SINGAPORE PATENTS ACT(CHAPTER 221)
PATENTS RULES

PATENTS FORM 1	Request for the Grant of a Patent
2001	

Pre-requisites:

a. *Among other requirements, a Date of Filing will be issued only if a description of the invention is filed or if a complete and valid statement is made under Part 9 of this form.*

Estimated Time:

This form may take approximately15 - 20minutes to complete.

General:

a. * *denotes mandatory field.*

PART 1 Reference

Applicant/Agent Reference	2200840/MIT/MK/ANG/ely

PART 2 Title of Invention*

Title of Invention*	HEMAGGLUTININ POLYPEPTIDES, AND REAGENTS AND METHODS RELATING THERETO

PART 3 Details of Applicant*

Note: If there is insufficient space, please use the continuation sheet CS 1.

Applicants 1

Type of Applicant	Corporate
UEN/ Company Code	E0169511A
Name	MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Address Type	Foreign
Address	77 MASSACHUSETTS AVENUE, CAMBRIDGE, MASSACHUSETTS 02139-4307, UNITED STATES OF AMERICA Massachusetts-United States of America
Is SME	
Country of Incorporation	United States of America
State of Incorporation	Massachusetts
Sole Proprietor or Partners' Names	

PART 4 Declaration of Priority under Section 17(2)

Declaration of Priority(s) under Section 17(2)

Application No.	Country	Filing Date
61/389,639	United States of America	04/10/2010

PART 5 Inventor*

A. Are all applicants named above also inventors? Yes ☐ No ☒

B. If yes, are there any other inventors?	Yes ☐ No ☐

PART 6 Claiming the Filing Date of an Earlier Singapore Application

This application is	☒ A divisional of an earlier Singapore application.
	☐ An application filed in response to an order by the Registrar after determination of a question regarding the entitlement of an earlier Singapore application. Application No. 10201709806V Filing Date *(DD/MM/YYYY)* 04/10/2011

PART 7 Disclosure of Invention Prior to Filing

☐ The invention was displayed at an international exhibition on a date later than the start of the period of 12 months before the date of filing this application.

PART 8 Micro-Organism Deposited under the Budapest Treaty

☐ The invention requires the use of a micro-organism and for disclosure purposes, a culture of the micro-organism has been deposited for with an international depository authority under the Budapest Treaty.

PART 9 Description Incorporated by Reference to an Earlier Relevant Application

☐ No description is filed with this request. Instead the description of the invention for which the patent is sought is incorporated in the application by reference to, and is completely contained in, the following earlier relevant application, as filed:

Country

Application No.

Filing Date *(DD/MM/YYYY)*

PART 10 Checklist*

Checklist A — The application consists of the following no. of sheets

Description	513	sheet(s)
Claim	5	sheet(s)
Drawing	54	sheet(s)
Abstract	1	sheet(s)

(Note:Please indicate on the abstract the figure of the drawing, if any, which should accompany the abstract when published)

Total number of sheets	573	sheet(s)

Checklist B — The application as filed is accompanied by:

	☒ Statement of inventorship & right to grant *(Patents Form 8)*
	☐ International exhibition certificate

PART 11 Contact Details*

Note:
a. This part need not be filled up if the person's address in Part 3 is to be used as an address for service in Singapore.
b. The email address and telephone number provided are for contact purposes in case of query.
c. The official correspondence will be sent electronically via IP2SG.

Singapore registered patent agent, or an advocate and solicitor, who has in force a practising certificate. Please refer to Singapore Patents Act Section 105 for details.

Agent UEN/Company Code	T14LL0024B
Agent Name	DONALDSON & BURKINSHAW LLP
Representative or C/O Name	
	Address for Service in Singapore
	Block/ House No.
	Street Name
	Level - Unit 3667
	Building Name ROBINSON ROAD POST OFFICE
	Postal Code 905667
Contact Person Direct Telephone No. Email Address	

PART 12 Declaration*

Declaration	**By Person Filing the Application** I do hereby declare that the information furnished above is true to the best of my knowledge. **By Agent** I do hereby declare that : i. I have been duly authorised to act as an agent, for the purposes of this application, on behalf of the person(s) filing this form. ii. The information furnished above on behalf of the person(s) filing this form is true to the best of the person(s)' knowledge.
Name	EVELINE ANGSTETRA

Signature		
Date *(DD/MM/YYYY)*	27/5/2022	
No. of Extra Sheets Attached to this Form	0	sheet(s)

GST Information

The tax invoice will be issued to:

Name: MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Abstract

HEMAGGLUTININ POLYPEPTIDES, AND REAGENTS AND METHODS RELATING THERETO

The present invention provides a system for analyzing interactions between glycans and interaction partners that bind to them. The present invention also provides HA polypeptides that bind to umbrella-topology glycans, and reagents and methods relating thereto.

FIG. 6A

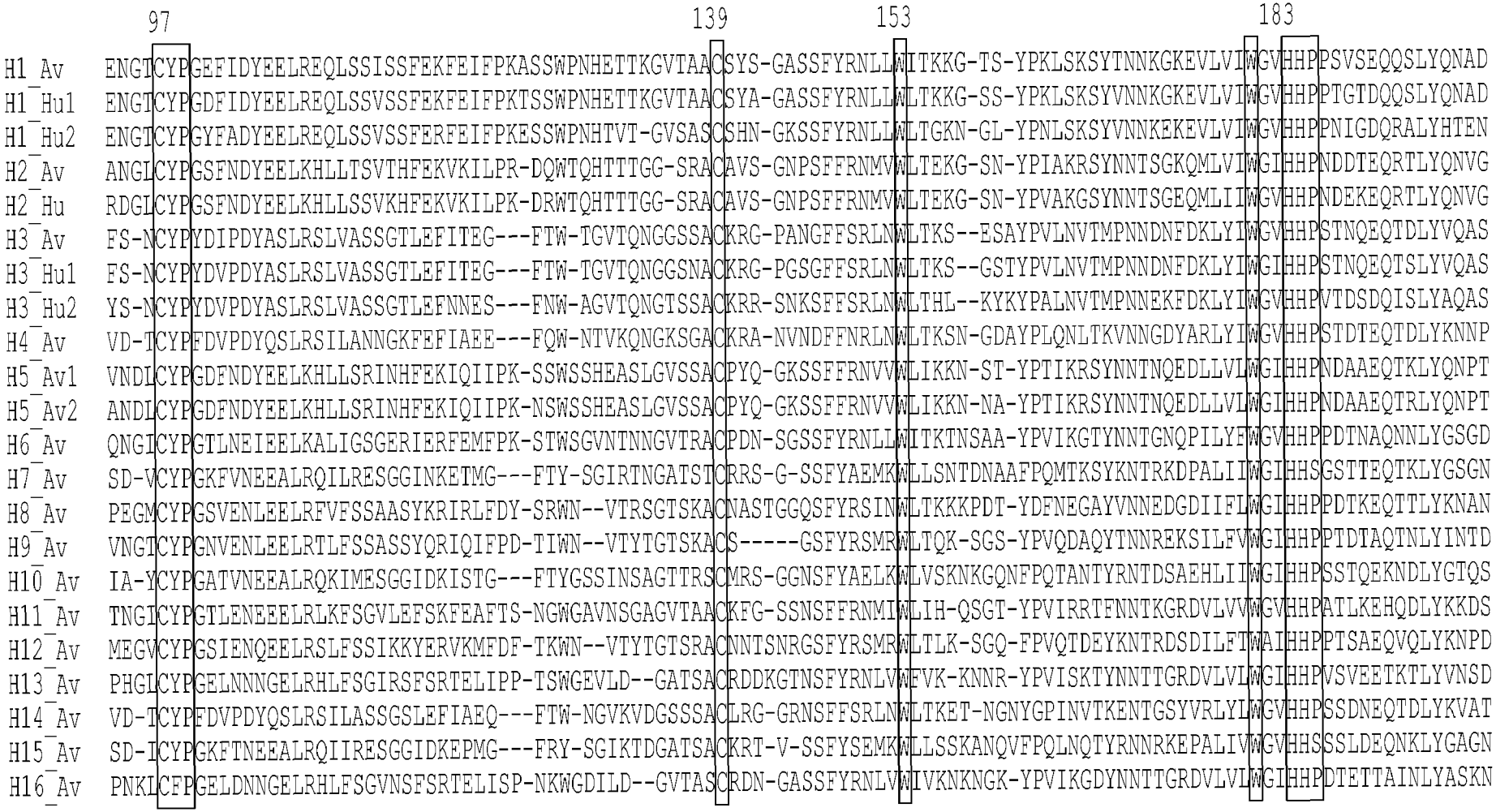

97
139
153
183
H1_Av ENGTCYPGEFIDYEELREQLSSISSFEKFEIFPKASSWPNHETTKGVTAACSYS-GASSFYRNLLWITKKG-TS-YPKLSKSYTNNKGKEVLVIWGVHHPPSVSEQQSLYQNAD
H1_Hu1 ENGTCYPGDFIDYEELREQLSSVSSFEKFEIFPKTSSWPNHETTKGVTAACSYA-GASSFYRNLLWLTKKG-SS-YPKLSKSYVNNKGKEVLVIWGVHHPPTGTDQQSLYQNAD
H1_Hu2 ENGTCYPGYFADYEELREQLSSVSSFERFEIFPKESSWPNHTVT-GVSASCSHN-GKSSFYRNLLWLTGKN-GL-YPNLSKSYVNNKEKEVLVIWGVHHPPNIGDQRALYHTEN
H2_Av ANGLCYPGSFNDYEELKHLLTSVTHFEKVKILPR-DQWTQHTTTGG-SRACAVS-GNPSFFRNMVWLTEKG-SN-YPIAKRSYNNTSGKQMLVIWGIHHPNDDTEQRTLYQNVG
H2_Hu RDGLCYPGSFNDYEELKHLLSSVKHFEKVKILPK-DRWTQHTTTGG-SRACAVS-GNPSFFRNMVWLTEKG-SN-YPVAKGSYNNTSGEQMLIIWGVHHPNDEKEQRTLYQNVG
H3_Av FS-NCYPYDIPDYASLRSLVASSGTLEFITEG---FTW-TGVTQNGGSSACKRG-PANGFFSRLNWLTKS--ESAYPVLNVTMPNNDNFDKLYIWGVHHPSTNQEQTDLYVQAS
H3_Hu1 FS-NCYPYDVPDYASLRSLVASSGTLEFITEG---FTW-TGVTQNGGSNACKRG-PGSGFFSRLNWLTKS--GSTYPVLNVTMPNNDNFDKLYIWGIHHPSTNQEQTSLYVQAS
H3_Hu2 YS-NCYPYDVPDYASLRSLVASSGTLEFNNES---FNW-AGVTQNGTSSACKRR-SNKSFFSRLNWLTHL--KYKYPALNVTMPNNEKFDKLYIWGVHHPVTDSDQISLYAQAS
H4_Av VD-TCYPFDVPDYQSLRSILANNGKFEFIAEE---FQW-NTVKQNGKSGACKRA-NVNDFFNRLNWLTKSN-GDAYPLQNLTKVNNGDYARLYIWGVHHPSTDTEQTDLYKNNP
H5_Av1 VNDLCYPGDFNDYEELKHLLSRINHFEKIQIIPK-SSWSSHEASLGVSSACPYQ-GKSSFFRNVVWLIKKN-ST-YPTIKRSYNNTNQEDLLVLWGIHHPNDAAEQTKLYQNPT
H5_Av2 ANDLCYPGDFNDYEELKHLLSRINHFEKIQIIPK-NSWSSHEASLGVSSACPYQ-GKSSFFRNVVWLIKKN-NA-YPTIKRSYNNTNQEDLLVLWGIHHPNDAAEQTRLYQNPT
H6_Av QNGICYPGTLNEIEELKALIGSGERIERFEMFPK-STWSGVNTNNGVTRACPDN-SGSSFYRNLLWITKTNSAA-YPVIKGTYNNTGNQPILYFWGVHHPPDTNAQNNLYGSGD
H7_Av SD-VCYPGKFVNEEALRQILRESGGINKETMG---FTY-SGIRTNGATSTCRRS-G-SSFYAEMKWLLSNTDNAAFPQMTKSYKNTRKDPALIIWGIHHSGSTTEQTKLYGSGN
H8_Av PEGMCYPGSVENLEELRFVFSSAASYKRIRLFDY-SRWN--VTRSGTSKACNASTGGQSFYRSINWLTKKKPDT-YDFNEGAYVNNEDGDIIFLWGIHHPPDTKEQTTLYKNAN
H9_Av VNGTCYPGNVENLEELRTLFSSASSYQRIQIFPD-TIWN--VTYTGTSKACS-----GSFYRSMRWLTQK-SGS-YPVQDAQYTNNREKSILFVWGIHHPPTDTAQTNLYINTD
H10_Av IA-YCYPGATVNEEALRQKIMESGGIDKISTG---FTYGSSINSAGTTRSCMRS-GGNSFYAELKWLVSKNKGQNFPQTANTYRNTDSAEHLIIWGIHHPSSTQEKNDLYGTQS
H11_Av TNGICYPGTLENEEELRLKFSGVLEFSKFEAFTS-NGWGAVNSGAGVTAACKFG-SSNSFFRNMIWLIH-QSGT-YPVIRRTFNNTKGRDVLVVWGVHHPATLKEHQDLYKKDS
H12_Av MEGVCYPGSIENQEELRSLFSSIKKYERVKMFDF-TKWN--VTYTGTSRACNNTSNRGSFYRSMRWLTLK-SGQ-FPVQTDEYKNTRDSDILFTWAIHHPPTSAEQVQLYKNPD
H13_Av PHGLCYPGELNNNGELRHLFSGIRSFSRTELIPP-TSWGEVLD--GATSACRDDKGTNSFYRNLVWFVK-KNNR-YPVISKTYNNTTGRDVLVLWGIHHPVSVEETKTLYVNSD
H14_Av VD-TCYPFDVPDYQSLRSILASSGSLEFIAEQ---FTW-NGVKVDGSSSACLRG-GRNSFFSRLNWLTKET-NGNYGPINVTKENTGSYVRLYLWGVHHPSSDNEQTDLYKVAT
H15_Av SD-ICYPGKFTNEEALRQIIRESGGIDKEPMG---FRY-SGIKTDGATSACKRT-V-SSFYSEMKWLLSSKANQVFPQLNQTYRNNRKEPALIVWGVHHSSSLDEQNKLYGAGN
H16_Av PNKLCFPGELDNNGELRHLFSGVNSFSRTELISP-NKWGDILD--GVTASCRDN-GASSFYRNLVWIVKNKNGK-YPVIKGDYNNTTGRDVLVLWGIHHPDTETTAINLYASKN

FIG. 1A

```
H1_Av    AYVSVGSSKYNRRFAPEIAARPEVRGQAGRMNYYWTLLDQGDTITFEATGNLIAPWYAFALNKGSD--------SGIITS-DAPVH-NCDTRCQTPHGALNSSLPFQNVHPIT
H1_Hu1   AYVSVGSSKYNRRFTPEIAARPKVRDQAGRMNYYWTLLEPGDTITFEATGNLIAPWYAFALNRGSG--------SGIITS-DAPVH-DCNTKCQTPHGAINSSLPFQNIHPVT
H1_Hu2   AYVSVVSSHYSRRFTPEIAKRPKVRDQEGRINYYWTLLEPGDTIIFEANGNLIAPWYAFALSRGFG--------SGIITS-NAPMD-ECDAKCQTPQGAINSSLPFQNVHPVT
H2_Av    TYVSVGTSTLNKRSIPEIATRPKVNGQGGRMEFSWTLLETWDVINFESTGNLIAPEYGFKISKRGS--------SGIMKT-EKTLE-NCETKCQTPLGAINTTLPFHNIHPLT
H2_Hu    TYVSVGTSTLNKRSTPDIATRPKVNGLGSRMEFSWTLLDMWDTINFESTGNLIAPEYGFKISKRGS--------SGIMKT-EGTLE-NCETKCQTPLGAINTTLPFHNVHPLT
H3_Av    GRVTVSTRRSQQTIIPNIGSRPWVRGQPGRISIYWTIVKPGDVLVINSNGNLIAPRGYFKMRTG---------KSSIMRS-DAPID-TCISECITPNGSIPNDKPFQNVNKIT
H3_Hu1   GRVTVSTRRSQQTIIPNIGSRPWVRGLSSRISIYWTIVKPGDVLVINSNGNLIAPRGYFKMRTG---------KSSIMRS-DAPID-TCISECITPNGSIPNDKPFQNVNKIT
H3_Hu2   GRITVSTKRSQQTVIPNIGYRPRVRDISSRISIYWTIVKPGDILLINSTGNLIAPRGYFKIRSG---------KSSIMRS-DAPIG-KCNSECITPNGSIPNDKPFQNVNRIT
H4_Av    GRVTVSTKTSQTSVVPNIGSRPWVRGQSGRISFYWTIVDPGDIIVFNTIGNLIAPRGHYKLNSQK--------KSTILNT-AVPIG-SCVSKCHTDRGSITTTKPFQNISRIS
H5_Av1   TYISVGTSTLNQRLVPRIATRSKVNGQSGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKGD--------STIMKS-ELEYG-NCNTKCQTPMGAINSSMPFHNIHPLT
H5_Av2   TYISVGTSTLNQRLVPKIATRSKVNGQNGRMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKGD--------SAIMKS-ELEYG-NCNTKCQTPMGAINSSMPFHNIHPLT
H6_Av    RYVRMGTESMNFAKGPEISARPVVNGQRGRIDYYWSVLKPGETLNVESNGNLIAPWYAYKFVSTNSK-------GAVFKS-NLPIE-NCDATCQTIAGVLRTNKTFQNVSPLW
H7_Av    KLITVESSNYQQSFVPSPGARPKVDGQSGRIDFHWLMLNPNDTITFSFNGAFIAPDRASFLR----------GKSMGIQS-GVQVDDNCEGDCYHSGGTIISNLPFQNINSRA
H8_Av    TLSSVTTNTINRSFQPNIGPRPLVRGQQGRMDYYWGILKRGETLKIRTNGNLIAPEFGYLLKGESY--------GRIIQNEDIPIG-NCNTKCQTYAGAINSSKPFQNASRHY
H9_Av    TTTSVTTEDLNRIFKPVIGPRPLVNGLQGRINYYWSLVKPGQTLRVRSNGNLIAPWYGHVLSGGSH--------GRILKT-DLNSG-NCVVQCQTEKGGLMSTLPFHNISKYA
H10_Av   LSISVGSSTYQNNFVPVVGARPQVNGQSGRIDFHWTMVQPGDNITFSHNGGLIAPSRVSKLK----------GRGLGIQS-GASVDNDCESKCFWKGGSINTKLPFQNLSPRT
H11_Av   SYVAVDSESYNRRFTPEISTRPKVNGQAGRMTFYWTIVKPGEAITFESNGAFLAPRYAFELVSLGN--------GKLFRS-DLNIE-SCSTKCQSEIGGINTNRSFHNVHRNT
H12_Av   TLSSVTTDEINRSFKPNIGPRPLVRGQQGRMDYYWAVLKPGQTVKIQTNGNLIAPEYGHLITGKSH--------GRILKN-NLPVG-QCVTECQLNEGVMNTSKPFQNTSKHY
H13_Av   PYTLVSTKSWSEKYKLETGVRPGYNGQRSWMKIYWSLLHPGEMITFESNGGLLAPRYGYIIEEYGK--------GRIFQS-RIRMS-KCNTKCQTSVGGINTNRTFQNIDKNA
H14_Av   GRVTVSTRSDQISIVPNIGSRPRVRNQSGRISIYWTLVNPGDSIIFNSIGNLIAPRGHYKISKST--------KSTVLKS-DKRIG-SCTSPCLTDKGSIQSDKPFQNVSRIA
H15_Av   KLITVGSSKYQQSFSPSPGDRPKVNGQAGRIDFHWMLLDPGDTVTFTFNGAFIAPDRATFLRSNAPSGVEYNGKSLGIQS-DAQIDESCEGECFYSGGTINSPLPFQNIDSWA
H16_Av   PYTLVSTKEWSKRYELEIGTRIG-DGQRSWMKLYWHLMRPGERIMFESNGGLIAPRYGYIIEKYGT--------GRIFQS-GVRMA-KCNTKCQTSLGGINTNKTFQNIERNA
```

FIG. 1B

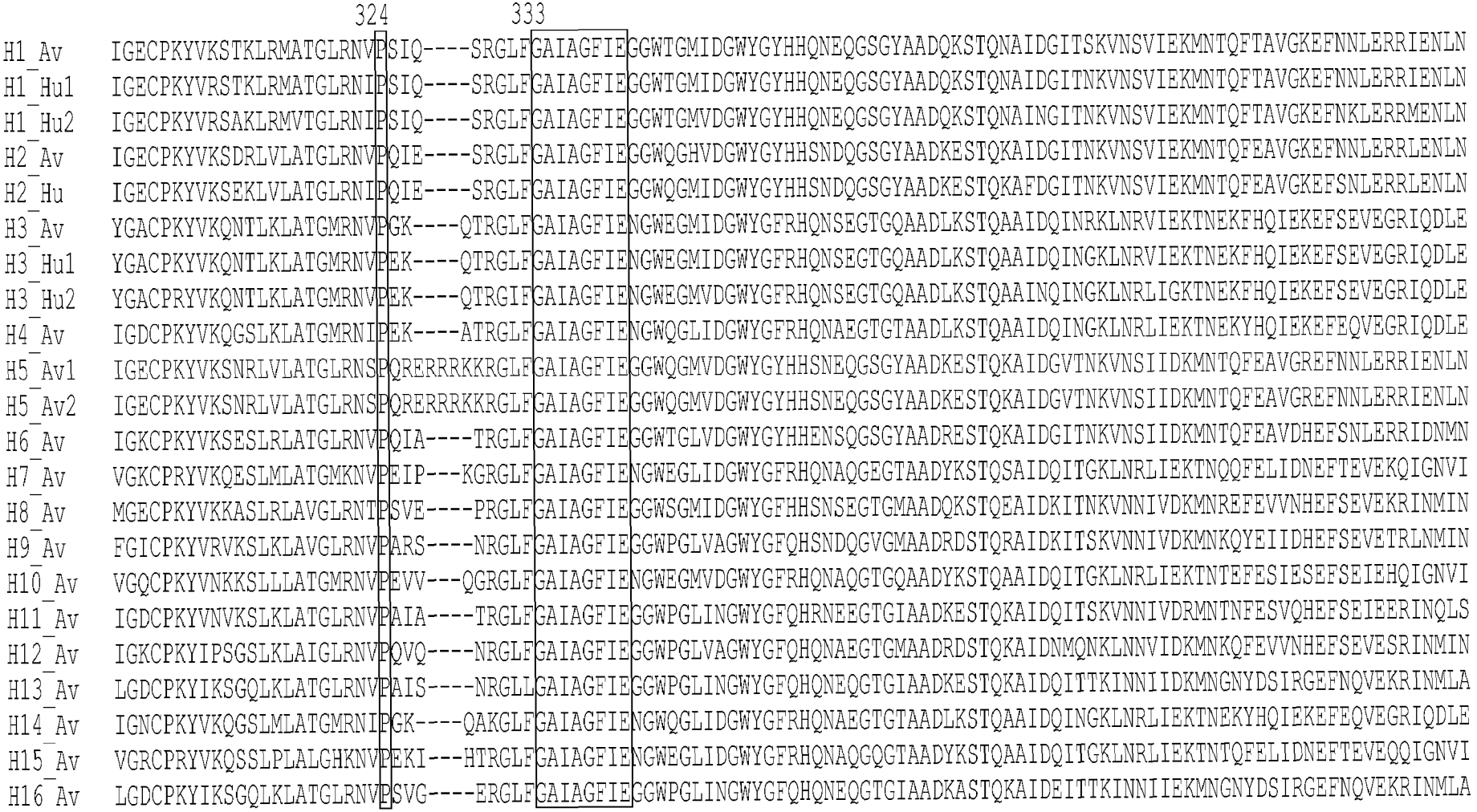

324
333
H1_Av IGECPKYVKSTKLRMATGLRNVPSIQ----SRGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAIDGITSKVNSVIEKMNTQFTAVGKEFNNLERRIENLN
H1_Hu1 IGECPKYVRSTKLRMATGLRNIPSIQ----SRGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAIDGITNKVNSVIEKMNTQFTAVGKEFNNLERRIENLN
H1_Hu2 IGECPKYVRSAKLRMVTGLRNIPSIQ----SRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNSVIEKMNTQFTAVGKEFNKLERRMENLN
H2_Av IGECPKYVKSDRLVLATGLRNVPQIE----SRGLFGAIAGFIEGGWQGHVDGWYGYHHSNDQGSGYAADKESTQKAIDGITNKVNSVIEKMNTQFEAVGKEFNNLERRLENLN
H2_Hu IGECPKYVKSEKLVLATGLRNIPQIE----SRGLFGAIAGFIEGGWQGMIDGWYGYHHSNDQGSGYAADKESTQKAFDGITNKVNSVIEKMNTQFEAVGKEFSNLERRLENLN
H3_Av YGACPKYVKQNTLKLATGMRNVPGK----QTRGLFGAIAGFIENGWEGMIDGWYGFRHQNSEGTGQAADLKSTQAAIDQINRKLNRVIEKTNEKFHQIEKEFSEVEGRIQDLE
H3_Hu1 YGACPKYVKQNTLKLATGMRNVPEK----QTRGLFGAIAGFIENGWEGMIDGWYGFRHQNSEGTGQAADLKSTQAAIDQINGKLNRVIEKTNEKFHQIEKEFSEVEGRIQDLE
H3_Hu2 YGACPRYVKQNTLKLATGMRNVPEK----QTRGIFGAIAGFIENGWEGMVDGWYGFRHQNSEGTGQAADLKSTQAAINQINGKLNRLIGKTNEKFHQIEKEFSEVEGRIQDLE
H4_Av IGDCPKYVKQGSLKLATGMRNIPEK----ATRGLFGAIAGFIENGWQGLIDGWYGFRHQNAEGTGTAADLKSTQAAIDQINGKLNRLIEKTNEKYHQIEKEFEQVEGRIQDLE
H5_Av1 IGECPKYVKSNRLVLATGLRNSPQRERRRKKRGLFGAIAGFIEGGWQGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFNNLERRIENLN
H5_Av2 IGECPKYVKSNRLVLATGLRNSPQRERRRKKRGLFGAIAGFIEGGWQGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFNNLERRIENLN
H6_Av IGKCPKYVKSESLRLATGLRNVPQIA----TRGLFGAIAGFIEGGWTGLVDGWYGYHHENSQGSGYAADRESTQKAIDGITNKVNSIIDKMNTQFEAVDHEFSNLERRIDNMN
H7_Av VGKCPRYVKQESLMLATGMKNVPEIP---KGRGLFGAIAGFIENGWEGLIDGWYGFRHQNAQGEGTAADYKSTQSAIDQITGKLNRLIEKTNQQFELIDNEFTEVEKQIGNVI
H8_Av MGECPKYVKKASLRLAVGLRNTPSVE----PRGLFGAIAGFIEGGWSGMIDGWYGFHHSNSEGTGMAADQKSTQEAIDKITNKVNNIVDKMNREFEVVNHEFSEVEKRINMIN
H9_Av FGICPKYVRVKSLKLAVGLRNVPARS----NRGLFGAIAGFIEGGWPGLVAGWYGFQHSNDQGVGMAADRDSTQRAIDKITSKVNNIVDKMNKQYEIIDHEFSEVETRLNMIN
H10_Av VGQCPKYVNKKSLLLATGMRNVPEVV---QGRGLFGAIAGFIENGWEGMVDGWYGFRHQNAQGTGQAADYKSTQAAIDQITGKLNRLIEKTNTEFESIESEFSEIEHQIGNVI
H11_Av IGDCPKYVNVKSLKLATGLRNVPAIA----TRGLFGAIAGFIEGGWPGLINGWYGFQHRNEEGTGIAADKESTQKAIDQITSKVNNIVDRMNTNFESVQHEFSEIEERINQLS
H12_Av IGKCPKYIPSGSLKLAIGLRNVPQVQ----NRGLFGAIAGFIEGGWPGLVAGWYGFQHQNAEGTGMAADRDSTQKAIDNMQNKLNNVIDKMNKQFEVVNHEFSEVESRINMIN
H13_Av LGDCPKYIKSGQLKLATGLRNVPAIS----NRGLLGAIAGFIEGGWPGLINGWYGFQHQNEQGTGIAADKESTQKAIDQITTKINNIIDKMNGNYDSIRGEFNQVEKRINMLA
H14_Av IGNCPKYVKQGSLMLATGMRNIPGK----QAKGLFGAIAGFIENGWQGLIDGWYGFRHQNAEGTGTAADLKSTQAAIDQINGKLNRLIEKTNEKYHQIEKEFEQVEGRIQDLE
H15_Av VGRCPRYVKQSSLPLALGHKNVPEKI---HTRGLFGAIAGFIENGWEGLIDGWYGFRHQNAQGQGTAADYKSTQAAIDQITGKLNRLIEKTNTQFELIDNEFTEVEQQIGNVI
H16_Av LGDCPKYIKSGQLKLATGLRNVPSVG----ERGLFGAIAGFIEGGWPGLINGWYGFQHQNEQGTGIAADKASTQKAIDEITTKINNIIEKMNGNYDSIRGEFNQVEKRINMLA

FIG. 1C

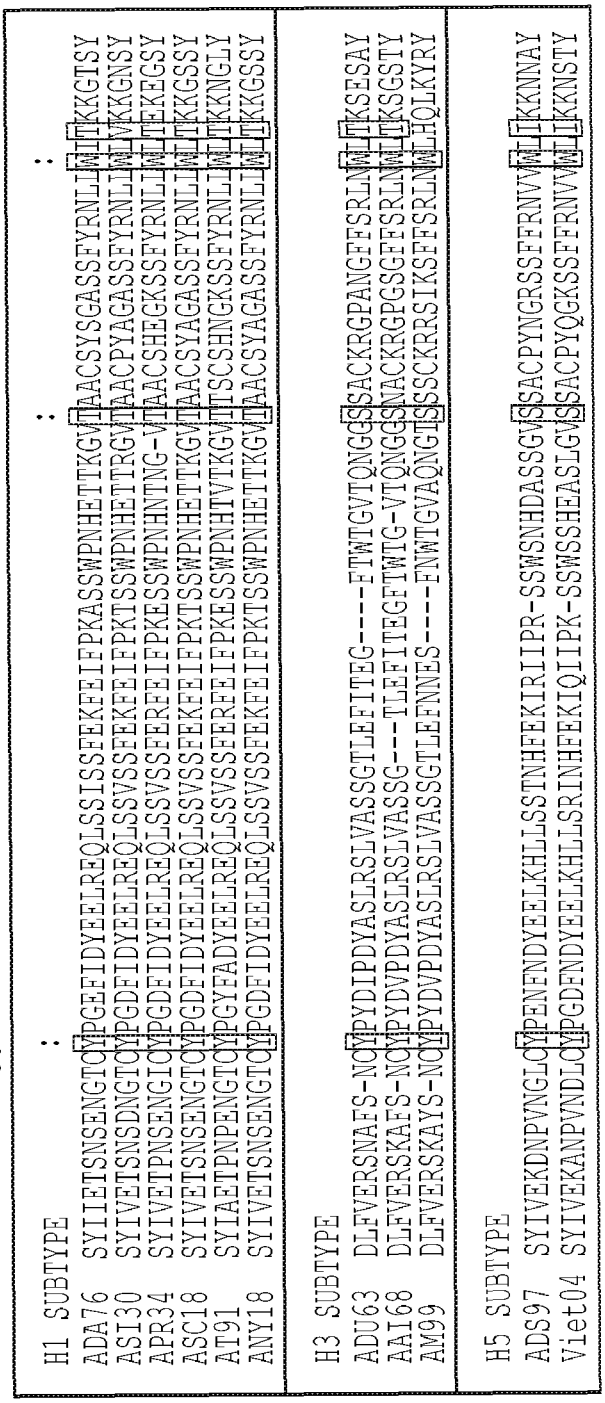

98
136
153
H1 SUBTYPE
ADA76
AST30
APR34
ASC18
AT91
ANY18
H3 SUBTYPE
ADU63
AAI68
AM99
H5 SUBTYPE
ADS97
Viet04

FIG. 2A

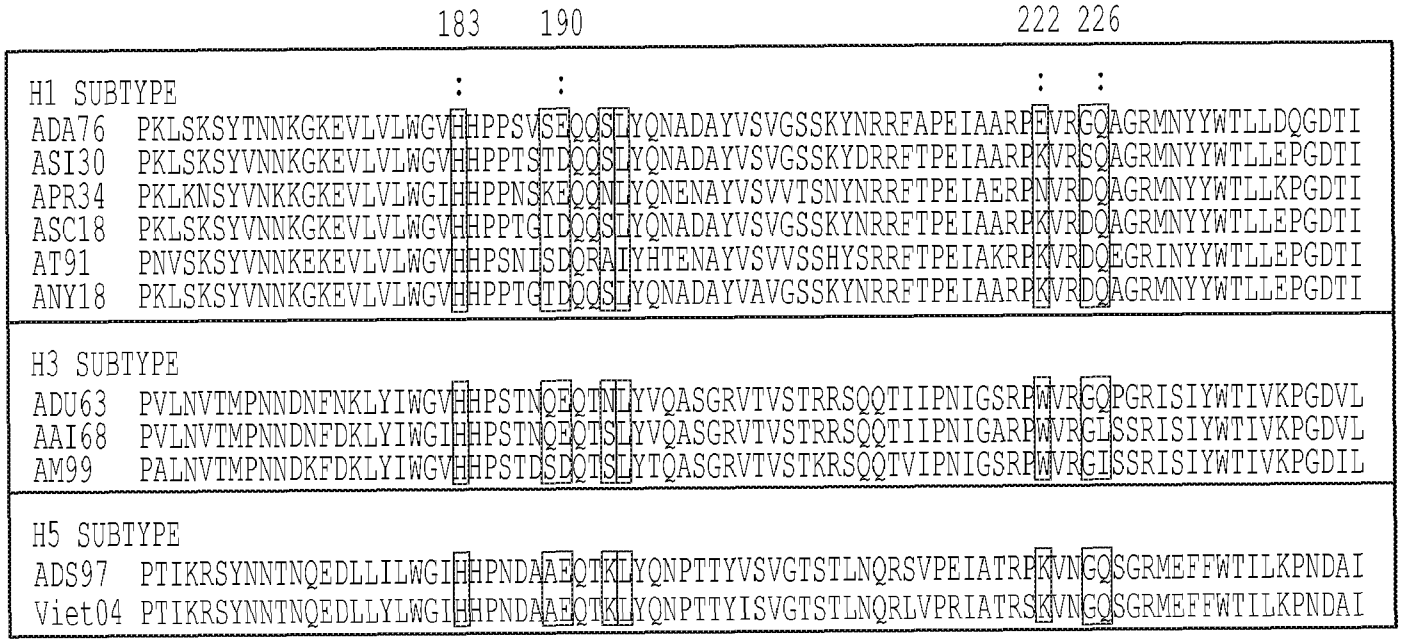

183 190 222 226
H1 SUBTYPE
ADA76 PKLSKSYTNNKGKEVLVLWGVHHPPSVSEQQSLYQNADAYVSVGSSKYNRRFAPEIAARPEVRGQAGRMNYYWTLLDQGDTI
ASI30 PKLSKSYVNNKGKEVLVLWGVHHPPTSTDQQSLYQNADAYVSVGSSKYDRRFTPEIAARPKVRSQAGRMNYYWTLLEPGDTI
APR34 PKLKNSYVNKKGKEVLVLWGIHHPPNSKEQQNLYQNENAYVSVVTSNYNRRFTPEIAERPNVRDQAGRMNYYWTLLKPGDTI
ASC18 PKLSKSYVNNKGKEVLVLWGVHHPPTGIDQQSLYQNADAYVSVGSSKYNRRFTPEIAARPKVRDQAGRMNYYWTLLEPGDTI
AT91 PNVSKSYVNNKEKEVLVLWGVHHPSNISDQRAIYHTENAYVSVVSSHYSRRFTPEIAKRPKVRDQEGRINYYWTLLEPGDTI
ANY18 PKLSKSYVNNKGKEVLVLWGVHHPPTGTDQQSLYQNADAYVAVGSSKYNRRFTPEIAARPKVRDQAGRMNYYWTLLEPGDTI
H3 SUBTYPE
ADU63 PVLNVTMPNNDNFNKLYIWGVHHPSTNQEQTNLYVQASGRVTVSTRRSQQTIIPNIGSRPWVRGQPGRISIYWTIVKPGDVL
AAI68 PVLNVTMPNNDNFDKLYIWGIHHPSTNQEQTSLYVQASGRVTVSTRRSQQTIIPNIGARPWVRGLSSRISIYWTIVKPGDVL
AM99 PALNVTMPNNDKFDKLYIWGVHHPSTDSDQTSLYTQASGRVTVSTKRSQQTVIPNIGSRPWVRGLSSRISIYWTIVKPGDIL
H5 SUBTYPE
ADS97 PTIKRSYNNTNQEDLLILWGIHHPNDAAEQTKLYQNPTTYVSVGTSTLNQRSVPEIATRPKVNGQSGRMEFFWTILKPNDAI
Viet04 PTIKRSYNNTNQEDLLYLWGIHHPNDAAEQTKLYQNPTTYISVGTSTLNQRLVPRIATRSKVNGQSGRMEFFWTILKPNDAI

FIG. 2B

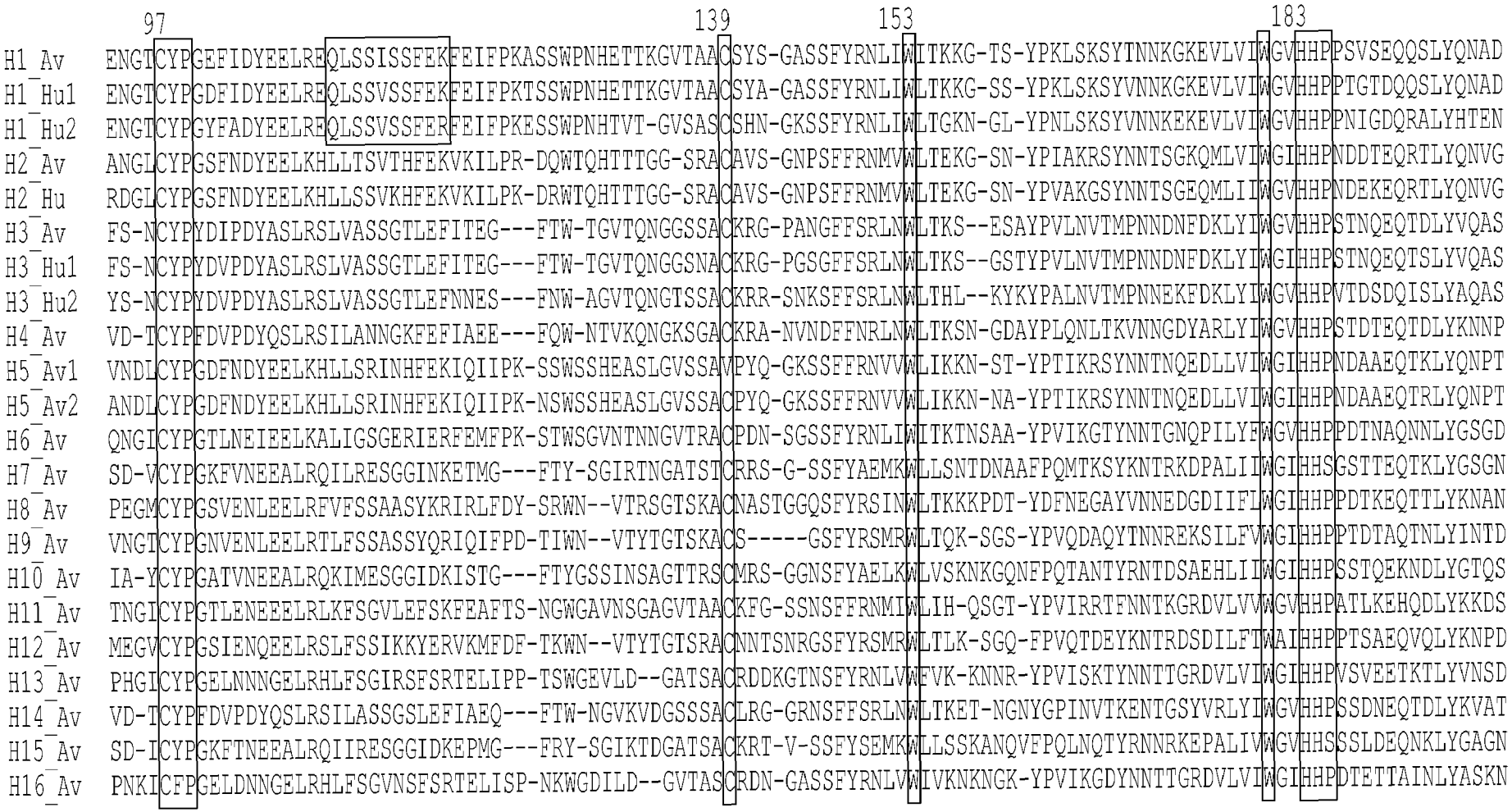

97
139
153
183
H1_Av ENGTCYPGEFIDYEELREQLSSISSFEKFEIFPKASSWPNHETTKGVTAACSYS-GASSFYRNLIWITKKG-TS-YPKLSKSYTNNKGKEVLVIWGVHHPPSVSEQQSLYQNAD
H1_Hu1 ENGTCYPGDFIDYEELREQLSSVSSFEKFEIFPKTSSWPNHETTKGVTAACSYA-GASSFYRNLIWLTKKG-SS-YPKLSKSYVNNKGKEVLVIWGVHHPPTGTDQQSLYQNAD
H1_Hu2 ENGTCYPGYFADYEELREQLSSVSSFERFEIFPKESSWPNHTVT-GVSASCSHN-GKSSFYRNLIWLTGKN-GL-YPNLSKSYVNNKEKEVLVIWGVHHPPNIGDQRALYHTEN
H2_Av ANGLCYPGSFNDYEELKHLLTSVTHFEKVKILPR-DQWTQHTTTGG-SRACAVS-GNPSFFRNMVWLTEKG-SN-YPIAKRSYNNTSGKQMLVIWGIHHPNDDTEQRTLYQNVG
H2_Hu RDGLCYPGSFNDYEELKHLLSSVKHFEKVKILPK-DRWTQHTTTGG-SRACAVS-GNPSFFRNMVWLTEKG-SN-YPVAKGSYNNTSGEQMLIIWGVHHPNDEKEQRTLYQNVG
H3_Av FS-NCYPYDIPDYASLRSLVASSGTLEFITEG---FTW-TGVTQNGGSSACKRG-PANGFFSRLNWLTKS--ESAYPVLNVTMPNNDNFDKLYIWGVHHPSTNQEQTDLYVQAS
H3_Hu1 FS-NCYPYDVPDYASLRSLVASSGTLEFITEG---FTW-TGVTQNGGSNACKRG-PGSGFFSRLNWLTKS--GSTYPVLNVTMPNNDNFDKLYIWGIHHPSTNQEQTSLYVQAS
H3_Hu2 YS-NCYPYDVPDYASLRSLVASSGTLEFNNES---FNW-AGVTQNGTSSACKRR-SNKSFFSRLNWLTHL--KYKYPALNVTMPNNEKFDKLYIWGVHHPVTDSDQISLYAQAS
H4_Av VD-TCYPFDVPDYQSLRSILANNGKFEFIAEE---FQW-NTVKQNGKSGACKRA-NVNDFFNRLNWLTKSN-GDAYPLQNLTKVNNGDYARLYIWGVHHPSTDTEQTDLYKNNP
H5_Av1 VNDLCYPGDFNDYEELKHLLSRINHFEKIQIIPK-SSWSSHEASLGVSSAVPYQ-GKSSFFRNVVWLIKKN-ST-YPTIKRSYNNTNQEDLLVIWGIHHPNDAAEQTKLYQNPT
H5_Av2 ANDLCYPGDFNDYEELKHLLSRINHFEKIQIIPK-NSWSSHEASLGVSSACPYQ-GKSSFFRNVVWLIKKN-NA-YPTIKRSYNNTNQEDLLVIWGIHHPNDAAEQTRLYQNPT
H6_Av QNGICYPGTLNEIEELKALIGSGERIERFEMFPK-STWSGVNTNNGVTRACPDN-SGSSFYRNLIWITKTNSAA-YPVIKGTYNNTGNQPILYFWGVHHPPDTNAQNNLYGSGD
H7_Av SD-VCYPGKFVNEEALRQILRESGGINKETMG---FTY-SGIRTNGATSTCRRS-G-SSFYAEMKWLLSNTDNAAFPQMTKSYKNTRKDPALIIWGIHHSGSTTEQTKLYGSGN
H8_Av PEGMCYPGSVENLEELRFVFSSAASYKRIRLFDY-SRWN--VTRSGTSKACNASTGGQSFYRSINWLTKKKPDT-YDFNEGAYVNNEDGDIIFLWGIHHPPDTKEQTTLYKNAN
H9_Av VNGTCYPGNVENLEELRTLFSSASSYQRIQIFPD-TIWN--VTYTGTSKACS-----GSFYRSMRWLTQK-SGS-YPVQDAQYTNNREKSILFVWGIHHPPTDTAQTNLYINTD
H10_Av IA-YCYPGATVNEEALRQKIMESGGIDKISTG---FTYGSSINSAGTTRSCMRS-GGNSFYAELKWLVSKNKGQNFPQTANTYRNTDSAEHLIIWGIHHPSSTQEKNDLYGTQS
H11_Av TNGICYPGTLENEEELRLKFSGVLEFSKFEAFTS-NGWGAVNSGAGVTAACKFG-SSNSFFRNMIWLIH-QSGT-YPVIRRTFNNTKGRDVLVVWGVHHPATLKEHQDLYKKDS
H12_Av MEGVCYPGSIENQEELRSLFSSIKKYERVKMFDF-TKWN--VTYTGTSRACNNTSNRGSFYRSMRWLTLK-SGQ-FPVQTDEYKNTRDSDILFTWAIHHPPTSAEQVQLYKNPD
H13_Av PHGICYPGELNNNGELRHLFSGIRSFSRTELIPP-TSWGEVLD--GATSACRDDKGTNSFYRNLVWFVK-KNNR-YPVISKTYNNTTGRDVLVIWGIHHPVSVEETKTLYVNSD
H14_Av VD-TCYPFDVPDYQSLRSILASSGSLEFIAEQ---FTW-NGVKVDGSSSACLRG-GRNSFFSRLNWLTKET-NGNYGPINVTKENTGSYVRLYIWGVHHPSSDNEQTDLYKVAT
H15_Av SD-ICYPGKFTNEEALRQIIRESGGIDKEPMG---FRY-SGIKTDGATSACKRT-V-SSFYSEMKWLLSSKANQVFPQLNQTYRNNRKEPALIVWGVHHSSSLDEQNKLYGAGN
H16_Av PNKICFPGELDNNGELRHLFSGVNSFSRTELISP-NKWGDILD--GVTASCRDN-GASSFYRNLVWIVKNKNGK-YPVIKGDYNNTTGRDVLVIWGIHHPDTETTAINLYASKN

FIG. 3A

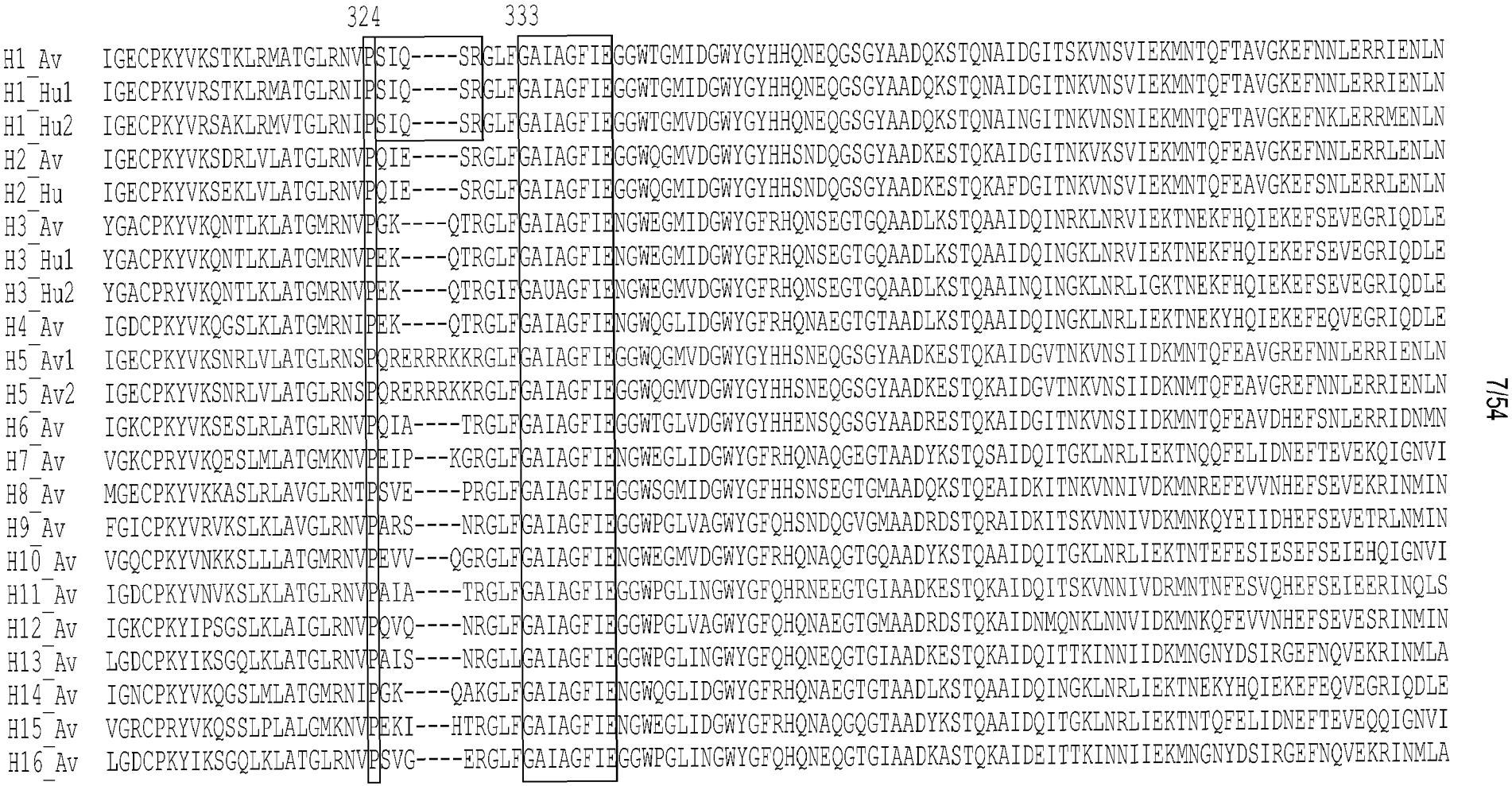

324
333
H1_Av IGECPKYVKSTKLRMATGLRNVPSIQ----SRGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAIDGITSKVNSVIEKMNTQFTAVGKEFNNLERRIENLN
H1_Hu1 IGECPKYVRSTKLRMATGLRNIPSIQ----SRGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAIDGITNKVNSVIEKMNTQFTAVGKEFNNLERRIENLN
H1_Hu2 IGECPKYVRSAKLRMVTGLRNIPSIQ----SRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNSNIEKMNTQFTAVGKEFNKLERRMENLN
H2_Av IGECPKYVKSDRLVLATGLRNVPQIE----SRGLFGAIAGFIEGGWQGMVDGWYGYHHSNDQGSGYAADKESTQKAIDGITNKVKSVIEKMNTQFEAVGKEFNNLERRLENLN
H2_Hu IGECPKYVKSEKLVLATGLRNVPQIE----SRGLFGAIAGFIEGGWQGMIDGWYGYHHSNDQGSGYAADKESTQKAFDGITNKVNSVIEKMNTQFEAVGKEFSNLERRLENLN
H3_Av YGACPKYVKQNTLKLATGMRNVPGK----QTRGLFGAIAGFIENGWEGMIDGWYGFRHQNSEGTGQAADLKSTQAAIDQINRKLNRVIEKTNEKFHQIEKEFSEVEGRIQDLE
H3_Hu1 YGACPKYVKQNTLKLATGMRNVPEK----QTRGLFGAIAGFIENGWEGMIDGWYGFRHQNSEGTGQAADLKSTQAAIDQINGKLNRVIEKTNEKFHQIEKEFSEVEGRIQDLE
H3_Hu2 YGACPRYVKQNTLKLATGMRNVPEK----QTRGIFGAUAGFIENGWEGMVDGWYGFRHQNSEGTGQAADLKSTQAAINQINGKLNRLIGKTNEKFHQIEKEFSEVEGRIQDLE
H4_Av IGDCPKYVKQGSLKLATGMRNIPEK----QTRGLFGAIAGFIENGWQGLIDGWYGFRHQNAEGTGTAADLKSTQAAIDQINGKLNRLIEKTNEKYHQIEKEFEQVEGRIQDLE
H5_Av1 IGECPKYVKSNRLVLATGLRNSPQRERRRKKRGLFGAIAGFIEGGWQGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFNNLERRIENLN
H5_Av2 IGECPKYVKSNRLVLATGLRNSPQRERRRKKRGLFGAIAGFIEGGWQGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKNMTQFEAVGREFNNLERRIENLN
H6_Av IGKCPKYVKSESLRLATGLRNVPQIA----TRGLFGAIAGFIEGGWTGLVDGWYGYHHENSQGSGYAADRESTQKAIDGITNKVNSIIDKMNTQFEAVDHEFSNLERRIDNMN
H7_Av VGKCPRYVKQESLMLATGMKNVPEIP---KGRGLFGAIAGFIENGWEGLIDGWYGFRHQNAQGEGTAADYKSTQSAIDQITGKLNRLIEKTNQQFELIDNEFTEVEKQIGNVI
H8_Av MGECPKYVKKASLRLAVGLRNTPSVE----PRGLFGAIAGFIEGGWSGMIDGWYGFHHSNSEGTGMAADQKSTQEAIDKITNKVNNIVDKMNREFEVVNHEFSEVEKRINMIN
H9_Av FGICPKYVRVKSLKLAVGLRNVPARS----NRGLFGAIAGFIEGGWPGLVAGWYGFQHSNDQGVGMAADRDSTQRAIDKITSKVNNIVDKMNKQYEIIDHEFSEVETRLNMIN
H10_Av VGQCPKYVNKKSLLLATGMRNVPEVV---QGRGLFGAIAGFIENGWEGMVDGWYGFRHQNAQGTGQAADYKSTQAAIDQITGKLNRLIEKTNTEFESIESEFSEIEHQIGNVI
H11_Av IGDCPKYVNVKSLKLATGLRNVPAIA----TRGLFGAIAGFIEGGWPGLINGWYGFQHRNEEGTGIAADKESTQKAIDQITSKVNNIVDRMNTNFESVQHEFSEIEERINQLS
H12_Av IGKCPKYIPSGSLKLAIGLRNVPQVQ----NRGLFGAIAGFIEGGWPGLVAGWYGFQHQNAEGTGMAADRDSTQKAIDNMQNKLNNVIDKMNKQFEVVNHEFSEVESRINMIN
H13_Av LGDCPKYIKSGQLKLATGLRNVPAIS----NRGLLGAIAGFIEGGWPGLINGWYGFQHQNEQGTGIAADKESTQKAIDQITTKINNIIDKMNGNYDSIRGEFNQVEKRINMLA
H14_Av IGNCPKYVKQGSLMLATGMRNIPGK----QAKGLFGAIAGFIENGWQGLIDGWYGFRHQNAEGTGTAADLKSTQAAIDQINGKLNRLIEKTNEKYHQIEKEFEQVEGRIQDLE
H15_Av VGRCPRYVKQSSLPLALGMKNVPEKI---HTRGLFGAIAGFIENGWEGLIDGWYGFRHQNAQGQGTAADYKSTQAAIDQITGKLNRLIEKTNTQFELIDNEFTEVEQQIGNVI
H16_Av LGDCPKYIKSGQLKLATGLRNVPSVG----ERGLFGAIAGFIEGGWPGLINGWYGFQHQNEQGTGIAADKASTQKAIDEITTKINNIIEKMNGNYDSIRGEFNQVEKRINMLA

FIG. 3B

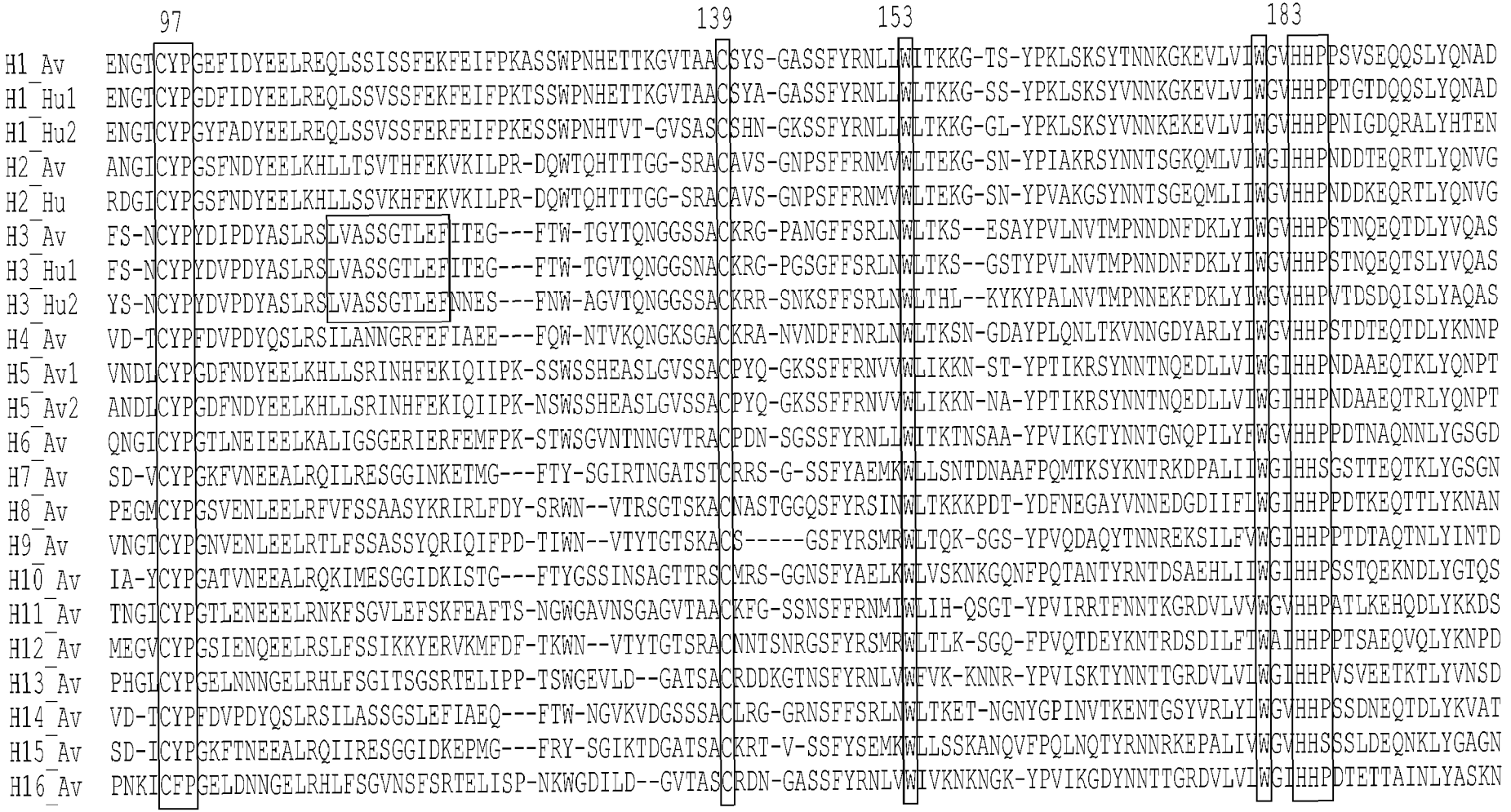

97
139
153
183
H1_Av ENGTCYPGEFIDYEELREQLSSISSFEKFEIFPKASSWPNHETTKGVTAACSYS-GASSFYRNLLWITKKG-TS-YPKLSKSYTNNKGKEVLVIWGVHHPPSVSEQQSLYQNAD
H1_Hu1 ENGTCYPGDFIDYEELREQLSSVSSFEKFEIFPKTSSWPNHETTKGVTAACSYA-GASSFYRNLLWLTKKG-SS-YPKLSKSYVNNKGKEVLVIWGVHHPPTGTDQQSLYQNAD
H1_Hu2 ENGTCYPGYFADYEELREQLSSVSSFERFEIFPKESSWPNHTVT-GVSASCSHN-GKSSFYRNLLWLTKKG-GL-YPKLSKSYVNNKEKEVLVIWGVHHPPNIGDQRALYHTEN
H2_Av ANGICYPGSFNDYEELKHLLTSVTHFEKVKILPR-DQWTQHTTTGG-SRACAVS-GNPSFFRNMVWLTEKG-SN-YPIAKRSYNNTSGKQMLVIWGIHHPNDDTEQRTLYQNVG
H2_Hu RDGICYPGSFNDYEELKHLLSSVKHFEKVKILPR-DQWTQHTTTGG-SRACAVS-GNPSFFRNMVWLTEKG-SN-YPVAKGSYNNTSGEQMLIIWGVHHPNDDKEQRTLYQNVG
H3_Av FS-NCYPYDIPDYASLRSLVASSGTLEFITEG---FTW-TGYTQNGGSSACKRG-PANGFFSRLNWLTKS--ESAYPVLNVTMPNNDNFDKLYIWGVHHPSTNQEQTDLYVQAS
H3_Hu1 FS-NCYPYDVPDYASLRSLVASSGTLEFITEG---FTW-TGVTQNGGSNACKRG-PGSGFFSRLNWLTKS--GSTYPVLNVTMPNNDNFDKLYIWGVHHPSTNQEQTSLYVQAS
H3_Hu2 YS-NCYPYDVPDYASLRSLVASSGTLEFNNES---FNW-AGVTQNGGSSACKRR-SNKSFFSRLNWLTHL--KYKYPALNVTMPNNEKFDKLYIWGVHHPVTDSDQISLYAQAS
H4_Av VD-TCYPFDVPDYQSLRSILANNGRFEFIAEE---FQW-NTVKQNGKSGACKRA-NVNDFFNRLNWLTKSN-GDAYPLQNLTKVNNGDYARLYIWGVHHPSTDTEQTDLYKNNP
H5_Av1 VNDLCYPGDFNDYEELKHLLSRINHFEKIQIIPK-SSWSSHEASLGVSSACPYQ-GKSSFFRNVVWLIKKN-ST-YPTIKRSYNNTNQEDLLVIWGIHHPNDAAEQTKLYQNPT
H5_Av2 ANDLCYPGDFNDYEELKHLLSRINHFEKIQIIPK-NSWSSHEASLGVSSACPYQ-GKSSFFRNVVWLIKKN-NA-YPTIKRSYNNTNQEDLLVIWGIHHPNDAAEQTRLYQNPT
H6_Av QNGICYPGTLNEIEELKALIGSGERIERFEMFPK-STWSGVNTNNGVTRACPDN-SGSSFYRNLLWITKTNSAA-YPVIKGTYNNTGNQPILYFWGVHHPPDTNAQNNLYGSGD
H7_Av SD-VCYPGKFVNEEALRQILRESGGINKETMG---FTY-SGIRTNGATSTCRRS-G-SSFYAEMKWLLSNTDNAAFPQMTKSYKNTRKDPALIIWGIHHSGSTTEQTKLYGSGN
H8_Av PEGMCYPGSVENLEELRFVFSSAASYKRIRLFDY-SRWN--VTRSGTSKACNASTGGQSFYRSINWLTKKKPDT-YDFNEGAYVNNEDGDIIFLWGIHHPPDTKEQTTLYKNAN
H9_Av VNGTCYPGNVENLEELRTLFSSASSYQRIQIFPD-TIWN--VTYTGTSKACS-----GSFYRSMRWLTQK-SGS-YPVQDAQYTNNREKSILFVWGIHHPPTDTAQTNLYINTD
H10_Av IA-YCYPGATVNEEALRQKIMESGGIDKISTG---FTYGSSINSAGTTRSCMRS-GGNSFYAELKWLVSKNKGQNFPQTANTYRNTDSAEHLIIWGIHHPSSTQEKNDLYGTQS
H11_Av TNGICYPGTLENEEELRNKFSGVLEFSKFEAFTS-NGWGAVNSGAGVTAACKFG-SSNSFFRNMIWLIH-QSGT-YPVIRRTFNNTKGRDVLVVWGVHHPATLKEHQDLYKKDS
H12_Av MEGVCYPGSIENQEELRSLFSSIKKYERVKMFDF-TKWN--VTYTGTSRACNNTSNRGSFYRSMRWLTLK-SGQ-FPVQTDEYKNTRDSDILFTWAIHHPPTSAEQVQLYKNPD
H13_Av PHGLCYPGELNNNGELRHLFSGITSGSRTELIPP-TSWGEVLD--GATSACRDDKGTNSFYRNLVWFVK-KNNR-YPVISKTYNNTTGRDVLVLWGIHHPVSVEETKTLYVNSD
H14_Av VD-TCYPFDVPDYQSLRSILASSGSLEFIAEQ---FTW-NGVKVDGSSSACLRG-GRNSFFSRLNWLTKET-NGNYGPINVTKENTGSYVRLYIWGVHHPSSDNEQTDLYKVAT
H15_Av SD-ICYPGKFTNEEALRQIIRESGGIDKEPMG---FRY-SGIKTDGATSACKRT-V-SSFYSEMKWLLSSKANQVFPQLNQTYRNNRKEPALIVWGVHHSSSLDEQNKLYGAGN
H16_Av PNKICFPGELDNNGELRHLFSGVNSFSRTELISP-NKWGDILD--GVTASCRDN-GASSFYRNLVWIVKNKNGK-YPVIKGDYNNTTGRDVLVLWGIHHPDTETTAINLYASKN

FIG. 4A

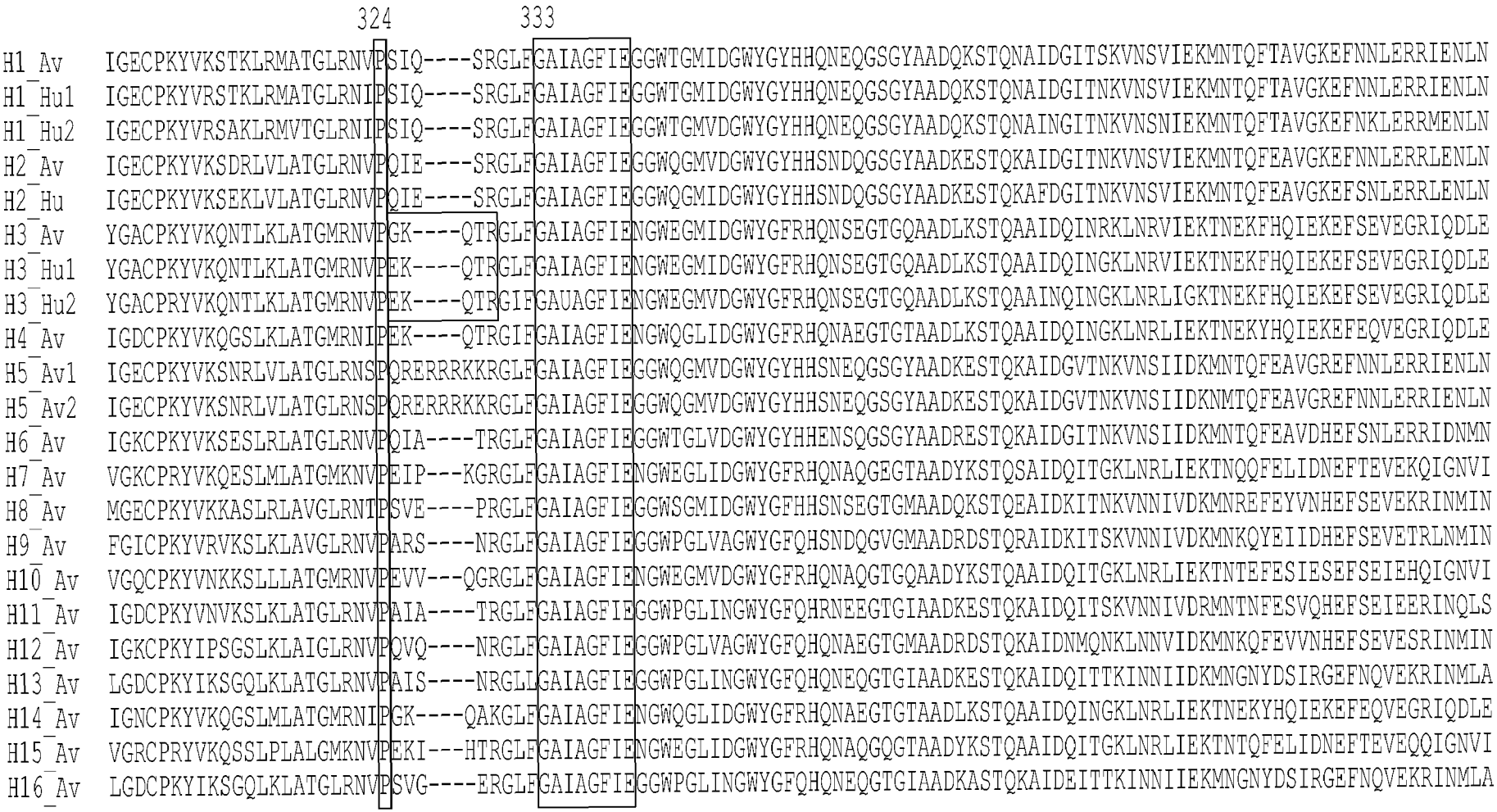

324
333
H1_Av IGECPKYVKSTKLRMATGLRNVPSIQ----SRGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAIDGITSKVNSVIEKMNTQFTAVGKEFNNLERRIENLN
H1_Hu1 IGECPKYVRSTKLRMATGLRNIPSIQ----SRGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAIDGITNKVNSVIEKMNTQFTAVGKEFNNLERRIENLN
H1_Hu2 IGECPKYVRSAKLRMVTGLRNIPSIQ----SRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNSNIEKMNTQFTAVGKEFNKLERRMENLN
H2_Av IGECPKYVKSDRLVLATGLRNVPQIE----SRGLFGAIAGFIEGGWQGMVDGWYGYHHSNDQGSGYAADKESTQKAIDGITNKVNSVIEKMNTQFEAVGKEFNNLERRLENLN
H2_Hu IGECPKYVKSEKLVLATGLRNVPQIE----SRGLFGAIAGFIEGGWQGMIDGWYGYHHSNDQGSGYAADKESTQKAFDGITNKVNSVIEKMNTQFEAVGKEFSNLERRLENLN
H3_Av YGACPKYVKQNTLKLATGMRNVPGK----QTRGLFGAIAGFIENGWEGMIDGWYGFRHQNSEGTGQAADLKSTQAAIDQINRKLNRVIEKTNEKFHQIEKEFSEVEGRIQDLE
H3_Hu1 YGACPKYVKQNTLKLATGMRNVPEK----QTRGLFGAIAGFIENGWEGMIDGWYGFRHQNSEGTGQAADLKSTQAAIDQINGKLNRVIEKTNEKFHQIEKEFSEVEGRIQDLE
H3_Hu2 YGACPRYVKQNTLKLATGMRNVPEK----QTRGIFGAUAGFIENGWEGMVDGWYGFRHQNSEGTGQAADLKSTQAAINQINGKLNRLIGKTNEKFHQIEKEFSEVEGRIQDLE
H4_Av IGDCPKYVKQGSLKLATGMRNIPEK----QTRGIFGAIAGFIENGWQGLIDGWYGFRHQNAEGTGTAADLKSTQAAIDQINGKLNRLIEKTNEKYHQIEKEFEQVEGRIQDLE
H5_Av1 IGECPKYVKSNRLVLATGLRNSPQRERRRKKRGLFGAIAGFIEGGWQGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFNNLERRIENLN
H5_Av2 IGECPKYVKSNRLVLATGLRNSPQRERRRKKRGLFGAIAGFIEGGWQGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKNMTQFEAVGREFNNLERRIENLN
H6_Av IGKCPKYVKSESLRLATGLRNVPQIA----TRGLFGAIAGFIEGGWTGLVDGWYGYHHENSQGSGYAADRESTQKAIDGITNKVNSIIDKMNTQFEAVDHEFSNLERRIDNMN
H7_Av VGKCPRYVKQESLMLATGMKNVPEIP---KGRGLFGAIAGFIENGWEGLIDGWYGFRHQNAQGEGTAADYKSTQSAIDQITGKLNRLIEKTNQQFELIDNEFTEVEKQIGNVI
H8_Av MGECPKYVKKASLRLAVGLRNTPSVE----PRGLFGAIAGFIEGGWSGMIDGWYGFHHSNSEGTGMAADQKSTQEAIDKITNKVNNIVDKMNREFEYVNHEFSEVEKRINMIN
H9_Av FGICPKYVRVKSLKLAVGLRNVPARS----NRGLFGAIAGFIEGGWPGLVAGWYGFQHSNDQGVGMAADRDSTQRAIDKITSKVNNIVDKMNKQYEIIDHEFSEVETRLNMIN
H10_Av VGQCPKYVNKKSLLLATGMRNVPEVV---QGRGLFGAIAGFIENGWEGMVDGWYGFRHQNAQGTGQAADYKSTQAAIDQITGKLNRLIEKTNTEFESIESEFSEIEHQIGNVI
H11_Av IGDCPKYVNVKSLKLATGLRNVPAIA----TRGLFGAIAGFIEGGWPGLINGWYGFQHRNEEGTGIAADKESTQKAIDQITSKVNNIVDRMNTNFESVQHEFSEIEERINQLS
H12_Av IGKCPKYIPSGSLKLAIGLRNVPQVQ----NRGLFGAIAGFIEGGWPGLVAGWYGFQHQNAEGTGMAADRDSTQKAIDNMQNKLNNVIDKMNKQFEVVNHEFSEVESRINMIN
H13_Av LGDCPKYIKSGQLKLATGLRNVPAIS----NRGLLGAIAGFIEGGWPGLINGWYGFQHQNEQGTGIAADKESTQKAIDQITTKINNIIDKMNGNYDSIRGEFNQVEKRINMLA
H14_Av IGNCPKYVKQGSLMLATGMRNIPGK----QAKGLFGAIAGFIENGWQGLIDGWYGFRHQNAEGTGTAADLKSTQAAIDQINGKLNRLIEKTNEKYHQIEKEFEQVEGRIQDLE
H15_Av VGRCPRYVKQSSLPLALGMKNVPEKI---HTRGLFGAIAGFIENGWEGLIDGWYGFRHQNAQGQGTAADYKSTQAAIDQITGKLNRLIEKTNTQFELIDNEFTEVEQQIGNVI
H16_Av LGDCPKYIKSGQLKLATGLRNVPSVG----ERGLFGAIAGFIEGGWPGLINGWYGFQHQNEQGTGIAADKASTQKAIDEITTKINNIIEKMNGNYDSIRGEFNQVEKRINMLA

FIG. 4B

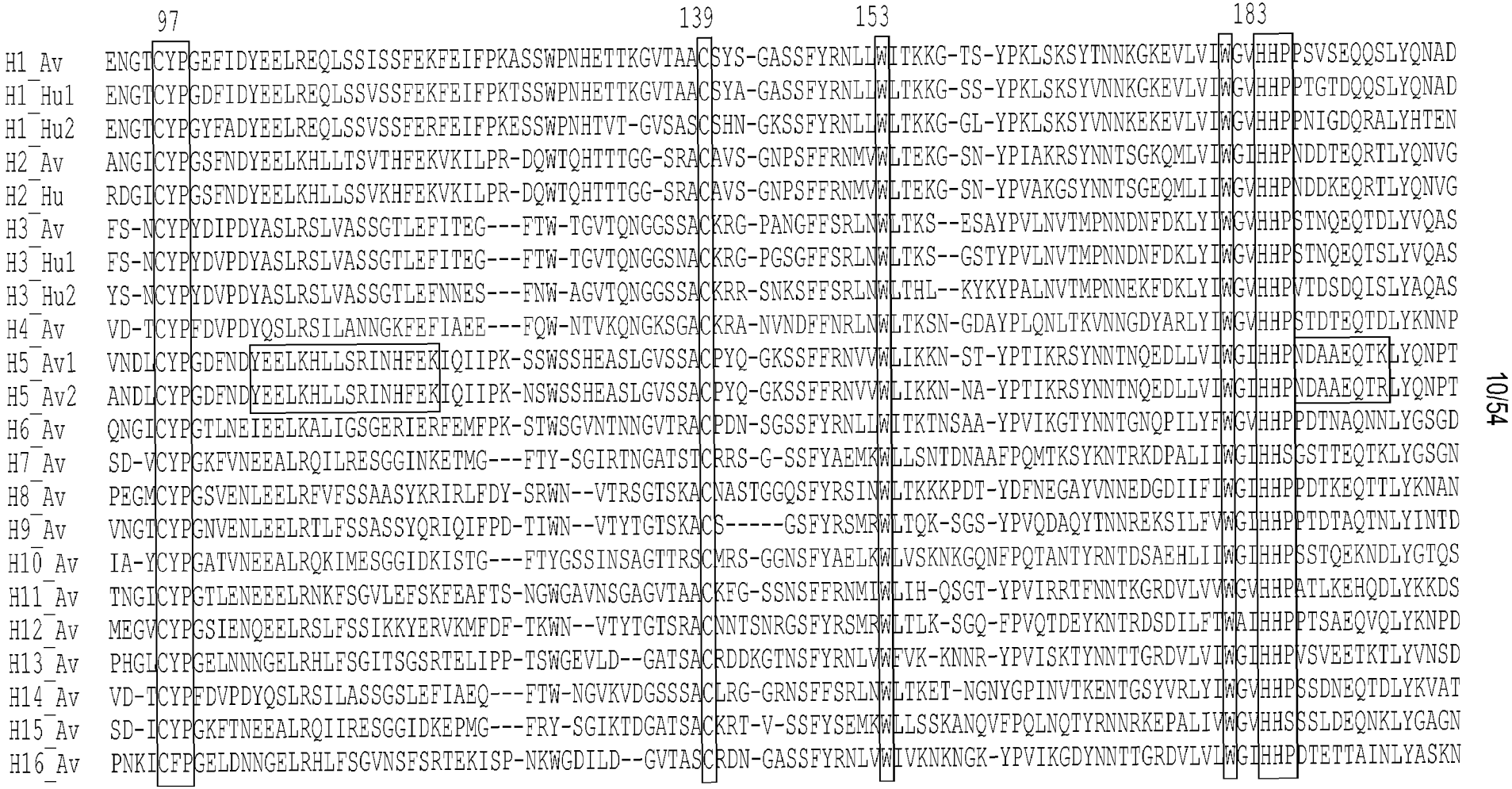

97
139
153
183
H1_Av ENGTCYPGEFIDYEELREQLSSISSFEKFEIFPKASSWPNHETTKGVTAACSYS-GASSFYRNLLWLTKKG-TS-YPKLSKSYTNNKGKEVLVLWGVHHPPSVSEQQSLYQNAD
H1_Hu1 ENGTCYPGDFIDYEELREQLSSVSSFEKFEIFPKTSSWPNHETTKGVTAACSYA-GASSFYRNLLWLTKKG-SS-YPKLSKSYVNNKGKEVLVLWGVHHPPTGTDQQSLYQNAD
H1_Hu2 ENGTCYPGYFADYEELREQLSSVSSFERFEIFPKESSWPNHTVT-GVSASCSHN-GKSSFYRNLLWLTKKG-GL-YPKLSKSYVNNKEKEVLVLWGVHHPPNIGDQRALYHTEN
H2_Av ANGICYPGSFNDYEELKHLLTSVTHFEKVKILPR-DQWTQHTTTGG-SRACAVS-GNPSFFRNMVWLTEKG-SN-YPIAKRSYNNTSGKQMLVLWGIHHPNDDTEQRTLYQNVG
H2_Hu RDGICYPGSFNDYEELKHLLSSVKHFEKVKILPR-DQWTQHTTTGG-SRACAVS-GNPSFFRNMVWLTEKG-SN-YPVAKGSYNNTSGEQMLILWGVHHPNDDKEQRTLYQNVG
H3_Av FS-NCYPYDIPDYASLRSLVASSGTLEFITEG---FTW-TGVTQNGGSSACKRG-PANGFFSRLNWLTKS--ESAYPVLNVTMPNNDNFDKLYIWGVHHPSTNQEQTDLYVQAS
H3_Hu1 FS-NCYPYDVPDYASLRSLVASSGTLEFITEG---FTW-TGVTQNGGSNACKRG-PGSGFFSRLNWLTKS--GSTYPVLNVTMPNNDNFDKLYIWGVHHPSTNQEQTSLYVQAS
H3_Hu2 YS-NCYPYDVPDYASLRSLVASSGTLEFNNES---FNW-AGVTQNGGSSACKRR-SNKSFFSRLNWLTHL--KYKYPALNVTMPNNEKFDKLYIWGVHHPVTDSDQISLYAQAS
H4_Av VD-TCYPFDVPDYQSLRSILANNGKFEFIAEE---FQW-NTVKQNGKSGACKRA-NVNDFFNRLNWLTKSN-GDAYPLQNLTKVNNGDYARLYIWGVHHPSTDTEQTDLYKNNP
H5_Av1 VNDLCYPGDFNDYEELKHLLSRINHFEKIQIIPK-SSWSSHEASLGVSSACPYQ-GKSSFFRNVVWLIKKN-ST-YPTIKRSYNNTNQEDLLVIWGIHHPNDAAEQTKLYQNPT
H5_Av2 ANDLCYPGDFNDYEELKHLLSRINHFEKIQIIPK-NSWSSHEASLGVSSACPYQ-GKSSFFRNVVWLIKKN-NA-YPTIKRSYNNTNQEDLLVIWGIHHPNDAAEQTRLYQNPT
H6_Av QNGICYPGTLNEIEELKALIGSGERIERFEMFPK-STWSGVNTNNGVTRACPDN-SGSSFYRNLLWITKTNSAA-YPVIKGTYNNTGNQPILYFWGVHHPPDTNAQNNLYGSGD
H7_Av SD-VCYPGKFVNEEALRQILRESGGINKETMG---FTY-SGIRTNGATSTCRRS-G-SSFYAEMKWLLSNTDNAAFPQMTKSYKNTRKDPALIIWGIHHSGSTTEQTKLYGSGN
H8_Av PEGMCYPGSVENLEELRFVFSSAASYKRIRLFDY-SRWN--VTRSGTSKACNASTGGQSFYRSINWLTKKKPDT-YDFNEGAYVNNEDGDIIFIWGIHHPPDTKEQTTLYKNAN
H9_Av VNGTCYPGNVENLEELRTLFSSASSYQRIQIFPD-TIWN--VTYTGTSKACS-----GSFYRSMRWLTQK-SGS-YPVQDAQYTNNREKSILFVWGIHHPPTDTAQTNLYINTD
H10_Av IA-YCYPGATVNEEALRQKIMESGGIDKISTG---FTYGSSINSAGTTRSCMRS-GGNSFYAELKWLVSKNKGQNFPQTANTYRNTDSAEHLIIWGIHHPSSTQEKNDLYGTQS
H11_Av TNGICYPGTLENEEELRNKFSGVLEFSKFEAFTS-NGWGAVNSGAGVTAACKFG-SSNSFFRNMIWLIH-QSGT-YPVIRRTFNNTKGRDVLVVWGVHHPATLKEHQDLYKKDS
H12_Av MEGVCYPGSIENQEELRSLFSSIKKYERVKMFDF-TKWN--VTYTGTSRACNNTSNRGSFYRSMRWLTLK-SGQ-FPVQTDEYKNTRDSDILFTWAIHHPPTSAEQVQLYKNPD
H13_Av PHGLCYPGELNNNGELRHLFSGITSGSRTELIPP-TSWGEVLD--GATSACRDDKGTNSFYRNLVWFVK-KNNR-YPVISKTYNNTTGRDVLVIWGIHHPVSVEETKTLYVNSD
H14_Av VD-TCYPFDVPDYQSLRSILASSGSLEFIAEQ---FTW-NGVKVDGSSSACLRG-GRNSFFSRLNWLTKET-NGNYGPINVTKENTGSYVRLYIWGVHHPSSDNEQTDLYKVAT
H15_Av SD-ICYPGKFTNEEALRQIIRESGGIDKEPMG---FRY-SGIKTDGATSACKRT-V-SSFYSEMKWLLSSKANQVFPQLNQTYRNNRKEPALIVWGVHHSSSLDEQNKLYGAGN
H16_Av PNKICFPGELDNNGELRHLFSGVNSFSRTEKISP-NKWGDILD--GVTASCRDN-GASSFYRNLVWIVKNKNGK-YPVIKGDYNNTTGRDVLVLWGIHHPDTETTAINLYASKN

FIG. 5A-1

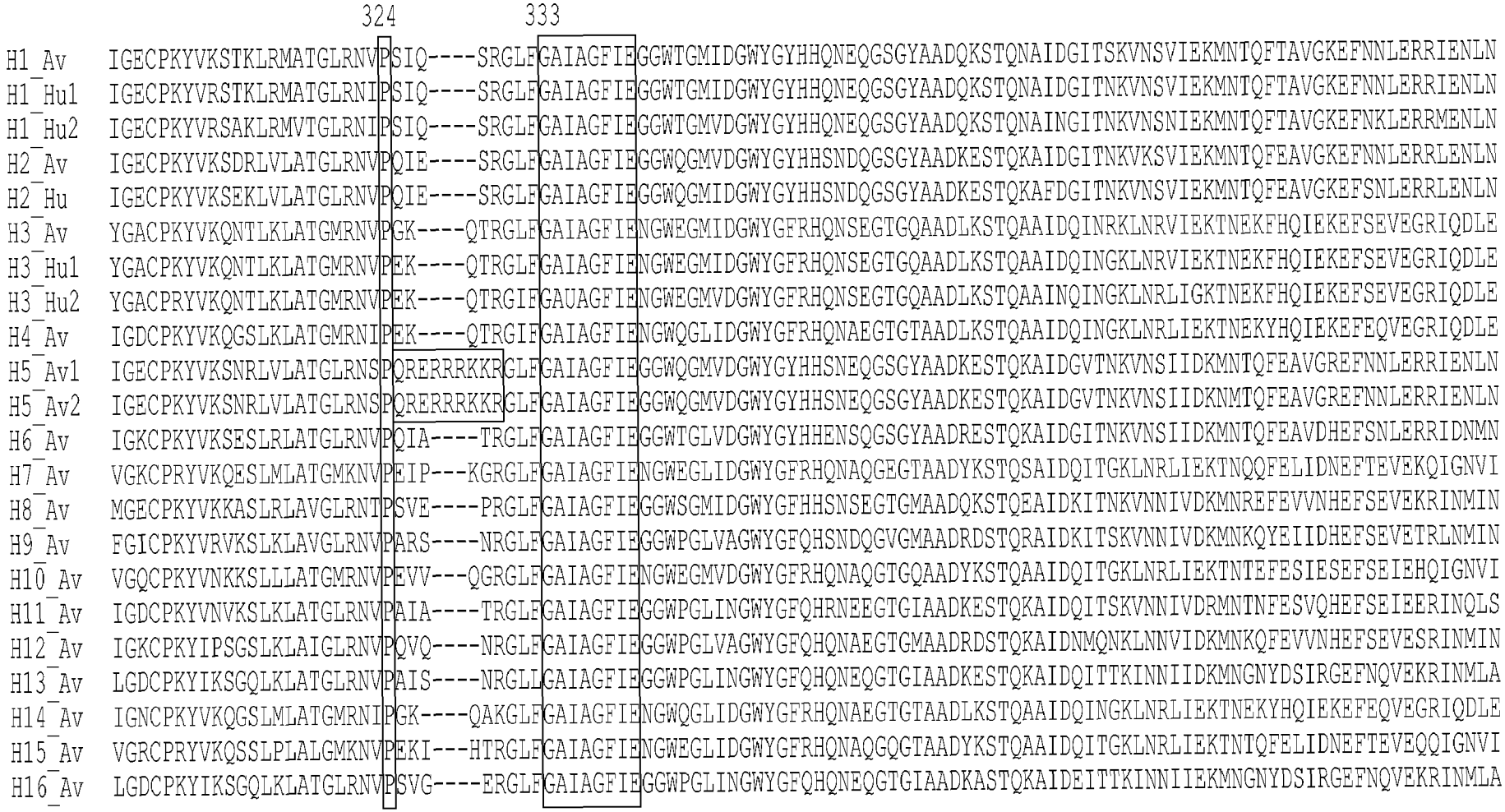

324
333
H1_Av IGECPKYVKSTKLRMATGLRNVPSIQ----SRGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAIDGITSKVNSVIEKMNTQFTAVGKEFNNLERRIENLN
H1_Hu1 IGECPKYVRSTKLRMATGLRNIPSIQ----SRGLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAIDGITNKVNSVIEKMNTQFTAVGKEFNNLERRIENLN
H1_Hu2 IGECPKYVRSAKLRMVTGLRNIPSIQ----SRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNSNIEKMNTQFTAVGKEFNKLERRMENLN
H2_Av IGECPKYVKSDRLVLATGLRNVPQIE----SRGLFGAIAGFIEGGWQGMVDGWYGYHHSNDQGSGYAADKESTQKAIDGITNKVKSVIEKMNTQFEAVGKEFNNLERRLENLN
H2_Hu IGECPKYVKSEKLVLATGLRNVPQIE----SRGLFGAIAGFIEGGWQGMIDGWYGYHHSNDQGSGYAADKESTQKAFDGITNKVNSVIEKMNTQFEAVGKEFSNLERRLENLN
H3_Av YGACPKYVKQNTLKLATGMRNVPGK----QTRGLFGAIAGFIENGWEGMIDGWYGFRHQNSEGTGQAADLKSTQAAIDQINRKLNRVIEKTNEKFHQIEKEFSEVEGRIQDLE
H3_Hu1 YGACPKYVKQNTLKLATGMRNVPEK----QTRGLFGAIAGFIENGWEGMIDGWYGFRHQNSEGTGQAADLKSTQAAIDQINGKLNRVIEKTNEKFHQIEKEFSEVEGRIQDLE
H3_Hu2 YGACPRYVKQNTLKLATGMRNVPEK----QTRGIFGAUAGFIENGWEGMVDGWYGFRHQNSEGTGQAADLKSTQAAINQINGKLNRLIGKTNEKFHQIEKEFSEVEGRIQDLE
H4_Av IGDCPKYVKQGSLKLATGMRNIPEK----QTRGIFGAIAGFIENGWQGLIDGWYGFRHQNAEGTGTAADLKSTQAAIDQINGKLNRLIEKTNEKYHQIEKEFEQVEGRIQDLE
H5_Av1 IGECPKYVKSNRLVLATGLRNSPQRERRRKKRGLFGAIAGFIEGGWQGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFNNLERRIENLN
H5_Av2 IGECPKYVKSNRLVLATGLRNSPQRERRRKKRGLFGAIAGFIEGGWQGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKNMTQFEAVGREFNNLERRIENLN
H6_Av IGKCPKYVKSESLRLATGLRNVPQIA----TRGLFGAIAGFIEGGWTGLVDGWYGYHHENSQGSGYAADRESTQKAIDGITNKVNSIIDKMNTQFEAVDHEFSNLERRIDNMN
H7_Av VGKCPRYVKQESLMLATGMKNVPEIP---KGRGLFGAIAGFIENGWEGLIDGWYGFRHQNAQGEGTAADYKSTQSAIDQITGKLNRLIEKTNQQFELIDNEFTEVEKQIGNVI
H8_Av MGECPKYVKKASLRLAVGLRNTPSVE----PRGLFGAIAGFIEGGWSGMIDGWYGFHHSNSEGTGMAADQKSTQEAIDKITNKVNNIVDKMNREFEVVNHEFSEVEKRINMIN
H9_Av FGICPKYVRVKSLKLAVGLRNVPARS----NRGLFGAIAGFIEGGWPGLVAGWYGFQHSNDQGVGMAADRDSTQRAIDKITSKVNNIVDKMNKQYEIIDHEFSEVETRLNMIN
H10_Av VGQCPKYVNKKSLLLATGMRNVPEVV---QGRGLFGAIAGFIENGWEGMVDGWYGFRHQNAQGTGQAADYKSTQAAIDQITGKLNRLIEKTNTEFESIESEFSEIEHQIGNVI
H11_Av IGDCPKYVNVKSLKLATGLRNVPAIA----TRGLFGAIAGFIEGGWPGLINGWYGFQHRNEEGTGIAADKESTQKAIDQITSKVNNIVDRMNTNFESVQHEFSEIEERINQLS
H12_Av IGKCPKYIPSGSLKLAIGLRNVPQVQ----NRGLFGAIAGFIEGGWPGLVAGWYGFQHQNAEGTGMAADRDSTQKAIDNMQNKLNNVIDKMNKQFEVVNHEFSEVESRINMIN
H13_Av LGDCPKYIKSGQLKLATGLRNVPAIS----NRGLLGAIAGFIEGGWPGLINGWYGFQHQNEQGTGIAADKESTQKAIDQITTKINNIIDKMNGNYDSIRGEFNQVEKRINMLA
H14_Av IGNCPKYVKQGSLMLATGMRNIPGK----QAKGLFGAIAGFIENGWQGLIDGWYGFRHQNAEGTGTAADLKSTQAAIDQINGKLNRLIEKTNEKYHQIEKEFEQVEGRIQDLE
H15_Av VGRCPRYVKQSSLPLALGMKNVPEKI---HTRGLFGAIAGFIENGWEGLIDGWYGFRHQNAQGQGTAADYKSTQAAIDQITGKLNRLIEKTNTQFELIDNEFTEVEQQIGNVI
H16_Av LGDCPKYIKSGQLKLATGLRNVPSVG----ERGLFGAIAGFIEGGWPGLINGWYGFQHQNEQGTGIAADKASTQKAIDEITTKINNIIEKMNGNYDSIRGEFNQVEKRINMLA

FIG. 5A-2

```
POSITIONS FROM 1 TILL 60
CONSENSUS      MEKIVLLLAIVSLVKSDQICIGYHANNSTEQVDTIMEKNVTVTHAQDILEKTHNGKLCDL
AAL59142       ............................................................
AAZ29963       .......F....................................................
ABA70758       .......F....................................................
ABB87042       ..R..IA...I.I..G.............K.....................E......S.
ABD14810       ............................................................
ABD46740       ----........................................................
ABD85144       -------F....................................................
ABE97569       ............................................................

POSITIONS FROM 61 TILL 120
CONSENSUS      DGVKPLILRDCSVAGWLLGNPMCDEFINVPEWSYIVEKANPANDLCYPGDFNDYEELKEL
AAL59142       ......................................S.D...................
AAZ29963       ........................................V...................
ABA70758       ........................................V...................
ABB87042       K..R...K..................L...........D..I.G................
ABD14810       .......K.........................................I..........
ABD46740       ..........................L...........I..........N..........
ABD85144       ..........................L...........I..........N..........
ABE97569       ...............................................N..........
```

FIG. 5B-1

```
POSITIONS FROM 121 TILL 180
CONSENSUS       LSRINHFEKIQIIPKSSWSDHEASSGVSSACPYQGKSSFFRNVVWLIKKNSAYPTIKRSY
AAL59142        ..................N.............H.........................
AAZ29963        ..................S....L.........................T........
ABA70758        ..................S....L..........P.............T........
ABB87042        M.ST.........R....N.D..........N.R.............N.......T.
ABD14810        ..................................................I..
ABD46740        ...................................R.............DN.........
ABD85144        ...................................R.............DN.........
ABE97569        ..................................L..........................

POSITIONS FROM 181 TILL 240
CONSENSUS       NNTNQEDLLVLWGIHHPNDAAEQTKLYQNPTTYISVGTSTLNQRLVPKIATRSKVNGQSG
AAL59142        ................................................................
AAZ29963        ...............................................R............
ABA70758        ........V.......................................R............
ABB87042        ....I....I..................SN..V..........SI.E....P.......
ABD14810        ........................R.......................................
ABD46740        ........................R.......................................
ABD85144        ........................R.......................................
ABE97569        ........................R.......................................
```

FIG. 5B-2

```
POSITIONS FROM 241 TILL 300
CONSENSUS      RMEFFWTILKPNDAINFESNGNFIAPEYAYKIVKKGDSTIMKSBLEYGNCNTKCQTPMGA
AAL59142       ....................................A.....................
AAZ29963       ............................................................
ABA70758       ............................................................
ABB87042       ...............S.......................A...........D......V..
ABD14810       ..............V..................................D...........
ABD46740       ............................N.............................I..
ABD85144       ..........S................N.............................I..
ABE97569       ....................................A.....................

POSITIONS FROM 301 TILL 360
CONSENSUS      INSSMPFHNIHPLTIGECPKYVKSNRLVLATGLRNSPQRERRRKKRGLFGAIAGFIEGGW
AAL59142       ..................................T....G...................
AAZ29963       ............................................................
ABA70758       ........................................K...................
ABB87042       .........V..............DK.........V....----T...............
ABD14810       ........................................-...................
ABD46740       ....................................G.....................
ABD85144       ....................................G.....................
ABE97569       ............................................................
```

FIG. 5B-3

```
POSITIONS FROM 361 TILL 420
CONSENSUS       QGMVDGWYGYHHSNEQGSGYAADKESTQKAIDGVTNKVNSIIDKMNTQFEAVGREFNNLE
AAL59142        ..........................................................K..
AAZ29963        ..........................................................
ABA70758        ..............K...........................................
ABB87042        ................................I................T..K......
ABD14810        ..........................................................
ABD46740        ..........................................................
ABD85144        ..........................................................
ABE97569        ..........................................................

POSITIONS FROM 421 TILL 480
CONSENSUS       RRIENLNKKMEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRLQLRDNAKELG
AAL59142        ..........................................................
AAZ29963        ..........................................................
ABA70758        ..........................................................
ABB87042        ..........................................................
ABD14810        ..........................................................
ABD46740        ..........................................................
ABD85144        ..........................................................
ABE97569        ..........................................................
```

FIG. 5B-4

```
POSITIONS FROM 481 TILL 540
CONSENSUS       NGCFEFYHKCDNECMESVRNGTYDYPQYSEEARLKREEISGVKLESIGTYQILSIYSTVA
AAL59142        .................K.............N..........M.............
AAZ29963        ................................................I...........
ABA70758        ................................................I...........
ABB87042        ..............................S..N....D......M.............
ABD14810        ..................................................................
ABD46740        ........R.........................................................
ABD85144        ........R.........................................................
ABE97569        .................I....N.......................I...........

POSITIONS FROM 541 TILL 568
CONSENSUS       SSLALAIMVAGLSLWMCSNGSLQCRICI
AAL59142        ............................
AAZ29963        .............------------
ABA70758        ............................
ABB87042        ............F...............
ABD14810        ............................
ABD46740        .................--------
ABD85144        ............F........VTM----
ABE97569        ........M....---------------
```

FIG. 5B-5

AAL59142: 568 Avian 4 (HA) H5N1 Hong Kong 2000 Influenza A virus (A/Goose/Hong Kong/385.3.2000(H5N1))

AAZ29963: 556 Avian 4 (HA) H5N1 Thailand 2004 Influenza A virus (A/Ostrich/Samut Prakan/Thailand/CU-19/04(H5N1))

ABA70758: 568 Avian 4 (HA) H5N1 Belgium 2004 Influenza A virus (A/crested eagle/Belgium/01/2004(H5N1))

ABB87042: 564 Avian 4 (HA) H5N2 Canada 1976/08/12 Influenza A virus (A/mallard duck/ALB/57/1976(H5N2))

ABD14810: 567 Avian 4 (HA) H5N1 China 2004 Influenza A virus (A/duck/Guangxi/13/2004(H5N1))

ABD46740: 556 Avian 4 (HA) H5N1 Nigeria 2006/01/17 Influenza A virus (A/chicken/Nigeria/641/2006(H5N1))

ABD85144: 557 Avian 4 (HA) H5N1 Egypt 2006 Influenza A virus (A/chicken/Egypt/960N3-004/2006(H5N1))

ABE97569: 553 Avian 4 (HA) H5N1 Indonesia 2004 Influenza A virus (A/turkey/Kedaton/BPPV3/2004(H5N1))

FIG. 5B-6

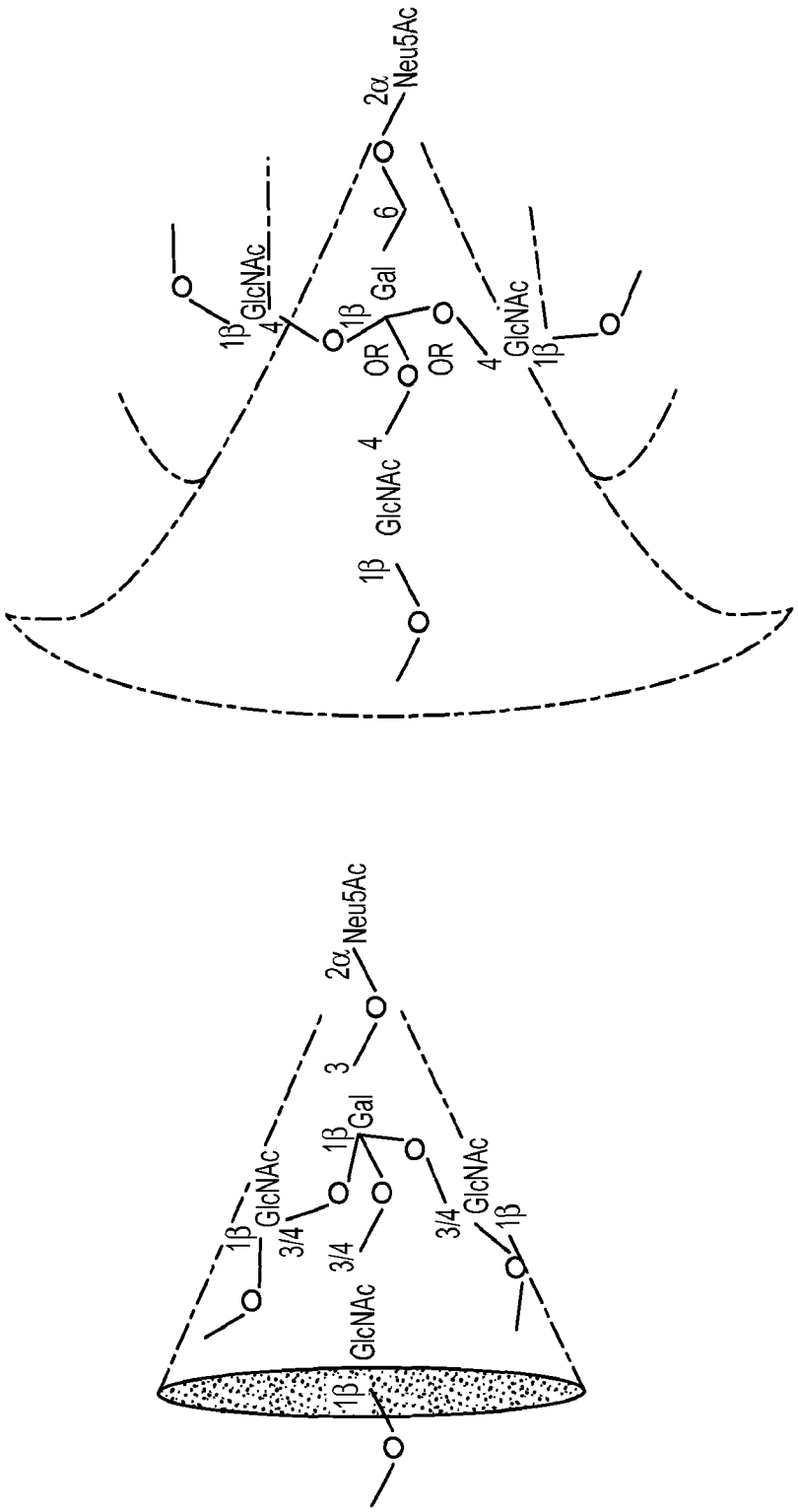

2α Neu5Ac
6
1β GlcNAc 4
1β Gal
OR
OR
4 GlcNAc 1β
4 GlcNAc 1β
2α Neu5Ac
3
1β GlcNAc 3/4
1β Gal
3/4 GlcNAc 1β
3/4 GlcNAc 1β

FIG. 6A

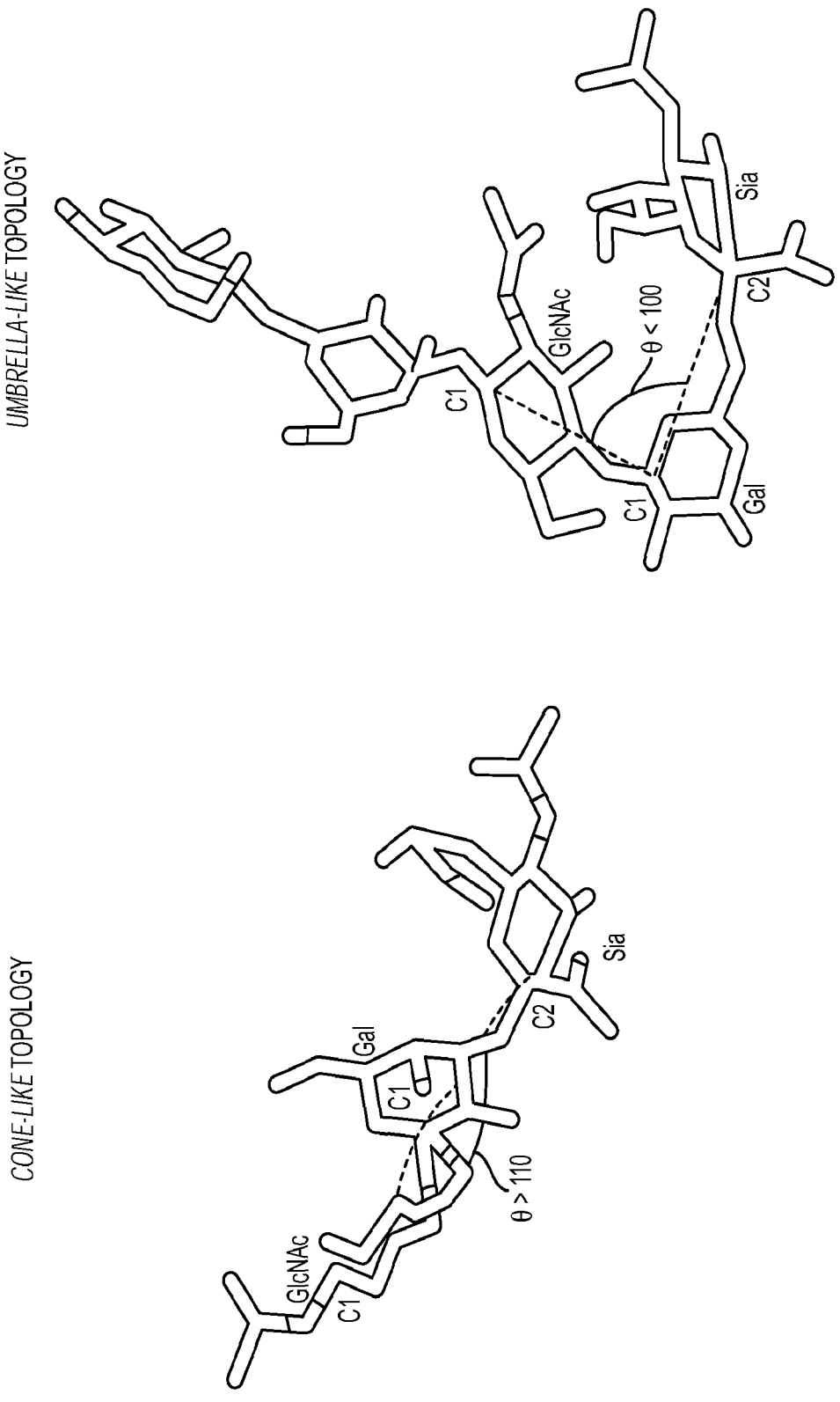

UMBRELLA-LIKE TOPOLOGY
CONE-LIKE TOPOLOGY
GlcNAc
C1
Gal
C1
C2
Sia
θ > 110
GlcNAc
C1
θ < 100
C2
Sia
C1
Gal

FIG. 6B-1

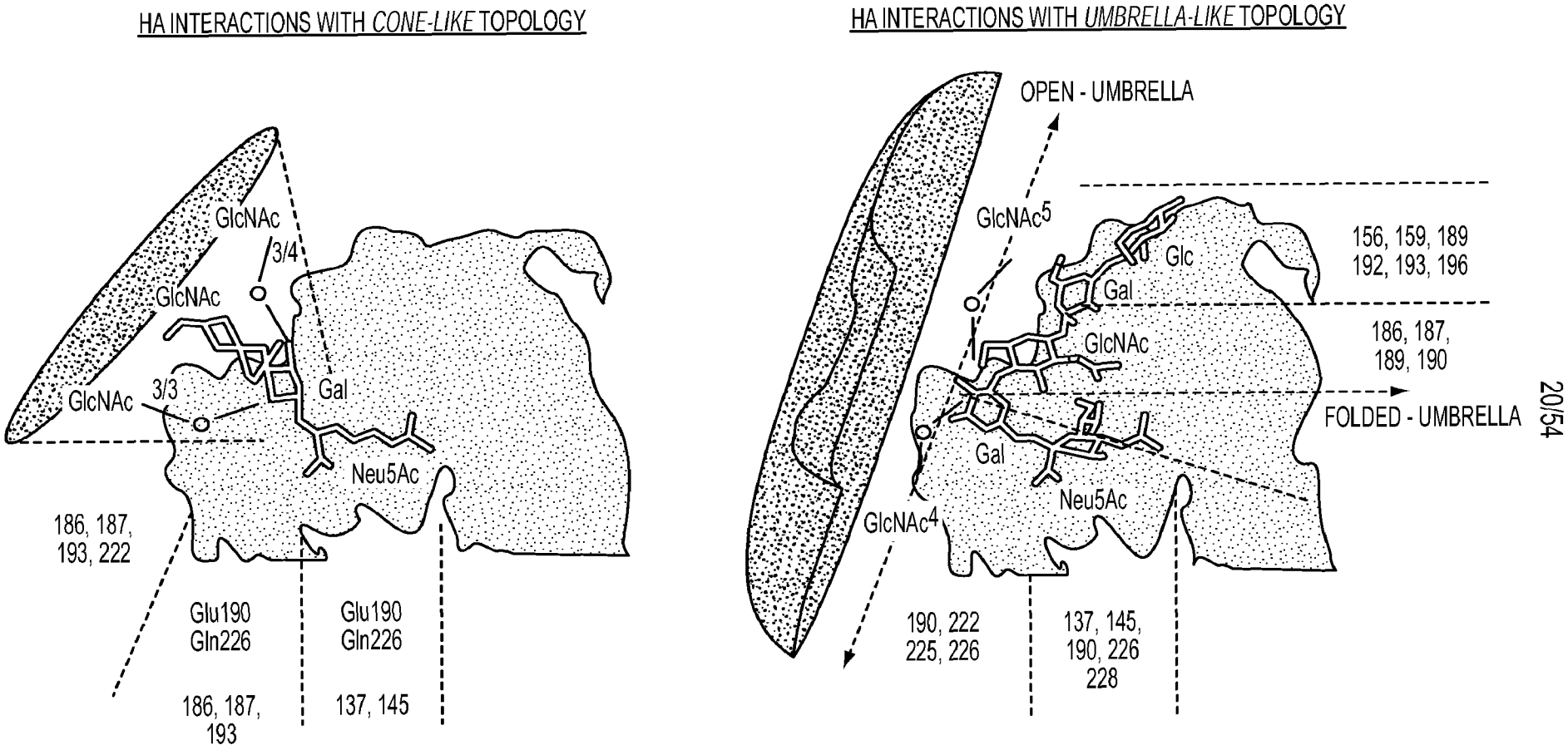

HA INTERACTIONS WITH *CONE-LIKE* TOPOLOGY
GlcNAc
3/4
GlcNAc
GlcNAc
3/3
Gal
Neu5Ac
186, 187,
193, 222
Glu190
Gln226
186, 187,
193
Glu190
Gln226
137, 145
HA INTERACTIONS WITH *UMBRELLA-LIKE* TOPOLOGY
OPEN - UMBRELLA
GlcNAc5
Glc
Gal
GlcNAc
Gal
Neu5Ac
GlcNAc4
156, 159, 189
192, 193, 196
186, 187,
189, 190
FOLDED - UMBRELLA
190, 222
225, 226
137, 145,
190, 226
228

FIG. 6B-2

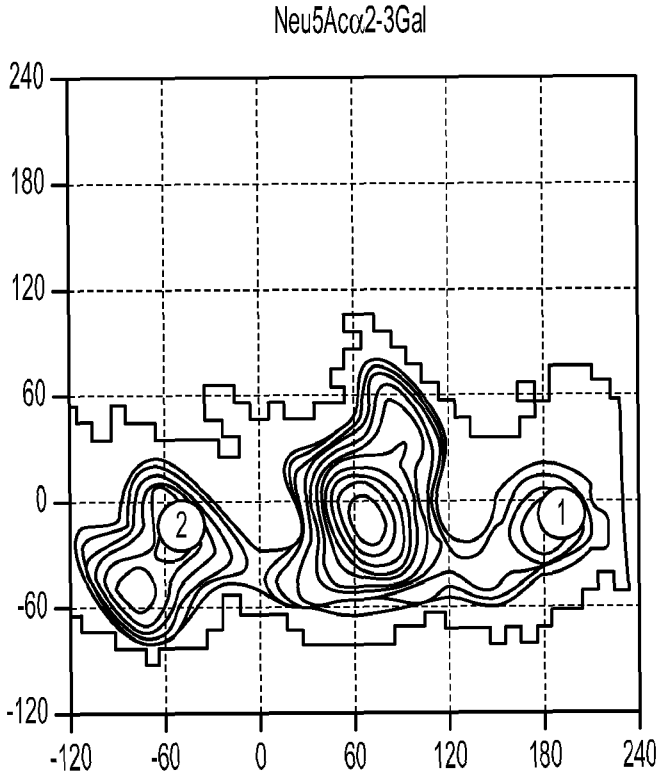


FIG. 6C-A

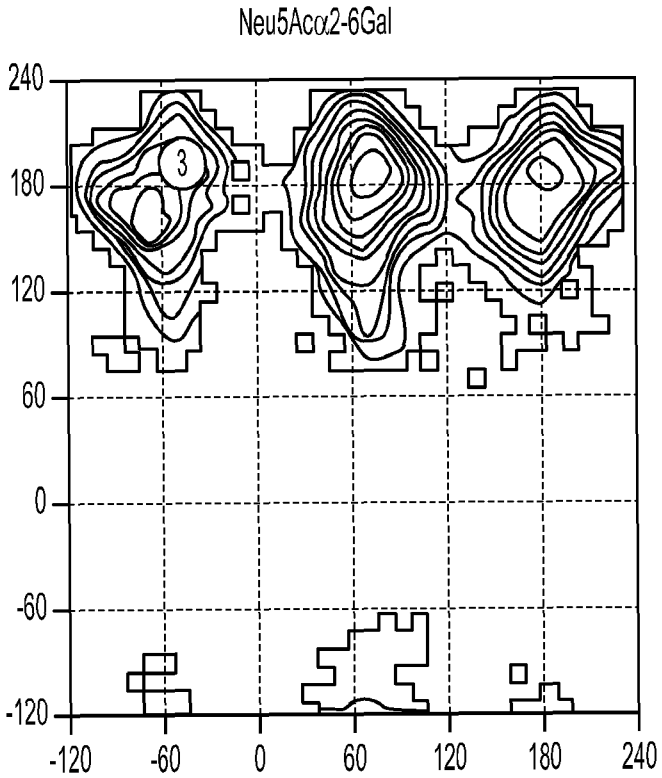


FIG. 6C-B

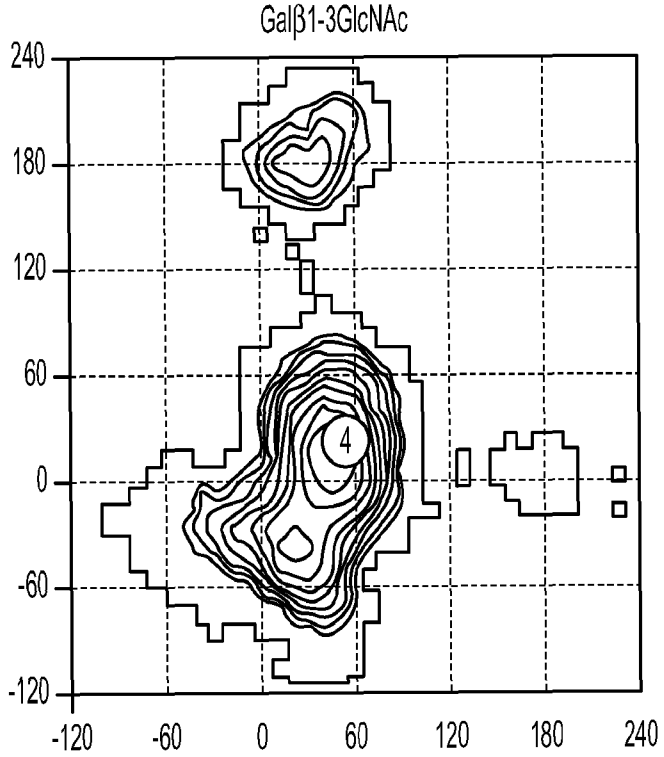


FIG. 6C-C

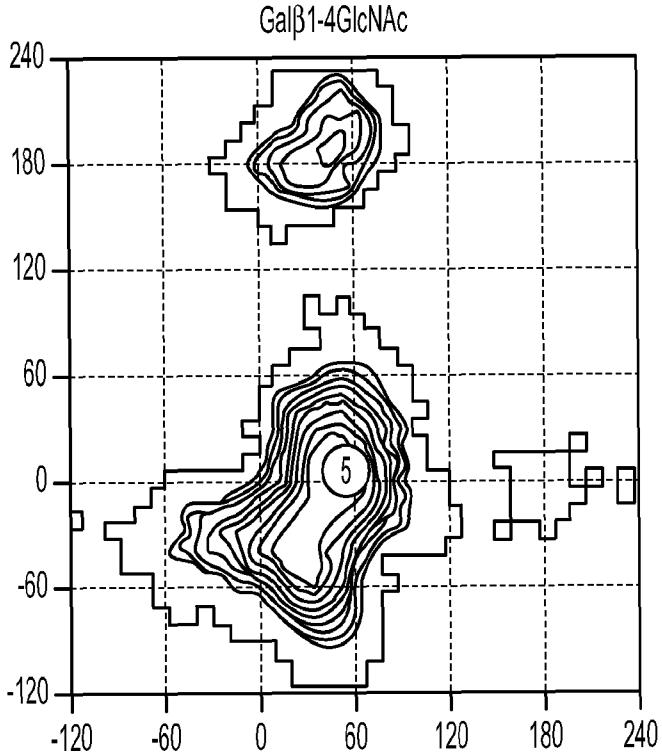


FIG. 6C-D

CONFORMATIONAL SAMPLING OF α2-3 LINKAGE

CONE-LIKE	100%
UMBRELLA-LIKE	0%

FIG. 6C-E

CONFORMATIONAL SAMPLING OF α2-6 LINKAGE

	UMBRELLA-LIKE (%) : *CONE*-LIKE (%)
$\omega = -60^{\circ}$	60 : 20
$\omega = +60^{\circ}$	10 : 40
$\omega = 180^{\circ}$	30 : 30

FIG. 6C-F

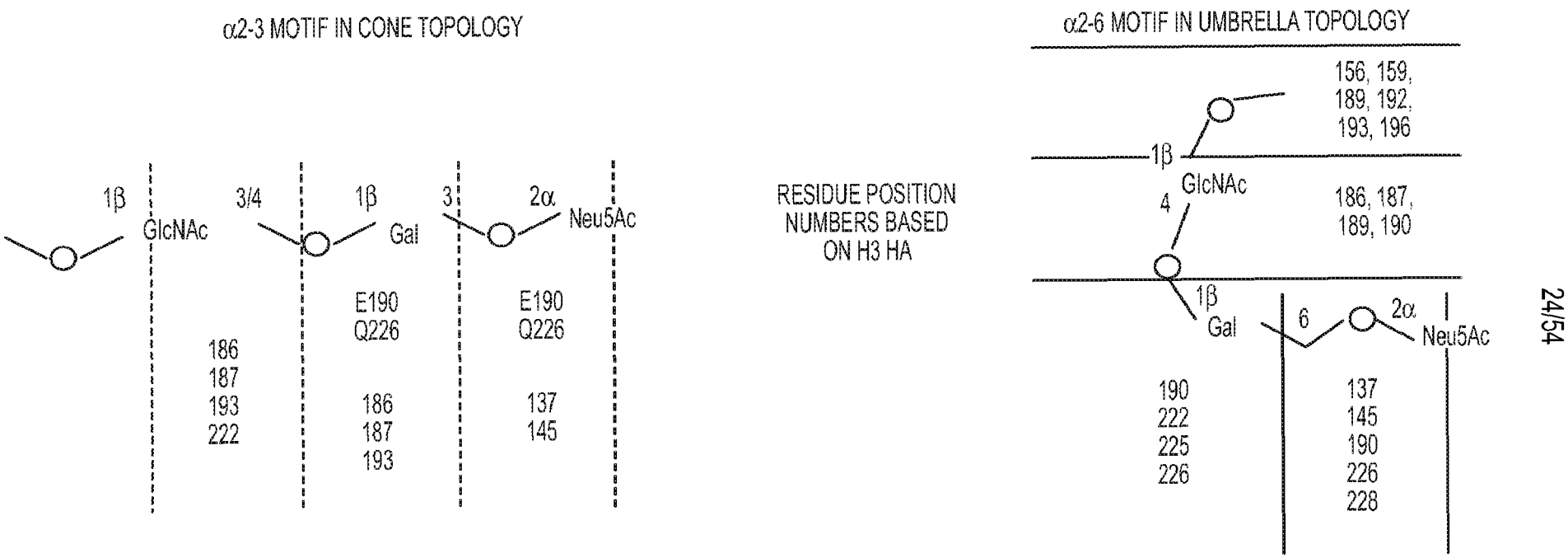

α2-3 MOTIF IN CONE TOPOLOGY
1β
GlcNAc
3/4
1β
Gal
3
2α
Neu5Ac
E190
Q226
186
187
193
222
E190
Q226
186
187
193
137
145
RESIDUE POSITION NUMBERS BASED ON H3 HA
α2-6 MOTIF IN UMBRELLA TOPOLOGY
156, 159,
189, 192,
193, 196
1β
GlcNAc
4
186, 187,
189, 190
1β
Gal
6
2α
Neu5Ac
190
222
225
226
137
145
190
226
228

FIG. 7

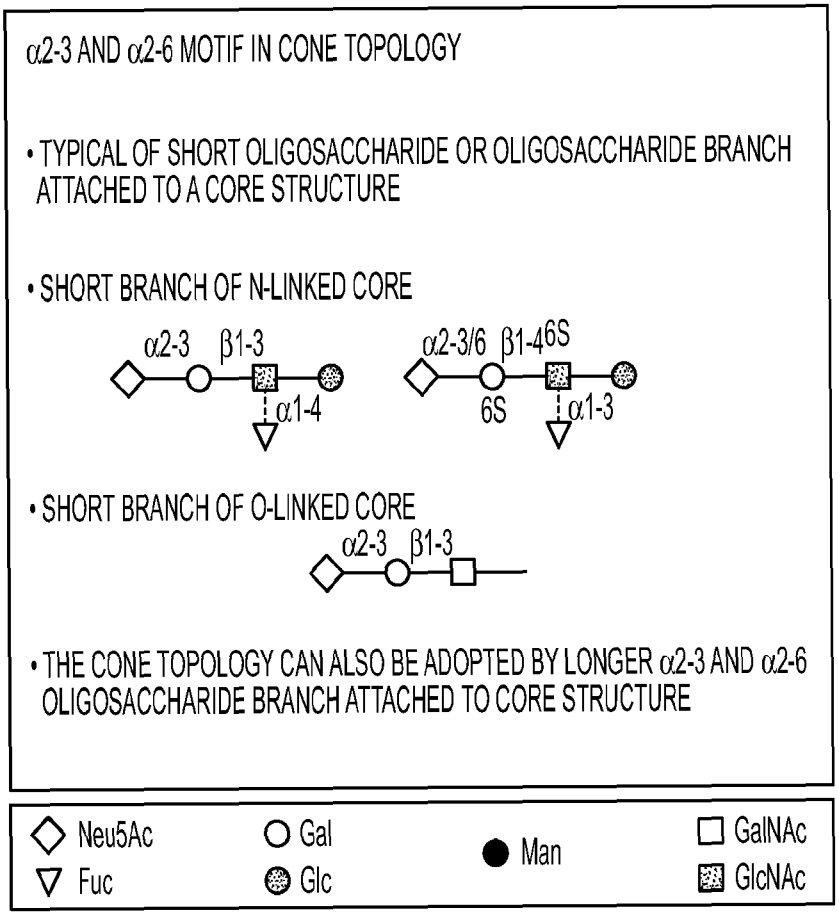

α2-3 AND α2-6 MOTIF IN CONE TOPOLOGY
• TYPICAL OF SHORT OLIGOSACCHARIDE OR OLIGOSACCHARIDE BRANCH ATTACHED TO A CORE STRUCTURE
• SHORT BRANCH OF N-LINKED CORE
α2-3
β1-3
α1-4
α2-3/6
β1-4
6S
6S
α1-3
• SHORT BRANCH OF O-LINKED CORE
α2-3
β1-3
• THE CONE TOPOLOGY CAN ALSO BE ADOPTED BY LONGER α2-3 AND α2-6 OLIGOSACCHARIDE BRANCH ATTACHED TO CORE STRUCTURE
Neu5Ac
Fuc
Gal
Glc
Man
GalNAc
GlcNAc
DOTTED GRAY LINES, 4S AND 6S INDICATE POTENTIAL SITES FOR FUCOSYLATION AND SULFATION MODIFICATIONS

FIG. 8

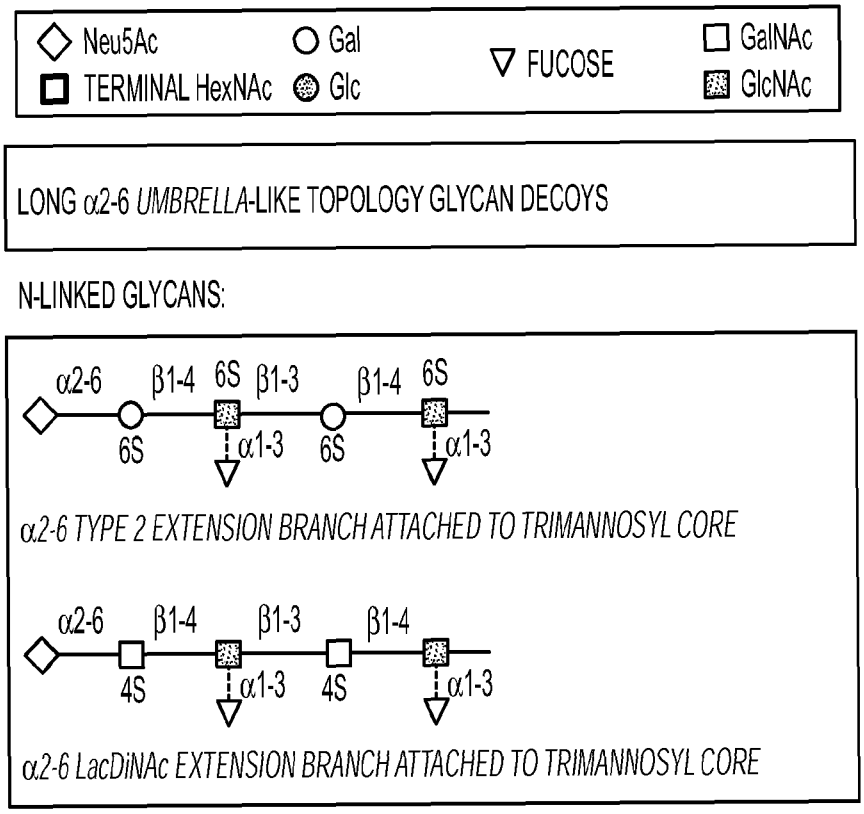
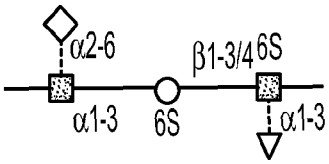

Neu5Ac
Gal
FUCOSE
GalNAc
TERMINAL HexNAc
Glc
GlcNAc
LONG α2-6 UMBRELLA-LIKE TOPOLOGY GLYCAN DECOYS
N-LINKED GLYCANS:
α2-6
β1-4
6S
β1-3
β1-4
6S
6S
α1-3
6S
α1-3
α2-6 TYPE 2 EXTENSION BRANCH ATTACHED TO TRIMANNOSYL CORE
α2-6
β1-4
β1-3
β1-4
4S
α1-3
4S
α1-3
α2-6 LacDiNAc EXTENSION BRANCH ATTACHED TO TRIMANNOSYL CORE

FIG. 9A-1

LONG α2-6 *UMBRELLA*-LIKE TOPOLOGY GLYCANS THAT ARE NOT DECOYS

N-LINKED GLYCANS:

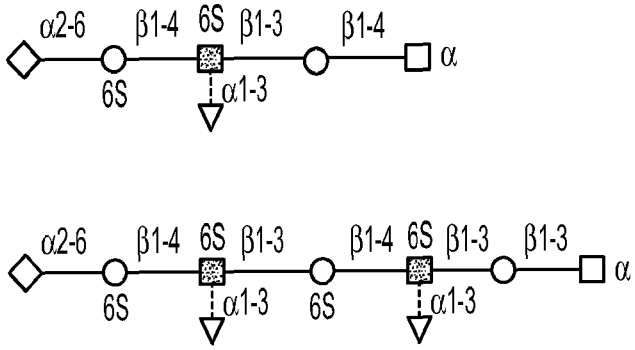


α2-6 LINKAGE ON GlcNAc OF TYPE 1/TYPE 2 EXTENSION

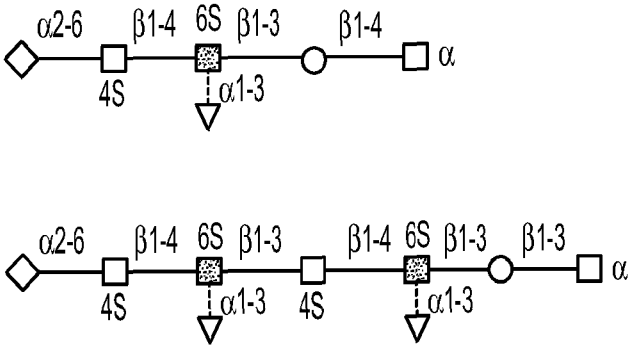
FIG. 9A-2

LONG α2-6 *UMBRELLA*-LIKE TOPOLOGY GLYCAN DECOYS

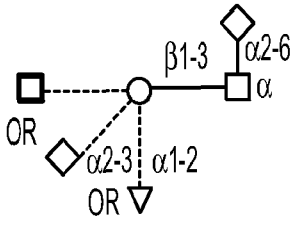
O-LINKED GLYCANS:



α2-6 TYPE 2 EXTENSION BRANCH IN A CORE 1 TYPE STRUCTURE



α2-6 LacDiNAc EXTENSION BRANCH IN A CORE 1 TYPE STRUCTURE

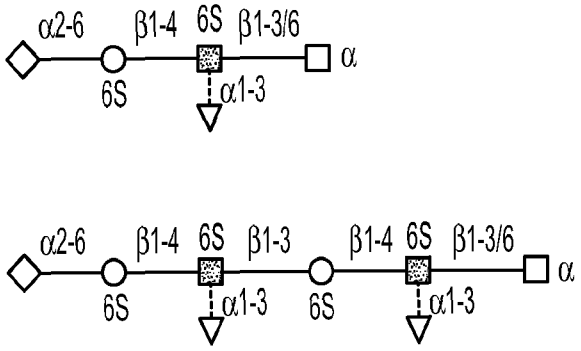


α2-6 ATTACHED TO CORE GalNAc IN CORE 1 TYPE STRUCTURE

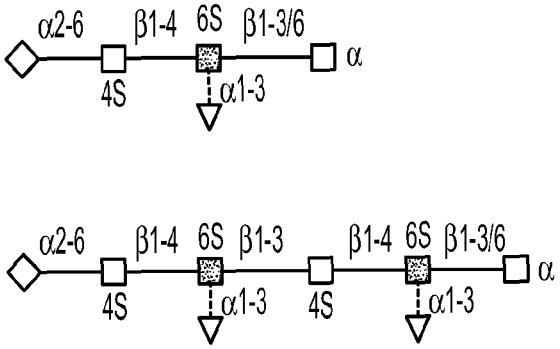
FIG. 9A-3

LONG α2-6 *UMBRELLA*-LIKE TOPOLOGY GLYCAN DECOYS

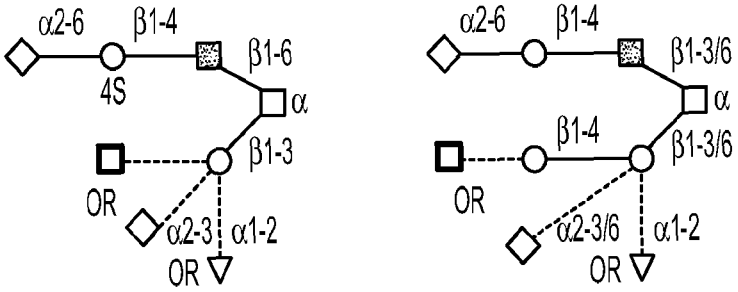
O-LINKED GLYCANS:



α2-6 TYPE 2 EXTENSION BRANCH IN A CORE 2 OR 3 OR 4 TYPE STRUCTURE



α2-6 LacDiNAc EXTENSION BRANCH IN A CORE 2 OR 3 OR 4 TYPE STRUCTURE

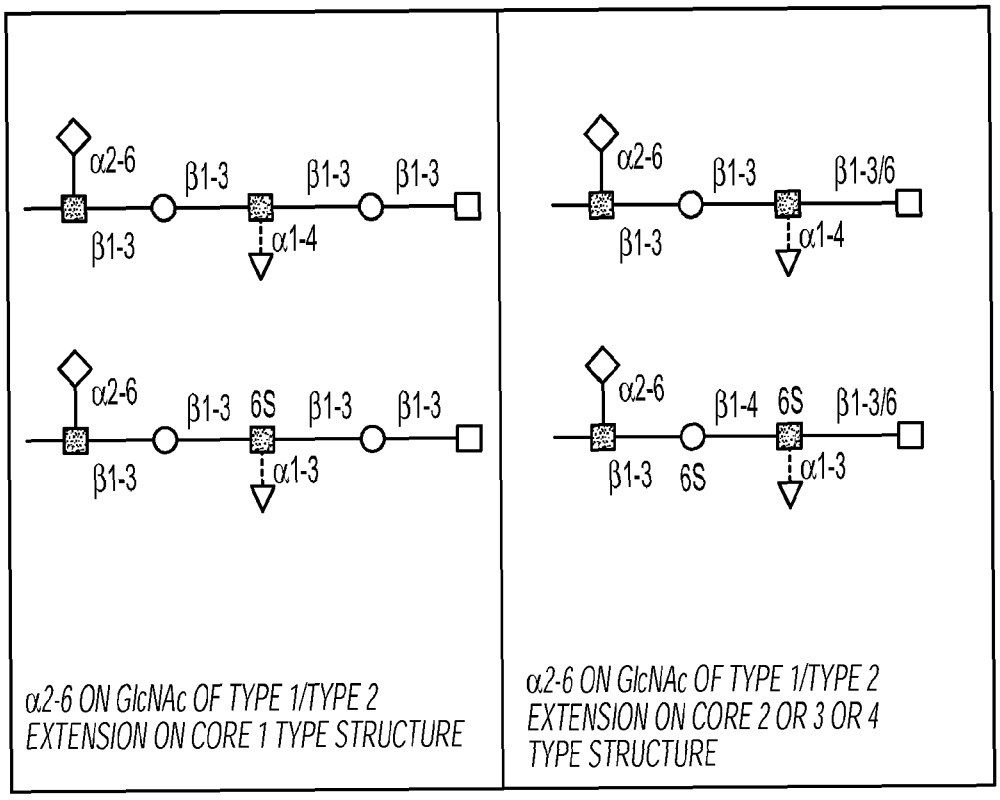


α2-6 ATTACHED TO BRANCHED CORE 2 AND CORE 4 STRUCTURES

FIG. 9A-4

LONG α2-6 *UMBRELLA*-LIKE TOPOLOGY GLYCAN THAT ARE NOT DECOYS

O-LINKED GLYCANS:



α2-6 ON GlcNAc OF TYPE 1/TYPE 2 EXTENSION ON CORE 1 TYPE STRUCTURE

α2-6 ON GlcNAc OF TYPE 1/TYPE 2 EXTENSION ON CORE 2 OR 3 OR 4 TYPE STRUCTURE

FIG. 9A-5

LONG α2-6 *UMBRELLA*-LIKE TOPOLOGY GLYCAN DECOYS

GLYCOLIPIDS:

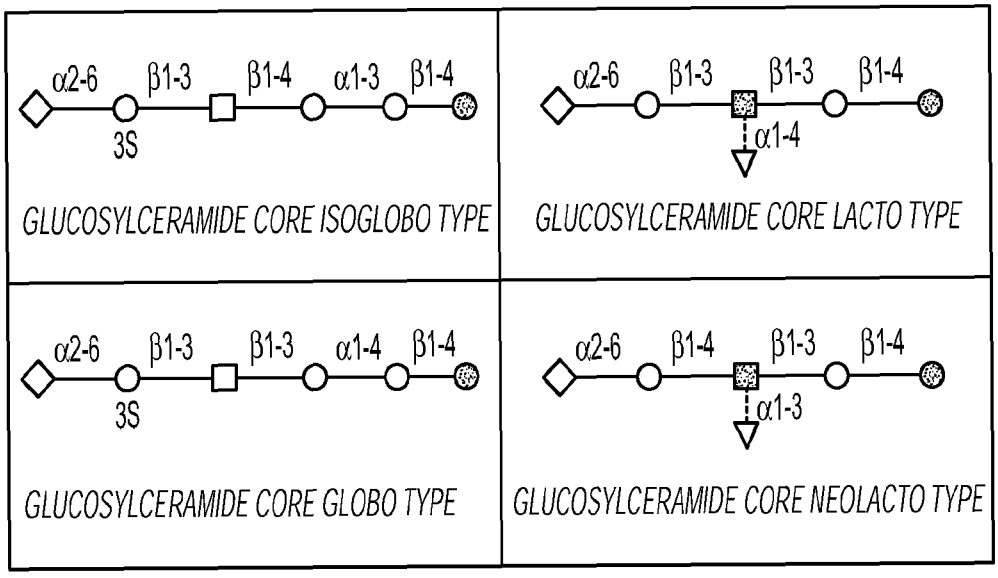
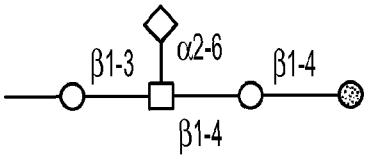


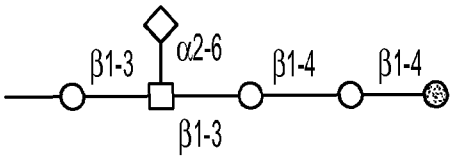
FIG. 9A-6

LONG α2-6 *UMBRELLA*-LIKE TOPOLOGY GLYCAN THAT ARE NOT DECOYS

GLYCOLIPIDS:



GLUCOSYLCERAMIDE CORE GANGLIO TYPE



GLUCOSYLCERAMIDE CORE GLOBO TYPE

FIG. 9A-7

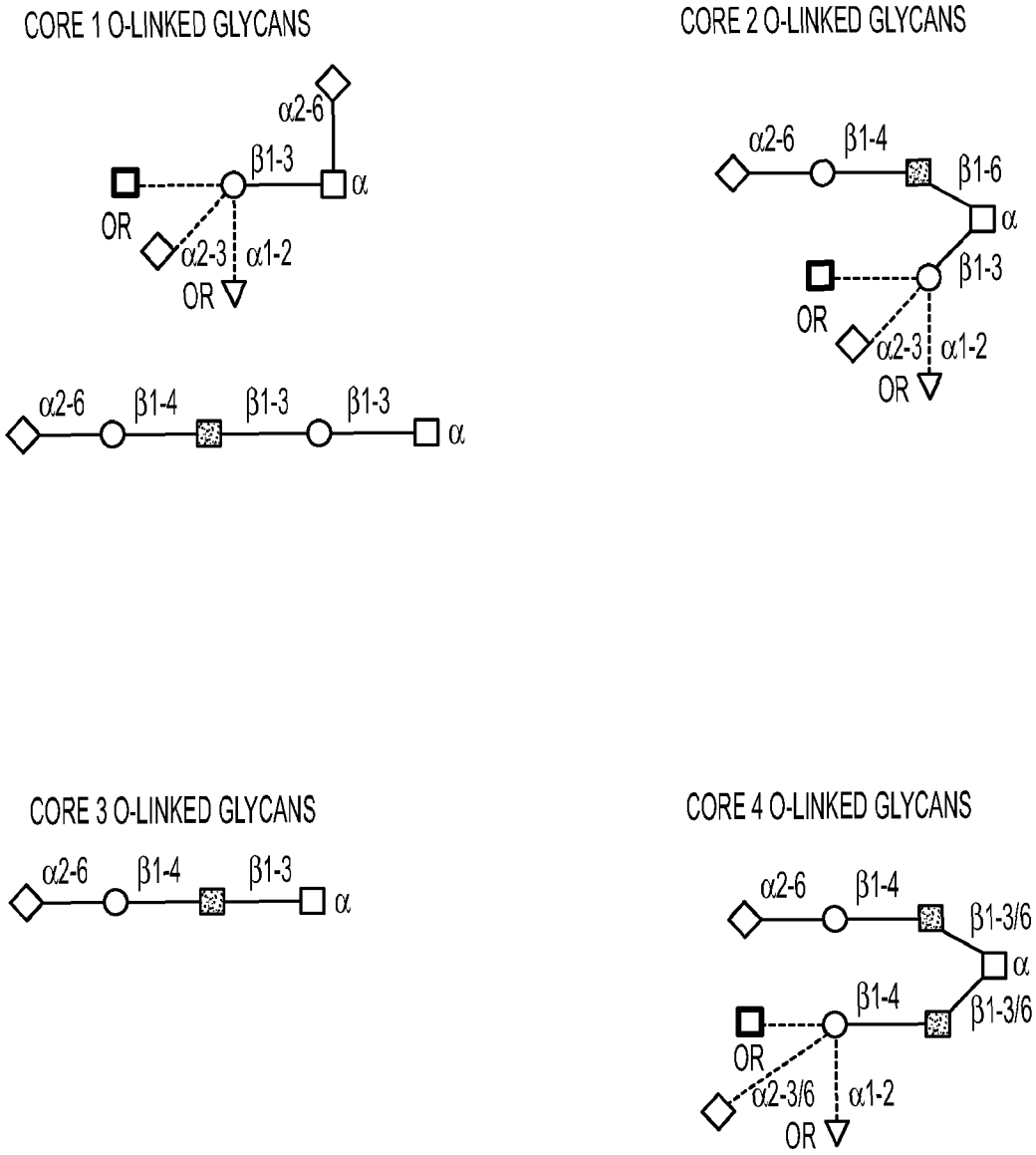

CORE 1 O-LINKED GLYCANS
α2-6
β1-3
α
OR
α2-3
α1-2
OR
α2-6
β1-4
β1-3
β1-3
α
CORE 2 O-LINKED GLYCANS
α2-6
β1-4
β1-6
α
β1-3
OR
α2-3
α1-2
OR
CORE 3 O-LINKED GLYCANS
α2-6
β1-4
β1-3
α
CORE 4 O-LINKED GLYCANS
α2-6
β1-4
β1-3/6
α
β1-4
β1-3/6
OR
α2-3/6
α1-2
OR
GlcNAc
α GalNAc O-LINKED TO PROTEIN
Neu5Ac
FUCOSE
Gal
TERMINAL HexNAc

FIG. 9B

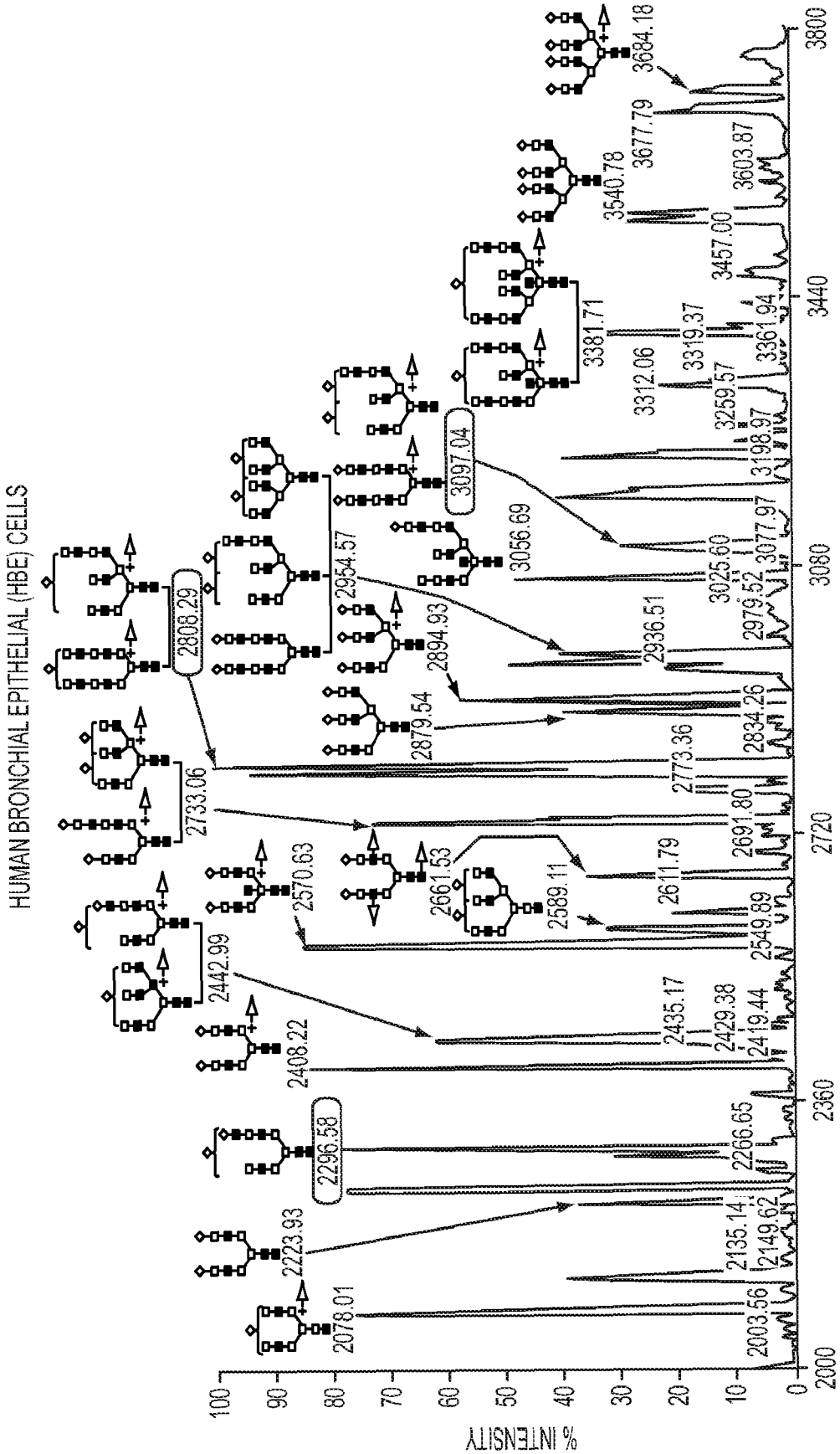

HUMAN BRONCHIAL EPITHELIAL (HBE) CELLS
% INTENSITY
MASS (m/z)
2000
2360
2720
3080
3440
3800
2078.01
2003.56
2135.14
2149.62
2223.93
2296.58
2266.65
2408.22
2442.99
2435.17
2429.38
2419.44
2549.89
2570.63
2661.53
2589.11
2611.79
2691.80
2733.06
2773.36
2808.29
2834.26
2879.54
2894.93
2954.57
2936.51
3025.60
2979.52
3077.97
3056.69
3097.04
3198.97
3259.57
3312.06
3319.37
3361.94
3381.71
3457.00
3540.78
3603.87
3677.79
3684.18

FIG. 10A

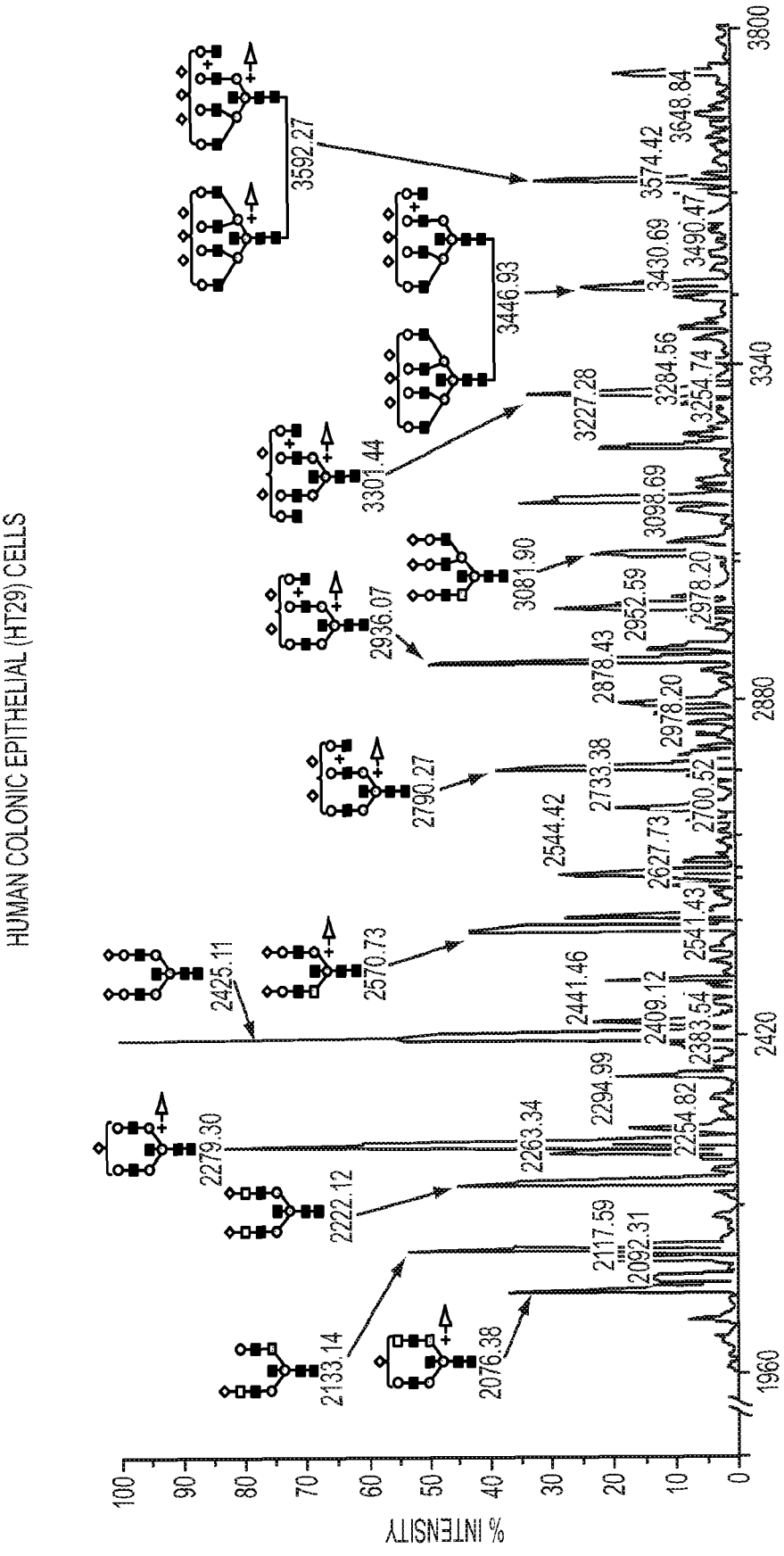

HUMAN COLONIC EPITHELIAL (HT29) CELLS
% INTENSITY
MASS (m/z)
100
90
80
70
60
50
40
30
20
10
0
1960
2420
2880
3340
3800
2133.14
2076.38
2279.30
2222.12
2425.11
2570.73
2790.27
2936.07
3301.44
3446.93
3592.27
2117.59
2092.31
2263.34
2294.99
2254.82
2441.46
2409.12
2383.54
2541.43
2544.42
2733.38
2627.73
2700.52
2978.20
2878.43
2952.59
3081.90
3098.69
3227.28
3284.56
3254.74
3430.69
3490.47
3574.42
3648.84

FIG. 10B

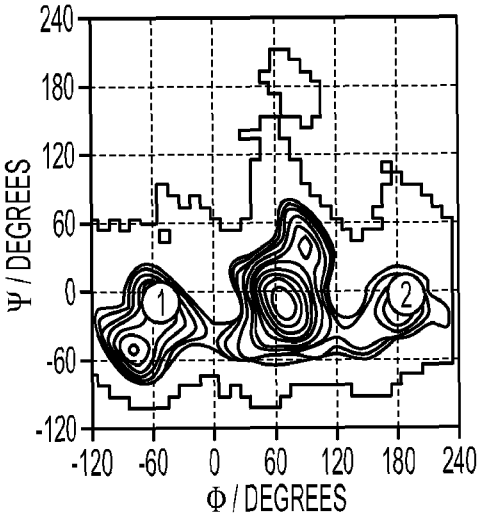


FIG. 11A

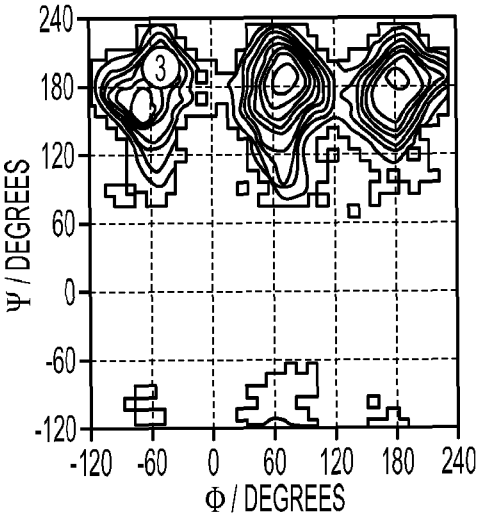


FIG. 11B

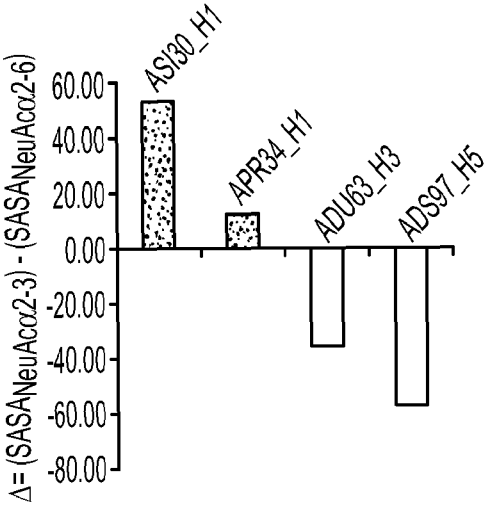


FIG. 11C

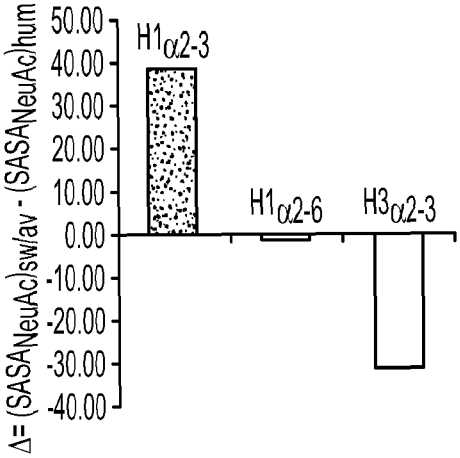
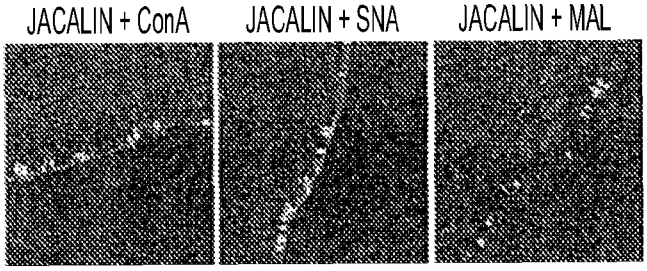


FIG. 11D



LECTIN STAINING OF UPPER RESPIRATORY TISSUE SECTIONS

FIG. 12A

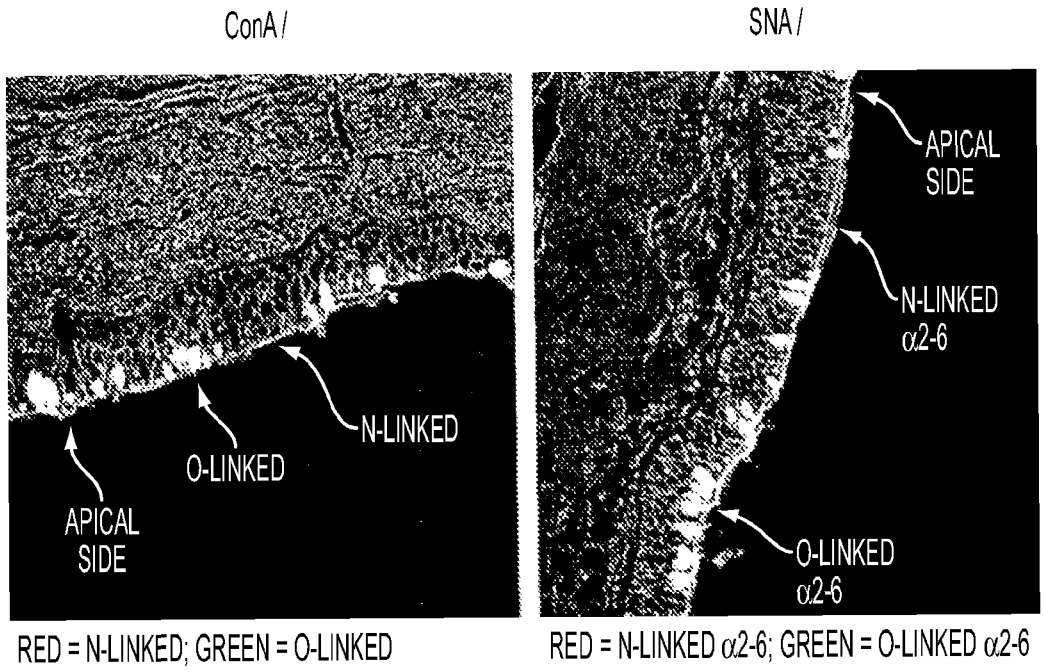

ConA /
SNA /
APICAL SIDE
O-LINKED
N-LINKED
APICAL SIDE
N-LINKED α2-6
O-LINKED α2-6
RED = N-LINKED; GREEN = O-LINKED
RED = N-LINKED α2-6; GREEN = O-LINKED α2-6

FIG. 12B

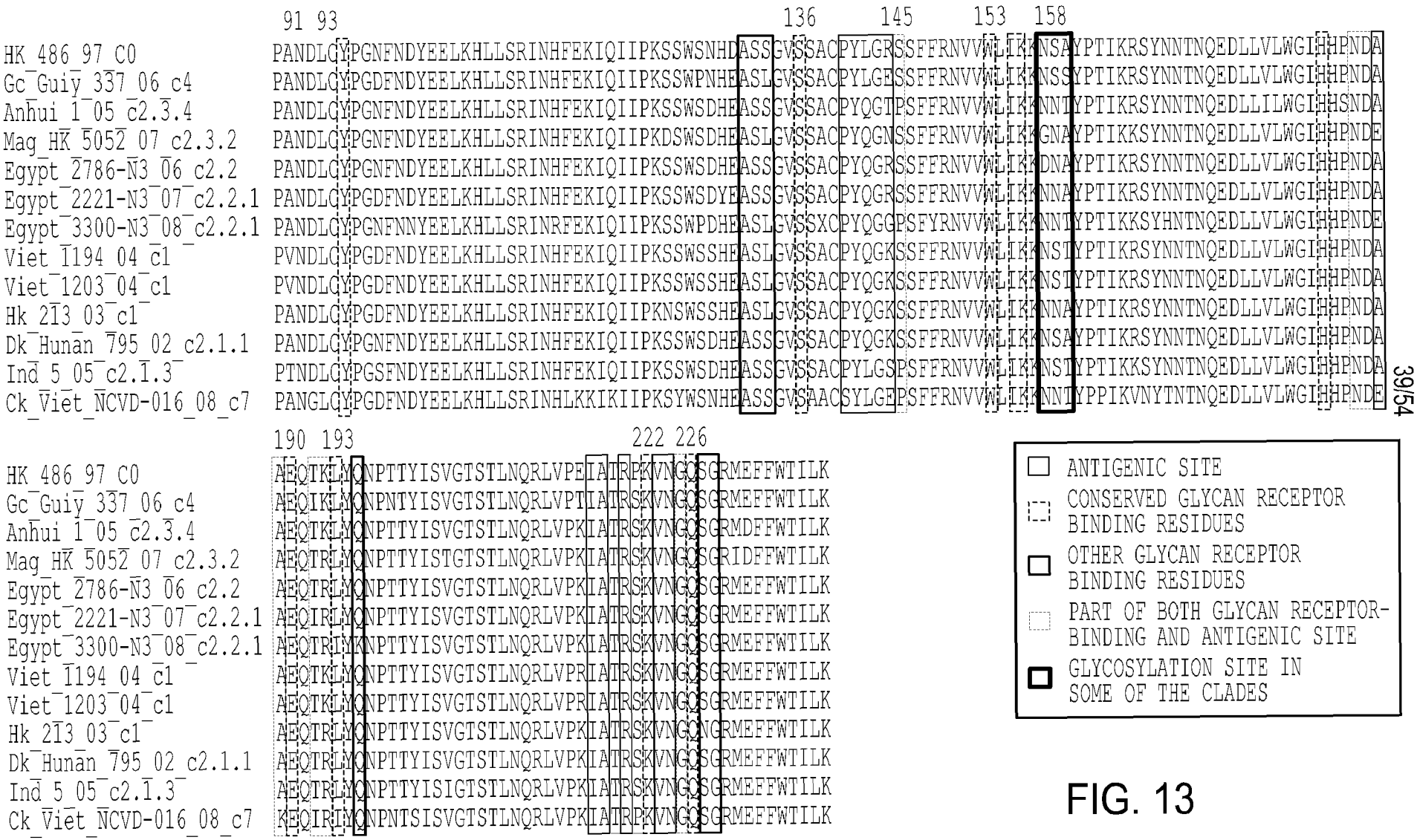

91 93
136
145
153
158
HK_486_97_C0
Gc_Guiy_337_06_c4
Anhui_1_05_c2.3.4
Mag_HK_5052_07_c2.3.2
Egypt_2786-N3_06_c2.2
Egypt_2221-N3_07_c2.2.1
Egypt_3300-N3_08_c2.2.1
Viet_1194_04_c1
Viet_1203_04_c1
Hk_213_03_c1
Dk_Hunan_795_02_c2.1.1
Ind_5_05_c2.1.3
Ck_Viet_NCVD-016_08_c7
39/54
190 193
222 226
ANTIGENIC SITE
CONSERVED GLYCAN RECEPTOR BINDING RESIDUES
OTHER GLYCAN RECEPTOR BINDING RESIDUES
PART OF BOTH GLYCAN RECEPTOR-BINDING AND ANTIGENIC SITE
GLYCOSYLATION SITE IN SOME OF THE CLADES

FIG. 13

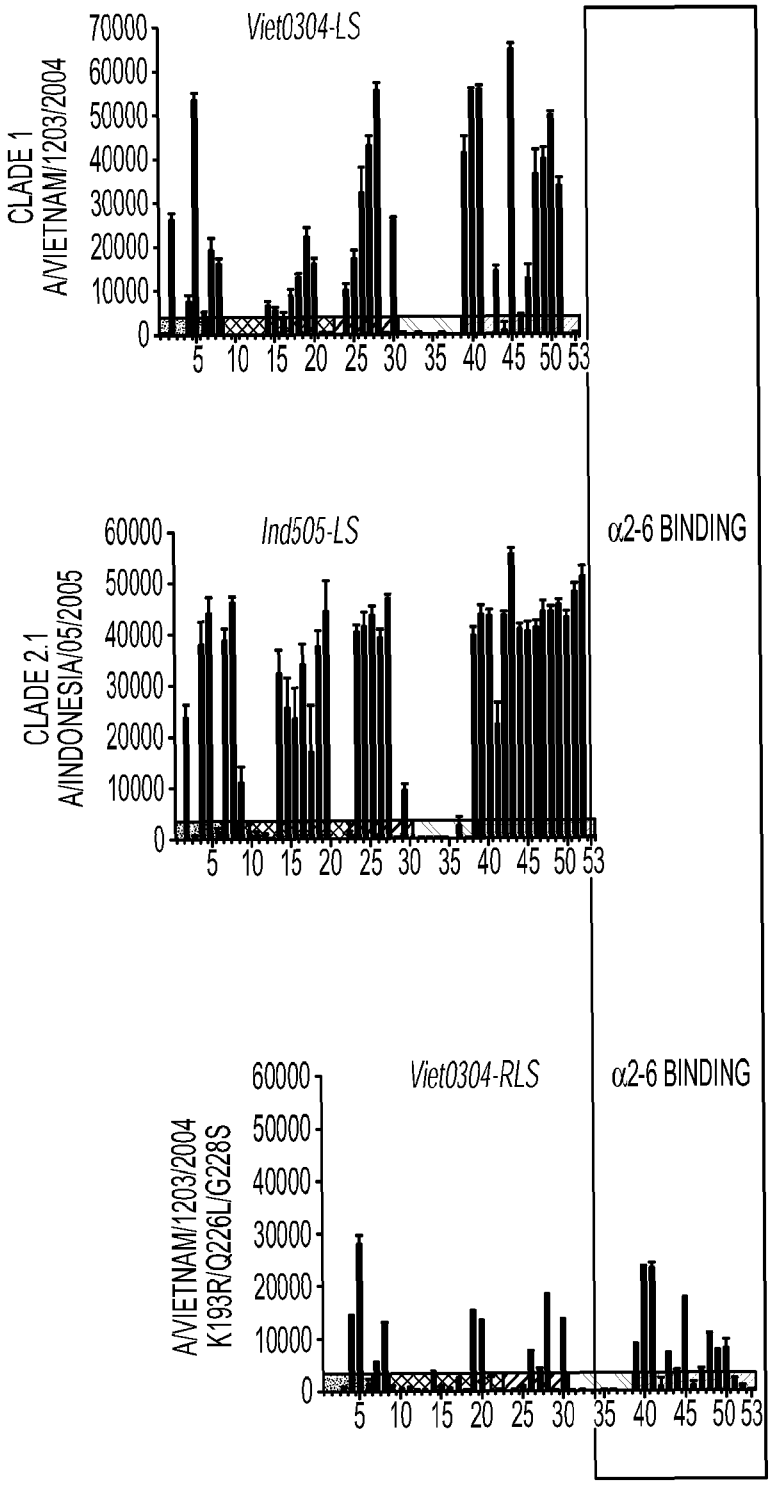

CLADE 1
A/VIETNAM/1203/2004
Viet0304-LS
70000
60000
50000
40000
30000
20000
10000
0
5 10 15 20 25 30 35 40 45 50 53
CLADE 2.1
A/INDONESIA/05/2005
Ind505-LS
α2-6 BINDING
60000
50000
40000
30000
20000
10000
0
5 10 15 20 25 30 35 40 45 50 53
A/VIETNAM/1203/2004
K193R/Q226L/G228S
Viet0304-RLS
α2-6 BINDING
60000
50000
40000
30000
20000
10000
0
5 10 15 20 25 30 35 40 45 50 53

FIG. 14

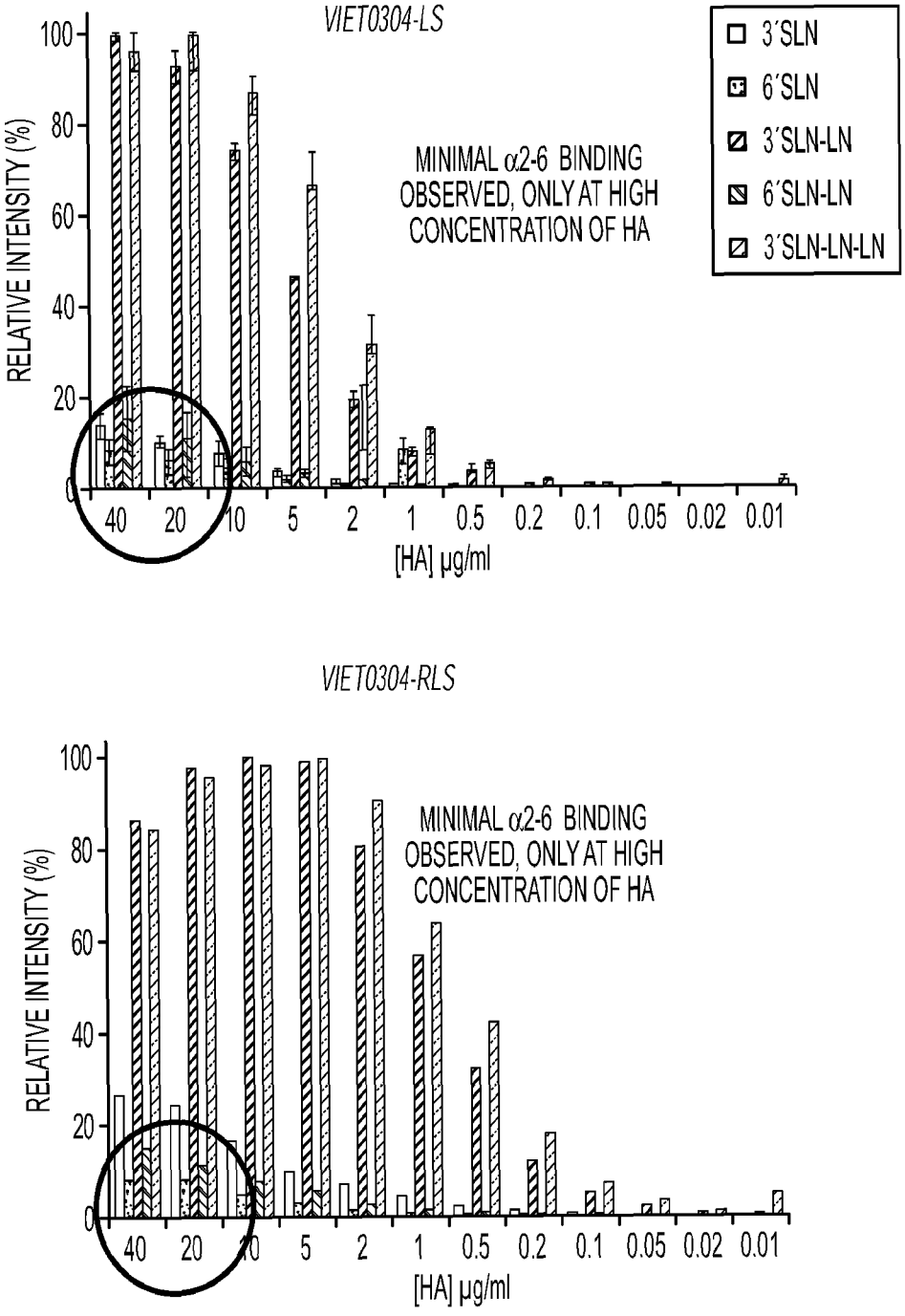

VIET0304-LS
3´SLN
6´SLN
3´SLN-LN
6´SLN-LN
3´SLN-LN-LN
MINIMAL α2-6 BINDING OBSERVED, ONLY AT HIGH CONCENTRATION OF HA
RELATIVE INTENSITY (%)
100
80
60
40
20
0
40 20 10 5 2 1 0.5 0.2 0.1 0.05 0.02 0.01
[HA] µg/ml
VIET0304-RLS
MINIMAL α2-6 BINDING OBSERVED, ONLY AT HIGH CONCENTRATION OF HA
RELATIVE INTENSITY (%)
100
80
60
40
20
0
40 20 10 5 2 1 0.5 0.2 0.1 0.05 0.02 0.01
[HA] µg/ml

FIG. 15

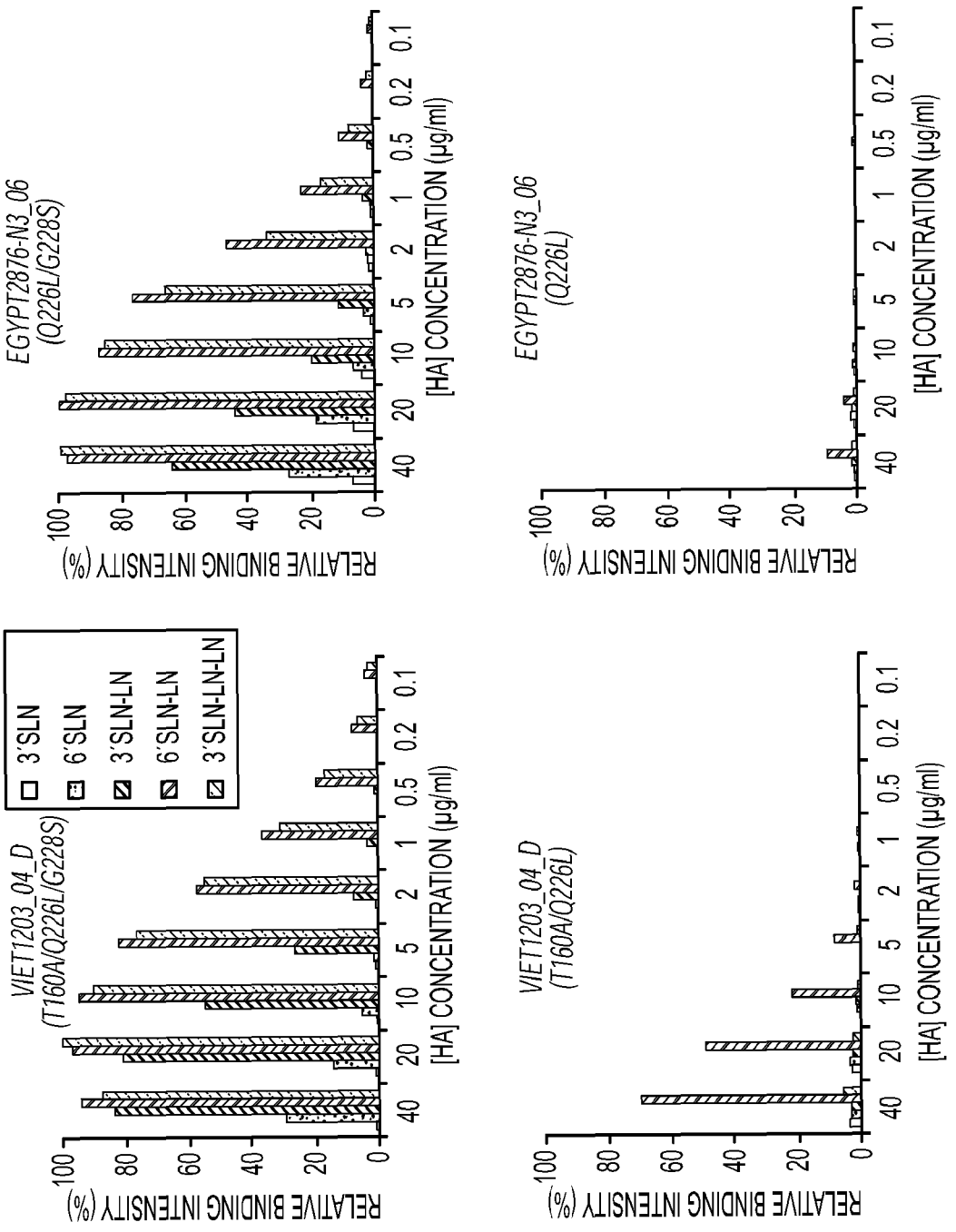

VIET1203_04_D
(T160A/Q226L/G228S)
EGYPT2876-N3_06
(Q226L/G228S)
VIET1203_04_D
(T160A/Q226L)
EGYPT2876-N3_06
(Q226L)
3'SLN
6'SLN
3'SLN-LN
6'SLN-LN
3'SLN-LN-LN
RELATIVE BINDING INTENSITY (%)
[HA] CONCENTRATION (µg/ml)
0
20
40
60
80
100
40
20
10
5
2
1
0.5
0.2
0.1

FIG. 16

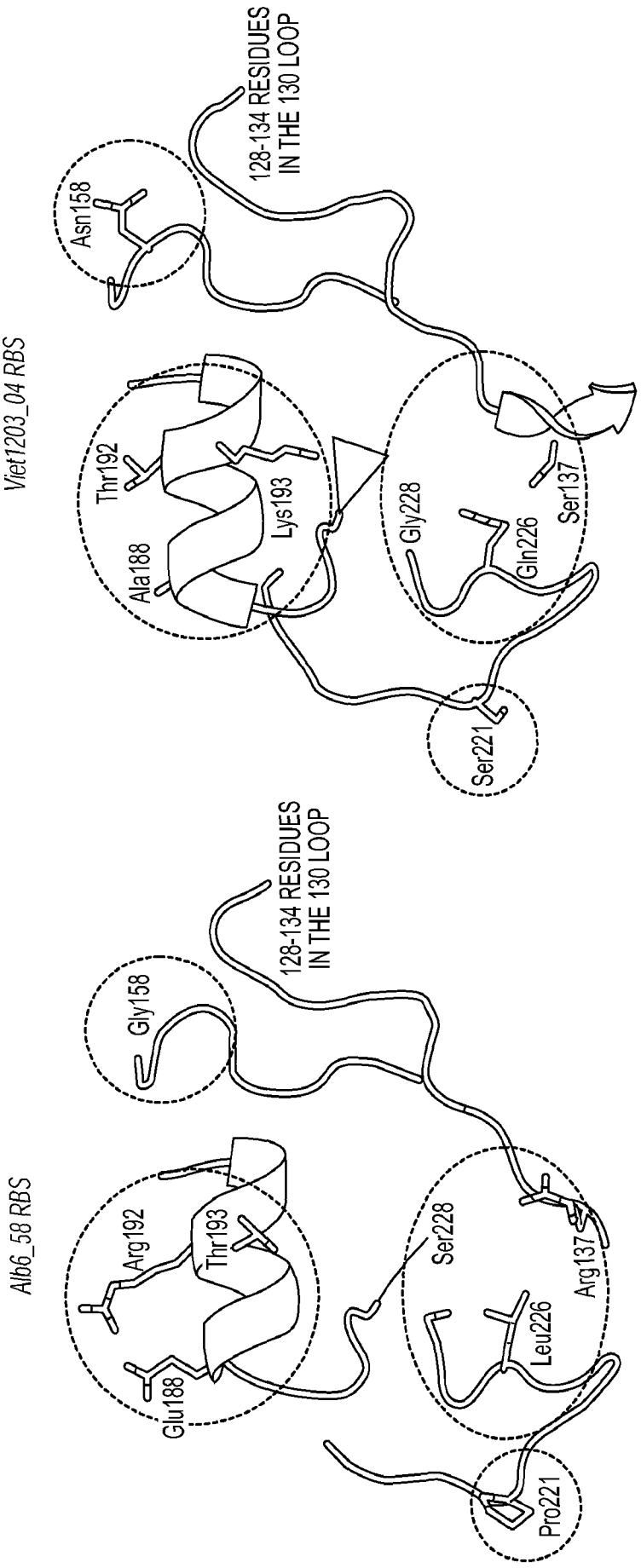

Alb6_58 RBS
Glu188
Arg192
Thr193
Gly158
128-134 RESIDUES
IN THE 130 LOOP
Pro221
Ser228
Leu226
Arg137
Viet1203_04 RBS
Ala188
Thr192
Lys193
Asn158
128-134 RESIDUES
IN THE 130 LOOP
Ser221
Gly228
Gln226
Ser137

FIG. 17

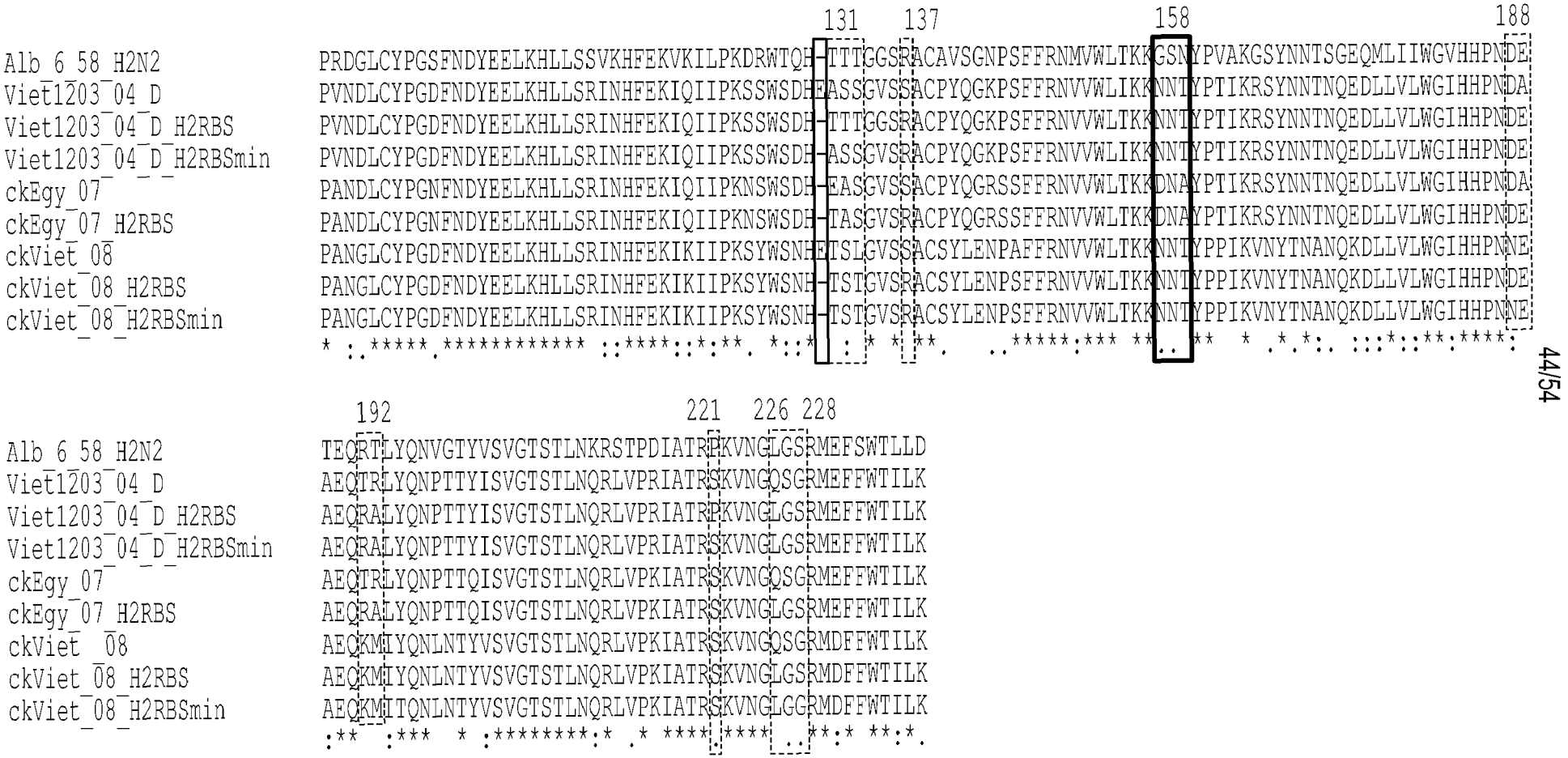

131
137
158
188
Alb_6_58_H2N2
Viet1203_04_D
Viet1203_04_D_H2RBS
Viet1203_04_D_H2RBSmin
ckEgy_07
ckEgy_07_H2RBS
ckViet_08
ckViet_08_H2RBS
ckViet_08_H2RBSmin
192
221
226
228
Alb_6_58_H2N2
Viet1203_04_D
Viet1203_04_D_H2RBS
Viet1203_04_D_H2RBSmin
ckEgy_07
ckEgy_07_H2RBS
ckViet_08
ckViet_08_H2RBS
ckViet_08_H2RBSmin

FIG. 18

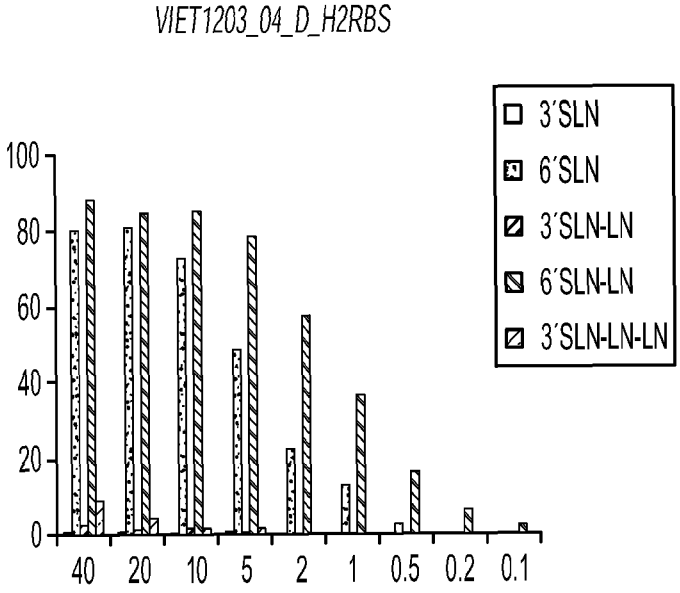

VIET1203_04_D_H2RBS
3´SLN
6´SLN
3´SLN-LN
6´SLN-LN
3´SLN-LN-LN
100
80
60
40
20
0
40
20
10
5
2
1
0.5
0.2
0.1

FIG. 19

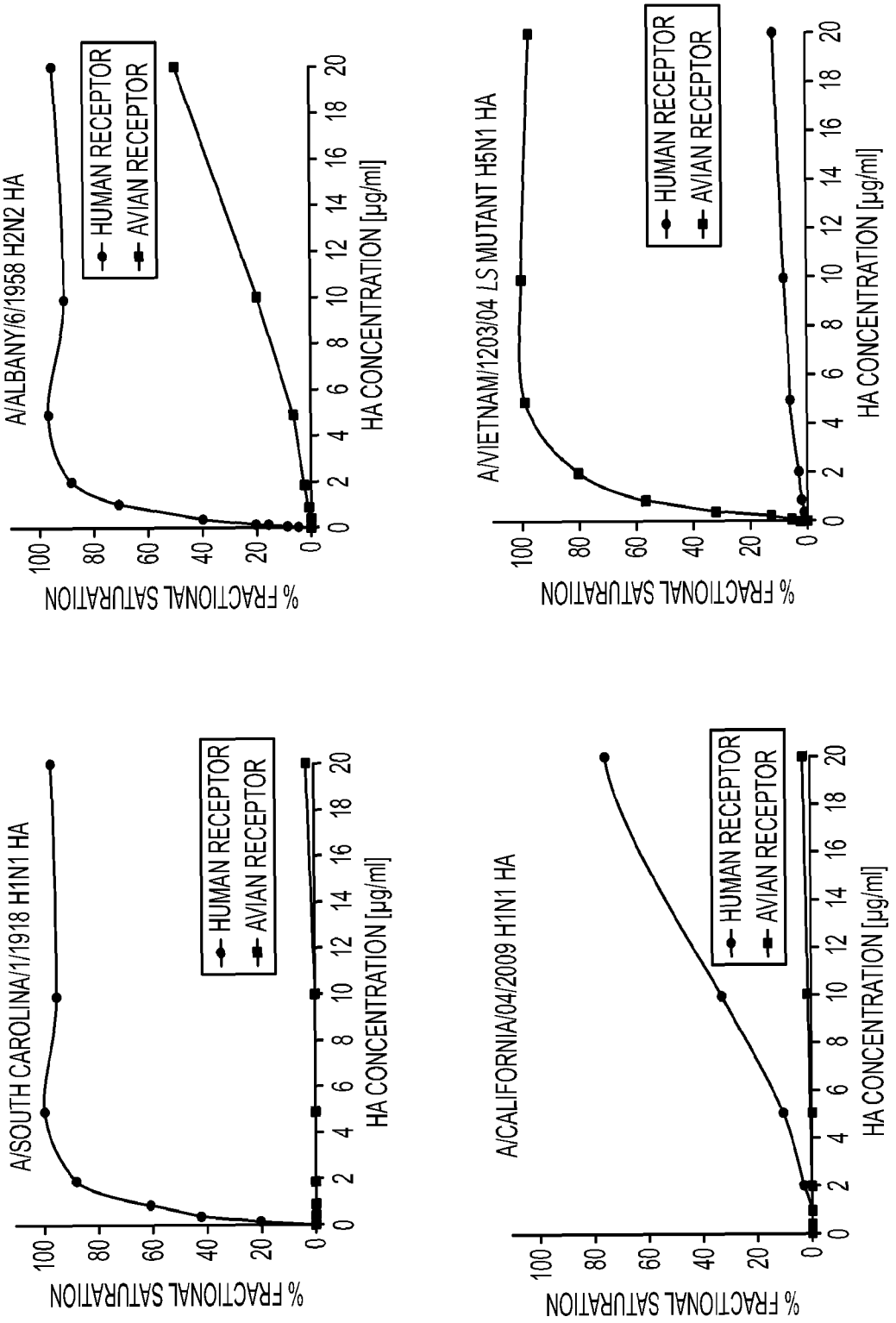

A/ALBANY/6/1958 H2N2 HA
A/VIETNAM/1203/04 LS MUTANT H5N1 HA
A/SOUTH CAROLINA/1/1918 H1N1 HA
A/CALIFORNIA/04/2009 H1N1 HA
HUMAN RECEPTOR
AVIAN RECEPTOR
% FRACTIONAL SATURATION
HA CONCENTRATION [µg/ml]
0
20
40
60
80
100
0
2
4
6
8
10
12
14
16
18
20

FIG. 20

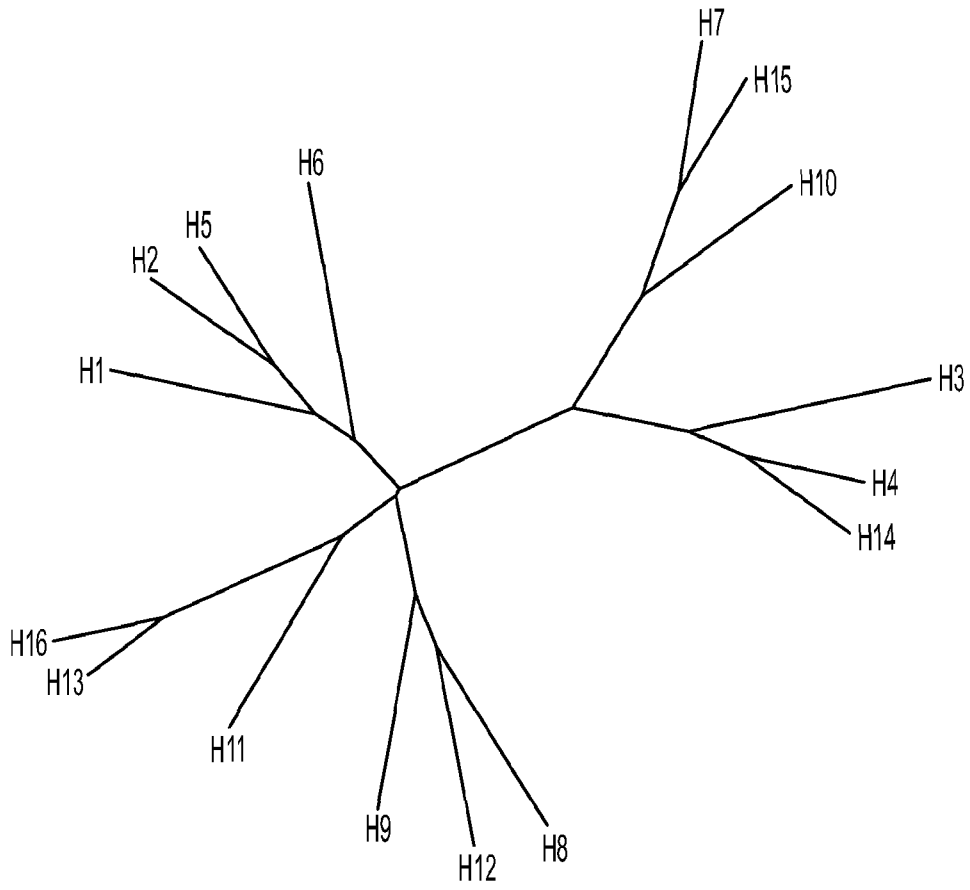


FIG. 21

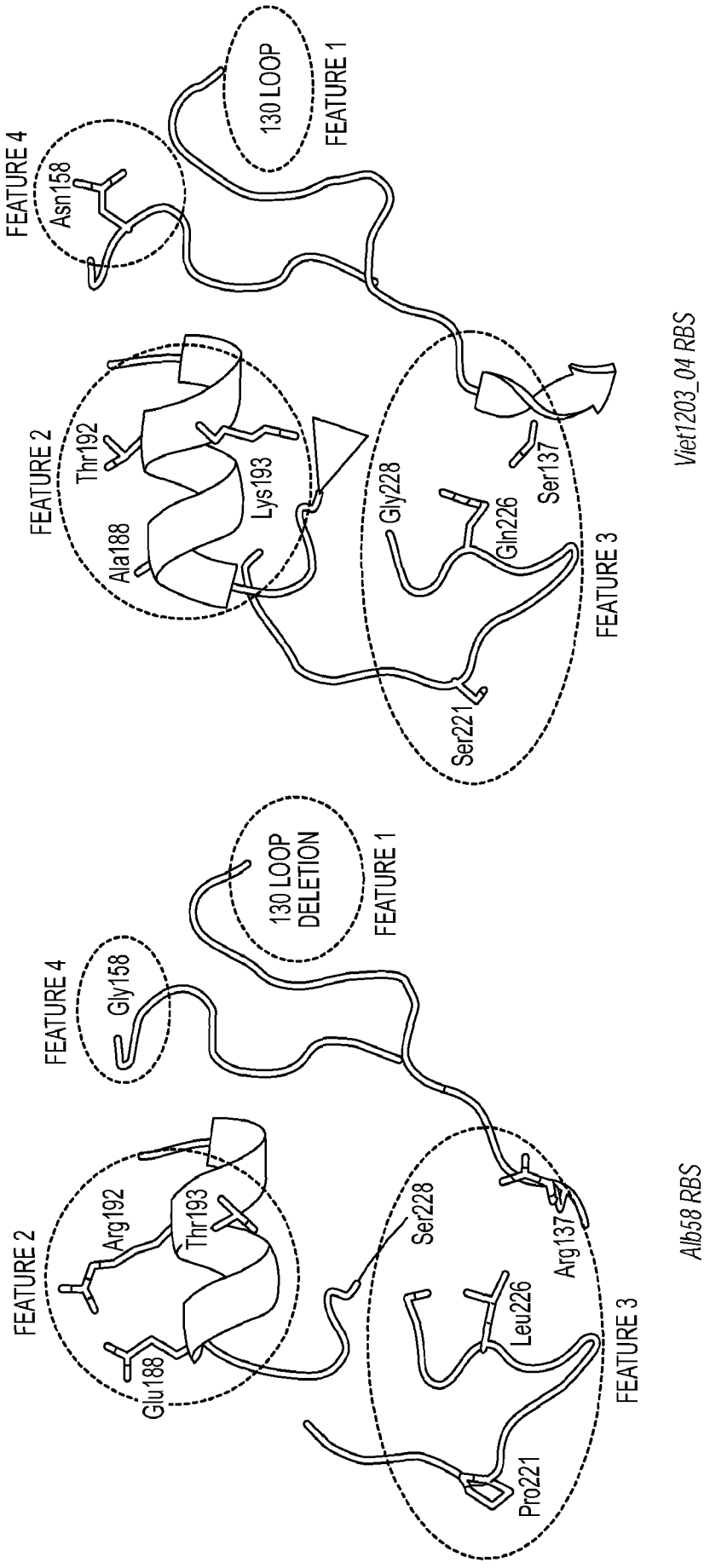

FEATURE 4
Asn158
130 LOOP
FEATURE 1
FEATURE 2
Thr192
Ala188
Lys193
Gly228
Gln226
Ser137
FEATURE 3
Ser221
Viet1203_04 RBS
FEATURE 4
Gly158
130 LOOP
DELETION
FEATURE 1
FEATURE 2
Arg192
Thr193
Ser228
Arg137
Leu226
Glu188
FEATURE 3
Pro221
Alb58 RBS

FIG. 22

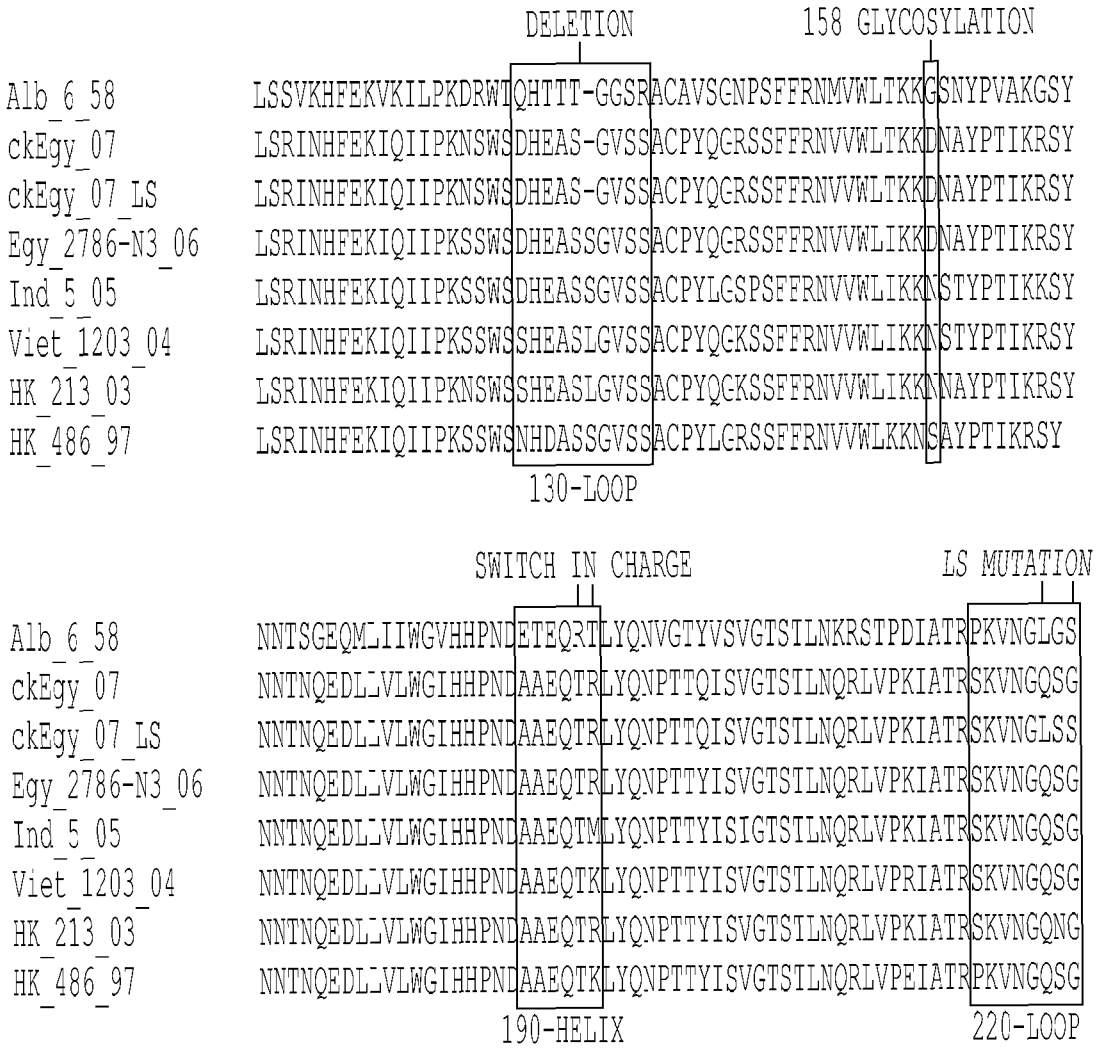

DELETION
158 GLYCOSYLATION
Alb_6_58 LSSVKHFEKVKILPKDRWTQHTTT-GGSRACAVSGNPSFFRNMVWLTKKGSNYPVAKGSY
ckEgy_07 LSRINHFEKIQIIPKNSWSDHEAS-GVSSACPYQGRSSFFRNVVWLTKKDNAYPTIKRSY
ckEgy_07_LS LSRINHFEKIQIIPKNSWSDHEAS-GVSSACPYQGRSSFFRNVVWLTKKDNAYPTIKRSY
Egy_2786-N3_06 LSRINHFEKIQIIPKSSWSDHEASSGVSSACPYQGRSSFFRNVVWLIKKDNAYPTIKRSY
Ind_5_05 LSRINHFEKIQIIPKSSWSDHEASSGVSSACPYLGSPSFFRNVVWLIKKNSTYPTIKKSY
Viet_1203_04 LSRINHFEKIQIIPKSSWSSHEASLGVSSACPYQGKSSFFRNVVWLIKKNSTYPTIKRSY
HK_213_03 LSRINHFEKIQIIPKNSWSSHEASLGVSSACPYQGKSSFFRNVVWLIKKNNAYPTIKRSY
HK_486_97 LSRINHFEKIQIIPKSSWSNHDASSGVSSACPYLGRSSFFRNVVWLKKNSAYPTIKRSY
130-LOOP
SWITCH IN CHARGE
LS MUTATION
Alb_6_58 NNTSGEQMLIIWGVHHPNDETEQRTLYQNVGTYVSVGTSILNKRSTPDIATRPKVNGLGS
ckEgy_07 NNTNQEDLLVLWGIHHPNDAAEQTRLYQNPTTQISVGTSILNQRLVPKIATRSKVNGQSG
ckEgy_07_LS NNTNQEDLLVLWGIHHPNDAAEQTRLYQNPTTQISVGTSILNQRLVPKIATRSKVNGLSS
Egy_2786-N3_06 NNTNQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISVGTSILNQRLVPKIATRSKVNGQSG
Ind_5_05 NNTNQEDLLVLWGIHHPNDAAEQTMLYQNPTTYISIGTSILNQRLVPKIATRSKVNGQSG
Viet_1203_04 NNTNQEDLLVLWGIHHPNDAAEQTKLYQNPTTYISVGTSILNQRLVPRIATRSKVNGQSG
HK_213_03 NNTNQEDLLVLWGIHHPNDAAEQTRLYQNPTTYISVGTSILNQRLVPKIATRSKVNGQNG
HK_486_97 NNTNQEDLLVLWGIHHPNDAAEQTKLYQNPTTYISVGTSILNQRLVPEIATRPKVNGQSG
190-HELIX
220-LOOP

FIG. 23

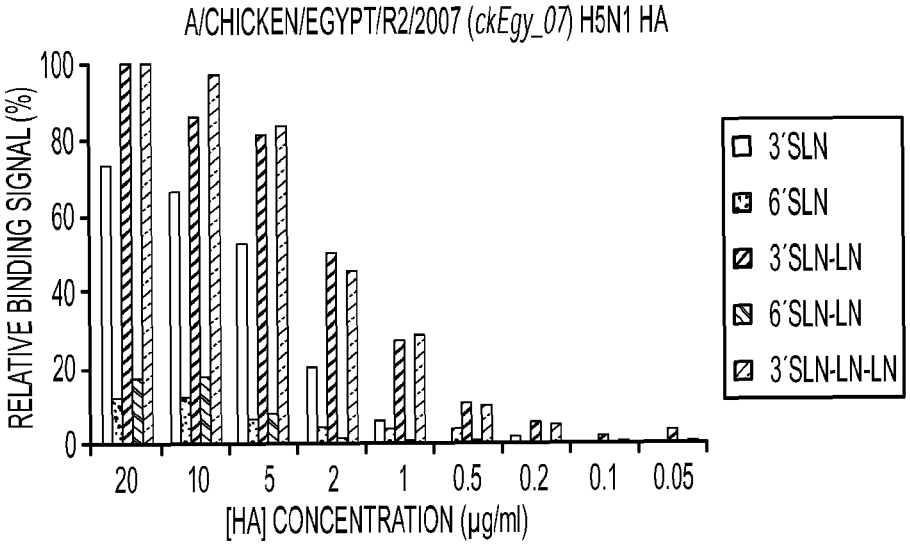


FIG. 24A

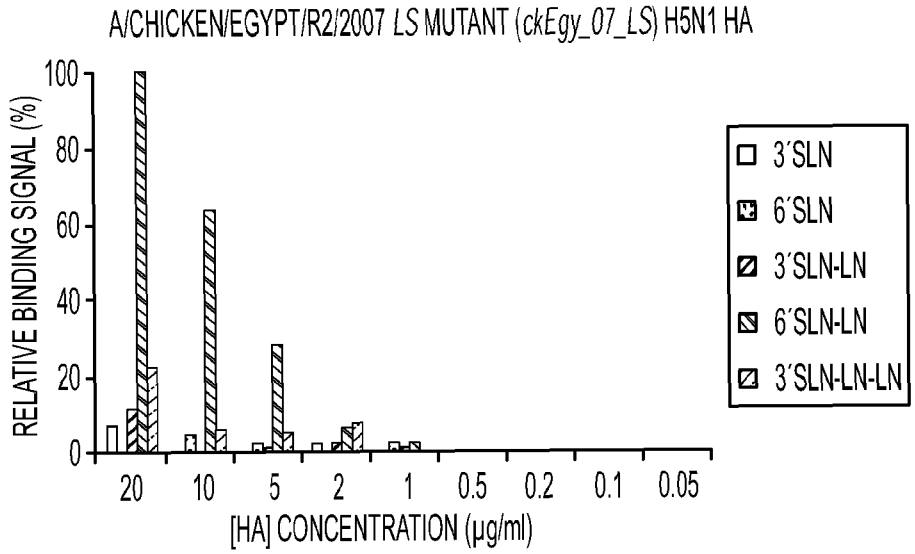
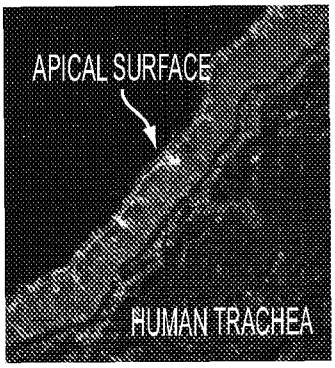


FIG. 24B



ckEgy_07_LS MUTANT HA/ PI

FIG. 24C

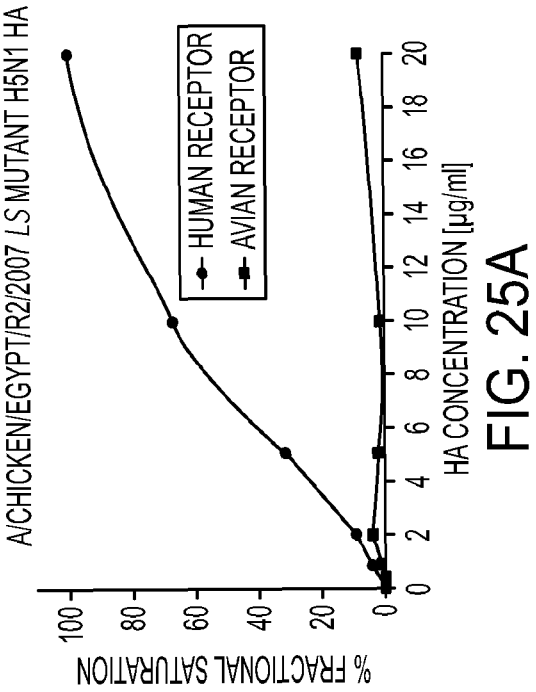


FIG. 25A

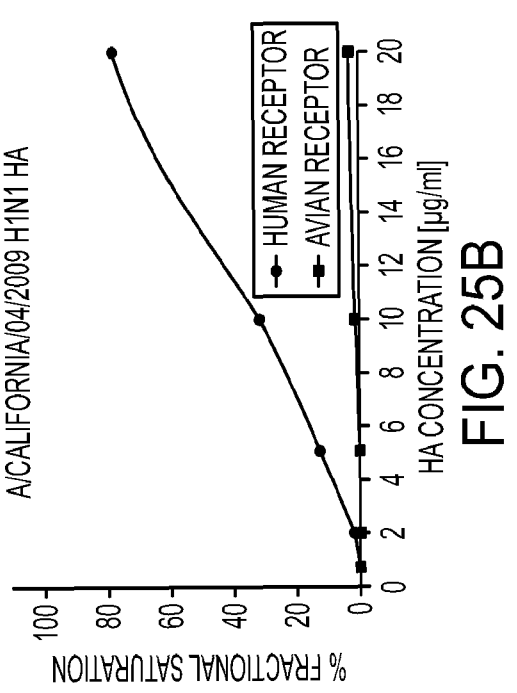


FIG. 25B

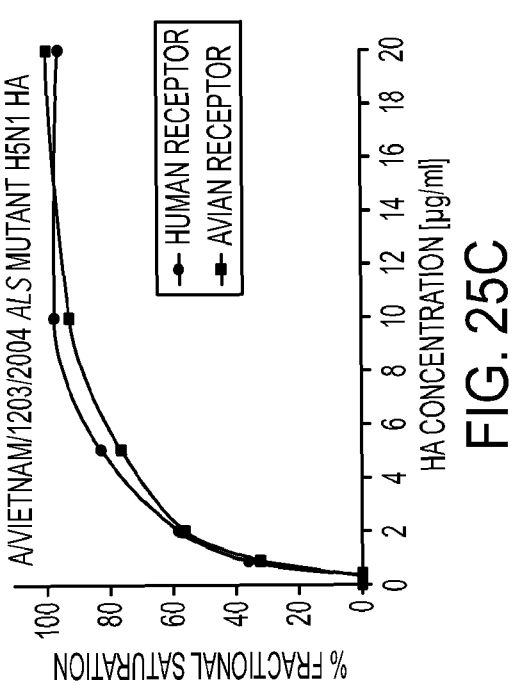


FIG. 25C

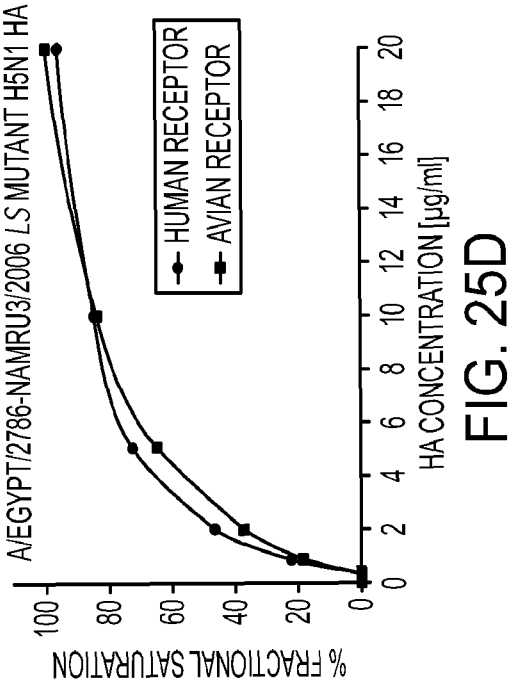


FIG. 25D

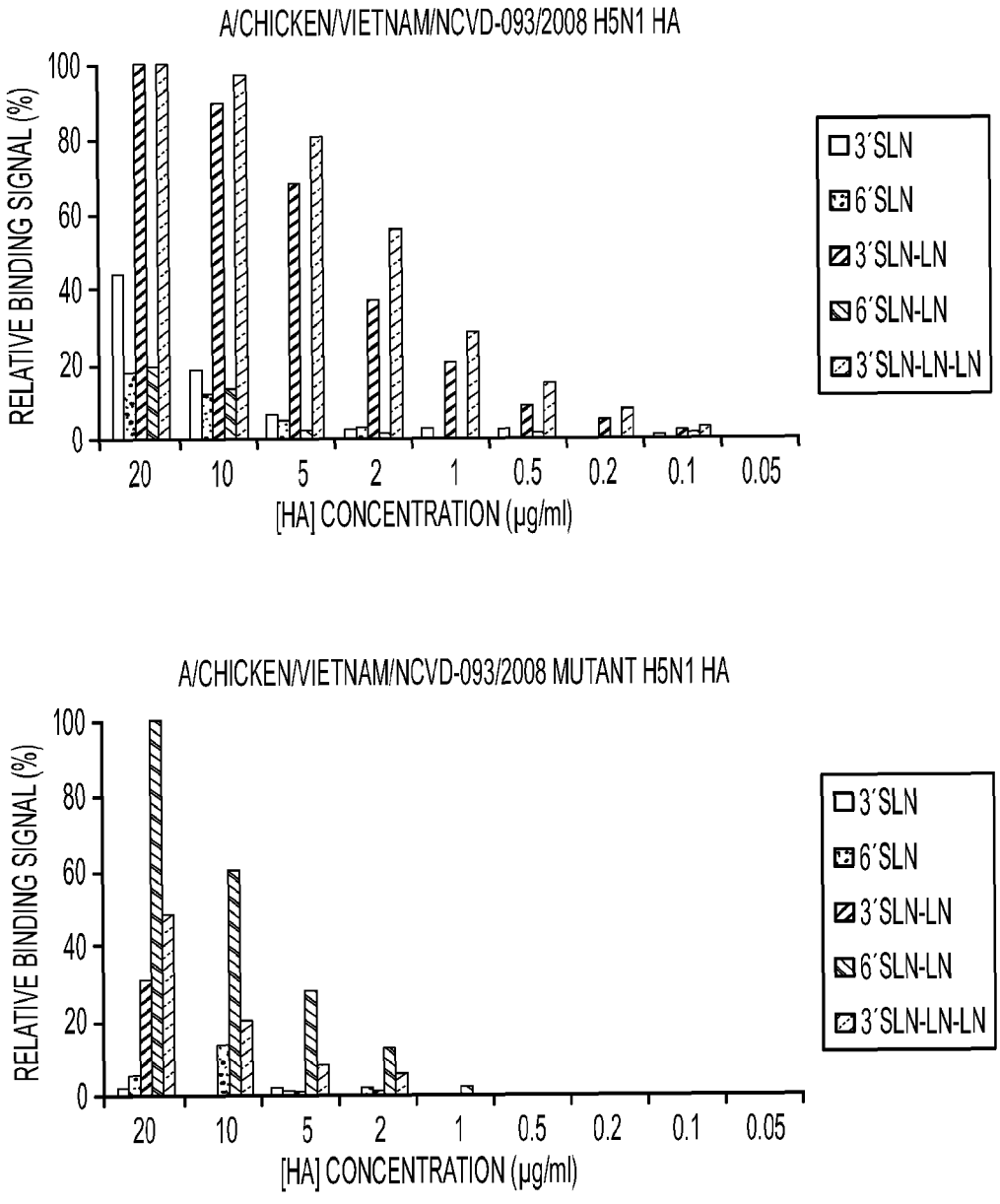

A/CHICKEN/VIETNAM/NCVD-093/2008 H5N1 HA
RELATIVE BINDING SIGNAL (%)
100
80
60
40
20
0
20
10
5
2
1
0.5
0.2
0.1
0.05
[HA] CONCENTRATION (µg/ml)
3´SLN
6´SLN
3´SLN-LN
6´SLN-LN
3´SLN-LN-LN
A/CHICKEN/VIETNAM/NCVD-093/2008 MUTANT H5N1 HA
RELATIVE BINDING SIGNAL (%)
100
80
60
40
20
0
20
10
5
2
1
0.5
0.2
0.1
0.05
[HA] CONCENTRATION (µg/ml)
3´SLN
6´SLN
3´SLN-LN
6´SLN-LN
3´SLN-LN-LN

FIG. 26

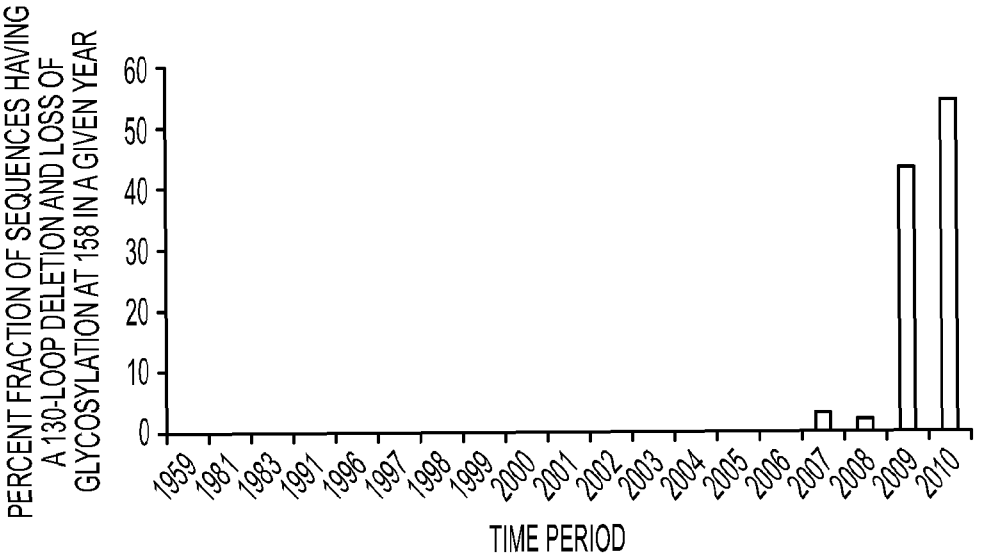


FIG. 27A

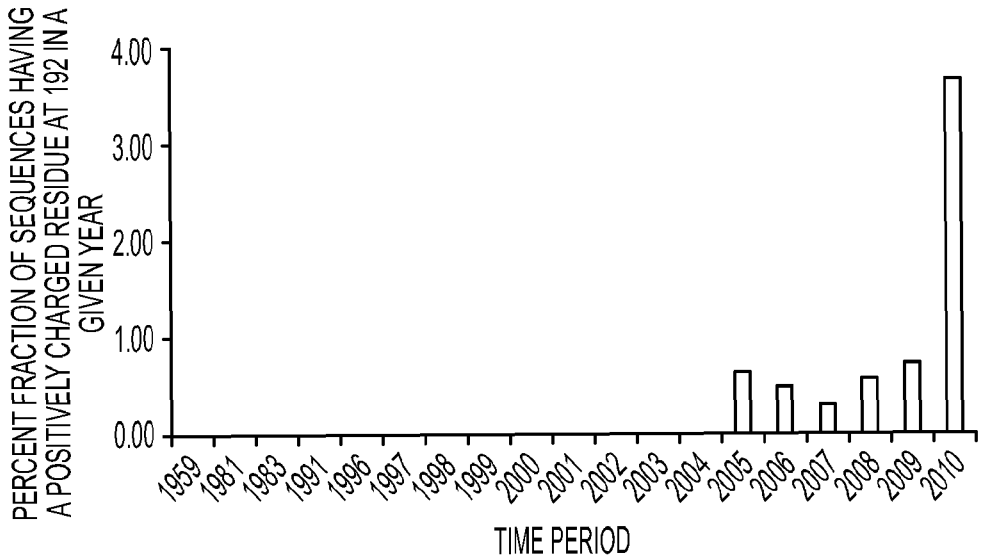


FIG. 27B

EXPANDED NOMENCLATURE OF GLYCANS USED IN THE GLYCAN ARRAY

GLYCAN	EXPANDED NOMENCLATURE
3´SLN	Neu5Acα2-3Galβ1-4GlcNAcβ1-
6´SLN	Neu5Acα2-6Galβ1-4GlcNAcβ1-
3´SLN-LN	Neu5Acα2-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-
6´SLN-LN	Neu5Acα2-6Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-
3´SLN-LN-LN	Neu5Acα2-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1

KEY: Neu5Ac: N-ACETYL D-NEURAMINIC ACID; Gal: D-GALACTOSE; GlcNAc: N-ACETYL D-GLUCOSAMINE. α / β: ANOMERIC CONFIGURATION OF THE PYRANOSE SUGARS. ALL THE SUGARS ARE LINKED VIA A SPACER TO BIOTIN (-Sp-LC-LC-BIOTIN AS DESCRIBED IN http://www.functionalglycomics.org/static/consortium/resources/resourcecored5.shtml)

FIG. 28

SINGAPORE PATENTS ACT (CHAPTER 221)
PATENTS RULES

PATENTS FORM 8	Statement of Inventorship and of Right to the Grant of a Patent

Estimated Time:
This form may take approximately 9 - 13 minutes to complete.

General:
*a. * denotes mandatory field.*
b. Please note that any fee paid is not refundable.

PART 1 Reference

Applicant/ Agent Reference	2200840/MIT/MK/ANG/ely

PART 2 Application No.*

Application No.*	

PART 3 Title of Invention*

Title of Invention*	HEMAGGLUTININ POLYPEPTIDES, AND REAGENTS AND METHODS RELATING THERETO

PART 4 Name of Applicant*

Note: If there is insufficient space, please use the continuation sheet CS 4.

Applicants	1	**Type of Applicant**	Corporate
		UEN/ Company Code	E0169511A
		Name	MASSACHUSETTS INSTITUTE OF TECHNOLOGY
		Address Type	Foreign
		Address	77 MASSACHUSETTS AVENUE, CAMBRIDGE, MASSACHUSETTS 02139-4307, UNITED STATES OF AMERICA Massachusetts-United States of America
		Is SME	
		Country of Incorporation	United States of America
		State of Incorporation	Massachusetts
		Sole Proprietor or Partners' Names	

PART 5 Details of Inventor*

Note: If there is insufficient space, please use the continuation sheet CS 5.

Inventor	1	**Name**	THARAKARAMAN, Kannan
		Company UEN	E00266561A
		CareOf	
		Address Type	Foreign
		Address	285 Massachusetts Avenue #11 Arlington, Massachusetts 02474

		Nationality	India
		Country of Residency	United States of America
		The above inventor was resident of Singapore at any time during the period of invention	No
	2	**Name**	VISWANATHAN, Karthik
		Company UEN	E00282940A
		CareOf	
		Address Type	Foreign
		Address	3114 Stearns Hill Road, Waltham, MA 02451, United States of America
		Nationality	India
		Country of Residency	United States of America
		The above inventor was resident of Singapore at any time during the period of invention	No
	3	**Name**	RAMAN, RAHUL
		Company UEN	E00216484A
		CareOf	
		Address Type	Foreign
		Address	132 CLOCKTOWER DRIVE, UNIT 311WALTHAM, MA 02452
		Nationality	India
		Country of Residency	United States of America
		The above inventor was resident of Singapore at any time during the period of invention	No
	4	**Name**	SASISEKHARAN, RAM
		Company UEN	E00232089A
		CareOf	
		Address Type	Foreign
		Address	C/O MASSACHUSETTS INSTITUTE OF TECHNOLOGY45 CARLETON STREET, E25-519CAMBRIDGE, MA 02139
		Nationality	United States of America
		Country of Residency	United States of America
		The above inventor was resident of Singapore at any time during the period of invention	No

PART 6 Derivation of Right*	
The Applicants Derived the Right from the Inventor to be Granted the Said Application* (*Note: Only one checkbox may be crossed)*	☐ By Contract of Employment ☒ By Assignment of Invention ☐ Others *(Please specify below)*

PART 7 Contact Details*	
Note: *a. This part need not be filled up if the person's address in Part 5 is to be used as an address for service in Singapore.* *b. The email address and telephone number provided are for contact purposes in case of query.* *c. The official correspondence will be sent electronically via IP2SG.*	
Agent UEN/ Company Code	T14LL0024B
Agent Name	DONALDSON & BURKINSHAW LLP
Representative or C/O Name	
	Address for Service in Singapore Block/ House No. Street Name Level - Unit 3667 Building Name ROBINSON ROAD POST OFFICE Postal Code 905667
Contact Person	
Direct Telephone No.	
Email Address	

PART 8 Statement

Statement	I/ We believe that the person(s) stated in part 5 is/ are the inventor(s) of the invention in respect of which the patent application is made. I/ We also consent to the publication of the details contained herein to each of the inventors as stated in part 5.
PART 9 Declaration*	
Declaration	**By Person Filing the Application** I do hereby declare that the information furnished above is true to the best of my knowledge. **By Agent** I do hereby declare that : i. I have been duly authorised to act as an agent, for the purposes of this application, on behalf of the person(s) filing this request. ii. The information furnished above on behalf of the person(s) filing this request is true to the best of the person(s)' knowledge.
Name	EVELINE ANGSTETRA
Signature	
Date *(DD/MM/YYYY)*	27/5/2022
No. of Extra Sheets Attached to this Form	sheet(s)