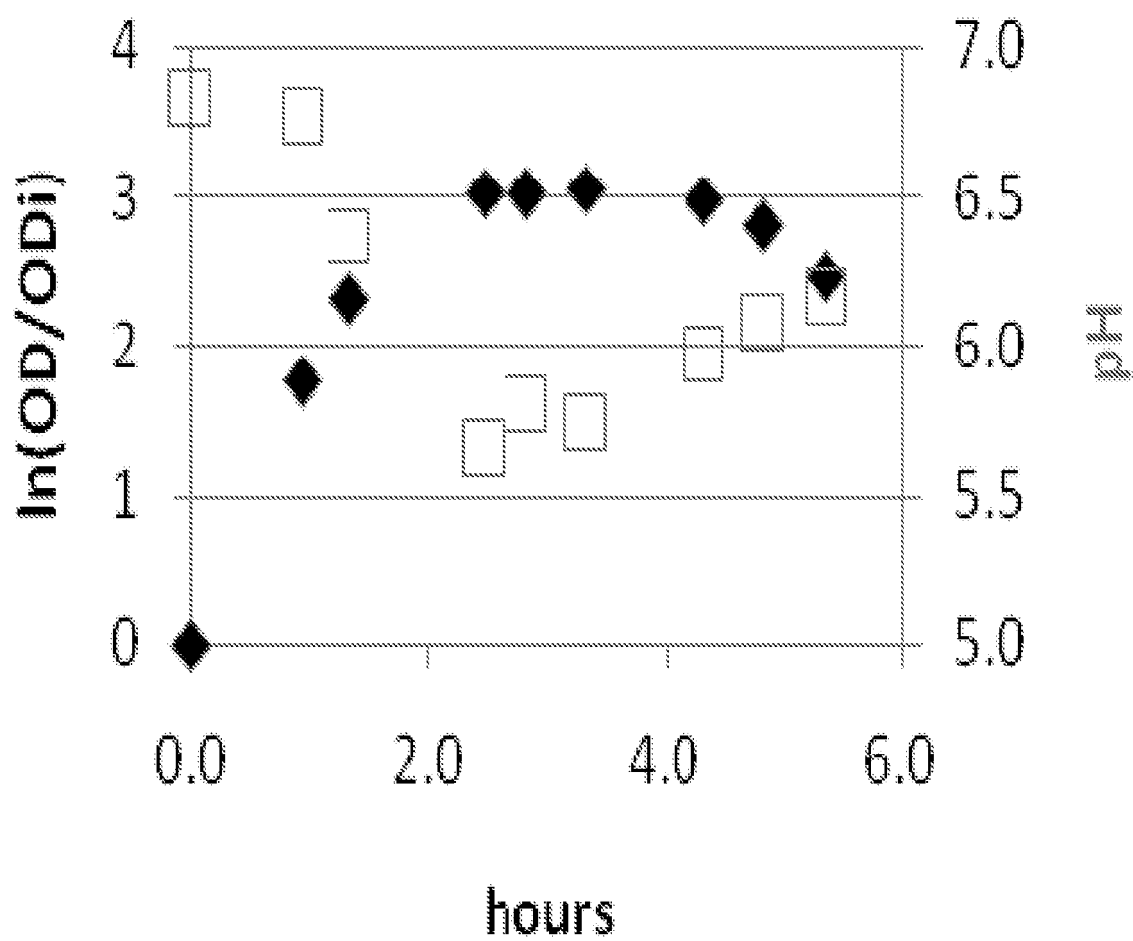


1/34

**FIG. 1**

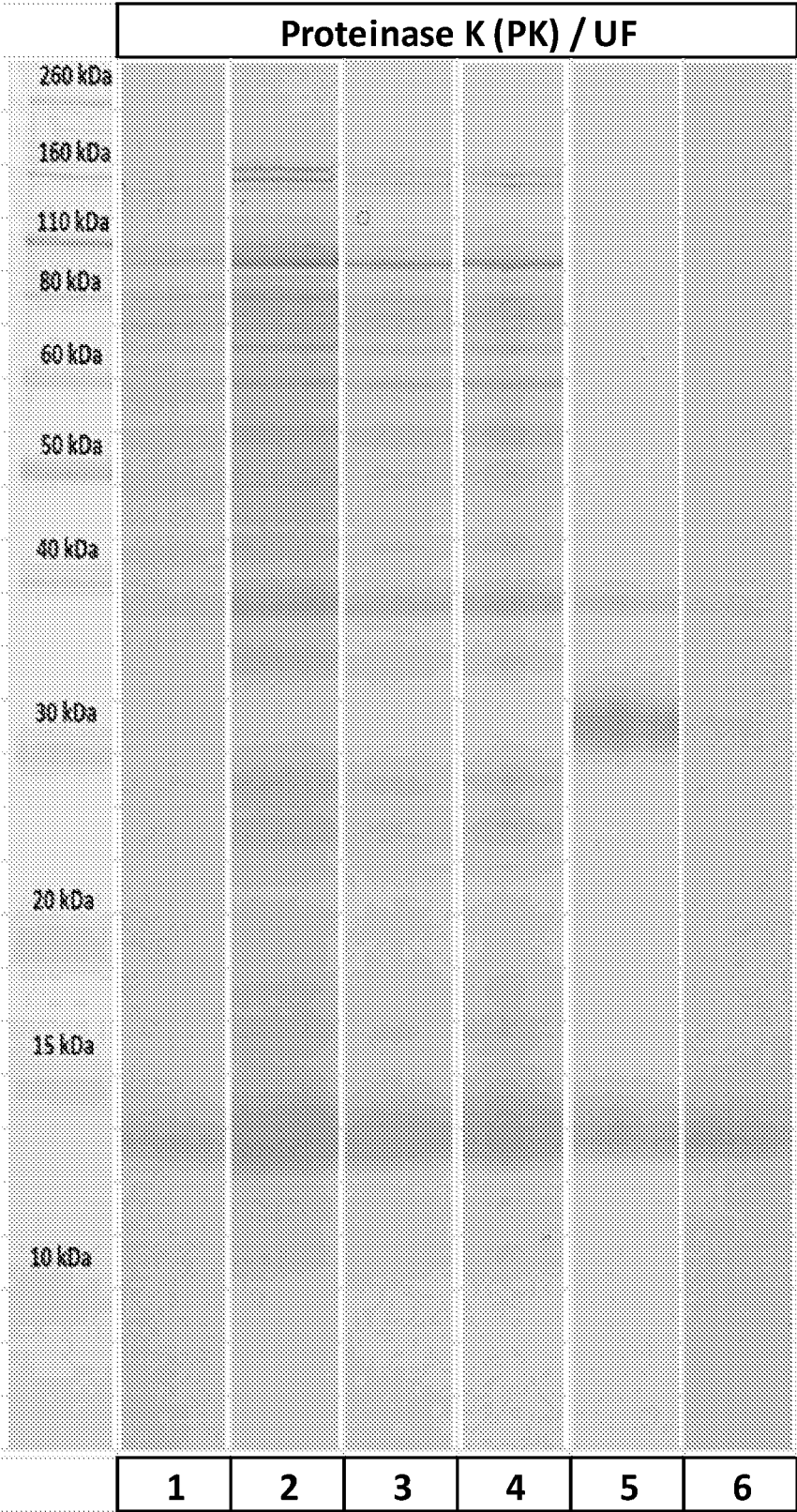


FIG. 2

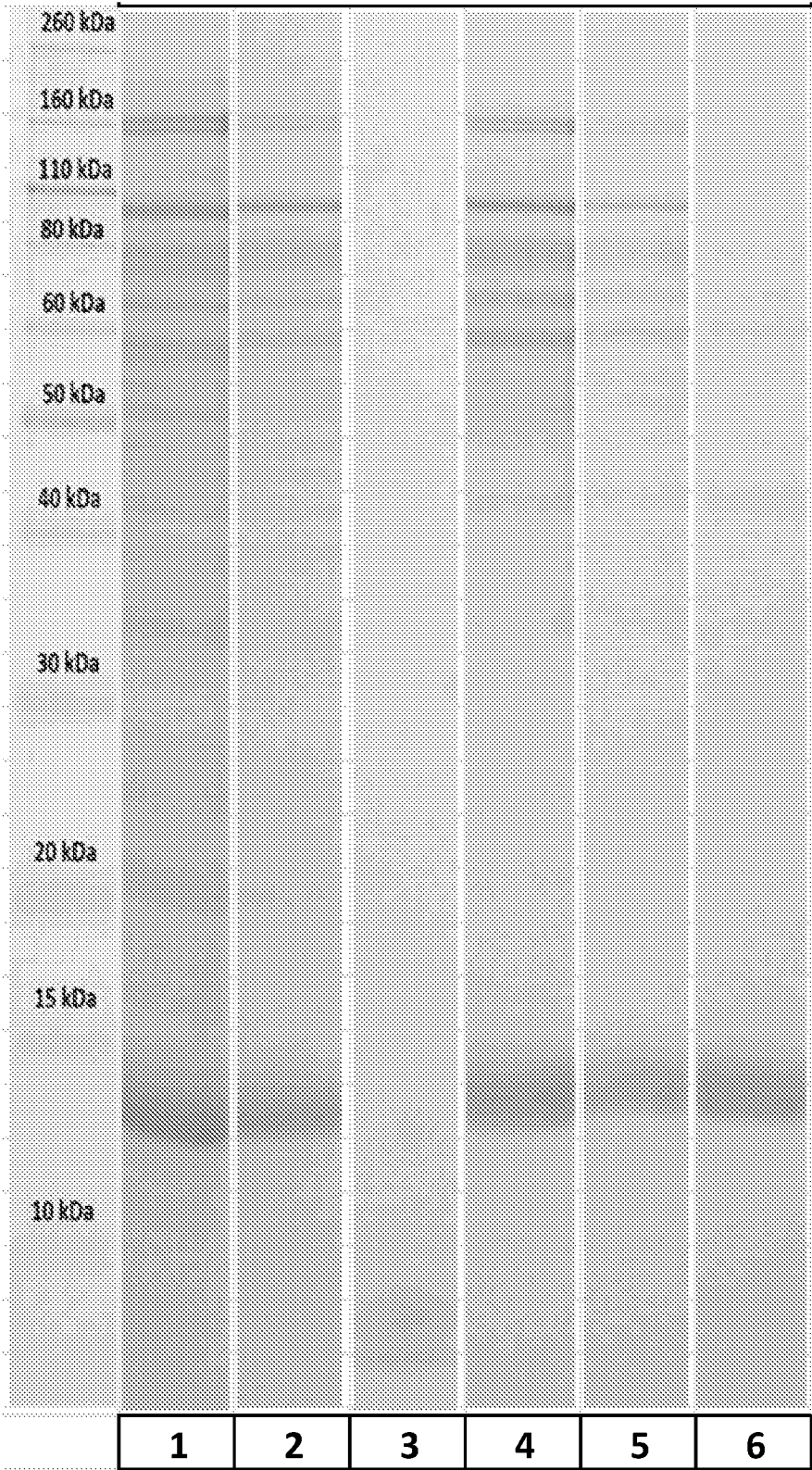


FIG. 3

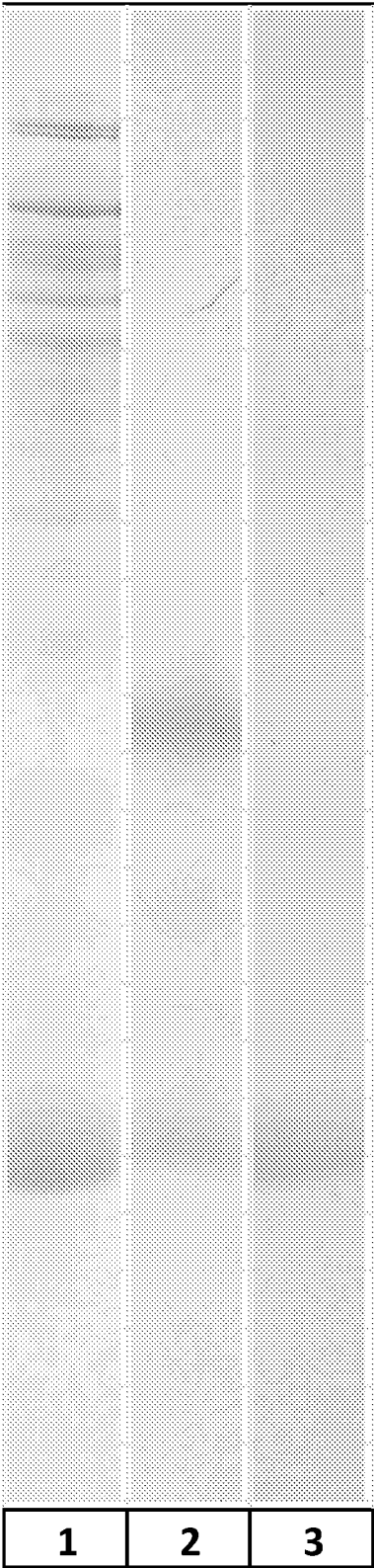


FIG. 4

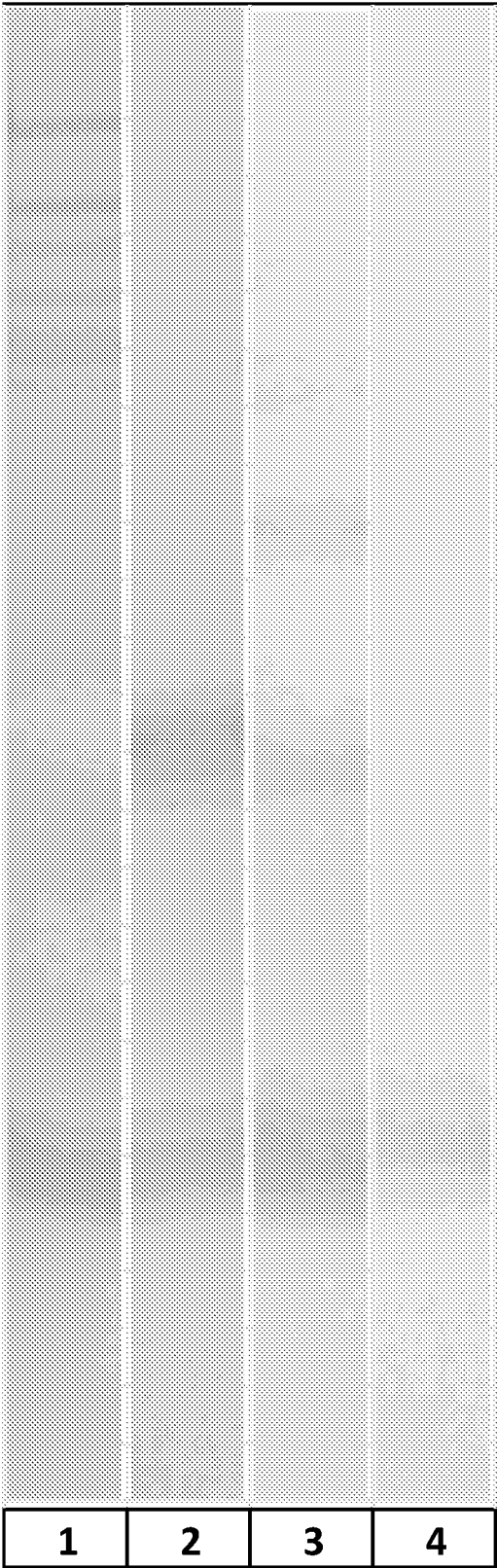
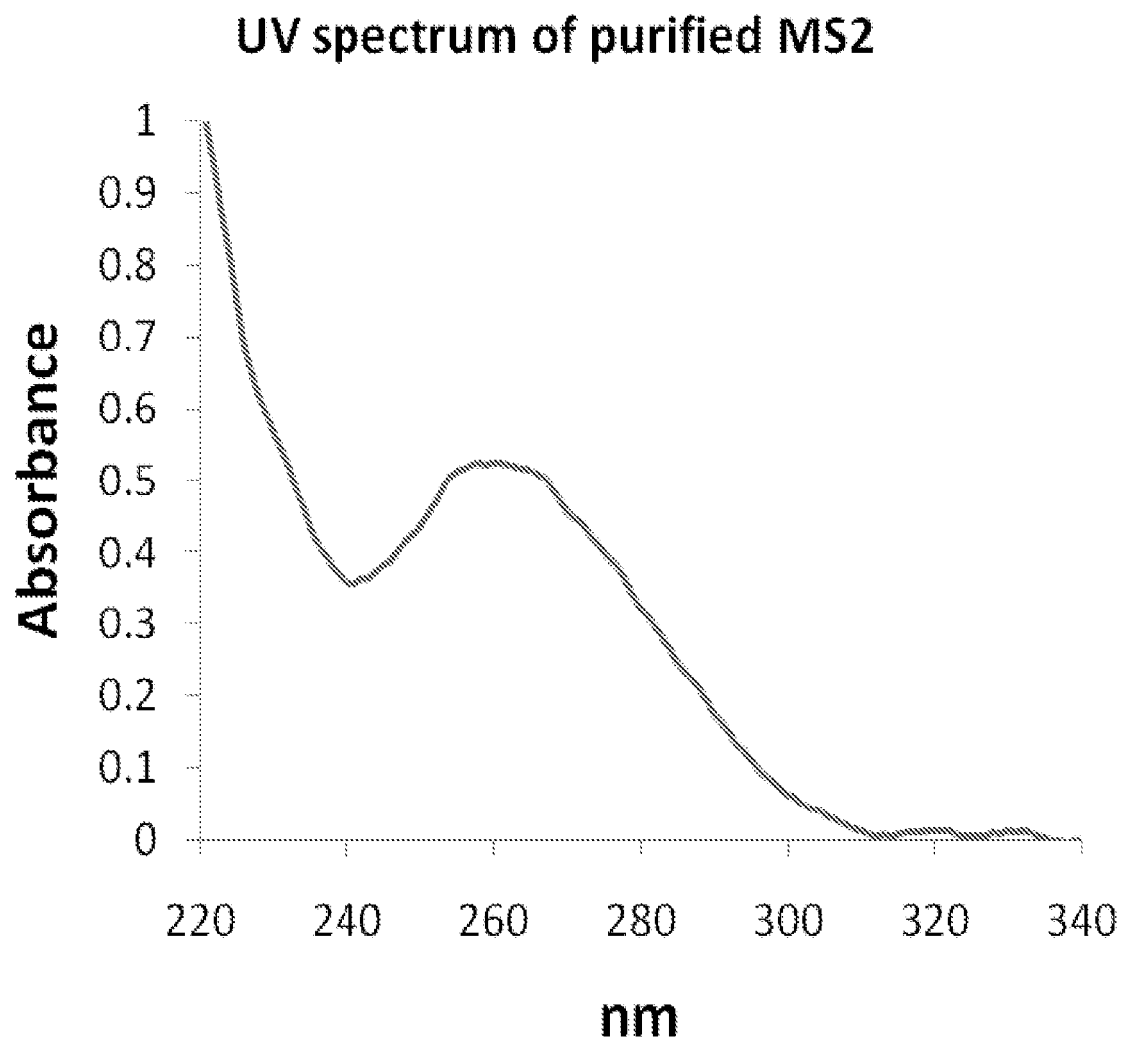
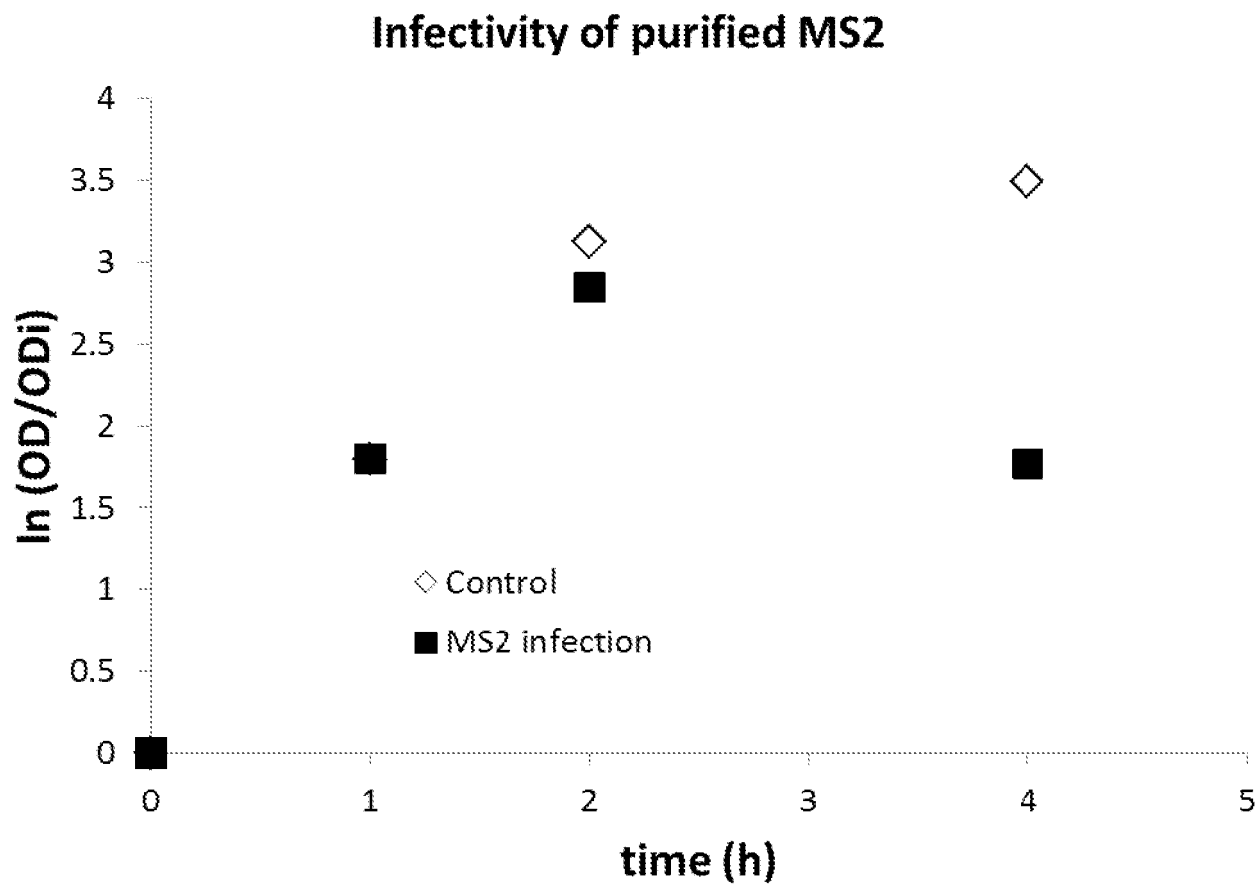


FIG. 5

6/34

**FIG. 6**

7/34

**FIG. 7**

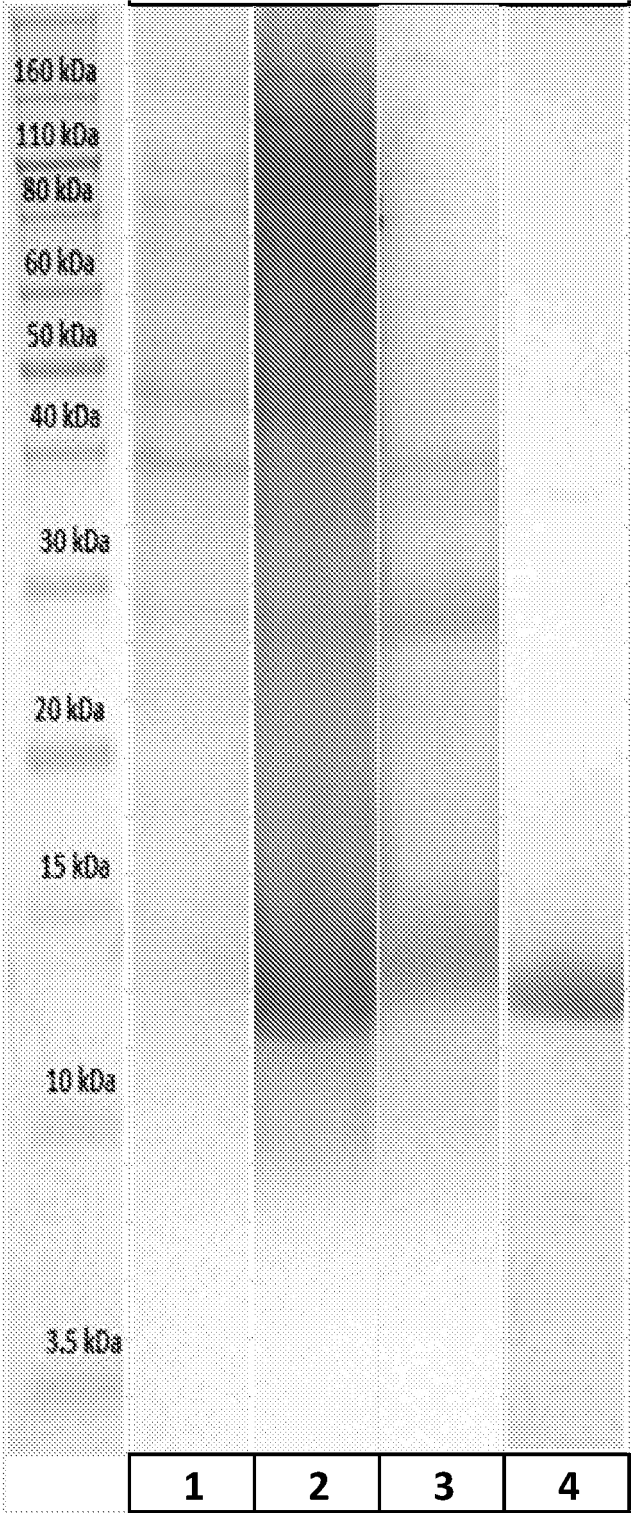


FIG. 8

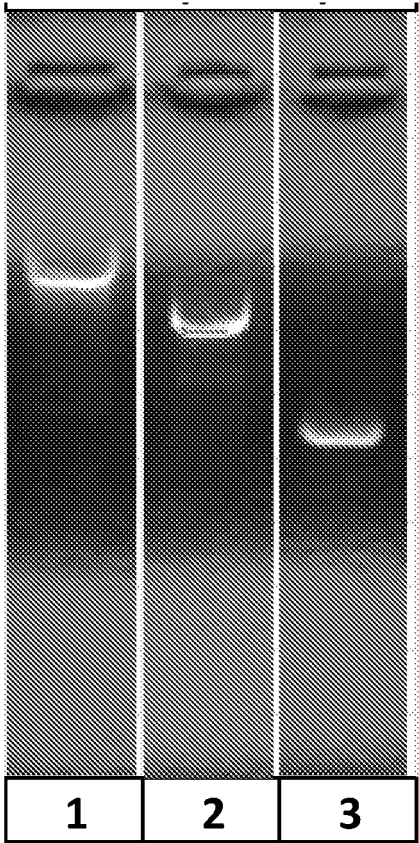


FIG. 9

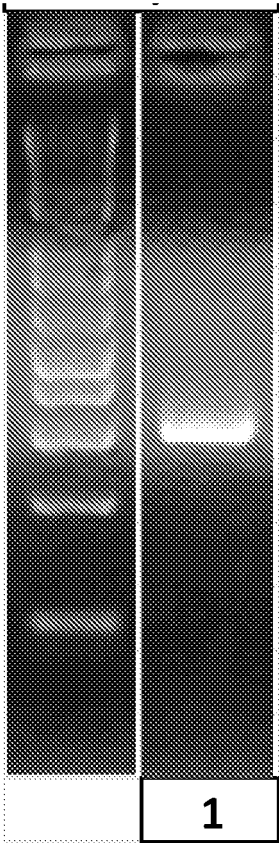


FIG. 10

10/34

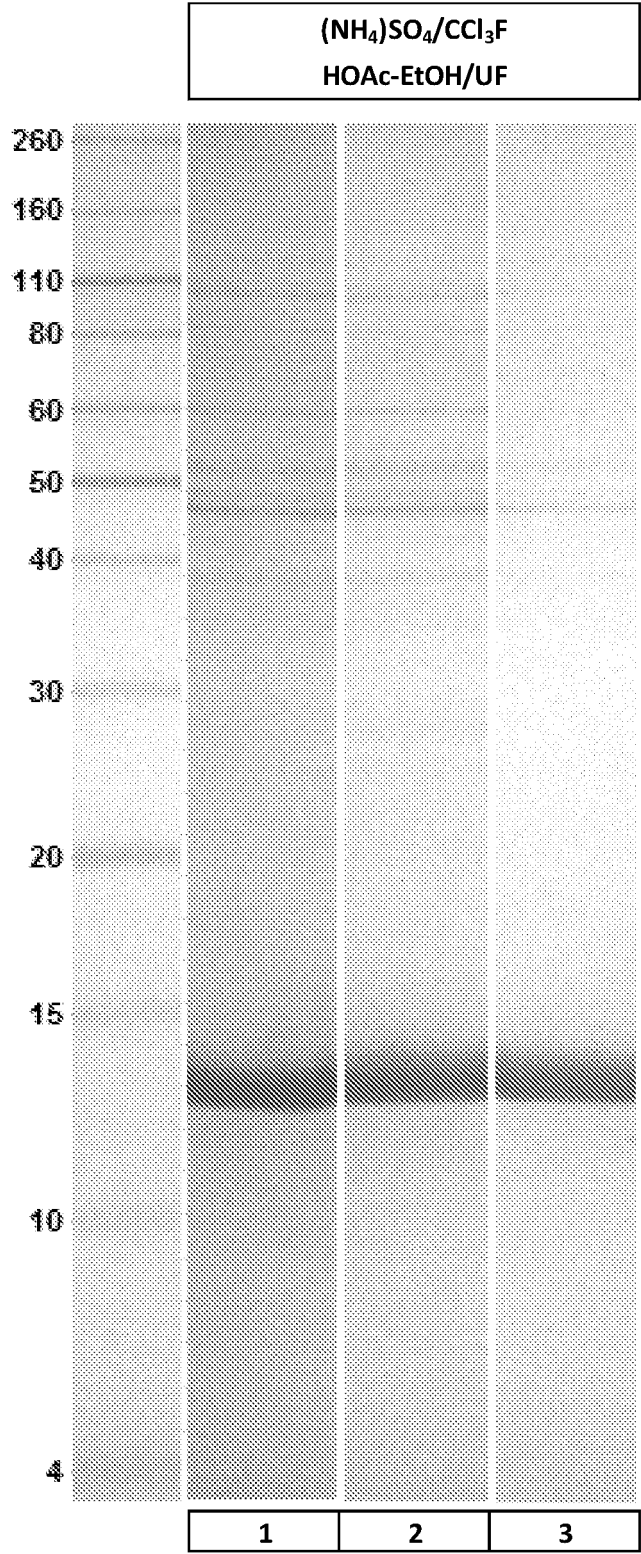


FIG. 11

11/34

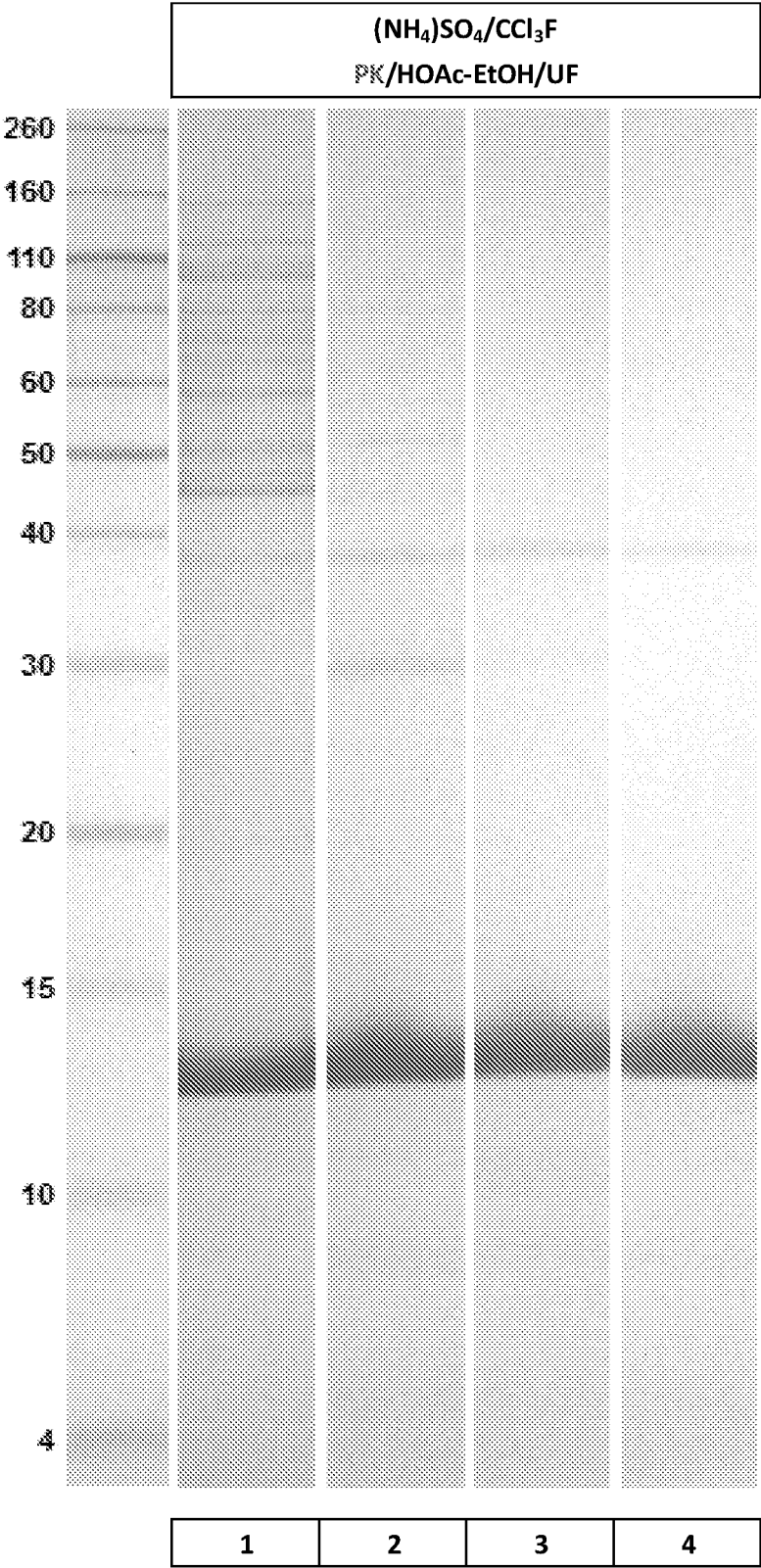
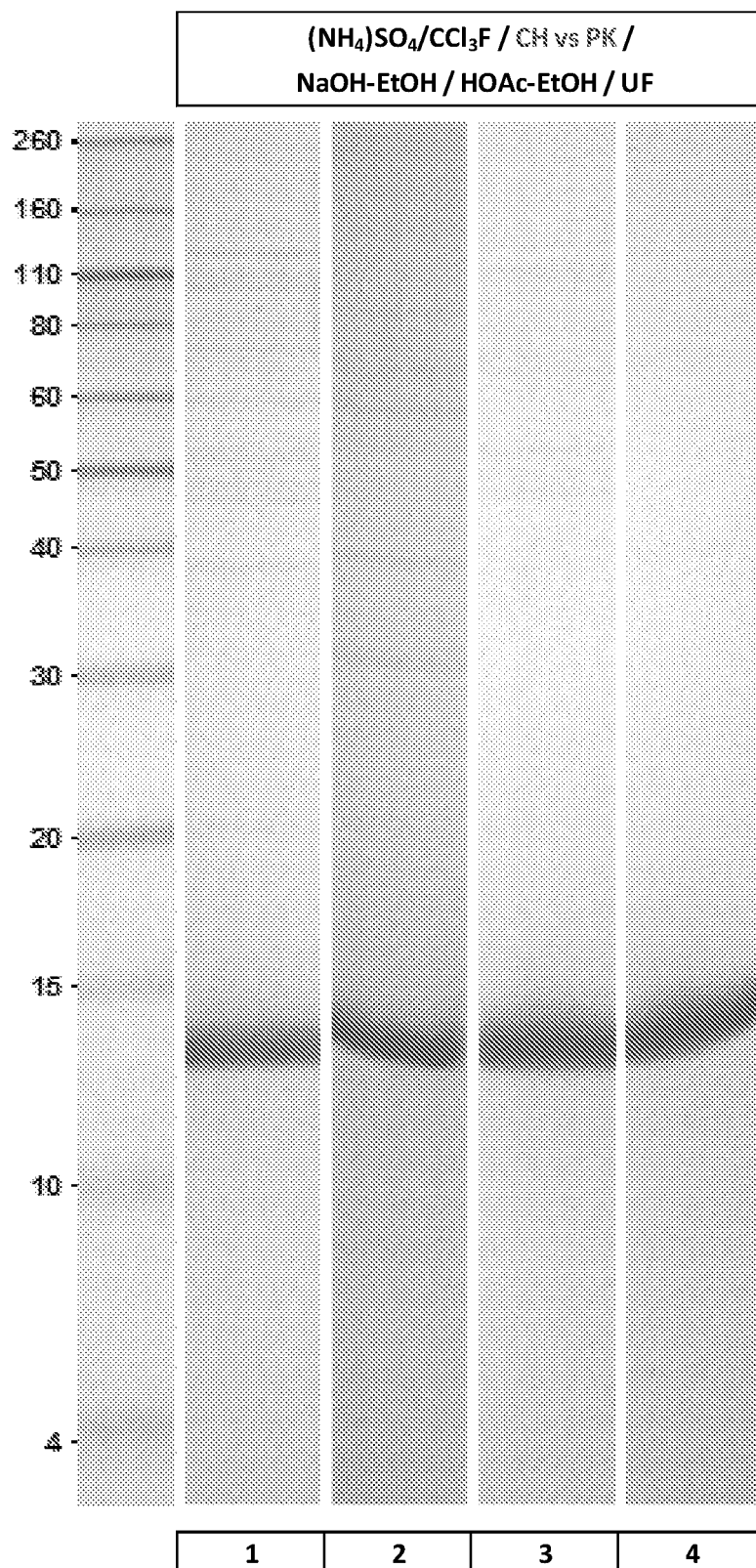


FIG. 12

12/34

**FIG. 13**

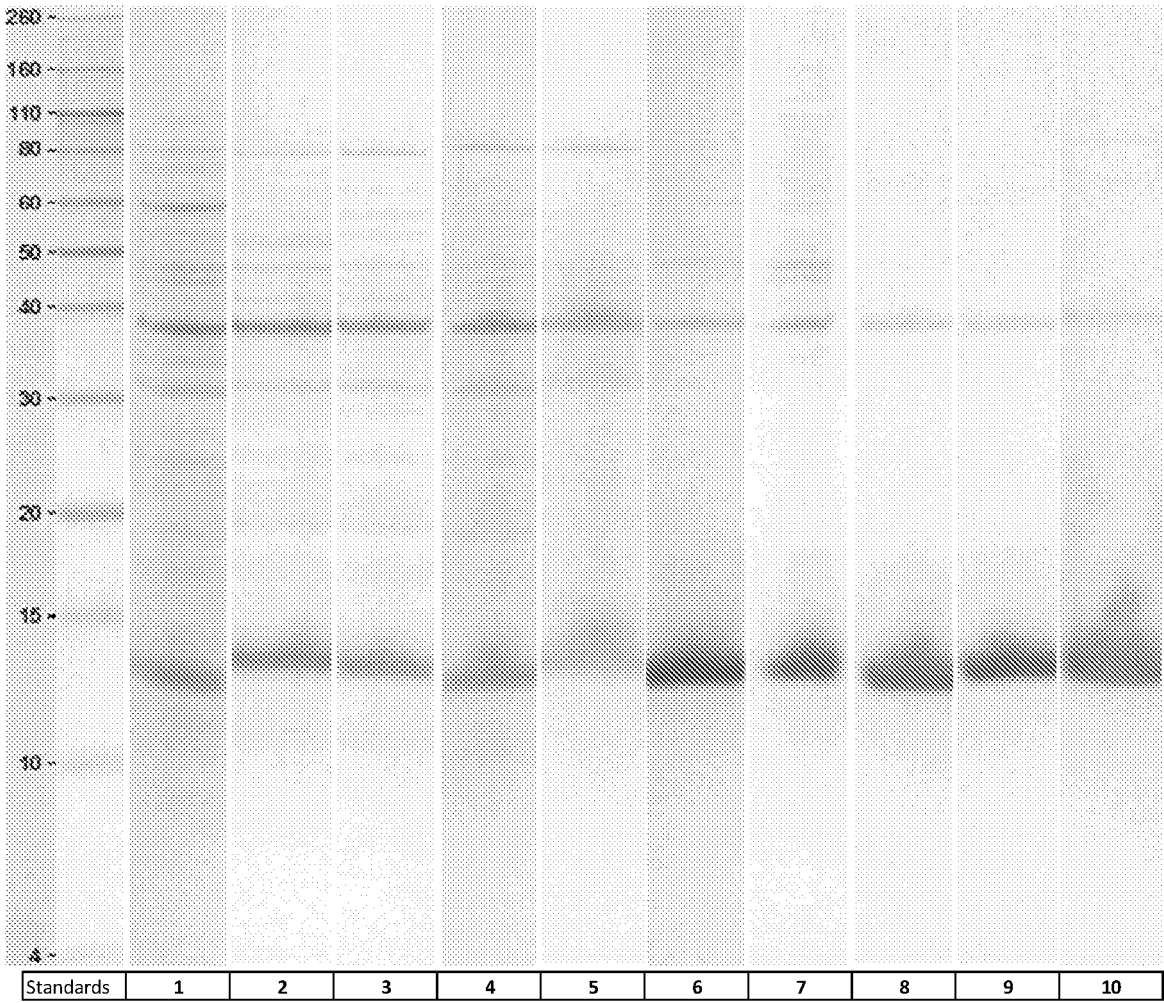


FIG. 14

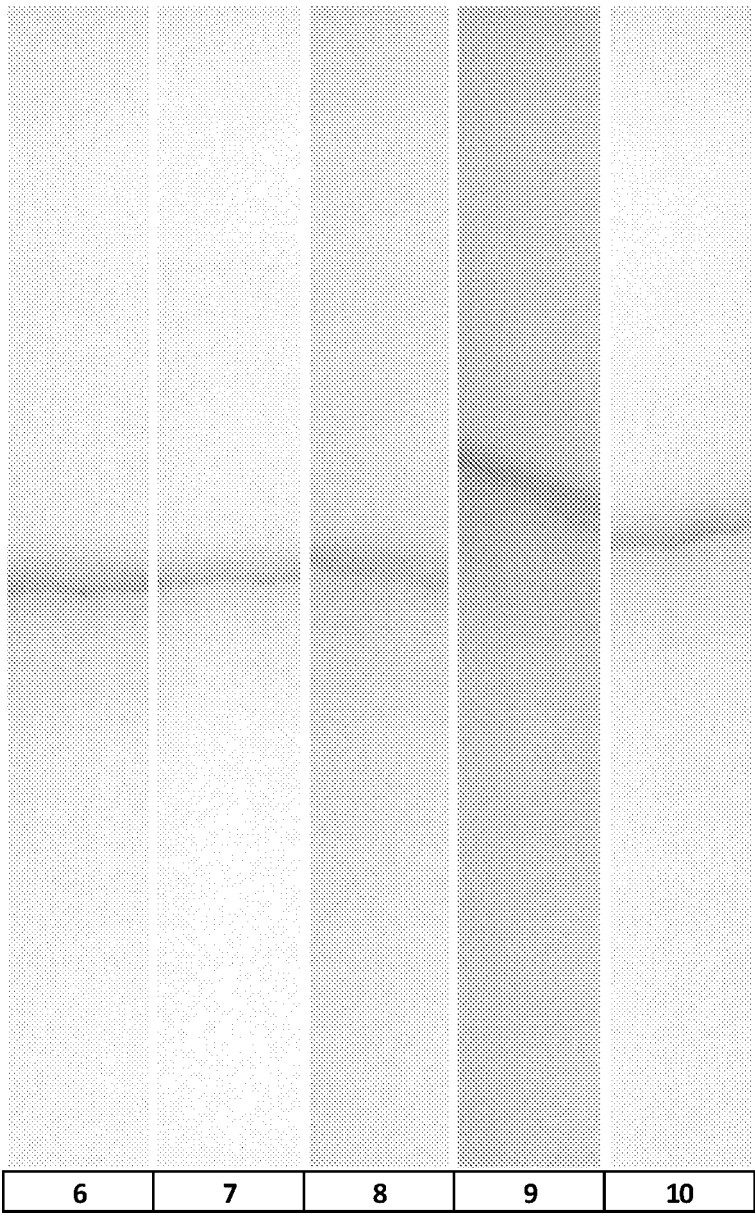


FIG. 15

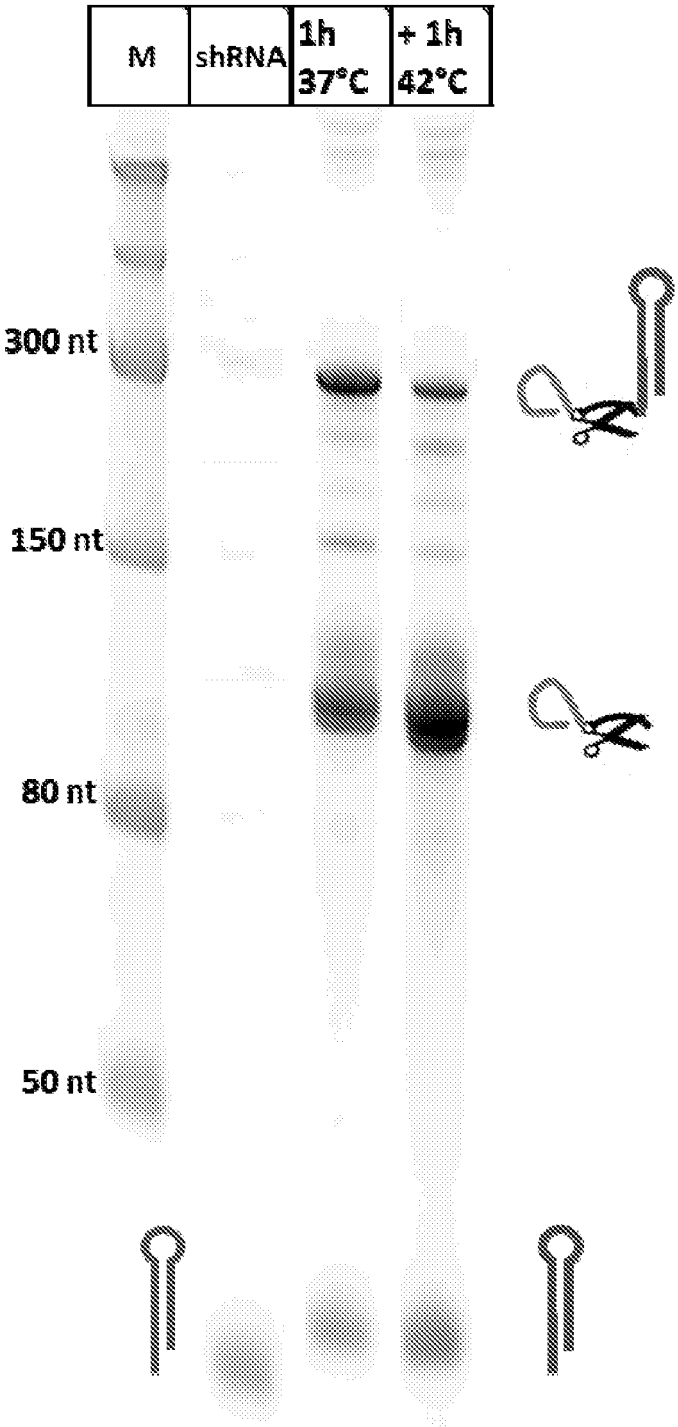


FIG. 16

16/34

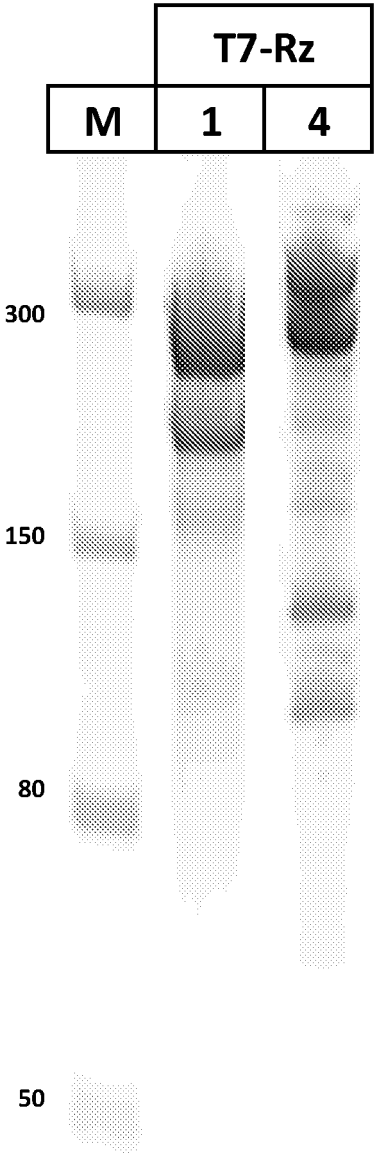


FIG. 17

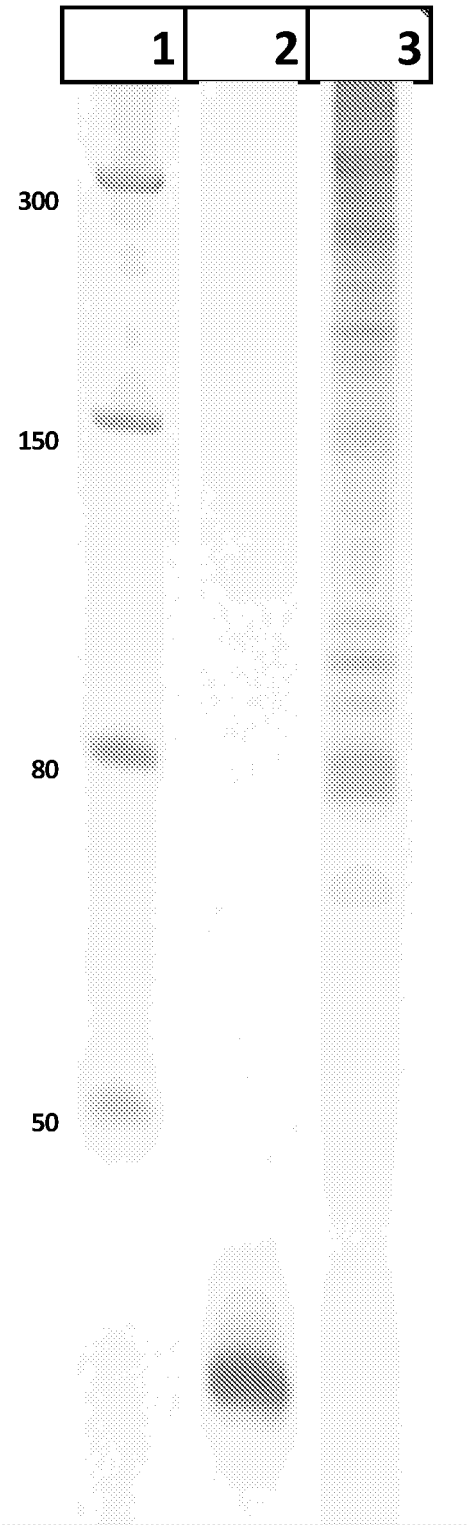
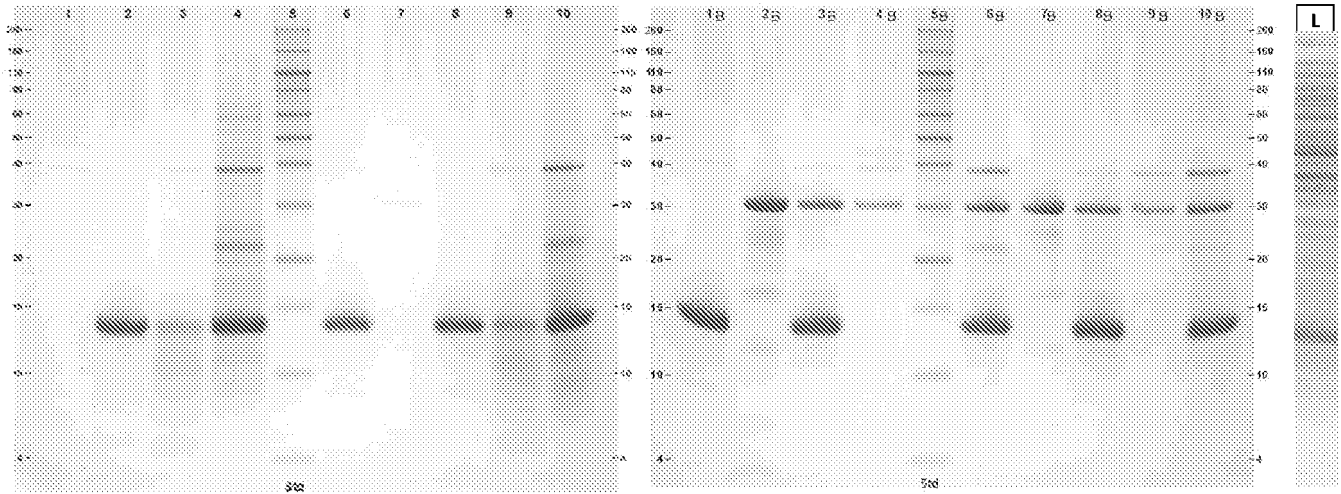


FIG. 18

18/34

**FIG. 19**

19/34

selected enterobacteria phage MS2 capsid protein alignments

```

123456789*123456789*123456789*123456789*123456789*1234
ASNFTQFVLVDNGGTGDTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAQ SEQ-ID 3 ms2
ASNFTQFVLVDNGGTGDTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAQ taxid 12022 ms2 T59S (1AQ3)
-ATLRSFVLVDNGGTGNTVVPVSNANGVAEWLSNNSRSQAYRVTASYRASGAD taxid 12018 ga
-ATLRSFVLVDNGGTGNTVVPVSNANGVAEWLSNNSRSQAYRVTASYRASGAD taxid 12018 ga A59T G79V (1GAV)
ASNFEFVLVDNGGTGDKVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAN taxid 12017 fr (1FRS)
!h!! h ! ! !!! h ! h ! h ! ! h
xx xx x x x

```

60 70 80 90 100 110

```

56789*123456789*123456789*123456789*123456789*123456789*1234
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMEITIPFATNSDCELVKAMQGLLKD SEQ-ID 3 ms2
NRKYSIKVEVPKVATQTVGGVELPVAAWRSYLNMEITIPFATNSDCELVKAMQGLLKD taxid 12022 ms2 T59S (1AQ3)
KRKYAIKLEVPKIQTQVVGVELPGSAWKAYASIDLTIPIFAATDDVTVISKSLAGLKFV taxid 12018 ga
KRKYTIKLEVPKIQTQVVGVELPVAWKAYASIDLTIPIFAATDDVTVISKSLAGLKFV taxid 12018 ga A59T G79V (1GAV)
NRKYTVKVEVPKVATQVQGGVELPVAAWRSYNNMEITIPVFATNDDCALIVKALQGTFTK taxid 12017 fr (1FRS)
! !h h h! !!! !! h! !!!h h !!! !h ! !h !
x x xx x x x xx

```

```

56789*123456789*
GNPIPSAIAANSIY- SEQ-ID 3 ms2
GNPIPSAIAANSIY- taxid 12022 ms2 T59S (1AQ3)
GNPIAEAISSQSGFYA taxid 12018 ga
GNPIAEAISSQSGFYA taxid 12018 ga A59T G79V (1GAV)
GNPIATAIAANSIY- taxid 12017 fr (1FRS)
!! !!h h
x x

```

P07234 UniProt public domain data base sequence identifier for
levivirus enterobacteria phage GA viral coat protein

P03614 UniProt public domain data base sequence identifier for
levivirus enterobacteria phage FR viral coat protein

! nonhomologous amino acid substitution (SEQ ID NO :3, taxid 12018, taxid 12017)

h homologous amino acid substitution (sidechains have similar size
and electrostatic requirements)

homologous sets (Ser,Thr), (Val,Ile,Leu), (Asp,Asn,Glu,Gln),
(Phe,Tyr,Trp), (Asn,Gln,His), (Arg,Lys),
(Phe,Tyr,Val,Ile,Leu)

x calculated to be substitutable based on sidechain solvent accessibility,
hydrogen bonding requirements and backbone conformational constraints

FIG. 20

20/34

1	2	3	4	5	
123456789012345678901	234567890123456	789012345678901234			
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				D0U1D6 D0U1D6__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				G4WZU0 G4WZU0__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM2U4 COM2U4__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM2S8 COM2S8__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM212 COM212__BPMS2
ASNFTQFVLVDNGGTGDDAVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM1M2 COM1M2__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM2L4 COM2L4__BPMS2
ASNFTKXFLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM220 COM220__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				Q2V0S8 Q2V0S8__BPBO1
ASNFTQFVLVDNDGTTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM216 COM216__BPMS2
ASNFTQFVLVDNGGTGKVTVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM1Y0 COM1Y0__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				D0U1E4 D0U1E4__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM309 COM309__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM325 COM325__BPMS2
ASNFTQFVLVDNGGTGDTVXX----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				Q9T1C7 Q9T1C7__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM2Z1 COM2Z1__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM1N8 COM1N8__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				J9QBW2 J9QBW2__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				C8XPC9 C8XPC9__BPMS21
--NFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM2Y4 COM2Y4__BPMS21
ASNFTQFVLVDNGGTGNVTVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				P69171 COAT__BPZE
ASNFTQFVLVDNGGTGNVTVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				P69170 COAT__BPR171
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				P03612 COAT__BPMS21
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				COM1L4 COM1L4__BPMS21
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				C8XPD7 C8XPD7__BPMS2
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				Q2V0T1 Q2V0T1__BPER1
ASNFTFVLVDNGGTGNVTVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				Q9MCD7 Q9MCD7__BPJP51
ASNFTQFVLVDNGGTGNVTVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				P03611 COAT__BPF21
ATLRSFVLVDNGGTGDTVTVV----	PVSNANGVAEWLSEN-SRSQAYRVTASYRASGAD				P34700 COAT__BPJF31
ATLRSFVLVDNGGTGDTVTVV----	PVSNANGVAEWLSEN-SRSQAYRVTASYRASGAD				Q2V0U0 Q2V0U0__BEB211
ATLRSFVLVDNGGTGNVTTVV----	PVSNANGVAEWLSEN-SRSQAYRVTASYRASGAD				Q2V0T7 Q2V0T7__BEB211
ATLRSFVLVDNGGTGNVTTVV----	PVSNANGVAEWLSEN-SRSQAYRVTASYRASGAD				Q9MBL2 Q9MBL2__BPKU11
ATLRSFVLVDNGGTGNVTTVV----	PVSNANGVAEWLSEN-SRSQAYRVTASYRASGAD				P07234 COAT__BPGAL
ATLRSFVLVDNGGTGNVTTVV----	PVSNANGVAEWLSEN-SRSQAYRVTASYRASGAD				C8YJG7 C8YJG7__BPE211
ATLRSFVLVDNGGTGNVTTVV----	PVSNANGVAEWLSEN-SRSQAYRVTASYRASGAD				C8YJH1 C8YJH1__BPE211
ATLRSFVLVDNGGTGNVTTVV----	PVSNANGVAEWLSEN-SRSQAYRVTASYRASGAD				C8YJH5 C8YJH5__BPE21
ATLRSFVLVDNGGTGNVTTVV----	PVSNANGVAEWLSEN-SRSQAYRVTASYRASGAD				Q2V0T4 Q2V0T4__BPTH11
ATLRSFVLVDNGGTGNVTTVV----	PVSNANGVAEWLSEN-SRSQAYRVTASYRASGAD				Q2V0U3 Q2V0U3__BPE211
ASNFEFVLVDNGGTGDKVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAN				P03614 COAT__BPR1
ASNFTQFVLVDNGGTGDTVVA----	PSNFANGVAEWISSN-SRSQAYKVTCSVRQSSAQ				ef108465 (GenBank)
****	**** *	***** *	***** *	***** *	***** *
xxx	x				x

FIG. 21A

lewi	6	7	8	9	100	110
56789012345678901234567890123456789012345678901234						
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNLELTIPIFATNSDCELI VKAMQGLLED					G4WZU0	G4WZU0__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNLELTIPIFATNSDCELI VKAMQGLLED					D0U1D6	D0U1D6__BPMS256
NRKYTIKVELPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM2U4	COM2U4__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM2S8	COM2S8__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM212	COM212__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM1M2	COM1M2__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM2L4	COM2L4__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM220	COM220__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					Q2V0S8	Q2V0S8__BPB01
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM216	COM216__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM1Y0	COM1Y0__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNLELTIPIFATNSDCELI VKAMQGLLED					D0U1E4	D0U1E4__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM309	COM309__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM325	COM325__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					Q9T1C7	Q9T1C7__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM2Z1	COM2Z1__BPMS2
NRKYTIKXVVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM1N8	COM1N8__BPMS256
NRKYTIKVEVPKVATQTVGGVQLPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					J9QBW2	J9QBW2__BPMS256
NRKYTIKVEVPKVATQTVGGVQLPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					C8XPC9	C8XPC9__BPMS253
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM2Y4	COM2Y4__BPMS255
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					P69171	COAT__BPB255
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					P69170	COAT__BPB1756
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					P03612	COAT__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					COM1L4	COM1L4__BPMS2
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					C8XPD7	C8XPD7__BPMS256
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					Q2V0T1	Q2V0T1__BPB256
NRKYTIKVAVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					Q9MCD7	Q9MCD7__BPB255
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNLELTIPIFATNSDCELI VKAMQGLLED					P03611	COAT__BPB255
KRKYTIKLBVPKIVTQVVNGVELPVSAMKAYASIDLTIPIFAATDDVTVISKSLAGLFXV					P34700	COAT__BPB255
KRKYTIKLBVPKIVTQVVNGVELPVSAMKAYASIDLTIPIFAATDDVTVISKSLAGLFXV					Q2V0U0	Q2V0U0__BPB2155
KRKYTIKLBVPKIVTQVVNGVELPVSAMKAYASIDLTIPIFAATDDVTVISKSLAGLFXV					Q2V0T7	Q2V0T7__BPB21
KRKYTIKLBVPKIVTQVVNGVELPVSAMKAYASIDLTIPIFAATDDVTVISKSLAGLFXV					Q9MBL2	Q9MBL2__BPB2155
KRKYTIKLBVPKIVTQVVNGVELPVSAMKAYASIDLTIPIFAATDDVTVISKSLAGLFXV					P07234	COAT__BPB255
KRKYTIKLBVPKIVTQVVNGVELPVSAMKAYASIDLTIPIFAATDDVTVISKSLAGLFXV					C8YJG7	C8YJG7__BPB2155
KRKYTIKLBVPKIVTQVVNGVELPVSAMKAYASIDLTIPIFAATDDVTVISKSLAGLFXV					C8YJH1	C8YJH1__BPB2155
KRKYTIKLBVPKIVTQVVNGVELPVSAMKAYASIDLTIPIFAATDDVTVISKSLAGLFXV					C8YJH5	C8YJH5__BPB2155
KRKYTIKLBVPKIVTQVVNGVELPVSAMKAYASIDLTIPIFAATDDVTVISKSLAGLFXV					Q2V0T4	Q2V0T4__BPB2155
KRKYTIKLBVPKIVTQVVNGVELPVSAMKAYASIDLTIPIFAATDDVTVISKSLAGLFXV					Q2V0U3	Q2V0U3__BPB2155
NRKYTVFVEVPKVATQVQGGVELPVAAWRSYMNMBELTIPIFATNSDCELI VKALQCTFFET					P03614	COAT__BPB2
NRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMBELTIPIFATNSDCELI VKAMQGLLED					ef108465	(GenBank)
***	**	**	**	**	****	* * * * *
	x	x	xx	x		xx

FIG. 21B

levi

120	130	
5678901234567890		
GNPIPSAIAANSGLY-	G4WZU0	G4WZU0_BPMS2116
GNPIPSAIAANSGLY-	DOU1D6	DOU1D6_BPMS2116
GNPIPSAIAANSGLY-	COM2U4	COM2U4_BPMS2116
GNPIPSAIAANSGLY-	COM2S8	COM2S8_BPMS2116
GNPIPSAIAANSGLY-	COM212	COM212_BPMS2116
GNPIPSAIAANSGLY-	COM1M2	COM1M2_BPMS2116
GNPIPSAIAANSGLY-	COM2L4	COM2L4_BPMS2116
GNPIPSAIAANSGLY-	COM220	COM220_BPMS2116
GNPIPSAIAANSGLY-	Q2V0S8	Q2V0S8_BPBO1116
GNPIPSAIAANSGLY-	COM216	COM216_BPMS2116
GNPIPSAIAANSGLY-	COM1Y0	COM1Y0_BPMS2116
GNPIPSAIAANSGLY-	DOU1E4	DOU1E4_BPMS2116
GNPIPSAIAANSGLY-	COM309	COM309_BPMS2116
GNPIPSAIAANSGLY-	COM325	COM325_BPMS2116
GNPIPSAIAANSGLY-	Q9T1C7	Q9T1C7_BPMS2116
GNPIPSAIAANSGLY-	COM2Z1	COM2Z1_BPMS2116
GNPIPSAIAANSGLY-	COM1N8	COM1N8_BPMS2116
GNPIPSAIAANSGLY-	J9QBW2	J9QBW2_BPMS2116
GNPIPSAIAANSGLY-	C8XPC9	C8XPC9_BPMS2113
GNPIPSAIAANSGLY-	COM2Y4	COM2Y4_BPMS2115
GNPIPSAIAANSGLY-	P69171	COAT_BPZR115
GNPIPSAIAANSGLY-	P69170	COAT_EPR17116
GNPIPSAIAANSGLY-	P03612	COAT_BPMS2
GNPIPSAIAANSGLY-	COM1L4	COM1L4_BPMS2116
GNPIPSAIAANSGLY-	C8XPD7	C8XPD7_BPMS2116
GNPIPSAIAANSGLY-	Q2V0T1	Q2V0T1_BPZR116
GNPIPSAIAANSGLY-	Q9MCD7	Q9MCD7_BPJP5115
GNPIPSAIAANSGLY-	P03611	COAT_BFP2115
GNPIADA1SSQSGFYA	P34700	COAT_BPJF3115
GNPIADA1SSQSGFYA	Q2V0U0	Q2V0U0_BPBZ1115
GNPIADA1SSQSGFYA	Q2V0T7	Q2V0T7_BPBZ1115
GNPVADA1SSQSGFYA	Q9MBL2	Q9MBL2_BPKV1115
GNPIABA1SSQSGFYA	P07234	COAT_BPGA115
GNPIABA1SSQSGFYA	C8YJG7	C8YJG7_BPBZ1115
GNPVADA1SSQSGFYA	C8YJH1	C8YJH1_BPBZ1115
GNPIADA1SSQSGFYA	C8YJH5	C8YJH5_BPBZ1115
GNPIADA1SSQSGFYA	Q2V0T4	Q2V0T4_BPTH1115
GNPIADA1SSQSGFYA	Q2V0U3	Q2V0U3_BPBZ1116
GNPIATAIAANSGLY-	P03614	COAT_BPFR116
GNPIPSAIAANSGLY-	ef108465	(GenBank, not UniProt)
*** ** *		
x	x	

FIG. 21C

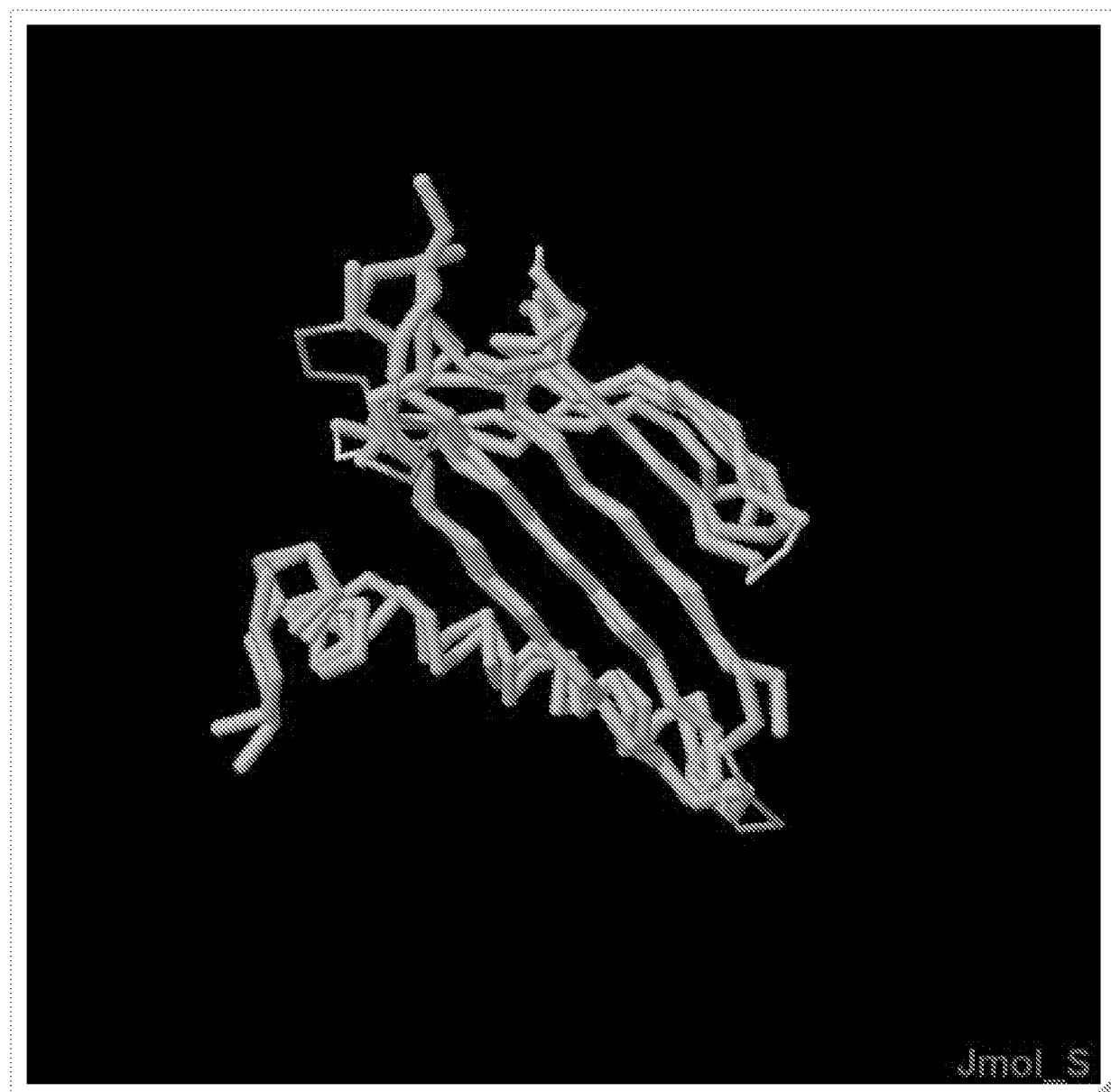


FIG. 22

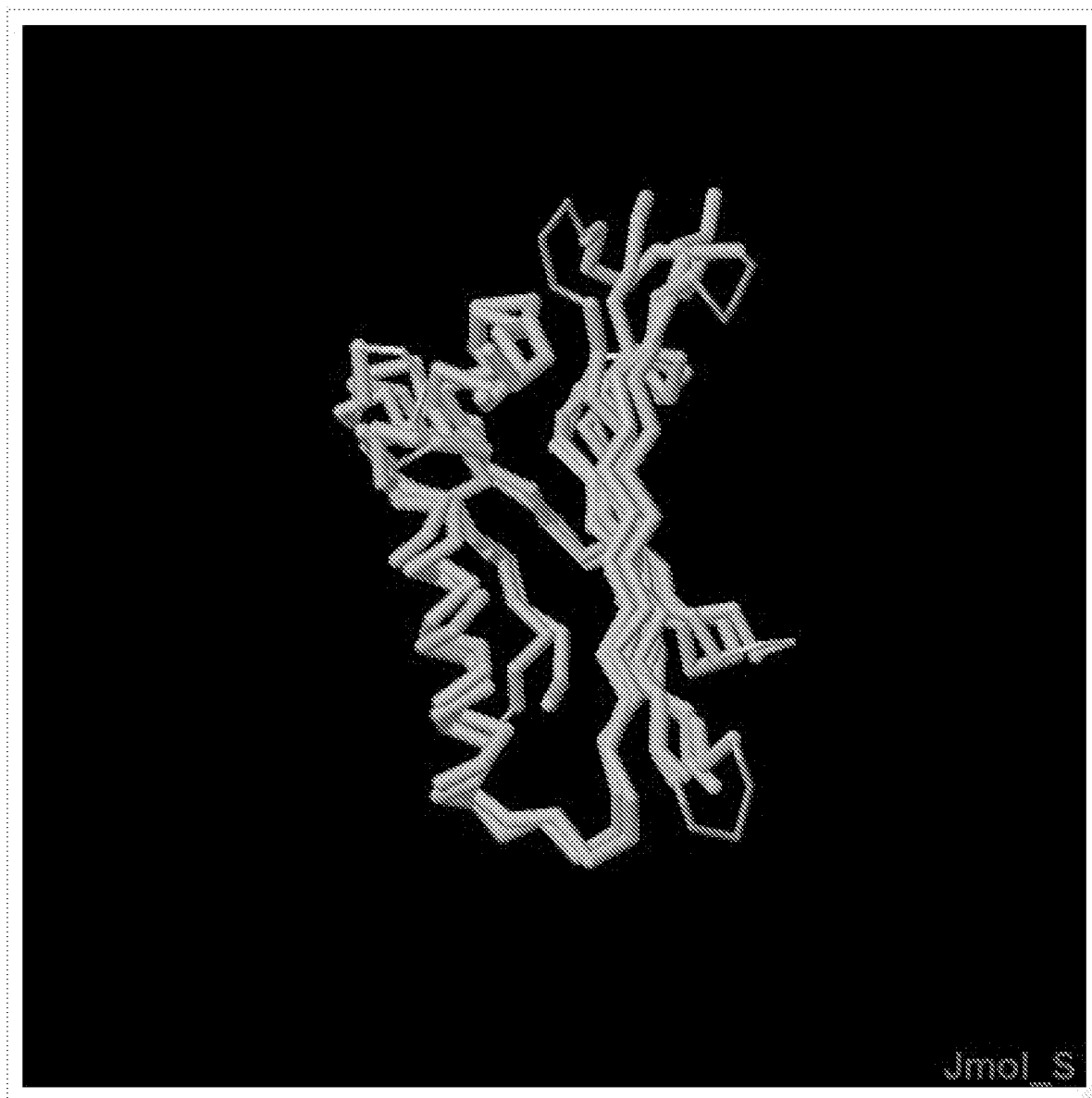


FIG. 23

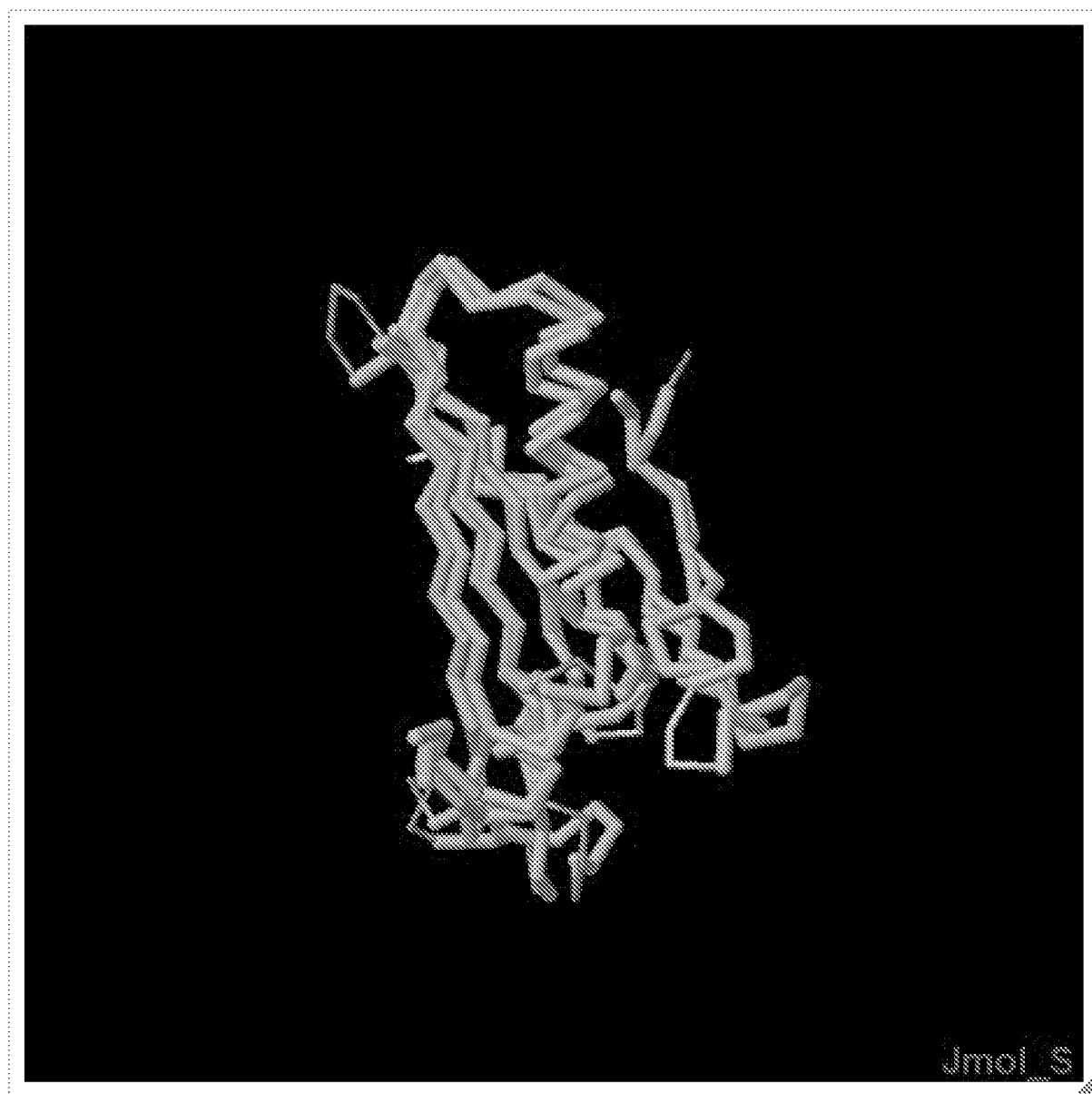


FIG. 24

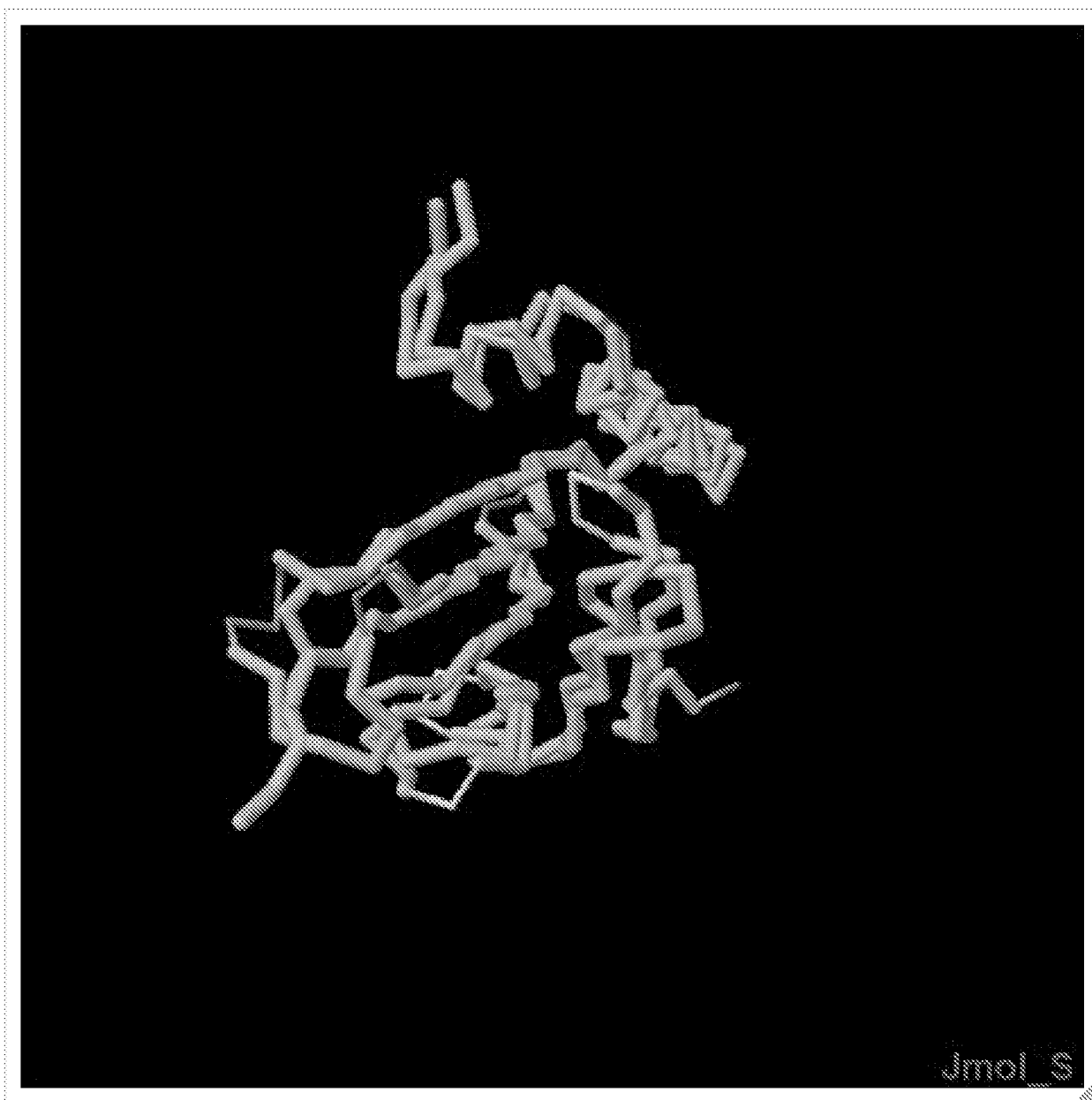


FIG. 25

27/34

structural (jFATCAT rigid) sequence alignment

```

      10      20      30      40      50
      ssss  tt  sssssss 33 3 ssss  ttt  sssssss
123456789*12 3456789*1234567 89*123456 789*123456789*123 4
MASNFTQFVLVDN--GGTGDVTVAPSNFAN--GVAEWISSN-SRSQAYKVTCSVRQSSA--Q
snftqFVLVDN--GGTGDVTVAPSNFAN--GVAEWISSN-SRSQAYKVTCSVRQSSA--Q
aSNFTQFVLVDN--GGTGDVTVAPsnfaN--GVAEWISSN-SRSQAYKVTCSVRQSSA--Q
asnftQFVLVDN--GGTGDVTVAPSNFAN--GVAEWISSN-SRSQAYKVTCSVRQSSA--Q
MAKLETVTLGNIKGDKGQTLVLNPRGVNPTNGVASLSQAGAVPALEKRVTVSVSQPsrnrK
***  **  **  *  *  *****  **  *  *****  **  *****
*  *  h  *  *  h  *  h  ***  h**  *  *  *

      60      70      80      90      100      110
      sssssssssss  tt  tt  333  ssssssssssstt  hhhhhhhhhhhhhhh t
56789*123456789*123456789*123456789*123456789*123456789*1234
NRKYTIKVEVPKvatqtvvggvelpvaaWRSYLNMEltiPIFATNSDCElIVKAMQGLLKD
NRKYTIKVEVPKvatqtvvggvelpvAAWRSYLNMEltiPIFATNSDCElIVKAMQGLLKD
NRKYTIKVEVPKvatQtvvggvelpvaaWrsylnMEltiPIfatNSDCElIVKAMQGLLKD
NRKYTIKVEVPKvatQTVGGVELPVAawRSYLNMEltiPIFATNSDCElIVKAMQGLLKD
NYKVQVKIQNPTACTan--gsCDPSVTRQAYADVTFSFTQYSTDEERAFVrTELAALLAS
*  *  *  **  *  ***  ***  *  *  ***  ***  *  *  ***  ***
*  *  *hh  *  *  *  *  *  h  hhh  h  *  h  hh  **

120
t hhhhhhhht 1aq3 DSSP
56789*123456789*123456789*
GNPIPSAIAANSgiY 1aq3:ms2 T59S
GNPIPSAIAANSgiy 2vtu:j vs 2vtu:j
GNPIPSAIAANSgiy 1aq3:a vs 2vtu:j
GNPIPSAIAANSgiY 1aq3:a vs 2vtu:l
P-LLIDAIDQLNPAY 1qbe:Qbeta
***  *****  conserved in alloviridae
h  **  *  1QBE vs SEQ NO 3

```

Residues whose backbone differ with respect to 1AQ3 are shown in lower case, a composite of all possible monomer-monomer comparisons. Lower case in the 1AQ3 sequence refers to backbone conformational differences between its 3 independently refined monomers.

DSSP secondary structure assignments (s=strand, h=helix, 3=

* conserved amino acid (SEQ ID NO :3)

h homologous amino acid substitution (sidechains have similar size and electrostatic requirements)

homologous sets (Ser,Thr), (Val,Ile,Leu), (Asp,Asn,Glu,Gln),
(Phe,Tyr,Trp), (Asn,Gln,His), (Arg,Lys),
(Phe,Tyr,Val,Ile,Leu)

x calculated to be substitutable based on sidechain solvent accessibility,
hydrogen bonding requirements and backbone conformational constraints

FIG. 26

complete alloverviviridae viral coat protein sequences retrieved from the UniProt database and aligned using their BLAST multiple alignment with default values for weighting array choice, gap penalties, etc.

sequence

```

-MAKLETVTTLGNIGKDGKQTLVLNPRGVNPTNGVASLSQAGAVPALEKRVTVSVSQPSRN P03615 COAT_BPQBE1
-MAKLQAITLSGIGKNGDVTLLNPRGVNPTNGVAALSEAGAVPALEKRVTVSVSQPSRN Q9T0R9 CAPSD_BPMX11
-MAKLNQVTLSKIGKNGDQTLTLTPRGVNPTNGVASLSEAGAVPALEKRVTVSVAQPSRN P09673 COAT_BPSP
-MAKLETVTLSNIGKDGKQTLVLNPRGVNPTNGVASLSEAGAVPALEKRVTVSVSQPSRN C8YJJ2 C8YJJ2_BPQBE1
-MAKLQAITLSGIGKKGDTVLDLNPGRVNPTNGVAALSEAGAVPALEKRVTVSVSQPSRN Q9T0S0 Q9T0S0_9VIRU1
-MAKLNKVTLTGIGKAGNQTLLTPRGVNPTNGVASLSEAGAVPALEKRVTVSVAQPSRN Q9T0R8 Q9T0R8_9VIRU1
-MAKLETVTLSNIGKDGKQTLVLNPRGVNPTNGVASLSEAGAVPALEKRVTVSVSQPSRN C8YJI8 C8YJI8_BPQBE1
-MAKLETVTLSNIGKDGKQTLVLNPRGVNPTNGVASLSEAGAVPALEKRVTVSVSQPSRN C8YJJ6 C8YJJ6_BPQBE1
-MAKLETVTLSNIGKDGQTLVLNPRGVNPTNGVAALSQAGAVPALEKRVTVSVSQPSRN C8YJI0 C8YJI0_BPQBE1
-MAKLNVTTLTKLGKEANKTMTLTPRGVNPTNGVATLSEAGAVPALEKRVTVSVAQPSRN A0T121 A0T121_9VIRU1
-MAKLNQVTTLTKLGKAGDQTLTLTPRGVNPTNGVASLSEAGAVPALEKRVTVSVAQPSRN C8YL75 C8YL75_9VIRU
-MAKLNQVTLSKIGKNGDQTLTLTPRGVNPTNGVASLSEAGAVPALEKRVTVSVAQPSRN C8YL83 C8YL83_9VIRU1
-MAKLNKVTLTGIGKAGDQTLTLTPRGVNPTNGVASLSEAGAVPALEKRVTVSVAQPSRN C8YL79 C8YL79_9VIRU1
-MAKLNQVTLSKLGKNGDQTLTLTPRGVNPTNGVASLSEAGAVPALEKRVTVSVAQPSRN C8YL87 C8YL87_9VIRU1
-MAKLNQVTLSKIGKNGDQTLTLTPRGVNPTNGVASLSEAGAVPALEKRVTVSVAQPSRN G4WZS5 G4WZS5_BPSP1
-MAKLNQVTLSKIGKNGDQTLTLTPRGVNPTNGVASLSEAGAVPALEKRVTVSVAQPSRN G4WZS9 G4WZS9_BPSP1
-MAKLETVTTLGNIGKDGKQTLVLNPRGVNPTNGVSSLSQAGAVPALEKRVTVSVSQPSRN G4WZR7 G4WZR7_BPQBE1
-MAKLETVTTLGNIGKDGKQTLVLNPRGVNPTNGVASLSQAGAVPALEKRVTVSVSQPSRN Q774G0 Q774G0_BPQBE
-MAKLNQVTLSKIGKNGDQTLTLTPRGVNPTNGVASLSEAGAVPALEKRVTVSVAQPSRN D0U1H3 D0U1H3_BPSP1
-MAKLETVTTLGNIGKDGKQTLVLNPRGVNPTNGVASLSQAGAVPALEKRVTVSVSQPSRN D0U1F3 D0U1F3_BPQBE1
-MAKLNQVTLSKIGKNGDQTLTLTPRGVNPTNGVASLSEAGAVPALEKRVTVSVAQPSRN D0U1H7 D0U1H7_BPSP1
-MAKLNQVTLSKIGKNGDQTLTLTPRGVNPTNGVASLSEAGVPALEKRVTVSVAQPSRN D0U1G9 D0U1G9_BPSP
***h h** h** * * ***** **h** *****h** *****

```

FIGURE 27A

RKNYKVQVKIQNPTACTANGSCDPSVTRQAYADVTFSTQYSTDEERAFVRTELAALLAS	P03615	COAT_BPQBE60
RKNYKVQVKIQNPTSCTASGTCPSVTRSAFADVTFSFTQYSTDEERALVRTELKALLAD	Q9T0R9	CAPSD_BPMX160
RKNFKVQIKLQNPACTR-DACDPSVTRSAFADVTLSFTSYSTDEERALIRTELAALLAD	P09673	COAT_BPSP60
RKNYKVQVKIQNPTACSANGSCDPSVTRQAYADVTFSTQYSTDEERAFVRTELAALLAG	C8YJJ2	C8YJJ2_BPQBE60
RKNYKVQVKIQNPTSCTASGTCPSVTRSAFADVTFSFTQYSTVEERALVRTELQALLAD	Q9T0S0	Q9T0S0_9VIRU60
RKNYKVQIKLQNPACTK-DACDPSVTRSGSRDVTLSFTSYSTERERALIRTELAALLKD	Q9T0R8	Q9T0R8_9VIRU
RKNYKVQVKIQNPTACPANGSCDPSVTRQAYADVTFSTQYSTDEERAFVRTELAALLAG	C8YJI8	C8YJI8_BPQBE60
RKNYKVQVKIQNPTACSANGSCDPSVTRQAYADVTFSTQYSTDEERAFVRTELAALLAD	C8YJJ6	C8YJJ6_BPQBE60
RKNYKVQVKIQNPTACTANGSCDPSVTRQAYADVTFSTQYSTDEERAFVRTELIALLAS	C8YJI0	C8YJI0_BPQBE60
RKNFKVQIKLQNPACTK-DACDPSVTRSAFADLTLSFTSYSTDEERALIRTELAALLAD	A0T121	A0T121_9VIRU60
RKNYKVQIKLQNPACTK-DACDPSVTRSAFADLTLSFTSYSTDAERALIRTELAALLQD	C8YL75	C8YL75_9VIRU60
RKNFKVQIKLQNPACTK-DACDPSVTRSAFADVTLSFTSYSTDEERALIRTELAALLQD	C8YL83	C8YL83_9VIRU60
RKNYKIQIKLQNPACTK-DACDPSVTRSAFADLTLSFTSYSTDVERALVRTELAALLKD	C8YL79	C8YL79_9VIRU60
RKNFKVQIKLQNPACTK-DACDPSVTRSAFADVTLSFTSYSTDAERALIRTELAALLQD	C8YL87	C8YL87_9VIRU60
RKNFKVQINLQNPACTK-DACDPSVTRSAFADVTLSFTSYSTDEELALIRTELAALLAD	G4WZS5	G4WZS5_BPSP
RKNFKVQIKLQNPACTK-DACDPSVTRSAFADVTLSFTSYSTDEERALIRTELAALLAD	G4WZS9	G4WZS9_BPSP60
RKNYKVQVKIQNPTACTANGSCDPSVTRQAYADVTFSTQYSTDEERAFVRTELAALLAS	G4WZR7	G4WZR7_BPQBE60
RKNYKVQVKIQNPTACTANGSCDPSVTRQAYADVTFSTQYSTDEERAFVRTELAALLAS	Q774G0	Q774G0_BPQBE60
RKNFKVQIKLQNPACTK-DACDPSVTRSAFADVTLSFTSYSTDEERALIRTELAALLAD	D0U1H3	D0U1H3_BPSP60
RKNYKVQVKIQNPTACAANGSCDPSVTRQAYADVTFSTQYSTDEERAFVRTELAALLAS	D0U1F3	D0U1F3_BPQBE60
RKNFKVQMKLQKPTACKK-DACDPFVTRSAFADVTLSFTSYSTDEERALIRTELAALLAD	D0U1H7	D0U1H7_BPSP60
RKNFKVQIKLQNPACTK-DACDPSVTRSAFADVTLSFTSYSTDEERALIRTELAALLAD	D0U1G9	D0U1G9_BPSP
h*h* hh ** * *** ** h *h*h ** * *hh**** **		

FIGURE 27B

PLL-IDAIDQLNPAY-	P03615	COAT_BPQBE120	enterobacteria phage Qbeta	TaxId 39803
PML-IDAIDNLNPAY-	Q9T0R9	CAPSD_BPMX1	enterobacteria phage mx1	TaxId 75723
PLI-VDAIDNLNPAY-	P09673	COAT_BPSP120	enterobacteria phage sp	TaxId 12027
PLL-IDAIDRLNPAY-	C8YJJ2	C8YJJ2_BPQBE120	enterobacteria phage Qbeta	TaxId 39803
PML-VNAIDNLNPAY-	Q9T0S0	Q9T0S0_9VIRU119	enterobacteria phage m11	TaxId 74336
DLI-VDAIDNLNPAY-	Q9T0R8	Q9T0R8_9VIRU120	enterobacteria phage n195	TaxId 75725
PLL-IDAIDRLNPAY-	C8YJI8	C8YJI8_BPQBE120	enterobacteria phage Qbeta	TaxId 39803
PLL-IDAIDQLNPAY-	C8YJJ6	C8YJJ6_BPQBE120	enterobacteria phage Qbeta	TaxId 39803
PLL-IDAIDQLNPAY-	C8YJI0	C8YJI0_BPQBE119	enterobacteria phage Qbeta	TaxId 39803
PLI-TDAIDNLNPAY-	A0T121	A0T121_9VIRU119	enterobacteria phage f1	TaxId 39804
PLI-VDAIDNLNPAY-	C8YL75	C8YL75_9VIRU119	enterobacteria phage f1	TaxId 39804
NLI-IDAIDNLNPAY-	C8YL83	C8YL83_9VIRU119	enterobacteria phage f1	TaxId 39804
DLI-VDAIDNLNPAY-	C8YL79	C8YL79_9VIRU119	enterobacteria phage f1	TaxId 39804
PLI-VDAIDNLNPAY-	C8YL87	C8YL87_9VIRU119	enterobacteria phage f1	TaxId 39804
PLI-VDAIDNLNPAY-	G4WZS5	G4WZS5_BPSP119	enterobacteria phage sp	TaxId 12027
PLI-VDAIDNLNPAY-	G4WZS9	G4WZS9_BPSP120	enterobacteria phage sp	TaxId 12027
PLL-IDAIDQLNPAY-	G4WZR7	G4WZR7_BPQBE120	enterobacteria phage Qbeta	TaxId 39803
PLL-IDAIDQLNPAY-	Q774G0	Q774G0_BPQBE119	enterobacteria phage Qbeta	TaxId 39803
PLI-VDAIDNLNPAY-	D0U1H3	D0U1H3_BPSP120	enterobacteria phage sp	TaxId 12027
PLL-IDAIDQLNPAY-	D0U1F3	D0U1F3_BPQBE119	enterobacteria phage Qbeta	TaxId 39803
PLI-VDAIDNLNPAY-	D0U1H7	D0U1H7_BPSP119	enterobacteria phage sp	TaxId 12027
PLI-VDAIDNLNPAY-	D0U1G9	D0U1G9_BPSP	enterobacteria phage sp	TaxId 12027

h h*** *****

FIG. 27C

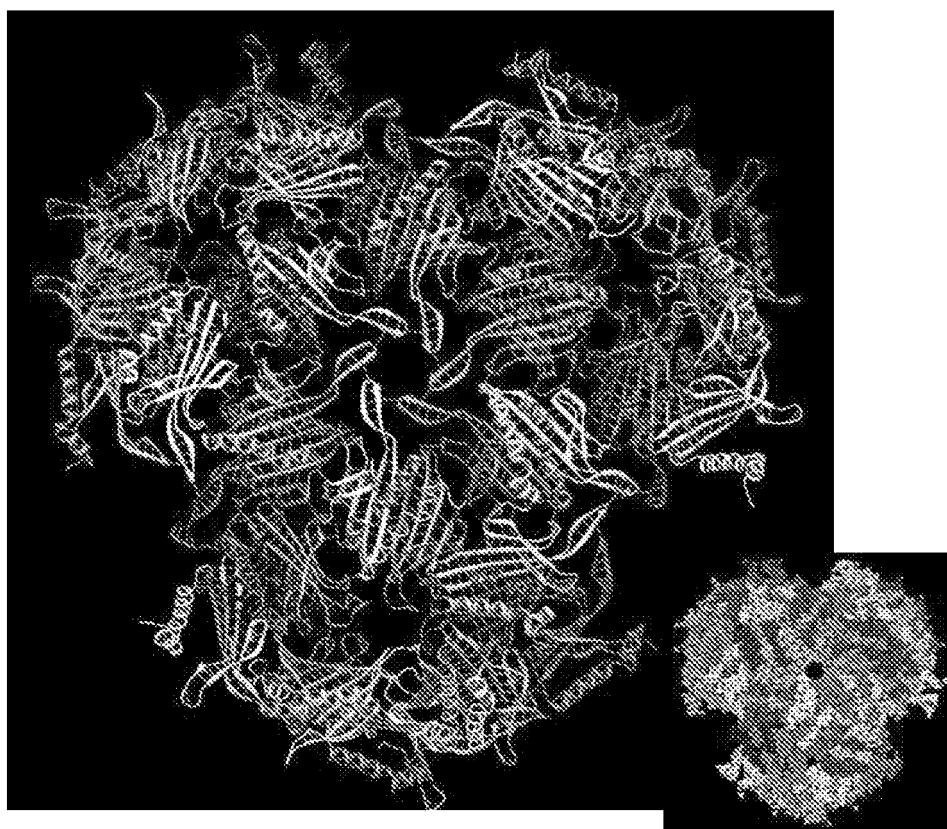


FIG. 28



FIG. 29



FIG. 30



FIG. 31

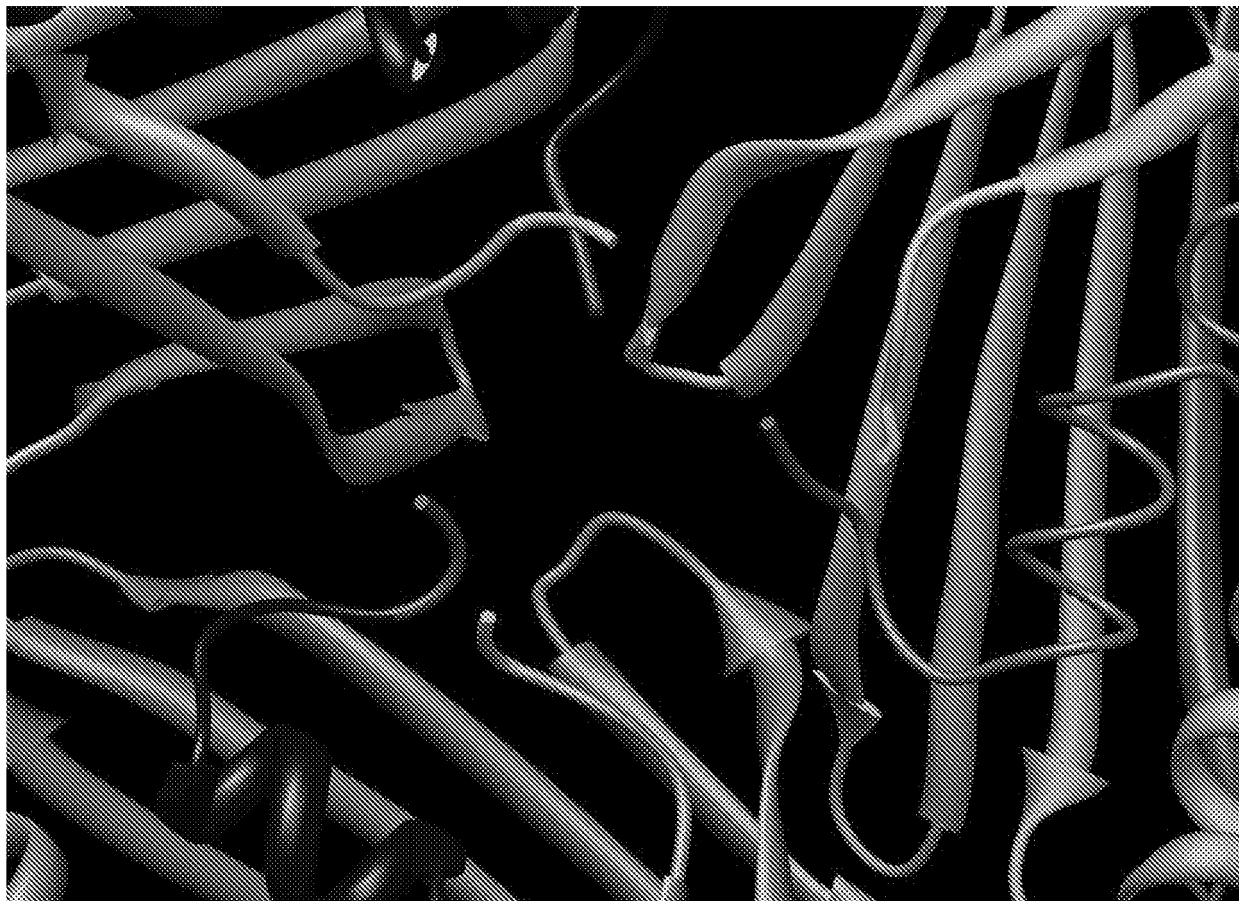


FIG. 32