

SEQUENCE LISTING

<110> MERCK SHARP & DOHME CORP.

Altman, Michael D.

Cash , Brandon D.

Chang, Wonsuk

Cumming, Jared N.

Haidle, Andrew M.

Henderson, Timothy J.

Jewell, James P.

Larsen, Matthew A.

Liang, Rui

Lim, Jongwon

Lu, Min

Otte, Ryan D.

Siu, Tony

Trotter, Benjamin Wesley

Tyagarajan, Sriram

<120> BENZO[b]THIOPHENE COMPOUNDS AS STING AGONISTS

<130> 24170

<150> 62/404,062

<151> 2016-10-04

<160> 4

<170> PatentIn version 3.5

<210> 1

<211> 414

<212> PRT

<213> Homo sapiens

<400> 1

Met Pro His Ser Ser Leu His Pro Ser Ile Pro Cys Pro Arg Gly His
1 5 10 15

Gly Ala Gln Lys Ala Ala Leu Val Leu Leu Ser Ala Cys Leu Val Thr
20 25 30

Leu Trp Gly Leu Gly Glu Pro Pro Glu His Thr Leu Arg Tyr Leu Val
35 40 45

Leu His Leu Ala Ser Leu Gln Leu Gly Leu Leu Leu Asn Gly Val Cys

50

55

60

Ser Leu Ala Glu Glu Leu His His Ile His Ser Arg Tyr Arg Gly Ser
65 70 75 80

Tyr Trp Arg Thr Val Arg Ala Cys Leu Gly Cys Pro Leu Arg Arg Gly
85 90 95

Ala Leu Leu Leu Leu Ser Ile Tyr Phe Tyr Tyr Ser Leu Pro Asn Ala
100 105 110

Val Gly Pro Pro Phe Thr Trp Met Leu Ala Leu Leu Gly Leu Ser Gln
115 120 125

Ala Leu Asn Ile Leu Leu Gly Leu Lys Gly Leu Ala Pro Ala Glu Ile
130 135 140

Ser Ala Val Cys Glu Lys Gly Asn Phe Asn Val Ala His Gly Leu Ala
145 150 155 160

Trp Ser Tyr Tyr Ile Gly Tyr Leu Arg Leu Ile Leu Pro Glu Leu Gln
165 170 175

Ala Arg Ile Arg Thr Tyr Asn Gln His Tyr Asn Asn Leu Leu Arg Gly
180 185 190

Ala Val Ser Gln Arg Leu Tyr Ile Leu Leu Pro Leu Asp Cys Gly Val
195 200 205

Pro Asp Asn Leu Ser Met Ala Asp Pro Asn Ile Arg Phe Leu Asp Lys
210 215 220

Leu Pro Gln Gln Thr Ala Asp Arg Ala Gly Ile Lys Asp Arg Val Tyr
225 230 235 240

Ser Asn Ser Ile Tyr Glu Leu Leu Glu Asn Gly Gln Arg Ala Gly Thr
245 250 255

Cys Val Leu Glu Tyr Ala Thr Pro Leu Gln Thr Leu Phe Ala Met Ser
260 265 270

Gln Tyr Ser Gln Ala Gly Phe Ser Arg Glu Asp Arg Leu Glu Gln Ala
275 280 285

Lys Leu Phe Cys Gln Thr Leu Glu Asp Ile Leu Ala Asp Ala Pro Glu
290 295 300

Ser Gln Asn Asn Cys Arg Leu Ile Ala Tyr Gln Glu Pro Ala Asp Asp
305 310 315 320

Ser Ser Phe Ser Leu Ser Gln Glu Val Leu Arg His Leu Arg Gln Glu
325 330 335

Glu Lys Glu Glu Val Thr Val Gly Ser Leu Lys Thr Ser Ala Val Pro
340 345 350

Ser Thr Ser Thr Met Ser Gln Glu Pro Glu Leu Leu Ile Ser Gly Met
355 360 365

Glu Lys Pro Leu Pro Leu Arg Thr Asp Phe Ser Gly Gly Gly Leu Asn
370 375 380

Asp Ile Phe Glu Ala Gln Lys Ile Glu Trp His Glu Gly Ser Leu Glu
385 390 395 400

Val Leu Phe Gln Gly Pro His His His His His His His His
405 410

<210> 2
<211> 6482
<212> DNA
<213> Homo sapiens

<400> 2
ggaacggctc cgcccactat taatgaaatt aaaaattcca attttaaaaa acgcagcaag 60

agaaacattt gtatgaaaga atgcgtagaa ggaaagaaaa atgtcgtcga catgctgaac 120

aacaagatta atatgcctcc gtgtataaaa aaaatattga acgatttgaa agaaaacaat	180
gtaccgcgcg gcggtatgta caggaagagg tttatactaa actgttacat tgcaaacgtg	240
gtttcgtgtg ccaagtgtga aaaccgatgt ttaatcaagg ctctgacgca tttctacaac	300
cacgactcca agtgtgtggg tgaagtcatg catcttttaa tcaaatccca agatgtgtat	360
aaaccaccaa actgccaaaa aatgaaaact gtcgacaagc tctgtccgtt tgctggcaac	420
tgcaagggtc tcaatcctat ttgtaattat tgaataataa aacaattata aatgctaaat	480
ttgtttttta ttaacgatac aaaccaaacg caacaagaac atttgtagta ttatctataa	540
ttgaaaacgc gtagttataa tcgctgaggt aatatTTaaa atcattttca aatgattcac	600
agttaatttg cgacaatata attttatttt cacataaact agacgccttg tcgtcttctt	660
cttcgtattc cttctctttt tcatttttct cttcataaaa attaacatag ttattatcgt	720
atccatatat gtatctatcg tatagagtaa attttttggt gtcataaata tatatgtctt	780
ttttaatggg gtgtatagta ccgctgcgca tagtttttct gtaatttaca acagtgtctat	840
tttctggtag ttcttcggag tgtgttgctt taattattaa atttatataa tcaatgaatt	900
tgggatcgtc ggTTTTgtac aatatgttgc cggcatagta cgcagcttct tctagttcaa	960
ttacaccatt ttttagcagc accggattaa cataactttc caaaatgttg tacgaaccgt	1020
taaacaaaaa cagttcacct cctttttcta tactattgtc tgcgagcagt tgtttgttgt	1080
taaaaataac agccattgta atgagacgca caaactaata tcacaaactg gaaatgtcta	1140
tcaatatata gttgctgac agatctgac atggagataa ttaaaatgat aaccatctcg	1200
caaataaata agtatTTTtac tgttttcgta acagtTTTgt aataaaaaaa cctataaata	1260
taggatccat gccccactcc agcctgcac catccatccc gtgtcccagg ggtcacgggg	1320
cccagaaggc agccttggtt ctgctgagtg cctgcctggt gacccttttg gggctaggag	1380
agccaccaga gcacactctc cggtacctgg tgctccacct agcctccctg cagctgggac	1440
tgctgttaaa cggggctctgc agcctggctg aggagctgca ccacatccac tccaggtacc	1500
ggggcagcta ctggaggact gtgcgggcct gcctgggctg cccctccgc cgtggggccc	1560
tgttgctgct gtccatctat ttctactact cctcccaaa tgcggtcggc ccgcccttca	1620

cttggatgct tgccctcctg ggccctctgc aggcactgaa catcctcctg ggccctcaagg	1680
gcctggcccc agctgagatc tctgcagtgt gtgaaaaagg gaatttcaac gtggcccatg	1740
ggctggcatg gtcataattac atcggatatc tgcggctgat cctgccagag ctccaggccc	1800
ggattcgaac ttacaatcag cattacaaca acctgctacg ggggtgcagtg agccagcggc	1860
tgtatatctt cctcccattg gactgtgggg tgcctgataa cctgagtatg gctgacccca	1920
acattcgctt cctggataaa ctgccccagc agaccgctga ccgtgctggc atcaaggatc	1980
gggtttacag caacagcatc tatgagcttc tggagaacgg gcagcgggcg ggcacctgtg	2040
tcctggagta cgccaccccc ttgcagactt tgtttgccat gtcacaatac agtcaagctg	2100
gcttttagccg ggaggatagg cttgagcagg ccaaactctt ctgccagaca cttgaggaca	2160
tcctggcaga tgcccctgag tctcagaaca actgccgcct cattgcctac caggaaacctg	2220
cagatgacag cagcttctcg ctgtcccagg aggttctccg gcacctgcgg caggaggaaa	2280
aggaagaggt tactgtgggc agcttgaaga cctcagcggg gcccagtacc tccacgatgt	2340
cccaagagcc tgagctcctc atcagtggaa tggaaaagcc cctccctctc cgcacggatt	2400
tctctggcgg tggcctgaac gacatcttcg aagcccagaa aatcgaatgg catgaaggca	2460
gcctggaagt gctgttccag ggcccacacc accatcatca ccatcaccat taatgagcgg	2520
ccgcactcga gcaccaccac caccaccact aacctaggta gctgagcgca tgcaagctga	2580
tccgggttat tagtacattt attaagcgct agattctgtg cgttgttgat ttacagacaa	2640
ttgttgtagc tattttaata attcattaaa ttataatct ttaggggtgg atgttagagc	2700
gaaaatcaaa tgattttcag cgtctttata tctgaattta aatattaaat cctcaataga	2760
tttgtaaaat aggtttcgat tagtttcaaa caagggttgt ttttccgaac cgatggctgg	2820
actatctaata ggattttcgc tcaacgccac aaaacttgcc aaatcttgta gcagcaatct	2880
agctttgtcg atattcgttt gtgttttggt ttgtaataaa gggtcgacgt cgttcaaaat	2940
attatgcgct tttgtatttc tttcatcact gtcgttagtg tacaattgac tcgacgtaaa	3000
cacgttaaat agagcttgga catatttaac atcgggcgtg ttagctttat taggccgatt	3060
atcgtcgtcg tcccaaccct cgtcgttaga agttgcttcc gaagacgatt ttgccatagc	3120

cacacgacgc ctattaattg tgtcggctaa cacgtccgcg atcaaatttg tagttgagct	3180
ttttggaatt atttctgatt gcgggcgttt ttgggcgggt ttcaatctaa ctgtgcccga	3240
ttttaattca gacaacacgt tagaaagcga tgggtgcaggc ggtggtaaca tttcagacgg	3300
caaatctact aatggcggcg gtgggtggagc tgatgataaa tctaccatcg gtggaggcgc	3360
aggcggggct ggccggcgag gcggaggcgg aggtggtggc ggtgatgcag acggcggttt	3420
aggctcaaat gtctcttttag gcaacacagt cggcacctca actattgtac tggtttcggg	3480
cgccgttttt ggtttgaccg gtctgagacg agtgcgattt ttttcgtttc taatagcttc	3540
caacaattgt tgtctgtcgt ctaaagggtgc agcgggttga ggttccgtcg gcattgggtg	3600
agcgggcggc aattcagaca tcgatggtgg tgggtggtgg ggaggcgctg gaatgttagg	3660
cacgggagaa ggtgggtggcg gcggtgccgc cgggtataatt tgttctgggt tagtttgttc	3720
gcgcacgatt gtgggcaccg gcgcaggcgc cgctggctgc acaacggaag gtcgtctgct	3780
tcgaggcagc gcttgggggtg gtggcaattc aatattataa ttggaataca aatcgtaaaa	3840
atctgctata agcattgtaa tttcgctatc gtttaccgtg ccgatattta acaaccgctc	3900
aatgtaagca attgtattgt aaagagattg tctcaagctc ggatcgatcc cgcacgccga	3960
taacaagcct tttcattttt actacagcat tgtagtggcg agacacttcg ctgtcgtcga	4020
ggtttaaacg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg	4080
gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga	4140
aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg	4200
gcgtttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag	4260
aggtggcgaa acccgacagg actataaaga taccaggcgt tttcccctgg aagctccctc	4320
gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg	4380
ggaagcgtgg cgcttttctca tagctcacgc tgtaggtatc tcagttcggg gtaggtcggt	4440
cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc	4500
ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc	4560
actggtaaca ggattagcag agcgagggtat gtaggcgggtg ctacagagtt cttgaagtgg	4620

tggcctaact acggctacac tagaaggaca gtatTTtgta tctgcgctct gctgaagcca	4680
gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc	4740
ggtggTTTTT ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat	4800
cctttgatct tttctacggg gtctgacgct cagtggaaacg aaaactcacg ttaagggatt	4860
ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt	4920
tttaaatacaa tctaaagtat atatgagtaa acttggctctg acagttacca atgcttaatc	4980
agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc	5040
gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtg tgcaatgata	5100
ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg	5160
gccgagcgca gaagtggTcc tgcaacttta tccgcctcca tccagtctat taattgttgc	5220
cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct	5280
acaggcatcg tgggtgtcacg ctcgctgttt ggtatggctt cattcagctc cggttcccaa	5340
cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt	5400
cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca	5460
ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac	5520
tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca	5580
atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt	5640
tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc	5700
actcgtgcac ccaactgac ttcagcatct tttactttca ccagcgtttc tgggtgagca	5760
aaaacaggaa ggcaaaatgc cgcaaaaaaag ggaataaggg cgacacggaa atgttgaata	5820
ctcatactct tcctttttca atattattga agcatttatc agggttattg tctcatgagc	5880
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc	5940
cgaaaagtgc cacctgacgc gccctgtagc ggcgcattaa gcgcggcggg tgtgggtggtt	6000
acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc	6060
ccttcctttc tcgccacgtt cgccggcttt ccccgTcaag ctctaaatcg ggggctccct	6120

ttagggttcc gatttagtgc tttagcgac ctcgaccca aaaaacttga ttagggtgat	6180
ggttcacgta gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc	6240
acgttcttta atagtggact cttgttccaa actggaacaa cactcaaccc tatctcggtc	6300
tattcttttg atttataagg gattttgccg atttcggcct attgggttaaa aaatgagctg	6360
atttaacaaa aatttaacgc gaattttaac aaaatattaa cgtttacaat ttcccattcg	6420
ccattcaggc tgcgcaactg ttgggaaggg cgatcgggtgc gggcctcttc gctattacgc	6480
ca	6482

<210> 3
 <211> 414
 <212> PRT
 <213> Homo sapiens

<400> 3

Met	Pro	His	Ser	Ser	Leu	His	Pro	Ser	Ile	Pro	Cys	Pro	Arg	Gly	His
1				5					10					15	

Gly	Ala	Gln	Lys	Ala	Ala	Leu	Val	Leu	Leu	Ser	Ala	Cys	Leu	Val	Thr
		20						25					30		

Leu	Trp	Gly	Leu	Gly	Glu	Pro	Pro	Glu	His	Thr	Leu	Arg	Tyr	Leu	Val
		35					40					45			

Leu	His	Leu	Ala	Ser	Leu	Gln	Leu	Gly	Leu	Leu	Leu	Asn	Gly	Val	Cys
	50					55					60				

Ser	Leu	Ala	Glu	Glu	Leu	Arg	His	Ile	His	Ser	Arg	Tyr	Arg	Gly	Ser
65					70					75					80

Tyr	Trp	Arg	Thr	Val	Arg	Ala	Cys	Leu	Gly	Cys	Pro	Leu	Arg	Arg	Gly
			85						90					95	

Ala	Leu	Leu	Leu	Leu	Ser	Ile	Tyr	Phe	Tyr	Tyr	Ser	Leu	Pro	Asn	Ala
			100					105					110		

Val Gly Pro Pro Phe Thr Trp Met Leu Ala Leu Leu Gly Leu Ser Gln
115 120 125

Ala Leu Asn Ile Leu Leu Gly Leu Lys Gly Leu Ala Pro Ala Glu Ile
130 135 140

Ser Ala Val Cys Glu Lys Gly Asn Phe Asn Val Ala His Gly Leu Ala
145 150 155 160

Trp Ser Tyr Tyr Ile Gly Tyr Leu Arg Leu Ile Leu Pro Glu Leu Gln
165 170 175

Ala Arg Ile Arg Thr Tyr Asn Gln His Tyr Asn Asn Leu Leu Arg Gly
180 185 190

Ala Val Ser Gln Arg Leu Tyr Ile Leu Leu Pro Leu Asp Cys Gly Val
195 200 205

Pro Asp Asn Leu Ser Met Ala Asp Pro Asn Ile Arg Phe Leu Asp Lys
210 215 220

Leu Pro Gln Gln Thr Gly Asp Arg Ala Gly Ile Lys Asp Arg Val Tyr
225 230 235 240

Ser Asn Ser Ile Tyr Glu Leu Leu Glu Asn Gly Gln Arg Ala Gly Thr
245 250 255

Cys Val Leu Glu Tyr Ala Thr Pro Leu Gln Thr Leu Phe Ala Met Ser
260 265 270

Gln Tyr Ser Gln Ala Gly Phe Ser Arg Glu Asp Arg Leu Glu Gln Ala
275 280 285

Lys Leu Phe Cys Arg Thr Leu Glu Asp Ile Leu Ala Asp Ala Pro Glu
290 295 300

Ser Gln Asn Asn Cys Arg Leu Ile Ala Tyr Gln Glu Pro Ala Asp Asp

305 310 315 320

Ser Ser Phe Ser Leu Ser Gln Glu Val Leu Arg His Leu Arg Gln Glu
325 330 335

Glu Lys Glu Glu Val Thr Val Gly Ser Leu Lys Thr Ser Ala Val Pro
340 345 350

Ser Thr Ser Thr Met Ser Gln Glu Pro Glu Leu Leu Ile Ser Gly Met
355 360 365

Glu Lys Pro Leu Pro Leu Arg Thr Asp Phe Ser Gly Gly Gly Leu Asn
370 375 380

Asp Ile Phe Glu Ala Gln Lys Ile Glu Trp His Glu Gly Ser Leu Glu
385 390 395 400

Val Leu Phe Gln Gly Pro His His His His His His His His
405 410

<210> 4
<211> 6482
<212> DNA
<213> HOMO SAPIENS

<400> 4
ggaacggctc cgcccactat taatgaaatt aaaaattcca attttaaaaa acgcagcaag 60
agaaacattt gtatgaaaga atgcgtagaa ggaaagaaaa atgtcgtcga catgctgaac 120
aacaagatta atatgcctcc gtgtataaaa aaaatattga acgatttgaa agaaaacaat 180
gtaccgcgcg gcggtatgta caggaagagg ttataactaa actgttacat tgcaaactg 240
gtttcgtgtg ccaagtgtga aaaccgatgt ttaatcaagg ctctgacgca tttctacaac 300
cacgactcca agtgtgtggg tgaagtcatg catcttttaa tcaaattcca agatgtgtat 360
aaaccaccaa actgccaaaa aatgaaaact gtcgacaagc tctgtccgtt tgctggcaac 420
tgcaagggtc tcaatcctat ttgtaattat tgaataataa aacaattata aatgtcaa 480
ttgtttttta ttaacgatac aaaccaaacg caacaagaac attttagtagta ttatctataa 540

ttgaaaacgc gtagttataa tcgctgaggt aatattttaa atcattttca aatgattcac	600
agttaatttg cgacaatata attttatttt cacataaaact agacgccttg tcgtcttctt	660
cttcgtattc cttctctttt tcatttttct cttcataaaa attaacatag ttattatcgt	720
atccatatat gtatctatcg tatagagtaa attttttggt gtcataaata tatatgtctt	780
ttttaatggg gtgtatagta ccgctgcgca tagtttttct gtaatttaca acagtgcctat	840
tttctggtag ttcttcggag tgtgttgctt taattattaa atttatataa tcaatgaatt	900
tgggatcgtc ggttttgtac aatatgttgc cggcatagta cgcagcttct tctagttcaa	960
ttacaccatt ttttagcagc accggattaa cataactttc caaaatgttg tacgaaccgt	1020
taaacaaaaa cagttcacct cctttttcta tactattgtc tgcgagcagt tgtttgttgt	1080
taaaaataac agccattgta atgagacgca caaactaata tcacaaactg gaaatgtcta	1140
tcaatatata gttgctgac agatctgac atggagataa ttaaaatgat aaccatctcg	1200
caaataaata agtattttac tgttttcgta acagttttgt aataaaaaaa cctataaata	1260
taggatccat gccccactcc agcctgcac catccatccc gtgtcccagg ggtcacgggg	1320
cccagaaggc agccttggtt ctgctgagtg cctgcctggt gaccctttgg gggctaggag	1380
agccaccaga gcacactctc cggtaacctg tgctccacct agcctccctg cagctgggac	1440
tgctgttaaa cggggctctg agcctggctg aggagctgcg ccacatccac tccaggtacc	1500
ggggcagcta ctggaggact gtgcgggcct gcctgggctg cccctccgc cgtggggccc	1560
tgttgctgct gtccatctat ttctactact cctcccaaa tgcggtcggc ccgcccctca	1620
cttggatgct tgccctcctg ggcctctgc aggcactgaa catcctcctg ggcctcaagg	1680
gcctggcccc agctgagatc tctgcagtgt gtgaaaaagg gaatttcaac gtggcccatg	1740
ggctggcatg gtcattattac atcggatatc tgcggctgat cctgccagag ctccaggccc	1800
ggattcgaac ttacaatcag cattacaaca acctgctacg gggcgcagt agccagcggc	1860
tgtatattct cctcccattg gactgtgggg tgcctgataa cctgagtatg gctgaccca	1920
acattcgctt cctggataaa ctgccccagc agaccggtga ccgtgctggc atcaaggatc	1980
gggtttacag caacagcatc tatgagcttc tggagaacgg gcagcgggcg ggcacctgtg	2040

tcctggagta cgccaccccc ttgcagactt tgtttgccat gtcacaatac agtcaagctg	2100
gcttttagccg ggaggatagg cttgagcagg ccaaactctt ctgccggaca cttgaggaca	2160
tcctggcaga tgccccctgag tctcagaaca actgccgcct cattgcctac caggaacctg	2220
cagatgacag cagcttctcg ctgtcccagg aggttctccg gcacctgcgg caggaggaaa	2280
aggaagaggt tactgtgggc agcttgaaga cctcagcggg gcccagtacc tccacgatgt	2340
cccaagagcc tgagctcctc atcagtggaa tggaaaagcc cctccctctc cgcacggatt	2400
tctctggcgg tggcctgaac gacatcttcg aagcccagaa aatcgaatgg catgaaggca	2460
gcctggaagt gctgttccag ggcccacacc accatcatca ccatcaccat taatgagcgg	2520
ccgcactcga gcaccaccac caccaccact aacctaggta gctgagcgcga tgcaagctga	2580
tccgggttat tagtacattt attaagcgct agattctgtg cgttgttgat ttacagacaa	2640
ttgttgtagc tattttaata attcattaaa ttataatct ttaggggtgg atgttagagc	2700
gaaaatcaaa tgattttcag cgtctttata tctgaattta aatattaaat cctcaataga	2760
tttgtaaaat aggtttcgat tagtttcaaa caagggttgt ttttccgaac cgatggctgg	2820
actatctaata ggattttcgc tcaacgccac aaaacttgcc aaatcttgta gcagcaatct	2880
agctttgtcg atattcgttt gtgttttggt ttgtaataaa gggtcgacgt cgttcaaaat	2940
attatgcgct tttgtatttc tttcatcact gtcgttagtg tacaattgac tcgacgtaaa	3000
cacgttaaat agagcttgga catatttaac atcgggcgtg ttagctttat taggccgatt	3060
atcgtcgtcg tcccaaccct cgtcgttaga agttgcttcc gaagacgatt ttgccatagc	3120
cacacgacgc ctattaattg tgtcggctaa cacgtccgcg atcaaatttg tagttgagct	3180
ttttggaatt atttctgatt gcgggcgttt ttgggcgggt ttcaatctaa ctgtgcccga	3240
ttttaattca gacaacacgt tagaaagcga tgggtgcaggc ggtggtaaca tttcagacgg	3300
caaatctact aatggcggcg gtgggtggagc tgatgataaa tctaccatcg gtggaggcgc	3360
aggcggggct ggccggcgag gcggaggcgg aggtgggtggc ggtgatgcag acggcggttt	3420
aggctcaaat gtctcttttag gcaacacagt cggcacctca actattgtac tggtttcggg	3480
cgccgttttt ggtttgaccg gtctgagacg agtgcgattt ttttcgtttc taatagcttc	3540

caacaattgt tgtctgtcgt ctaaagggtgc agcgggttga ggttccgtcg gcattgggtgg	3600
agcggggcggc aattcagaca tcgatgggtg tggtgggtgg ggaggcgctg gaatgttagg	3660
cacgggagaa ggtgggtggcg gcgggtgccgc cgggtataatt tgttctgggt tagtttggttc	3720
gcgcacgatt gtggggcaccg gcgcaggcgc cgctggctgc acaacggaag gtcgtctgct	3780
tcgaggcagc gcttgggggtg gtggcaattc aatattataa ttggaatata aatcgtaaaa	3840
atctgctata agcattgtaa tttcgctatc gtttaccgtg ccgatattta acaaccgctc	3900
aatgtaagca attgtattgt aaagagattg tctcaagctc ggatcgatcc cgcacgccga	3960
taacaagcct tttcatTTTT actacagcat tgtagtggcg agacacttcg ctgtcgtcga	4020
ggtttaaacg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg	4080
gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga	4140
aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg	4200
gcgtttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag	4260
agggtggcgaa acccgacagg actataaaga taccaggcgt tttccccctgg aagctccctc	4320
gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg	4380
ggaagcgtgg cgcttttctca tagctcacgc tgtaggtatc tcagttcggg gtaggtcggt	4440
cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc	4500
ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc	4560
actggtaaca ggattagcag agcgaggatg gtaggcgggtg ctacagagtt cttgaagtgg	4620
tggcctaact acggctacac tagaaggaca gtatttggtg tctgcgctct gctgaagcca	4680
gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc	4740
ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat	4800
cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggtt	4860
ttgggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt	4920
tttaaatacaa tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc	4980
agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc	5040

gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata	5100
ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg	5160
gccgagcgca gaagtgggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc	5220
cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct	5280
acaggcatcg tgggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa	5340
cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcggtttag ctccttcggt	5400
cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca	5460
ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac	5520
tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca	5580
atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt	5640
tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc	5700
actcgtgcac ccaactgac ttcagcatct tttactttca ccagcgtttc tgggtgagca	5760
aaaacaggaa ggcaaaatgc cgcaaaaaaag ggaataaggg cgacacggaa atgttgaata	5820
ctcatactct tcctttttca atattattga agcattttatc agggttattg tctcatgagc	5880
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc	5940
cgaaaagtgc cacctgacgc gccctgtagc ggcgcattha gcgcggcggg tgtgggtggt	6000
acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc	6060
ccttcctttc tcgccacgtt cgccggcttt ccccgtaag ctctaaatcg ggggctcct	6120
ttagggttcc gatttagtgc tttacggcac ctcgaccca aaaaacttga ttagggatgat	6180
ggttcacgta gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc	6240
acgttcttta atagtggact cttgttccaa actggaacaa cactcaacc tatctcggtc	6300
tattcttttg atttataagg gatcttgccg atttcggcct attgggttaa aaatgagctg	6360
atttaacaaa aatttaacgc gaattttaac aaaatattaa cgtttacaat ttcccatcgc	6420
ccattcaggc tgcgcaactg ttgggaaggg cgatcgggtgc gggcctcttc gctattacgc	6480
ca	6482

