

Project Name: 240119

Status: generated
Creation date: 2024-01-19
Last modified: 2024-01-19

General Information

Current application

Application number: before

IP Office:
Application number:
Filing date:
Applicant file reference: ESP1V231585JW-US

Invention Title

Language	Invention title
en	UNNATURAL AMINO ACID, AND USE THEREOF, RECOMBINANT PROTEIN CONTAINING SAME, AND RECOMBINANT PROTEIN CONJUGATE

Applicant & Inventor:

Applicant name: Novocodex Biopharmaceuticals Co., Ltd.
Language: en
Name Latin:
Residence Address:
Correspondence Address:

Sequences

Sequence 1: "Growth hormone"

Length	Molecule Type	Organism	Contains DNA and RNA fragments	Skipped Sequence
191	AA	Homo sapiens	No	No

Features

Feature Key	Feature Location	Qualifiers
source	1..191	mol_type = protein organism = Homo sapiens

Residues

FPTIPLSRFL	DNAMLAHRL	HOLAFDTYOE	FEEAYIPKEO	KYSFLONPOT	SLCFSES IPT	60
PSNREETOOK	SNLELLRISL	LLIQSWLEPV	OFLRSVFANS	LVYGASDSNV	YDLLKDLEEG	120
IQTLMGRLED	GSPRTGQIFK	QTYSKFDTNS	HNDDALLKNY	GLLYCFRKDM	DKVETFLRIV	180
QCRSVEGSCG	F					191

Sequence 2: "Recombinant human growth hormone"

Length	Molecule Type	Organism	Contains DNA and RNA fragments	Skipped Sequence
603	DNA	synthetic construct	No	No

Features

Feature Key	Feature Location	Qualifiers
misc_feature	1..603	note = Recombinant human growth hormone
source	1..603	mol_type = other DNA organism = synthetic construct

Residues

```

atgtttccta ctataccact atctcgtcta ttcgataacg ctatgcttcg ggcccatcgt 60
cttcatcagc tggcctttga cacctaccag gagtttgaag aagcctatat cccaaaggaa 120
cagaagtatt cattcctgca gaacccccag acctccctct gtttctcaga gtctatttccg 180
acaccctcca acaggaggga aacacaacag aaatccaacc tagagctgct cgcctctctcc 240
ctgctgctca tccagtctgt gctggagccc gtgcagttcc tcaggagtgt cttcgccaac 300
agcctggtgt acggcgcttc ttagagcaac gtctatgacc tcctaaaggga cctagaggaa 360
ggcatccaaa cgctgatggg gaggctggaa gatggcagcc cccggactgg gcagatcttc 420
aagcagacct acagcaagtt cgaacaaaac tcacacaacg atgacgact actcaagaac 480
tacgggctgc tctactgctt caggaaggac atggacaagg tcgagacatt cctgcgcac 540
gtgcagtgcc gctctgtgga gggcagctgt ggcttcctcg agcaccacca ccaaccacc 600
taa                                     603

```

Sequence 3: "Wild-type green fluorescent protein containing purification tag"

Length	Molecule Type	Organism	Contains DNA and RNA fragments	Skipped Sequence
753	DNA	synthetic construct	No	No

Features

Feature Key	Feature Location	Qualifiers
misc_feature	1..753	note = Wild-type green fluorescent protein containing the purification tag
source	1..753	mol_type = other DNA organism = synthetic construct

Residues

```

atgggtgagca agggcgagga gctgttcacc ggggtggtgc ccatcctggt cgagctggac 60
ggcgacgttaa acggccacaa gttcagcgtg tccggcgaag gcgagggcga tgcacctaag 120
ggcaagctga cctgaagtt catctgcacc accggcaagc tgcccggtgc ctggccacc 180
ctcgtgacca cctgacctt cggcgtgcag tgcttcagcc gctaccccga ccacatgaag 240
cagcacgact tcttcaagtc cggcatggcc gaaggctacg tccaggagcg caccatcttc 300
ttcaagcgag acggcaacta cgaacccccc gccgaggtga agttcgaggg cgacacccctc 360
gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcacaac cctggggcac 420
aagctggagt acaactacaa cagccacaac gtctatatac tggccgacaa gcagaagaac 480
ggcatcaagg agcagaacac ccccatcggc gacggccccc tgctgctgcc cgacaaccac 540
gaccactacc cccagtccgc cctgagcaaa gaccccaacg agaagcgcga tcacatgttc 600
ctgctggagt cgtgaccgc cgccgggata actctcggca tggacgagct gtacaagaag 660
tacgaagatt acaaggatga cgacgataag taa                                     720
                                     753

```

Sequence 4: "Expression plasmid pEGFP40"

Length	Molecule Type	Organism	Contains DNA and RNA fragments	Skipped Sequence
6170	DNA	synthetic construct	No	No

Features

Feature Key	Feature Location	Qualifiers
misc_feature	1..6170	note = Expression plasmid pEGFP40
source	1..6170	mol_type = other DNA organism = synthetic construct

Residues

gttaggcggt	ttgcgcgtgct	tcgcgatgta	cgggccagat	atacgcgttg	acattgatta	60
ttgactagtt	atfaaatagt	atcaattacg	gggtcattag	ttcatagcc	atataatggag	120
ttccgcgtta	cataactfac	ggtaaatggc	ccgcctggct	gaccgcgcc	cgaccccgcc	180
ccattgacgt	caataatgac	gtatgttccc	atagtaacgc	caataaggac	tttccattga	240
cgtcaatggg	tggagctatt	acggtaaaact	gcccacttgg	cagtaacatca	agtgtatcat	300
atgccaaagta	cgccccctat	tgacgtcaaat	gacggtaaaat	ggcccgccctg	gcaattatgcc	360
cagtacatga	ccttatggga	ctttcctact	tggcagtaga	tctacgtatt	agtcacgtct	420
attaccatgg	tgtatggcgt	ttggcagtag	atcaatgggg	gtggatagcg	gtttgactca	480
cgggggatttc	caagtcctcca	ccccattgac	gtcaatggga	gtttgttttg	gcaccaaaaat	540
caaacgggact	ttcccaaaaat	tcgttaacaa	tccgccccat	tgaacgaaaat	gggcggtagg	600
cgtgtacgggt	gggagggtcta	tataagcaga	gctcgtttta	tgaaccgta	gatcgccctgg	660
agacgcgcac	cacgctgtttt	tgacctccat	agaaagacacc	gggacccgatc	cagccctccgg	720
actctagagg	atcccccgcca	ccatgggtgag	caaggggcgag	gagctgttca	ccgggggtgg	780
gccccatcctg	gtcggagctgg	acgggcgacgt	aaacggccac	aaagtccagcg	tgtccggcga	840
ggggcgagggg	gatggccacct	agggcgaagct	gacctgaaag	ttcatctgca	ccaccggcga	900
gctgcgccctg	cccttggccca	ccctcgtgac	cacctgacc	tacggcgtgc	agtgtctcag	960
ccgctacccc	gaccacatga	acagcagcga	cttcttcaag	ggcgcctatg	ccgaaggcta	1020
cgtccagggag	cgccaccatct	tgttcaaggga	cgacggcaac	tacaagacc	gcgcggaggt	1080
gaagtctcag	ggcgacaccc	gtcggaaccg	catcgagctc	aaaggcgatcg	acttcaaggga	1140
ggacggcaca	atcctggggg	acaaagcgtga	gtacaactac	aaacagccaca	acgtctatat	1200
catggcccgac	aagcagaaga	acggcatcaa	ggtgaacttc	aaagatccgcc	acaacatcga	1260
ggacggcgac	gtgcagctcg	ccgaccacta	ccagcagaac	acccccatcg	gcgacggccc	1320
cgtgctcgtg	cccgaacaac	actacactgag	caccagctcc	ccgctgagca	aaagacccca	1380
cgagaaagcgc	gatcacatgg	tctcgtcgtga	gttcgtgacc	ggccgcggga	tcaactctcgg	1440
catggacagag	gtgtacaaga	agtacgaaga	ttacaaggat	gacgacgata	agtaagaatt	1500
ccggaaagggg	tccgacccta	ccgggttagta	atgagtttaa	acggggggagg	ctaactgaaa	1560
cacgggaaggga	gacaaatccg	gaagggaaccc	gcgctatgac	ggcaataaaa	agacagaata	1620
aaaacgcacgg	gtgttgggttc	gtttgtttcat	aaacgcgggg	ttcgggtccca	gggctggcac	1680
tctgttcgata	ccccacccgag	accccatggg	ggccaataacg	cccgcgtttc	ttccttttcc	1740
ccaccccacc	ccccaaagttc	gggtgaagggt	ccagggtctcg	cagcccaacgt	cgggggcgga	1800
ggccctcgcca	tagcagatctt	gcgcgagctgg	ggctctaggg	gggtatcccca	cgcccccctgt	1860
agcggcgcat	taagcggcgc	gggtgtgggtg	gttaacgcgca	gcgtgacccg	tacacttgct	1920
atgcgccctag	cgcccgctccc	tttcggtcttc	ttcccttctct	ttctcgccac	gttcgcggcg	1980
tttcccccgtc	aaagctctaaa	cccgggcatct	cttttaggggt	cccttagttag	tgctttacgg	2040
cacctcgacc	ccaaaataact	tgatttaggg	gatgggttcc	gtagtggggc	atcgccctga	2100
tagacgtttt	ttcgccctttt	gcagtttggag	tccacgtttct	tttaatagtg	actcttgttc	2160
caaaactggaa	caacactcaa	ccccatcttcg	gtctattctt	ttgatttata	agggatttttg	2220
gggattttcgg	cctatttggtt	ccaaaatgag	ctgattttaac	aaaaatttta	gcggaatttaa	2280
ttctgttgaaa	tgtgtgtcag	tttaggtgtg	gaaaagtcgcc	aggctcccca	gcagggcagaa	2340
gtatgcaaaag	catgcacttc	aattagtcaag	caaccaggtg	tggaaaagtc	ccaggtctcc	2400
cagcaggcgag	aagtatgcgaa	agcatgcata	tcaatttagtc	agcaaccata	gtcccgcccc	2460
taactccggc	cafcccgccc	ctcaactccgc	ccagttccgc	ccattctccg	ccccatggct	2520
gactaatttt	ttttattttt	cgagaggccg	aggccgcctc	tgccctctgag	ctattccaga	2580
agttagtgag	aggctttttt	ggagggcctag	gcttttgcaa	aaagctcccg	ggagcttgta	2640
tatccattgca	cgcatgtgat	caagagacag	gatgaggatc	gtttcgcatg	attgaacaaag	2700
atggattgca	cgaggtttct	ccggcccgctt	gggtggagag	gctattcggc	tatgactggg	2760
cacaaacagac	aatcggtctg	tctgtagccg	ccgtgtttccg	gctgtcagcg	caggggcgcc	2820
ccggtttcttt	tgtcaagaac	gacctgtccg	gtgccctgaa	tgaactgca	gacgagcgag	2880
cgcggtctatc	gtgggttgccc	acgacggggc	ttccttgccg	agctgtgctc	gacgttgtca	2940
ctgaagccggg	aaagggaactg	ctgtcatttg	gcgaagtgcc	ggggcaggat	ctcctgtcat	3000
ctcacctctgc	tcctgcccag	aaagtatccca	tcattggctga	tgcattgcgg	cggtcgcata	3060
cgcttgatcc	ggctaacctgt	ccattcgagc	accaaagcgaa	acatcgcgac	gagcggagcac	3120
gfactcggat	ggaaagccgct	cttgtcgatc	aggatgatct	ggacgaaaga	catcaggggc	3180
tccgcccagc	cgaaactgttc	gccaggtctca	aggcggcgat	ggccgacggc	gaggatctctg	3240
tcgtgaccca	tggcgatggc	tgcttgccga	atatcatggt	ggaaaatggc	cgcttttctg	3300
gatttcatcga	ctgttgccgg	ctgggtgttg	cggaacgcta	ctcaggacata	gcgtttggcta	3360
cccggtgatat	tgttgaaagag	cttggcgccg	aatgggctga	ccgcttctct	gtgcttttacg	3420
gtatcgcggc	ctccgatttcg	cagcgccttcg	ccctctatcg	cccttcttgc	gagttcttct	3480
gagcgggact	tccgggttcg	gccaaatgacc	gaccaaagcg	cgcccaacct	gccatcacga	3540
gattttcgatt	ccaccggccg	cttctatgaa	aggttgggct	tccgaatcgt	tttccgggac	3600
gcccggctgga	tgtactctcca	gcgcggggat	ctcatgctgg	agttcttccg	ccaccccac	3660
ttgtttattg	cagctctataa	tggtttacaaa	taaaagcaat	gcatacaca	tttcacaaat	3720
aaagcatttt	tttccactgaa	ttctagtgtt	ggtttgctca	aaactcatca	tgatctttat	3780
catgtctgta	taccgtcgac	ctctcgttag	agctttggcg	aatcatgttc	atagctgttt	3840
ctcgtgtgaa	attgttatcc	gctcacaaat	ccacacaaac	tacgagcccg	aagcataaaa	3900
tgtaaaagcct	gggggtgccta	atgagtggag	taactcacat	taattgcggt	gcgctcactg	3960
cccgctctcc	agttcgggaaa	ctcgtctgtc	cagctgcatt	aatgaatcgg	ccaacgcggc	4020
gggagagggc	gttttgctgat	tgggcgctct	tccgcttctc	cgctcactga	ctcgcgtgcg	4080
tccgtctgttc	ggctgcggcg	gctcactcaa	aggcggtaat	agggcgtaat	acggttatcc	4140
acagaaatcag	gggataaacgc	aggaagaac	atgtgagcaa	aaaggccagca	aaaggccagc	4200
aaccgtaaaa	aggccgcggtt	ctggcggttt	ttccataggc	tccgcccccc	tgaaggacat	4260
caaaaaaatc	gaCgctcaag	tcagaggttg	cgaaaaccga	caggactata	aaagataccag	4320
gcgctttcccc	ctggaaagctc	cctcgtggcg	tctcctgttc	cgaacccctgc	gcttaccggga	4380
taactctccg	ccctttctccc	ttcgggaaag	gtggcgcttt	ctcaatgctc	accgtgttag	4440
tatctcaagt	cggtgtaggt	cgcttcgctc	aaagctgggt	gtgtgcacga	accccccggt	4500
cagccccgac	gtctgccttc	atccggtaac	tatcgtcttg	agtcacaacc	ggtaagacac	4560
gactttafcg	cactggcagc	agccactggt	aaacagatta	gcagagcgag	gtatgtaggc	4620
ggtgtctacag	agttctctgaa	gtgttggtct	aaactacggct	acactagaag	gacagtattt	4680
ggtatctcgg	ctctcgtgaa	gccagttacc	ttccgaaaaa	gagttggtag	ctcttgatcc	4740
ggcaaacaaa	ccaccgctgg	tacgggtggt	ttttttgttt	gcgaagcagca	gattacgcgc	4800
agaaaaaaag	gatctcaaga	agatctcttt	atcttttcta	cggggtctga	cgctcagttg	4860
aaacgaaaaa	cacgtttaagg	gatttttggt	atgagattat	caaaaaaggat	cttcacctaag	4920
atcctttttaa	attaaaaatg	aaattttttaa	tcaatctaaa	gtatatatga	gtaaaacttg	4980
tctgacagtt	accaatcgtt	aaactagtga	gcaacctatc	cagcgatctg	tctatttcgt	5040
tcatccatag	gtgcctgact	ccccgtcgtg	tagataacta	cgatacggga	gggctttacca	5100
tctggcccca	tgtcgtgcaat	gataccgcga	gacccacgct	caccggctcc	agattttatca	5160
gcaataaaac	agccagccgg	aaagggccgag	cgcagaagtg	gtcctgcaac	tttatccgcc	5220
tccatccagt	ctattaatgt	tttgcgggaa	gctagagtaa	gtagtctcgc	agtttaagt	5280
ttgcgcgaag	tgtttggcct	tgctaaggg	atcgtgggtg	cacgctcgtc	gttttggtatg	5340
gcttcatctca	gttctgggtt	ccaacgattca	aggcgagtta	catgatcccc	catgtttgtg	400
aaaaaacggg	tttagctctct	cggtctctccg	atcgttgtca	gaagttaagtt	ggcccgagtg	4460
ttatcatctca	tggtttatgg	agcactgcta	aattctcttta	ctgtcatgcc	atccgtaaaa	4520
tgcttttctg	tgactgggtga	gfactcaacc	aaagtattct	gagaaatagtg	tatgctgggga	4580
ccgagttgct	cttgcggcggt	gtcaaacacg	gataataacc	cgccacataag	cagaaacttga	4640
aaagtgcctca	tcatttggaaa	acgttctctc	ggggcaaaaa	tctcaaggga	cttaccgctg	4700
ttcagatcca	gttcgagtga	accactcgt	gcacccaact	gatcttcagc	atcttttact	4760
ttcacccagg	ttttctgggtg	agcaaaaaaa	ggaaggcaaa	atccgcacaa	aaaagggaaa	4820
atgggcagac	ggaaaatgttg	aaatactcata	ctcttctctt	ttcaatatata	ttgaagcatt	4880
tatcaggggt	attgtctcat	gagcggatga	atatttgaat	gtatttagaa	aaaataacaa	4940
ataggggttc	cgcgacacat	tccccgaaaa	gtgccacctg	acgtcgcagc	atcgggagat	5000
ctcccgatcc	cctatggtcg	actctcagta	caatctgctc	gtatgcgcga	tagttaagcc	5060
agtatctgct	ccctgcttgt	gtgttgagg	tcgctgagta	gtgcgcgagc	aaaatttaag	5120

Sequence 5: "Heavy-chain of trastuzumab"

Length	Molecule Type	Organism	Contains DNA and RNA fragments	Skipped Sequence
450	AA	synthetic construct	No	No

Features

Feature Key	Feature Location	Qualifiers
REGION	1..450	note = Heavy-chain of trastuzumab
source	1..450	mol_type = protein organism = synthetic construct

Residues

```

EVOLVESGGG LVDPGGSLRL SCAASGFNIK DTYIHWVROA PGKGLEWVAR IYPTNGYTRY 60
ADSVKGRFTI SADTSKNTAY LOMNSLRAED TAVYYCSRWG GDGFYAMDYV GOGTLVTVSS 120
ASTKGPSVFP LAPSSKSTSG GTAALGCLVK DYFPEFVTVS WNSGALTSGV HTFPAVLQSS 180
GLYSLSVVVT VPSSSLGTQT YICNVNHKPS NTKVDKKVEP KSCDKTHTCP PCPAPPELLGG 240
PSVFLFPPKP KDTLMISRTP EVTCVVVDVS HEDPEVKFNW YVDGVEVHNA KTKPREEOYN 300
STYRVVSVLT VLHODWLNK EYCKVSNKA LPAPIEKTIK KAKGQPREPO VYTLPPSREE 360
MTKNQVSLTC LKGGFYPDSI AVEWESNGQP ENNYKTTTPV LDSDGSFFLY SKLTVDKSRW 420
QQGNVFSCSV MHEALHNHYT QKSLSLSPGK 450

```

Sequence 6: "Light-chain of trastuzumab"

Length	Molecule Type	Organism	Contains DNA and RNA fragments	Skipped Sequence
214	AA	synthetic construct	No	No

Features

Feature Key	Feature Location	Qualifiers
REGION	1..214	note = Light-chain of trastuzumab
source	1..214	mol_type = protein organism = synthetic construct

Residues

```

DIOMTQSPSS LSASVGRVDT ITCRASQDVN TAVAWYQOKP GKAPKLLIYS ASFLYSGVPS 60
RESGSRSGTD FTLTISSLOP EDFATYYCOO HYTTPPTFGQ GTKVEIKRTV AAPSVEFIEPP 120
SDEQLKSGTA SVVCLLNNFY PREAKVQWKV DNALQSGNSQ ESVTEQDSKD STYLSSTLT 180
LSKADYEKKH VYACEVTHQG LSSPVTKSFN RGECL 214

```

Sequence 7: "Expression plasmid pCDNA3.1-Trastuzumab-UAG142"

Length	Molecule Type	Organism	Contains DNA and RNA fragments	Skipped Sequence
8071	DNA	synthetic construct	No	No

Features

Feature Key	Feature Location	Qualifiers
misc_feature	1..8071	note = Expression plasmid pCDNA3.1-Trastuzumab-UAG142

Feature Key	Feature Location	Qualifiers
source	1..8071	mol_type = other DNA organism = synthetic construct

Residues

gacgggatcgg	gagatctccc	gatccccctat	ggtgcactct	cagtacaatc	tgctctgatg	60
cgcgcatagtt	aaagccagtat	ctgtctccctg	cttgctgtgtt	ggagggtcgtc	gagtagtgccg	120
cgaagcaaaat	ttaaagctaca	acaaggccaag	gcttgaccgga	caatttgcatg	aaagaaatctgc	180
ttaggggttag	gcggttttgcg	ctgcttcgcgg	atgttacgggc	cagatatacag	cgttgacattt	240
gattatttgac	tagttatataa	tagtaaatcaa	ttacgggggtc	attagttcat	agcccatata	300
tggaattccg	gtttacataaa	ctttacggtta	atggcccggcc	tggtcgaccg	cccaacgacc	360
cccgcgccatt	gacgtcaata	atgacgtatg	ttcccataagt	aacgcgaata	gggacttttcc	420
atttgacgtca	atgggtgggag	tatttacgggt	aaactgcccc	cttggcagta	catcaagtgt	480
atcctatggcc	aaatgacggcc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgccccagta	catgaccttta	tggggactttc	ctacttggcca	gtacatctac	gtattagtca	600
tgcgtatfac	catggtgtatg	cgggttttggc	agtaacatcaa	tggtcggtgga	tagcgggtttg	660
actcacggggg	attttccaaagt	ctcccacccca	ttgacgtcaa	tggtgagtttg	ttttggcaccc	720
aaaatccaacg	gggactttcca	aaatgtcgtta	acaaactccgc	cccattgacg	caaatgggcg	780
gtaggcgtgt	acggtggggag	gtctatataaa	gcagagctct	ctggctaact	agagaaccca	840
ctgcttactg	gctttatcgaa	atataatacga	ctcactatag	ggagaccctaa	gcttgcttagc	900
gtttaaaactt	aaagcttatgt	accggatgca	gctgctgagc	tgtatcgccc	tgtctctggc	960
cctcgtgacc	aacagcgaagt	tgcagctggt	ggaaaagcggc	ggaggacttg	tgcagccttg	1020
cggatctctgt	agactgagct	cggcttccaa	ggacttcaa	atcaaggaca	ctcatatcca	1080
ctgggtgcgcg	caggccctctg	gcagcgggac	ggaaatgggtg	gccagaatct	accccaccaa	1140
cgcctacaacc	agatacgcgg	acagcgtgaa	gggcccgtttc	accatcagcg	ccgacaccag	1200
caagaaacacc	gcctactctgc	agatgaacag	cctgcccggcc	gaggacacgg	ccgtgtactta	1260
ctgttagtaga	tgggggagggc	acggctttcta	ggcctatgga	tattgggggc	aggcgaacctt	1320
cgtgacagtg	tcttagtgctg	agaccaaagg	ggcctcgggtc	ttcccctctg	catcctctctc	1380
caagagacacc	tctggggggca	cagcggccctt	gggctgcctt	gtcgaagact	acttcccctga	1440
accggtgacg	gtgtcgttgg	actcagggcgc	cctgaaccagc	ggcgtgcaca	ccttcccgggc	1500
tgtcctacag	ctctcaggac	tctcatccctt	cagcagcgtg	gtgaccgtgc	cctccagcag	1560
cftgggcacc	cagacctaca	cttgcacact	gaatcacaa	cccagcaaca	ccaaggctgga	1620
caagaaaagtt	gagcccaaat	cttgtgacaa	aaactcacaca	tgcccaccgt	gcccagcacc	1680
tgaactcctgt	gggggaccgt	cagctcttctc	cttcccccca	aaaaccgaag	acacctcat	1740
gatctcccg	acccctgagg	tcacatgcgt	ggtggtggac	gtgagccaag	aaagaccctga	1800
gggtcaagtgc	aaactggtaag	tggaacggcgt	ggaggtgcat	aatgccaaag	caaaagccggc	1860
ggaggagcag	tacaaacagca	gcctacgggt	gggtcagcgtc	ctcaccgttc	tgcaaccagg	1920
ctggctggaat	gccaagggagt	acaaagtcca	ggctctccaa	aaagcccttc	cagcccccat	1980
cgagaaaacc	atctccaaag	ccaaaaggga	gccccgagaa	ccaaggtgt	acacccctgcc	2040
cccatcccg	gatgactgga	ccaaagaaaca	ggtcagcctg	acctgcctg	tcaaaggctt	2100
ctatcccgac	gacatcgccg	tggagtggga	gagcaaatgg	cagccggaga	acaactacaa	2160
gaccacgcct	cccgctctgg	actccgacgg	ctcctctctt	ctctacagca	agctcacctg	2220
ggacaagagc	agggtggcagc	aggggaacgt	cttctcatgc	tccgtgatgc	atgaggctct	2280
gcacaaacga	tacacccaga	cctgctctctc	cctgtctctcc	ggcaagtgat	aaaggccggcc	2340
cctctcccttc	ccccccctct	aaagcttactg	ggccgaagccg	cttggaaata	ggccgggtgtg	2400
ctgtttgtcta	ctgtttatatt	tccaccataat	tgcgcttctt	tggcaatgtg	agggccccgga	2460
aaactggccc	tgtcttctttg	acggagcattc	cttaggggtct	ttcccctctt	ggcgaaggaaa	2520
tgcgaagtctt	gttgaattctt	gtgaaaggaa	cagttctctct	ggaagcttct	tgaagacaaa	2580
caacgctctgt	agcggaacctt	tgcgaagcag	ggaaacccccc	actctggcgac	agggtgctctt	2640
ggcgcccaaaa	ggcaccgtgta	ttaagatacac	cttgcgaaggc	ggcacaaccc	cgagggccaag	2700
ttgtgagtgtg	gataagttgtg	gaaaagagta	aatggctctct	ctcaagcgta	ttcaacaaagg	2760
ggctggaagga	tggcccaagaa	gtaccaccatt	gtatggggtac	tgatctgggg	cctcggtaaca	2820
catgtctttat	ctgtgttttag	tcgaggtttaa	aaaaaacgtct	aggccccccg	aaccacgggg	2880
acgtgtgtttt	ccctttgaaaa	acacagatgat	aatatggcca	caaccatgta	ccgcattgcaa	2940
ctcctgtctct	cgatttgcact	aaagtcttgca	cttgttcaaca	acagtgtat	ccagatgacc	3000
cagagcccca	gcagctctgtc	tgccagcgtg	ggcgacagag	tgaacctcac	ctgtagagcc	3060
agccagggag	tgaaccccg	cctgcctctg	tatcagcaga	agcctggcaa	ggcccccaag	3120
ctgctgatat	acagcggccag	cttctctgtac	agcggcgtgc	ccagcagatt	cagcggcagc	3180
agatccgggca	ccgactctcac	cctgacacctc	agctccctgc	agcccgagg	cttccgcaac	3240
tactactggc	agcagcacta	accacaccccc	cccacatttg	ggcaggggac	caaggtggaa	3300
atcaagcgtta	cgggtctctgc	accactctgtc	ttcatcttcc	ggccatctga	tgaagcagttg	3360
aaatctctgaa	ctgcctctgtg	tgtgtgcctg	ctgaataact	tctaccccag	agaagccaaa	3420
gtgcagtgga	agggtgcgaa	gcctctgcag	agcggaaaaca	gccaggaag	ggtgacagag	3480
caggatctcca	aggattccac	atcacagcctg	agccagcacac	tgaactgtgt	caaagccgac	3540
tacagagaagc	acaaggtgta	cgctctcgaa	gtgacacacc	agggactgtc	ctccctctgtg	3600
acaaaagagct	tcaaacagag	agaattctgta	tgatctagag	ggcccgttta	aaacccgctga	3660
tcagccttcca	ctgtgccttct	taagtggccag	ccatctgtttg	tttgcctctt	ccccgtgctt	3720
tccttgcacc	tggaaagggtg	cactctccact	gtccttttctt	aatataaatga	ggaaatttga	3780
tgcgactgatt	tgaattaggtg	tattctctatt	ctgggggggtg	gggtggggga	ggacagcaag	3840
ggggagaggtt	gggaagacaa	tagcagagcat	gctgggggat	cgggtgggtc	tatggcttctt	3900
gaagcggaaa	gaaccagctgt	gggtctctagg	gggttatccc	acggccctct	tagggcgcca	3960
ttaaagcggg	cgggtgtggtg	gggttaagcgc	agcgtgaccc	ctacacttgc	cagcgcctta	4020
cgcccgccctc	cttttcgcttt	cttccctttcc	tttctcgcca	ggttcgcccg	ctttcccgtg	4080
caagctcttaa	atccgggggct	cccttttaggg	ttccgatttta	gtgtctttacg	gcacctcgac	4140
cccaaaaaaac	ttgacttaggg	tgatggtttca	cgtatgtgggc	catcgccctt	atagacggtt	4200
tttgcgctct	tggacttggga	gtccacgtttc	tttaatatgtt	gactctttgt	ccaaacttga	4260
acaaacattca	acccctatctc	ggtctatttct	tttgattttat	aaagggaattt	gccgatttgc	4320
gcctatttgtt	taaaaaaatga	ctgtgattttt	caaaaaatttt	acggcgaatta	attctgttgg	4380
atgtgtgtgca	gttagaggtgt	ggaaaagctcc	caggctcccc	agcaggcaga	agtatgcaaa	4440
gcattgcattc	gaatttagtgt	gtggaaagtc	gtggaaagtc	cccaggtctc	ccagcaggga	4500
gaagattgca	aagcatgtcat	ctcaattagt	cagcaaccat	agtcctcgcc	ctaaactccg	4560
ccatcccgcc	ccctaacctcg	cccagttccg	ccctattctc	ggcccatggc	tgactaattt	4620
tttttattttt	tgcagagggcc	gagggccgct	ctgcctctga	gctattccag	aaagtagtag	4680
gattgcttttt	tggagggctta	cttcttttga	aaaagctctcc	gggagcttgt	atatccattt	4740
tccgattctga	tcaagagaca	ggatgaggat	cggttcgcgat	gattgaaaca	gatggattgc	4800
acgcaggttc	tccggccgctc	tgggtggaga	ggctattctcg	ctatgactgg	gcacaaacaga	4860
taatccgctgt	ctctgatggc	ggcgtgttcc	ggctgtcagc	gcaggggcgc	ccggttcttt	4920
ttgtcaagaa	cgactgttcc	gttgcctctga	atgaactgca	ggacagggca	gcggcgctat	4980
cgtgggttgg	cacgaggggc	gttctcttgc	cagctgtgct	cgaagtgtgc	actgaagcgg	5040
gaagggactgt	gctgtcatatt	ggcgaagtg	cggggcagga	tctctgtgca	tctcaccttg	5100
ctccttgcgga	gaaagtatcc	atcgaatggc	atggcaatgg	ggcgtgtcat	acgcttgatc	5160
cgggtacctgt	cccatctcgac	acccaagcga	aaatatcgcat	cgaagcagga	cgtaactcga	5220
tggaagccgg	ctttgtcgat	caggatgatc	tggacggaaga	cgatcagggg	ctcggccag	5280
cgaacttgtt	cgccagggctc	aaaggcggca	tgcccagacg	cagggatctc	gtcgtgacc	5340
atggcgatgc	ctgcttgcgg	aatatctatg	tggaaaaatg	cgctttttct	ggattctatg	5400
actgtgtggcg	gctgggtgtgt	cgggacggct	atcaggacat	agcgtttggc	acccgtgata	5460
ttgtctgaaga	gcttggcggc	gaattgggtgt	acccgttctc	cgtgttttac	ggatatcgccg	5520
cttcccgatcc	gcagcgcctac	gcctctctatc	gcctctcttga	cgagtctctc	tgaagcggac	5580
tctgggggttc	gaaattgaccg	accctacggac	gccccaacctg	ccatcacaga	atttcgattc	5640
caccgcgcgcc	tctctatggat	ttctgtggctt	cggaatcggt	ttccgggacg	ccggctggat	5700
gatccctccag	cgcggcgaaat	tcattgctgga	gttcttctgcc	caccctcaact	tgtttatttg	5760
agctttataat	ggtttacaaat	aaaggcaatag	catcacaaat	ttcacaaaata	aaagcattttt	5820
ttcactgcat	tctagtttgtg	ttttgttccaa	actcatcaat	gtatctttatc	attgtctgtat	5880
accgttgcacc	tctagcttaga	cttttggcgta	atcatgggtca	tagctgtttt	ctgtgtgaaa	5940
ttgtttatccg	ctcacaaactt	cacacaaactt	acgagccgga	agcatataagt	gtataagcctg	6000
gggtgtcctaa	tgagttagctt	aaactcacatt	aatgtcggttg	cgctcactgc	ccgctttcca	6060
gtcgggaaac	ctctgtctgcc	agctgcattt	atgaatcggc	caaagcgccg	ggagagggcg	6120
tttgcgttatt	ggcgctctctt	ccgcttccctc	gctcactgac	tccgttgcgt	cggctcgttgc	6180
ctgctcgccga	ggcgctatcag	ctcactcaaa	ggcggtaata	cggttatcca	cagaaatcag	6240
ggataaacga	ggaaagaaca	tgtgagcaaa	aggccagcaa	aaagccagga	accgtaaaaa	6300

ggccgcgttg	ctggcggtttt	tccataggct	cgcgccccct	gacgagcatc	acaaaaatcg	6360
acgctcgaag	cagaggtggc	gaaacccgac	aggactataa	agataccagg	cgtttccccc	6420
tggaagctcc	ctcgtgcgct	ctcctgttcc	gacctggccg	cttaccggat	acctgtccgc	6480
ctttctccct	tcggggaagcg	tggtcggttc	tcatagtctca	cgctgtaggt	atctcagttc	6540
ggtgtaggtc	gttcgctcca	agctgggctg	tgtgcacgaa	ccccccgttc	agcccgaaccg	6600
ctgcgcttta	tccggttaact	atcgtcttga	gtccaacccg	gtaagacacg	acttatcgcc	6660
actggcagca	gccactggta	acaggattag	cagagcgagg	tatgtaggcg	gtgctacaga	6720
gttcttgaag	tggtggccta	actacggcta	cactagaaga	acagtatttg	gtatctgcgc	6780
tctgctgaag	ccagttaacct	toggaaaaag	agttggtagc	tcttgatccg	gcaaacaaaac	6840
caccgcgtgt	agcgggttttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	6900
tcaagaaagt	cttttctatct	tttctacggg	gtctgacgct	cagtggaaacg	aaaactcacg	6960
ttaagggaatt	tttgtcatga	gattatcaaa	aggatcttcc	acctagatcc	ttttaaatat	7020
aaaaatgaagt	tttaaatcaa	tctaaagtat	atatgagtaa	acttggtctg	acagttaacca	7080
atgcttaaat	agtgaaggcac	ctatctcagc	gatctgtcta	tttcgttcat	ccatagtgtgc	7140
ctgactcccc	gtcgtgtaga	taactacgat	acgggagggc	ttaccatctg	gccccagtgc	7200
tgcgaatgata	ccgcgagacc	cacgctcacc	ggctccagat	ttaccagcaa	taaacacagcc	7260
agccgggaagg	gccgagcgca	gaagtgggtcc	tgcaacttta	tccgcctcca	tccagtctat	7320
taattgttgc	cggggaagcta	gagtaagtag	ttcgccagtt	aatagtttgc	gcaacgttgt	7380
tgccattgct	acagggatcg	tggtgtcacg	ctcgtcggtt	ggtagggctt	catccagctc	7440
cggttcccaa	cgatcaaggc	gagttacatg	atcccctatg	ttgtgcaaaa	aaggcgttag	7500
ctccttcgggt	cttccgaltcg	ttgtcagaag	taagtctggc	gcagtggttat	cactcatggt	7560
tatggcagca	ctgcataaatt	ctcttactgt	catgccatcc	gtaagatgct	tttctgtgac	7620
tggtgagtac	tcaaccgaagt	cattctgaga	atagtgtatg	cggcgaccga	gttgctcttg	7680
cccggcgcta	ataccgggata	ataccgcgcc	acatagcaga	actttaaaaag	tgctcatcat	7740
tggaataacgt	tcttcggggc	gaaaaactctc	aaggatctta	ccgctgttga	gatccagttc	7800
gatgtaaccc	actcgtgcac	ccaactgatc	ttcagcatct	tttactttta	ccagcgtttc	7860
tgggtgagca	aaaacaggaa	ggcaaaaatgc	cgcaaaaaag	ggaataaggg	cgaacaggaa	7920
atggttgata	ctcatactct	tcctttttca	atattattga	agcatttatc	agggttattg	7980
tctcatgagc	ggatacatat	ttgaatgtat	ttagaaaaat	aaacaaatag	gggttccgcg	8040
cacatttccc	cgaaaagtgc	cacctgacgt	c			8071