

Sequence_Listing
SEQUENCE LISTING

<110> UNIVERSITY OF PITTSBURGH - OF THE COMMONWEALTH SYSTEM OF
HIGHER EDUCATION
Ross, Ted M.
Crevar, Corey J.
Giles, Brendan M.

<120> COMPUTATIONALLY OPTIMIZED BROADLY REACTIVE ANTIGENS FOR HUMAN AND
AVIAN H5N1 INFLUENZA

<130> 8123-88295-02

<150> US 61/597,998
<151> 2012-02-13

<160> 4

<170> PatentIn version 3.5

<210> 1
<211> 566
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic polypeptide

<400> 1

Met Glu Lys Ile Val Leu Leu Leu Ala Ile Val Ser Leu Val Lys Ser
1 5 10 15

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
20 25 30

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
35 40 45

Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
50 55 60

Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
65 70 75 80

Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
85 90 95

Glu Lys Ala Ser Pro Ala Asn Asp Leu Cys Tyr Pro Gly Asp Phe Asn
100 105 110

Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
115 120 125

Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asn His Glu Ala Ser
130 135 140

Ser Gly Val Ser Ser Ala Cys Pro Tyr Gln Gly Lys Ser Ser Phe Phe
145 150 155 160

Sequence_Listing

Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Ala Tyr Pro Thr Ile
 165 170 175
 Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
 180 185 190
 Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Lys Leu Tyr Gln
 195 200 205
 Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gln Arg
 210 215 220
 Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
 225 230 235 240
 Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
 245 250 255
 Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
 260 265 270
 Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly
 275 280 285
 Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
 290 295 300
 Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
 305 310 315 320
 Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
 325 330 335
 Pro Gln Arg Glu Arg Arg Arg Lys Lys Arg Gly Leu Phe Gly Ala Ile
 340 345 350
 Ala Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr
 355 360 365
 Gly Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys
 370 375 380
 Glu Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser
 385 390 395 400
 Ile Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe
 405 410 415
 Asn Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp
 420 425 430

Sequence_Listing

Gly Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met
435 440 445

Glu Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu
450 455 460

Tyr Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly
465 470 475 480

Asn Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu
485 490 495

Ser Val Arg Asn Gly Thr Tyr Asp Tyr Pro Gln Tyr Ser Glu Glu Ala
500 505 510

Arg Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly
515 520 525

Thr Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala
530 535 540

Leu Ala Ile Met Val Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly
545 550 555 560

Ser Leu Gln Cys Arg Ile
565

<210> 2
<211> 1736
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic polynucleotide

<400> 2
atggaaaaga tcgtgctgct gctggccatc gtgtccctgg tgaagagcga ccagatttgc 60
attggctacc acgccaacaa tagcacagag caggtggaca ccatcatgga gaaaaacgtg 120
accgtgaccc acgcccagga catcctggag aaaaccacaca acggcaagct gtgtgacctg 180
gacggcgtga agcccctgat cctgagagac tgctccgtgg ccggctggct gctgggcaac 240
cccattgtgtg acgagttcat caacgtgccc gagtggagct acatcgtgga gaaggccagc 300
cccgccaacg acctgtgcta ccccggcgac ttcaacgact acgaggagct gaagcacctg 360
ctgtcccgga tcaaccactt cgagaagatc cagatcatcc ccaagagcag ctggagcaac 420
cacgaggcca gcagcggcgt gtccagcgcg tccccctacc agggcaagag cagcttcttc 480
cggaacgtgg tctggctgat caagaagaac tctgcctatc ccaccatcaa gcggagctac 540
aacaacacca accaggagga tctgctggtc ctgtggggca tccaccaccc caacgacgcc 600
gccgagcaga ccaagctgta ccagaacccc accacctaca tcagcgtggg caccagcacc 660

Sequence_Listing

```

ctgaaccagc ggctggtgcc caagatcgcc acccggcca aagtgaacgg ccagagcggc 720
cggatggaat tcttctggac catcctgaag cccaacgatg ccatcaactt cgagagcaac 780
ggcaacttca tcgcccccca gtacgcctac aagatcgtga agaagggcga cagcgccatc 840
atgaagagcg agctggaata cggcaactgc aacaccaagt gccagacccc catgggcgcc 900
atcaacagca gcatgccctt ccacaacatc caccctctga ccatcggcga gtgccccaaag 960
tacgtgaaga gcaacaggct ggtgctggcc accggcctgc ggaacagccc ccagcgggag 1020
cggcggagga agaagcgggg cctgttcggc gccatcgccg gcttcacga gggcggctgg 1080
cagggcatgg tggacgggtg gtacggctac caccacagca atgagcaggg cagcggctac 1140
gccgccgaca aagagagcac ccagaaggcc atcgacggcg tcaccaacaa ggtgaacagc 1200
atcatcgaca agatgaacac ccagttcgag gccgtgggcc gggagttcaa caacctggaa 1260
cggcggatcg agaacctgaa caagaaaatg gaagatggct tcctggacgt gtggacctac 1320
aacgccgaac tcctggtgct gatggaaaac gagcggaccc tggacttcca cgacagcaac 1380
gtgaagaacc tgtacgacaa agtgcggctg cagctgcggg acaacgcaa agagctgggc 1440
aacggctgct tcgagttcta ccacaagtgc gacaacgagt gcatggaaag cgtgcggaac 1500
ggcacctacg actaccccca gtacagcgag gaagcccgcc tgaagcggga ggaaatcagc 1560
ggcgtgaaac tggaaagcat cggcacctac cagatcctga gcatctacag caccgtggcc 1620
agcagcctcg ctctggccat tatggtggcc ggcctgagcc tgtggatgtg cagcaacggc 1680
agcctgcagt gccggatcgg atccagatct gctagcgctg actctagatt aattaa 1736

```

<210> 3
 <211> 567
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic polypeptide

<400> 3

Met Glu Lys Ile Val Leu Leu Leu Ala Ile Val Ser Leu Val Lys Ser
 1 5 10 15

Asp Gln Ile Cys Ile Gly Tyr His Ala Asn Asn Ser Thr Glu Gln Val
 20 25 30

Asp Thr Ile Met Glu Lys Asn Val Thr Val Thr His Ala Gln Asp Ile
 35 40 45

Leu Glu Lys Thr His Asn Gly Lys Leu Cys Asp Leu Asp Gly Val Lys
 50 55 60

Pro Leu Ile Leu Arg Asp Cys Ser Val Ala Gly Trp Leu Leu Gly Asn
 65 70 75 80

Sequence_Listing

Pro Met Cys Asp Glu Phe Ile Asn Val Pro Glu Trp Ser Tyr Ile Val
85 90 95

Glu Lys Ala Asn Pro Ala Asn Asp Leu Cys Tyr Pro Gly Asn Phe Asn
100 105 110

Asp Tyr Glu Glu Leu Lys His Leu Leu Ser Arg Ile Asn His Phe Glu
115 120 125

Lys Ile Gln Ile Ile Pro Lys Ser Ser Trp Ser Asp His Glu Ala Ser
130 135 140

Ser Gly Val Ser Ser Ala Cys Pro Tyr Gln Gly Lys Ser Ser Phe Phe
145 150 155 160

Arg Asn Val Val Trp Leu Ile Lys Lys Asn Ser Ala Tyr Pro Thr Ile
165 170 175

Lys Arg Ser Tyr Asn Asn Thr Asn Gln Glu Asp Leu Leu Val Leu Trp
180 185 190

Gly Ile His His Pro Asn Asp Ala Ala Glu Gln Thr Arg Leu Tyr Gln
195 200 205

Asn Pro Thr Thr Tyr Ile Ser Val Gly Thr Ser Thr Leu Asn Gln Arg
210 215 220

Leu Val Pro Lys Ile Ala Thr Arg Ser Lys Val Asn Gly Gln Ser Gly
225 230 235 240

Arg Met Glu Phe Phe Trp Thr Ile Leu Lys Pro Asn Asp Ala Ile Asn
245 250 255

Phe Glu Ser Asn Gly Asn Phe Ile Ala Pro Glu Tyr Ala Tyr Lys Ile
260 265 270

Val Lys Lys Gly Asp Ser Ala Ile Met Lys Ser Glu Leu Glu Tyr Gly
275 280 285

Asn Cys Asn Thr Lys Cys Gln Thr Pro Met Gly Ala Ile Asn Ser Ser
290 295 300

Met Pro Phe His Asn Ile His Pro Leu Thr Ile Gly Glu Cys Pro Lys
305 310 315 320

Tyr Val Lys Ser Asn Arg Leu Val Leu Ala Thr Gly Leu Arg Asn Ser
325 330 335

Pro Gln Arg Glu Arg Arg Arg Lys Arg Gly Leu Phe Gly Ala Ile Ala
340 345 350

Sequence_Listing

Gly Phe Ile Glu Gly Gly Trp Gln Gly Met Val Asp Gly Trp Tyr Gly
355 360 365

Tyr His His Ser Asn Glu Gln Gly Ser Gly Tyr Ala Ala Asp Lys Glu
370 375 380

Ser Thr Gln Lys Ala Ile Asp Gly Val Thr Asn Lys Val Asn Ser Ile
385 390 395 400

Ile Asp Lys Met Asn Thr Gln Phe Glu Ala Val Gly Arg Glu Phe Asn
405 410 415

Asn Leu Glu Arg Arg Ile Glu Asn Leu Asn Lys Lys Met Glu Asp Gly
420 425 430

Phe Leu Asp Val Trp Thr Tyr Asn Ala Glu Leu Leu Val Leu Met Glu
435 440 445

Asn Glu Arg Thr Leu Asp Phe His Asp Ser Asn Val Lys Asn Leu Tyr
450 455 460

Asp Lys Val Arg Leu Gln Leu Arg Asp Asn Ala Lys Glu Leu Gly Asn
465 470 475 480

Gly Cys Phe Glu Phe Tyr His Lys Cys Asp Asn Glu Cys Met Glu Ser
485 490 495

Val Arg Asn Gly Thr Tyr Asp Tyr Pro Gln Tyr Ser Glu Glu Ala Arg
500 505 510

Leu Lys Arg Glu Glu Ile Ser Gly Val Lys Leu Glu Ser Ile Gly Thr
515 520 525

Tyr Gln Ile Leu Ser Ile Tyr Ser Thr Val Ala Ser Ser Leu Ala Leu
530 535 540

Ala Ile Met Val Ala Gly Leu Ser Leu Trp Met Cys Ser Asn Gly Ser
545 550 555 560

Leu Gln Cys Arg Ile Cys Ile
565

<210> 4
<211> 1739
<212> DNA
<213> Artificial Sequence

<220>
<223> syntheric polynucleotide

<400> 4
atggagaaga tcgtgctgct gctggccatc gtgtccctgg tgaagagcga ccagatttgc 60

atcggctacc acgccaacaa tagcaccgag caagtggaca ccatcatgga gaaaaacgtg 120
Page 6

Sequence_Listing

accgtgaccc	acgctcagga	catcctcgaa	aaaaccacaca	acggcaagct	gtgcatctg	180
gacggcgtga	agcccctgat	cctgagagac	tgcagcgtgg	ccggctggct	gctgggcaat	240
cccatgtgcg	acgagttcat	caacgtgccc	gagtggagct	acatcgtgga	gaaggccaac	300
cccggccaacg	acctgtgcta	ccccggcaac	ttcaacgact	acgaggagct	gaagcacctg	360
ctgagccgga	tcaaccactt	cgagaagatc	cagatcatcc	ccaagagcag	ctggagcgac	420
catgaggcaa	gcagcggcgt	gtccagcgcc	tgtccctacc	agggcaagtc	cagcttcttc	480
cgcaacgttg	tgtggctgat	caagaagaac	agcgcctacc	ccaccatcaa	gcgagactac	540
aacaacacca	accaggagga	cctgctgggtc	ctgtggggca	tccaccaccc	caacgacgcc	600
gccgagcaga	cccggctgta	ccagaacccc	accacctaca	tctctgtggg	caccagcacc	660
ctgaaccagc	ggctgggtgcc	caagatcgcc	acccggagca	aggtgaacgg	ccagagcggc	720
cggatggagt	tcttctggac	catcctgaag	cccaacgatg	ccatcaactt	cgagagcaac	780
ggcaacttca	tcgccccga	gtacgcctac	aagatcgtga	agaagggcga	cagcgccatc	840
atgaagtccg	agctggagta	cggcaactgt	aacaccaagt	gccagacccc	catgggcgcc	900
atcaacagca	gcatgccctt	ccacaacatc	caccccctga	ccatcggcga	gtgccccaa	960
tacgtgaaga	gcaacaggct	ggtgctggcc	accggcctga	gaaacagccc	ccagagagag	1020
cggagaagaa	agagaggcct	gttcggcgcc	attgccggct	tcatcgaggg	cggctggcag	1080
ggcatggtgg	acgggtggta	cggctaccac	cactccaacg	agcagggcag	cggctacgcc	1140
gccgacaaag	agagcaccca	gaaagctatc	gacggcgtga	ccaacaaagt	gaacagcatc	1200
atcgacaaga	tgaataccca	gttcgaggcc	gtgggcagag	agttcaacaa	cctggaaaga	1260
agaatcgaga	acctgaacaa	gaaaatggaa	gatggctttc	tggatgtgtg	gacctacaac	1320
gccgagctgc	tgggtgctgat	ggaaaacgag	cggaccctgg	acttccacga	cagcaacgtg	1380
aagaatctgt	acgacaaagt	gcggctgcag	ctgagagaca	acgccaaaga	gctgggcaac	1440
ggctgcttcg	agttctacca	caagtgcgac	aatgagtgcg	tggaaagcgt	gcggaacggc	1500
acctacgact	acccccagta	cagcgaggaa	gcccggctga	agagagaaga	gatttccggc	1560
gtgaagctgg	aaagcatcgg	cacctaccag	atcctgagca	tctacagcac	cgtggccagc	1620
agcctggccc	tggccatcat	ggtggccggc	ctgagcctgt	ggatgtgcag	caacggcagc	1680
ctgcagtgcc	ggatctgcat	cggatccaga	tctgctagcg	tcgactctag	attaattaa	1739