

103399PC sequ list.txt
SEQUENCE LISTING

<110> GlycoVaxyn AG
Eidgenossische Materialprüfungs-und Forschungsanstalt

<120> METHODS AND COMPOSITIONS RELATING TO CRM197

<130> 103399PC

<140> to be assigned

<141>

<150> 61/746,366

<151> 2012-12-27

<160> 6

<170> PatentIn version 3.5

<210> 1

<211> 7093

<212> DNA

<213> Artificial Sequence

<220>

<223> expression plasmid p932

<400> 1

gtc	gag	ctag	taaa	agcatt	ttaa	aataagg	agga	aataaca	catat	gaaaa	agatt	tggt		60
ggc	gct	ggct	ggtt	tagttt	tag	cggtttag	cgcc	agcgcc	gcag	atgacg	ttgt	tgacag		120
cag	caa	atcc	ttcg	ttatgg	aaa	acttctc	ctct	tatcac	ggc	accaa	ac	cggt	tatgt	180
gg	acagcatt	caga	aaggt	tcc	aaaa	aacc	gaa	atctggc	acg	cagggt	act	acgatga		240
cg	attggaaa	gaatt	ctaca	gc	accg	acaa	caa	atatgat	gcg	gccggt	act	cagttga		300
ca	acgaaa	at	ccgt	gtcgg	gcaa	agccgg	cggt	gtgggt	aa	agtcacgt	atcc	gggcct		360
ga	ccaaa	agtc	ctgg	ccctga	aagt	ggataa	tgc	agaaacc	atc	aaaaa	ag	aact	gggtct	420
gag	cctgacg	ga	accgctga	tg	gaacaggt	tgg	caccgaa	ga	atttatca	aac	gcttcgg			480
cg	atgggtgcc	agt	cgtgtcg	tg	ctgtccct	gcc	gttcgca	ga	aggtagct	ct	agtgtcga			540
at	atattaac	aatt	gggaac	aag	cgaagc	cct	gtccgtg	ga	actggaaa	tca	actttga			600
a	acccgcggc	aa	acgtggtc	agg	atgcgat	gt	atgaatac	at	ggcacaag	ctt	gcgcggg			660
ta	atcgcggt	cg	tcgcagcg	tc	ggctcctc	act	gtcttgt	at	caacctgg	act	gggatgt			720
t	atccgtgat	aaa	acaaaa	cg	aaaatcga	a	agtctgaaa	ga	acacggcc	cg	atcaaaaa			780
c	aaaatgagc	ga	atctccga	at	aaaacggt	gt	ccgaagaa	aa	agctaaac	ag	tatctgga			840
a	gaattccac	caa	accgcac	t	ggaacatcc	g	gaactgtca	ga	actgaaaa	cc	gtcacggg			900
t	accaacccg	gt	gtttgccg	gc	gcaaatta	cg	cagcttgg	g	ctgtgaacg	tt	gcgcaagt			960
g	attgactcg	gaa	acggccg	ata	atctgga	aaa	aaccacg	gc	ggccctga	gt	attctgcc			1020
g	ggcatcggt	tcc	gtgatgg	gt	attgccga	t	ggcgagtt	cat	cacaaca	cc	gaagaaat			1080
t	gtcgcccag	t	ctatcgac	t	gtcgagcct	gat	ggttgct	ca	agcgattc	cg	ctggttgg			1140
c	gaactgggt	gat	atcggct	tt	gcagctta	ca	acttcgtg	gaa	agtatca	tca	aacctggt			1200

103399PC sequ list.txt

tcagggttgtc	cataactcat	ataatcgccc	ggcctactcg	ccgggtcaca	aaaccaacc	1260
gttcctgcat	gacggctacg	cggttagctg	gaatacggtc	gaagattcta	ttatccgtac	1320
cggctttcag	ggtgaatctg	gccacgacat	taaaatcacg	gctgaaaaca	ccccgctgcc	1380
gattgccggt	gttctgctgc	cgaccatccc	gggtaaactg	gatgtgaata	aatcaaaaac	1440
ccatatctcg	gttaacggtc	gcaaaattcg	tatgcgctgc	cgtgcgatcg	acggcgatgt	1500
gaccttctgt	cgtcggaaaa	gcccgggtcta	tgtgggcaac	ggtgttcatg	ctaactctgca	1560
cgtcgcgttt	catcgctcta	gttccgaaaa	aatccacagt	aacgaaatct	catcggactc	1620
cattgggtgtg	ctgggctacc	agaaaacggg	ggatcatacc	aaagttaata	gcaaactgtc	1680
actgttcttc	gaaatcaaat	caggctcgca	tcatcatcat	caccactaat	ctagaggatc	1740
cccgggtacc	gagctcgaat	tcttgaagac	gaaagggcct	cgtgatacgc	ctatTTTTAT	1800
aggttaatgt	catgataata	atggtttctt	agacgtcagg	tggcactttt	cggggaaatg	1860
tgcgcggaac	ccctatttgt	ttatTTTTct	aaatacattc	aaatatgtat	ccgctcatga	1920
gacaataacc	ctgataaatg	cttcaataat	attgaaaaag	gaagagtatg	agtattcaac	1980
atttccgtgt	cgcccttatt	ccctTTTTtg	cggcattttg	ccttcctggt	tttgctcacc	2040
cagaaacgct	ggtgaaagta	aaagatgctg	aagatcagtt	gggtgcacga	gtgggttaca	2100
tcgaactgga	tctcaacagc	ggtaagatcc	ttgagagttt	tcgccccgaa	gaacgttttc	2160
caatgatgag	cacttttaaa	gttctgctat	gtggcgcggt	attatcccgt	gttgacgccg	2220
ggcaagagca	actcggtcgc	cgcatacact	attctcagaa	tgacttggtt	gagtactcac	2280
cagtcacaga	aaagcatctt	acggatggca	tgacagtaag	agaattatgc	agtgctgcca	2340
taaccatgag	tgataacact	gcggccaaact	tacttctgac	aacgatcgga	ggaccgaagg	2400
agctaaccgc	ttttttgcac	aacatggggg	atcatgtaac	tcgccttgat	cgttgggaac	2460
cggagctgaa	tgaagccata	ccaaacgacg	agcgtgacac	cacgatgcct	gcagcaatgg	2520
caacaacgtt	gcgcaaacta	ttaactggcg	aactacttac	tctagcttcc	cggcaacaat	2580
taatagactg	gatggaggcg	gataaagttg	caggaccact	tctgcgctcg	gcccttccgg	2640
ctggctgggt	tattgctgat	aaatctggag	ccggtgagcg	tgggtctcgc	ggtatcattg	2700
cagcactggg	gccagatggt	aagccctccc	gtatcgtagt	tatctacacg	acggggagtc	2760
aggcaactat	ggatgaacga	aatagacaga	tcgctgagat	agggtgcctca	ctgattaagc	2820
attggtaact	gtcagaccaa	gtttactcat	atatacttta	gattgattta	aaacttcatt	2880
tttaatttaa	aaggatctag	gtgaagatcc	tttttgataa	tctcatgacc	aaaatccctt	2940
aacgtgagtt	ttcgttccac	tgagcgtcag	accccgtaga	aaagatcaaa	ggatcttctt	3000
gagatccttt	ttttctgcgc	gtaatctgct	gcttgcaaac	aaaaaaacca	ccgctaccag	3060
cggtggtttg	tttgccggat	caagagctac	caactctttt	tccgaaggta	actggcttca	3120
gcagagcgca	gataccaaat	actgtccttc	tagtgtagcc	gtagttaggc	caccacttca	3180
agaactctgt	agcaccgcct	acatacctcg	ctctgctaata	cctgtttacca	gtggctgctg	3240

103399PC sequ list.txt

ccagtggcga taagtcgtgt cttaccgggt tggactcaag acgatatgta ccggataagg	3300
cgagcggtc gggctgaacg ggggggttcgt gcacacagcc cagcttggag cgaacgacct	3360
acaccgaact gagataccta cagcgtgagc tatgagaaag cgccacgctt cccgaaggga	3420
gaaaggcggg caggtatccg gtaagcggca gggtcggaaac aggagagcgc acgagggagc	3480
ttccaggggg aaacgcctgg tatctttata gtcctgtcgg gtttcgccac ctctgacttg	3540
agcgtcgatt tttgtgatgc tcgtcagggg ggcggagcct atggaaaaac gccagcaacg	3600
cggccttttt acggttcctg gccttttgct ggccttttgc tcacatgttc tttcctgcgt	3660
tatcccctga ttctgtggat aaccgtatta ccgcctttga gtgagctgat accgctcgcc	3720
gcagccgaac gaccgagcgc agcagagtcag tgagcgagga agcgggaagag cgcctgatgc	3780
ggtattttct cttacgcat ctgtgcggta tttcacaccg catggtgcac tctcagtaca	3840
atctgctctg atgccgcata gtttaagccag tatacactcc gctatcgcta cgtgactggg	3900
tcattggctgc gccccgacac ccgccaacac ccgctgacgc gccctgacgg gcttgtctgc	3960
tcccggcatc cgcttacaga caagctgtga ccgtctccgg gagctgcatg tgtcagaggt	4020
tttcaccgtc atcaccgaaa cgcgcgaggc agctgcggta aagctcatca gcgtggctgt	4080
gaagcgattc acagatgtct gcctgttcat ccgcgtccag ctctgttgagt ttctccagaa	4140
gcgttaatgt ctggcttctg ataaagcggg ccatgttaag ggcggttttt tcctgtttgg	4200
tcactgatgc ctccgtgtaa gggggatttc tgttcattgg ggtaatgata ccgatgaaac	4260
gagagaggat gtcacgata cgggttactg atgatgaaca tgcccgggtta ctggaacggt	4320
gtgagggtaa acaactggcg gtatggatgc ggcgggacca gagaaaaatc actcagggtc	4380
aatgccagcg cttcgttaat acagatgtag gtgttcacac gggtagccag cagcatcctg	4440
cgatgcagat ccggaacata atggtgcagg gcgctgactt ccgcgtttcc agactttacg	4500
aaacacggaa accgaagacc attcatgttg ttgctcaggt cgagacggt ttgcagcagc	4560
agtcgcttca cgttcgctcg cgtatcggtg attcattctg ctaaccagta aggcaacccc	4620
gccagcctag ccgggtcctc aacgacagga gcacgatcat gcgcacccgt ggccaggacc	4680
caacgctgcc cgagcgtcaa cggcgccaga tacagcaaac ggctgcgggg gaaatacgcg	4740
gttaaacgat cgactgccgc tttgccgctg cgccacagcc gccagcatag ccagcctccg	4800
accacagca gcaacgccgt cgccagcagc agccatttga aatctccgct ctgcatatcg	4860
gaaggaatat cgattgccgc tcccgccaga atgcccggca ggaaataaaa cggcgccac	4920
agcaaacagc caatcaagtt cggcccaata aatttcgcc cgggaagatc cagcatccct	4980
gccaccatcg gcaccagcg cctcgtcggg ccgacaaaac gtccgaccag gatcgtgaac	5040
atactgtgct gatgcagcgc gtgttcggtt ttatccagca gcgacttggt ctttttcata	5100
aaagaccagc ggtgtagcgg ctttttaaaag cgccacccca gccagaacga aatccagtcg	5160
cccatcagac agccgataat acccaccagc caggcatgcc aaaaattgag ctcgccgctg	5220
ccgataagcg cgccagccc cgccatcagt accgtgccgg gtaaaatcaa ccccaccagc	5280

103399PC sequ list.txt

gccagcgatt ccaggaaggc gaccagcaac acggcgatga gccaatacag agtggattgg	5340
gtgataaagt gttccagcag tgcttgcata gtgtgtccgt cagcgtgatg aagcagggat	5400
tctgcttacc ccgtccccct tcgtcaagcc gtcaattatc cgaatagtta cggcttatga	5460
catctttgtg gacacatcat tcacttttta ttcacatccg gccctgaact cgctaggact	5520
tgccccggtg catttttttaa ataccgcga aaaatagagc tgatcgtcaa atccaacatt	5580
gcgccaacg gtcgctatcg gcattcgctg agtgctaagc agaagtttcg cctggctgat	5640
acgctgatct tcgcgccagc tcaatacgct aatgcctaac tgctggcgga acagatgtga	5700
taaccgggag ggcgacaggc agacatgctg ggcgacgctg gcgatataca aatggctgtc	5760
cgccagatgg tcgctgatat actggcaggc atcgcgca ca cggctatcca tcggcgggtg	5820
caacgactca ttaattaccg ccatacgtct gagcaacaac tgctccagca gattgatcgc	5880
cagtagctca gaatagcgac cttccccttg cccggcgctg atgatctgcc cgaacagttc	5940
gctgaaatgc ggctggcgcg cctcgtccgg gcggaataat cctgtctggg caaagattgt	6000
cggccaggtc agccactcct gccagtaggc gcgaggccgg aaataaacc actggtgata	6060
ccactcgctg gcgtccggat gccgtccata gtgatgaatc tcgcccggcg gaaacaataa	6120
tatatcgcca ggccgacaga caaactgctc gccattatta ttaatgacgc cctctccgcg	6180
gatggtcagg ttaagaatat atcccttcat gcccaacgga cgatcgataa aaaaatccag	6240
atatccattc gcttcaattg gcgtcagccc ggcgaccaga tgggcattaa atgaatatcc	6300
cggcaatagc ggatcatttt gcgtttcagc catgatttct ctaccccccg atgttcagag	6360
aagaaacaaa ttgtccatat cgaccaggac gacagagctt ccgtctccgc aagactttgc	6420
gcttgatgaa agcacgtatc aaccccgctt gtgaaaagcg ctttgtaaca aaagcgtaca	6480
gttcaggcga taaaattaag taacagaagt gtctataact atggctggaa tgtccacatt	6540
gaatatttgc acagcgtcac actttgcaaa gcattagcat ttttgtccat aagattagcg	6600
gatcctgcct gacggttttt gccgcgactc tctactgttt ctccatacct gtttttctgg	6660
atggagtaag acgatggcaa ttgcaattgg cctcgatttt ggcagtgatt cagtgcgcgc	6720
tctggcagtg gactgcgcca ccggcgacga gatcgccacc agcgtagagt ggtatccgcg	6780
ctggcaagaa ggccgttatt gcgacggccc gaacaaccag ttccgtcatc atccgcgcga	6840
ctacatggag tcaatggagg ccgcgctgaa agccgttctg gcacaattaa gcgccgcga	6900
acgcgcaa at gtcgttgga ttggcggtga cagcaccggc tctacgccag cgccgattga	6960
cgccgacggt aacgtcctgg cgctgcgtcc agagttcgcc gagaaccga atgcgatgtt	7020
tgtgctgtgg aaagatcaca ccgccgtgga agaggccgac gaaatcactc gtctgtgcca	7080
taagccaggc aag	7093

<210> 2
 <211> 7096
 <212> DNA
 <213> Artificial Sequence

103399PC sequ list.txt

<220>

<223> expression plasmid p933

<400> 2

gtcgcgctag taaaagcatt ttaaataagg aggaataaca catatgaaaa agatttggct	60
ggcgcctggct ggtttagttt tagcgtttag cgccatggca ggcgcagatg acgttgttga	120
cagcagcaaa tccttcgtta tggaaaactt ctctcttat cacggcacca aaccgggcta	180
tgtggacagc attcagaaag gtatccaaaa accgaaatct ggcacgcagg gtaactacga	240
tgacgattgg aaagaattct acagcaccga caacaaatat gatgcggccg gttactcagt	300
tgacaacgaa aatccgctgt cgggcaaagc cggcgggtgtg gttaaagtca cgtatccggg	360
cctgacaaa gtcttgcccc tgaaagtgga taatgcagaa accatcaaaa aagaactggg	420
tctgagcctg acggaaccgc tgatggaaca ggttggcacc gaagaattta tcaaacgctt	480
cggcgatggg gccagtcgtg tcgtgctgtc cctgccgttc gcagaaggta gctctagtgt	540
cgaatatatt aacaattggg aacaagcgaa agccctgtcc gtggaactgg aaatcaactt	600
tgaaacccgc ggcaaactg gtcaggatgc gatgtatgaa tacatggcac aagcttgccg	660
gggtaatcgc gttcgtcgca gcgtcggctc ctactgtct tgtatcaacc tggactggga	720
tgttatccgt gataaaacca aaacgaaaat cgaaagtctg aaagaacacg gcccgatcaa	780
aaacaaaatg agcgaatctc cgaataaaac ggtgtccgaa gaaaaagcta aacagtatct	840
ggaagaattc caccaaaccg cactggaaca tccggaactg tcagaactga aaaccgtcac	900
gggtaccaac ccggtgtttg ccggcgcaaa ttacgcagct tgggctgtga acgttgccga	960
agtgattgac tcggaaacgg ccgataatct ggaaaaaacc acggcggccc tgagtattct	1020
gccgggcatc gggtccgtga tgggtattgc cgatggcgca gttcatcaca acaccgaaga	1080
aattgtcgcc cagtctatcg cactgtcgag cctgatgggt gctcaagcga ttccgctggg	1140
tggcgaactg gttgatatcg gctttgcagc ttacaacttc gtggaaagta tcatcaacct	1200
gtttcagggt gtccataact catataatcg cccggcctac tcgccgggtc acaaaaccca	1260
accgttcctg catgacggct acgcgggttag ctggaatacg gtcgaagatt ctattatccg	1320
taccggcttt cagggtgaat ctggccacga cattaaaatc acggctgaaa acaccccgct	1380
gccgattgcc ggtgttctgc tgccgaccat cccgggtaaa ctggatgtga ataaatcaaa	1440
aacccatata tcggttaacg gtcgcaaaat tcgtatgcgc tgccgtgcga tcgacggcga	1500
tgtgaccttc tgtcgtccga aaagcccggg ctatgtgggc aacgggtgtt atgctaattct	1560
gcacgtcgcg tttcatcgct ctagtccga aaaaatccac agtaacgaaa tctcatcgga	1620
ctccattggg gtgctgggct accagaaaac ggtggatcat accaaagtta atagcaaact	1680
gtcactgttc ttcgaaatca aatcaggctc gcatcatcat catcaccact aatctagagg	1740
atccccgggt accgagctcg aattcttgaa gacgaaaggg cctcgtgata cgcctatttt	1800
tatagggttaa tgtcatgata ataattggtt cttagacgtc aggtggcact tttcggggaa	1860
atgtgcgcgg aaccctatt tgtttatttt tctaaatata ttcaaatatg tatccgctca	1920

103399PC sequ list.txt

tgagacaata accctgataa atgcttcaat aatattgaaa aaggaagagt atgagtattc	1980
aacatttccg tgtcgccctt attccctttt ttgcggcatt ttgccttcct gtttttgctc	2040
accagaaaac gctggtgaaa gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt	2100
acatcgaact ggatctcaac agcggtaaga tccttgagag ttttcgcccc gaagaacgtt	2160
ttccaatgat gagcactttt aaagtctctgc tatgtggcgc ggtattatcc cgtgttgacg	2220
ccgggcaaga gcaactcggg cgccgcatac actattctca gaatgacttg gttgagtact	2280
caccagtcac agaaaagcat cttacggatg gcatgacagt aagagaatta tgcagtgctg	2340
ccataaccat gagtataac actgcggcca acttacttct gacaacgatc ggaggaccga	2400
aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgcctt gatcgttggg	2460
aaccggagct gaatgaagcc ataccaaagc acgagcgtga caccacgatg cctgcagcaa	2520
tggcaacaac gttgcgcaaa ctattaactg gcgaactact tactctagct tcccggcaac	2580
aattaataga ctggatggag gcggataaag ttgcaggacc acttctgcgc tcggcccttc	2640
cggctggctg gtttattgct gataaatctg gagccggtga gcgtgggtct cgcggtatca	2700
ttgcagcact ggggcccagat ggtaagccct cccgtatcgt agttatctac acgacgggga	2760
gtcaggcaac tatggatgaa cgaaatagac agatcgtga gataggtgcc tcaactgatta	2820
agcattggta actgtcagac caagtttact catatatact ttagattgat ttaaaacttc	2880
atttttaatt taaaaggatc taggtgaaga tcctttttga taatctcatg accaaaatcc	2940
cttaacgtga gttttcgttc cactgagcgt cagaccccggt agaaaagatc aaaggatcctt	3000
cttgagatcc tttttttctg cgcgtaatct gctgcttgca aacaaaaaaaa ccaccgtac	3060
cagcgggtggg ttgtttgccg gatcaagagc taccaactct ttttccgaag gtaactggct	3120
tcagcagagc gcagatacca aatactgtcc ttctagtgtg gccgtagtta ggccaccact	3180
tcaagaactc tgtagcaccg cctacatacc tcgctctgct aatcctgtta ccagtggctg	3240
ctgccagtgg cgataagtcg tgtcttaccg gggtggactc aagacgatag ttaccggata	3300
aggcgcagcg gtcgggctga acgggggggt cgtgcacaca gcccagcttg gagcgaacga	3360
cctacaccga actgagatac ctacagcgtg agctatgaga aagcgccacg cttcccgaag	3420
ggagaaaaggc ggacaggtat ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg	3480
agcttccagg gggaaacgcc tggatatctt atagtcctgt cgggtttcgc cacctctgac	3540
ttgagcgtcg atttttgtga tgctcgtcag gggggcggag cctatggaaa aacgccagca	3600
acgcggcctt tttacggttc ctggcctttt gctggccttt tgctcacatg ttctttcctg	3660
cgttatcccc tgattctgtg gataaccgta ttaccgcctt tgagttagct gataccgctc	3720
gccgcagccg aacgaccgag cgcagcagat cagttagcga ggaagcggaa gagcgcctga	3780
tgcggtatct tctccttacg catctgtgcg gtatttcaca ccgcatgggt cactctcagt	3840
acaatctgct ctgatgccgc atagttaagc cagtatacac tccgctatcg ctacgtgact	3900
gggtcatggc tgcgccccga caccgcgcaa caccgcgtga cgcgccctga cgggcttgct	3960

103399PC sequ list.txt

tgctccccggc atccgcttac agacaagctg tgaccgtctc cgggagctgc atgtgtcaga	4020
ggttttcacc gtcatcaccg aaacgcgcga ggcagctgcg gtaaagctca tcagcgtggt	4080
cgtgaagcga ttcacagatg tctgcctgtt catccgcgtc cagctcgttg agtttctcca	4140
gaagcgttaa tgtctggctt ctgataaagc gggccatgtt aagggcggtt ttttcctgtt	4200
tggtcactga tgcctccgtg taagggggat ttctgttcat gggggtaatg ataccgatga	4260
aacgagagag gatgctcacg atacgggtta ctgatgatga acatgcccgg ttactggaac	4320
gttgtgaggg taaacaactg gcggtatgga tgcggcggga ccagagaaaa atcactcagg	4380
gtcaatgcca gcgcttcgtt aatacagatg taggtgttcc acagggtagc cagcagcatc	4440
ctgcgatgca gatccggaac ataatggtgc agggcgctga cttccgcgtt tccagacttt	4500
acgaaacacg gaaaccgaag accattcatg ttgttgctca ggtcgcagac gttttgcagc	4560
agcagtcgct tcacgttcgc tcgcgatatc gtgattcatt ctgctaacca gtaaggcaac	4620
cccgccagcc tagccgggtc ctcaacgaca ggagcacgat catgcgcacc cgtggccagg	4680
acccaacgct gcccagcgt caacggcgcc agatacagca aacggctgcg ggggaaatac	4740
gcggttaaac gatcgactgc cgctttgccc ctgcgccaca gccgccagca tagccagcct	4800
ccgaccaca gcagcaacgc cgtcgccagc agcagccatt tgaaatctcc gctctgcata	4860
tcggaaggaa tatcgattgc cgctcccgcc agaatgcccg gcaggaaata aaacggcggc	4920
cacagcaaac agccaatcaa gttcggccca ataaatttcg ccacgggaag atccagcatc	4980
cctgccacca tcggcaccag cggcctcgtc ggaccgacaa aacgtccgac caggatcgtg	5040
aacatactgt gctgatgcag cgcggtttcg gttttatcca gcagcgactt gttctttttc	5100
ataaaagacc agcgggtgtag cggcttttta aagcgccacc ccagccagaa cgaaatccag	5160
tcgccccatca gacagccgat aatacccacc agccaggcat gccaaaaatt gagctcgccg	5220
ctgccgataa gcgcgcccag ccccgccatc agtaccgtgc cgggtaaaat caaccccacc	5280
agcgccagcg attccaggaa ggcgaccagc aacacggcga tgagcgaata cagagtggat	5340
tgggtgataa agtgttccag cagtgttgc atagtgtgtc cgtcagcgtg atgaagcagg	5400
gattctgctt accccgtccc ccttcgtcaa gccgtcaatt atccgaatag ttacggctta	5460
tgacatcttt gtggacacat cattcacttt ttattcacat ccggccctga actcgctagg	5520
acttgccccg gtgcattttt taaatacccg cgaaaaatag agctgatcgt caaatccaac	5580
attgcgcca acggtcgcta tcggcattcg cgtagtgcta agcagaagtt tcgcctggct	5640
gatacgctga tcttcgcgc agctcaatac gctaattgcct aactgctggc ggaacagatg	5700
tgataaccgg gagggcgaca ggcagacatg ctgggcgacg ctggcgatat caaatggct	5760
gtccgccaga tggtcgctga tatactggca ggcacgcgc acacggctat ccatcggcgg	5820
gtgcaacgac tcattaatta ccgccatacg tctgagcaac aactgctcca gcagattgat	5880
cgccagtagc tcagaatagc gaccttcccc ttgcccggcg ctgatgatct gcccgaaacag	5940
ttcgtgaaa tgcggctggc gcgcctcgtc cgggcggaaa aatcctgtct gggcaaagat	6000

103399PC sequ list.txt

tgctggccag gtcagccact cctgccagta ggcgcgaggc cggaaataaa cccactggtg	6060
ataccactcg ctggcgctccg gatgccgtcc atagtgatga atctcgcccg gcggaacaa	6120
taatatatcg ccaggccgac agacaaactg ctcgccatta ttattaatga cgccctctcc	6180
gcggatggtc aggttaagaa tatatccctt catgcccaac ggacgatcga taaaaaatc	6240
cagatatcca ttcgcttcaa ttggcgtcag cccggcgacc agatgggcat taaatgaata	6300
tcccggcaat agcggatcat tttgcgtttc agccatgatt tctctacccc ccgatgttca	6360
gagaagaaac aaattgtcca tatcgaccag gacgacagag cttccgtctc cgcaagactt	6420
tgcgcttgat gaaagcacgt atcaaccccg cttgtgaaaa gcgcttttga acaaaagcgt	6480
acagttcagg cgataaaatt aagtaacaga agtgtctata actatggctg gaatgtccac	6540
attgaatatt tgcacagcgt cacactttgc aaagcattag cttttttgtc cataagatta	6600
gcggatcctg cctgacggtt tttgccgca ctctctactg tttctccata cctgtttttc	6660
tggatggagt aagacgatgg caattgcaat tggcctcgat tttggcagtg attcagtgcg	6720
cgctctggca gtggactgcg ccaccggcga cgagatcgcc accagcgtag agtggtatcc	6780
gcgctggcaa gaaggccggt attgcgacgg cccgaacaac cagttccgtc atcatccgcg	6840
cgactacatg gagtcaatgg aggccgcgct gaaagccggt ctggcacaat taagcgccgc	6900
gcaacgcgca aatgtcgttg gcattggcgt tgacagcacc ggctctacgc cagcgccgat	6960
tgacgcccgc ggtaacgtcc tggcgctgcg tccagagtcc gccgagaacc cgaatgcat	7020
gtttgtgctg tggaaagatc acaccgccgt ggaagaggcc gacgaaatca ctcgtctgtg	7080
ccataagcca ggcaag	7096

<210> 3
 <211> 561
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Signal peptide containing Crm197 amino acid sequence
 expressed from p932

<400> 3

Met	Lys	Lys	Ile	Trp	Leu	Ala	Leu	Ala	Gly	Leu	Val	Leu	Ala	Phe	Ser
1				5					10					15	

Ala	Ser	Ala	Ala	Asp	Asp	Val	Val	Asp	Ser	Ser	Lys	Ser	Phe	Val	Met
			20					25					30		

Glu	Asn	Phe	Ser	Ser	Tyr	His	Gly	Thr	Lys	Pro	Gly	Tyr	Val	Asp	Ser
		35					40					45			

Ile	Gln	Lys	Gly	Ile	Gln	Lys	Pro	Lys	Ser	Gly	Thr	Gln	Gly	Asn	Tyr
	50					55					60				

Asp	Asp	Asp	Trp	Lys	Glu	Phe	Tyr	Ser	Thr	Asp	Asn	Lys	Tyr	Asp	Ala
65					70					75					80

103399PC sequ list.txt

Ala Gly Tyr Ser Val Asp Asn Glu Asn Pro Leu Ser Gly Lys Ala Gly
85 90 95

Gly Val Val Lys Val Thr Tyr Pro Gly Leu Thr Lys Val Leu Ala Leu
100 105 110

Lys Val Asp Asn Ala Glu Thr Ile Lys Lys Glu Leu Gly Leu Ser Leu
115 120 125

Thr Glu Pro Leu Met Glu Gln Val Gly Thr Glu Glu Phe Ile Lys Arg
130 135 140

Phe Gly Asp Gly Ala Ser Arg Val Val Leu Ser Leu Pro Phe Ala Glu
145 150 155 160

Gly Ser Ser Ser Val Glu Tyr Ile Asn Asn Trp Glu Gln Ala Lys Ala
165 170 175

Leu Ser Val Glu Leu Glu Ile Asn Phe Glu Thr Arg Gly Lys Arg Gly
180 185 190

Gln Asp Ala Met Tyr Glu Tyr Met Ala Gln Ala Cys Ala Gly Asn Arg
195 200 205

Val Arg Arg Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp
210 215 220

Asp Val Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu
225 230 235 240

His Gly Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val
245 250 255

Ser Glu Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala
260 265 270

Leu Glu His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn
275 280 285

Pro Val Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala
290 295 300

Gln Val Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala
305 310 315 320

Ala Leu Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp
325 330 335

Gly Ala Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala
340 345 350

103399PC sequ list.txt

Leu Ser Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu
355 360 365

Val Asp Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn
370 375 380

Leu Phe Gln Val Val His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser Pro
385 390 395 400

Gly His Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp
405 410 415

Asn Thr Val Glu Asp Ser Ile Ile Arg Thr Gly Phe Gln Gly Glu Ser
420 425 430

Gly His Asp Ile Lys Ile Thr Ala Glu Asn Thr Pro Leu Pro Ile Ala
435 440 445

Gly Val Leu Leu Pro Thr Ile Pro Gly Lys Leu Asp Val Asn Lys Ser
450 455 460

Lys Thr His Ile Ser Val Asn Gly Arg Lys Ile Arg Met Arg Cys Arg
465 470 475 480

Ala Ile Asp Gly Asp Val Thr Phe Cys Arg Pro Lys Ser Pro Val Tyr
485 490 495

Val Gly Asn Gly Val His Ala Asn Leu His Val Ala Phe His Arg Ser
500 505 510

Ser Ser Glu Lys Ile His Ser Asn Glu Ile Ser Ser Asp Ser Ile Gly
515 520 525

Val Leu Gly Tyr Gln Lys Thr Val Asp His Thr Lys Val Asn Ser Lys
530 535 540

Leu Ser Leu Phe Phe Glu Ile Lys Ser Gly Ser His His His His His
545 550 555 560

His

<210> 4
<211> 562
<212> PRT
<213> Artificial Sequence

<220>
<223> Signal peptide containing Crm197 amino acid sequence
expressed from p933

<400> 4

103399PC sequ list.txt

Met Lys Lys Ile Trp Leu Ala Leu Ala Gly Leu Val Leu Ala Phe Ser
1 5 10 15

Ala Met Ala Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val
20 25 30

Met Glu Asn Phe Ser Ser Tyr His Gly Thr Lys Pro Gly Tyr Val Asp
35 40 45

Ser Ile Gln Lys Gly Ile Gln Lys Pro Lys Ser Gly Thr Gln Gly Asn
50 55 60

Tyr Asp Asp Asp Trp Lys Glu Phe Tyr Ser Thr Asp Asn Lys Tyr Asp
65 70 75 80

Ala Ala Gly Tyr Ser Val Asp Asn Glu Asn Pro Leu Ser Gly Lys Ala
85 90 95

Gly Gly Val Val Lys Val Thr Tyr Pro Gly Leu Thr Lys Val Leu Ala
100 105 110

Leu Lys Val Asp Asn Ala Glu Thr Ile Lys Lys Glu Leu Gly Leu Ser
115 120 125

Leu Thr Glu Pro Leu Met Glu Gln Val Gly Thr Glu Glu Phe Ile Lys
130 135 140

Arg Phe Gly Asp Gly Ala Ser Arg Val Val Leu Ser Leu Pro Phe Ala
145 150 155 160

Glu Gly Ser Ser Ser Val Glu Tyr Ile Asn Asn Trp Glu Gln Ala Lys
165 170 175

Ala Leu Ser Val Glu Leu Glu Ile Asn Phe Glu Thr Arg Gly Lys Arg
180 185 190

Gly Gln Asp Ala Met Tyr Glu Tyr Met Ala Gln Ala Cys Ala Gly Asn
195 200 205

Arg Val Arg Arg Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp
210 215 220

Trp Asp Val Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys
225 230 235 240

Glu His Gly Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr
245 250 255

Val Ser Glu Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr
260 265 270

103399PC sequ list.txt

Ala Leu Glu His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr
275 280 285

Asn Pro Val Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val
290 295 300

Ala Gln Val Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr
305 310 315 320

Ala Ala Leu Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala
325 330 335

Asp Gly Ala Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser Ile
340 345 350

Ala Leu Ser Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu
355 360 365

Leu Val Asp Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile
370 375 380

Asn Leu Phe Gln Val Val His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser
385 390 395 400

Pro Gly His Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser
405 410 415

Trp Asn Thr Val Glu Asp Ser Ile Ile Arg Thr Gly Phe Gln Gly Glu
420 425 430

Ser Gly His Asp Ile Lys Ile Thr Ala Glu Asn Thr Pro Leu Pro Ile
435 440 445

Ala Gly Val Leu Leu Pro Thr Ile Pro Gly Lys Leu Asp Val Asn Lys
450 455 460

Ser Lys Thr His Ile Ser Val Asn Gly Arg Lys Ile Arg Met Arg Cys
465 470 475 480

Arg Ala Ile Asp Gly Asp Val Thr Phe Cys Arg Pro Lys Ser Pro Val
485 490 495

Tyr Val Gly Asn Gly Val His Ala Asn Leu His Val Ala Phe His Arg
500 505 510

Ser Ser Ser Glu Lys Ile His Ser Asn Glu Ile Ser Ser Asp Ser Ile
515 520 525

Gly Val Leu Gly Tyr Gln Lys Thr Val Asp His Thr Lys Val Asn Ser
530 535 540

103399PC sequ list.txt

Lys Leu Ser Leu Phe Phe Glu Ile Lys Ser Gly Ser His His His His
545 550 555 560

His His

<210> 5
<211> 561
<212> PRT
<213> Artificial Sequence

<220>
<223> translated protein sequence of expression plasmid p722

<400> 5

Met Lys Lys Ile Trp Leu Ala Leu Ala Gly Leu Val Leu Ala Phe Ser
1 5 10 15

Ala Met Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val Met
20 25 30

Glu Asn Phe Ser Ser Tyr His Gly Thr Lys Pro Gly Tyr Val Asp Ser
35 40 45

Ile Gln Lys Gly Ile Gln Lys Pro Lys Ser Gly Thr Gln Gly Asn Tyr
50 55 60

Asp Asp Asp Trp Lys Glu Phe Tyr Ser Thr Asp Asn Lys Tyr Asp Ala
65 70 75 80

Ala Gly Tyr Ser Val Asp Asn Glu Asn Pro Leu Ser Gly Lys Ala Gly
85 90 95

Gly Val Val Lys Val Thr Tyr Pro Gly Leu Thr Lys Val Leu Ala Leu
100 105 110

Lys Val Asp Asn Ala Glu Thr Ile Lys Lys Glu Leu Gly Leu Ser Leu
115 120 125

Thr Glu Pro Leu Met Glu Gln Val Gly Thr Glu Glu Phe Ile Lys Arg
130 135 140

Phe Gly Asp Gly Ala Ser Arg Val Val Leu Ser Leu Pro Phe Ala Glu
145 150 155 160

Gly Ser Ser Ser Val Glu Tyr Ile Asn Asn Trp Glu Gln Ala Lys Ala
165 170 175

Leu Ser Val Glu Leu Glu Ile Asn Phe Glu Thr Arg Gly Lys Arg Gly
180 185 190

Gln Asp Ala Met Tyr Glu Tyr Met Ala Gln Ala Cys Ala Gly Asn Arg

195

Val Arg Arg Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp
210 215 220

Asp Val Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu
225 230 235 240

His Gly Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val
245 250 255

Ser Glu Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala
260 265 270

Leu Glu His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn
275 280 285

Pro Val Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala
290 295 300

Gln Val Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala
305 310 315 320

Ala Leu Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp
325 330 335

Gly Ala Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala
340 345 350

Leu Ser Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu
355 360 365

Val Asp Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn
370 375 380

Leu Phe Gln Val Val His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser Pro
385 390 395 400

Gly His Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp
405 410 415

Asn Thr Val Glu Asp Ser Ile Ile Arg Thr Gly Phe Gln Gly Glu Ser
420 425 430

Gly His Asp Ile Lys Ile Thr Ala Glu Asn Thr Pro Leu Pro Ile Ala
435 440 445

Gly Val Leu Leu Pro Thr Ile Pro Gly Lys Leu Asp Val Asn Lys Ser
450 455 460

Lys Thr His Ile Ser Val Asn Gly Arg Lys Ile Arg Met Arg Cys Arg

465 470 475 480
 Ala Ile Asp Gly Asp Val Thr Phe Cys Arg Pro Lys Ser Pro Val Tyr
 485 490 495
 Val Gly Asn Gly Val His Ala Asn Leu His Val Ala Phe His Arg Ser
 500 505 510
 Ser Ser Glu Lys Ile His Ser Asn Glu Ile Ser Ser Asp Ser Ile Gly
 515 520 525
 Val Leu Gly Tyr Gln Lys Thr Val Asp His Thr Lys Val Asn Ser Lys
 530 535 540
 Leu Ser Leu Phe Phe Glu Ile Lys Ser Gly Ser His His His His His
 545 550 555 560

His

<210> 6
 <211> 535
 <212> PRT
 <213> Corynebacterium diphtheriae
 <220>
 <223> secreted mature CRM197 protein
 <400> 6

Gly Ala Asp Asp Val Val Asp Ser Ser Lys Ser Phe Val Met Glu Asn
 1 5 10 15
 Phe Ser Ser Tyr His Gly Thr Lys Pro Gly Tyr Val Asp Ser Ile Gln
 20 25 30
 Lys Gly Ile Gln Lys Pro Lys Ser Gly Thr Gln Gly Asn Tyr Asp Asp
 35 40 45
 Asp Trp Lys Glu Phe Tyr Ser Thr Asp Asn Lys Tyr Asp Ala Ala Gly
 50 55 60
 Tyr Ser Val Asp Asn Glu Asn Pro Leu Ser Gly Lys Ala Gly Gly Val
 65 70 75 80
 Val Lys Val Thr Tyr Pro Gly Leu Thr Lys Val Leu Ala Leu Lys Val
 85 90 95
 Asp Asn Ala Glu Thr Ile Lys Lys Glu Leu Gly Leu Ser Leu Thr Glu
 100 105 110
 Pro Leu Met Glu Gln Val Gly Thr Glu Glu Phe Ile Lys Arg Phe Gly
 115 120 125

103399PC sequ list.txt

Asp Gly Ala Ser Arg Val Val Leu Ser Leu Pro Phe Ala Glu Gly Ser
 130 135 140
 Ser Ser Val Glu Tyr Ile Asn Asn Trp Glu Gln Ala Lys Ala Leu Ser
 145 150 155 160
 Val Glu Leu Glu Ile Asn Phe Glu Thr Arg Gly Lys Arg Gly Gln Asp
 165 170 175
 Ala Met Tyr Glu Tyr Met Ala Gln Ala Cys Ala Gly Asn Arg Val Arg
 180 185 190
 Arg Ser Val Gly Ser Ser Leu Ser Cys Ile Asn Leu Asp Trp Asp Val
 195 200 205
 Ile Arg Asp Lys Thr Lys Thr Lys Ile Glu Ser Leu Lys Glu His Gly
 210 215 220
 Pro Ile Lys Asn Lys Met Ser Glu Ser Pro Asn Lys Thr Val Ser Glu
 225 230 235 240
 Glu Lys Ala Lys Gln Tyr Leu Glu Glu Phe His Gln Thr Ala Leu Glu
 245 250 255
 His Pro Glu Leu Ser Glu Leu Lys Thr Val Thr Gly Thr Asn Pro Val
 260 265 270
 Phe Ala Gly Ala Asn Tyr Ala Ala Trp Ala Val Asn Val Ala Gln Val
 275 280 285
 Ile Asp Ser Glu Thr Ala Asp Asn Leu Glu Lys Thr Thr Ala Ala Leu
 290 295 300
 Ser Ile Leu Pro Gly Ile Gly Ser Val Met Gly Ile Ala Asp Gly Ala
 305 310 315 320
 Val His His Asn Thr Glu Glu Ile Val Ala Gln Ser Ile Ala Leu Ser
 325 330 335
 Ser Leu Met Val Ala Gln Ala Ile Pro Leu Val Gly Glu Leu Val Asp
 340 345 350
 Ile Gly Phe Ala Ala Tyr Asn Phe Val Glu Ser Ile Ile Asn Leu Phe
 355 360 365
 Gln Val Val His Asn Ser Tyr Asn Arg Pro Ala Tyr Ser Pro Gly His
 370 375 380
 Lys Thr Gln Pro Phe Leu His Asp Gly Tyr Ala Val Ser Trp Asn Thr
 385 390 395 400

103399PC sequ list.txt

Val Glu Asp Ser Ile Ile Arg Thr Gly Phe Gln Gly Glu Ser Gly His
405 410 415

Asp Ile Lys Ile Thr Ala Glu Asn Thr Pro Leu Pro Ile Ala Gly Val
420 425 430

Leu Leu Pro Thr Ile Pro Gly Lys Leu Asp Val Asn Lys Ser Lys Thr
435 440 445

His Ile Ser Val Asn Gly Arg Lys Ile Arg Met Arg Cys Arg Ala Ile
450 455 460

Asp Gly Asp Val Thr Phe Cys Arg Pro Lys Ser Pro Val Tyr Val Gly
465 470 475 480

Asn Gly Val His Ala Asn Leu His Val Ala Phe His Arg Ser Ser Ser
485 490 495

Glu Lys Ile His Ser Asn Glu Ile Ser Ser Asp Ser Ile Gly Val Leu
500 505 510

Gly Tyr Gln Lys Thr Val Asp His Thr Lys Val Asn Ser Lys Leu Ser
515 520 525

Leu Phe Phe Glu Ile Lys Ser
530 535