

AMINOACIDOS HIDROFOBICOS EN HEPTADAS REPETIDAS EN LA GLICOPROTEINA DE RABDOVIRUS

G. LORENZO

A. ESTEPA

J. M. COLL

Centro de Investigación en Sanidad Animal (CISA)-INIA
Valdeolmo. 28130 Madrid

RESUMEN

En las secuencias de aminoácidos (aa) de la glicoproteína G de todos los rabdovirus examinados (rabia, estomatitis vesicular, rabdovirus de peces y plantas) se han localizado dos o tres regiones que contienen varias héptadas consecutivas de repeticiones de aa hidrofóbicos en posiciones a-d nunca descritas antes. Estas nuevas héptadas repetidas permiten la presencia de cualquier aa hidrofóbico en las posiciones a-d y no sólo de algunos como sucedió en las héptadas encontradas anteriormente en otros virus y que además formaban hélices superenrolladas. Las dos-tres regiones de nuevas héptadas repetidas poseen secuencias muy conservadas entre los miembros de cada uno de los géneros rabdovirales, pero no mostraban ninguna semejanza de secuencias entre los distintos géneros. Todas estas nuevas héptadas repetidas se detectaron bien dentro o en la proximidad de regiones altamente hidrofóbicas en cada uno de los géneros de rabdovirus. La conservación de posición y estructura de estas nuevas héptadas repetidas entre todos los rabdovirus sugiere su participación en funciones comunes, probablemente relacionadas con la fusión viral con las membranas celulares.

PALABRAS CLAVE: Héptadas hidrofóbicas
Rabdovirus
Glicoproteína

INTRODUCCION

La familia Rabdoviridae está mayoritariamente formada por el género *Lisavirus* (rabia) y el género *Vesiculovirus* (VSV) con dos serotipos principales, New-Jersey (NJ) e Indiana (Ind) (Bourhy *et al.*, 1993; Nichol *et al.*, 1989), pero también comprende los rabdovirus de peces, que incluyen el virus de la septicemia hemorrágica viral VSHV, el virus de la necrosis hematopoyética infecciosa (VNHI) y los rabdovirus de plantas. Un caso interesante y quizás único entre los virus para estudiar virología comparativa.

Los rabdovirus poseen una glicoproteína (G) como una proteína de membrana homotrimérica que forma espículas de 83 Å (Gaudin *et al.*, 1992, 1993). La G

Recibido: 3-2-95

Aceptado para su publicación: 1-6-95

inicia la unión del virus a los receptores celulares (Bunner *et al.*, 1984), reacciona con anticuerpos neutralizantes (Benmansour *et al.*, 1991) y lleva las propiedades de fusión a bajo pH (Gaudin *et al.*, 1992, 1993; Bourhy *et al.*, 1993; Lecocq-Xhonneux *et al.*, 1994). Las héptadas de aa repetidas con probabilidad de formar hélices superenrolladas se definieron como secuencias consecutivas de siete aminoácidos (aa), **a b c d e f g**, en las que el aa en cada posición se definió por una matriz de probabilidades, siendo los aa en las posiciones a-d mayoritariamente, F, Y, I, L, V, M y A (Lupas *et al.*, 1991). Su presencia ha sido también descrita en las secuencias de glicoproteínas de varios virus con envuelta membranosa (Chambers *et al.*, 1990), y relacionados con regiones hidrofóbicas, o con la fusión viral (Bullough *et al.*, 1994; Carr, Kim, 1993). Sin embargo, héptadas repetidas similares no se han encontrado antes en ningún rabdovirus (Zhang, Ghosh, 1994). En el presente trabajo buscamos posibles héptadas repetidas en las secuencias de la G de especies de rabdovirus después de encontrar que una región del VHSV que liga fosfolípidos tenía una estructura repetida de héptadas que incluía tres aa hidrofóbicos más que aquellos definidos por Lupas (Coll, 1995).

MATERIAL Y METODOS

Selección de las secuencias de la G de rabdovirus

La selección de las secuencias de G para este estudio se realizó con el objeto de encontrar secuencias representativas del mayor número de miembros como fuera posible de la familia Rabdoviridae. La selección estuvo limitada primero por el número de secuencias de aa derivados del ADNc de la G rabdoviral publicados actualmente, y segundo, por un número lo bastante pequeño como para hacer el estudio viable. Las secuencias de la G actualmente disponibles de rabdovirus se dividieron de acuerdo con los huéspedes que infectan peces (tres secuencias), artrópodos (dos secuencias), plantas (una secuencia), lisavirus (siete secuencias) y vesiculovirus (tres secuencias seleccionadas de treinta y cuatro secuencias de VSV-NJ, New Jersey, y tres de veintiséis de VSV-Ind, Indiana) (Tabla 1).

Análisis de la G de rabdovirus

El perfil de hidrofiliidad que se utilizó para predecir regiones de péptido señal y regiones transmembrana de las secuencias de aa derivadas del ADNc de la G de rabdovirus se obtuvo empleando los programas SOAP, ANTIGEN y PSIGNAL (PCGene, Intelligenetics Inc, Ginebra, Suiza) utilizando una ventana de 9 aa de longitud. Las secuencias hidrofóbicas de las héptadas repetidas recién definidas se buscaron con el programa PSEARCH. Los aa con unos valores de $-\Delta G \geq 0,4$ Kcal/mol para transferir la cadena de los aa del agua al etanol se definieron como aa hidrofóbicos (en el código de aa con una letra, W, F, Y, I, L, V, M, A, H, T) (Schulz, Schimer, 1984). La subsecuencia buscada fue: (aa hidrofóbico) XX (aa hidrofóbico) XXX, repetido dos veces (X para cualquier aa). De este modo, se seleccionaron las posiciones donde se encontraron más de dos

TABLA 1
SECUENCIAS RABDOVIRALES DE LA GLICOPROTEINA G
SELECCIONADAS PARA ESTE ESTUDIO

Rhabdoviral sequences of the glycoprotein G selected for this study

Huésped	Virus	Descripción	#aa	Referencia
Peces	VHSGER	VHSV 07.71	507	Thiry <i>et al.</i> , 1991b
	VHSDK	VHSV DK	507	Lorenzen <i>et al.</i> , 1993
	IHNVP	IHNVP Cedar	507	Koener <i>et al.</i> , 1987
Mamíferos	RABMOK	rab Mokila	524	Tordo <i>et al.</i> , 1993
	RABPV	rab PV	524	Tordo <i>et al.</i> , 1986
	RHRBGD	rab ERA	524	Anilionis <i>et al.</i> , 1981
	RABSAD	rab SADB19	524	Conzelmann <i>et al.</i> , 1990
	RABMEP	rab MEP	524	Morimoto <i>et al.</i> , 1989
	RABLEP	rab LEP	524	Tordo, 1991
	TABCVS	rab CVS	524	Prehaud <i>et al.</i> , 1989
Vacuno	VSVGPN08	VSV N.J.-I	517	Nichol <i>et al.</i> , 1989
Caballo	VSGPNJA	VSV N.J.-II	517	Nichol <i>et al.</i> , 1989
Cerdo	VSVGPN29	VSV N.J.-III	517	Nichol <i>et al.</i> , 1989
	RHGPORS	VSV Orsay	511	Gallione, Rose, 1985
	RHGM	VSV San Juan	511	Rose, Gallione, 1981
	RHVSVGR	VSV Indiana	511	Vandepol <i>et al.</i> , 1986
Insecto	SIGMA	Drosophila	526	Teninges <i>et al.</i> , 1993
	BEFV	Bov. eph. fev.	623	Walker <i>et al.</i> , 1992
Plantas	SYNV	ATCC-PV263	628	Goldberg <i>et al.</i> , 1991

VSHV, septicemia hemorrágica viral, *viral haemorrhagic septicaemia*. NHI, necrosis hematopoyética infecciosa, *infectious haematopoietic necrosis*. RAB, rabies, *rabies*. VSV, virus de la estomatitis vesicular, *vesicular stomatitis virus*. BEFV, virus de la fiebre efímera bovina, *bovine ephemeral fever virus*. SYNV, sonchus yellow net virus, *sonchus yellow net virus*

héptadas repetidas con la subsecuencia definida en la secuencia de la glicoproteína G (Tabla 2). Las héptadas repetidas con probabilidad para formar hélices superenrolladas se examinaron con el programa COIL (Lupas *et al.*, 1991).

RESULTADOS

El número total de aa de la glicoproteína G varía de 507 a 526 para la mayoría de los rhabdovirus, excepto para el BEFV y SYNV, que tienen 623 y 628 aa, respectivamente (Tabla 1). Al menos dos regiones con aa hidrofóbicos en las posiciones a-d de héptadas repetidas tres o más veces se encontraron en todas las secuencias de la glicoproteína G seleccionadas por búsqueda en el ordenador. No se prestó atención a la composición del resto de los aa de las posiciones (b, c, e, f, g) dentro de las héptadas, al contrario que el método de predicción del programa

TABLA 2
HEPTADAS REPETIDAS DE aa HIDROFOBICOS DE LA
GLICOPROTEINA G DE RABDOVIRUS

Heptads of hydrophobic aa repeats of the glycoprotein G of rhabdoviruses

Repeticiones hidrofóbicas amino-terminal									
VHSGER	68	IHKV	FEDINKG	LVSVPTR	IIHLPLS	VTSVSAV	ASGHYLH	102	
VHSDK	68		FEDINKG	LVSVPTR	IIHLPLS	VTSVSAV	ASGHYLH	102	
IHNPG	99	IHKV	LYRTICS	TGFFGGQ	TIE			119	
SIGMA	151		VKDHPVM	LDPYTNN	YVDAIFP			171	
BEFV	149		LIQHKPF	LNPDYNI	IYD			165	
SYNVG	90		ITGARWN	YVGISIP	VFKIVTN			110	
RABMOK	140	WLRT	VTTTKES	LLIISPS	IVEMDIY			164	
RABPV	140	WLRT	VTTTKES	LVIISPS	VADLDPY			164	
RHRBGD	140	WLRT	VTTTKES	LVIISPS	VADLDPY			164	
RABSAD	140	WLRT	VTTTKES	LVIISPS	VADLDPY			164	
RABHEP	140	WLRT	VTTTKES	LVIISPS	VADLDPY			164	
RABLEP	140	WLRT	VTTTKES	LVIISPS	VTDLDPY			164	
RABCVS	140	WLRT	VTTTKES	LVIISPS	VTDLDPY			164	
VSVGPN08	134	TVTD	AEAHIT	VTPHSVK	VDEYTGE	WID		161	
VSVGPNJA	134	TVTD	AEAHIVT	VTPHSVK	VDEYTGE	WID		161	
VSVGPN29	138		AEAHIT	VTPHSVK	VDEYTGE	WID		161	
RHGORS	134	TVTD	AEAHIVQ	VTPHHVL	VDEYTGE	WVD		161	
RHGM	134	TVTD	AEAHIVQ	VTPHHVL	VDEYTGE	WVD		161	
RHVSVGR	134	TVTD	AEAHIVQ	VTPHHVL	VDEYTGE	WVD		161	
Repeticiones hidrofóbicas carboxi-terminal									
VHSGER	288	TDIQ	MRGATDD	FSYLNHL	ITNMAQR	TECLDAH		319	
VHSDK	288	ADVQ	MRGATDD	FSYLNHL	ITNMAQR	TECLDAH		319	
IHNPG	327	TPYL	LSKFRSP	HPGINDV	YAMHKGS	IYH		354	
SIGMA	316	ISKM	VSGLPTS	VFDLSYL	IQV			336	
BEFV	385		IGSYKRA	WCEYRPF	VDK			401	
SYNVG	320		IEGVNRA	FEDLELT	YCSATCD	LFA		343	
RABMOK	345		TNYYYKR	VDKWADI	LPS			361	
RABPV	330		FGKAYTI	FNKTLME	ADAHYKS	VRTWNEI	IPS	360	
RHRBGD	330		FGKAYTI	FNKTLME	ADAHYKS	VRTWNEI	IPS	360	
RABSAD	330		FGKAYTI	FNKTLME	ADAHYKS	VRTWNEI	IPS	360	
RABHEP	330		FGKAYTI	FNKTLME	ADAHYKS	VRTWNEI	IPS	360	
RABLEP	330		FGKAYTI	FNKTLME	ADAHYKS	VRTWNEI	IPS	360	
RABCVS	330		FGKAYTI	FNKTLME	ADAHYKS	VRTWNEI	IPS	360	
VSVGPN08	332	VGPV	FTINGS	LHYFTSK	YLRVELE			356	
VSVGPNJA	332	VGPV	FTINGS	LHYFTSK	YLRVELE			356	
VSVGPN29	332	VGPV	FTINGS	LHYFTSK	YLRVELE			356	
RHGORS	328	TGPV	FTINGT	LKYFETR	YIRVDIA	APILSRM	VGMISGT	TTE	369
RHGM	328	TGPA	FTINGT	LKYFETR	YIRVDIA	APILSRM	VGMISGT	TTE	369
RHVSVGR	328	TGPV	FTINGT	LKYFETR	YIRVDIA	APILSRM	VGMISGT	TTE	369

Las posiciones a y d de los aa hidrofóbicos (W, F, Y, I, L, V, M, A, H, T) en la héptada están en negrita. Cada héptada está separada por un espacio. Los números a la izquierda y a la derecha de las secuencias son las posiciones amino terminal y carboxi terminal del primer y último aa de las héptadas.

COIL (Lupas *et al.*, 1991), que asocia una probabilidad a cada aa en cada posición de la héptada (siendo los aa en a-d, por ejemplo, principalmente = F, Y, I, L, V, M, A, ya que esos aa eran los que tenían una probabilidad más alta de estar en estructuras demostradas de hélices superenrolladas).

La existencia de aa a-d hidrofóbicos en las héptadas repetidas de la especie definida en este trabajo (aa en las posiciones a-d de las héptadas = F, Y, I, L, V, M, A + W, H, T, aa en el código de una letra), por lo tanto, no es por sí mismo un indicador de una estructura de hélice superenrollada. En todo caso lo sería de una hélice anfipática. La inclusión de los aa W, H y T como nuevos candidatos para las posiciones a ó d fue sugerida por la aparición de tales aa en las posiciones a ó d de las héptadas del dominio de unión a fosfolípidos en VSHV (Coll, 1995). Análisis anteriores presentados de la secuencia de la glicoproteína G de VSV por el método de Lupas, buscando la presencia de cuatro héptadas repetidas consecutivas (los péptidos más cortos que aún manifiestan una estructura de hélices superenrolladas estable en solución) fracasaron al tratar de identificar los dominios predichos con hélices superenrolladas (Cambers *et al.*, 1990; Lupas *et al.*, 1991; Zhang, Ghosh, 1994). Incluso el acortamiento de la longitud de la ventana de aa a tres héptadas repetidas en el programa COIL no era capaz de reconocer regiones con hélices superenrolladas con una probabilidad suficientemente alta en ninguna de las secuencias rabdovirales de la Tabla 1 (no mostrado).

Las dos regiones de al menos tres héptadas consecutivas en la mayoría de las secuencias rabdovirales estudiadas se situaron alrededor de los aa 100-150 (amino-terminal) y alrededor del aa 350 (carboxi-terminal). Por el contrario, la localización de las héptadas repetidas en las proteínas de otros 20 tipos de virus con envuelta fue en una única región situada alrededor del aa 150 (cerca del aa amino-terminal) o alrededor de los aa 400-500 (cerca del aa carboxi-terminal), pero nunca en ambas posiciones (Chambers *et al.*, 1990). Los rabdovirus SVHV y Sigma mostraron aún otra región carboxi-terminal de héptadas repetidas situada entre los aa 377 al 400 y 416 al 443, respectivamente (Fig. 1). Las héptadas repetidas carboxi-terminales de rabia y vesiculovirus y la segunda héptada repetida carboxi-terminal de los rabdovirus VSHV y Sigma contenían una de las posibles posiciones de glicosilación. Todas las héptadas repetidas estaban fuera y alrededor de un núcleo interno donde siete cisteínas (alrededor del aa 170 al 300) están altamente conservadas (alineamiento máximo con minoría de huecos, no mostrado) (Fig. 1).

Las secuencias de las héptadas repetidas no muestran homología de secuencias entre los diferentes géneros, pero estaban altamente conservadas en la secuencia de aa y su posición relativa en la G entre los miembros del mismo género, incluyendo VSV-NJ e Ind (Tabla 2). Para estudiar la variabilidad de los aa de las héptadas repetidas dentro de cada género se escogieron los casos donde había un gran número de secuencias de aa publicadas, rabia (siete secuencias), VSV-NJ (34 secuencias) y VSV-Ind (26 secuencias). La variación máxima en el número de aa diferentes que aparecían en una posición dada en VSV-NJ (Nichol *et al.*, 1989), VSV-Ind (Bilsel, Nichol, 1990) y rabia (Tordo *et al.*, 1993), se obtuvieron en las posiciones carboxi-terminales (del aa 400 al 500), incluyendo las regiones de los tramos hidrofóbicos transmembrana predichos, la región citoplasmática y el péptido señal amino-terminal (Fig. 2). No había variaciones de

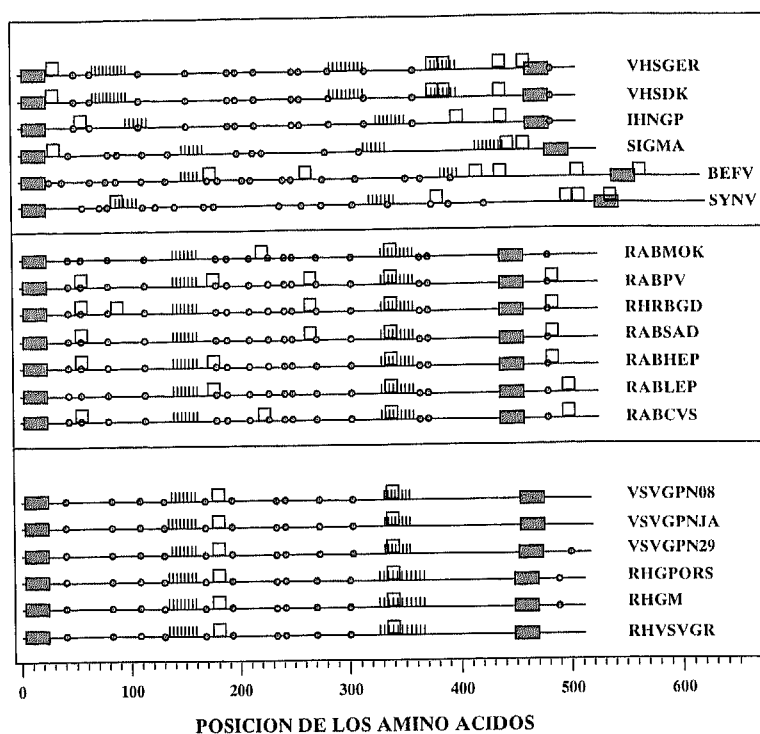


Fig. 1.—Esquema de las héptadas repetidas a-d hidrofóbicas en la glicoproteína G de rabdovirus
Sequences of amino-terminal and carboxy-terminal a-d heptad-repeats in the glycoprotein G of rabdoviruses

Las héptadas repetidas (||||) se localizaron mediante el programa PSEARCH (PCGene package, Intelligenetics). Valores de $-\Delta G > 0.4$ Kcalorías/mol para transferir la cadena lateral de los aa del agua al etanol se emplearon para definir los aa hidrofóbicos (en código de una letra los aa, W, F, Y, I, L, V, M, A, H, T) (Schulz, Schimer, 1984). La subsecuencia utilizada para buscar las héptadas era: (aa hidrofóbico) XX (aa hidrofóbico) XXX, repetido dos veces (X para cualquier aa). De este modo, se seleccionaron las posiciones donde se encontraron más de dos héptadas repetidas en la secuencia de la glicoproteína G. Las secuencias que contienen aa altamente hidrofóbicos y continuos de las regiones del péptido señal y el dominio transmembranal no se consideraron. Las posiciones en la glicoproteína G rabdoviral de las cisteínas (●), carbohidratos (□) y péptido señal y transmembrana (■) se han señalado en la figura. Los nombres de los rabdovirus corresponden a la Tabla 1

The heptad-repeats (||||) were located by using the program PSEARCH (PCGene package, Intelligenetics). $-\Delta G$ values > 0.4 Kcal/mol for transfer of the side chain from water to ethanol were used to define hydrophobic aa (single letter code aa, W, F, Y, I, L, V, M, A, H, T) (Schulz, Schimer, 1984). The subsequence used to search was (Hydrophobic aa) XX (hydrophobic aa) XXX, repeated two times (X for any aa). Then the positions where more than two heptad-repeats were found in the sequence of the whole glycoprotein G were selected. The highly and continuous hydrophobic aa sequences of the predicted transmembrane and signal peptide regions were not considered. Cysteine (●), putative carbohydrate (□) and predicted transmembrane and signal peptide (■), relative positions in the rabdoviral glycoprotein G. The names of the rabdoviruses correspond to Table 1

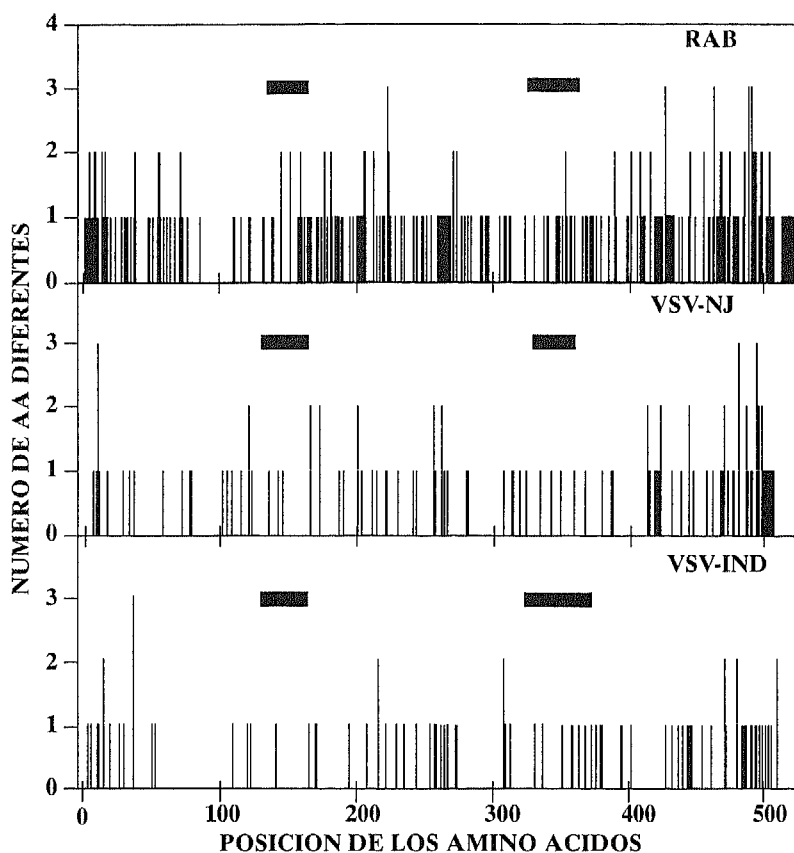


Fig. 2.—Rango de variabilidad de las secuencias de la glicoproteína G

Variability plot of glycoprotein G sequences

La variabilidad se definió como el número de aa diferentes que ocurren en cada posición de la glicoproteína G. Las barras horizontales negras indican las posiciones de las héptadas repetidas. Las secuencias de la glicoproteína G empleadas fueron siete de rabia (Tordo et al., 1993), treinta y cuatro de VSV-NJ (Nichol et al., 1989) y 26 de VSV-Ind (Bilsel, Nichol, 1990).

The variability was defined as the number of aa different from one occurring at any given position. Glycoprotein G sequences used were, 7 from rabies (Tordo et al., 1993), 34 from VSV-NJ (Nichol et al., 1989) and 26 from VSV-Ind (Bilsel, Nichol, 1990).

los aa en las posiciones a-d en la héptada repetida 68-102 y solamente uno (aa 288, T ó A) en la héptada repetida 288-319 de las dos secuencias disponibles de VSHV. Del mismo modo, no había variaciones de aa en las posiciones a-d de las héptadas repetidas en la posición 330-360 para rabia, excepto aquellas para RABMOK (aa 158, 161, 345, 348 y 359, todos también hidrofóbicos) (Tabla 2). Se encontraron pocas variaciones de aa (aa 134, T ó S; aa 141, H ó Q y aa 332, V

ó A) en las héptadas repetidas 134-161 o la 332-356 de los treinta y cuatro aislados de VSV-NJ estudiados (Nichol *et al.*, 1989). Finalmente, sólo se encontró una variación en un aa (aa 141, V ó A ó aa 335, I ó V) en cada una de las héptadas repetidas 134-161 o la 328-369 de los veintiséis aislados de VSV-Ind estudiados (Bilsel, Nichol, 1990).

Para investigar las posibles relaciones entre las héptadas repetidas y las posiciones de los aa extra en BEFV y SYNIV, se realizó un alineamiento múltiple de secuencias con RABMOK, RABPV, VSHV-07.71 y VSVGPN08 utilizando el programa CLUSTAL (PCGene package, Intelligenetics, S.A.). Como muestra la Figura 3, el exceso relativo de aa de BEFV y SYNIV se distribuyó en cortos tramos de 10 a 20 aa colocados alrededor de las posiciones pertenecientes a la parte carboxi-terminal de la molécula (alrededor del aa 400) y al péptido señal, pero generalmente no coincidían con las posiciones de sus héptadas repetidas (Fig. 1), excepto para las héptadas repetidas carboxi-terminales de BEFV. Incluidos como controles, VSHV 07.71 y VSVGPN08 no mostraron grandes inserciones de aa (> 10 aa) en las posiciones alrededor del aa 400.

En la mayoría de los casos, las héptadas repetidas de rabdovirus en la G están seguidas, precedidas o dentro de cortas regiones (10-15 aa) hidrofóbicas (Fig. 4), tal y como se ha descrito en otros virus con envuelta (Chambers *et al.*, 1990).

DISCUSION

Las funciones de las héptadas repetidas definidas en la G de rabdovirus no son conocidas actualmente. El procedimiento de búsqueda utilizado reconoce cualquiera de los 10 aa hidrofóbicos (Schulz, Schimer, 1984) en cada posición a ó d de la héptada especificada, y puede detectar series de aa no relacionados con estructuras de hélices superenrolladas, particularmente si éstos incluyen α -hélices. La presencia de numerosos residuos de prolina en las hipotéticas α -hélices dispersas por las nuevas héptadas repetidas es el principal factor que existe contra la existencia de todas estas estructuras como α -hélices, si bien en algunos casos se pueden predecir por ordenador dentro de las héptadas (Tablada 1). Aun así podría tratarse de α -hélices deformadas que tendrían carácter anfipático debido al agrupamiento de los aa hidrofóbicos a un solo lado de la α -hélices.

Varias líneas de evidencias indirectas sugieren que las nuevas héptadas repetidas situadas en la región amino-terminal de la G podrían estar de alguna manera relacionadas con fusión viral, neutralización y/o unión a fosfolípidos. Por mutagénesis dirigida se identificó la secuencia del aa 123 al 137 (aa 134-161 contiene las héptadas hidrofóbicas repetidas) de la G de VSV como un péptido fusogénico dependiente de bajo pH (Li *et al.*, 1993; Zhang, Ghosh, 1994). Del resto de los mutantes defectivos para fusión de VSV identificados en las posiciones de los aa 120-150, 190-210, 300-360 y 409-419, solamente los 300-360 también mapean cerca de otra región que va del aa 332-356 rica en nuevas héptadas repetidas en VSV-NJ o aa 328-369 en VSV-Ind.

Por otra parte, algunos mutantes de rabia resistentes a anticuerpos monoclonales (MAR) mapean en los aa 330-338 y aa 342-343, ambos dentro de las nuevas héptadas repetidas de rabia aa 330-360 (Benmansour *et al.*, 1991; Lafon *et al.*, 1983; Seif *et al.*, 1985).

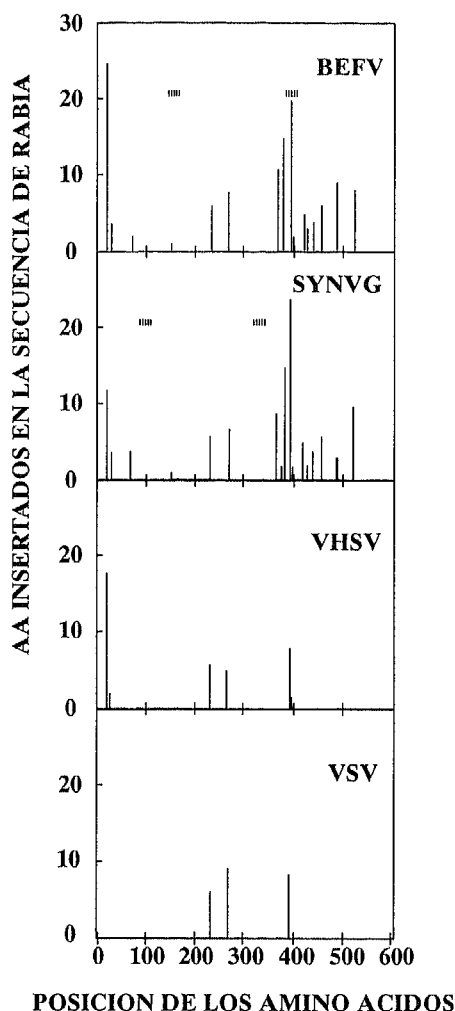


Fig. 3.—Número de aa insertados en la secuencia de la glicoproteína G de rabia para alineamiento máximo de secuencias con BEFV y SYN

Number of aa inserted in the sequence of glycoprotein G from rabies for maximal sequences alignment

Las secuencias mostradas en la figura se alinearon con las secuencias de la glicoproteína G de RABMOK y RABPV, mediante el programa CLUSTAL de PCGene (Intelligenetics). (| | | |), posición de las héptadas repetidas. BEFV, fiebre efímera bovina; SYN, sonchus yellow net virus; VHSV, virus 07.71 de la septicemia hemorrágica viral; VSV, virus VSVGPN08 de la estomatitis vesicular. La posición de los aa corresponden a RABMOK y RABPV

The sequences shown in the figure were aligned with the glycoprotein G sequences from RABMOK and RABPV, by the CLUSTAL program from PCGene (Intelligenetics). (| | | |), position of heptad-repeats. BEFV, bovine ephemeral fever virus; SYN, sonchus yellow net virus; VHSV, viral haemorrhagic septicaemia virus 07.71; VSV, vesicular stomatitis virus VSVGPN08. The position number of aa correspond to RABMOK and RABPV

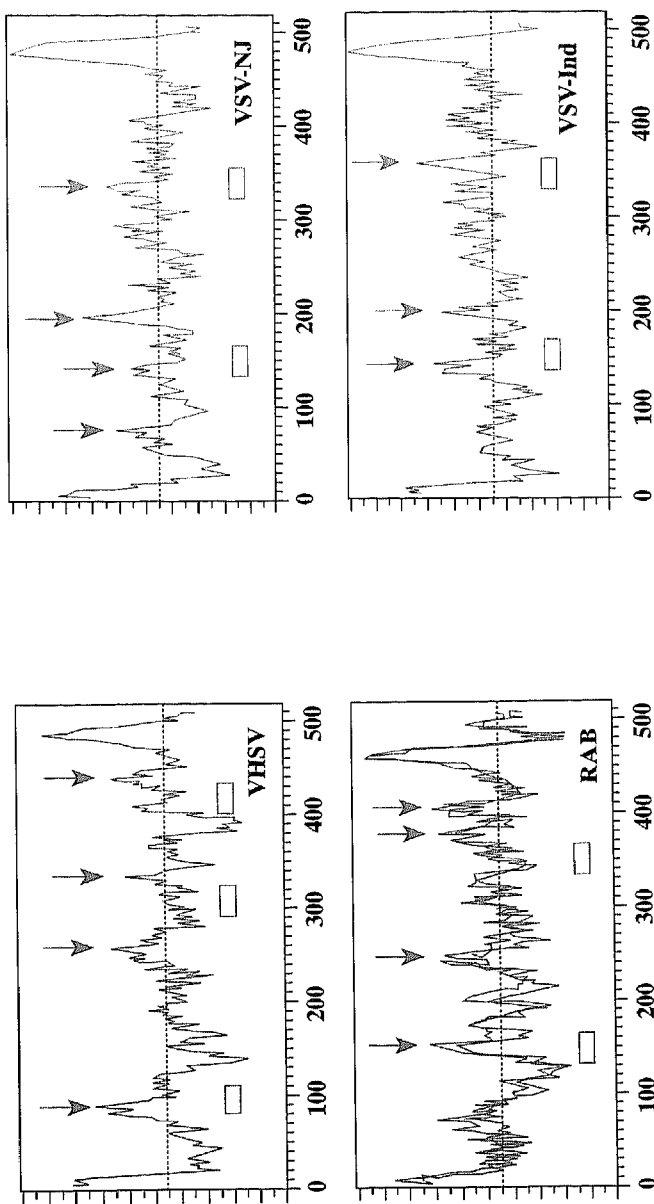


Fig. 4.—Posición relativa de los perfiles hidrofóbicos y de las heptádas repetidas.

Hydrophobic profiles and location of the heptad-repeats

Los perfiles hidrofóbicos de todas las secuencias de la Tabla 1 se obtuvieron con el programa SOAP del PC-Gene (Intelligenetics). Los perfiles de VHSV, rabia, VSV-NJ y VSV-Indiana se superpusieron y las figuras obtenidas se muestran en la figura. Las flechas indican la posición de los picos hidrofóbicos mayoritarios, mientras que en las barras horizontales se muestra la posición de las heptádas repetidas. Eje-Y, escala hidrofóbica; eje-X, posición amino-terminal de los aa de la glicoproteína G de rhabdovirus

The hydrophobic profiles from all the sequences of Table 1 were obtained with the program SOAP from PCGene (Intelligenetics). The profiles from VHSV, rabies, VSV-NJ and VSV-Indiana were superimposed and the figures obtained are shown. Arrows indicate the position of the most hydrophobic peaks where as in horizontal bars are shown the position of the heptad-repeats. Y-axis, hydrophobic scales; X-axis, amino-terminal position of the aa of the glycoprotein G of rhabdoviruses

Finalmente, una región de unión a fosfolípidos (P2) de la G de VHSV ha sido identificado mediante ensayos de unión a fosfolípidos en fase sólida con pepsan, péptidos sintéticos, G purificada o recombinante (Estepa *et al.*, 1994) y VHSV purificado. La región p2 (aa 82 al 109) se encontraba dentro de cinco héptadas repetidas con aa hidrofóbicos en a y d (aa 68 al 102) de los aa 82 a 102 tienen estructura probable de α -hélice. No se sabe todavía si las regiones inmediatas podrían estar relacionadas con la fusión de membranas (Chambers *et al.*, 1990) inducidas por la G de VHSV, sin embargo, existe una dependencia de pH similar en la unión a fosfolípidos por VHSV y en la fusión con membranas celulares mediada por la G de VHSV (Lecocq-Xhonneus *et al.*, 1994).

SUMMARY

Hydrophobic heptads in the glycoprotein of rhabdoviruses

Two or three regions containing three or more successive newly defined heptads of a-d hydrophobic amino acid (aa) repeats have been located in the cDNA-derived aa sequences of the glycoprotein G of all the rhabdoviruses examined (rabies, vesicular stomatitis, fish and plant rhabdoviruses) by computer search. These new heptad-repeats differ from those previously reported in other viruses because of the presence of all the hydrophobic aa in positions a d and because they are not predicted to form coiled-coils by current methods and thus they were not detected before in any rhabdoviruses. The two three heptad-repeats regions were the only parts of the glycoprotein with at least three successive heptad-repeats in all the rhabdoviral sequences studied and had low sequence variability among the members of each of the rhabdoviral genus but show no sequence similarity among the different genus. All these newly detected heptad-repeats were inside or in the vicinity of some of the higher hydrophobic regions in each of the rhabdovirus genus and were found mostly but not always outside the extra aa sequences that occur in the longer insect or plant rhabdovirus glycoprotein G. The correspondence of position and structure of these heptad-repeats among all the rhabdoviruses suggests its participation in common function(s), most probably related to the viral fusion with cellular membranes.

KEY WORDS: Hydrophobic heptad
Rhabdovirus
Glycoprotein

REFERENCIAS BIBLIOGRÁFICAS

- ANILIONIS A., WUNNER W. H., CURTIS P. J., 1981. Structure of the glycoprotein gene in rabies virus. *Nature*, 294, 274-278.
- BENMANSOUR H., LEBLOIS H., COULON P., TUFFEUREAU C., GAUDIN Y., FLAMAND A., LAFAY F., 1991. Antigenicity of rabies virus glycoprotein. *Journal Virology*, 65, 4198-4203.
- BILSEL P. A., NICHOL S. T., 1990. Polymerase errors accumulating during natural evolution of the glycoprotein gene of vesicular stomatitis virus indiana serotype isolates. *Journal Virology*, 64, 4873-4883.
- BOURHY H., KISSI B., TORDO N., 1993. Molecular diversity of the lyssavirus genus. *Virology*, 194, 70-81.
- BULLOUGH P. A., HUGHSON F. M., SKEHEL J. J., WILEY D. C., 1994. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature*, 371, 37-43.
- CAR C. M., KIM P. S., 1993. A spring-loaded mechanism for the conformational change of influenza hemagglutinin. *Cell*, 73, 823-832.
- CHAMBERS P., PRINGLE C. R., EASTON A. J., 1990. Heptad repeat sequences are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins. *Journal General Virology*, 71, 3075-3080.
- COLL J. M., 1995. Synthetic peptides reveal a phospholipid-binding domain in the glycoprotein of

- salmonid rhabdoviruses. *Virus Res* (in press).
- CONZELMANN K. K., COX J. H., SCHNEIDER L. G., THIEL H. J., 1990. Molecular cloning and complete nucleotide sequence of the attenuated rabies virus SAD B19. *Virology*, 175, 485-489.
- ESTEPA A., THIRY M., COLL J. M., 1994. Recombinant protein fragments from haemorrhagic septicaemia rhabdovirus stimulate trout leucocyte anamnestic in vitro responses. *Journal General Virology*, 75, 1329-1338.
- GALLIONE C. J., ROSE J. K., 1985. A single amino acid substitution in a hydrophobic domain causes temperature-sensitive cell-surface transport of a mutant viral glycoprotein. *Journal Virology*, 54, 374-382.
- GAUDIN Y., RUIGROK R. W. H., KNOSSOW M., FLAMAND A., 1993. Low-pH conformational changes of rabies virus glycoprotein and their role in membrane fusion. *Journal Virology*, 67, 1365-1372.
- GAUDIN Y., RUIGROK R. W. H., TUFFEREAU C., KNOSSOW M., FLAMAND A., 1992. Rabies virus glycoprotein is a trimer. *Virology*, 187, 627-632.
- GOLDBERG B. K., MODRELL B., HILLMAN B. I., HEATON A. L., CHOI T., JACKSON O. A., 1991. Structure of the Glycoprotein Gene of *Sonchus yellow net virus*, a plant Rhabdovirus. *Virology*, 185, 32-38.
- KOENER J. F., PASSAVANT C. W., KURATH G., LEONG J., 1987. Nucleotide sequence of a cDNA clone carrying the glycoprotein gene of Infectious Haematopoietic Necrosis virus, a fish rhabdovirus. *Journal Virology*, 61, 1342-1349.
- LAFON M., WIKTOR T. J., MACFARLAN R. I., 1983. Antigenic sites on the CVS rabies virus glycoprotein: analysis with monoclonal antibodies. *Journal General Virology*, 64, 843-851.
- LECOCQ-XHONNEUX F., THIRY M., DHEUR I., ROSSUS M., VANDERHEIJDEN N., MARTIAL J., DEKINKELIN P., 1994. A recombinant viral haemorrhagic septicaemia virus glycoprotein expressed in insect cells induces protective immunity in rainbow trout. *Journal Virology*, 75, 1579-1587.
- LI Y., DRONE C., SAT E., GHOSH H. P., 1993. Mutational analysis of the vesicular stomatitis virus glycoprotein G for membrane fusion domains. *Journal Virology*, 67, 4070-4077.
- LORENZEN N., OLESEN N. J., VESTERGAARD-JORGENSEN P. E., ETZERODT M., HOLTET T. L., THORGERSEN M. C., 1993. Molecular cloning and expression in *Escherichia coli* of the glycoprotein gene of VHS virus and immunization of rainbow trout with the recombinant protein. *Journal General Virology*, 74, 623-630.
- LUPAS A., VAN DYKE M., STOCK J., 1991. Predicting coiled coils from protein sequences. *Science*, 252, 1162-1169.
- MORIMOTO K., OHKUBO A., KAWAI A., 1989. Structure and transcription of the glycoprotein gene of attenuated HEP-Flury strain of rabies virus. *Virology*, 173, 465-477.
- NICHOL S. T., ROWE J. E., FITCH W. M., 1989. Glycoprotein evolution of vesicular stomatitis virus New Jersey. *Virology*, 168, 281-291.
- prehaud c., takehara k., flamand a., bishop d. h. l., 1989. Immunogenic and protective properties of rabies virus glycoprotein expressed by baculovirus vectors. *Virology*, 173, 390-399.
- ROSE J. K., GALLIONE C. J., 1981. Stomatitis virus G and M proteins determined from cDNA clones containing the complete coding regions. *Journal Virology*, 39, 519-528.
- SCHULZ G. E., SCHIMER R. H., 1984. Principles of protein structure. ed. Springer-Verlag, N. Y., 2.
- SEIF G., COULON P., ROLIN P. E., FLAMAND A., 1985. Effect on pathogenicity and sequence characterization of rabies virus mutations affecting antigenic site III of the glycoproteins. *Journal Virology*, 53, 926-934.
- TENINGES D., BRAS F., DEZELEE S., 1993. Genome organization of the sigma rhabdovirus: Six genes and Gene overlap. *Virology*, 193, 1018-1023.
- thiry m., lecoq-xhonneux f., dheur i., renard a., kinkelin d., 1991a. Molecular cloning of the m-RNA coding for the G protein of the viral haemorrhagic septicaemia (VHS) of salmonids. *Journal Veterinary Microbiology*, 23, 221-226.
- THIRY M., LECOQ-XHONNEUX F., DHEUR I., RENARD A., KINKELIN D., 1991b. Sequence of a cDNA carrying the glycoprotein gene and part of the matrix protein M2 gene of viral haemorrhagic septicaemia virus, a fish rhabdovirus. *Biochimica Biophysica Acta*, 1090, 345-347.
- TORDO N., 1991. Contribution of molecular biology to vaccine development and molecular epidemiology of rabies disease. *Memories Institute Butantan*, 53, 31-51.
- TORDO N., BOURHY H., SATHER S., OLLO R., 1993. Structure and expression in baculovirus of the Mokola virus glycoprotein: an efficient recombinant vaccine. *Virology*, 194, 59-69.
- TORDEO N., POCH O., ERMINE A., KEITH G., ROUGEON F., 1986. Walking along the rabies

- genome: is the large G-L intergenic region a remnant gene? *Proceedings National Academy Sciences USA*, 83, 3914-3918.
- VANDEPOL S. B., LE FRANCOIS L., HOLLAND J. J., 1986. Sequences of the major antibody binding epitopes of the Indiana serotype of vesicular stomatitis virus. *Virology*, 148, 312-325.
- WALKER P. J., BYRNE K. A., RIDING G. A., COWLEY J. A., WANG Y., MCWILLIAM S., 1992. The genome of bovine ephemeral fever rhabdovirus contains two related glycoprotein genes. *Virology*, 191, 49-61.
- WUNNER W. H., REAGAN K. J., KOPROWSKI H., 1984. Characterization of saturable binding sites for rabies virus. *Journal Virology*, 50, 691-697.
- ZHANG L., GHOSH H. P., 1994. Characterization of the putative fusogenic domain in vesicular stomatitis virus glycoprotein G. *Journal Virology*, 68, 2186-2193.