividad anticonvulsiva de esta clase de compuestos. Por consiguiente, es de esperar que el enlace de los compuestos análogos del levetiracetam al LBS dé lugar a una modificación de la función del o de los componentes proteicos del LBS en el cerebro, lo que conduce al efecto terapéutico deseado resultado de la actividad anticonvulsiva.

20 La familia de proteínas 2 de la vesícula sináptica (SV2) de proteínas de la vesícula sináptica fue identificada en primer lugar con un anticuerpo monoclonal preparado contra vesículas colinérgicas del órgano eléctrico de la raya marina *D. ommata* (Buckley *et al.*, J. Cell. Biol. 100:1284-1294, (1985)). La clonación de los miembros individuales de la familia marcados con el anticuerpo dio lugar a la identificación de tres isoformas diferentes, SV2A (Bajjalieh *et al.*, Science 257:1271-1273 (1992)), SV2B (Feany *et al.*, Cell 70(5):861-867, 1992) y SV2C (Janz y Sudhof, Neuroscience 94(4):1279-1290 (1999)), todos los cuales reaccionan con el anticuerpo original. La homología global entre las tres isoformas es aproximadamente 60%, siendo SV2A y SV2C más similares entre sí que SV2B (Janz y Sudhof, Neuroscience 94(4):1279-1290 (1999)).

30 Las proteínas SV2 son proteínas integrales de las membranas y tienen una homología significativa pero de bajo nivel (20-30%) con la familia doce transmembranal de proteínas transportadoras bacterianas y fúngicas que transportan azúcar, citrato y agentes xenobióticos (Bajjalieh *et al.*, Science 257:1271-1273 (1992)). Como miembros putativos de superfamilia 12 TM, las proteínas SV2 muestran varias características únicas. Tienen unos terminales N- y C- relativamente cortos y bucles cortos que conectan los segmentos Tm. Sin embargo, dos notables excepciones son el bucle citoplasmático largo entre las regiones transmembranales 6 y 7 y el bucle intravesicular entre las regiones transmembranales 7 y 8 (que contiene 3 sitios de N-glicosilación). Aún no se ha descubierto ningún homólogo próximo de las
35 proteínas SV2 en levaduras o invertebrados, aunque una proteína de la vesícula sináptica distantemente relacionada conocida como SVOP tiene homólogos en *Drosophila* y *C. elegans* (Janz y Sudhof, Neuroscience 18(22):9269-9281 (1998)).

40 Como familia, las proteínas SV2 están ampliamente distribuidas en el cerebro y en las células endocrinas. Las tres isoformas se solapan significativamente en su distribución, y pueden encontrarse coexpresadas en la misma neurona, e incluso en la misma vesícula sináptica. Una u otra isoforma de las proteínas SV2 parece estar presente en todas la vesículas sinápticas y probablemente no están limitadas a neuronas que contienen algún neurotransmisor específico, aunque un estudio informa que las vesículas colinérgicas pueden no contener SV2 (Blumber *et al.*, J. Neurochem. 58 (3):801-810 (1992)). Las proteínas SV2 son por lo tanto una de las proteínas más comunes de las vesículas sinápticas y han sido implicadas en el control de la exocitosis mediada por el calcio de proteínas SV2 de las vesículas sinápticas. También se ha mostrado que las proteínas SV2 se expresan en células endocrinas y, junto con las proteínas integrales adicionales p38 y p65 de la membrana de la vesícula sináptica se ha demostrado que están presentes en membranas endocrinas con gránulos de núcleo denso (Lowe *et al.*, J. Cell. Biol. 106(1):51-59 (1988). SV2A, la isoforma más
50 común de SV2, se expresa ubicuamente por todo el cerebro y está bien presente en los gránulos secretores de las células endocrinas. Aunque SV2B está ampliamente distribuida en el cerebro, es indetectable en varias estructuras del cerebro, incluyendo la circunvolución dentada del hipocampo, el globo pálido, los núcleos reticulares del tálamo y la parte reticular del núcleo de Soemmering, el mesencéfalo y el tallo encefálico del bulbo olfatorio. Es indetectable en la corteza cerebral y en el hipocampo, y se encuentra en bajas concentraciones en la corteza cerebelosa (Janz y Sudhof, Neuroscience 94(4):1279-1290 (1999)).
55

Además de la proteína SV2, la sinapsis contiene otras proteínas reguladoras únicas tales como la sinapsina, la sinaptotagmina y CAPS, las cuales pueden mediar en la fusión o el brote de las vesículas. SV2A puede ser una proteína reguladora del Ca^{2+} esencial para la formación de complejos de prefusión llamados complejos SNARE (Xu *et al.*, Cell 99(7):713-722 (1999)), que incluyen las proteína VAMP/sinaptobrevina asociada a la vesícula sináptica y las proteínas de la membrana plasmática syntaxina y SNAP-25. Tras la acumulación del Ca^{2+} en la sinapsis se inhibe el enlace de la sinaptotagmina a SV2A y se estimula la dimerización de dos dominios de enlace sinaptotagmina Ca^{2+} (Bajjalieh, Curr. Opin. Neurobiol. 9(3):321-328 (1999)). Esta dimerización puede jugar un papel en organizar el complejo SNARE y promover la fusión de las vesículas, ya que a bajas concentraciones de Ca^{2+} SV2A permanece
65 enlazada a la sinaptotagmina y la fusión no se produce.

La afinidad de SV2A por la sinaptotagmina está regulada por la fosforilación del término amino de SV2 (Pyle *et al.*, J.Biol.Chem. 275(22):17195-17200 (2000)). La posibilidad de que las proteínas SV2 jueguen un papel en el

transporte del Ca^{2+} o en la regulación en la vesícula sináptica ha sido soportada por estudios de animales sin SV2A y SV2B (Janz *et al.*, Neuron 24:1003-1016 (1999)). Una hipótesis alternativa es que las proteínas SV2, aunque derivadas de proteínas de transporte, cumplen ahora una función diferente en la vesícula, ya un papel estructural, ya un papel en la regulación de la fusión o el reciclado de las vesículas y en la liberación exocitótica de sus contenidos (Janz y Sudhof, Neuroscience 94(4):1279-1290 (1999)).

Ha habido dos informes de ratones sin proteínas SV2: uno que examina solo los que no tienen SV2A (Crowder *et al.*, Proc. Nat. Acad. Sci. USA 96(26):15268-15273 (1999)) y el otro que contempla tanto animales sin SV2A como sin SV2B así como sin ambas proteínas ((Janz *et al.*, Neuron 24:1003-1016 (1999)).

Los animales homocigotos respecto al trastorno del gen de SV2A parecen normales al nacer, pero no crecen, experimentan ataques graves, y mueren a las pocas semanas de nacer. Los animales homocigotos sin SV2A experimentan ataques que son de mayor duración, más fuertes y más debilitantes que en otras cepas de ratones (Janz *et al.*, Neuron 24:1003-1016 (1999)). A pesar de la aparición de ataques postnatales, todos los animales sin SV2A tienen una morfología del conjunto del cerebro completamente normal, incluyendo concentraciones normales de las proteínas sinápticas ensayadas. Además, los cultivos neuronales del hipocampo tanto de ratones sin SV2A como sin SV2A/SV2B formaron sinapsis que fueron ultraestructuralmente normales, y que tuvieron un tamaño, número y localización de vesículas sinápticas inalterado (Janz *et al.*, Neuron 24:1003-1016 (1999); Crowder *et al.*, Proc. Nat. Acad. Sci. USA 96(26):15268-15273 (1999)). Es interesante advertir que, al contrario que en los frecuentemente observados ataques provocados por anomalías estructurales y del desarrollo fácilmente detectadas en muchos otros tipos de animales sin estas proteínas, los ratones sin SV2A muestran un fenotipo de ataque fuerte sin ninguna anomalía asociada del cerebro o las sinapsis en macro o micro escala. Esta observación sugiere un papel directo y específico de SV2A y el fenotipo observado. Como otro marcador de la función cerebral, estudios de la transmisión sináptica en cultivos neuronales primarios de ratones sin SV2A, SV2B y SV2A/SV2B indican que los tamaños y las frecuencias de las sIPSCs y de las corrientes postsinápticas excitantes espontáneas (sEPSCs) son normales. La estimulación eléctrica indujo EPSCs e IPSCs robustas en las neuronas cultivadas de todos los genotipos.

En contraste con SV2A, los ratones sin SV2B no revelan ninguna patología abierta (Janz *et al.*, 1999). Se sugiere que una posible razón de esta carencia de consecuencias de la pérdida de SV2B es que puede ser funcionalmente reemplazada por SV2A, la cual parece que se coexpresa en cualquier lugar en el que SV2B se expresa normalmente.

Aunque la función de SV2A y de otros miembros de la familia permanece desconocida, la hipótesis favorecida es que este transportador homólogo es un transportador funcional de alguna molécula común de la vesícula sináptica. Más específicamente, hay evidencia que une SV2A a la regulación de la exocitosis de las vesículas mediada por el calcio y como resultado se piensa que puede ser una proteína transportadora del Ca^{2+} . SV2A y otros miembros de la familia también tienen papeles en la función de las vesículas sinápticas. Tales papeles pueden incluir aspectos moduladores de su formación, carga con el neurotransmisor, fusión con la membrana plasmática, reciclado e interacciones con otras proteínas y compartimentos y orgánulos celulares. Por ejemplo, se ha mostrado que las proteínas SV2 pueden interactuar con la proteína de las vesículas sinápticas sinaptotagmina y la proteína de la matriz extracelular laminina-1 (Carlsson, Perspect. Dev. Neurobiol. 3(4):373-386 (1996)). Las proteínas SV2 pueden jugar importantes papeles en la regulación de las concentraciones citoplasmáticas o en los orgánulos de calcio en la terminal presináptica, y también pueden interactuar con los canales del calcio del tipo N de la membrana plasmática, directa o indirectamente.

Los presentes inventores han descubierto que SV2A es el sitio de enlace del fármaco antiataques LEV y de sus análogos. El alto grado de correlación entre las afinidades relativas de enlace de una serie de análogos de levetiracetam y sus potencias anticonvulsivas en ciertos modelos animales de epilepsia proporciona una fuerte evidencia de que el enlace de estos análogos a las proteínas SV2 modifica su función para proporcionar efectos anticonvulsivos.

En una realización preferida, la invención incluye un método para identificar un socio de enlace de una proteína SV2. El método comprende incubar una proteína o un fragmento de proteína SV2 con levetiracetam o uno de sus análogos o derivados y un potencial socio de enlace. El método comprende además determinar si el potencial socio de enlace modula el enlace de levetiracetam o de uno de sus análogos o derivados a la proteína o un fragmento de proteína SV2, identificando de este modo un socio de enlace de la proteína SV2.

En otra realización preferida, la invención incluye un método para identificar un socio de enlace de una proteína SV2. El método comprende exponer a una proteína o un fragmento de proteína SV2 a un potencial socio de enlace e incubar la proteína o el fragmento de proteína y el potencial socio de enlace con (2S)-2-[4-(3-azidofenil)-2-oxopirrolidin-1-il]butanamida. El método comprende además determinar si el enlace de (2S)-2-[4-(3-azidofenil)-2-oxopirrolidin-1-il]butanamida a la proteína es inhibido por el potencial socio de enlace, identificando de este modo un socio de enlace de la proteína.

En aún otra realización preferida, la invención incluye un método para identificar un agente útil para el tratamiento de un trastorno neurológico o endocrinológico. El método comprende exponer a una proteína o un fragmento de proteína SV2 al agente y a levetiracetam o uno de sus análogos o derivados. El método además comprende determinar si el enlace de levetiracetam o uno de sus análogos o derivados a la proteína es modulado por el agente, identificando de este modo un agente útil para el tratamiento de un trastorno neurológico o endocrinológico.

ES 2 311 667 T3

En aún otra realización preferida, la invención incluye un método para identificar un agente útil para el tratamiento de un trastorno neurológico o endocrinológico. El método comprende exponer a una proteína o un fragmento de proteína SV2 al agente e incubar la proteína o el fragmento de proteína y el agente con (2S)-2-[4-(3-azidofenil)-2-oxopirrolidin-1-il]butanamida. El método comprende además determinar si el enlace de (2S)-2-[4-(3-azidofenil)-2-oxopirrolidin-1-il]butanamida a la proteína es inhibido por el agente, identificando de este modo socios de enlace de la proteína.

La figura 1 representa el enlace reversible del análogo de LEV ucb 30889 a LBS en la corteza cerebral de rata.

La figura 2 representa las curvas de saturación del enlace de ucb 30889.

La figura 3 muestra que el enlace específico no pudo detectarse en los tejidos periféricos.

La figura 4 representa las curvas de competición por el enlace que muestran que ucb 30889 se enlaza a LBS con una afinidad aproximadamente 10 veces superior que LEV.

La figura 5 representa los valores de pIC_{50} de ucb 30889 frente a levetiracetam.

La figura 6 representa la inhibición del enlace de $[^3H]$ ucb 30889 dependiente de la concentración por levetiracetam sin marcar en autorradiografía de cerebro de rata.

La figura 7 representa una autorradiografía del enlace de $[^3H]$ ucb 30889 a secciones coronales de cerebro de rata.

La figura 8 representa la distribución subcelular del enlace de $[^3H]$ ucb 30889 dentro del cerebro de rata.

La figura 9 representa el subfraccionamiento de la fracción sinaptosomal por centrifugación en gradiente de sacarosa.

La figura 10 representa el fotomarcaje del LBS mediante $[^3H]$ ucb 30889 y la irreversibilidad del complejo.

La figura 11 representa una electroforesis en gel de proteínas de la membrana marcadas con $[^3H]$ ucb 30889.

La figura 12 (A y B) representa lisados inmunoteñidos de las células COS-7 transfectadas con SV2A, membranas brutas de cerebro de ratón y varios lisados diferentes de PC 12 con diferentes concentraciones de LBS.

La figura 13 representa el enlace específico de $[^3H]$ ucb 30889 a células COS-7 transfectadas con SV2A-12.2, transfectadas con el vector testigo que expresa el gen β -gal, o células que no han sido transfectadas.

La figura 14 representa una representación gráfica de IC_{50} que compara tres ligandos diferentes que se enlazan a SV2A en presencia de 3H-ucb 30889.

La figura 15 representa la estructura de (A) levetiracetam y (B) ucb 30889.

En los ensayos descritos en la presente memoria puede usarse cualquier proteína SV2 que se enlaza LEV o a uno de sus derivados o análogos.

Cuando se usa en la presente memoria, las proteínas SV2 incluyen proteínas aisladas, variantes alélicas de las proteínas y sustituciones conservadoras de aminoácidos de las proteínas. Cuando se usa en la presente memoria, la “proteína” o el “polipéptido” se refiere, en parte, a SV2A, una proteína codificada por la secuencia de ácido nucleico de SEQ ID NO: 1 ó que tiene la secuencia de aminoácidos de ser humano representada en la secuencia SEQ ID NO: 2 ó uno de sus fragmentos; a SV2B, que incluye a la proteína de ser humano codificada por la secuencia de ácido nucleico de SEQ ID NO: 3 ó la secuencia de aminoácidos representada en la secuencia SEQ ID NO: 4 o uno de sus fragmentos; a SV2C, que incluye a la proteína de ser humano codificada por la secuencia de ácido nucleico de SEQ ID NO: 5 ó la secuencia de aminoácidos representada en la secuencia SEQ ID NO: 6 ó uno de sus fragmentos; y a SVOP, que incluye a la proteína de ser humano codificada por la secuencia de ácido nucleico de SEQ ID NO: 7 ó la secuencia de aminoácidos representada en la secuencia SEQ ID NO: 8 ó uno de sus fragmentos. Los términos también se refieren a variantes alélicas y proteínas naturales que tiene una secuencia de aminoácidos ligeramente diferente que la específicamente enumerada anteriormente. Las variantes alélicas, aunque poseen una secuencia de aminoácidos ligeramente diferente que las enumeradas anteriormente, aún tendrán las mismas o similares funciones biológicas asociadas con estas proteínas.

Cuando se usa en la presente memoria, la familia de proteínas SV2 relacionada con la secuencia de aminoácidos de ser humano SEQ ID NO: 2, 4, 6 u 8, se refiere en parte a proteínas que han sido aisladas de organismos además de seres humanos. Por ejemplo, se han identificado y están incluidos en la presente memoria homólogos de rata del ácido nucleico (SEQ ID NO: 9) y la proteína (SEQ ID NO: 10) SV2A, del ácido nucleico (SEQ ID NO: 11) y la proteína (SEQ ID NO: 12) SV2B, del ácido nucleico (SEQ ID NO: 13) y la proteína (SEQ ID NO: 14) SV2C y del ácido nucleico (SEQ ID NO: 15) y la proteína (SEQ ID NO: 16) SVOP. Los métodos usados para identificar y aislar otros miembros de la familia de proteínas relacionada con estas proteínas se describen más adelante.

Las proteínas SV2 usadas en la presente invención están preferiblemente en forma aislada en parte de un fragmento de la membrana celular de las vesículas, se expresan en una célula huésped transformada, o se expresan naturalmente en una célula o tipo de tejido dado. Cuando se usa en la presente memoria, se dice que una proteína está aislada cuando se emplean métodos físicos, mecánicos o químicos para separar la proteína de sus constituyentes celulares que normalmente están asociados con la proteína. Un experto puede emplear fácilmente métodos estándar de purificación para obtener una proteína aislada.

Las proteínas SV2 que pueden usarse en los métodos de la invención incluyen además variantes derivadas de inserciones, supresiones, sustituciones conservadoras de aminoácidos o empalmes de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14 ó 16. Cuando se usa en la presente memoria, una variante “conservadora” se refiere a alteraciones de la secuencia de aminoácidos que no afectan adversamente a las funciones biológicas de la proteína.

Se dice que una sustitución, inserción o supresión afecta adversamente a la proteína cuando la secuencia alterada impide o trastorna una función biológica asociada con la proteína. Por ejemplo, la carga total, la estructura o las propiedades hidrófobas/hidrófilas de la proteína pueden ser alteradas sin afectar adversamente a una actividad biológica. Por consiguiente, la secuencia de aminoácidos puede ser alterada, por ejemplo para hacer que el péptido sea más hidrófobo o hidrófilo, sin afectar adversamente las actividades biológicas de la proteína. Cuando se usa en la presente memoria, una “supresión” se define como un cambio en la secuencia nucleótida o de aminoácidos en las cuales están ausentes uno o más residuos nucleótidos o aminoácidos, respectivamente; una “inserción” o “adición” es el cambio de la secuencia nucleótida o de aminoácidos que ha resultado de la adición de uno o más residuos nucleótidos o aminoácidos, respectivamente, en comparación con la proteína SV2 natural y una “sustitución” resulta del reemplazamiento de uno o más nucleótidos o aminoácidos por diferentes nucleótidos o aminoácidos, respectivamente.

Las proteínas SV2 de la presente invención incluyen además proteínas de fusión, en las que una proteína SV2, o uno de sus fragmentos, se fusiona terminalmente por los extremos N- o C- con otra proteína SV2 o uno de sus fragmentos, que puede ser la misma que o diferente de la primera proteína SV2 o uno de sus fragmentos, y/o con un socio de fusión tipo péptido heterólogo. El péptido heterólogo puede ser una secuencia de polipéptido útil para la expresión, purificación, solubilidad, identificación, antigenicidad o extensión de la estabilidad de la proteína SV2 o uno de sus fragmentos. Los socios de fusión heterólogos útiles en la presente invención incluyen, pero no se limitan a, glutatión-S-transferasa (GST), etiquetas de polihistidina, proteína verde fluorescente (GFP), albúmina y ovoalbúmina o uno de sus fragmentos.

Ordinariamente, las variantes alélicas, las variantes con sustituciones conservadoras y los miembros de la familia de proteínas SV2 tendrán una secuencia de aminoácidos que al menos tiene aproximadamente 35%, 40%, 50%, 60%, 65%, 70% ó 75% de identidad de la secuencia de aminoácidos con la secuencia de longitud completa puesta de manifiesto en SEQ ID NO: 2, 4, 6, 8, 10, 12, 14 ó 16, más preferiblemente al menos aproximadamente 80%, incluso más preferiblemente al menos aproximadamente 90%, y mucho más preferiblemente al menos aproximadamente 95%, 97% ó 99% de identidad de la secuencia. En la presente memoria, la identidad u homología con respecto a tales secuencias se define como el porcentaje de residuos aminoácidos en la secuencia candidata que son idénticos con los péptidos conocidos, después de alienar las secuencia se introducir huecos, si es necesario, para conseguir el máximo porcentaje de homología, y no considerando ninguna sustitución conservadora como parte de la identidad de la secuencia (véase la sección B para los parámetros relevantes). Las proteínas de fusión, o las extensiones, supresiones o inserciones N-terminales, C-terminales o internas en la secuencia del péptido no se interpretarán como que afecten a la homología.

Las variantes contempladas incluyen además las que contienen mutaciones predeterminadas mediante, por ejemplo, recombinación de homólogos, mutagénesis por PCR o dirigida al sitio, y las correspondientes proteínas de otras especies de animales, que incluyen, pero no se limitan a, conejo, ratón, rata, porcino, bovino, ovino, equino y especies de primates no humanos, y los alelos u otras variantes naturales de la familia de proteínas; y derivados en los que la proteína ha sido covalentemente modificada por sustitución química, enzimática o por otros medios apropiados con un resto diferente que un aminoácido natural (por ejemplo un resto detectable tal como una enzima o un radioisótopo).

En los métodos de la invención también pueden usarse fragmentos de las proteínas SV2. En particular, pueden usarse fragmentos que comprenden el sitio de enlace de LEV. Tales fragmentos pueden tener al menos aproximadamente 6 ó 10, 15 ó 20, ó 25 ó 30 residuos aminoácidos, más preferiblemente 35 ó 40 residuos aminoácidos, incluso más preferiblemente 45 ó 50 residuos aminoácidos, aún más preferiblemente 55 ó 60, aún más preferiblemente 65 ó 70 residuos aminoácidos y mucho más preferiblemente 75 ó más residuos aminoácidos.

Los métodos de la presente invención también pueden utilizar moléculas de ácidos nucleicos que codifican miembros de la familia de proteínas SV2, que incluyen, pero no se limitan a, proteínas tanto de rata como de ser humano conocidas como SV2A, SV2B, SV2C y la proteína relacionada de las vesículas sinápticas SVOP, tales como las que consisten en o que comprenden las secuencias SEQ ID NO: 2, 4, 6, 8, 10, 12, 14 ó 16 y las proteínas relacionadas descritas en la presente memoria, preferiblemente en forma aislada. Para producir una proteína SV2 también pueden usarse vectores, plásmidos y células huésped transformadas. Cuando se usa en la presente memoria, “ácido nucleico” se define como RNA o DNA o moléculas relacionadas que codifican una proteína o péptido que se definieron anteriormente, son complementarias a una secuencia de ácido nucleico que codifica tales péptidos, hibridan a tal ácido nucleico y permanecen establemente enlazadas a él en condiciones severas apropiadas, o codifican un polipéptido que al menos comparte aproximadamente 35%, 40%, 50%, 60%, 65%, 70% ó 75% de identidad de la secuencia, pre-

feriblemente al menos aproximadamente 80%, más preferiblemente al menos aproximadamente 85%, e incluso más preferiblemente al menos aproximadamente 90%, 95%, 97% ó 99% o más de identidad con las secuencias de péptidos de longitud completa de SEQ ID NO: 2, 4, 6, 8 ó 10. Las “moléculas de ácidos nucleicos” útiles en la invención incluyen además moléculas de ácidos nucleicos que al menos comparten aproximadamente 70% ó 75% de identidad de las secuencias, preferiblemente al menos aproximadamente 80%, más preferiblemente al menos aproximadamente 85%, e incluso más preferiblemente al menos aproximadamente 90% y mucho más preferiblemente 95%, 97%, 99% o más de identidad con las secuencias de nucleótidos de SEQ ID NO: 1, 3, 5, 7 ó 9. Los ácidos nucleicos de la presente invención también incluyen los que codifican proteínas de fusión que comprenden una proteína SV2 N- o C-terminalmente fundida a una secuencia heteróloga de proteína o a otra secuencia de proteína SV2.

La homología o la identidad al nivel de la secuencia de nucleótidos o de aminoácidos se determina mediante análisis BLAST (Herramienta Básica de Búsqueda de Alineamiento Local) usando el algoritmo empleado por los programas blastp, blastn, blastx, tblastn y tblastx (Altschul *et al.*, Nucleic Acids Res. 25: 3389-3402 (1997); Karlin *et al.*, Proc. Natl. Acad. Sci. USA 87:2264-2268 (1990)) que están diseñados para la búsqueda de la similitud de las secuencias. El enfoque usado por el programa BLAST es considerar en primer lugar segmentos similares, con y sin huecos, entre una secuencia duda y una secuencia base de datos, evaluar a continuación la significancia estadística de todos los emparejamientos que son idénticos y finalmente sumarizar sólo los emparejamientos que satisfagan un umbral de significancia preseleccionado. Para una discusión de los temas básicos en la búsqueda de la similitud de las bases de datos de secuencias véase Altschul *et al.* (Nature Genetics 6, 119-129 (1994)). Los parámetros de búsqueda para el histograma, las descripciones, alineaciones, la expectativa (es decir, el umbral de significancia estadística para informar de emparejamientos frente a las secuencias de la base de datos), el punto de corte, la matriz y el filtro (baja complejidad) están en las posiciones de por defecto. La matriz de puntuación por defecto usada por blastp, blastx, blastn, tblastn y tblastx es la matriz BLOSUM62 (Henikoff *et al.*, Proc. Natl. Acad. Sci. USA 89:10915-10919 (1992)), recomendada para las secuencias duda por encima de una longitud de 85 (bases nucleótidas o aminoácidos).

Para blastn, la matriz de puntuación se establece mediante las relaciones de M (es decir, la puntuación recompensa para un par de residuos que se emparejan) a N (es decir, la puntuación de penalización para residuos que no se emparejan), en la que los valores por defecto para M y N son +5 y -4, respectivamente. Cuatro parámetros de blastn se ajustaron como sigue: Q = 10 (penalización por creación de huecos); R = 10 (penalización por extensión de huecos); wink = 1 (genera aciertos de palabras en cada posición wink a lo largo de la secuencia duda); y gapw = 16 (establece la anchura de la ventana dentro de la cual se generan alineamientos con intersticios). Las posiciones equivalentes de los parámetros Blastp fueron Q = 9; R = 2; wink = 1; y gapw = 32. Una comparación de huecos entre secuencias, disponible en la versión 10.2 del programa Winsconsin Package de Accelrys, usa parámetros para DNA GAP = 50 (penalización por creación de huecos) y LEN = 3 (penalización por extensión de huecos) y las posiciones equivalentes en comparaciones de proteínas son GAP = 8 y LEN = 2.

“Condiciones severas” incluyen las que: (1) emplean una baja fuerza iónica y una alta temperatura de lavado, por ejemplo, NaCl 0,015 M/citrato de sodio 0,0015 M/SDS 0,1% a 50°C, o (2) durante la hibridación emplean un agente desnaturizante tal como formamida, por ejemplo, formamida al 50% (vol/vol) con albúmina de suero de bovino al 0,1%/Ficoll al 0,1%/polivinilpirrolidona al 0,1%/tampón de fosfato de sodio 50 mM a pH 6,5 con NaCl 750 mM, citrato de sodio 75 mM a 42°C. Otro ejemplo es la hibridación en formamida al 50%, 5 x SSC (NaCl 0,75 M, citrato de sodio 0,075 M), fosfato de sodio 50 mM (pH 6,8), pirofosfato de sodio al 0,1%, 5 x disolución de Denhardt, DNA de esperma de salmón sonificado (50 µg/mL), SDS al 0,1% y sulfato de dextrano al 10% a 42°C, con lavados a 42°C en 0,2 x SSC y SDS al 0,1%. Un experto puede determinar y variar fácilmente las condiciones severas apropiadamente para obtener una señal de hibridación clara y detectable. Las moléculas preferidas son las que hibridan en las anteriores condiciones al complemento de SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 ó 15 y las que codifican una proteína funcional. Incluso, las moléculas que se hibridan más preferidas son las que se hibridan en las anteriores condiciones a la hebra del complemento del marco de lectura abierto de SEQ ID NO: 1, 3, 5, 7, 9, 11, 13 ó 15. Cuando se usa en la presente memoria, se dice que una molécula de ácido nucleico está “aislada” cuando la molécula de ácido nucleico está sustancialmente separada de las moléculas de ácidos nucleicos contaminantes que codifican otros polipéptidos.

A. SV2A y sitio de enlace de levetiracetam (LBS)

La invención incluye la caracterización y el uso del LBS localizado en la proteína SV2A.

Como se describió anteriormente, “SV2A” incluye la proteína de ser humano que se describe en SEQ ID NO: 2, la proteína de ser humano codificada por SEQ ID NO: 1, especies homólogas de SV2A de ser humano, variantes de SEQ ID NO: 2 que se describen en la presente memoria y fragmentos de SV2A que comprenden el LBS.

II. Levetiracetam y compuestos análogos

Los métodos de la invención incluyen el uso de LEV y compuestos análogos de LEV o sus derivados en ensayos para identificar nuevos agentes farmacológicos. En una realización preferida, los métodos de la presente invención identifican compuestos o agentes que compiten con LEV y los compuestos análogos de LEV o sus derivados por enlazarse al LBS de SV2. Cuando se usa en la presente memoria, los términos “compite” y “enlace competitivo” se refieren a agentes o compuestos que ocupan el mismo sitio de enlace en el LBS que LEV o los compuestos análogos de LEV o sus derivados; desplazan, o son desplazados por, LEV o los compuestos análogos de LEV o sus derivados en el enlace al LBS; o inhiben, o son inhibidos por, LEV o los compuestos análogos de LEV o sus derivados en el enlace

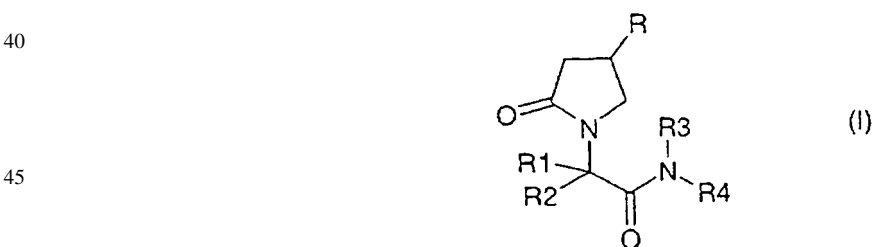
al LBS. En otra realización preferida, la invención incluye la identificación de compuestos o agentes que modulan la actividad de SV2A. En otra realización preferida, los métodos de la presente invención identifican compuestos o agentes que tienen menos, aproximadamente la misma, o mayor actividad por el LBS que LEV. En aún otra realización preferida, los métodos de la presente invención identifican compuestos o agentes que en una cantidad efectiva modulan la actividad de SV2A durante un período de tiempo más largo que una cantidad efectiva de LEV. En incluso otra realización preferida, los métodos de la presente invención identifican compuestos o agentes que modulan en una cantidad efectiva la actividad de SV2A durante un período de tiempo más corto que una cantidad efectiva de LEV.

Cuando se usa en la presente memoria, “levetiracetam” (figura 15A; LEV) se refiere al nombre internacional de uso libre del compuesto (S)- α -etil-oxo-1-pirrolidina acetamida que se describe en la patente europea nº 0162036B1, que se incorpora en su totalidad a la presente memoria a modo de referencia. LEV es un compuesto levorrotatorio que es un agente protector para el tratamiento y la prevención de agresiones al sistema nervioso central de tipo hipóxico e isquémico. Este compuesto es también efectivo en el tratamiento de la epilepsia. La α -etil-oxo-1-pirrolidina acetamida racémica y sus compuestos análogos son conocidos por la patente británica nº 1309692. La patente de EE.UU. nº 3.459.738 describe derivados de 2-oxo-1-pirrolidina acetamida.

Cuando se usa en la presente memoria, la expresión “compuestos análogos de LEV o sus derivados” incluye derivados de 2-oxo-pirrolidina N-alquilados opcionalmente sustituidos. Preferiblemente, esos compuestos son derivados de alquilamidas sustituidos en las posiciones 4 y/o 5 del anillo de pirrolidona. Ejemplos de derivados de 2-oxo-pirrolidina N-alquilados opcionalmente sustituidos incluyen, pero no se limitan a, compuestos tales como los descritos en la solicitud de patente internacional PCT/EP01/01992 tales como (2S)-2-[(4S)-4-(2,2-difluorovinil)-2-oxopirrolidinil]butanamida, (2S)-2-[(4R)-2-oxo-4-propilpirrolidinil]butanamida, (2S)-2-[(4S)-2-oxo-4-propilpirrolidinil]butanamida y (2S)-2-[4-(3-azidofenil)-2-oxo-4-propilpirrolidinil]butanamida y (2S)-2-[4-(3-azidofenil)-2-oxopirrolidinil-1-il]butanamida.

Cuando se usa en la presente memoria, la expresión “compuestos análogos de LEV o sus derivados” incluye además derivados de 2-oxo-piperidinilo N-alquilados opcionalmente sustituidos. Preferiblemente, esos compuestos son derivados de alquilamidas sustituidos en las posiciones 4 y/o 5 y/o 6 del anillo de 2-oxo-piperidinilo. Ejemplos de derivados de 2-oxo-piperidinilo N-alquilados opcionalmente sustituidos incluyen, pero no se limitan a, compuestos tales como los descritos en la solicitud de patente internacional PCT/EP02/05503 tales como (2S)-2-[5-(yodometil)-2-oxo-1-piperidinil]butanamida, (2S)-2-[5-(azidometil)-2-oxo-1-piperidinil]butanamida, 2-(2-oxo-5-fenil-1-piperidinil)butanamida, (2S)-2-[4-(yodometil)-2-oxo-1-piperidinil]butanamida y (2S)-2-[4-(2-fluoro-2-metilpropil)-2-oxo-1-piperidinil]butanamida.

Cuando se usa en la presente memoria, la expresión “compuestos análogos de LEV o sus derivados” incluye cualquier compuesto tipo acetam de la fórmula I, en forma racémica o isómera, o una de sus sales farmacéuticamente aceptables.



en la que:

- R representa un átomo de hidrógeno o un grupo hidroxilo;
- R¹ y R² representan independientemente un átomo de hidrógeno o un grupo alquilo de 1-4 átomos de carbono; y
- R³ y R⁴ representan independientemente un átomo de hidrógeno, un grupo alquilo de 1-4 átomos de carbono o un grupo -(CH₂)_n-NR⁵R⁶ en el que n es 1, 2 ó 3 y R⁵ y R⁶ representan independientemente un átomo de hidrógeno o un grupo alquilo de 1-4 átomos de carbono.

Un ejemplo de tal compuesto tipo acetam incluye, pero no se limita a, un compuesto de fórmula I en la que R, R¹, R², R³ y R⁴ son un átomo de hidrógeno, 2-oxo-pirrolidina-acetamida, conocido por el nombre genérico de piracetam que se describe en las patentes UK nºs 1.039.113 y 1.309.692.

Cuando se usa en la presente memoria, la expresión “compuestos análogos de LEV o sus derivados” también incluye derivados de 2-oxo-azepanilo N-alquilados opcionalmente sustituidos. Preferiblemente, esos compuestos son derivados de alquilamidas sustituidos en las posiciones 4 y/o 5 y/o 6 y/o 7 del anillo de 2-oxo-azepanilo. Ejemplos de derivados de 2-oxo-azepanilo N-alquilados opcionalmente sustituidos incluyen, pero no se limitan a, compuestos

tales como los descritos en la solicitud de patente internacional PCT/EP02/05503 tales como 2-[5-(yodometil)-2-oxo-1-azepanil]butanamida.

En otra realización, la presente invención incluye compuestos o agentes que son derivados o compuestos análogos de piracetam los cuales se enlazan al LBS. Tales compuestos también incluirían moléculas tales como aniracetam y nefiracetam. En una realización preferida, los derivados o compuestos análogos de piracetam son los que modulan la actividad de SV2A o de otros miembros de la familia SV2.

III. Formatos de ensayo

Los ensayos de la presente invención incluyen métodos para identificar agentes o compuestos que sean útiles para el tratamiento de trastornos neurológicos, tales como ataques, epilepsia, enfermedad de Parkinson, disquinesias de Parkinson, migrañas, enfermedad de Alzheimer, dolor neuropático, temblor esencial, trastornos cognitivos, trastornos del movimiento, endocrinopatía y enfermedades relacionadas con la médula adrenal relacionada, tales como la hipoglucemia y el choque circulatorio. Los ensayos de la presente invención también incluyen métodos para identificar agentes o compuestos que tengan efectos que potencian la cognición, tal como, por ejemplo, podría medirse en modelos animales de cognición. En particular, los ensayos de la presente invención incluyen métodos para identificar agentes o compuestos que compiten con LEV o sus compuestos análogos o derivados por el enlace al LBS de SV2A, desplazan, o son desplazados por, LEV o los compuestos análogos de LEV o sus derivados en el enlace al LBS; o inhiben, o son inhibidos por, LEV o los compuestos análogos de LEV o sus derivados en el enlace al LBS. LEV, ucb 30889 (figura 15B) y otros derivados o compuestos análogos de LEV que se describieron anteriormente son útiles en los métodos de la invención como agentes enlazantes en ensayos para seleccionar nuevos compuestos o agentes que se enlazan al LBS de SV2A. En tales realizaciones de ensayo, pueden usarse LEV, ucb 30889 y los derivados o compuestos análogos sin modificación o pueden modificarse de una diversidad de maneras; por ejemplo, por marcaje, tal como unir covalente o no covalentemente un resto que directa o indirectamente proporcione una señal detectable. En cualquiera de estos ensayos, los materiales pueden marcarse directa o indirectamente. Las posibilidades del marcaje directo incluye grupos marcadores tales como: radiomarcadores que incluyen, pero no se limitan a, [^3H], [^{14}C], [^{32}P], [^{35}S] o [^{125}I], enzimas tales como la peroxidasa y la fosfatasa alcalina y marcadores fluorescentes capaces de monitorizar el cambio de la intensidad de fluorescencia, el desplazamiento de la longitud de onda o la polarización de la fluorescencia, que incluyen, pero no se limitan a, fluoresceína o rodamina. Además, podrían usarse técnicas FRET para analizar interacciones entre ligandos y el LBS de SV2A. Las posibilidades del marcaje indirecto incluyen la biotilación de un constituyente seguida por el enlace a la avidina condensada con uno de los grupos marcadores anteriores o el uso de anticuerpos antiligando. Los compuestos también pueden incluir agentes espaciadores o de unión en los casos en los que los compuestos se han de unir a un soporte sólido.

Para identificar agentes o compuestos que compitan o interaccionen con LEV y ucb 30889 y derivados por el enlace al LBS de SV2A pueden usarse células intactas, fragmentos celulares o de las membranas que contengan SV2A o la proteína SV2A completa o un fragmento que comprenda el LBS de la proteína SV2A. El agente o compuesto puede incubarse con las células, membranas, la proteína o el fragmento de proteína SV2 antes de, al mismo tiempo que o después de la incubación con LEV o uno de sus compuestos análogos o derivados. Los ensayos de la presente invención pueden medir cualquier propiedad o función de las proteínas SV2, vesículas sinápticas, transmisión neural y/o función de las células endocrinas, así como la acumulación presináptica de cationes divalentes, incluyendo los iones Ca^{2+} . Ejemplos de propiedades o funciones de una proteína SV2 que pueden medirse como punto final de un ensayo incluyen, pero no se limitan a, el estado de la fosforilación, el enlace de cationes divalentes, incluyendo los iones Ca^{2+} ; el transporte por las membranas; el transporte de cationes divalentes (incluyendo los iones Ca^{2+}) hacia el interior y/o hacia afuera de las vesículas sinápticas; el transporte de neurotransmisores (que incluyen, pero no se limitan a, aminas, acetilcolina, neurotransmisores excitantes, GABA, serotonina y glicina) hacia el interior y/o hacia afuera de las vesículas sinápticas; la interacción con otras proteínas (que incluyen, pero no se limitan a, las lamininas y la sinaptotagmina); los cambios conformacionales, que se miden por la sensibilidad a la proteólisis u otros cambios en propiedades bioquímicas o biofísicas; la formación de canales para cationes divalentes; la formación o disociación de complejos de proteínas; la función de las vesículas sinápticas; la fusión; la exocitosis y el reciclado de las vesículas sinápticas.

Los ensayos de la invención pueden modificarse o prepararse en cualquier formato disponible, incluyendo ensayos de alta productividad que monitorizan el enlace de LEV o el enlace de sus derivados o análogos a SV2A o al LBS de la proteína SV2A. En muchos programas de selección de fármacos que ensayan bibliotecas de compuestos, los ensayos de alta productividad son deseables con el fin de maximizar el número de compuestos examinados en un periodo de tiempo dado. Tales ensayos de selección pueden usar células intactas, fragmentos celulares o de las membranas que contengan SV2A así como sistemas exentos de células o de membranas, tales como los que pueden derivarse de proteínas purificadas o semipurificadas. La ventaja del ensayo con un fragmento de membrana que contenga proteínas SV2A o proteínas y péptidos SV2A purificados es que los efectos de la toxicidad celular y/o de la biodisponibilidad del compuesto de ensayo pueden en general ignorarse, focalizándose el ensayo en su lugar principalmente en el efecto del fármaco sobre la diana molecular que puede manifestarse en una inhibición de, por ejemplo, el enlace entre dos moléculas.

En una realización de un ensayo competitivo de selección, el ensayo puede formularse para detectar la capacidad del agente o compuesto de ensayo para inhibir el enlace de ucb 30889 a SV2A o un fragmento de SV2A que comprenda el LBS, o de LEV, o de sus derivados o compuestos análogos, a SV2A o a un fragmento de SV2 que comprenda el

LBS. En otra realización de un ensayo competitivo de selección, el ensayo puede formularse para detectar la capacidad de ucb 30889 o de LEV, o de sus derivados o compuestos análogos, para inhibir el enlace de un agente o compuesto de ensayo a SV2A o un fragmento de SV2A que comprenda el LBS. La inhibición de la formación del complejo puede detectarse mediante una variedad de técnicas. Por ejemplo, la modulación de la formación puede cuantificarse usando, por ejemplo, ucb 30889, LEV, o sus derivados o compuestos análogos detectablemente marcados. La inhibición de la formación de un complejo puede detectarse usando una versión detectablemente marcada del agente o compuesto que se está ensayando respecto al enlace competitivo al LBS de SV2A. Alternativamente, el enlace entre la proteína SV2A y un ligando puede detectarse sin ninguna necesidad de una sonda marcada. Por ejemplo, los instrumentos a escoger para tales ensayos de enlace son la resonancia de plasmón superficial, la resonancia magnética nuclear o la espectrometría de masas. Otro método es medir cambios en la sensibilidad de las proteínas SV2 a proteasas inducidos por el enlace a un ligando.

En ciertos casos, será deseable inmovilizar uno de los LBS (SV2A o un fragmento de SV2A que comprenda el LBS) o el ligando (LEV, ucb 30889 o el agente o compuesto de ensayo) para facilitar la separación de los complejos de las formas no complejadas, así como para acomodar la automatización del ensayo. El enlace de un ligando al LBS, por ejemplo el enlace de un agente o compuesto candidato a SV2A, en presencia o ausencia de LEV o de ucb 30889, puede conseguirse en un depósito adecuado para contener los reaccionantes. Ejemplos incluyen las placas de microtitulación, los tubos de ensayo y los tubos de microcentrífuga. En una realización, puede proporcionarse una proteína de fusión la cual añade un dominio que permite que el LBS se enlace a una matriz. Por ejemplo, las proteínas de fusión de glutatión-S-transferasa (GST) pueden adsorberse sobre bolas de glutatión sefarosa (Sigma Chemical, St Louis, Mo.) o placas de microtitulación derivatizadas con glutatión, las cuales a continuación se combinan con LEV, ucb 30889 o derivados de LEV sin marcar y el agente o compuesto marcado. A continuación, la mezcla se incuba en condiciones que conducen a la formación del complejo. Tras la incubación, las bolas se lavan para separar cualesquiera de los reaccionantes sin enlazar y la marca inmovilizada sobre la matriz se determina directamente, o en el sobrenadante después de que los complejos LBS/ligando sean subsiguientemente disociados. Cuando es posible, los complejos pueden disociarse de la matriz, separarse por SDS-PAGE, y la marca del ligando encontrada en la fracción de las bolas cuantificarse a partir del gel usando técnicas electroforéticas estándar.

También están disponibles otras técnicas para inmovilizar proteínas sobre matrices para usar en el ensayo objeto. Por ejemplo, el LBS puede inmovilizarse usando la conjugación de la biotina y de la estreptavidina. Las moléculas biotiniladas pueden prepararse a partir de biotin-NHS (N-hidroxisuccinimida) usando técnicas bien conocidas en la técnica (por ejemplo, el kit de biotinilación, Pierce Chemicals, III) e inmovilizarse en los pocillos de placas de 96 pocillos revestidos de estreptavidina (Pierce Chemicals). Alternativamente, en los pocillos de la placa pueden derivatizarse anticuerpos reactivos con el LBS pero que no interfieren con el enlace al ligando y atraparse el LBS en los pocillos mediante conjugación de anticuerpos. Como antes, las preparaciones de un ligando y de un compuesto de ensayo se incuban en los pocillos de la placa que tienen proteína y puede cuantificarse la cantidad de complejo proteína/ligando atrapada en el pocillo. Métodos ejemplos para detectar tales complejos, además de los descritos anteriormente, incluyen la inmunodetección de complejos usando anticuerpos reactivos con el ligando, o que son reactivos con la proteína y compiten por el enlace con el ligando.

En otra realización de la invención, pueden llevarse a cabo ensayos de enlace competitivo usando extractos celulares de células o tejidos que comprenden el LBS para identificar socios de enlace de SV2. Cuando se usa en la presente memoria, un extracto celular se refiere a una preparación o fracción que se hace de un lisado o célula desbaratada. La fuente preferida de extractos celulares será las células derivadas de tejido nervioso central o de tejidos endocrinos de ser humano. En particular, los extractos celulares pueden prepararse de una región particular, que incluye, pero no se limita a, el hipocampo, el cerebelo, el cerebro, la corteza cerebral, la pituitaria, la médula y la glándula adrenal. Además, pueden prepararse extractos celulares a partir de un aislado particular de células primarias del sistema nervioso central origen o de los sistemas endocrinos que incluyen, pero no se limitan a, neuronas, astrocitos y células endocrinas de la médula. Alternativamente, pueden prepararse extractos celulares a partir de líneas celulares disponibles, particularmente líneas celulares de origen neurológico o endocrino. Las líneas celulares contempladas en la presente memoria incluyen, pero no se limitan a, células PC 12 de feocromocitoma de rata, células AtT-20, GH3 y HIT.

Para obtener un extracto de una célula pueden usarse una variedad de métodos. Las células pueden romperse usando métodos de desmoronamiento físicos y químicos. Ejemplos de métodos de desmoronamiento físicos incluyen, pero no se limitan a, lisis con detergentes y lisis con enzimas. Un experto adaptará fácilmente métodos para preparar extractos celulares con el fin de obtener extractos para usar en los presentes métodos.

Una vez que se ha preparado un extracto de una célula, el extracto se mezcla con la proteína SV2 o un fragmento y otros componentes del ensayo en condiciones en las que puede ocurrir la asociación de la proteína con el socio de enlace, seguido por la adición de LEV o uno de sus compuestos análogos o derivados. Alternativamente, puede añadirse el LEV o uno de sus compuestos análogos o derivados al extracto celular antes o incluso a la vez que el agente o compuesto de ensayo. Pueden usarse una variedad de condiciones, siendo las condiciones más preferidas las que se parecen estrechamente a las encontradas en el citoplasma de una célula de ser humano. Características tales como la osmolaridad, el pH, la temperatura y la concentración del extracto celular usado, pueden variar para optimizar la asociación de la proteína con el socio de enlace.

Después de mezclar en condiciones apropiadas, el complejo enlazado se separa de la mezcla. Para separar la mezcla pueden usarse una variedad de técnicas. Por ejemplo, pueden usarse anticuerpos específicos para SV2A para inmunoprecipitar el complejo del socio de enlace. Alternativamente, pueden usarse técnicas estándar de separación química tales como la cromatografía y la centrifugación en gradiente de densidad/sedimento.

5

Después de la separación de los constituyentes celulares no asociados encontrados en el extracto, el socio de enlace puede disociarse del complejo usando métodos convencionales. Por ejemplo, la disociación puede conseguirse alterando la concentración de sales o el pH de la mezcla.

10

Para ayudar a la separación de pares asociados del socio de enlace del extracto mezclado, el LBS puede inmovilizarse sobre un soporte sólido. Por ejemplo, el LBS puede unirse a una matriz de nitrocelulosa o a bolas acrílicas. La unión del LBS a un soporte sólido ayuda a la separación de los pares péptido/socio de enlace de los otros constituyentes encontrados en el extracto. Los socios de enlace identificados pueden ser una única proteína o un complejo constituido por dos o más proteínas. Alternativamente, los socios de enlace pueden identificarse usando un ensayo Far-Western según los procedimientos de Takayama *et al.*, *Methods Mol. Biol.* 69:171-184 (1997) o Sauder *et al.*, *J. Gen. Virol.* 77:991-996 (1996), o identificarse por medio del uso de proteínas etiquetadas con epítopes o proteínas de fusión con GST.

15

Alternativamente, el sistema de dos híbridos en levadura puede ser una herramienta para la identificación de interacciones proteína-proteína.

20

La invención incluye el uso de compuestos o de agentes identificados por los métodos de la invención para el tratamiento de trastornos neurológicos o endocrinológicos. En una realización preferida, los agentes identificados por los métodos de la presente invención se usan para el tratamiento de la epilepsia, al epileptogénesis, los trastornos por ataques, las convulsiones y los síndromes de abstinencia. En otras realizaciones preferidas los agentes identificados por los métodos de la presente invención se usan para el tratamiento de trastornos neurológicos que incluyen trastornos bipolares, manía, depresión, ansiedad, migraña, estados clínicos de dolores crónicos trigeminales y otras neuralgias, dolor neuropático, hiperexcitabilidad relacionada con la anestesia, isquemia cerebral, trauma de cabeza, miotonía, abuso de la cocaína y del alcohol, apoplejía, mioclonus, temblor esencial, tics, síndrome de Tourette, disquinesia, espasticidad y otros trastornos del movimiento, hemorragia cerebral neonatal, esclerosis amiotrófica lateral, enfermedad de Parkinson y de Alzheimer y otros trastornos neurodegenerativos que incluyen la demencia. En aún otras realizaciones preferidas, los agentes identificados por los métodos de la presente invención se usan para el tratamiento de trastornos endocrinológicos que incluyen las endocrinopatías que implican hiper o hiposecreción de una o varias hormonas y enfermedades relacionadas con la médula adrenal, tales como la hipoglucemia y el choque circulatorio. En otras realizaciones preferidas, la invención incluye el uso de compuestos o agentes identificados por los métodos de la invención para el tratamiento de estados de excitación provocados por enfermedades que incluyen, pero no se limitan a, el abuso dependencia y/o abstinencia del alcohol o las drogas; y la emergencia de la anestesia general.

30

Sin más descripción, se cree que un experto en la técnica puede, usando la descripción precedente y los ejemplos ilustrativos siguientes, fabricar y utilizar los compuestos de la presente invención y practicar los métodos reivindicados. Por lo tanto, los siguientes ejemplos de trabajo indican específicamente realizaciones preferidas de la presente invención, y no se pretende que limiten de ninguna manera el resto de la descripción.

40

Ejemplo 1

Desarrollo de un compuesto análogo de levetiracetam para estudios de enlace

50

Se ha mostrado que LEV se enlaza a un sitio de enlace específico localizado preferencialmente en el cerebro (sitio de enlace de levetiracetam o LBS: Noyer *et al.*, *Euro J. Pharmacol.* 286:137-146 (1995)). Sin embargo, [³H]LEV mostró sólo una afinidad micromolar por este sitio, haciéndole inadecuado para una caracterización profunda. Este ejemplo describe las propiedades de enlace de [³H]ucb 30889, (2S)-2-[4-(3-azidofenil)-2-oxopirrolidin-1-il]butanamida, un compuesto análogo de levetiracetam. Los experimentos de enlace se realizaron en membrana brutas de cerebro de rata a 4°C como se describe en Noyer *et al.* (*Euro. J. Pharmacol.* 286:137-146 (1995)). El tiempo de incubación para los estudios de equilibrio fue 120 min. Para los estudios cinéticos y de competición se usó [³H]ucb 30889 (30 Ci/mmol) a una concentración de 1,3 nM en 0,5 mL de tampón Tris-HCl (pH 7,4) que contenía Mg²⁺ 2 mM. La localización del LBS en subestructuras de cerebro se evaluó por autorradiografía en cortes gruesos de 25 µm incubados en condiciones similares. Los cortes se lavaron a continuación dos veces durante 10 min a 4°C en Tris-HCl 50 mM (pH 7,4) que contenía BSA al 0,5%, se secaron y se expusieron durante 3 semanas a [³H]Hyperfilm a -20°C. El enlace no específico (NBS) se determinó mediante la inclusión de LEV 1 mM durante el periodo de incubación.

60

La figura 1 muestra que [³H]ucb 30889 se enlaza reversiblemente a LBS en la corteza de cerebro de rata. La cinética de enlace fue bifásica: los tiempos medios de asociación y disociación fueron respectivamente 3 ± 2 min y 4 ± 1 min para el componente rápido (25 a 50% de los sitios) y 47 ± 13 min y 61 ± 15 min para el componente lento. A 25°C, la cinética aumentó dramáticamente y sólo permaneció un componente.

65

La figura 2 muestra que las curvas de saturación del enlace de $[^3\text{H}]\text{ucb 30889}$ fueron compatibles con el marcaje de una población homogénea de sitios de enlace. K_D y B_{max} fueron respectivamente $42 \pm 10 \text{ nM}$ y $5054 \pm 704 \text{ fmol/mg}$ de proteína. La B_{max} fue similar al valor estimado usando $[^3\text{H}]\text{levetiracetam}$ como radioligando en preparaciones similares de membranas ($4718 \pm 413 \text{ fmol/mg}$ de proteína).

En los tejidos periféricos examinados podría no detectarse el enlace específico (figura 3). El límite de detección en las condiciones experimentales ($150 \mu\text{g}$ de proteína/ensayo y $1,3 \text{ nM}$ de radioligando) fue un B_{max} de 200 fmol/mg de proteína. Esto sugiere que al menos hay 25 veces más sitios de enlace en la corteza cerebral en comparación con la periferia.

Las curvas de competición por el enlace mostraron que ucb 30889 se enlaza a LBS con una afinidad aproximadamente 10 veces mayor que a LEV (figura 4). El pK_i de ucb 30889 ($7,1 \pm 0,2$) concuerda bien con el K_D de ucb 30889 que se determinó por la curva de saturación del enlace (figura 2). Los valores de pIC_{50} de una diversidad de compuestos análogos de levetiracetam y de otros compuestos que se sabe interaccionan con el LBS, tales como pentilenotetrazol o bemegrida (Noyer *et al.*, 1995), fueron idénticos tanto si se obtuvieron con $[^3\text{H}]\text{ucb 30889}$ como con $[^3\text{H}]\text{levetiracetam}$ (figura 5).

Secciones de cerebro de rata incubadas con $[^3\text{H}]\text{ucb 30889}$ (figura 6) muestran que los LBS marcados con $[^3\text{H}]\text{ucb 30889}$ están localizados difusamente por todo el cerebro y que este enlace puede ser inhibido por el levetiracetam en concentraciones equivalentes a las observadas en el enlace *in vitro* (figura 4).

Este ejemplo demuestra a través de estudios de competición por el enlace y distribución en los tejidos que ucb 30889 y LEV están ambos marcando los mismos sitios, es decir el LBS que está localizado por todo el sistema nervioso central. Comparado con LEV, ucb 30889 se enlaza a LBS con una afinidad 10 veces mayor y con un enlace no específico muy bajo. Estos criterios junto con cinéticas de enlace adecuadas a 4°C hicieron posible usar este radioligando para realizar estudios de enlace por autorradiografía en cortes de cerebro (figura 6) y mostrar la distribución anatómica de LBS en cerebro de rata.

Ejemplo 2

Distribución celular y subcelular del LBS

Para identificar y caracterizar el LBS *in situ*, se usó $[^3\text{H}]\text{ucb 30889}$ para realizar un mapa de la distribución del LBS dentro del cerebro y para estudiar tanto su distribución celular como subcelular. Para la autorradiografía de cerebro de rata se incubaron cortes de $25 \mu\text{m}$ con $[^3\text{H}]\text{ucb 30889}$ $1,3 \text{ nM}$ durante 120 min a 4°C en tampón Tris-HCl 50 mM ($\text{pH } 7,4$). Los ensayos de enlace con membranas de cerebro de rata y varias líneas de células neuronales se realizaron en condiciones similares. El enlace no específico se determinó mediante la inclusión de levetiracetam 1 mM en el ensayo. Para el fotomarcaje, las membranas se incubaron con $[^3\text{H}]\text{ucb 30889}$ 40 nM durante 120 min a 4°C en el mismo tampón, seguido por irradiación con luz UV durante 30 min .

Para la autorradiografía de cerebro de rata se incubaron cortes de $25 \mu\text{m}$ con $[^3\text{H}]\text{ucb 30889}$ $1,3 \text{ nM}$ durante 120 min a 4°C en tampón Tris-HCl 50 mM ($\text{pH } 7,4$). La figura 7 muestra que los sitios de enlace de ucb 30889 están heterogéneamente distribuidos en el cerebro de rata. Aunque no hay ningún enlace aparente, en la materia blanca hay una alta concentración de enlace en el giro dentado, el colículo superior, varios núcleos talámicos y en la capa molecular del cerebelo. El enlace es menos pronunciado en la corteza cerebral, el hipotálamo y el cuerpo estriado. Abreviaciones: cc, cuerpo calloso; Aca, comisura anterior; ic, cápsula interna; Mtg, tracto mamilotegmental; ML, capa molecular; Hi, hipocampo; DG, giro dentado; sc, colículo superior; CG, sustancia gris central; Pu, caudate putamen; Pv, núcleo paraventricular; MG, núcleo geniculado; Po hy, áreas hipotalámicas posteriores; Hb, habénula; Pi, corteza piriforme.

El enlace de $[^3\text{H}]\text{ucb 30889}$ en las neuronas de los gránulos cerebelosos y en células PC 12 mostraron altos niveles de enlace específico (tabla 1). K_d fue similar al valor medido en la corteza cerebral de rata (42 nM , véase el ejemplo 1). El mismo sitio de enlace específico no pudo detectarse en astrocitos primarios y en un intervalo de líneas celulares relacionadas con el SNC y líneas celulares no neuronales. Abreviación: nd, no detectado.

TABLA 1

Densidad y afinidad del enlace de $[^3\text{H}]\text{ucb 30889}$ en varios tipos de células

Tipo de célula	B_{max}	K_d
Neuronas de gránulos cerebelosos de rata	0,7 pmol/mg proteína	59 nM
Neuronas corticales de ratón		
Astrocitos corticales de ratón	1,4 pmol/mg proteína	34 nM
PC 12	nd	nd
SK-N-SH	79000 sitios/célula	50 nM
NG108-15	nd	nd
N1E-115	nd	nd
HCN-1a	nd	nd
CHO-K1	nd	nd
COS-7	nd	nd
	nd	nd

Se separaron membranas de cerebro de rata por centrifugación diferencial (figura 8). El enlace a LBS (8A), receptores muscarínicos (8B), NMDA (8C) y a los receptores periféricos de la benzodiazepina (8D) se determinó usando $[^3\text{H}]\text{ucb 30889}$, $[^3\text{H}]\text{NMS}$, $[^3\text{H}]\text{MK801}$ ó $[^3\text{H}]\text{PK11195}$, respectivamente. Este estudio muestra que el sitio de enlace del levetiracetam está presente en sinaptosomas brutos (P2), membranas microsomales (P3) y está enriquecido en las vesículas sinápticas (LP2). En contraste, los otros receptores usados no son más abundantes en LP2 comparados con P2 ó P3. P1 es un pelet de baja velocidad que contiene núcleo y grandes residuos.

Para aislar los compartimentos subcelulares de los sinaptosomas brutos se usó un fraccionamiento en gradiente de sacarosa. El LBS se encontró en membranas sinápticas purificadas pero no estaba presente en el pelet de sacarosa 1,2 M que contenía la fracción mitocondrial purificada (figura 9). Como testigo de la pureza de las fracciones subcelulares, también se analizó la distribución de los receptores muscarínicos y de los receptores periféricos de la benzodiazepina. Los datos se expresan como porcentaje del enlace total específico.

Los sinaptosomas brutos (fracción P2) se preincubaron con $[^3\text{H}]\text{ucb 30889}$ 40 nM, a continuación se irradiaron con luz UV y se lavaron. En el minuto 0 se añadió levetiracetam 1 mM. La figura 10B muestra el mismo experimento pero realizado en ausencia de irradiación con luz UV. Estos resultados indican que durante la irradiación con luz UV el radioligando se inserta covalentemente en el dominio de enlace del LBS.

El marcaje de fotoafinidad se realizó en ausencia o en presencia de levetiracetam 1 mM. Las proteínas se resolvieron por SDS-PAGE usando una concentración de acrilamida de 7,5% (p/p) y la radioactividad se evaluó en cada corte del gel. El sitio principal de incorporación se produce a un peso molecular de 97.000 (figura 11).

En este ejemplo se muestra que el sitio de enlace de $[^3\text{H}]\text{ucb 30889}$ en cerebro de rata tiene un único perfil de distribución y no parece que se correlacione con ningún sistema neurotransmisor específico que esté típicamente asociado con la epilepsia. Este nuevo sitio de enlace está restringido a tipos celulares neuronales y a varias áreas del cerebro. Este nuevo radioligando puede usarse como una marca de fotoafinidad y se enlaza covalentemente a una proteína de la membrana de alto peso molecular que está principalmente localizada en las vesículas sinápticas.

Ejemplo 3

El LBS está en SV2A

La caracterización bioquímica del LBS en cerebro de rata conduce a estudios para identificar proteínas potenciales candidatas de LBS para clonarlas y caracterizar el enlace. Sobre la base de la naturaleza de la proteína de la membrana integral, la expresión específica en el cerebro, el tamaño aparente y la localización de las vesículas sinápticas, se analizó la familia de proteínas SV2 como candidata para la localización del LBS. Por consiguiente, las proteínas SV2 se clonaron y ensayaron respecto al enlace de ligandos al LBS.

La SV2A de ser humano se clonó a partir de una biblioteca de cDNA de cerebro fetal de ser humano como un producto 3609 bp por PCR que comprendía la región codificante y regiones flanco significativas del transcrito. Usando un vector que contenía la región que codificaba a SV2A más DNA flanco significativo como fuente, la región codificante se amplificó por PCR sin las regiones flanco. Este producto se clonó en un vector donante GATEWAY (Invitrogen) para facilitar la subclonación. Solamente el uso de un vector de clonación con fuertes sitios para la transcripción directamente aguas arriba del sitio de clonación dio lugar a la clonación con éxito de cDNA de SV2A con solo la región codificante. Esto sugiere que este producto puede ser tóxico para *E. coli*, incluso en pequeñas cantidades. El secuenciado del clon final pDONR GATEWAY de SV2A mostró que había dos mutaciones; una silenciosa y una mutación de Leu a Pro. La mutación no silenciosa se corrigió y la secuenciación que se clonó la secuencia codificante correcta de SV2A de ser humano de longitud completa.

La región codificante de SV2A de ser humano se transfirió desde el vector de clonación pDONR GATEWAY a un vector de expresión pDEST 12.2 GATEWAY. Este vector fue un promotor CMV que dirigía el gen introducido, y un SV40 ori, el cual permite muy altos niveles de replicación en la línea celular COS-7, que contiene el antígeno T grande. Además, la región codificante de SV2A de ser humano se transfirió en el vector de expresión pDEST 40 GATEWAY. Este vector es muy similar al anterior vector 12.2, con un promotor CMV que dirige la expresión de hSV2A, y un SV40 ori y un gen de resistencia a la neomicina.

Los ensayos iniciales de expresión de SV2A usando el vector pDEST 12.2 se realizaron en la línea celular COS-7, la cual había previamente demostrado la expresión con éxito de proteínas SV2. La línea celular COS-7 se ensayó respecto al enlace de [³H]ucb 30889, no observándose ningún enlace por encima del ruido de fondo, y por tanto ninguna presencia significativa medible del sitio de enlace del levetiracetam (LBS). Además, para establecer un conjunto de células PC 13 que expresaban hSV2A en una selección estable de antibióticos se usó un subclon de la línea celular PC 12, PC12a, que tiene un bajo contenido de LBS.

Para transfectar DNA en células COS-7 90% confluentes se usó el reactivo de transfección lipofectamina 2000 (Invitrogen). Asimismo, se usó el mismo reactivo para transfectar el vector que contenía hSV2A en la línea celular PC12a y seleccionar respecto a la resistencia a los antibióticos. Para ensayar la expresión del producto SV2A en células COS-7 transfectadas o en células PC12a transfectadas se usó anticuerpo policlonal anti-SV2 (CalBiochem). Se recogieron lisados de las células COS-7 18 h después de la transfección en un gel SDS-PAGE, se transfirieron a una membrana y se introdujo una sonda con un anticuerpo policlonal contra SV2A, en comparación con membranas brutas de cerebro de ratón (figura 12A). También se muestran lisados de células COS-7 no transfectadas, células PC12a no transfectadas (de bajo contenido de LBS), células PC12b (de alto contenido de LBS) o células PC12a transfectadas con hSV2A (figura 12B). En el testigo COS-7 sin transfectar no se observó ningún marcaje de bandas de proteínas, mientras que las células COS-7 transfectadas muestran múltiples bandas, con la mayor densidad en el intervalo de 80 - 120 kD, quizás debido a los estados múltiples de glicosilación de la proteína expresada. Además, la inmunoreactividad SV2A está presente en las muestras de células PC12b y PC12a/hSV2A, pero prácticamente ausente en las células PC12a de bajo contenido de LBS (figura 12B).

En un experimento de enlace, se midió el enlace específico de ³H-30889 a células COS-7 que habían sido transfectadas con SV2A-12.2, o como testigos, un vector que expresaba β -gal, o células que no habían sido transfectadas (figura 13). Se incubaron por triplicado pocillos de una placa de 24 pocillos con ³H-30889 1 nM (marcado "Caliente") o ³H-30889 más un exceso de levetiracetam frío (50 μ M) (marcado "Caliente + Frío"). Las células se incubaron a 4°C durante 2 horas y a continuación se lavaron rápidamente con PBS enfriada en hielo. Las células se lisaron sobre la placa, se transfirieron a viales de centelleo con un fluido de centelleo y se contaron respecto a la declinación de la emisión de ³H. Estos resultados muestran que las células COS-7 transfectadas con SV2A han adquirido la capacidad de enlazar específicamente ³H-30889. En idénticos experimentos de enlace a células intactas usando células PC12bs, que se sabe expresan el LBS, se ve típicamente una diferencia de 1,5 a 2 veces en CPM entre las muestras "caliente" y "caliente+ frío" en comparación con la diferencia de 5 veces vista en la presente memoria.

Estudios adicionales caracterizaron con más detalle el enlace de ³H-30889 a SV2A expresada en células COS-7. Se transfectaron células COS-7 en una placa de 24 pocillos y se ensayaron como antes respecto a su capacidad enlazante. Se añadieron una serie de concentraciones de levetiracetam, su estereoisómero L060, o 30889 frío con el fin de generar IC50s de estos compuestos contra SV2A expresada en células COS-7 (figura 14). Los IC50s medidos son del mismo orden de magnitud que los vistos en los estudios de LBS en cerebro de rata y con IC50s muy similares medidos para los tres compuestos. Estos resultados indican que SV2A es funcionalmente equivalente con el sitio de enlace del cerebro para levetiracetam que se ha observado en cerebro de rata y subclones PC 12, soportando que la proteína SV2A de las vesículas sinápticas es el sitio de enlace nativo del compuesto antiepiléptico levetiracetam, sugiriendo además una unión entre la función y la modulación de la vesícula sináptica y los trastornos neurológicos. Debido a que la correlación entre la afinidad del enlace al LBS y las propiedades antiataques del levetiracetam y sus compuestos análogos, el enlace de estos compuestos a las proteínas miembro de la familia SV2 es probable que juegue un importante papel en sus propiedades anticonvulsivas.

Aunque la presente invención se ha descrito en detalle con referencia a los ejemplos anteriores, se entiende que las diversas modificaciones pueden hacerse sin apartarse del espíritu de la invención. Por consiguiente, la invención está sólo limitada por las siguientes reivindicaciones.

REIVINDICACIONES

1. Un método para identificar un socio de enlace de una proteína SV2, que comprende:

- a) incubar una proteína o fragmento 2 (SV2) de las vesículas sinápticas con levetiracetam o uno de sus compuestos análogos o derivados y un potencial socio de enlace; y
- b) determinar si el potencial socio de enlace modula el enlace del levetiracetam o uno de sus compuestos análogos o derivados a la proteína o fragmento SV2, identificando de este modo un socio de enlace de la proteína SV2.

2. Un método según la reivindicación 1, en el que el compuesto análogo de levetiracetam es (2S)-2-[4-(3-azidofenil)-2-oxopirrolidin-1-il]butanamida.

3. Un método según la reivindicación 1, en el que la proteína SV2 es SV2A de ser humano.

4. Un método según la reivindicación 1, en el que la proteína SV2A comprende SEQ ID NO: 2.

5. Un método según la reivindicación 1, en el que la proteína o fragmento SV2 está inmovilizada.

6. Un método según la reivindicación 1, en el que la proteína o fragmento SV2 está purificada.

7. Un método según la reivindicación 1, en el que la proteína o fragmento SV2 se expresa en una célula huésped transformada.

8. Un método según la reivindicación 1, en el que el levetiracetam o uno de sus compuestos análogos o derivados está directa o indirectamente marcado.

9. Un método según la reivindicación 1, en el que la proteína o fragmento SV2 se incuba con el levetiracetam o uno de sus compuestos análogos o derivados antes de la adición del potencial socio de enlace.

10. Un método según la reivindicación 1, en el que la proteína o fragmento SV2 se incuba con el levetiracetam o uno de sus compuestos análogos o derivados después de la adición del potencial socio de enlace.

11. Un método según la reivindicación 1, en el que la proteína o fragmento SV2 se incuba con el levetiracetam o uno de sus compuestos análogos o derivados concurrentemente con el potencial socio de enlace.

12. Un método según la reivindicación 1, en el que la proteína o fragmento SV2 se incuba con levetiracetam.

13. Un método según la reivindicación 1 ó 12, en el que el levetiracetam o uno de sus compuestos análogos o derivados está directa o indirectamente marcado.

14. Un método según la reivindicación 13, en el que la marca es una radiomarca.

15. Un método según la reivindicación 14, en el que la radiomarca es ^3H .

16. Un método según la reivindicación 1, en el que el potencial socio de enlace es un agente o un compuesto.

17. Un método para identificar un agente o un compuesto que se enlace a una proteína 2A (SV2A) de las vesículas sinápticas de ser humano, que comprende:

- a) incubar la proteína SV2A con levetiracetam o uno de sus compuestos análogos o derivados en presencia de un agente o de un compuesto;
- b) detectar el enlace de levetiracetam o uno de sus compuestos análogos o derivados a la proteína SV2A; y
- c) determinar si el agente o el compuesto modula el enlace del levetiracetam o de uno de sus compuestos análogos o derivados a la proteína SV2A en comparación con un testigo, identificando de este modo un agente o un compuesto que se enlaza a la proteína SV2A.

Figura 1

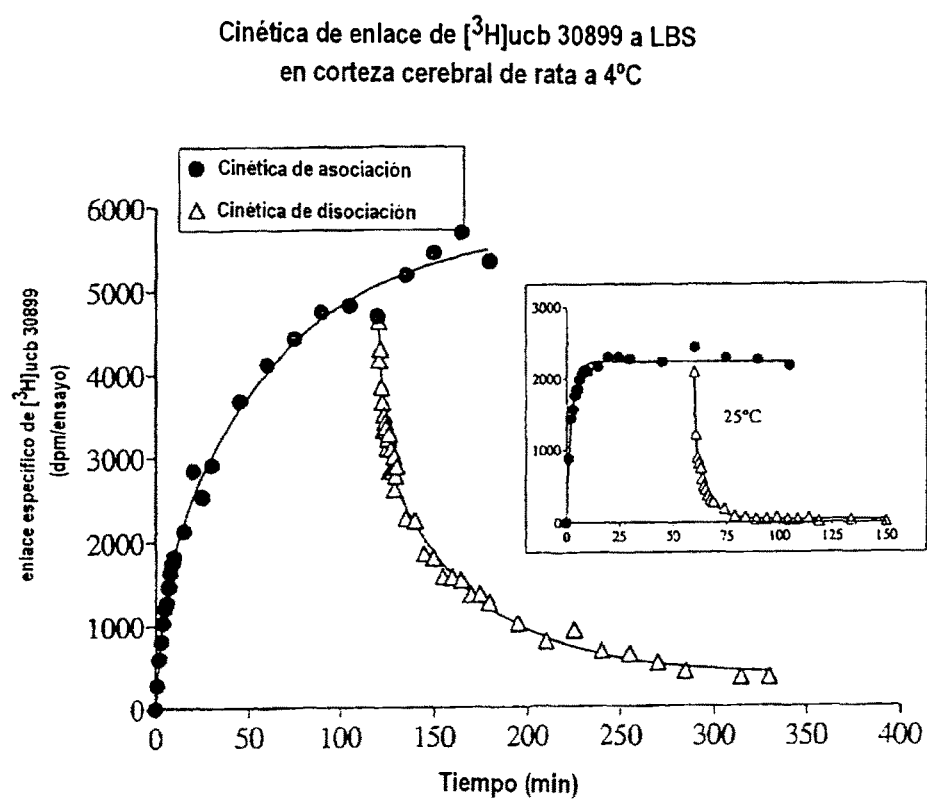


Figura 2

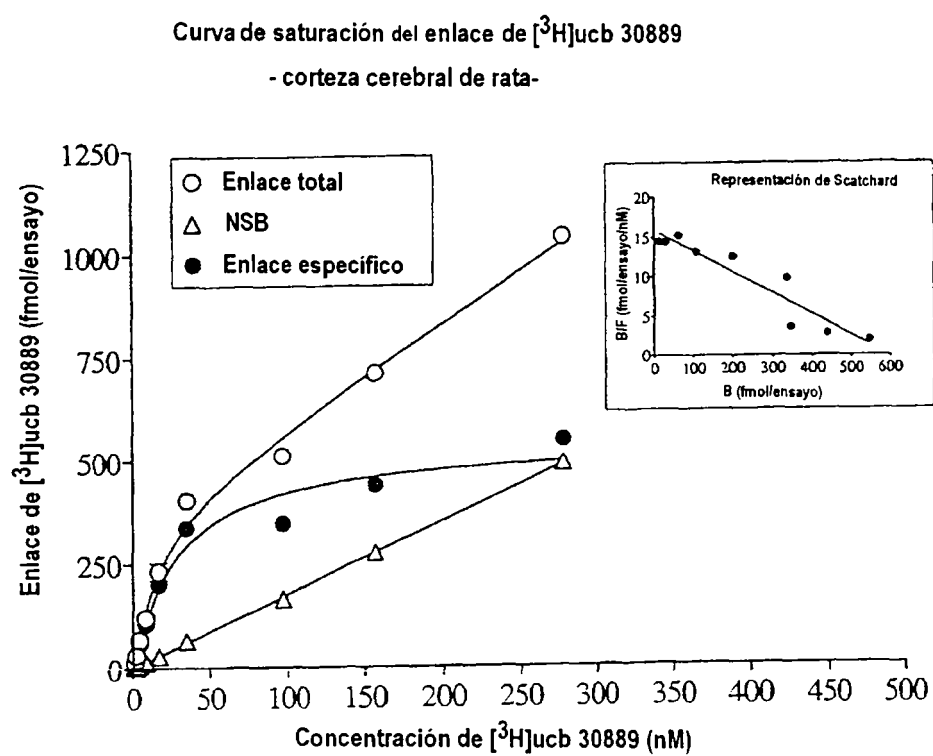


Figura 3

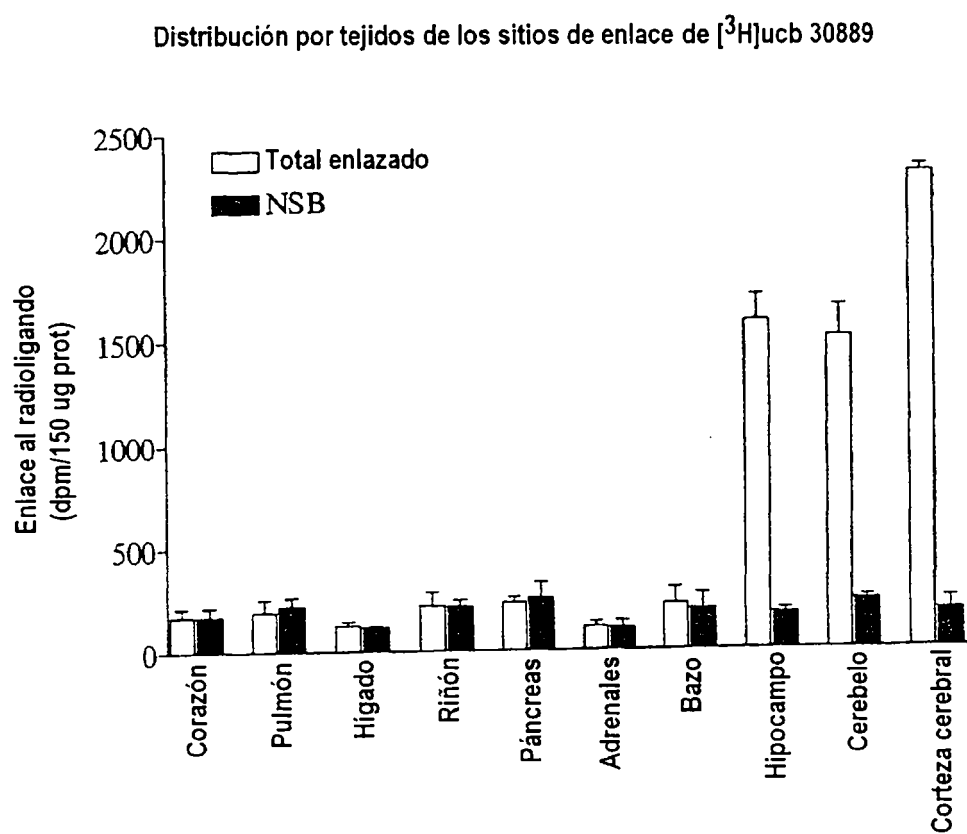


Figura 4

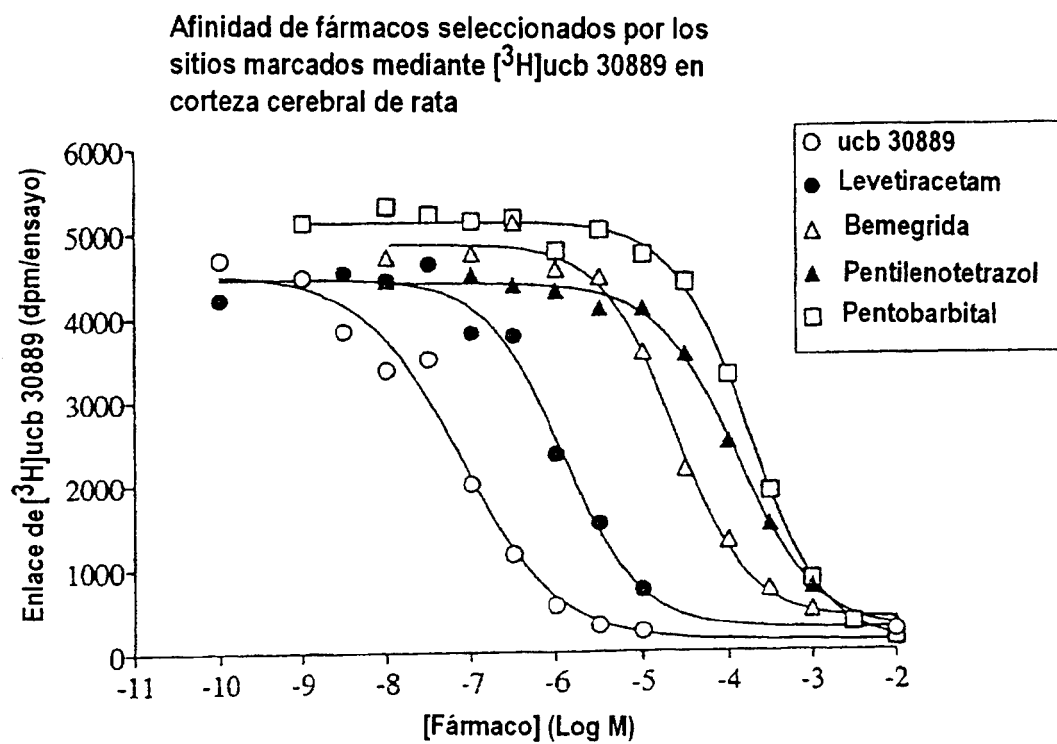


Figura 5

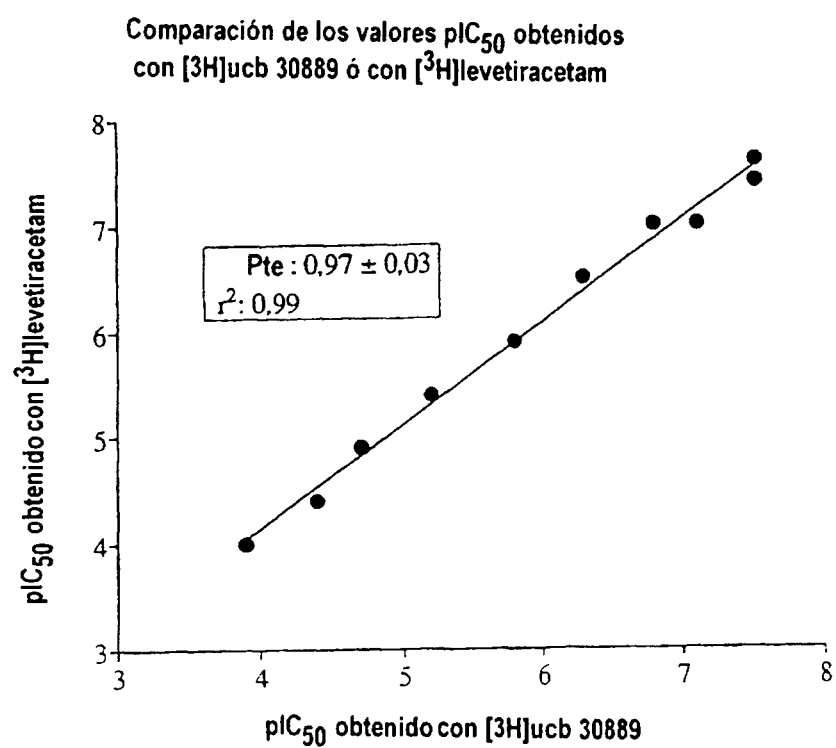


Figura 6

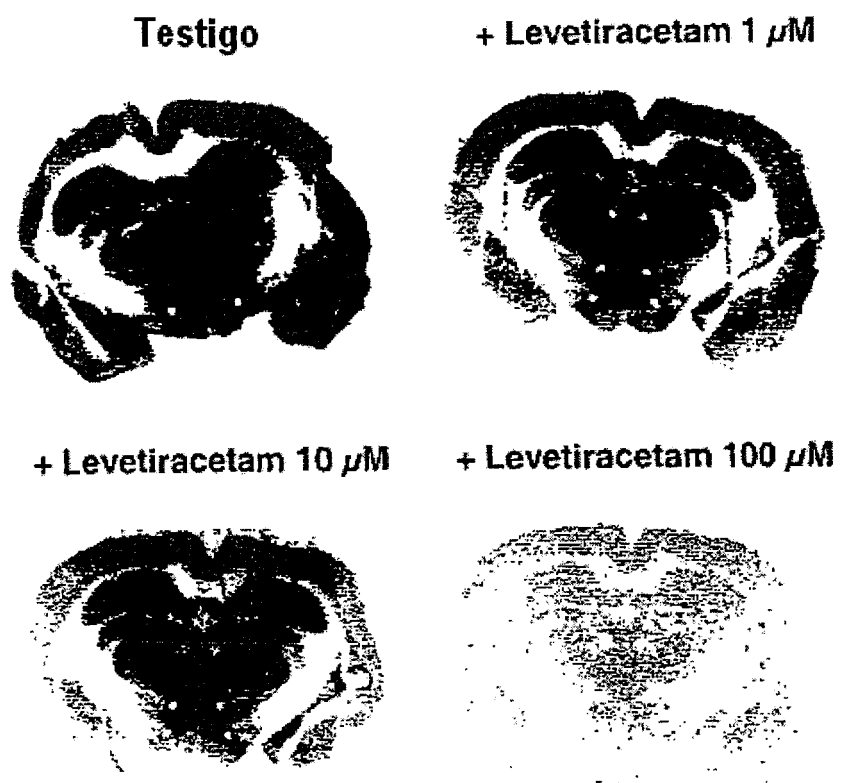


Figura 7

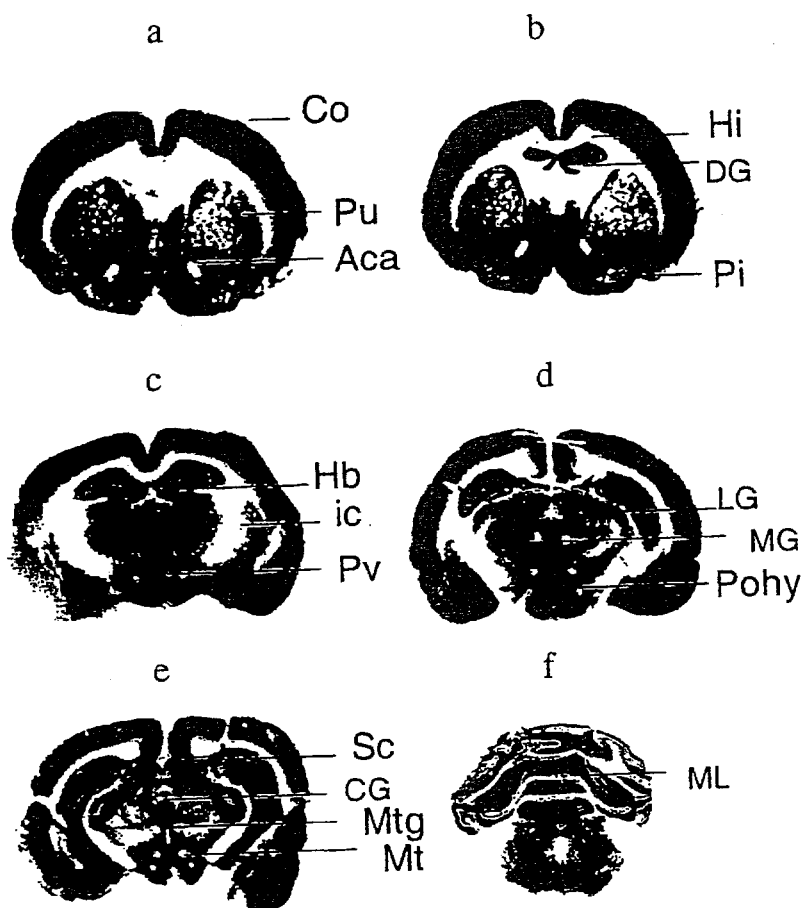


Figura 8

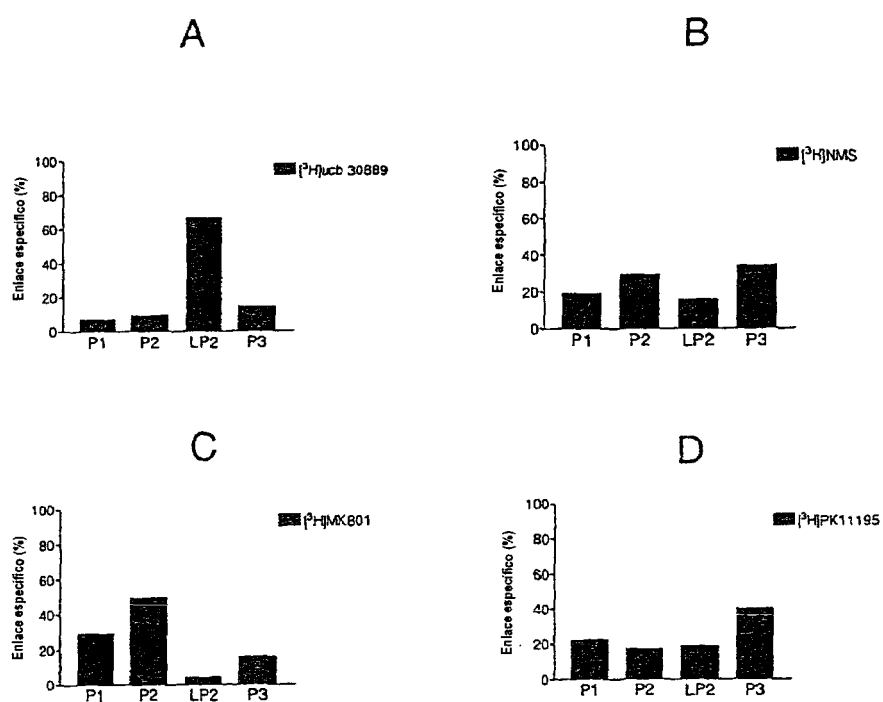


Figura 9

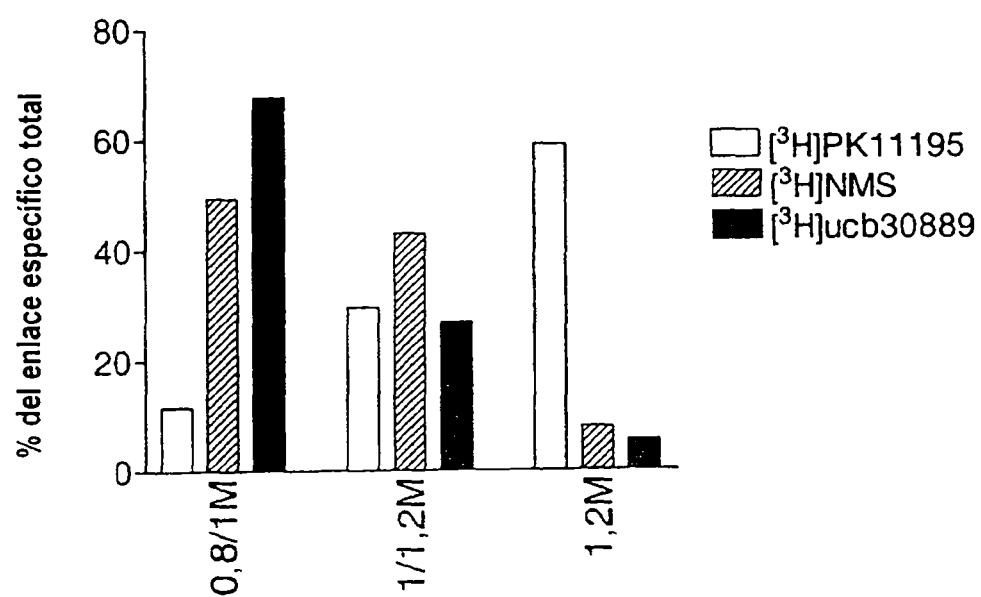


Figura 10

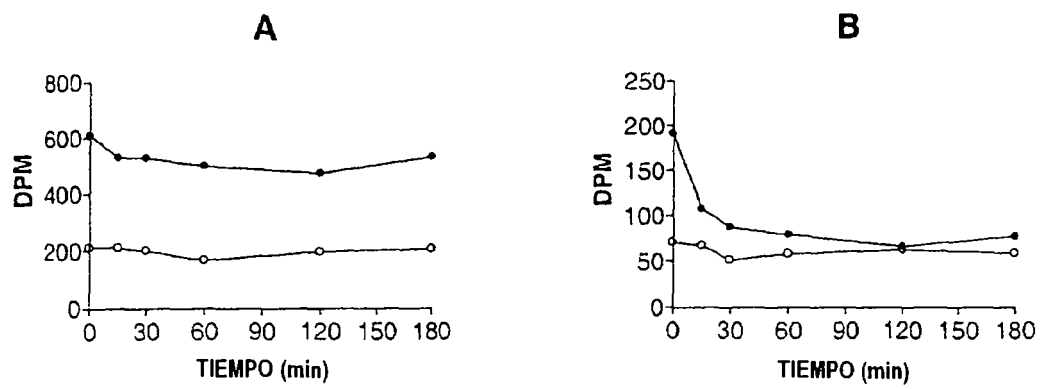


Figura 11

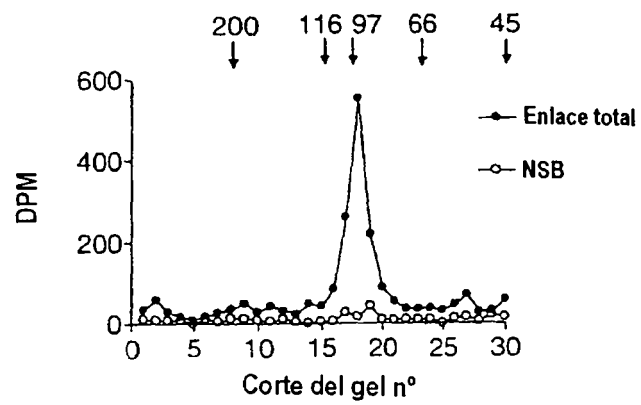


Figura 12 (A)

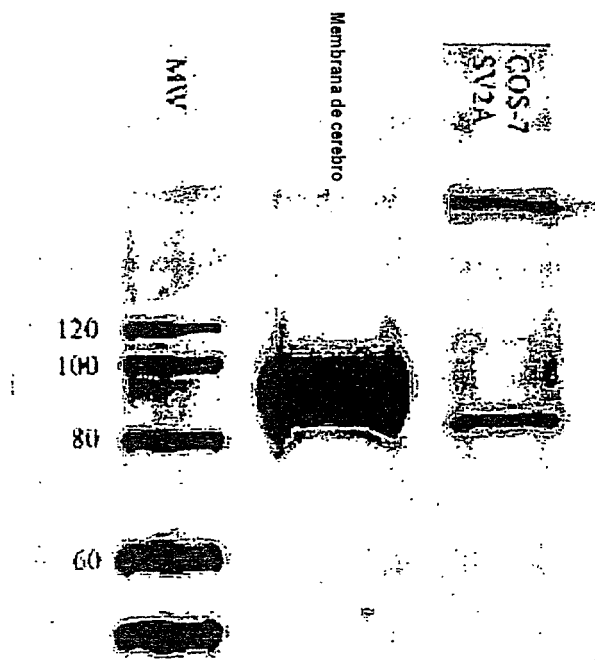


Figura 12 (B)

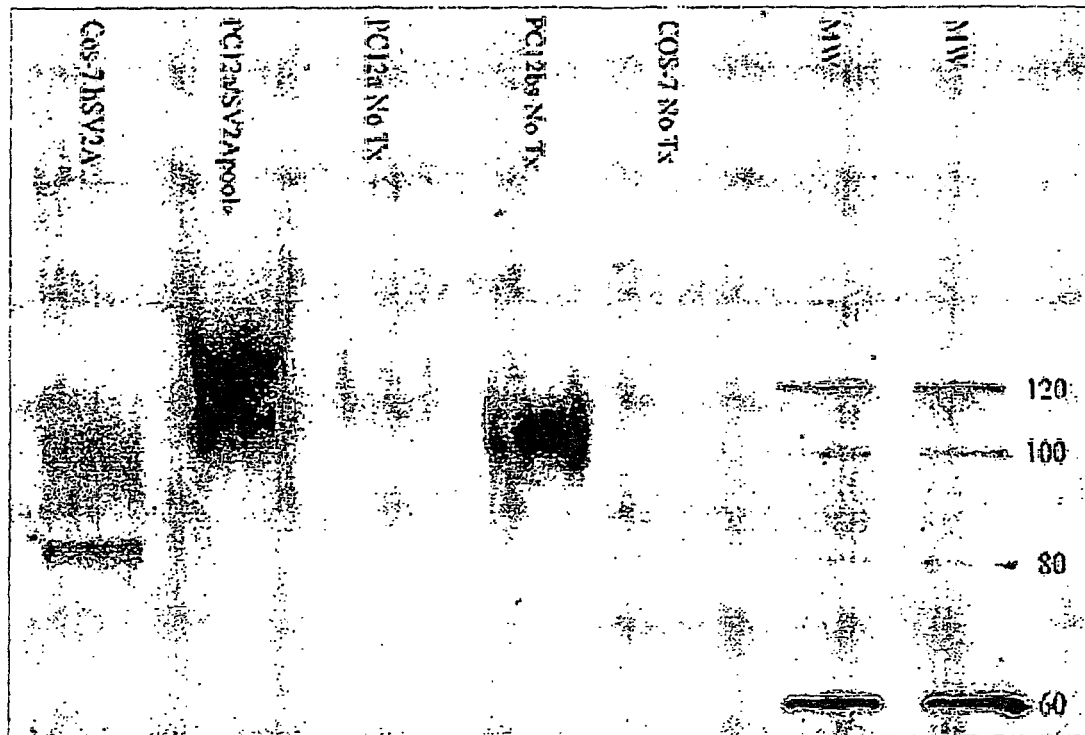


Figura 13

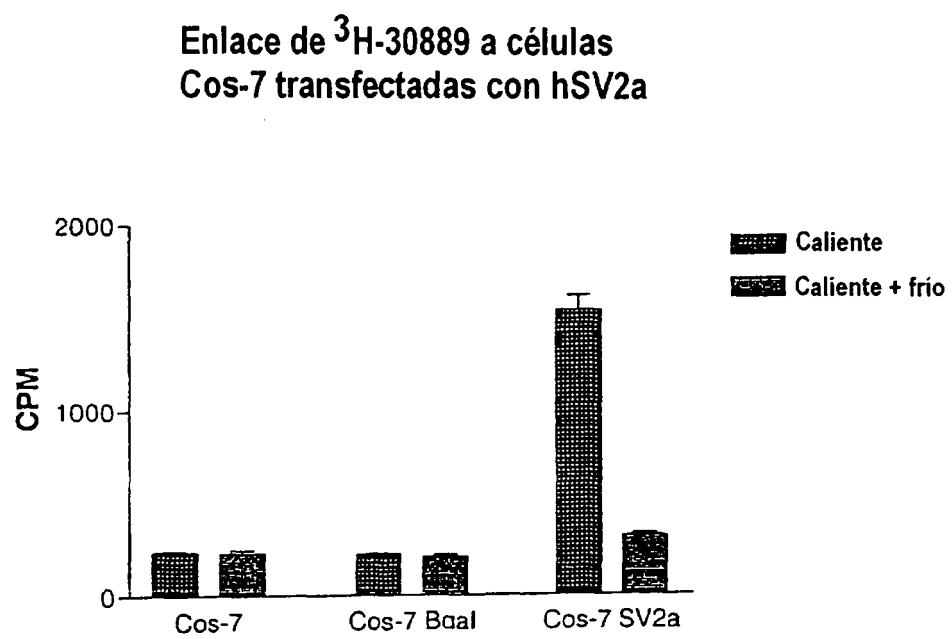


Figura 14

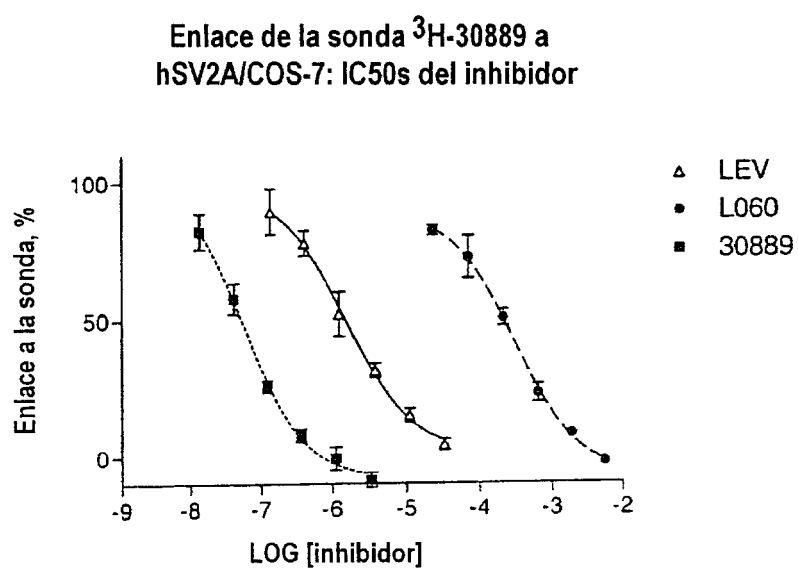
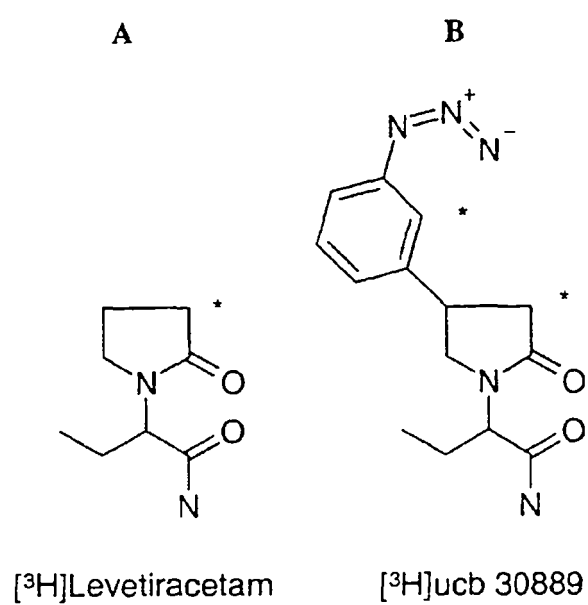


Figura 15



ES 2 311 667 T3

LISTA DE SECUENCIAS

UCB, S.A.
LYNCH, Berkley
NOCKA, Karl
FUKS, Bruno

Métodos para la identificación de agentes para el tratamiento de ataques, trastornos neurológicos, endocrinopatías y trastornos hormonales

17.60.EP

16

PatentIn version 3.1

1
2229
ADN
Homo sapiens

CDS
(1)..(2229)

atg gaa gag ggc ttc cga gac cgg gca gct ttc atc cgt ggg gcc aaa	48
Met Glu Glu Gly Phe Arg Asp Arg Ala Ala Phe Ile Arg Gly Ala Lys	
1 5 10 15	

gac att gct aag gaa gtc aaa aag cat gcg gcc aag aag gtg gtg aag	96
Asp Ile Ala Lys Glu Val Lys Lys His Ala Ala Lys Lys Val Val Lys	
20 25 30	

ggc ctg gac aga gtc cag gac gaa tat tcc cga aga tcg tac tcc cgc	144
---	-----

ES 2 311 667 T3

	Gly	Leu	Asp	Arg	Val	Gln	Asp	Glu	Tyr	Ser	Arg	Arg	Ser	Tyr	Ser	Arg	
			35					40					45				
5	ttt	gag	gag	gag	gat	gat	gat	gat	gac	ttc	cct	gct	ccc	agt	gat	ggg	192
	Phe	Glu	Glu	Glu	Asp	Asp	Asp	Asp	Asp	Phe	Pro	Ala	Pro	Ser	Asp	Gly	
		50					55					60					
10	tat	tac	cga	gga	gaa	ggg	acc	cag	gat	gag	gag	gaa	ggg	ggg	gca	tcc	240
	Tyr	Tyr	Arg	Gly	Glu	Gly	Thr	Gln	Asp	Glu	Glu	Glu	Gly	Gly	Ala	Ser	
	65					70					75					80	
15	agt	gat	gct	act	gag	ggc	cat	gac	gag	gat	gat	gag	atc	tat	gaa	ggg	288
	Ser	Asp	Ala	Thr	Glu	Gly	His	Asp	Glu	Asp	Asp	Glu	Ile	Tyr	Glu	Gly	
					85				90						95		
20	gaa	tat	cag	ggc	att	ccc	cgg	gca	gag	tct	ggg	ggc	aaa	ggc	gag	cgg	336
	Glu	Tyr	Gln	Gly	Ile	Pro	Arg	Ala	Glu	Ser	Gly	Gly	Lys	Gly	Glu	Arg	
				100					105					110			
25	atg	gca	gat	ggg	gcg	ccc	ctg	gct	gga	gta	agg	ggg	ggc	ttg	agt	gat	384
	Met	Ala	Asp	Gly	Ala	Pro	Leu	Ala	Gly	Val	Arg	Gly	Gly	Leu	Ser	Asp	
				115				120					125				
30	ggg	gag	ggg	ccc	cct	ggg	ggc	cgg	ggg	gag	gca	caa	cga	cgg	aaa	gaa	432
	Gly	Glu	Gly	Pro	Pro	Gly	Gly	Arg	Gly	Glu	Ala	Gln	Arg	Arg	Lys	Glu	
		130				135						140					
35	cga	gaa	gaa	ctg	gcc	caa	cag	tat	gaa	gcc	atc	cta	cgg	gag	tgt	ggc	480
	Arg	Glu	Glu	Leu	Ala	Gln	Gln	Tyr	Glu	Ala	Ile	Leu	Arg	Glu	Cys	Gly	
					145		150				155					160	
40	cac	ggc	cgc	ttc	cag	tgg	aca	ctg	tat	ttt	gtg	ctt	ggg	ctg	gcg	ctg	528
	His	Gly	Arg	Phe	Gln	Trp	Thr	Leu	Tyr	Phe	Val	Leu	Gly	Leu	Ala	Leu	
					165					170					175		
45	atg	gct	gac	ggg	gtg	gag	gtc	ttt	gtg	gtg	ggc	ttc	gtg	ctg	ccc	agc	576
	Met	Ala	Asp	Gly	Val	Glu	Val	Phe	Val	Val	Gly	Phe	Val	Leu	Pro	Ser	
				180				185						190			
50	gct	gag	aaa	gac	atg	tgc	ctg	tcc	gac	tcc	aac	aaa	ggc	atg	cta	ggc	624
	Ala	Glu	Lys	Asp	Met	Cys	Leu	Ser	Asp	Ser	Asn	Lys	Gly	Met	Leu	Gly	
			195				200						205				
55	ctc	atc	gtc	tac	ctg	ggc	atg	atg	gtg	gga	gcc	ttc	ctc	tgg	gga	ggg	672
	Leu	Ile	Val	Tyr	Leu	Gly	Met	Met	Val	Gly	Ala	Phe	Leu	Trp	Gly	Gly	
				210			215					220					
60	ctg	gct	gac	cgg	ctg	ggg	cgg	agg	cag	tgt	ctg	ctc	atc	tgc	ctc	tca	720
	Leu	Ala	Asp	Arg	Leu	Gly	Arg	Arg	Gln	Cys	Leu	Leu	Ile	Ser	Leu	Ser	
					225		230				235					240	
65	gtc	aac	agc	gtc	ttc	gcc	ttc	ttc	tca	tct	ttt	gtc	cag	ggg	tac	ggc	768
	Val	Asn	Ser	Val	Phe	Ala	Phe	Phe	Ser	Ser	Phe	Val	Gln	Gly	Tyr	Gly	
					245				250						255		
70	act	ttc	ctc	ttc	tgc	cgc	cta	ctt	tct	ggg	gtt	ggg	att	gga	ggg	tcc	816
	Thr	Phe	Leu	Phe	Cys	Arg	Leu	Leu	Ser	Gly	Val	Gly	Ile	Gly	Gly	Ser	
				260				265						270			
75	atc	ccc	att	gtc	ttc	tcc	tat	ttc	tcc	gag	ttt	ctg	gcc	cag	gag	aaa	864
	Ile	Pro	Ile	Val	Phe	Ser	Tyr	Phe	Ser	Glu	Phe	Leu	Ala	Gln	Glu	Lys	

ES 2 311 667 T3

	275	280	285	
5	cga ggg gag cat ttg agc tgg ctc tgc atg ttt tgg atg att ggt ggc Arg Gly Glu His Leu Ser Trp Leu Cys Met Phe Trp Met Ile Gly Gly 290 295 300			912
10	gtg tac gca gct gct atg gcc tgg gcc atc atc ccc cac tat ggg tgg Val Tyr Ala Ala Ala Met Ala Trp Ala Ile Ile Pro His Tyr Gly Trp 305 310 315 320			960
15	agt ttt cag atg ggt tct gcc tac cag ttc cac agc tgg agg gtc ttc Ser Phe Gln Met Gly Ser Ala Tyr Gln Phe His Ser Trp Arg Val Phe 325 330 335			1008
20	gtc ctc gtc tgc gcc ttt cct tct gtg ttt gcc att ggg gct ctg acc Val Leu Val Cys Ala Phe Pro Ser Val Phe Ala Ile Gly Ala Leu Thr 340 345 350			1056
25	acg cag cct gag agc ccc cgt ttc ttc cta gag aat gga aag cat gat Thr Gln Pro Glu Ser Pro Arg Phe Phe Leu Glu Asn Gly Lys His Asp 355 360 365			1104
30	gag gcc tgg atg gtg ctg aag cag gtc cat gat acc aac atg cga gcc Glu Ala Trp Met Val Leu Lys Gln Val His Asp Thr Asn Met Arg Ala 370 375 380			1152
35	aaa gga cat cct gag cga gtg ttc tca gta acc cac att aag acg att Lys Gly His Pro Glu Arg Val Phe Ser Val Thr His Ile Lys Thr Ile 385 390 395 400			1200
40	cat cag gag gat gaa ttg att gag atc cag tcg gac aca ggg acc tgg His Gln Glu Asp Glu Leu Ile Glu Ile Gln Ser Asp Thr Gly Thr Trp 405 410 415			1248
45	tac cag cgc tgg ggg gtc cgg gcc ttg agc cta ggg ggg cag gtt tgg Tyr Gln Arg Trp Gly Val Arg Ala Leu Ser Leu Gly Gly Gln Val Trp 420 425 430			1296
50	ggg aat ttt ctc tcc tgt ttt ggt ccc gaa tat cgg cgc atc act ctg Gly Asn Phe Leu Ser Cys Phe Gly Pro Glu Tyr Arg Arg Ile Thr Leu 435 440 445			1344
55	atg atg atg ggt gtg tgg ttc acc atg tca ttc agc tac tat ggc ctg Met Met Met Gly Val Trp Phe Thr Met Ser Phe Ser Tyr Tyr Gly Leu 450 455 460			1392
60	acc gtc tgg ttt cct gac atg atc cgc cat ctc cag gca gtg gac tac Thr Val Trp Phe Pro Asp Met Ile Arg His Leu Gln Ala Val Asp Tyr 465 470 475 480			1440
65	gca tcc cgc acc aaa gtg ttc ccc ggg gag cgc gta gag cat gta act Ala Ser Arg Thr Lys Val Phe Pro Gly Glu Arg Val Glu His Val Thr 485 490 495			1488
70	ttt aac ttc acg ttg gag aat cag atc cac cga ggc ggg cag tac ttc Phe Asn Phe Thr Leu Glu Asn Gln Ile His Arg Gly Gly Gln Tyr Phe 500 505 510			1536
75	aat gac aag ttc att ggg ctg cgg ctc aag tca gtg tcc ttt gag gat Asn Asp Lys Phe Ile Gly Leu Arg Leu Lys Ser Val Ser Phe Glu Asp 515 520 525			1584

ES 2 311 667 T3

	tcc	ctg	ttt	gaa	gag	tgt	tat	ttt	gag	gat	gtc	aca	tcc	agc	aac	acg	1632
	Ser	Leu	Phe	Glu	Glu	Cys	Tyr	Phe	Glu	Asp	Val	Thr	Ser	Ser	Asn	Thr	
	530						535					540					
5	ttt	ttc	cgc	aac	tgc	aca	ttc	atc	aac	act	gtg	ttc	tat	aac	act	gac	1680
	Phe	Phe	Arg	Asn	Cys	Thr	Phe	Ile	Asn	Thr	Val	Phe	Tyr	Asn	Thr	Asp	
	545					550					555					560	
10	ctg	ttc	gag	tac	aag	ttt	gtg	aac	agc	cgt	ctg	ata	aac	agt	aca	ttc	1728
	Leu	Phe	Glu	Tyr	Lys	Phe	Val	Asn	Ser	Arg	Leu	Ile	Asn	Ser	Thr	Phe	
					565					570					575		
15	ctg	cac	aac	aag	gag	ggc	tgc	ccg	cta	gac	gtg	aca	ggg	acg	ggc	gaa	1776
	Leu	His	Asn	Lys	Glu	Gly	Cys	Pro	Leu	Asp	Val	Thr	Gly	Thr	Gly	Glu	
				580				585						590			
20	ggg	gcc	tac	atg	gta	tac	ttt	gtg	agc	ttc	ctg	ggg	aca	ctg	gca	gtg	1824
	Gly	Ala	Tyr	Met	Val	Tyr	Phe	Val	Ser	Phe	Leu	Gly	Thr	Leu	Ala	Val	
			595				600						605				
25	ctt	cct	ggg	aat	atc	gtg	tct	gcc	ctg	ctc	atg	gac	aag	atc	ggc	agg	1872
	Leu	Pro	Gly	Asn	Ile	Val	Ser	Ala	Leu	Leu	Met	Asp	Lys	Ile	Gly	Arg	
		610					615					620					
30	ctc	aga	atg	ctt	gct	ggc	tcc	agc	gtg	atg	tcc	tgt	gtc	tcc	tgc	ttc	1920
	Leu	Arg	Met	Leu	Ala	Gly	Ser	Ser	Val	Met	Ser	Cys	Val	Ser	Cys	Phe	
		625				630					635					640	
35	ttc	ctg	tct	ttt	ggg	aac	agt	gag	tgc	gcc	atg	atc	gct	ctg	ctc	tgc	1968
	Phe	Leu	Ser	Phe	Gly	Asn	Ser	Glu	Ser	Ala	Met	Ile	Ala	Leu	Leu	Cys	
				645						650					655		
40	ctt	ttt	ggc	ggg	gtc	agc	att	gca	tcc	tgg	aat	gcg	ctg	gac	gtg	ttg	2016
	Leu	Phe	Gly	Gly	Val	Ser	Ile	Ala	Ala	Trp	Asn	Ala	Leu	Asp	Val	Leu	
				660					665					670			
45	act	gtt	gaa	ctc	tac	ccc	tca	gac	aag	agg	acc	aca	gct	ttt	ggc	ttc	2064
	Thr	Val	Glu	Leu	Tyr	Pro	Ser	Asp	Lys	Arg	Thr	Thr	Ala	Phe	Gly	Phe	
			675					680					685				
50	ctg	aat	gcc	ctg	tgt	aag	ctg	gca	gct	gtg	ctg	ggg	atc	agc	atc	ttc	2112
	Leu	Asn	Ala	Leu	Cys	Lys	Leu	Ala	Ala	Val	Leu	Gly	Ile	Ser	Ile	Phe	
		690					695					700					
55	aca	tcc	ttc	gtg	gga	atc	acc	aag	gct	gca	ccc	atc	ctc	ttt	gcc	tca	2160
	Thr	Ser	Phe	Val	Gly	Ile	Thr	Lys	Ala	Ala	Pro	Ile	Leu	Phe	Ala	Ser	
		705				710					715					720	
60	gct	gcc	ctt	gcc	ctt	ggc	agc	tct	ctg	gcc	ctg	aag	ctg	cct	gag	acc	2208
	Ala	Ala	Leu	Ala	Leu	Gly	Ser	Ser	Leu	Ala	Leu	Lys	Leu	Pro	Glu	Thr	
				725						730					735		
65	cgg	ggg	cag	gtg	ctg	cag	tga										2229
	Arg	Gly	Gln	Val	Leu	Gln											
				740													

ES 2 311 667 T3

2

742

PRT

5 *Homo sapiens*

10	Met Glu Glu Gly Phe Arg Asp Arg Ala Ala Phe Ile Arg Gly Ala Lys	1 5 10 15
15	Asp Ile Ala Lys Glu Val Lys Lys His Ala Ala Lys Lys Val Val Lys	20 25 30
20	Gly Leu Asp Arg Val Gln Asp Glu Tyr Ser Arg Arg Ser Tyr Ser Arg	35 40 45
25	Phe Glu Glu Glu Asp Asp Asp Asp Asp Phe Pro Ala Pro Ser Asp Gly	50 55 60
30	Tyr Tyr Arg Gly Glu Gly Thr Gln Asp Glu Glu Glu Gly Gly Ala Ser	65 70 75 80
35	Ser Asp Ala Thr Glu Gly His Asp Glu Asp Asp Glu Ile Tyr Glu Gly	85 90 95
40	Glu Tyr Gln Gly Ile Pro Arg Ala Glu Ser Gly Gly Lys Gly Glu Arg	100 105 110
45	Met Ala Asp Gly Ala Pro Leu Ala Gly Val Arg Gly Gly Leu Ser Asp	115 120 125
50	Gly Glu Gly Pro Pro Gly Gly Arg Gly Glu Ala Gln Arg Arg Lys Glu	130 135 140
55	Arg Glu Glu Leu Ala Gln Gln Tyr Glu Ala Ile Leu Arg Glu Cys Gly	145 150 155 160
60	His Gly Arg Phe Gln Trp Thr Leu Tyr Phe Val Leu Gly Leu Ala Leu	165 170 175
65	Met Ala Asp Gly Val Glu Val Phe Val Val Gly Phe Val Leu Pro Ser	180 185 190
	Ala Glu Lys Asp Met Cys Leu Ser Asp Ser Asn Lys Gly Met Leu Gly	195 200 205

ES 2 311 667 T3

Leu Ile Val Tyr Leu Gly Met Met Val Gly Ala Phe Leu Trp Gly Gly
 210 215 220
 5
 Leu Ala Asp Arg Leu Gly Arg Arg Gln Cys Leu Leu Ile Ser Leu Ser
 225 230 235 240
 10
 Val Asn Ser Val Phe Ala Phe Phe Ser Ser Phe Val Gln Gly Tyr Gly
 245 250 255
 15
 Thr Phe Leu Phe Cys Arg Leu Leu Ser Gly Val Gly Ile Gly Gly Ser
 260 265 270
 20
 Ile Pro Ile Val Phe Ser Tyr Phe Ser Glu Phe Leu Ala Gln Glu Lys
 275 280 285
 25
 Arg Gly Glu His Leu Ser Trp Leu Cys Met Phe Trp Met Ile Gly Gly
 290 295 300
 30
 Val Tyr Ala Ala Ala Met Ala Trp Ala Ile Ile Pro His Tyr Gly Trp
 305 310 315 320
 35
 Ser Phe Gln Met Gly Ser Ala Tyr Gln Phe His Ser Trp Arg Val Phe
 325 330 335
 40
 Val Leu Val Cys Ala Phe Pro Ser Val Phe Ala Ile Gly Ala Leu Thr
 340 345 350
 45
 Thr Gln Pro Glu Ser Pro Arg Phe Phe Leu Glu Asn Gly Lys His Asp
 355 360 365
 50
 Glu Ala Trp Met Val Leu Lys Gln Val His Asp Thr Asn Met Arg Ala
 370 375 380
 55
 Lys Gly His Pro Glu Arg Val Phe Ser Val Thr His Ile Lys Thr Ile
 385 390 395 400
 60
 His Gln Glu Asp Glu Leu Ile Glu Ile Gln Ser Asp Thr Gly Thr Trp
 405 410 415
 65
 Tyr Gln Arg Trp Gly Val Arg Ala Leu Ser Leu Gly Gly Gln Val Trp
 420 425 430
 Gly Asn Phe Leu Ser Cys Phe Gly Pro Glu Tyr Arg Arg Ile Thr Leu
 435 440 445
 Met Met Met Gly Val Trp Phe Thr Met Ser Phe Ser Tyr Tyr Gly Leu

ES 2 311 667 T3

	450	455	460
5	Thr Val Trp Phe Pro Asp Met Ile Arg His Leu Gln Ala Val Asp Tyr 465 470 475 480		
10	Ala Ser Arg Thr Lys Val Phe Pro Gly Glu Arg Val Glu His Val Thr 485 490 495		
15	Phe Asn Phe Thr Leu Glu Asn Gln Ile His Arg Gly Gly Gln Tyr Phe 500 505 510		
20	Asn Asp Lys Phe Ile Gly Leu Arg Leu Lys Ser Val Ser Phe Glu Asp 515 520 525		
25	Ser Leu Phe Glu Glu Cys Tyr Phe Glu Asp Val Thr Ser Ser Asn Thr 530 535 540		
30	Phe Phe Arg Asn Cys Thr Phe Ile Asn Thr Val Phe Tyr Asn Thr Asp 545 550 555 560		
35	Leu Phe Glu Tyr Lys Phe Val Asn Ser Arg Leu Ile Asn Ser Thr Phe 565 570 575		
40	Leu His Asn Lys Glu Gly Cys Pro Leu Asp Val Thr Gly Thr Gly Glu 580 585 590		
45	Gly Ala Tyr Met Val Tyr Phe Val Ser Phe Leu Gly Thr Leu Ala Val 595 600 605		
50	Leu Pro Gly Asn Ile Val Ser Ala Leu Leu Met Asp Lys Ile Gly Arg 610 615 620		
55	Leu Arg Met Leu Ala Gly Ser Ser Val Met Ser Cys Val Ser Cys Phe 625 630 635 640		
60	Phe Leu Ser Phe Gly Asn Ser Glu Ser Ala Met Ile Ala Leu Leu Cys 645 650 655		
65	Leu Phe Gly Gly Val Ser Ile Ala Ser Trp Asn Ala Leu Asp Val Leu 660 665 670		
70	Thr Val Glu Leu Tyr Pro Ser Asp Lys Arg Thr Thr Ala Phe Gly Phe 675 680 685		
75	Leu Asn Ala Leu Cys Lys Leu Ala Ala Val Leu Gly Ile Ser Ile Phe 690 695 700		
80	Thr Ser Phe Val Gly Ile Thr Lys Ala Ala Pro Ile Leu Phe Ala Ser 705 710 715 720		
85	Ala Ala Leu Ala Leu Gly Ser Ser Leu Ala Leu Lys Leu Pro Glu Thr 725 730 735		
90	Arg Gly Gln Val Leu Gln 740		

ES 2 311 667 T3

3

2052

ADN

5 *Homo sapiens*

CDS

(1)..(2052)

10

15

20

25

30

35

40

45

50

55

60

65

atg gat gac tac aag tat cag gac aat tat ggg ggc tat gct ccc agt	48
Met Asp Asp Tyr Lys Tyr Gln Asp Asn Tyr Gly Gly Tyr Ala Pro Ser	
1 5 10 15	
gat ggc tat tac cgc ggc aat gag tcc aac cca gaa gaa gat gca cag	96
Asp Gly Tyr Tyr Arg Gly Asn Glu Ser Asn Pro Glu Glu Asp Ala Gln	
20 25 30	
agt gat gtc acc gaa ggc cat gat gag gaa gac gag atc tat gag ggc	144
Ser Asp Val Thr Glu Gly His Asp Glu Glu Asp Glu Ile Tyr Glu Gly	
35 40 45	
gag tac cag ggt atc cct cac cca gat gat gtc aag gcc aag cag gcc	192
Glu Tyr Gln Gly Ile Pro His Pro Asp Asp Val Lys Ala Lys Gln Ala	
50 55 60	
aag atg gcg ccc tcc aga atg gac agc ctt cgg ggc cag aca gac ctg	240
Lys Met Ala Pro Ser Arg Met Asp Ser Leu Arg Gly Gln Thr Asp Leu	
65 70 75 80	
atg gct gag agg ctg gaa gat gag gag cag ttg gcc cac cag tac gag	288
Met Ala Glu Arg Leu Glu Asp Glu Glu Gln Leu Ala His Gln Tyr Glu	
85 90 95	
acc atc atg gat gag tgt ggc cat ggc cgc ttc cag tgg atc ctc ttt	336
Thr Ile Met Asp Glu Cys Gly His Gly Arg Phe Gln Trp Ile Leu Phe	

ES 2 311 667 T3

	100	105	110	
5	ttc gtc ttg ggt ttg gcc ctg atg gcc gat ggg gtg gaa gtg ttc gtg Phe Val Leu Gly Leu Ala Leu Met Ala Asp Gly Val Glu Val Phe Val 115 120 125	384		
	gtg agt ttt gcc ctg ccc agt gca gag aag gac atg tgt ctg tcc agt Val Ser Phe Ala Leu Pro Ser Ala Glu Lys Asp Met Cys Leu Ser Ser 130 135 140	432		
10	tcc aaa aaa gga atg cta ggg atg ata gtc tac ttg gga atg atg gcg Ser Lys Lys Gly Met Leu Gly Met Ile Val Tyr Leu Gly Met Met Ala 145 150 155 160	480		
15	ggc gcc ttc atc ctg gga ggc ctg gct gat aag ctg gga agg aag cga Gly Ala Phe Ile Leu Gly Gly Leu Ala Asp Lys Leu Gly Arg Lys Arg 165 170 175	528		
20	gtc ctc agc atg tct ctg gcc gtc aat gcc tcc ttc gcc tcc ctc tct Val Leu Ser Met Ser Leu Ala Val Asn Ala Ser Phe Ala Ser Leu Ser 180 185 190	576		
	tcc ttc gtg cag gga tat gga gcc ttc ctc ttc tgc cga ctc atc tca Ser Phe Val Gln Gly Tyr Gly Ala Phe Leu Phe Cys Arg Leu Ile Ser 195 200 205	624		
25	ggc atc ggt att ggg ggt gct cta ccg att gtt ttt gcc tat ttt tct Gly Ile Gly Ile Gly Gly Ala Leu Pro Ile Val Phe Ala Tyr Phe Ser 210 215 220	672		
30	gaa ttc ttg tct cgg gag aag cga gga gaa cac ctc agt tgg ctg ggc Glu Phe Leu Ser Arg Glu Lys Arg Gly Glu His Leu Ser Trp Leu Gly 225 230 235 240	720		
	atc ttc tgg atg act ggg ggc ctg tac gca tct gcc atg gcc tgg agc Ile Phe Trp Met Thr Gly Gly Leu Tyr Ala Ser Ala Met Ala Trp Ser 245 250 255	768		
35	atc atc cca cac tat ggc tgg ggc ttc agc atg ggg acc aat tac cac Ile Ile Pro His Tyr Gly Trp Gly Phe Ser Met Gly Thr Asn Tyr His 260 265 270	816		
40	ttc cat agc tgg aga gtg ttt gtc atc gtc tgt gct ctg ccc tgc acc Phe His Ser Trp Arg Val Phe Val Ile Val Cys Ala Leu Pro Cys Thr 275 280 285	864		
	gtg tcc atg gtg gcc ctg aag ttc atg cca gag agc cca agg ttt ctg Val Ser Met Val Ala Leu Lys Phe Met Pro Glu Ser Pro Arg Phe Leu 290 295 300	912		
45	cta gag atg ggc aaa cat gat gaa gcc tgg atg att ctc aag caa gtc Leu Glu Met Gly Lys His Asp Glu Ala Trp Met Ile Leu Lys Gln Val 305 310 315 320	960		
50	cat gac acc aac atg aga gct aag ggg acc cca gag aaa gtg ttc acg His Asp Thr Asn Met Arg Ala Lys Gly Thr Pro Glu Lys Val Phe Thr 325 330 335	1008		
55	gtt tcc aac atc aaa act ccc aag caa atg gat gaa ttc att gag atc Val Ser Asn Ile Lys Thr Pro Lys Gln Met Asp Glu Phe Ile Glu Ile 340 345 350	1056		

ES 2 311 667 T3

	caa agt tca aca gga acc tgg tac cag cgc tgg ctg gtc aga ttc aag	1104
	Gln Ser Ser Thr Gly Thr Trp Tyr Gln Arg Trp Leu Val Arg Phe Lys	
	355 360 365	
5	acc att ttc aag cag gtc tgg gat aat gcc ctg tac tgt gtg atg ggg	1152
	Thr Ile Phe Lys Gln Val Trp Asp Asn Ala Leu Tyr Cys Val Met Gly	
	370 375 380	
10	ccc tac aga atg aat aca ctg att ctg gcc gtg gtt tgg ttt gcc atg	1200
	Pro Tyr Arg Met Asn Thr Leu Ile Leu Ala Val Val Trp Phe Ala Met	
	385 390 395 400	
15	gca ttc agt tac tat gga ctg aca gtt tgg ttt cct gat atg atc cgc	1248
	Ala Phe Ser Tyr Tyr Gly Leu Thr Val Trp Phe Pro Asp Met Ile Arg	
	405 410 415	
20	tat ttt caa gat gaa gaa tac aag tct aaa atg aag gtg ttt ttt ggt	1296
	Tyr Phe Gln Asp Glu Glu Tyr Lys Ser Lys Met Lys Val Phe Phe Gly	
	420 425 430	
25	gag cat gtg tac ggc gcc aca atc aac ttc acg atg gaa aat cag atc	1344
	Glu His Val Tyr Gly Ala Thr Ile Asn Phe Thr Met Glu Asn Gln Ile	
	435 440 445	
30	cac caa cat ggg aaa ctt gtg aat gat aag ttc aca aga atg tac ttt	1392
	His Gln His Gly Lys Leu Val Asn Asp Lys Phe Thr Arg Met Tyr Phe	
	450 455 460	
35	aaa cat gta ctc ttt gag gac aca ttc ttt gac gag tgc tat ttt gaa	1440
	Lys His Val Leu Phe Glu Asp Thr Phe Phe Asp Glu Cys Tyr Phe Glu	
	465 470 475 480	
40	gac gta aca tca aca gat acc tac ttc aaa aat tgt acc att gaa tca	1488
	Asp Val Thr Ser Thr Asp Thr Tyr Phe Lys Asn Cys Thr Ile Glu Ser	
	485 490 495	
45	acc atc ttt tac aac aca gac ctc tac gag cac aag ttc atc aac tgt	1536
	Thr Ile Phe Tyr Asn Thr Asp Leu Tyr Glu His Lys Phe Ile Asn Cys	
	500 505 510	
50	cgg ttt atc aac tcc acc ttc ctg gag cag aag gag ggc tgc cac atg	1584
	Arg Phe Ile Asn Ser Thr Phe Leu Glu Gln Lys Glu Gly Cys His Met	
	515 520 525	
55	gac ttg gag caa gat aat gac ttc ctg att tac ctc gtc agc ttc ctg	1632
	Asp Leu Glu Gln Asp Asn Asp Phe Leu Ile Tyr Leu Val Ser Phe Leu	
	530 535 540	
60	ggc agc ctg tct gtc tta ccc ggg aac atc att tct gcc ctg ctc atg	1680
	Gly Ser Leu Ser Val Leu Pro Gly Asn Ile Ile Ser Ala Leu Leu Met	
	545 550 555 560	
65	gat aga att gga agg ctc aag atg att ggt ggc tcc atg cta atc tct	1728
	Asp Arg Ile Gly Arg Leu Lys Met Ile Gly Gly Ser Met Leu Ile Ser	
	565 570 575	
70	gca gtc tgc tgc ttc ttc ctg ttt ttt ggc aac agt gag tct gca atg	1776
	Ala Val Cys Cys Phe Phe Leu Phe Phe Gly Asn Ser Glu Ser Ala Met	
	580 585 590	

ES 2 311 667 T3

atc ggc tgg cag tgc ctg ttc tgt ggg aca agc att gca gcc tgg aat 1824
 Ile Gly Trp Gln Cys Leu Phe Cys Gly Thr Ser Ile Ala Ala Trp Asn
 595 600 605
 5 gct ctg gat gtg atc aca gtg gag ctg tat ccc acc aac cag aga gca 1872
 Ala Leu Asp Val Ile Thr Val Glu Leu Tyr Pro Thr Asn Gln Arg Ala
 610 615 620
 10 aca gcc ttc ggc att ctc aat gga tta tgc aaa ttt ggc gcc atc ctg 1920
 Thr Ala Phe Gly Ile Leu Asn Gly Leu Cys Lys Phe Gly Ala Ile Leu
 625 630 635 640
 15 gga aac acc atc ttt gct tct ttt gtt ggg ata acc aaa gtg gtc ccc 1968
 Gly Asn Thr Ile Phe Ala Ser Phe Val Gly Ile Thr Lys Val Val Pro
 645 650 655
 atc ctt ctg gct gct gct tct ctg gtt ggg ggt ggc ctg att gcc ctt 2016
 Ile Leu Leu Ala Ala Ser Leu Val Gly Gly Gly Leu Ile Ala Leu
 660 665 670
 20 cga ctg cca gag act cga gaa cag gtc ctg atg tga 2052
 Arg Leu Pro Glu Thr Arg Glu Gln Val Leu Met
 675 680

25 4
 683
 PRT
 Homo sapiens

35 Met Asp Asp Tyr Lys Tyr Gln Asp Asn Tyr Gly Gly Tyr Ala Pro Ser
 1 5 10 15
 Asp Gly Tyr Tyr Arg Gly Asn Glu Ser Asn Pro Glu Glu Asp Ala Gln
 20 25 30
 40 Ser Asp Val Thr Glu Gly His Asp Glu Glu Asp Glu Ile Tyr Glu Gly
 35 40 45
 45 Glu Tyr Gln Gly Ile Pro His Pro Asp Asp Val Lys Ala Lys Gln Ala
 50 55 60
 50 Lys Met Ala Pro Ser Arg Met Asp Ser Leu Arg Gly Gln Thr Asp Leu
 65 70 75 80
 Met Ala Glu Arg Leu Glu Asp Glu Glu Gln Leu Ala His Gln Tyr Glu
 85 90 95

ES 2 311 667 T3

	Thr	Ile	Met	Asp	Glu	Cys	Gly	His	Gly	Arg	Phe	Gln	Trp	Ile	Leu	Phe
				100					105					110		
5	Phe	Val	Leu	Gly	Leu	Ala	Leu	Met	Ala	Asp	Gly	Val	Glu	Val	Phe	Val
			115					120					125			
10	Val	Ser	Phe	Ala	Leu	Pro	Ser	Ala	Glu	Lys	Asp	Met	Cys	Leu	Ser	Ser
		130					135					140				
15	Ser	Lys	Lys	Gly	Met	Leu	Gly	Met	Ile	Val	Tyr	Leu	Gly	Met	Met	Ala
	145					150					155					160
20	Gly	Ala	Phe	Ile	Leu	Gly	Gly	Leu	Ala	Asp	Lys	Leu	Gly	Arg	Lys	Arg
				165						170					175	
25	Val	Leu	Ser	Met	Ser	Leu	Ala	Val	Asn	Ala	Ser	Phe	Ala	Ser	Leu	Ser
			180						185					190		
30	Ser	Phe	Val	Gln	Gly	Tyr	Gly	Ala	Phe	Leu	Phe	Cys	Arg	Leu	Ile	Ser
		195						200					205			
35	Gly	Ile	Gly	Ile	Gly	Gly	Ala	Leu	Pro	Ile	Val	Phe	Ala	Tyr	Phe	Ser
	210						215						220			
40	Glu	Phe	Leu	Ser	Arg	Glu	Lys	Arg	Gly	Glu	His	Leu	Ser	Trp	Leu	Gly
	225					230					235					240
45	Ile	Phe	Trp	Met	Thr	Gly	Gly	Leu	Tyr	Ala	Ser	Ala	Met	Ala	Trp	Ser
				245						250					255	
50	Ile	Ile	Pro	His	Tyr	Gly	Trp	Gly	Phe	Ser	Met	Gly	Thr	Asn	Tyr	His
			260						265					270		
55	Phe	His	Ser	Trp	Arg	Val	Phe	Val	Ile	Val	Cys	Ala	Leu	Pro	Cys	Thr
			275					280					285			
60	Val	Ser	Met	Val	Ala	Leu	Lys	Phe	Met	Pro	Glu	Ser	Pro	Arg	Phe	Leu
	290						295					300				
65	Leu	Glu	Met	Gly	Lys	His	Asp	Glu	Ala	Trp	Met	Ile	Leu	Lys	Gln	Val
	305					310					315					320
70	His	Asp	Thr	Asn	Met	Arg	Ala	Lys	Gly	Thr	Pro	Glu	Lys	Val	Phe	Thr
				325						330					335	
75	Val	Ser	Asn	Ile	Lys	Thr	Pro	Lys	Gln	Met	Asp	Glu	Phe	Ile	Glu	Ile

ES 2 311 667 T3

	340		345		350
5	Gln Ser Ser Thr Gly Thr Trp Tyr Gln Arg Trp Leu Val Arg Phe Lys	355	360	365	
	Thr Ile Phe Lys Gln Val Trp Asp Asn Ala Leu Tyr Cys Val Met Gly	370	375	380	
10	Pro Tyr Arg Met Asn Thr Leu Ile Leu Ala Val Val Trp Phe Ala Met	385	390	395	400
15	Ala Phe Ser Tyr Tyr Gly Leu Thr Val Trp Phe Pro Asp Met Ile Arg	405	410	415	
	Tyr Phe Gln Asp Glu Glu Tyr Lys Ser Lys Met Lys Val Phe Phe Gly	420	425	430	
20	Glu His Val Tyr Gly Ala Thr Ile Asn Phe Thr Met Glu Asn Gln Ile	435	440	445	
25	His Gln His Gly Lys Leu Val Asn Asp Lys Phe Thr Arg Met Tyr Phe	450	455	460	
	Lys His Val Leu Phe Glu Asp Thr Phe Phe Asp Glu Cys Tyr Phe Glu	465	470	475	480
30	Asp Val Thr Ser Thr Asp Thr Tyr Phe Lys Asn Cys Thr Ile Glu Ser	485	490	495	
35	Thr Ile Phe Tyr Asn Thr Asp Leu Tyr Glu His Lys Phe Ile Asn Cys	500	505	510	
40	Arg Phe Ile Asn Ser Thr Phe Leu Glu Gln Lys Glu Gly Cys His Met	515	520	525	
	Asp Leu Glu Gln Asp Asn Asp Phe Leu Ile Tyr Leu Val Ser Phe Leu	530	535	540	
45	Gly Ser Leu Ser Val Leu Pro Gly Asn Ile Ile Ser Ala Leu Leu Met	545	550	555	560
50	Asp Arg Ile Gly Arg Leu Lys Met Ile Gly Gly Ser Met Leu Ile Ser	565	570	575	
	Ala Val Cys Cys Phe Phe Leu Phe Phe Gly Asn Ser Glu Ser Ala Met	580	585	590	

ES 2 311 667 T3

Ile Gly Trp Gln Cys Leu Phe Cys Gly Thr Ser Ile Ala Ala Trp Asn
595 600 605

Ala Leu Asp Val Ile Thr Val Glu Leu Tyr Pro Thr Asn Gln Arg Ala
610 615 620

Thr Ala Phe Gly Ile Leu Asn Gly Leu Cys Lys Phe Gly Ala Ile Leu
625 630 635 640

Gly Asn Thr Ile Phe Ala Ser Phe Val Gly Ile Thr Lys Val Val Pro
645 650 655

Ile Leu Leu Ala Ala Ala Ser Leu Val Gly Gly Gly Leu Ile Ala Leu
660 665 670

Arg Leu Pro Glu Thr Arg Glu Gln Val Leu Met
675 680

5
2184
ADN
Homo sapiens

CDS
(1)..(2184)

atg gaa gac tct tac aag gat agg act tca ctg atg aag ggt gcc aag	48
Met Glu Asp Ser Tyr Lys Asp Arg Thr Ser Leu Met Lys Gly Ala Lys	
1 5 10 15	
gac att gcc aga gag gtg aag aaa caa aca gta aag aag gtg aat caa	96
Asp Ile Ala Arg Glu Val Lys Lys Gln Thr Val Lys Lys Val Asn Gln	
20 25 30	
gct gtg gac cga gcc cag gat gaa tac acc cag agg tcc tac agt cgg	144
Ala Val Asp Arg Ala Gln Asp Glu Tyr Thr Gln Arg Ser Tyr Ser Arg	
35 40 45	
ttc caa gat gaa gaa gat gat gat gac tac tac ccg gct gga gaa acc	192
Phe Gln Asp Glu Glu Asp Asp Asp Asp Tyr Tyr Pro Ala Gly Glu Thr	

ES 2 311 667 T3

	50	55	60	
5	tat aat ggt gag gcc aac gat gac gaa ggc tca agt gaa gcc act gag Tyr Asn Gly Glu Ala Asn Asp Asp Glu Gly Ser Ser Glu Ala Thr Glu 65 70 75 80	240		
	ggg cat gat gaa gat gat gag atc tat gag ggg gag tat cag ggc atc Gly His Asp Glu Asp Asp Glu Ile Tyr Glu Gly Glu Tyr Gln Gly Ile 85 90 95	288		
10	ccc agt atg aac caa gcg aag gac agc atc gtg tca gtg ggg cag ccc Pro Ser Met Asn Gln Ala Lys Asp Ser Ile Val Ser Val Gly Gln Pro 100 105 110	336		
15	aag ggc gat gag tac aag gac cga cgg gag ctg gaa tca gaa agg aga Lys Gly Asp Glu Tyr Lys Asp Arg Arg Glu Leu Glu Ser Glu Arg Arg 115 120 125	384		
20	gct gac gag gaa gag tta gcc cag cag tat gag ctg ata atc caa gaa Ala Asp Glu Glu Glu Leu Ala Gln Gln Tyr Glu Leu Ile Ile Gln Glu 130 135 140	432		
	tgc ggt cat ggt cgt ttt cag tgg gcc ctt ttc ttc gtc ctg ggc atg Cys Gly His Gly Arg Phe Gln Trp Ala Leu Phe Phe Val Leu Gly Met 145 150 155 160	480		
25	gct ctt atg gca gac ggt gta gag gtg ttt gtc gtt ggc ttc gtg tta Ala Leu Met Ala Asp Gly Val Glu Val Phe Val Val Gly Phe Val Leu 165 170 175	528		
30	ccc agt gct gag aca gac ctc tgc atc cca aat tca gga tct gga tgg Pro Ser Ala Glu Thr Asp Leu Cys Ile Pro Asn Ser Gly Ser Gly Trp 180 185 190	576		
	cta ggc agc ata gtg tac ctc ggg atg atg gtg ggg gcg ttc ttc tgg Leu Gly Ser Ile Val Tyr Leu Gly Met Met Val Gly Ala Phe Phe Trp 195 200 205	624		
35	gga gga ctg gca gac aaa gtg gga agg aaa cag tct ctt ctg att tgc Gly Gly Leu Ala Asp Lys Val Gly Arg Lys Gln Ser Leu Leu Ile Cys 210 215 220	672		
40	atg tct gtc aac gga ttc ttt gcc ttc ctt tct tca ttt gtc caa ggt Met Ser Val Asn Gly Phe Phe Ala Phe Leu Ser Ser Phe Val Gln Gly 225 230 235 240	720		
45	tat ggc ttc ttt ctc ttc tgt cgc tta ctt tct gga ttc ggg att gga Tyr Gly Phe Phe Leu Phe Cys Arg Leu Leu Ser Gly Phe Gly Ile Gly 245 250 255	768		
	gga gcc ata ccc act gtg ttc tcg tac ttt gct gaa gtc ctg gcc cgg Gly Ala Ile Pro Thr Val Phe Ser Tyr Phe Ala Glu Val Leu Ala Arg 260 265 270	816		
50	gaa aag cgg ggc gaa cac ttg agc tgg ctc tgc atg ttc tgg atg atc Glu Lys Arg Gly Glu His Leu Ser Trp Leu Cys Met Phe Trp Met Ile 275 280 285	864		
55	ggg ggc atc tac gcc tct gcc atg gcc tgg gcc atc atc ccg cac tac Gly Gly Ile Tyr Ala Ser Ala Met Ala Trp Ala Ile Ile Pro His Tyr 290 295 300	912		
60				
65				

ES 2 311 667 T3

	ggg tgg agc ttc agc atg gga tgc gcc tac cag ttt cac agt tgg cgt Gly Trp Ser Phe Ser Met Gly Ser Ala Tyr Gln Phe His Ser Trp Arg 305 310 315 320	960
5	gtg ttt gtc atc gtc tgt gca ctc ccc tgt gtc tcc tcc gtg gtg gcc Val Phe Val Ile Val Cys Ala Leu Pro Cys Val Ser Ser Val Val Ala 325 330 335	1008
10	ctc aca ttc atg cct gaa agc cca cga ttc ttg ttg gag gtt gga aaa Leu Thr Phe Met Pro Glu Ser Pro Arg Phe Leu Leu Glu Val Gly Lys 340 345 350	1056
15	cat gat gaa gct tgg atg att ctg aag tta att cat gac acc aac atg His Asp Glu Ala Trp Met Ile Leu Lys Leu Ile His Asp Thr Asn Met 355 360 365	1104
	aga gcc cgg ggt cag cct gag aag gtc ttc acg gta aac aaa ata aaa Arg Ala Arg Gly Gln Pro Glu Lys Val Phe Thr Val Asn Lys Ile Lys 370 375 380	1152
20	act cct aaa caa ata gat gag ctg att gaa att gag agt gac aca gga Thr Pro Lys Gln Ile Asp Glu Leu Ile Glu Ile Glu Ser Asp Thr Gly 385 390 395 400	1200
25	aca tgg tat agg agg tgt ttt gtt cgg atc cgc acc gag ctg tac gga Thr Trp Tyr Arg Arg Cys Phe Val Arg Ile Arg Thr Glu Leu Tyr Gly 405 410 415	1248
	att tgg ttg act ttt atg aga tgt ttc aac tac cca gtc agg gat aat Ile Trp Leu Thr Phe Met Arg Cys Phe Asn Tyr Pro Val Arg Asp Asn 420 425 430	1296
30	aca ata aag ctt aca att gtt tgg ttc acc ctg tcc ttt ggg tac tat Thr Ile Lys Leu Thr Ile Val Trp Phe Thr Leu Ser Phe Gly Tyr Tyr 435 440 445	1344
35	gga tta tcc gtt tgg ttc cct gat gtc att aaa cct ctg cag tcc gat Gly Leu Ser Val Trp Phe Pro Asp Val Ile Lys Pro Leu Gln Ser Asp 450 455 460	1392
40	gaa tat gca ttg cta acc aga aat gtg gag aga gat aaa tat gca aat Glu Tyr Ala Leu Leu Thr Arg Asn Val Glu Arg Asp Lys Tyr Ala Asn 465 470 475 480	1440
	ttc act att aac ttt aca atg gaa aat cag att cat act gga atg gaa Phe Thr Ile Asn Phe Thr Met Glu Asn Gln Ile His Thr Gly Met Glu 485 490 495	1488
45	tac gac aat ggc aga ttc ata ggg gtc aag ttc aaa tct gta act ttc Tyr Asp Asn Gly Arg Phe Ile Gly Val Lys Phe Lys Ser Val Thr Phe 500 505 510	1536
50	aaa gac tct gtt ttt aag tcc tgc acc ttt gag gat gta act tca gtg Lys Asp Ser Val Phe Lys Ser Cys Thr Phe Glu Asp Val Thr Ser Val 515 520 525	1584
	aac acc tac ttc aag aac tgc aca ttt att gac act gtt ttt gac aac Asn Thr Tyr Phe Lys Asn Cys Thr Phe Ile Asp Thr Val Phe Asp Asn 530 535 540	1632

ES 2 311 667 T3

	aca gat ttt gag cca tat aaa ttc att gac agt gaa ttt aaa aac tgc	1680
	Thr Asp Phe Glu Pro Tyr Lys Phe Ile Asp Ser Glu Phe Lys Asn Cys	
	545 550 555 560	
5	tcg ttt ttt cac aac aag acg gga tgt cag att acc ttt gat gat gac	1728
	Ser Phe Phe His Asn Lys Thr Gly Cys Gln Ile Thr Phe Asp Asp Asp	
	565 570 575	
10	tat agt gcc tac tgg att tat ttt gtc aac ttt ctg ggg aca ttg gca	1776
	Tyr Ser Ala Tyr Trp Ile Tyr Phe Val Asn Phe Leu Gly Thr Leu Ala	
	580 585 590	
15	gta ttg cca ggg aac att gtg tct gct ctg ctg atg gac aga att ggg	1824
	Val Leu Pro Gly Asn Ile Val Ser Ala Leu Leu Met Asp Arg Ile Gly	
	595 600 605	
20	cgc tta aca atg cta ggt ggc tct atg gtg ctt tcg ggg atc agc tgt	1872
	Arg Leu Thr Met Leu Gly Gly Ser Met Val Leu Ser Gly Ile Ser Cys	
	610 615 620	
25	ttc ttc ctt tgg ttc ggc acc agt gaa tcc atg atg ata ggc atg ctg	1920
	Phe Phe Leu Trp Phe Gly Thr Ser Glu Ser Met Met Ile Gly Met Leu	
	625 630 635 640	
30	tgt ctg tac aat gga ttg acc atc tca gcc tgg aac tct ctt gac gtg	1968
	Cys Leu Tyr Asn Gly Leu Thr Ile Ser Ala Trp Asn Ser Leu Asp Val	
	645 650 655	
35	gtc act gtg gaa ctg tac ccc aca gac cgg agg gca aca ggc ttt ggc	2016
	Val Thr Val Glu Leu Tyr Pro Thr Asp Arg Arg Ala Thr Gly Phe Gly	
	660 665 670	
40	ttc tta aat gcg cta tgc aag gca gca gcc gtc ctg gga aac tta ata	2064
	Phe Leu Asn Ala Leu Cys Lys Ala Ala Ala Val Leu Gly Asn Leu Ile	
	675 680 685	
45	ttt ggc tct ctg gtc agc atc acc aaa tca atc ccc atc ctg ctg gct	2112
	Phe Gly Ser Leu Val Ser Ile Thr Lys Ser Ile Pro Ile Leu Leu Ala	
	690 695 700	
50	tct act gtg ctc gtg tgt gga gga ctc gtt ggg ctg tgc ctg cct gac	2160
	Ser Thr Val Leu Val Cys Gly Gly Leu Val Gly Leu Cys Leu Pro Asp	
	705 710 715 720	
55	aca cga acc cag gtt ctg atg taa	2184
	Thr Arg Thr Gln Val Leu Met	
	725	

ES 2 311 667 T3

6

727

PRT

5 *Homo sapiens*

10	Met Glu Asp Ser Tyr Lys Asp Arg Thr Ser Leu Met Lys Gly Ala Lys	1 5 10 15
15	Asp Ile Ala Arg Glu Val Lys Lys Gln Thr Val Lys Lys Val Asn Gln	20 25 30
20	Ala Val Asp Arg Ala Gln Asp Glu Tyr Thr Gln Arg Ser Tyr Ser Arg	35 40 45
25	Phe Gln Asp Glu Glu Asp Asp Asp Asp Tyr Tyr Pro Ala Gly Glu Thr	50 55 60
30	Tyr Asn Gly Glu Ala Asn Asp Asp Glu Gly Ser Ser Glu Ala Thr Glu	65 70 75 80
35	Gly His Asp Glu Asp Asp Glu Ile Tyr Glu Gly Glu Tyr Gln Gly Ile	85 90 95
40	Pro Ser Met Asn Gln Ala Lys Asp Ser Ile Val Ser Val Gly Gln Pro	100 105 110
45	Lys Gly Asp Glu Tyr Lys Asp Arg Arg Glu Leu Glu Ser Glu Arg Arg	115 120 125
50	Ala Asp Glu Glu Glu Leu Ala Gln Gln Tyr Glu Leu Ile Ile Gln Glu	130 135 140
55	Cys Gly His Gly Arg Phe Gln Trp Ala Leu Phe Phe Val Leu Gly Met	145 150 155 160
60	Ala Leu Met Ala Asp Gly Val Glu Val Phe Val Val Gly Phe Val Leu	165 170 175
65	Pro Ser Ala Glu Thr Asp Leu Cys Ile Pro Asn Ser Gly Ser Gly Trp	180 185 190
70	Leu Gly Ser Ile Val Tyr Leu Gly Met Met Val Gly Ala Phe Phe Trp	195 200 205
75	Gly Gly Leu Ala Asp Lys Val Gly Arg Lys Gln Ser Leu Leu Ile Cys	210 215 220
80	Met Ser Val Asn Gly Phe Phe Ala Phe Leu Ser Ser Phe Val Gln Gly	225 230 235 240
85	Tyr Gly Phe Phe Leu Phe Cys Arg Leu Leu Ser Gly Phe Gly Ile Gly	

ES 2 311 667 T3

	245	250	255
5	Gly Ala Ile Pro Thr Val Phe Ser Tyr Phe Ala Glu Val Leu Ala Arg 260 265 270		
	Glu Lys Arg Gly Glu His Leu Ser Trp Leu Cys Met Phe Trp Met Ile 275 280 285		
10	Gly Gly Ile Tyr Ala Ser Ala Met Ala Trp Ala Ile Ile Pro His Tyr 290 295 300		
15	Gly Trp Ser Phe Ser Met Gly Ser Ala Tyr Gln Phe His Ser Trp Arg 305 310 315 320		
	Val Phe Val Ile Val Cys Ala Leu Pro Cys Val Ser Ser Val Val Ala 325 330 335		
20	Leu Thr Phe Met Pro Glu Ser Pro Arg Phe Leu Leu Glu Val Gly Lys 340 345 350		
25	His Asp Glu Ala Trp Met Ile Leu Lys Leu Ile His Asp Thr Asn Met 355 360 365		
	Arg Ala Arg Gly Gln Pro Glu Lys Val Phe Thr Val Asn Lys Ile Lys 370 375 380		
30	Thr Pro Lys Gln Ile Asp Glu Leu Ile Glu Ile Glu Ser Asp Thr Gly 385 390 395 400		
35	Thr Trp Tyr Arg Arg Cys Phe Val Arg Ile Arg Thr Glu Leu Tyr Gly 405 410 415		
	Ile Trp Leu Thr Phe Met Arg Cys Phe Asn Tyr Pro Val Arg Asp Asn 420 425 430		
40	Thr Ile Lys Leu Thr Ile Val Trp Phe Thr Leu Ser Phe Gly Tyr Tyr 435 440 445		
45	Gly Leu Ser Val Trp Phe Pro Asp Val Ile Lys Pro Leu Gln Ser Asp 450 455 460		
50	Glu Tyr Ala Leu Leu Thr Arg Asn Val Glu Arg Asp Lys Tyr Ala Asn 465 470 475 480		
	Phe Thr Ile Asn Phe Thr Met Glu Asn Gln Ile His Thr Gly Met Glu 485 490 495		
55			
60			
65			

ES 2 311 667 T3

	Tyr	Asp	Asn	Gly	Arg	Phe	Ile	Gly	Val	Lys	Phe	Lys	Ser	Val	Thr	Phe	
				500					505					510			
5	Lys	Asp	Ser	Val	Phe	Lys	Ser	Cys	Thr	Phe	Glu	Asp	Val	Thr	Ser	Val	
			515					520					525				
10	Asn	Thr	Tyr	Phe	Lys	Asn	Cys	Thr	Phe	Ile	Asp	Thr	Val	Phe	Asp	Asn	
		530					535						540				
	Thr	Asp	Phe	Glu	Pro	Tyr	Lys	Phe	Ile	Asp	Ser	Glu	Phe	Lys	Asn	Cys	
	545						550				555					560	
15	Ser	Phe	Phe	His	Asn	Lys	Thr	Gly	Cys	Gln	Ile	Thr	Phe	Asp	Asp	Asp	
				565						570					575		
20	Tyr	Ser	Ala	Tyr	Trp	Ile	Tyr	Phe	Val	Asn	Phe	Leu	Gly	Thr	Leu	Ala	
				580					585					590			
	Val	Leu	Pro	Gly	Asn	Ile	Val	Ser	Ala	Leu	Leu	Met	Asp	Arg	Ile	Gly	
			595					600					605				
25	Arg	Leu	Thr	Met	Leu	Gly	Gly	Ser	Met	Val	Leu	Ser	Gly	Ile	Ser	Cys	
		610					615						620				
30	Phe	Phe	Leu	Trp	Phe	Gly	Thr	Ser	Glu	Ser	Met	Met	Ile	Gly	Met	Leu	
	625					630					635					640	
	Cys	Leu	Tyr	Asn	Gly	Leu	Thr	Ile	Ser	Ala	Trp	Asn	Ser	Leu	Asp	Val	
					645					650					655		
35	Val	Thr	Val	Glu	Leu	Tyr	Pro	Thr	Asp	Arg	Arg	Ala	Thr	Gly	Phe	Gly	
				660					665					670			
40	Phe	Leu	Asn	Ala	Leu	Cys	Lys	Ala	Ala	Ala	Val	Leu	Gly	Asn	Leu	Ile	
			675					680						685			
	Phe	Gly	Ser	Leu	Val	Ser	Ile	Thr	Lys	Ser	Ile	Pro	Ile	Leu	Leu	Ala	
45		690					695					700					
	Ser	Thr	Val	Leu	Val	Cys	Gly	Gly	Leu	Val	Gly	Leu	Cys	Leu	Pro	Asp	
	705					710					715					720	
50	Thr	Arg	Thr	Gln	Val	Leu	Met										
					725												
55																	
60																	
65																	

ES 2 311 667 T3

7

1647

ADN

5 *Homo sapiens*

CDS

(1)..(1647)

10	atg gag gag gac tta ttc cag cta agg cag ctg ccg gtt gtg aaa ttc	48
	Met Glu Glu Asp Leu Phe Gln Leu Arg Gln Leu Pro Val Val Lys Phe	
	1 5 10 15	
15	cgt cgc aca ggc gag agt gca agg tca gag gac gac acg gct tca gga	96
	Arg Arg Thr Gly Glu Ser Ala Arg Ser Glu Asp Asp Thr Ala Ser Gly	
	20 25 30	
20	gag cat gaa gtc cag att gaa ggg gtc cac gtg ggc cta gag gct gtg	144
	Glu His Glu Val Gln Ile Glu Gly Val His Val Gly Leu Glu Ala Val	
	35 40 45	
25	gag ctg gat gat ggg gca gct gtg ccc aag gag ttt gcc aat ccc act	192
	Glu Leu Asp Asp Gly Ala Ala Val Pro Lys Glu Phe Ala Asn Pro Thr	
	50 55 60	
30	gat gat act ttc atg gtg gaa gat gca gtg gaa gcc att ggc ttt gga	240
	Asp Asp Thr Phe Met Val Glu Asp Ala Val Glu Ala Ile Gly Phe Gly	
	65 70 75 80	
35	aaa ttt cag tgg aag ctg tct gtt ctc act ggc ttg gct tgg atg gct	288
	Lys Phe Gln Trp Lys Leu Ser Val Leu Thr Gly Leu Ala Trp Met Ala	
	85 90 95	
40	gat gcc atg gag atg atg atc ctc agc atc ctg gca cca cag ctg cat	336
	Asp Ala Met Glu Met Met Ile Leu Ser Ile Leu Ala Pro Gln Leu His	
	100 105 110	
45	tgc gag tgg agg ctc cca agc tgg cag gtg gca ttg ctg acc tcg gtg	384
	Cys Glu Trp Arg Leu Pro Ser Trp Gln Val Ala Leu Leu Thr Ser Val	
	115 120 125	
50	gtc ttt gta ggc atg atg tcc agc tcc acg ctc tgg gga aat atc tca	432
	Val Phe Val Gly Met Met Ser Ser Ser Thr Leu Trp Gly Asn Ile Ser	
	130 135 140	
55	gac cag tac ggc agg aaa aca ggg ctg aag atc agc gtg ctg tgg act	480
	Asp Gln Tyr Gly Arg Lys Thr Gly Leu Lys Ile Ser Val Leu Trp Thr	
	145 150 155 160	

ES 2 311 667 T3

	ctg tac tat ggc atc ctt agt gca ttt gcg ccc gtg tat agc tgg atc Leu Tyr Tyr Gly Ile Leu Ser Ala Phe Ala Pro Val Tyr Ser Trp Ile 165 170 175	528
5	ctg gtg ctc cgg ggc ctg gtg ggc ttc ggg atc gga gga gtt ccc cag Leu Val Leu Arg Gly Leu Val Gly Phe Gly Ile Gly Gly Val Pro Gln 180 185 190	576
10	tcg gtg acg ctg tat gcc gag ttc ctt ccc atg aaa gcc aga gct aaa Ser Val Thr Leu Tyr Ala Glu Phe Leu Pro Met Lys Ala Arg Ala Lys 195 200 205	624
15	tgt att ttg ctg att gag gta ttc tgg gcc atc ggg aca gtg ttc gag Cys Ile Leu Leu Ile Glu Val Phe Trp Ala Ile Gly Thr Val Phe Glu 210 215 220	672
	gtc gtc ctg gct gtg ttc gtg atg ccc agc ctg ggc tgg cgt tgg ctg Val Val Leu Ala Val Phe Val Met Pro Ser Leu Gly Trp Arg Trp Leu 225 230 235 240	720
20	ctc atc ctc tca gct gtc ccg ctc ctc ctc ttt gcc gtg ctg tgt ttc Leu Ile Leu Ser Ala Val Pro Leu Leu Leu Phe Ala Val Leu Cys Phe 245 250 255	768
25	tgg ctg cct gaa agt gca agg tat gat gtg ctg tca ggg aac cag gaa Trp Leu Pro Glu Ser Ala Arg Tyr Asp Val Leu Ser Gly Asn Gln Glu 260 265 270	816
	aag gca atc gcc acc tta aag agg ata gca act gaa aac gga gct ccc Lys Ala Ile Ala Thr Leu Lys Arg Ile Ala Thr Glu Asn Gly Ala Pro 275 280 285	864
30	atg ccg ctg ggg aaa ctc atc atc tcc aga cag gaa gac cga ggc aaa Met Pro Leu Gly Lys Leu Ile Ile Ser Arg Gln Glu Asp Arg Gly Lys 290 295 300	912
35	atg agg gac ctt ttc aca ccc cat ttt aga tgg aca act ttg ctg ctg Met Arg Asp Leu Phe Thr Pro His Phe Arg Trp Thr Thr Leu Leu 305 310 315 320	960
40	tgg ttt ata tgg ttt tcc aat gca ttc tct tac tac ggg tta gtt cta Trp Phe Ile Trp Phe Ser Asn Ala Phe Ser Tyr Tyr Gly Leu Val Leu 325 330 335	1008
	ctc acc aca gaa ctc ttc cag gca gga gat gtc tgc ggc atc tcc agt Leu Thr Thr Glu Leu Phe Gln Ala Gly Asp Val Cys Gly Ile Ser Ser 340 345 350	1056
45	cgg aag aag gct gta gag gca aaa tgc agc ctg gcc tgc gag tac ctg Arg Lys Lys Ala Val Glu Ala Lys Cys Ser Leu Ala Cys Glu Tyr Leu 355 360 365	1104
50	agt gag gag gat tac atg gac ttg ctg tgg acc acc ctc tct gag ttt Ser Glu Glu Asp Tyr Met Asp Leu Leu Trp Thr Thr Leu Ser Glu Phe 370 375 380	1152
55	cca ggt gtc ctt gtg act ctg tgg att att gac cgc ctg ggg cgc aag Pro Gly Val Leu Val Thr Leu Trp Ile Ile Asp Arg Leu Gly Arg Lys 385 390 395 400	1200

ES 2 311 667 T3

	aag acc atg gcc ctg tgc ttt gtc atc ttc tcc ttc tgc agc ctc ctg	1248
	Lys Thr Met Ala Leu Cys Phe Val Ile Phe Ser Phe Cys Ser Leu Leu	
	405 410 415	
5	ctg ttt atc tgt gtt gga aga aat gtg ctc act ctg tta ctc ttc att	1296
	Leu Phe Ile Cys Val Gly Arg Asn Val Leu Thr Leu Leu Phe Ile	
	420 425 430	
10	gca aga gcg ttt att tct gga ggc ttt caa gcg gca tat gtt tac aca	1344
	Ala Arg Ala Phe Ile Ser Gly Gly Phe Gln Ala Ala Tyr Val Tyr Thr	
	435 440 445	
15	cct gag gtc tac ccc acg gca acg cgg gcc ctc ggc ctg ggc acc tgc	1392
	Pro Glu Val Tyr Pro Thr Ala Thr Arg Ala Leu Gly Leu Gly Thr Cys	
	450 455 460	
20	agc ggc atg gca aga gtg ggt gct ctc atc act ccg ttc atc gcc cag	1440
	Ser Gly Met Ala Arg Val Gly Ala Leu Ile Thr Pro Phe Ile Ala Gln	
	465 470 475 480	
25	gtg atg ctg gaa tcc tct gtg tac ctg act ctg gca gtt tac agt ggc	1488
	Val Met Leu Glu Ser Ser Val Tyr Leu Thr Leu Ala Val Tyr Ser Gly	
	485 490 495	
30	tgc tgc ctc ctg gct gcc ctg gcc tcc tgc ttt ttg ccc att gag acc	1536
	Cys Cys Leu Leu Ala Ala Leu Ala Ser Cys Phe Leu Pro Ile Glu Thr	
	500 505 510	
35	aaa ggc cga gga ctg cag gag tcc agc cac cgg gag tgg ggc cag gag	1584
	Lys Gly Arg Gly Leu Gln Glu Ser Ser His Arg Glu Trp Gly Gln Glu	
	515 520 525	
40	atg gtc ggc cga gga atg cac ggt gca ggt gtt acc agg tcg aac tct	1632
	Met Val Gly Arg Gly Met His Gly Ala Gly Val Thr Arg Ser Asn Ser	
	530 535 540	
45	ggc tct cag gaa tag	1647
	Gly Ser Gln Glu	
	545	

8

40 548

PRT

Homo sapiens

45	Met Glu Glu Asp Leu Phe Gln Leu Arg Gln Leu Pro Val Val Lys Phe
	1 5 10 15

50	Arg Arg Thr Gly Glu Ser Ala Arg Ser Glu Asp Asp Thr Ala Ser Gly
	20 25 30

55

60

65

ES 2 311 667 T3

	Glu	His	Glu	Val	Gln	Ile	Glu	Gly	Val	His	Val	Gly	Leu	Glu	Ala	Val	
			35					40					45				
5	Glu	Leu	Asp	Asp	Gly	Ala	Ala	Val	Pro	Lys	Glu	Phe	Ala	Asn	Pro	Thr	
		50					55					60					
10	Asp	Asp	Thr	Phe	Met	Val	Glu	Asp	Ala	Val	Glu	Ala	Ile	Gly	Phe	Gly	
	65					70					75				80		
15	Lys	Phe	Gln	Trp	Lys	Leu	Ser	Val	Leu	Thr	Gly	Leu	Ala	Trp	Met	Ala	
					85					90					95		
20	Asp	Ala	Met	Glu	Met	Met	Ile	Leu	Ser	Ile	Leu	Ala	Pro	Gln	Leu	His	
				100					105					110			
25	Cys	Glu	Trp	Arg	Leu	Pro	Ser	Trp	Gln	Val	Ala	Leu	Leu	Thr	Ser	Val	
			115					120					125				
30	Val	Phe	Val	Gly	Met	Met	Ser	Ser	Ser	Thr	Leu	Trp	Gly	Asn	Ile	Ser	
		130					135					140					
35	Asp	Gln	Tyr	Gly	Arg	Lys	Thr	Gly	Leu	Lys	Ile	Ser	Val	Leu	Trp	Thr	
	145					150				155					160		
40	Leu	Tyr	Tyr	Gly	Ile	Leu	Ser	Ala	Phe	Ala	Pro	Val	Tyr	Ser	Trp	Ile	
				165					170					175			
45	Leu	Val	Leu	Arg	Gly	Leu	Val	Gly	Phe	Gly	Ile	Gly	Gly	Val	Pro	Gln	
			180					185					190				
50	Ser	Val	Thr	Leu	Tyr	Ala	Glu	Phe	Leu	Pro	Met	Lys	Ala	Arg	Ala	Lys	
		195						200				205					
55	Cys	Ile	Leu	Leu	Ile	Glu	Val	Phe	Trp	Ala	Ile	Gly	Thr	Val	Phe	Glu	
	210					215						220					
60	Val	Val	Leu	Ala	Val	Phe	Val	Met	Pro	Ser	Leu	Gly	Trp	Arg	Trp	Leu	
	225				230						235				240		
65	Leu	Ile	Leu	Ser	Ala	Val	Pro	Leu	Leu	Leu	Phe	Ala	Val	Leu	Cys	Phe	
				245					250					255			
70	Trp	Leu	Pro	Glu	Ser	Ala	Arg	Tyr	Asp	Val	Leu	Ser	Gly	Asn	Gln	Glu	
			260					265						270			
75	Lys	Ala	Ile	Ala	Thr	Leu	Lys	Arg	Ile	Ala	Thr	Glu	Asn	Gly	Ala	Pro	

ES 2 311 667 T3

	275	280	285
5	Met Pro Leu Gly Lys Leu	Ile Ile Ser Arg Gln	Glu Asp Arg Gly Lys
	290	295	300
10	Met Arg Asp Leu Phe Thr	Pro His Phe Arg Trp	Thr Thr Leu Leu Leu
	305	310	315 320
15	Trp Phe Ile Trp Phe Ser	Asn Ala Phe Ser Tyr	Tyr Gly Leu Val Leu
	325	330	335
20	Leu Thr Thr Glu Leu Phe	Gln Ala Gly Asp Val	Cys Gly Ile Ser Ser
	340	345	350
25	Arg Lys Lys Ala Val Glu	Ala Lys Cys Ser Leu	Ala Cys Glu Tyr Leu
	355	360	365
30	Ser Glu Glu Asp Tyr Met	Asp Leu Leu Trp Thr	Thr Thr Leu Ser Glu Phe
	370	375	380
35	Pro Gly Val Leu Val Thr	Leu Trp Ile Ile Asp	Arg Leu Gly Arg Lys
	385	390	395 400
40	Lys Thr Met Ala Leu Cys	Phe Val Ile Phe Ser	Phe Cys Ser Leu Leu
	405	410	415
45	Leu Phe Ile Cys Val Gly	Arg Asn Val Leu Thr	Leu Leu Leu Phe Ile
	420	425	430
50	Ala Arg Ala Phe Ile Ser	Gly Gly Phe Gln Ala	Ala Tyr Val Tyr Thr
	435	440	445
55	Pro Glu Val Tyr Pro Thr	Ala Thr Arg Ala Leu	Gly Leu Gly Thr Cys
	450	455	460
60	Ser Gly Met Ala Arg Val	Gly Ala Leu Ile Thr	Pro Phe Ile Ala Gln
	465	470	475 480
65	Val Met Leu Glu Ser Ser	Val Tyr Leu Thr Leu	Ala Val Tyr Ser Gly
	485	490	495
70	Cys Cys Leu Leu Ala Ala	Leu Ala Ser Cys Phe	Leu Pro Ile Glu Thr
	500	505	510
75	Lys Gly Arg Gly Leu Gln	Glu Ser Ser His Arg	Glu Trp Gly Gln Glu
	515	520	525
80	Met Val Gly Arg Gly Met	His Gly Ala Gly Val	Thr Arg Ser Asn Ser
	530	535	540
85	Gly Ser Gln Glu		
	545		

ES 2 311 667 T3

9

3844

ADN

5 *Rattus norvegicus*

CDS

(400)..(2628)

10

```

cctcagccct gtggctggac cccctccctc acccggggac tccctgaccc ggggaaccaa      60
gctcaggtct ccagagcctc ccagaagaaa aataggcagc cctccctgaa atatcttggc      120
15 tctcagttt atcctttcca acctggetcc ccttcccag ttcctctccc tactcctgt      180
ctccctcccc aactcaccct actgaactgg gtgcagagca aagccctttt cgcccttttc      240
20 cccatctgga cttctctggc cagttcctct tagtccgatc ccaaagacac tggaacacat      300
ttctaaaggg tcttcttgat cctccaatt cattgagcaa agggctgaaa aagaagcaga      360
gagtaaggta gagccagtga ctgccccca agcccatc atg gaa gaa ggc ttt      414
25                               Met Glu Glu Gly Phe
                               1           5

cga gac cga gca gcg ttc atc cgt ggg gcc aaa gac att gcc aag gaa      462
Arg Asp Arg Ala Ala Phe Ile Arg Gly Ala Lys Asp Ile Ala Lys Glu
10           15           20

gtt aag aag cac gcg gcc aag aag gtg gtg aag ggt ctc gac aga gtc      510
Val Lys Lys His Ala Ala Lys Lys Val Val Lys Gly Leu Asp Arg Val
25           30           35

cag gat gaa tat tcc cga agg tcc tac tcc cgc ttt gag gag gag gag      558
Gln Asp Glu Tyr Ser Arg Arg Ser Tyr Ser Arg Phe Glu Glu Glu Glu
40           45           50

gat gat gat gac ttc cct gcc cct gct gac ggc tat tac cgc gga gaa      606
Asp Asp Asp Asp Phe Pro Ala Pro Ala Asp Gly Tyr Tyr Arg Gly Glu

```

40

45

50

55

60

65

ES 2 311 667 T3

	55	60	65	
5	ggg gcc cag gat gag gag gaa ggt ggc gct tcc agt gat gcc act gag Gly Ala Gln Asp Glu Glu Glu Gly Gly Ala Ser Ser Asp Ala Thr Glu 70 75 80 85	654		
	ggc cac gat gag gat gat gag atc tac gag gga gaa tat cag ggc atc Gly His Asp Glu Asp Asp Glu Ile Tyr Glu Gly Glu Tyr Gln Gly Ile 90 95 100	702		
10	ccc cgg gca gag tct ggg ggc aaa ggc gaa cgg atg gca gat ggg gca Pro Arg Ala Glu Ser Gly Gly Lys Gly Glu Arg Met Ala Asp Gly Ala 105 110 115	750		
15	ccc ctg gct gga gtg aga ggg ggc tta agt gat ggg gag ggt ccc cct Pro Leu Ala Gly Val Arg Gly Gly Leu Ser Asp Gly Glu Gly Pro Pro 120 125 130	798		
20	ggg ggt cgc ggg gag gcg cag cgg cgt aaa gat cgg gaa gaa ttg gct Gly Gly Arg Gly Glu Ala Gln Arg Arg Lys Asp Arg Glu Glu Leu Ala 135 140 145	846		
	cag cag tat gag acc atc ctc cgg gag tgc ggc cat ggt cgc ttc cag Gln Gln Tyr Glu Thr Ile Leu Arg Glu Cys Gly His Gly Arg Phe Gln 150 155 160 165	894		
25	tgg aca ctc tac ttc gtg ctg ggt ctg gcg ctg atg gcc gat ggt gta Trp Thr Leu Tyr Phe Val Leu Gly Leu Ala Leu Met Ala Asp Gly Val 170 175 180	942		
30	gag gtc ttt gtg gtg ggc ttt gtg ctg ccc agt gct gag aaa gat atg Glu Val Phe Val Val Gly Phe Val Leu Pro Ser Ala Glu Lys Asp Met 185 190 195	990		
	tgc ctg tcg gac tcc aac aaa ggc atg cta ggc ctc att gtg tac ctg Cys Leu Ser Asp Ser Asn Lys Gly Met Leu Gly Leu Ile Val Tyr Leu 200 205 210	1038		
35	ggc atg atg gtg ggg gcc ttc ctc tgg gga ggc ctg gct gat cgg ctg Gly Met Met Val Gly Ala Phe Leu Trp Gly Gly Leu Ala Asp Arg Leu 215 220 225	1086		
40	ggg cgg aga cag tgt ctg ctc atc tcg ctc tca gtc aac agc gtc ttc Gly Arg Arg Gln Cys Leu Leu Ile Ser Leu Ser Val Asn Ser Val Phe 230 235 240 245	1134		
45	gct ttc ttc tca tcc ttc gtc cag ggt tat ggc acc ttc ctt ttc tgc Ala Phe Phe Ser Ser Phe Val Gln Gly Tyr Gly Thr Phe Leu Phe Cys 250 255 260	1182		
	cgc ctc ctt tct ggg gtt ggg att ggt ggt tcc atc ccc att gtc ttc Arg Leu Leu Ser Gly Val Gly Ile Gly Gly Ser Ile Pro Ile Val Phe 265 270 275	1230		
50	tcc tat ttt tcg gag ttt ctg gcc cag gag aaa cgt ggg gag cat ttg Ser Tyr Phe Ser Glu Phe Leu Ala Gln Glu Lys Arg Gly Glu His Leu 280 285 290	1278		
55	agc tgg ctc tgt atg ttc tgg atg att ggt ggc gtg tat gca gct gca Ser Trp Leu Cys Met Phe Trp Met Ile Gly Gly Val Tyr Ala Ala Ala 295 300 305	1326		

ES 2 311 667 T3

	atg gcc tgg gcc atc atc ccc cac tat ggg tgg agt ttc cag atg ggc	1374
	Met Ala Trp Ala Ile Ile Pro His Tyr Gly Trp Ser Phe Gln Met Gly	
	310 315 320 325	
5	tct gct tac cag ttc cac agc tgg agg gtc ttt gtc ctc gtc ttt gcc	1422
	Ser Ala Tyr Gln Phe His Ser Trp Arg Val Phe Val Leu Val Phe Ala	
	330 335 340	
10	ttt ccc tct gtg ttt gcc atc ggg gct ctg act acg cag ccg gag agt	1470
	Phe Pro Ser Val Phe Ala Ile Gly Ala Leu Thr Thr Gln Pro Glu Ser	
	345 350 355	
15	ccc cgc ttc ttc tta gag aat ggg aag cac gat gag gcc tgg atg gtg	1518
	Pro Arg Phe Phe Leu Glu Asn Gly Lys His Asp Glu Ala Trp Met Val	
	360 365 370	
20	ctg aag cag gtt cat gac acc aac atg cga gcc aag ggc cat cct gag	1566
	Leu Lys Gln Val His Asp Thr Asn Met Arg Ala Lys Gly His Pro Glu	
	375 380 385	
25	cga gtc ttc tca gta acc cac att aaa acg att cat cag gag gat gaa	1614
	Arg Val Phe Ser Val Thr His Ile Lys Thr Ile His Gln Glu Asp Glu	
	390 395 400 405	
30	ttg att gag atc cag tca gac aca gga acc tgg tac cag cgc tgg gga	1662
	Leu Ile Glu Ile Gln Ser Asp Thr Gly Thr Trp Tyr Gln Arg Trp Gly	
	410 415 420	
35	gtg cgg gct ttg agc ctg ggg ggt cag gtt tgg ggg aac ttc ctc tcc	1710
	Val Arg Ala Leu Ser Leu Gly Gly Gln Val Trp Gly Asn Phe Leu Ser	
	425 430 435	
40	tgc ttc agt cca gag tac cgg cgc atc act ctg atg atg atg ggg gta	1758
	Cys Phe Ser Pro Glu Tyr Arg Arg Ile Thr Leu Met Met Met Gly Val	
	440 445 450	
45	tgg ttc acc atg tcc ttc agc tac tac ggt ttg act gtc tgg ttt ccc	1806
	Trp Phe Thr Met Ser Phe Ser Tyr Tyr Gly Leu Thr Val Trp Phe Pro	
	455 460 465	
50	gac atg atc cgc cat ctc cag gct gtg gac tat gca gcc cga acc aaa	1854
	Asp Met Ile Arg His Leu Gln Ala Val Asp Tyr Ala Ala Arg Thr Lys	
	470 475 480 485	
55	gtg ttc cca ggg gag cgc gtg gag cac gtg aca ttt aac ttc aca ctg	1902
	Val Phe Pro Gly Glu Arg Val Glu His Val Thr Phe Asn Phe Thr Leu	
	490 495 500	
60	gag aat cag atc cac cga ggg gga cag tac ttc aat gac aag ttc atc	1950
	Glu Asn Gln Ile His Arg Gly Gly Gln Tyr Phe Asn Asp Lys Phe Ile	
	505 510 515	
65	ggg ctg cgt ctg aag tca gtg tcc ttt gag gat tcc ctg ttt gag gaa	1998
	Gly Leu Arg Leu Lys Ser Val Ser Phe Glu Asp Ser Leu Phe Glu Glu	
	520 525 530	
70	tgt tac ttt gaa gat gtc aca tcc agc aac aca ttc ttc cgc aac tgc	2046
	Cys Tyr Phe Glu Asp Val Thr Ser Ser Asn Thr Phe Phe Arg Asn Cys	
	535 540 545	

ES 2 311 667 T3

	aca ttc atc aac acc gtg ttc tac aac acg gac ctg ttt gag tac aag	2094
	Thr Phe Ile Asn Thr Val Phe Tyr Asn Thr Asp Leu Phe Glu Tyr Lys	
	550 555 560 565	
5	ttc gtg aac agc cgc ctg gtg aac agc aca ttc ctg cac aat aag gaa	2142
	Phe Val Asn Ser Arg Leu Val Asn Ser Thr Phe Leu His Asn Lys Glu	
	570 575 580	
10	ggc tgc cca cta gat gtg aca ggc acg ggc gaa ggt gcc tac atg gtg	2190
	Gly Cys Pro Leu Asp Val Thr Gly Thr Gly Glu Gly Ala Tyr Met Val	
	585 590 595	
	tac ttt gtc agc ttc ttg ggg aca ctg gct gtg ctc cct gga aat att	2238
	Tyr Phe Val Ser Phe Leu Gly Thr Leu Ala Val Leu Pro Gly Asn Ile	
	600 605 610	
15	gtg tct gct ctg ctc atg gac aag att ggc agg ctc aga atg ctt gct	2286
	Val Ser Ala Leu Leu Met Asp Lys Ile Gly Arg Leu Arg Met Leu Ala	
	615 620 625	
20	ggc tcc agt gtg ttg tcc tgt gtt tcc tgc ttc ttc ctg tct ttt ggg	2334
	Gly Ser Ser Val Leu Ser Cys Val Ser Cys Phe Phe Leu Ser Phe Gly	
	630 635 640 645	
	aac agt gag tca gcc atg atc gct ctg ctc tgc ctt ttt ggg gga gtc	2382
	Asn Ser Glu Ser Ala Met Ile Ala Leu Leu Cys Leu Phe Gly Gly Val	
	650 655 660	
25	agt att gca tcc tgg aac gcg ctg gac gtg ctg act gtt gaa ctc tac	2430
	Ser Ile Ala Ser Trp Asn Ala Leu Asp Val Leu Thr Val Glu Leu Tyr	
	665 670 675	
30	cct tcc gac aag agg acg acg gcc ttc ggc ttc ctg aat gcc ctg tgt	2478
	Pro Ser Asp Lys Arg Thr Thr Ala Phe Gly Phe Leu Asn Ala Leu Cys	
	680 685 690	
35	aag ctg gca gct gta ctg ggc atc agc atc ttc acg tcc ttt gtg gga	2526
	Lys Leu Ala Ala Val Leu Gly Ile Ser Ile Phe Thr Ser Phe Val Gly	
	695 700 705	
	atc acc aag gcc gct ccc atc ctc ttc gcc tca gct gcg ctt gcc ctt	2574
	Ile Thr Lys Ala Ala Pro Ile Leu Phe Ala Ser Ala Ala Leu Ala Leu	
	710 715 720 725	
40	ggc agc tct ctg gct ctg aag ctg cct gag acc cgg gga cag gtg ctg	2622
	Gly Ser Ser Leu Ala Leu Lys Leu Pro Glu Thr Arg Gly Gln Val Leu	
	730 735 740	
45	cag tga gggatggggg agtgtctcag gggcttttagg gatggcaggc acactgtgac	2678
	Gln	
	caataatttc ttttatccct accctgccct gctgtcctgg tctactccg tgtttggtgt	2738
50	cttagctgtg tgcctgtgtg catgtgtgtg accctgacgg gcaggggcta cggggagggt	2798
	ccctttgtcc catgtttggg aggagggact cccacactgc tgccaccctc aactttgcac	2858
	aaggagaagg ctgagctgca tccttctctc cctcagtgtt agcaggggtg gggacgactg	2918
55	tttctctgct ccaggtgttc cagaatttct gcctttccca tcattccctc cgcctaggcc	2978
60		
65		

ES 2 311 667 T3

ctggtgaacc acaggtatgg agttatagtg ggggctgagg cttggaccaa aagaacttct 3038
 tgagtgggag cctcccaagg atgctgggga gtagcaataa accttagcct ccgttttcac 3098
 5 ctcaattcag gctacaagtg tgaagcctgg attttatgga attagttttc tgattcttat 3158
 ttatatgtaa gttctgaggc agcttagctg gactgtgtgt ggatgtatac atacactcat 3218
 atgtgtgtgt gtgtgcgcgg cgtgcgtgtg cgtgtgtgta tgtgtgcat ggggtagggg 3278
 10 taccactata ctgttcaatt ataagccaag agtagtagtt tcagttagca cacacacaac 3338
 actgtttttc tatcgtaact cccagaatct tgtacctgtg ttggggctgc aggcagaagt 3398
 ccttggtcag gctgggggtga gccctgcaat cttggggggac ctgagggcac ctgacaagga 3458
 15 ctttctcctt cctctagaga gggtctaccc actagccaca gccctcccat ctgacctgtc 3518
 cacacaggca gtgtatcaga ggaaagaaag ggaaaataac caggcacaca tggcacaacc 3578
 20 agcaggtctg aaagcacaaag aagctggggc aggggcagga agcaggggtc catcccctaa 3638
 cccttctcaa aaaggctggg tcgtgaggga cccctaattgc agggacctga agcctcagtt 3698
 tccccatctt gcccttccac agaacagcct ctgtaggtag agctgcccc cgctctaccc 3758
 25 tactcttggt gctgctttct ttggtactct tcccactccc accgtagctg tgacgtgttg 3818
 tagtttttag ctgtttgtaa aatggt 3844

30

10

742

PRT

Rattus norvegicus

35

Met Glu Glu Gly Phe Arg Asp Arg Ala Ala Phe Ile Arg Gly Ala Lys
 1 5 10 15

40

Asp Ile Ala Lys Glu Val Lys Lys His Ala Ala Lys Lys Val Val Lys
 20 25 30

45

Gly Leu Asp Arg Val Gln Asp Glu Tyr Ser Arg Arg Ser Tyr Ser Arg
 35 40 45

50

Phe Glu Glu Glu Glu Asp Asp Asp Asp Phe Pro Ala Pro Ala Asp Gly
 50 55 60

55

Tyr Tyr Arg Gly Glu Gly Ala Gln Asp Glu Glu Glu Gly Gly Ala Ser

60

65

ES 2 311 667 T3

	65	70	75	80
5	Ser Asp Ala Thr Glu Gly His Asp Glu Asp Asp Glu Ile Tyr Glu Gly	85	90	95
10	Glu Tyr Gln Gly Ile Pro Arg Ala Glu Ser Gly Gly Lys Gly Glu Arg	100	105	110
15	Met Ala Asp Gly Ala Pro Leu Ala Gly Val Arg Gly Gly Leu Ser Asp	115	120	125
20	Gly Glu Gly Pro Pro Gly Gly Arg Gly Glu Ala Gln Arg Arg Lys Asp	130	135	140
25	Arg Glu Glu Leu Ala Gln Gln Tyr Glu Thr Ile Leu Arg Glu Cys Gly	145	150	155
30	His Gly Arg Phe Gln Trp Thr Leu Tyr Phe Val Leu Gly Leu Ala Leu	165	170	175
35	Met Ala Asp Gly Val Glu Val Phe Val Val Gly Phe Val Leu Pro Ser	180	185	190
40	Ala Glu Lys Asp Met Cys Leu Ser Asp Ser Asn Lys Gly Met Leu Gly	195	200	205
45	Leu Ile Val Tyr Leu Gly Met Met Val Gly Ala Phe Leu Trp Gly Gly	210	215	220
50	Leu Ala Asp Arg Leu Gly Arg Arg Gln Cys Leu Leu Ile Ser Leu Ser	225	230	235
55	Val Asn Ser Val Phe Ala Phe Phe Ser Ser Phe Val Gln Gly Tyr Gly	245	250	255
60	Thr Phe Leu Phe Cys Arg Leu Leu Ser Gly Val Gly Ile Gly Gly Ser	260	265	270
65	Ile Pro Ile Val Phe Ser Tyr Phe Ser Glu Phe Leu Ala Gln Glu Lys	275	280	285
	Arg Gly Glu His Leu Ser Trp Leu Cys Met Phe Trp Met Ile Gly Gly	290	295	300
	Val Tyr Ala Ala Ala Met Ala Trp Ala Ile Ile Pro His Tyr Gly Trp	305	310	315
				320

ES 2 311 667 T3

	Ser	Phe	Gln	Met	Gly	Ser	Ala	Tyr	Gln	Phe	His	Ser	Trp	Arg	Val	Phe	
					325					330					335		
5	Val	Leu	Val	Phe	Ala	Phe	Pro	Ser	Val	Phe	Ala	Ile	Gly	Ala	Leu	Thr	
				340					345					350			
10	Thr	Gln	Pro	Glu	Ser	Pro	Arg	Phe	Phe	Leu	Glu	Asn	Gly	Lys	His	Asp	
			355					360					365				
15	Glu	Ala	Trp	Met	Val	Leu	Lys	Gln	Val	His	Asp	Thr	Asn	Met	Arg	Ala	
		370					375					380					
20	Lys	Gly	His	Pro	Glu	Arg	Val	Phe	Ser	Val	Thr	His	Ile	Lys	Thr	Ile	
	385					390					395					400	
25	His	Gln	Glu	Asp	Glu	Leu	Ile	Glu	Ile	Gln	Ser	Asp	Thr	Gly	Thr	Trp	
				405						410					415		
30	Tyr	Gln	Arg	Trp	Gly	Val	Arg	Ala	Leu	Ser	Leu	Gly	Gly	Gln	Val	Trp	
				420					425					430			
35	Gly	Asn	Phe	Leu	Ser	Cys	Phe	Ser	Pro	Glu	Tyr	Arg	Arg	Ile	Thr	Leu	
			435					440					445				
40	Met	Met	Met	Gly	Val	Trp	Phe	Thr	Met	Ser	Phe	Ser	Tyr	Tyr	Gly	Leu	
		450					455						460				
45	Thr	Val	Trp	Phe	Pro	Asp	Met	Ile	Arg	His	Leu	Gln	Ala	Val	Asp	Tyr	
	465					470					475				480		
50	Ala	Ala	Arg	Thr	Lys	Val	Phe	Pro	Gly	Glu	Arg	Val	Glu	His	Val	Thr	
					485					490					495		
55	Phe	Asn	Phe	Thr	Leu	Glu	Asn	Gln	Ile	His	Arg	Gly	Gly	Gln	Tyr	Phe	
				500				505						510			
60	Asn	Asp	Lys	Phe	Ile	Gly	Leu	Arg	Leu	Lys	Ser	Val	Ser	Phe	Glu	Asp	
			515					520					525				
65	Ser	Leu	Phe	Glu	Glu	Cys	Tyr	Phe	Glu	Asp	Val	Thr	Ser	Ser	Asn	Thr	
		530					535					540					
70	Phe	Phe	Arg	Asn	Cys	Thr	Phe	Ile	Asn	Thr	Val	Phe	Tyr	Asn	Thr	Asp	
	545					550					555					560	

ES 2 311 667 T3

	Leu	Phe	Glu	Tyr	Lys	Phe	Val	Asn	Ser	Arg	Leu	Val	Asn	Ser	Thr	Phe	
					565					570					575		
5	Leu	His	Asn	Lys	Glu	Gly	Cys	Pro	Leu	Asp	Val	Thr	Gly	Thr	Gly	Glu	
				580					585					590			
10	Gly	Ala	Tyr	Met	Val	Tyr	Phe	Val	Ser	Phe	Leu	Gly	Thr	Leu	Ala	Val	
			595					600					605				
15	Leu	Pro	Gly	Asn	Ile	Val	Ser	Ala	Leu	Leu	Met	Asp	Lys	Ile	Gly	Arg	
		610					615					620					
20	Leu	Arg	Met	Leu	Ala	Gly	Ser	Ser	Val	Leu	Ser	Cys	Val	Ser	Cys	Phe	
	625					630					635					640	
25	Phe	Leu	Ser	Phe	Gly	Asn	Ser	Glu	Ser	Ala	Met	Ile	Ala	Leu	Leu	Cys	
					645					650						655	
30	Leu	Phe	Gly	Gly	Val	Ser	Ile	Ala	Ser	Trp	Asn	Ala	Leu	Asp	Val	Leu	
				660					665					670			
35	Thr	Val	Glu	Leu	Tyr	Pro	Ser	Asp	Lys	Arg	Thr	Thr	Ala	Phe	Gly	Phe	
			675					680					685				
40	Leu	Asn	Ala	Leu	Cys	Lys	Leu	Ala	Ala	Val	Leu	Gly	Ile	Ser	Ile	Phe	
		690					695					700					
45	Thr	Ser	Phe	Val	Gly	Ile	Thr	Lys	Ala	Ala	Pro	Ile	Leu	Phe	Ala	Ser	
	705					710					715					720	
50	Ala	Ala	Leu	Ala	Leu	Gly	Ser	Ser	Leu	Ala	Leu	Lys	Leu	Pro	Glu	Thr	
					725					730					735		
55	Arg	Gly	Gln	Val	Leu	Gln											
				740													

ES 2 311 667 T3

11

2052

ADN

5 *Rattus norvegicus*

CDS

(1)..(2052)

10

atg gat gac tac agg tat cgg gac aac tat gag ggc tat gcc cct aat 48
Met Asp Asp Tyr Arg Tyr Arg Asp Asn Tyr Glu Gly Tyr Ala Pro Asn
1 5 10 15

15

gat ggc tac tac cgg ggc aat gag cag aac ccg gaa gaa gat gca cag 96
Asp Gly Tyr Tyr Arg Gly Asn Glu Gln Asn Pro Glu Glu Asp Ala Gln
20 25 30

20

agc gat gtt aca gaa ggc cac gat gaa gag gat gag atc tat gag ggc 144
Ser Asp Val Thr Glu Gly His Asp Glu Glu Asp Glu Ile Tyr Glu Gly
35 40 45

25

gag tac caa ggc atc cct cat cca gat gat gtc aag tct aag cag act 192
Glu Tyr Gln Gly Ile Pro His Pro Asp Asp Val Lys Ser Lys Gln Thr
50 55 60

aag atg gca ccg tcc aga gca gat ggc ctt cgg ggc cag gca gac ctg 240
Lys Met Ala Pro Ser Arg Ala Asp Gly Leu Arg Gly Gln Ala Asp Leu
65 70 75 80

30

atg gct gag aga atg gaa gat gag gag cag ctc gct cac cag tac gag 288
Met Ala Glu Arg Met Glu Asp Glu Glu Gln Leu Ala His Gln Tyr Glu
85 90 95

35

acc atc att gat gag tgt ggc cat ggg cgc ttc cag tgg acc ctc ttt 336
Thr Ile Ile Asp Glu Cys Gly His Gly Arg Phe Gln Trp Thr Leu Phe
100 105 110

40

ttc gtc ttg gtc ttg gcc ttg atg gct gac gga gtg gaa gtg ttt gtg 384
Phe Val Leu Val Leu Ala Leu Met Ala Asp Gly Val Glu Val Phe Val
115 120 125

gtg agc ttt gct ctg cca agt gca gag aaa gat atg tgt ctg tca agt 432
Val Ser Phe Ala Leu Pro Ser Ala Glu Lys Asp Met Cys Leu Ser Ser
130 135 140

45

tcc aag aaa gga atg ctc ggg ctg att gtc tac cta gga atg atg gca 480
Ser Lys Lys Gly Met Leu Gly Leu Ile Val Tyr Leu Gly Met Met Ala
145 150 155 160

50

gga gcc ttc atc ctg ggg ggc ctg gct gat aaa ctg gga agg aag aag 528
Gly Ala Phe Ile Leu Gly Gly Leu Ala Asp Lys Leu Gly Arg Lys Lys
165 170 175

gtc ctc agc atg tcc ttg gct atc aat gct tcc ttt gcc tcc ctc tcc 576
Val Leu Ser Met Ser Leu Ala Ile Asn Ala Ser Phe Ala Ser Leu Ser
180 185 190

55

tcc ttc gtg cag gga tat gga gct ttc ctc ttc tgc aga ctc atc tca 624
Ser Phe Val Gln Gly Tyr Gly Ala Phe Leu Phe Cys Arg Leu Ile Ser
195 200 205

60

65

ES 2 311 667 T3

	ggc ata ggt att ggg ggc tcc ctg cca att gtt ttt gcc tac ttt tct Gly Ile Gly Ile Gly Gly Ser Leu Pro Ile Val Phe Ala Tyr Phe Ser 210 215 220	672
5	gag ttc tta tca cgg gag aaa cgc ggt gag cat ctc agc tgg ctg ggt Glu Phe Leu Ser Arg Glu Lys Arg Gly Glu His Leu Ser Trp Leu Gly 225 230 235 240	720
10	atc ttc tgg atg act ggg ggc atc tac gca tct gcc atg gcc tgg agc Ile Phe Trp Met Thr Gly Gly Ile Tyr Ala Ser Ala Met Ala Trp Ser 245 250 255	768
15	atc att cca cac tat ggc tgg ggc ttc agc atg gga acc aat tat cac Ile Ile Pro His Tyr Gly Trp Gly Phe Ser Met Gly Thr Asn Tyr His 260 265 270	816
	ttc cac agc tgg aga gtg ttt gtc atc gtc tgt gct ctg cct gcc act Phe His Ser Trp Arg Val Phe Val Ile Val Cys Ala Leu Pro Ala Thr 275 280 285	864
20	gtg tcc atg gtg gcc ctg aag ttc atg cca gaa agc ccc agg ttc ctg Val Ser Met Val Ala Leu Lys Phe Met Pro Glu Ser Pro Arg Phe Leu 290 295 300	912
25	ctg gag atg ggc aag cat gat gaa gcc tgg atg att ctc aag caa gtc Leu Glu Met Gly Lys His Asp Glu Ala Trp Met Ile Leu Lys Gln Val 305 310 315 320	960
	cat gac acc aac atg aga gct aag ggg acc cct gag aag gtg ttc acg His Asp Thr Asn Met Arg Ala Lys Gly Thr Pro Glu Lys Val Phe Thr 325 330 335	1008
30	gtt tcc cac atc aaa act ccc aag caa atg gat gaa ttc att gag atc Val Ser His Ile Lys Thr Pro Lys Gln Met Asp Glu Phe Ile Glu Ile 340 345 350	1056
35	cag agt tca aca ggg act tgg tac cag cgc tgg ttg gtc agg ttc atg Gln Ser Ser Thr Gly Thr Trp Tyr Gln Arg Trp Leu Val Arg Phe Met 355 360 365	1104
	acc att ttc aaa cag gtg tgg gat aac gcc ttg tac tgt gtg atg gga Thr Ile Phe Lys Gln Val Trp Asp Asn Ala Leu Tyr Cys Val Met Gly 370 375 380	1152
40	ccc tac aga atg aac acc ctg att ctg gct gtg gtc tgg ttc acc atg Pro Tyr Arg Met Asn Thr Leu Ile Leu Ala Val Val Trp Phe Thr Met 385 390 395 400	1200
45	gcc tta agt tac tat ggc ctg aca gtg tgg ttc ccc gac atg atc cgg Ala Leu Ser Tyr Tyr Gly Leu Thr Val Trp Phe Pro Asp Met Ile Arg 405 410 415	1248
	tat ttc cag gat gaa gaa tat aag tct aaa atg aag gtg ttt ttt ggt Tyr Phe Gln Asp Glu Glu Tyr Lys Ser Lys Met Lys Val Phe Phe Gly 420 425 430	1296
50	gag cac gtg cat ggc gcc aca atc aac ttc acc atg gaa aac cag atc Glu His Val His Gly Ala Thr Ile Asn Phe Thr Met Glu Asn Gln Ile 435 440 445	1344
55	cac caa cat ggg aag ctt gtg aac gat aag ttc ata aag atg tac ttt	1392

ES 2 311 667 T3

	His	Gln	His	Gly	Lys	Leu	Val	Asn	Asp	Lys	Phe	Ile	Lys	Met	Tyr	Phe	
	450						455					460					
5	aag	cat	gtc	ctc	ttt	gag	gac	aca	ttc	ttt	gac	aaa	tgc	tat	ttt	gaa	1440
	Lys	His	Val	Leu	Phe	Glu	Asp	Thr	Phe	Phe	Asp	Lys	Cys	Tyr	Phe	Glu	
	465					470					475					480	
	gat	gtg	aca	tcc	aca	gat	act	tat	ttc	aag	aac	tgc	acc	att	gaa	tcg	1488
10	Asp	Val	Thr	Ser	Thr	Asp	Thr	Tyr	Phe	Lys	Asn	Cys	Thr	Ile	Glu	Ser	
					485					490					495		
	act	acc	ttc	tac	aac	aca	gac	ctc	tac	aaa	cac	aag	ttc	att	gac	tgt	1536
	Thr	Thr	Phe	Tyr	Asn	Thr	Asp	Leu	Tyr	Lys	His	Lys	Phe	Ile	Asp	Cys	
				500					505					510			
15	cgg	ttt	atc	aat	tcc	acc	ttt	ctg	gag	cag	aag	gag	ggc	tgc	cac	atg	1584
	Arg	Phe	Ile	Asn	Ser	Thr	Phe	Leu	Glu	Gln	Lys	Glu	Gly	Cys	His	Met	
		515						520					525				
20	gac	ttt	gaa	gag	gac	aat	gat	ttt	ctg	att	tac	ctc	gtc	agc	ttc	ctc	1632
	Asp	Phe	Glu	Glu	Asp	Asn	Asp	Phe	Leu	Ile	Tyr	Leu	Val	Ser	Phe	Leu	
		530					535					540					
	ggc	agc	ctg	tct	gtc	ttg	cct	ggg	aac	ata	att	tct	gcc	ctg	ctc	atg	1680
25	Gly	Ser	Leu	Ser	Val	Leu	Pro	Gly	Asn	Ile	Ile	Ser	Ala	Leu	Leu	Met	
	545					550				555						560	
	gac	aga	atc	gga	aga	ctt	aag	atg	att	ggg	ggc	tcc	atg	ctc	atc	tct	1728
	Asp	Arg	Ile	Gly	Arg	Leu	Lys	Met	Ile	Gly	Gly	Ser	Met	Leu	Ile	Ser	
					565					570					575		
30	gca	gtc	tgc	tgc	ttc	ttc	ctg	ttt	ttt	ggc	aac	agc	gag	tct	gcg	atg	1776
	Ala	Val	Cys	Cys	Phe	Phe	Leu	Phe	Phe	Gly	Asn	Ser	Glu	Ser	Ala	Met	
					580				585					590			
35	atc	ggc	tgg	caa	tgc	ctg	ttc	tgt	ggg	acc	agc	att	gca	gcc	tgg	aat	1824
	Ile	Gly	Trp	Gln	Cys	Leu	Phe	Cys	Gly	Thr	Ser	Ile	Ala	Ala	Trp	Asn	
			595					600					605				
	gct	ctg	gat	gtg	atc	aca	gta	gag	ctg	tat	ccc	acc	aac	cag	agg	gcc	1872
40	Ala	Leu	Asp	Val	Ile	Thr	Val	Glu	Leu	Tyr	Pro	Thr	Asn	Gln	Arg	Ala	
		610					615					620					
	act	gcc	ttc	ggc	atc	ctc	aat	gga	ctg	tgc	aaa	ctt	ggg	gcc	atc	ctg	1920
	Thr	Ala	Phe	Gly	Ile	Leu	Asn	Gly	Leu	Cys	Lys	Leu	Gly	Ala	Ile	Leu	
						630					635					640	
45	gga	aac	act	atc	ttt	gct	tct	ttt	gtt	ggg	atc	acc	aaa	gtg	gtc	ccc	1968
	Gly	Asn	Thr	Ile	Phe	Ala	Ser	Phe	Val	Gly	Ile	Thr	Lys	Val	Val	Pro	
					645					650					655		
	atc	ctt	ctg	gct	gct	gct	tct	ctg	gtt	gga	ggg	ggc	ttg	gtt	gcc	ctt	2016
50	Ile	Leu	Leu	Ala	Ala	Ala	Ser	Leu	Val	Gly	Gly	Gly	Leu	Val	Ala	Leu	
				660					665					670			
	cga	ctg	cca	gag	act	cga	gag	cag	gtc	ctg	atg	tga					2052
55	Arg	Leu	Pro	Glu	Thr	Arg	Glu	Gln	Val	Leu	Met						
			675					680									
60																	
65																	

ES 2 311 667 T3

12
683
PRT

5 *Rattus norvegicus*

10	Met Asp Asp Tyr Arg Tyr Arg Asp Asn Tyr Glu Gly Tyr Ala Pro Asn	1 5 10 15
15	Asp Gly Tyr Tyr Arg Gly Asn Glu Gln Asn Pro Glu Glu Asp Ala Gln	20 25 30
20	Ser Asp Val Thr Glu Gly His Asp Glu Glu Asp Glu Ile Tyr Glu Gly	35 40 45
25	Glu Tyr Gln Gly Ile Pro His Pro Asp Asp Val Lys Ser Lys Gln Thr	50 55 60
30	Lys Met Ala Pro Ser Arg Ala Asp Gly Leu Arg Gly Gln Ala Asp Leu	65 70 75 80
35	Met Ala Glu Arg Met Glu Asp Glu Glu Gln Leu Ala His Gln Tyr Glu	85 90 95
40	Thr Ile Ile Asp Glu Cys Gly His Gly Arg Phe Gln Trp Thr Leu Phe	100 105 110
45	Phe Val Leu Val Leu Ala Leu Met Ala Asp Gly Val Glu Val Phe Val	115 120 125
50	Val Ser Phe Ala Leu Pro Ser Ala Glu Lys Asp Met Cys Leu Ser Ser	130 135 140
55	Ser Lys Lys Gly Met Leu Gly Leu Ile Val Tyr Leu Gly Met Met Ala	145 150 155 160
60	Gly Ala Phe Ile Leu Gly Gly Leu Ala Asp Lys Leu Gly Arg Lys Lys	165 170 175
65	Val Leu Ser Met Ser Leu Ala Ile Asn Ala Ser Phe Ala Ser Leu Ser	180 185 190
	Ser Phe Val Gln Gly Tyr Gly Ala Phe Leu Phe Cys Arg Leu Ile Ser	195 200 205

ES 2 311 667 T3

	Gly	Ile	Gly	Ile	Gly	Gly	Ser	Leu	Pro	Ile	Val	Phe	Ala	Tyr	Phe	Ser	
	210						215					220					
5	Glu	Phe	Leu	Ser	Arg	Glu	Lys	Arg	Gly	Glu	His	Leu	Ser	Trp	Leu	Gly	
	225					230				235						240	
10	Ile	Phe	Trp	Met	Thr	Gly	Gly	Ile	Tyr	Ala	Ser	Ala	Met	Ala	Trp	Ser	
					245					250					255		
15	Ile	Ile	Pro	His	Tyr	Gly	Trp	Gly	Phe	Ser	Met	Gly	Thr	Asn	Tyr	His	
				260					265					270			
20	Phe	His	Ser	Trp	Arg	Val	Phe	Val	Ile	Val	Cys	Ala	Leu	Pro	Ala	Thr	
			275				280					285					
25	Val	Ser	Met	Val	Ala	Leu	Lys	Phe	Met	Pro	Glu	Ser	Pro	Arg	Phe	Leu	
	290					295				300							
30	Leu	Glu	Met	Gly	Lys	His	Asp	Glu	Ala	Trp	Met	Ile	Leu	Lys	Gln	Val	
	305				310					315						320	
35	His	Asp	Thr	Asn	Met	Arg	Ala	Lys	Gly	Thr	Pro	Glu	Lys	Val	Phe	Thr	
				325						330					335		
40	Val	Ser	His	Ile	Lys	Thr	Pro	Lys	Gln	Met	Asp	Glu	Phe	Ile	Glu	Ile	
			340				345					350					
45	Gln	Ser	Ser	Thr	Gly	Thr	Trp	Tyr	Gln	Arg	Trp	Leu	Val	Arg	Phe	Met	
			355				360					365					
50	Thr	Ile	Phe	Lys	Gln	Val	Trp	Asp	Asn	Ala	Leu	Tyr	Cys	Val	Met	Gly	
	370				375					380							
55	Pro	Tyr	Arg	Met	Asn	Thr	Leu	Ile	Leu	Ala	Val	Val	Trp	Phe	Thr	Met	
	385				390					395					400		
60	Ala	Leu	Ser	Tyr	Tyr	Gly	Leu	Thr	Val	Trp	Phe	Pro	Asp	Met	Ile	Arg	
				405			410							415			
65	Tyr	Phe	Gln	Asp	Glu	Glu	Tyr	Lys	Ser	Lys	Met	Lys	Val	Phe	Phe	Gly	
			420				425					430					
70	Glu	His	Val	His	Gly	Ala	Thr	Ile	Asn	Phe	Thr	Met	Glu	Asn	Gln	Ile	
			435				440					445					

ES 2 311 667 T3

	His	Gln	His	Gly	Lys	Leu	Val	Asn	Asp	Lys	Phe	Ile	Lys	Met	Tyr	Phe	
	450						455					460					
5	Lys	His	Val	Leu	Phe	Glu	Asp	Thr	Phe	Phe	Asp	Lys	Cys	Tyr	Phe	Glu	
	465					470					475					480	
10	Asp	Val	Thr	Ser	Thr	Asp	Thr	Tyr	Phe	Lys	Asn	Cys	Thr	Ile	Glu	Ser	
					485					490					495		
15	Thr	Thr	Phe	Tyr	Asn	Thr	Asp	Leu	Tyr	Lys	His	Lys	Phe	Ile	Asp	Cys	
				500					505					510			
20	Arg	Phe	Ile	Asn	Ser	Thr	Phe	Leu	Glu	Gln	Lys	Glu	Gly	Cys	His	Met	
			515					520					525				
25	Asp	Phe	Glu	Glu	Asp	Asn	Asp	Phe	Leu	Ile	Tyr	Leu	Val	Ser	Phe	Leu	
	530					535						540					
30	Gly	Ser	Leu	Ser	Val	Leu	Pro	Gly	Asn	Ile	Ile	Ser	Ala	Leu	Leu	Met	
	545					550					555					560	
35	Asp	Arg	Ile	Gly	Arg	Leu	Lys	Met	Ile	Gly	Gly	Ser	Met	Leu	Ile	Ser	
					565					570					575		
40	Ala	Val	Cys	Cys	Phe	Phe	Leu	Phe	Phe	Gly	Asn	Ser	Glu	Ser	Ala	Met	
				580					585					590			
45	Ile	Gly	Trp	Gln	Cys	Leu	Phe	Cys	Gly	Thr	Ser	Ile	Ala	Ala	Trp	Asn	
			595					600					605				
50	Ala	Leu	Asp	Val	Ile	Thr	Val	Glu	Leu	Tyr	Pro	Thr	Asn	Gln	Arg	Ala	
	610						615					620					
55	Thr	Ala	Phe	Gly	Ile	Leu	Asn	Gly	Leu	Cys	Lys	Leu	Gly	Ala	Ile	Leu	
	625					630					635					640	
60	Gly	Asn	Thr	Ile	Phe	Ala	Ser	Phe	Val	Gly	Ile	Thr	Lys	Val	Val	Pro	
				645					650					655			
65	Ile	Leu	Leu	Ala	Ala	Ala	Ser	Leu	Val	Gly	Gly	Gly	Leu	Val	Ala	Leu	
				660					665					670			
	Arg	Leu	Pro	Glu	Thr	Arg	Glu	Gln	Val	Leu	Met						
		675						680									

ES 2 311 667 T3

13

2622

ADN

5 *Rattus norvegicus*

CDS

(224)..(2407)

10

gcgcgctgca ggaagagtgg cagaccgaag cggcctcggg ctgcaaacgg aggggcgctc 60

gcgcggcgac ggctgcaggg ctgacaccgc tcagggcagg ggggtcccag gcggctggaa 120

15

cgctctattc tgaactgtga gtggatgatg ctgttcagc caagctgctg aacacactcc 180

gtggactctt ccctgctgtg ccttgcccat cggccgagat aaa atg gaa gac tcc 235
Met Glu Asp Ser
1

20

tac aag gat agg act tca ctg atg aag ggc gcc aag gac att gcc aaa 283
Tyr Lys Asp Arg Thr Ser Leu Met Lys Gly Ala Lys Asp Ile Ala Lys
5 10 15 20

25

gag gtg aag aag caa aca gtg aag aag gtg aac cag gca gtg gac cgg 331
Glu Val Lys Lys Gln Thr Val Lys Lys Val Asn Gln Ala Val Asp Arg
25 30 35

30

gcc cag gat gaa tac acc cag cgg tcc tac agt cga ttc cag gat gaa 379
Ala Gln Asp Glu Tyr Thr Gln Arg Ser Tyr Ser Arg Phe Gln Asp Glu
40 45 50

gat gat gat gat gac tac tac cca cct gga gaa acc tac agt ggg gag 427
Asp Asp Asp Asp Asp Tyr Tyr Pro Pro Gly Glu Thr Tyr Ser Gly Glu
55 60 65

35

gcc aat gat gat gaa ggc tca agt gaa gcc act gag ggt cac gat gaa 475
Ala Asn Asp Asp Glu Gly Ser Ser Glu Ala Thr Glu Gly His Asp Glu
70 75 80

40

gaa gac gag atc tat gaa ggg gaa tac cag ggc atc ccc agc acg aac 523
Glu Asp Glu Ile Tyr Glu Gly Glu Tyr Gln Gly Ile Pro Ser Thr Asn
85 90 95 100

45

caa ggg aag gac agc ata gtg tct gta gga caa ccc aaa gga gat gag 571
Gln Gly Lys Asp Ser Ile Val Ser Val Gly Gln Pro Lys Gly Asp Glu
105 110 115

tac aag gac cgc aga gag ctg gag tca gag agg agg gct gat gag gag 619
Tyr Lys Asp Arg Arg Glu Leu Glu Ser Glu Arg Arg Ala Asp Glu Glu

50

55

60

65

ES 2 311 667 T3

	120	125	130	
5	gag ctc gcc cag cag tat gag ctg ata atc caa gag tgt ggc cat ggc Glu Leu Ala Gln Gln Tyr Glu Leu Ile Ile Gln Glu Cys Gly His Gly 135 140 145	667		
10	cgt ttc cag tgg gcc ctt ttc ttc gtc ctg ggc atg gct ctc atg gca Arg Phe Gln Trp Ala Leu Phe Phe Val Leu Gly Met Ala Leu Met Ala 150 155 160	715		
15	gac ggc gtg gag gtg ttt gtg gtg ggc ttt gtg tta ccc agt gca gag Asp Gly Val Glu Val Phe Val Val Gly Phe Val Leu Pro Ser Ala Glu 165 170 175 180	763		
20	aca gac cta tgc ata ccg aat tca gga tct gga tgg cta ggc agc ata Thr Asp Leu Cys Ile Pro Asn Ser Gly Ser Gly Trp Leu Gly Ser Ile 185 190 195	811		
25	gtg tac ctc ggg atg atg gtg ggg ggc ttc ttc tgg gga gga ctg gca Val Tyr Leu Gly Met Met Val Gly Ala Phe Phe Trp Gly Gly Leu Ala 200 205 210	859		
30	gac aaa gtg gga agg aag cag tct ctt ctg att tgc atg tcc gtc aac Asp Lys Val Gly Arg Lys Gln Ser Leu Leu Ile Cys Met Ser Val Asn 215 220 225	907		
35	gga ttc ttt gcc ttc ctt tct tca ttt gtc caa ggt tac ggc ttc ttt Gly Phe Phe Ala Phe Leu Ser Ser Phe Val Gln Gly Tyr Gly Phe Phe 230 235 240	955		
40	ctc ctc tgt cgt ttg ctt tca gga ttc ggg att gga ggc gcc att ccc Leu Leu Cys Arg Leu Leu Ser Gly Phe Gly Ile Gly Gly Ala Ile Pro 245 250 255 260	1003		
45	act gtg ttc tcc tac ttt gct gaa gtc ctg gcc cgg gag aag cgc ggt Thr Val Phe Ser Tyr Phe Ala Glu Val Leu Ala Arg Glu Lys Arg Gly 265 270 275	1051		
50	gag cac ctc agt tgg ctc tgc atg ttc tgg atg att ggc ggt atc tat Glu His Leu Ser Trp Leu Cys Met Phe Trp Met Ile Gly Gly Ile Tyr 280 285 290	1099		
55	gct tca gcc atg gcc tgg gcc atc atc ccc cac tat ggg tgg agc ttc Ala Ser Ala Met Ala Trp Ala Ile Ile Pro His Tyr Gly Trp Ser Phe 295 300 305	1147		
60	agc atg ggc tca gcc tac cag ttc cac agc tgg cgc gtc ttc gtc atc Ser Met Gly Ser Ala Tyr Gln Phe His Ser Trp Arg Val Phe Val Ile 310 315 320	1195		
65	gtc tgt gcc ctc ccg tgc gtc tcc tcg gtg gtg gcc ctc acc ttc atg Val Cys Ala Leu Pro Cys Val Ser Ser Val Val Ala Leu Thr Phe Met 325 330 335 340	1243		
70	ccc gaa agc cct cgg ttc ttg ctg gag gta gga aaa cat gat gaa gcc Pro Glu Ser Pro Arg Phe Leu Leu Glu Val Gly Lys His Asp Glu Ala 345 350 355	1291		
75	tgg atg att ctg aag cta att cat gat acc aac atg aga gcc cgg ggc Trp Met Ile Leu Lys Leu Ile His Asp Thr Asn Met Arg Ala Arg Gly 360 365 370	1339		

ES 2 311 667 T3

	cag cca gag aag gtc ttc acg gta aat aaa atc aag act ccc aag caa	1387
	Gln Pro Glu Lys Val Phe Thr Val Asn Lys Ile Lys Thr Pro Lys Gln	
	375 380 385	
5	ata gat gag ctg att gag att gag agc gac aca gga acc tgg tac cgg	1435
	Ile Asp Glu Leu Ile Glu Ile Glu Ser Asp Thr Gly Thr Trp Tyr Arg	
	390 395 400	
10	agg tgt ttt gtt cgg atc cgc aca gaa ctg tac gga att tgg ttg act	1483
	Arg Cys Phe Val Arg Ile Arg Thr Glu Leu Tyr Gly Ile Trp Leu Thr	
	405 410 415 420	
15	ttt atg aga tgc ttc aac tac ccg gtc agg gaa aac acc ata aag ctt	1531
	Phe Met Arg Cys Phe Asn Tyr Pro Val Arg Glu Asn Thr Ile Lys Leu	
	425 430 435	
20	acg att gtt tgg ttc acc ctg tcc ttt ggg tac tat gga ctg tcc gtt	1579
	Thr Ile Val Trp Phe Thr Leu Ser Phe Gly Tyr Tyr Gly Leu Ser Val	
	440 445 450	
25	tgg ttc cca gat gtc att aaa cac ctc cag tct gac gag tac gcc ctg	1627
	Trp Phe Pro Asp Val Ile Lys His Leu Gln Ser Asp Glu Tyr Ala Leu	
	455 460 465	
30	ctg act cgg aat gtg cag aag gat aaa tat gca aac ttt agc att aac	1675
	Leu Thr Arg Asn Val Gln Lys Asp Lys Tyr Ala Asn Phe Ser Ile Asn	
	470 475 480	
35	ttc acc atg gaa aac cag gtc cac acc gga atg gaa tat gac aat ggc	1723
	Phe Thr Met Glu Asn Gln Val His Thr Gly Met Glu Tyr Asp Asn Gly	
	485 490 495 500	
40	agg ttc ctc gga gtc aaa ttc aaa tcg gta acc ttc aag gat tca gtg	1771
	Arg Phe Leu Gly Val Lys Phe Lys Ser Val Thr Phe Lys Asp Ser Val	
	505 510 515	
45	ttt aag tcc tgc acc ttt gac gat gtg acc tca gtc aac acc tac ttc	1819
	Phe Lys Ser Cys Thr Phe Asp Asp Val Thr Ser Val Asn Thr Tyr Phe	
	520 525 530	
50	aag aac tgc acg ttt att gat acc ctt ttt gag aac aca gat ttt gag	1867
	Lys Asn Cys Thr Phe Ile Asp Thr Leu Phe Glu Asn Thr Asp Phe Glu	
	535 540 545	
55	ccc tat aaa ttc ata gac agc gag ttt caa aac tgc tcg ttt ctt cac	1915
	Pro Tyr Lys Phe Ile Asp Ser Glu Phe Gln Asn Cys Ser Phe Leu His	
	550 555 560	
60	aat aag acg gga tgt cag att act ttt gac gac gac tac agt gcc tac	1963
	Asn Lys Thr Gly Cys Gln Ile Thr Phe Asp Asp Asp Tyr Ser Ala Tyr	
	565 570 575 580	
65	tgg att tac ttt gtc aac ttt ctc ggg aca ttg gca gtg tta cca gga	2011
	Trp Ile Tyr Phe Val Asn Phe Leu Gly Thr Leu Ala Val Leu Pro Gly	
	585 590 595	
70	aat atc gtg tct gct ctc ctg atg gac agg atc ggg cgc tta acg atg	2059
	Asn Ile Val Ser Ala Leu Leu Met Asp Arg Ile Gly Arg Leu Thr Met	
	600 605 610	

ES 2 311 667 T3

	cta ggt ggc tcc atg gtg ctc tgc ggg atc agc tgc ttc ttc ctg tgg	2107
	Leu Gly Gly Ser Met Val Leu Ser Gly Ile Ser Cys Phe Phe Leu Trp	
	615 620 625	
5	ttt ggc acc agc gaa tcc atg atg ata ggc atg ctg tgc ttg tac aac	2155
	Phe Gly Thr Ser Glu Ser Met Met Ile Gly Met Leu Cys Leu Tyr Asn	
	630 635 640	
10	gga ctg acc atc tca gcg tgg aac tct ctt gat gta gtc acg gtg gaa	2203
	Gly Leu Thr Ile Ser Ala Trp Asn Ser Leu Asp Val Val Thr Val Glu	
	645 650 655 660	
15	cta tac ccc aca gac cgg aga gca acg ggc ttt ggc ttc ttg aac gca	2251
	Leu Tyr Pro Thr Asp Arg Arg Ala Thr Gly Phe Gly Phe Leu Asn Ala	
	665 670 675	
20	ctc tgt aaa gca gcg gcc gtc ctg gga aac tta ata ttc ggc tcc ttg	2299
	Leu Cys Lys Ala Ala Ala Val Leu Gly Asn Leu Ile Phe Gly Ser Leu	
	680 685 690	
25	gtc agc atc acc aaa gca atc cct atc ctg ctg gct tcc acc gtg ctc	2347
	Val Ser Ile Thr Lys Ala Ile Pro Ile Leu Leu Ala Ser Thr Val Leu	
	695 700 705	
30	gtg tgt gga gga ctc gtg ggg ctg cgc ctg ccc gac aca aga acc cag	2395
	Val Cys Gly Gly Leu Val Gly Leu Arg Leu Pro Asp Thr Arg Thr Gln	
	710 715 720	
35	ggt ctg atg tga caaaagccat tctcttctct caccatgggt cagccctatt	2447
	Val Leu Met	
	725	
40	gcctgactca aggcttcaga gtttttatgt atagaaagggt ggccaagtat cagaactcaa	2507
	acttttgctg tgacgtaaat gtagctgtgt attgtccccg ccagtgtgat ttgcagggtc	2567
	ctccccctcc cccgcgcctt gttatctttt cctaattgtg atgttctctgc ttccg	2622

14

727

40 PRT

Rattus norvegicus

45	Met Glu Asp Ser Tyr Lys Asp Arg Thr Ser Leu Met Lys Gly Ala Lys
	1 5 10 15
50	Asp Ile Ala Lys Glu Val Lys Lys Gln Thr Val Lys Lys Val Asn Gln
	20 25 30
55	Ala Val Asp Arg Ala Gln Asp Glu Tyr Thr Gln Arg Ser Tyr Ser Arg
	35 40 45

60

65

ES 2 311 667 T3

	Phe	Gln	Asp	Glu	Asp	Asp	Asp	Asp	Asp	Tyr	Tyr	Pro	Pro	Gly	Glu	Thr	
	50						55					60					
5	Tyr	Ser	Gly	Glu	Ala	Asn	Asp	Asp	Glu	Gly	Ser	Ser	Glu	Ala	Thr	Glu	
	65					70					75					80	
	Gly	His	Asp	Glu	Glu	Asp	Glu	Ile	Tyr	Glu	Gly	Glu	Tyr	Gln	Gly	Ile	
10					85					90					95		
	Pro	Ser	Thr	Asn	Gln	Gly	Lys	Asp	Ser	Ile	Val	Ser	Val	Gly	Gln	Pro	
				100					105					110			
15	Lys	Gly	Asp	Glu	Tyr	Lys	Asp	Arg	Arg	Glu	Leu	Glu	Ser	Glu	Arg	Arg	
			115					120						125			
	Ala	Asp	Glu	Glu	Glu	Leu	Ala	Gln	Gln	Tyr	Glu	Leu	Ile	Ile	Gln	Glu	
20							135						140				
	Cys	Gly	His	Gly	Arg	Phe	Gln	Trp	Ala	Leu	Phe	Phe	Val	Leu	Gly	Met	
25	145					150					155					160	
	Ala	Leu	Met	Ala	Asp	Gly	Val	Glu	Val	Phe	Val	Val	Gly	Phe	Val	Leu	
					165					170					175		
30	Pro	Ser	Ala	Glu	Thr	Asp	Leu	Cys	Ile	Pro	Asn	Ser	Gly	Ser	Gly	Trp	
				180					185					190			
	Leu	Gly	Ser	Ile	Val	Tyr	Leu	Gly	Met	Met	Val	Gly	Ala	Phe	Phe	Trp	
35			195					200					205				
	Gly	Gly	Leu	Ala	Asp	Lys	Val	Gly	Arg	Lys	Gln	Ser	Leu	Leu	Ile	Cys	
40			210				215					220					
	Met	Ser	Val	Asn	Gly	Phe	Phe	Ala	Phe	Leu	Ser	Ser	Phe	Val	Gln	Gly	
	225					230					235					240	
45	Tyr	Gly	Phe	Phe	Leu	Leu	Cys	Arg	Leu	Leu	Ser	Gly	Phe	Gly	Ile	Gly	
					245					250					255		
	Gly	Ala	Ile	Pro	Thr	Val	Phe	Ser	Tyr	Phe	Ala	Glu	Val	Leu	Ala	Arg	
50				260					265					270			
	Glu	Lys	Arg	Gly	Glu	His	Leu	Ser	Trp	Leu	Cys	Met	Phe	Trp	Met	Ile	
55			275					280					285				
60																	
65																	

ES 2 311 667 T3

	Gly	Gly	Ile	Tyr	Ala	Ser	Ala	Met	Ala	Trp	Ala	Ile	Ile	Pro	His	Tyr	
	290						295					300					
5	Gly	Trp	Ser	Phe	Ser	Met	Gly	Ser	Ala	Tyr	Gln	Phe	His	Ser	Trp	Arg	
	305					310					315					320	
10	Val	Phe	Val	Ile	Val	Cys	Ala	Leu	Pro	Cys	Val	Ser	Ser	Val	Val	Ala	
					325					330					335		
15	Leu	Thr	Phe	Met	Pro	Glu	Ser	Pro	Arg	Phe	Leu	Leu	Glu	Val	Gly	Lys	
				340					345					350			
20	His	Asp	Glu	Ala	Trp	Met	Ile	Leu	Lys	Leu	Ile	His	Asp	Thr	Asn	Met	
		355						360					365				
25	Arg	Ala	Arg	Gly	Gln	Pro	Glu	Lys	Val	Phe	Thr	Val	Asn	Lys	Ile	Lys	
	370						375					380					
30	Thr	Pro	Lys	Gln	Ile	Asp	Glu	Leu	Ile	Glu	Ile	Glu	Ser	Asp	Thr	Gly	
	385					390				395						400	
35	Thr	Trp	Tyr	Arg	Arg	Cys	Phe	Val	Arg	Ile	Arg	Thr	Glu	Leu	Tyr	Gly	
				405					410						415		
40	Ile	Trp	Leu	Thr	Phe	Met	Arg	Cys	Phe	Asn	Tyr	Pro	Val	Arg	Glu	Asn	
			420						425					430			
45	Thr	Ile	Lys	Leu	Thr	Ile	Val	Trp	Phe	Thr	Leu	Ser	Phe	Gly	Tyr	Tyr	
		435						440					445				
50	Gly	Leu	Ser	Val	Trp	Phe	Pro	Asp	Val	Ile	Lys	His	Leu	Gln	Ser	Asp	
	450						455					460					
55	Glu	Tyr	Ala	Leu	Leu	Thr	Arg	Asn	Val	Gln	Lys	Asp	Lys	Tyr	Ala	Asn	
	465					470				475						480	
60	Phe	Ser	Ile	Asn	Phe	Thr	Met	Glu	Asn	Gln	Val	His	Thr	Gly	Met	Glu	
				485						490					495		
65	Tyr	Asp	Asn	Gly	Arg	Phe	Leu	Gly	Val	Lys	Phe	Lys	Ser	Val	Thr	Phe	
			500					505						510			
70	Lys	Asp	Ser	Val	Phe	Lys	Ser	Cys	Thr	Phe	Asp	Asp	Val	Thr	Ser	Val	
		515						520					525				

ES 2 311 667 T3

	Asn	Thr	Tyr	Phe	Lys	Asn	Cys	Thr	Phe	Ile	Asp	Thr	Leu	Phe	Glu	Asn	
	530						535					540					
5	Thr	Asp	Phe	Glu	Pro	Tyr	Lys	Phe	Ile	Asp	Ser	Glu	Phe	Gln	Asn	Cys	
	545					550					555					560	
10	Ser	Phe	Leu	His	Asn	Lys	Thr	Gly	Cys	Gln	Ile	Thr	Phe	Asp	Asp	Asp	
					565					570					575		
15	Tyr	Ser	Ala	Tyr	Trp	Ile	Tyr	Phe	Val	Asn	Phe	Leu	Gly	Thr	Leu	Ala	
				580					585					590			
20	Val	Leu	Pro	Gly	Asn	Ile	Val	Ser	Ala	Leu	Leu	Met	Asp	Arg	Ile	Gly	
			595					600					605				
25	Arg	Leu	Thr	Met	Leu	Gly	Gly	Ser	Met	Val	Leu	Ser	Gly	Ile	Ser	Cys	
	610					615						620					
30	Phe	Phe	Leu	Trp	Phe	Gly	Thr	Ser	Glu	Ser	Met	Met	Ile	Gly	Met	Leu	
	625					630					635					640	
35	Cys	Leu	Tyr	Asn	Gly	Leu	Thr	Ile	Ser	Ala	Trp	Asn	Ser	Leu	Asp	Val	
					645					650					655		
40	Val	Thr	Val	Glu	Leu	Tyr	Pro	Thr	Asp	Arg	Arg	Ala	Thr	Gly	Phe	Gly	
				660					665					670			
45	Phe	Leu	Asn	Ala	Leu	Cys	Lys	Ala	Ala	Ala	Val	Leu	Gly	Asn	Leu	Ile	
			675					680					685				
50	Phe	Gly	Ser	Leu	Val	Ser	Ile	Thr	Lys	Ala	Ile	Pro	Ile	Leu	Leu	Ala	
	690					695						700					
55	Ser	Thr	Val	Leu	Val	Cys	Gly	Gly	Leu	Val	Gly	Leu	Arg	Leu	Pro	Asp	
	705					710					715					720	
60	Thr	Arg	Thr	Gln	Val	Leu	Met										
					725												

ES 2 311 667 T3

15
1647
ADN

5 *Rattus norvegicus*

CDS
(1)..(1647)

10	atg gag gag gac ctg ttc cag ctc agg cag ttg ccg gtg gtg aaa ttc	48
	Met Glu Glu Asp Leu Phe Gln Leu Arg Gln Leu Pro Val Val Lys Phe	
	1 5 10 15	
15	cgc cgc aca gga gag agc gca cgg tca gag gac gac gcg gct tcc ggg	96
	Arg Arg Thr Gly Glu Ser Ala Arg Ser Glu Asp Asp Ala Ala Ser Gly	
	20 25 30	
20	gaa cat gat gtt cag att gag ggg gtc cga gtg ggc cta gaa gct gtc	144
	Glu His Asp Val Gln Ile Glu Gly Val Arg Val Gly Leu Glu Ala Val	
	35 40 45	
25	gag ctg gat gat gga gca gct gtc ccc aag gag ttt gcc aat ccc act	192
	Glu Leu Asp Asp Gly Ala Ala Val Pro Lys Glu Phe Ala Asn Pro Thr	
	50 55 60	
30	gat gac act ttc atg gtc gaa gat gcg gtg gaa gcc att ggg ttc gga	240
	Asp Asp Thr Phe Met Val Glu Asp Ala Val Glu Ala Ile Gly Phe Gly	
	65 70 75 80	
35	aga ttc cag tgg aag ctc tct gtt ctc acc ggc ttg gct tgg atg gcg	288
	Arg Phe Gln Trp Lys Leu Ser Val Leu Thr Gly Leu Ala Trp Met Ala	
	85 90 95	
40	gac gcc atg gag atg atg att ctg agc atc ctg gcg cct cag ctg cac	336
	Asp Ala Met Glu Met Met Ile Leu Ser Ile Leu Ala Pro Gln Leu His	
	100 105 110	
45	tgc gag tgg cga ctc ccc agc tgg cag gtg gcg ctg ctg act tcg gtg	384
	Cys Glu Trp Arg Leu Pro Ser Trp Gln Val Ala Leu Leu Thr Ser Val	
	115 120 125	
50	gtc ttc att ggt atg atg tcc agt tct acg ctc tgg gga aac atc tcg	432
	Val Phe Ile Gly Met Met Ser Ser Ser Thr Leu Trp Gly Asn Ile Ser	
	130 135 140	
55	gat cag tat ggc agg aaa aca ggg ctg aag atc agt gtg ttc tgg acc	480
	Asp Gln Tyr Gly Arg Lys Thr Gly Leu Lys Ile Ser Val Phe Trp Thr	
	145 150 155 160	
60	ctg tac tac ggc atc ctc agc gct ttc gcg cca gtg tat agc tgg atc	528
	Leu Tyr Tyr Gly Ile Leu Ser Ala Phe Ala Pro Val Tyr Ser Trp Ile	
	165 170 175	
65	ctg gtg ctc cga ggc ctc gtg ggc ttt ggg att gga ggg gtg cct cag	576
	Leu Val Leu Arg Gly Leu Val Gly Phe Gly Ile Gly Gly Val Pro Gln	
	180 185 190	
	tcg gtg acc ctg tac gcc gag ttc ctc ccc atg aag gcc aga gcc aag	624

ES 2 311 667 T3

	Ser	Val	Thr	Leu	Tyr	Ala	Glu	Phe	Leu	Pro	Met	Lys	Ala	Arg	Ala	Lys	
			195					200					205				
5	tgc	att	ttg	ctg	att	gag	gtt	ttc	tgg	gcc	atc	ggg	acc	gtg	ttc	gag	672
	Cys	Ile	Leu	Leu	Ile	Glu	Val	Phe	Trp	Ala	Ile	Gly	Thr	Val	Phe	Glu	
		210					215					220					
10	gtg	ctt	ctg	gct	gtg	ttt	gtg	atg	ccc	agc	ctg	ggc	tgg	cgc	tgg	ctg	720
	Val	Leu	Leu	Ala	Val	Phe	Val	Met	Pro	Ser	Leu	Gly	Trp	Arg	Trp	Leu	
		225				230					235					240	
15	ctg	ctg	ctg	tcg	gcc	gct	cca	cta	ctt	gtc	ttt	gct	gtt	ctg	tgt	ttc	768
	Leu	Leu	Leu	Ser	Ala	Ala	Pro	Leu	Leu	Val	Phe	Ala	Val	Leu	Cys	Phe	
					245					250					255		
20	tgg	ctg	cca	gag	agt	gct	agg	tac	gat	gtg	ctg	tct	ggg	aac	cag	gaa	816
	Trp	Leu	Pro	Glu	Ser	Ala	Arg	Tyr	Asp	Val	Leu	Ser	Gly	Asn	Gln	Glu	
				260				265						270			
25	aag	gcg	att	gct	acc	tta	aag	cgg	atc	gca	aca	gaa	aat	gga	gcc	ccc	864
	Lys	Ala	Ile	Ala	Thr	Leu	Lys	Arg	Ile	Ala	Thr	Glu	Asn	Gly	Ala	Pro	
			275					280					285				
30	atg	cct	ctg	ggg	aag	ctc	atc	atc	tcc	aga	cag	gaa	gac	cga	ggc	aaa	912
	Met	Pro	Leu	Gly	Lys	Leu	Ile	Ile	Ser	Arg	Gln	Glu	Asp	Arg	Gly	Lys	
		290					295					300					
35	atg	agg	gac	ctt	ttc	aca	ccc	cac	ttt	cgt	tgg	aca	act	ctg	ctg	ctg	960
	Met	Arg	Asp	Leu	Phe	Pro	His	Phe	Arg	Trp	Thr	Thr	Thr	Leu	Leu	Leu	
		305				310					315					320	
40	tgg	ttt	ata	tgg	ttc	tcc	aat	gcc	ttc	tct	tat	tac	ggc	ttg	gtt	ctg	1008
	Trp	Phe	Ile	Trp	Phe	Ser	Asn	Ala	Phe	Ser	Tyr	Tyr	Gly	Leu	Val	Leu	
					325					330					335		
45	ctc	acc	aca	gaa	ctc	ttc	cag	gcc	gga	gat	gtt	tgc	agc	atc	tcc	agc	1056
	Leu	Thr	Thr	Glu	Leu	Phe	Gln	Ala	Gly	Asp	Val	Cys	Ser	Ile	Ser	Ser	
				340					345					350			
50	cgg	aag	aag	gcg	gtg	gaa	gcc	aaa	tgc	agc	ctg	gct	tgt	gag	tac	ctc	1104
	Arg	Lys	Lys	Ala	Val	Glu	Ala	Lys	Cys	Ser	Leu	Ala	Cys	Glu	Tyr	Leu	
			355					360					365				
55	agc	aaa	gag	gat	tac	atg	gac	ctg	ctg	tgg	acc	acc	ctg	tct	gag	ttc	1152
	Ser	Lys	Glu	Asp	Tyr	Met	Asp	Leu	Leu	Trp	Thr	Thr	Leu	Ser	Glu	Phe	
		370					375						380				
60	cca	ggt	gtc	ctt	gtg	act	ctg	tgg	gtc	atc	gac	cgc	ctg	ggc	cgc	aag	1200
	Pro	Gly	Val	Leu	Val	Thr	Leu	Trp	Val	Ile	Asp	Arg	Leu	Gly	Arg	Lys	
		385				390					395					400	
65	aag	acc	atg	gct	ctg	tgt	ttc	gtc	atc	ttt	tcc	ctc	tgc	agc	ctc	ctg	1248
	Lys	Thr	Met	Ala	Leu	Cys	Phe	Val	Ile	Phe	Ser	Leu	Cys	Ser	Leu	Leu	
				405						410					415		
70	ctg	ttc	atc	tgc	att	gga	aga	aat	gtg	cta	acc	ctc	tta	ctg	ttc	att	1296
	Leu	Phe	Ile	Cys	Ile	Gly	Arg	Asn	Val	Leu	Thr	Leu	Leu	Leu	Phe	Ile	
				420					425					430			
75	gca	aga	gcg	ttt	att	tct	gga	ggc	ttc	caa	gca	gcc	tac	gtt	tac	acg	1344
	Ala	Arg	Ala	Phe	Ile	Ser	Gly	Gly	Phe	Gln	Ala	Ala	Tyr	Val	Tyr	Thr	

ES 2 311 667 T3

	435	440	445	
5	cct gag gtg tat cca acg gcg acg agg gcg ctg ggc ctg ggc acc tgc Pro Glu Val Tyr Pro Thr Ala Thr Arg Ala Leu Gly Leu Gly Thr Cys 450 455 460			1392
10	agc ggc atg gcg aga gtg ggc gcg ctc atc act cca ttc ata gct cag Ser Gly Met Ala Arg Val Gly Ala Leu Ile Thr Pro Phe Ile Ala Gln 465 470 475 480			1440
15	gtg atg ctg gaa tct tcc gtg tac ctg acc ctg gcc gtc tac agt ggc Val Met Leu Glu Ser Ser Val Tyr Leu Thr Leu Ala Val Tyr Ser Gly 485 490 495			1488
20	tgc tgc ctc ctt gct gcc ttg gcc tcc tgc ttt ctg ccc atc gag acc Cys Cys Leu Leu Ala Ala Leu Ala Ser Cys Phe Leu Pro Ile Glu Thr 500 505 510			1536
25	aaa ggc cga gca ctg cag gag tcc agc cac cgg gag tgg ggc cag gag Lys Gly Arg Ala Leu Gln Glu Ser Ser His Arg Glu Trp Gly Gln Glu 515 520 525			1584
30	atg gtt ggc cga ggg aca aac agc aca ggc gtc ccc agg tcg aac tct Met Val Gly Arg Gly Thr Asn Ser Thr Gly Val Pro Arg Ser Asn Ser 530 535 540			1632
35	ggc tct cag gag tag Gly Ser Gln Glu 545			1647

16
548
PRT
Rattus norvegicus

40	Met Glu Glu Asp Leu Phe Gln Leu Arg Gln Leu Pro Val Val Lys Phe 1 5 10 15
45	Arg Arg Thr Gly Glu Ser Ala Arg Ser Glu Asp Asp Ala Ala Ser Gly 20 25 30
50	Glu His Asp Val Gln Ile Glu Gly Val Arg Val Gly Leu Glu Ala Val 35 40 45
55	Glu Leu Asp Asp Gly Ala Ala Val Pro Lys Glu Phe Ala Asn Pro Thr 50 55 60
60	Asp Asp Thr Phe Met Val Glu Asp Ala Val Glu Ala Ile Gly Phe Gly 65 70 75 80

ES 2 311 667 T3

	Arg	Phe	Gln	Trp	Lys	Leu	Ser	Val	Leu	Thr	Gly	Leu	Ala	Trp	Met	Ala	
					85					90					95		
5	Asp	Ala	Met	Glu	Met	Met	Ile	Leu	Ser	Ile	Leu	Ala	Pro	Gln	Leu	His	
				100					105					110			
10	Cys	Glu	Trp	Arg	Leu	Pro	Ser	Trp	Gln	Val	Ala	Leu	Leu	Thr	Ser	Val	
			115					120					125				
15	Val	Phe	Ile	Gly	Met	Met	Ser	Ser	Ser	Thr	Leu	Trp	Gly	Asn	Ile	Ser	
		130					135					140					
20	Asp	Gln	Tyr	Gly	Arg	Lys	Thr	Gly	Leu	Lys	Ile	Ser	Val	Phe	Trp	Thr	
	145					150					155					160	
25	Leu	Tyr	Tyr	Gly	Ile	Leu	Ser	Ala	Phe	Ala	Pro	Val	Tyr	Ser	Trp	Ile	
				165						170					175		
30	Leu	Val	Leu	Arg	Gly	Leu	Val	Gly	Phe	Gly	Ile	Gly	Gly	Val	Pro	Gln	
				180					185					190			
35	Ser	Val	Thr	Leu	Tyr	Ala	Glu	Phe	Leu	Pro	Met	Lys	Ala	Arg	Ala	Lys	
			195					200					205				
40	Cys	Ile	Leu	Leu	Ile	Glu	Val	Phe	Trp	Ala	Ile	Gly	Thr	Val	Phe	Glu	
		210					215					220					
45	Val	Leu	Leu	Ala	Val	Phe	Val	Met	Pro	Ser	Leu	Gly	Trp	Arg	Trp	Leu	
	225				230						235					240	
50	Leu	Leu	Leu	Ser	Ala	Ala	Pro	Leu	Leu	Val	Phe	Ala	Val	Leu	Cys	Phe	
				245						250					255		
55	Trp	Leu	Pro	Glu	Ser	Ala	Arg	Tyr	Asp	Val	Leu	Ser	Gly	Asn	Gln	Glu	
				260					265					270			
60	Lys	Ala	Ile	Ala	Thr	Leu	Lys	Arg	Ile	Ala	Thr	Glu	Asn	Gly	Ala	Pro	
		275						280					285				
65	Met	Pro	Leu	Gly	Lys	Leu	Ile	Ile	Ser	Arg	Gln	Glu	Asp	Arg	Gly	Lys	
		290					295					300					
70	Met	Arg	Asp	Leu	Phe	Thr	Pro	His	Phe	Arg	Trp	Thr	Thr	Leu	Leu	Leu	
	305				310						315					320	

ES 2 311 667 T3

	Trp	Phe	Ile	Trp	Phe	Ser	Asn	Ala	Phe	Ser	Tyr	Tyr	Gly	Leu	Val	Leu
					325					330					335	
5	Leu	Thr	Thr	Glu	Leu	Phe	Gln	Ala	Gly	Asp	Val	Cys	Ser	Ile	Ser	Ser
				340					345					350		
10	Arg	Lys	Lys	Ala	Val	Glu	Ala	Lys	Cys	Ser	Leu	Ala	Cys	Glu	Tyr	Leu
			355					360					365			
15	Ser	Lys	Glu	Asp	Tyr	Met	Asp	Leu	Leu	Trp	Thr	Thr	Leu	Ser	Glu	Phe
		370					375					380				
20	Pro	Gly	Val	Leu	Val	Thr	Leu	Trp	Val	Ile	Asp	Arg	Leu	Gly	Arg	Lys
						390					395					400
25	Lys	Thr	Met	Ala	Leu	Cys	Phe	Val	Ile	Phe	Ser	Leu	Cys	Ser	Leu	Leu
					405					410					415	
30	Leu	Phe	Ile	Cys	Ile	Gly	Arg	Asn	Val	Leu	Thr	Leu	Leu	Leu	Phe	Ile
				420					425					430		
35	Ala	Arg	Ala	Phe	Ile	Ser	Gly	Gly	Phe	Gln	Ala	Ala	Tyr	Val	Tyr	Thr
			435					440					445			
40	Pro	Glu	Val	Tyr	Pro	Thr	Ala	Thr	Arg	Ala	Leu	Gly	Leu	Gly	Thr	Cys
		450					455					460				
45	Ser	Gly	Met	Ala	Arg	Val	Gly	Ala	Leu	Ile	Thr	Pro	Phe	Ile	Ala	Gln
	465					470					475					480
50	Val	Met	Leu	Glu	Ser	Ser	Val	Tyr	Leu	Thr	Leu	Ala	Val	Tyr	Ser	Gly
					485				490						495	
55	Cys	Cys	Leu	Leu	Ala	Ala	Leu	Ala	Ser	Cys	Phe	Leu	Pro	Ile	Glu	Thr
				500					505					510		
60	Lys	Gly	Arg	Ala	Leu	Gln	Glu	Ser	Ser	His	Arg	Glu	Trp	Gly	Gln	Glu
		515						520					525			
65	Met	Val	Gly	Arg	Gly	Thr	Asn	Ser	Thr	Gly	Val	Pro	Arg	Ser	Asn	Ser
		530					535					540				
	Gly	Ser	Gln	Glu												
	545															