

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

03 May 2018 (03.05.2018)



(10) International Publication Number

WO 2018/081590 A1

(51) International Patent Classification:

A61K 31/436 (2006.01) A61P 17/06 (2006.01)

A61P 17/00 (2006.01) A61P 25/00 (2006.01)

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(21) International Application Number:

PCT/US2017/058800

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(22) International Filing Date:

27 October 2017 (27.10.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/414,410 28 October 2016 (28.10.2016) US

(71) Applicant: WARP DRIVE BIO, INC. [US/US]; 400 Technology Square, Cambridge, MA 02139 (US).

(72) Inventors: GRAY, Daniel, C.; 246 Harvard Street Apt. 3, Medford, MA 02155 (US). LI, Enhui; 116 Pearl Street, Newton, MA 02458 (US). BOWMAN, Brian, R.; 199 Seton Drive, New Rochelle, NY 10804 (US). VERDINE, Gregory, L.; 20 Rowes Wharf Unit 410, Boston, MA 02110 (US). ROBISON, Keith; 73 Tewksbury Street, Andover, MA 01810 (US). CHEVRETTE, Marc; 2001 Monroe Street, Madison, WI 53711 (US). UDWARY, Dan; 3634 Ramita Terrace, Davis, CA 95618 (US). WANG, Pam Shouping; 29 Concord Avenue Apt. 809, Cambridge, MA 02138 (US). LI, Anna; 70 Bishop Allen Drive Apt. 6, Cambridge, MA 02139 (US). MORGENSTERN, Jay, P.; 452 Beacon Street #3, Boston, MA 02115 (US).

(74) Agent: BELLIVEAU, Michael, J.; Clark & Elbing LLP, 101 Federal Street, 15th Floor, Boston, MA 02110 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

(54) Title: COMPOSITIONS AND METHODS FOR THE PRODUCTION OF COMPOUNDS

(57) Abstract: The present disclosure provides proteins, nucleic acids, vectors, and host molecules useful for the production of compounds of interest, and methods for their use.

WO 2018/081590 A1

COMPOSITIONS AND METHODS FOR THE PRODUCTION OF COMPOUNDS

Background

Polyketide natural products are produced biosynthetically by polyketide synthases (PKSs), e.g., type I polyketide synthases, in conjunction with other tailoring enzymes. Polyketide synthases (PKSs) are a family of large, multi-domain proteins whose catalytic functions are organized into modules to produce polyketides. The basic functional unit of polyketide synthase clusters is the module, which encodes a 2-carbon extender unit, e.g., derived from malonyl-CoA. The modules generally present in a polyketide synthase include i) a loading module; ii) extending modules; and iii) releasing modules. Within the module, the minimal domain architecture required for polyketide chain extension and elongation includes the ketosynthase (KS), acyl-transferase (AT) and the ACP (acyl-carrier protein) domains, and the specific chemistry of each module is encoded by the AT domain and by the presence of the β -ketone processing domains: ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) domains. Polyketide synthase biosynthesis proceeds by two key mechanisms: polyketide chain elongation with a polyketide synthase extending module and translocation of the polyketide intermediate between modules. Productive chain elongation depends on the concerted function of the numerous catalytic domains both within and between modules.

Combinatorial biosynthesis is a general strategy that has been employed to engineer polyketide synthase (PKS) gene clusters to produce novel drug candidates (Weissman and Leadlay, Nature Reviews Microbiology, 2005). To date, these strategies have relied on engineering PKS domain deletions and/or domain swaps within a module or by swapping an entire module from another cluster to produce a chimeric cluster. The problem with this approach is that protein engineering of the polyketide megasynthases via wholesale domain and/or module replacement, insertion, or deletion can perturb the "assembly line" architecture of the PKS, thus drastically reducing the amount of polyketide synthesized.

Summary of the Invention

The present disclosure provides compositions and methods for use in combinatorial biosynthesis of polyketides without a significant loss of compound production by module swapping between polyketide synthase genes. Bioinformatics approaches may be used to predict module interface compatibility and therefore, the likelihood that a heterologous module may be swapped into a PKS gene. The resulting compatibility information may be used to engineer a polyketide synthase with an increased likelihood of functioning in assembly-line polyketide biosynthesis.

Accordingly, in one aspect, the disclosure provides an engineered polyketide synthase that includes one or more heterologous modules with altered enzymatic activity relative to a reference polyketide, wherein the engineered polyketide synthase is capable of producing a polyketide when expressed under conditions suitable to allow expression of a compound by the engineered polyketide synthase and wherein the one or more heterologous modules do not substantially inhibit polyketide translocation during polyketide biosynthesis.

In another aspect, the disclosure provides an engineered polyketide synthase including one or more heterologous modules with altered enzymatic activity relative to a reference polyketide, wherein the engineered polyketide synthase is capable of producing a polyketide when expressed under conditions

suitable to allow expression of a compound by the engineered polyketide synthase and wherein the one or more heterologous modules include linking sequences which are compatible to the linking sequences of the modules adjacent thereto.

In another aspect, the disclosure provides an engineered polyketide synthase including one or more heterologous modules with altered enzymatic activity relative to a reference polyketide, wherein the engineered polyketide synthase is capable of producing a polyketide when expressed under conditions suitable to allow expression of a compound by the engineered polyketide synthase and wherein the polyketide expression level of the engineered polyketide synthase is at least 1% (e.g., at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 100%, at least 110%, at least 120%, at least 130%, at least 140%, at least 150%) of the polyketide expression level of the reference polyketide synthase.

In some embodiments, the polyketide expression level of the engineered polyketide synthase is at least 1-10% (e.g. at least 1-10%, at least 11-20%, at least 21-30%, at least 31-40%, at least 41-50%, at least 51-60%, at least 61-70%, at least 71-80%, at least 81-90%, at least 91-100%, at least 101-110%, at least 111-120%, at least 121-130%, at least 131-140%, at least 141-150%). In some embodiments, the engineered polyketide synthase includes one or more heterologous modules with native linking sequences.

In some embodiments, the engineered polyketide synthase may include one, two, three, or more heterologous modules. In some embodiments in which the engineered polyketide synthase contains multiple heterologous modules, the heterologous modules may be adjacent in the engineered polyketide synthase. In some embodiments in which the polyketide synthase contains multiple heterologous modules, any of the modules may be separated by one or more native modules in the engineered polyketide synthase.

In some embodiments of any of the above described aspects, at least one of the one or more heterologous modules is an elongation module which modifies a β -carbonyl unit in the variable region of the polyketide.

In some embodiments of any of the above described aspects, at least one of the one or more heterologous modules includes a portion having at least 90% identity to any one of SEQ ID NO: 1-174.

In some embodiments of any of the above described aspects, at least one of the one or more heterologous modules includes a portion having the sequence of any one of SEQ ID NO: 1-174.

SEQ ID NO: 1

QPLAIVGMACRLPGGVSSPEDLWRLVESGTDALSGFPVDRGWDVDGLYDPDPDVPGKSYTVEGGFLDAVTGFDAPFF
 GISPREALAMDPQQLVLEASWEAFERAGIEPGSVRGS DTGVFMGAFFPGGYGTGADLGGFGMTGGAASVLSGRVSYF
 FGLEGPAMTVDTVCSSSLVALHQAGYALRHGEC SLALVGGVTVMSTPQTFVEFSRQRGLAADGRCKAFSDDADGTGW
 SEGVGVLLVERLSDAQARGHNILAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLTGADV DVVEAHGTGTT
 LGDPIEAQAVIATYGQDRDQPVLLGSLKSNLGH TQAAAGVSGVIKVMALQNGVVPRTLHV EEPSRHVDWTAGAVEL
 VTENQPWPELGRARRAAVSSFGLSGTNAHVILE SAPDQPPAPSTDSPVSAVTAGVVPLPISAKTLPALADLEDR LRT
 YLTTTPD TDLPAVASTLATRSLFEHRAVLLGEDT VTGT AIPDPRVVVFVFPQGQWQWQGMGSALLTSS TVFAERMAE
 CAAALSEFVDWDL LTVLDDPSVVD RVDVVPACWAVMISLA AVWQAAGIHPDIVLGHSQGEIAAACLAGAISLPDAA

RIVAQRSQLIAHQGLGHGAMASISLPADDIPTTDQVWIAAHNGTSTVIAGDPQAVEAVLATCETR GARVRKINVDYAS
HTPHVEQIRTELDDITTTGIEAHTPAVPWLSTTDNTWIDQPLDPTYWYRNLREPVRFGPAIDLLQTQDNNLFIEISAS
PVLLQTMDNAATVATLRRDEDTTHRLLTAF AEAHVHGATINWPTVLDTTTTTPVDLPTYPFQRQRYWATSNGHPADLT
PEALLKVVRDSAAMVLGHASADTVPTATAFQELGLDSLTAVELRNSLT KATGLRLPATMAFDYPTPDALAARL

5

SEQ ID NO: 2

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDAITDFPTDRGWDTDTLFDPDPDTPGKTTYTVHGGFLDDVAGFDAPFF
GISPREAVAMDPQQRLVLESSWEAFERAGIQPDSIRGSDTGVMGAYPDGYGIGADLAGFGVTAGAGSVLSGRVSYF
FGLEGPAMTVDTACSSSLVALHQAAAYALRQGECSLALVGGVTVMPSRPTFIEFSRQRGLAADGRSKAFADAADGTGF
10 SEGVGVLLVERLSDAQAKGHNILALVRSSAVNQDGASNGLTAPNGPSQQRVIQSALAGAGLTSADVDVVEAHGTGTT
LGDPIEAQAVLATYQGDRDPVLLGSLKSNLGHGTQAAAGVSGVIKVMALQHNTVPATLHVDAPSRHVDWTAGAVRL
ATENQPPWETNRPRRAGVSSFGVSGTNAHVILEQAPAASPVEPVDTTDVVVPLVVSARSSGSLSDQADRLAALVGSF
DAPALTSLADALLTRRTVFSQRAVVVAGSHEQAAAGLRALAAGDSHPALVTGAAGPARVVLVFPFGQGSQWAGMGAEL
LDASPVFAARIAECAEALRPWVDWSLDEVLRGDASADVLGRVDVVPQASFAVMVGLAAVWESAGVRPDAVLGHSQGE
15 IAAAYVAGALSLTDAKIVAVRSRLIAARLGRGGMASVALAPEEAAKLGRTELA AVNSPASVVIAGDAEALDETLAM
LEGEGVRVRRVAVDYASHTPHVEELEQSM AEALADVRSRQPRVRFLLSTVTGDWVTEAGALDGGYWRNL RQPVRFGP
AVASLAEAGYTVFVEASAHFVLVQPVAETLDRD TAVVTGTLRRQDGGLPRLLTSMAELFVGGVPVNWVLLPAGAVR
GWVDLPTYAFDHQRYWLENRELTP EALLKLVCGRAAAVLGHVDADAVPVA AAFRDLGVDSLTAVELRNSLAKATGLR
LPATLVFDYPTPTVLAGRL

20

SEQ ID NO: 3

EPLAIVGMACRLPGGVLSPEDLWRLVESGGDAISGFVPDRGWDVENLFD PDPDAAGRTYAVRGGFLDGAAGFDASFF
GISPREAQAMDPQQRLVLEVSWEAFERAGIEPGSVRGSDTGVMGAYPPGGYGMGTDLGGFGMTSVAVSVLAGRVSYF
FGLEGPAMTVDTACSSSLVALHQAGSALRQGECSLALVGGVTVMPTPQT FVEFSRQRGLAADGRCKAFADAADGTGF
25 SEGVGVLLVERLSDAQARGHNILAVVRGSAVNQDGASNGLTAPNGPSQQRVIQSALAGAGLTSADVDVVEAHGTGTT
LGDPIEAQAVIATYQGDRDQPVLLGSLKSNLGHGTQAAAGVSGVIKVMALQNGVVPRTLHIDEPSRHIDWTAGAVEL
VTENQSWPETGRDRRAAVSSFGISGTNAHVILESAPAQPVPVPTDTPVSDVTAGVVPLPISARTVPALADLEDRLRAY
LTTTTPETDLPAVASTLAMTRSVFEHRAVLLGEETVTGIAVSDPRVVFVFSGQGSQRVMGEELAAAFPLFARLHRQV
WDL LDVPDLEVDDTGYVQPALFALQVALFGLLESWGVRPQAVLGH SVGEVAAGYVAGVWSLEDACTLV SARARLMQA
30 LPAGGAMVAVPVSEEQARAVLVDGVEIAAVNGPASVVLSGDES AVLRAEGLGRWTRLSASHAFHSVRMEPMLEEFR
QVASELTYREPRIVMAAGEQVTTPEYWVRQVRD TVRFQDQVAAF GDAVFLEIGPDRTL SRLIDGIAMLDGDDEVRAA
VAALAVMHVQGVGDWPAILGTTTGRVLDLPTYAFQH ERYWMVIQELSPEALLKIVRDSAAMVLGHANADTVPTATA
FQELGLDSLTAVELRNSLT KATGLRLPATMAFDYPTPAALAGRL

35 SEQ ID NO: 4

EPLAIVGMACRLPGGVSTPEDLWRLVESGTDAITDFPTDRGWDTD DLFDPDPDTPGKTTYTVHGGFLDDVAGFDASFF
GISPREALAMDSQQRLVLEASWEAFERAGIEPGSVRGSDTGVMGAYPDGYGIGVDLGGFGATAGAGSVLSGRLSYF
FGLEGPAMTVDTACSSSLVALHQAGSALRQGECSLALVGGVTVIANPQIFVEFSRQRGLAADGRCKAFADNADGTGF
SEGVGVLLVERLSDAQARGHNILALVRSSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLTGA EVDVVEAHGTGTT
40 LGDPIEAQAVLATYQGDRDQPVLLGSLKSNLGHGTQAAAGVSGVIKVMALRHDTVPATLHIDEPSRHIDWTAGAVEL
VTENQPPVPLDRPRRAAVSAFGVSGTNAHVILESAPDQPPASATDTPAPAVTAGVVPLPISAKTVPALADLEDRLRA
YLT TTPETDLPAVASTLATTRS LFEHRAVLLGEDTVTGT TTPDPRIVFVFPGQGWQWQMGMSALLTSSTVFAERMAE

CAAALSEFVDWDLTLVLDLDPISIVDRVDVVQPACWAVMISLAAVWQAAGIHPDIVLGHSQGEIAAACLAGAISLPDAA
RIVAQRSQQLIAHQGLGHGAMASISLPADDIPTTDKVIWIAAHNGTSTVIAGDPQAVEAILATCETRGRARVRKINVDYAS
HTPHVEQIRTELLEDITTGIEAHTPTVPWLSTTDNTWIDQPLDPTYWYRNLREPVRFGPAIDLLQTQDNNLFIEISAS
PVLLQTMNDNATTVATLRRDEDTTQRLLTAFEAHVHGATINWPTVLNTTTTTPVDLPTYPFQQRQRYWATSNDRLNGRT
5 SVEQHRIMVELVLAHATSVLGHESPDAIAPDRAFKDLGMDSLTAIELRNHLVAETGVRLPATTAFDHPTADDLAKRL

SEQ ID NO: 5

EPIAIVSMSCRAPGGVDSPELWRLVESGTDAITDFPGDRGWDVAGLYSPDPTGYKTYCVQGGFLDAAADFDAFFG
ISPREALGMDPQQRLLETSWEAIERARIDPRSLPGRNVGVYVGGAAQGYGVGAIDQQRDNVITGSSISLLSGRLSY
10 ALGLEPGVTVDTACSSSLVALHLACQALRQRECSMALVSGVSVIPTPDVFVEFSRQRGLAADGRCKSFSASADGTI
WAEVGVLVLERLSEATRLGHRVLAVVRGSAVNSDGNLTAPNGVSQQRVIRQALTGAGLTAADVVDVVEAHGTGT
KLGDPIDAEAILATYQDRSTPVCLGSLKSNIGHAMAASGVLAVIKMVEAMRHGLIPRTLHVEEPSPHVDWASGDVA
LLTENQWPWDDAKLRRAGVSSFGLSGTNAHVVEQYRAPAAPDITTTTEHQPLAWTLSARDPKALREQAGRLHAALTE
SPWRPLDIGYSLATTRSNFAHRAVAVGSDRELLRALS KLADGSAWPALVTATAKDRRVAYLFDGQGSQRPDMGSGL
15 YERFPAFARAWDRISAEFGKHLHDHSLTDVYLGRGDAATADLVDDTLYAQAGLFTMEIALFELLAEWGVRPDFVSGHS
IGETAAYAAAGVLSLEDVTKLIVARGRALRQVPPGAMVALRAGEDAREFLGRTGAALDLAAVNSPTS VVVGASEA
VAGFRARWTESGREARTLNVRHAFHSRHVEAVLGEFREVLES LTFRTPALPVVSTVTGR LIEPTELSTSEYWL RQVR
QTVRFHDVAVRELSGQGVGTVEIGPSGALASAGLECLGDEASFHAVQRPGSPGDVCLMTAVALHAGGTTVDWATVL
AGGRATDLPVYPFHGHSYWLAPARPSAPEEPRTMLELVRLEAAIALSITDPGLIADDSFLDLGFD SIALRLSNRL
20 AAVTGLDLPPSLLFDHPTPAELARLD

SEQ ID NO: 6

EPLAIVGMACRLPGGVSSPDDLWRLVASGTD AISEFPADRGWDVDNLYDPDPDAPGKTYTVLGGFLDG VAGFDASFF
GISPREALAMPQQRMLLEVSWAEFEHAGIPPSVRGSDTGVMGAFPPSGYNAGLEEFGMTGD AVSVLSGRVSYFFG
25 LEGPAITVDTACSSSLVALHQASSALRQGECSLALVGGVTVLATPQTFVEFSRQRGLALDGRSKAFADAADGAGWAE
GVGVLVVERLS DARAKGHIWGVIRGSAVNQDGASNGLSAPNGPSQQRVIRQALANAGLAPHEVDVVEAHGTGTMLG
DPIEAQAVIATYQDREQPLLLGSLKSNVGHTQAAAGVSGVIKMVMALQHDTV PATLHVDAPS RHVDWTAGAVELVT
ENRPWPETGRVRRAGVSSFGISGTNAHVILESAPPEAVAPVVASDRVPLVISAKTPAALAEMENRLRAYLA
AAPGADPRAVASTLATARSVFEHRAVLLGENTITGT VAGADPRVVVFVFPQGQWQQLGMGRALRESSPVFAARMAECA
30 AALSEFVDWDLFTMLDDPAVIDRIDVLQPACWAVMMSLAAVWQAAGVRPDAVIGHSQGEIAAACVAGALS LRDAARI
VALRSQLLAREMGHGVMAAVALPADDIPLVDGVWIGARNGPSSTVISGTPEAVEVVVAACEERGARVRRI TA AVASH
SPLGEKIRTELLGISASIPSRTPVVPWLSTADGIWIEAPLDPAYWWRNLREPVGFGPAVDLLQARGENVFLEMSASP
VLLPAMNDAVTVATLRRDDDTDRMLTALAEAHAGVIVDWPRVFGSTTRVLDLPTYAFEHQRYWAVNGRPADLTPE
ALLKLVCGRAAAVLGHVDADAVPVAVAFRDLGVDSL TAVELRNSLAKATGLRLPATLVFDYPTPTVLAGRL
35

SEQ ID NO: 7

EPLAIVGMACRLPGGVLSPEDLWRLVESGGDAISGFVPVDRGWDVENLFD PDPDAAGRTYAVRGGFLDGAAGFDASFF
GISPREAQAMPQQRVLLEVSWEAXERAGIEPGSVRGSDTGVMGAYPGGXGXTDLGGFGMTSVAVSVLAGRVSYF
FGLEGPAMTVDTACSSSLVALHQAGSALRQGECSLALVGGVTVMP TPQTFVEFSRQRGLAADGRCKAFADAADGTGF
40 SEGVGVLLVERLSDAQARGHNILAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLAGAEVDVVEAHGTGTT
LGDPIEAQAVIATYQDRDQPVLLGSLKSNLGHTQAAAGVSGVIKMVMALRHDTV PATLHIDEPSRHIDWTAGAVEL
VTENQSWPETGRDRRAAVSSFGISGTNAHVILESAPAQVPPMDTPVSAVTAGVVPLPISARTVPALADLEDRLRAY

LTATPETDLPAVASTLAVTRSVFEHRAVLLGEETVTGTIAVSDPRVVFVFSGQGSQRVGMGEELAAAFPLFARLHRQV
 WDLLEDVDPDLEVDDTGYVQPALFALQVALFGLLESWGVRPQAVIGHSVGEVAAGYVAGVWSLEDACTLV SARARLMQA
 LPAGGAMVAVPVSEERARAVLVDGVEIAAVNGPASVVLSGDES AVLRVAEGLGRWTRLSASHAFHSVRMEPMLLEEFR
 QVASELTYREPRIVMAAGEQVTTPEYWVRQVRDTRVFGDQVAAGDAVFLEIGPDRTLSRLIDGIPTLHGDDQHAV
 5 VAALAEHLVQGVPIDWSSILGVNPARVDLPTYAFQHERYWMVIQELSPEALLKIVRDSAAMMLGHPNTDAIAATTAF
 RDLGVDSLIAVELRNSLAKATGLRLPATLVFDYPTPTVLGRRL

SEQ ID NO: 8

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDAITDFPTDRGWDTDLDLDPDPTPGKTYTVHGGFLDDVAGFDASFF
 10 GISPREAQAMDPQQLVLEAAWEAFERAGIEPGSVRGSDTGVMGAYPGGYGIGVDLGGFGATAGAGSVLSGRLSYF
 FGLEGPA MTDVTACSSSLVALHQAGSALRQGECSLALVGGVTVIANPQIFVEFSRQRLAADGRCKAFADSADGTGW
 SEGVGVLLVERLSDAQARGHNILAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLAGAEVDVVEAHGTGTT
 LGDPIEAQAVIATYQDRDQPVLLGSLKSNLGHTQAAAGVSGVIKVMALQNGVVPRTLHADQPSRHIDWTAGAVEL
 VTENQWPPELGRPRRAAVSAFGVSGTNAHVILESAPAQPVPVPTPVS AVTAGVVPLPISARTVPALADLEDRLRAY
 15 LTATPETDLPAVASTLATTRSVFEHRAVLLGEDTGTGTAIPDRIVFVFSGQGSQRVGMGEELAAAFPLFARLHRQV
 WDLLEDVDPDLDDVDDTGYVQPALFALQVALFGLLESWGVRPQAVIGHSVGEVAAGYVAGVWSLEDACTLV SARARLMQA
 LPAGGAMVAVPVSEEQARAVLVDGVEIAAVNGPASVVLSGDEAAVLRVAEGLGRWTRLSASHAFHSVRMEPMLLEEFR
 QVVSRLTYREPRIVMAAGEQVTTPEYWVRQVRETTRVFGDQVAAGDAVFLEIGPDRTLSRLIDGIAMLDGDDDEVRAA
 VAALAVMHVQGVGVDPAILGTTTGRVLDLPTYAFQHERYWMANNRGPADLTPEALLKVVRDSAAMVLGHANADTV
 20 AATAFQELGLDSLIAVELRNSLAKATGLRLPATMVFYPTPAALAGRL

SEQ ID NO: 9

EPLAIVGMACRLPGGVSSPEDLWRLVESGFDAITGFPTDRGWDVDNLYDPDPDAPGKSTTLHGGFLDDVAGFDASFF
 GISPREAVAMDPQQLAMEVSWEAERAGIEPGSVRGSDTGVMGAYPGGYGIGAEELGGFMLTGRAGSVLAGRVSYF
 25 FGLEGPA MTDVTACSSSLVALHQAAAYALRQGECSLALVGGVTVMPTPVMFVEFSQQQNLADDGRCKAFADSADGTGW
 SEGVGVLLVERLSDAQARGHNILAVVRGSAVNQDGASNGLTAPNGPSQQRVIRSALTSAGLTTADVDVVEAHGTGTT
 LGDPIEAQAVLATYQDRDQPVLLGSLKSNLGHTQAAAGVSGVIKVMALQNGVVPRTLHVVEEPSRHVDWTAGAVEL
 VTENQSWPETGRARRAAVSSFGVSGTNAHVILESAPAQPVPMPDTPAPT VTTGVVPLPISAKSLPALADLEDQLRAY
 LTATPETDLPAVASTLAMTRSVFEHRAVLLGEETVTGTGTAIPDRIVFVFSGQGSQRVGMGEELAAAFPLFARLHRQV
 30 WDLLEDVDPDLDDVDDTGYVQPALFALQVALFGLLESWGVRPQAVIGHSVGEVAAGYVAGVWSLEDACTLV SARARLMQA
 LPAGGAMVAVPVSEEQARAALVDGVEIAAVNGPASVVLSGDEAAVLRVAEGLGRWTRLSASHAFHSVRMEPMLLEEF
 QVASELTYQEPRIVMAAGEQVTTPEYWVRQVRDTRVFGDQVAAGDAVFLEIGPDRTLSRLIDGIAMLDGDDDEVRAA
 VAALAEHLVQGVPIDWPAAILGTTTGRVLDLPTYAFQHORYWAASWLAGLAPEEREALMKVVRDTAAVVLGHADAG
 TIPVTA AFKDLGLDSLTAVELRNSLAKSTGLRLPATMVFYPTPASLAARLD
 35

SEQ ID NO: 10

EPLAIVGMACRLPGGVSPEDLWRLVESGTD AISGF PADRGWADLSLRGGFLGDAAHFDAAFFGISPREALAMDPQQ
 RLILEASWEAFERAGIEPGSVRGSDTGVMGAFSGGYGAGADLAGFGVTAGAVSVLSGRVSYLFGLEGPA VTDVTAC
 SSSSLVALHQAGHALRQGECSLALVGGVTVMPTPDIFVEFSRQGG LASDGRCKAFADAADGT SWSEGAGVLVVERLS
 40 AERRGHTVLALVRGSAVNQDGASNGLTAPNGPSQQRV IQAALANAGLTPHEVDVVEAHGTGTRLGDPIEAQAVIATY
 GRDREHPLLLGSLKSNVGHTQAASGVSGLIKVMALRRGTVPRTLHVDEPSRHVDWTAGAVQLAIENQWPPELGRPR
 RAAVSSFGVSGTNAHVILEGVPEEPADSEEPAGLTPLLISAKTPAALAEFEDRLRARLTTEPNLSAVASTLVTRSL

FDHRAVLLDGETVSGMAEPDPRVVFVFSGQGSQRAGMGDDLAAAFPVFAKIRQQVWDLDDIPDLPVDETGHAPALF
ALQVALFGLLDSWGVRPDALVGHSIGELAAGYVAGIWSLEDACALVSARARLMETLPPGGVMVAVPVSEEQARAVLT
DGVEIAAVNGPASVVLSSGEETAVLQAAAAALGGRSKRLATSHAFHSARMEPMLDEFRAVAEQITYGSPRIPMAVGDGP
DYWVRQVRDTRVFEQVAAHDGAIFVELGPDGSLARLVDGIAVLDRDEPRAALTALARLHVRGVKVDWPIAAGRRE
5 LDLPTYPFQRQRYWAETPTARRAPTDLTLVRDTTATVLGYPDNTAVTPTTAFTDLGIDSLTAIELRNNMATTTGLR
LPATLVFDYPTPATLAARLD

SEQ ID NO: 11

EPLAIIGMACRLPGGVTTPEDLWQLVETGTDAISGFPTDRGWDVESLYDPDPAAGKSYCVEGGFLDAVADFASFF
10 GISPREALAMPQQRLILETSWEAFERAGIDPADARGSDTGVMGAFTSGYGADLEGFGGTAGALSVLSGRVSYFFG
LEGPAATVDTACSSSLVALHQAGYSLRHGECSLALVGGVTVMATPRTFVEFSRQRGLASDGRCKAFGDTADGTWSE
GVGVLLVERLSDAERNGHRVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIIQAALDNAGLAPQDQDVVEAHGTGTTLG
DPIEAQAVIATYQONREQPLLLGSLKSNVGHTQAAAGVSGVIKMIMALRHGVVPRTLHVDEPSRHVDWTAGAVHLVR
ENQWPDPVDRPRRAGVSSFGVSGTNAHIIIESPPSQPAPEPAPALSPLVISAKTPQALAAAYEDRLRITYLTAAPSTDA
15 RALAVTRSLFEHRAVLLGEDTVTGTALTEPRVVFVFPQGQWQLGMGAALMESVVFMAERMAECAAALSEFVDWNLTIT
VLNDPAVIDQVDVVQPACWAVMVSLAAVWQAAGVRPDAVIGHSQGEIAAACVAGAIISLRDAARIVALRSRLISERLG
KGAMASITLPADQITLAEGAWIAAYNGPTSTVVAGTPQAIEQMHERVRRIAVDYASHTPHVEQIRAEELDLTTDVS
SQTPTLPWYSTVDGTWIDSPLDGDYWRNLRQPVGFHFAVQTLQALGETVFVEVSASPVLPPAMDDAVTIATLRRDE
GTLTRMHTALAEAHVLGVTIDWPTVLGVTTTRHVDLPTYAFQRQRYWVAELASLGAERERARLKLVSDDTAAGILGHA
20 DSGTVPVTAAFRELGVDSLTAVELRNGLAKATGLRLPATMVFDDYPTPQALADRL

SEQ ID NO: 12

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDIAIDFPADRGWDVESLYDPDPAAGKSYCVRGGFLAAAAEFDAAFF
GISPREALAMPQQRLVLETSWEAFERAGIEPGSVRGSDTGVMGAFAGGYGAAVEGFGATAGATSVLSGRVSYFFG
25 LQGPATVDTACSSSLVALHQAGYSLRQGECSMALVGGVTVMATPQSFVEFSRQRGLAPDGRCKAFADTADGTWSE
GVGVLLVERLSDAERNGHRVLAVVRSSAVNQDGASNGLSAPNGPAQQRVIRQALANAGLAAADVDVVEAHGTGTTLG
DPIEAQAVLATYQDRERPLLLGSLKSNVGHTQAAAGVSGVIKMVMALRHGVVPRTLHVDEPSRHVDWTAGAVHLVT
ENQWPDPDTRPRRAGVSSFGVSGTNAHVIIIEGSPTSSPVAEPSPGDVPLVVSATPQALTAYEDRLRAFLAAAPVTD
TRAVASTLAVTRSLFEHRAVLVDNTVTGTALAEPRVVFVFPQGQWQLGMGAALMESVVFMAERMAECAAALGEFVD
30 WDLAVLDDSAVVDRVDVVQPACWAVMVSLAAVWQDAGVRPDAVIGHSQGEIAAACVAGAIISLRDAARIVALRSRLI
SERLGKGAMASITLPADQITLAEGAWIAAYNGPASTVVAGTPDAIEQMQRVRRIAVDYASHTPHVEQIRAEELDL
TAEVGSRTPTVPWYSTVDGTWIDSPLDGEYWRNLRQPVGFHFAVQTLQALGETVFVEVSASPVLPPAMDDAVTVAT
LRRDEGTLTRMHTALAEASHVLGVSIDWPHVLGDTGERMLDLPTYAFERHRYWSTARRNPSIAPDDLLTVVRDSAAVV
LGYADGGAVPVTGAFKDLGIDSLTAVELRNGLAKATGLRLPATVAFDDYPTPQALAARL
35

SEQ ID NO: 13

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDAITGFADRGWAEYSFQGGFLDDAADFDAAFFGISPREALAMPQQ
RLVLETAWEAFERAGIEPGSLRGSDTGVMGAYPGGYGIGADRAGFGATAGAGSVLSGRVSYFFGLEGPVAVTDTAC
SSSLVALHQAGHALRLGECSLALVGGVTVMATPDTFVEFSRQGGLAADGRSKAFADSADGAGFAEGAGVLLVERLS
40 AQRHGHQVLALVRGSAVNQDGASNGLTAPNGPSQQRVIIQAALDNAGLTAAEVDVVEAHGTGTTLGDPPIEAQAVIAAY
GQGRGEPLLLGSIKSNVGHTQAAAGVSGVIKVMALRHGVVPRTLHVDEPSRHVDWTAGAVRLATENQSWPETGRPR
RAGVSSFGISGTNAHVILEGVPEEPAGHEEPAGLTPLLIISAKTPAALAEFEDRLRAYLTTEPSLPVASTLARTRSL

FDHRAVVLGDGVVRGVAEPDRRVVVFVSGQGSQRAGMGDDLAAAFPVFAKIRQQVWDQLDIPDLVPDQTYAQPALF
 ALQVALFGLLDSWGVRPDALVGHSIGELAAGYVAGIWSLEDACALVSARARLMQALPPGGVMVAVPVSEQQARGALT
 DGVEIAAVNGPASVVLSGDEAAVLRAAAALGGRSKRLATSHAFHSARMEPMLDEFMRVAERLSYGSPRISMAVGDGP
 DYWVRQVREAVRFGEQVAAHGAVFVELGPDGSLARLIDGIAMLDLDRDDEPRAALTALARLHVQGVKVDWPIGAGRRV
 5 DLPTYPFQRQRYWIDRPTARRAPTDLTLVRDTAATVLGYPDSSAVPATTAFKDLGVDSLTAIELRNGMATTTGLRL
 PATLVFDYPTPAALAARL

SEQ ID NO: 14

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALSGFPTDRGWAHESFQGGFLDGAGDFDAPFFGISPREARVMDPQQ
 10 RLVLEASWEAFERAGIEPGTVRGS DTGVFMGAYSGGYAAGADLAGFAATAGAGSVLSGRVSYFFGLEGPAVTVDTAC
 SSSLVALHQAGHALRQGECSLALVGGVTVMATPDLFVEFARQQGLAADGRCKAFADNADGTGWSEGVGVLVERLSD
 AERNGHRVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIIQAALDNAGLTPADIDVVEAHGTGTTLGDPPIEAQAVIATY
 GQTREQPLLLGSLKSNVGHTQAAAGVSGVIKVMALRHGVVPRTLHVDEPSRHVDWTAGAVQLAVENQPWPNTGRPR
 RAGVSAFGVSGTNAHVIIIEGSPTPSPVAEPSGDVLPVLISAKTPQALTAYEDRLRITYLNATPEIDTRAVASTLAVTR
 15 SLFEHRAVLLGDNTVSGTALTEPRVVFVFPQGQWLGMGAALMESVVF AERMAECAAALSEFVDWNLITVLNDPAV
 VDQVDVVQPACWAVMVSLAAVWQDAGVRPAAVIGHSGQEI AACVAGAI SLRDAARIVALRSRLIGERLGRGAMASV
 ALPADEIALVDEVWVAAYNGPASTVIAGAPDAIEQMLGDRVRRIAVDYASHTPQVEQIRAE LLDLTAEVSSQAPTVP
 WYSTVDGTWIDGPLDSYWRNL RQPVGFHPAVEALGGLGETVTFVEVSASPVLLPAMDDAVTVATLRRDEGTLTRMH
 TALAEAHVLGVTIDWPAVVGDTGERMLDLPTYAFQHHRWTTATARLEGRTGAEKHRLLLDIVLANAATVLGHDTAD
 20 TIASDKPFKDLGIDSLTAVELRNSLARATELRLPATTAFDYPTPEALATRL

SEQ ID NO: 15

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALITGF PADRGWPDDSRQGGFLDDAADFDAAFFGISPREALAMPQQ
 RLVLEAAWEAFERAGIEPGSLRGSDTGVFMGAYPGGYGIGADQAGFGTTAGAGSVLSGRVSYLFGLEGPAVTVDTAC
 25 SSSLVALHQAGHALRLGECSLALVGGVTVMGTPDIFAEFSRQGGGLASDGRCKPFADAADGTGWAEGVGVLVERLSD
 AERHGHQVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIIQSALI QAGLAPHEVDVVEAHGTGTTLGDPPIEAQAVIAAY
 GQDRAQPLLLGSIKSNVGHTQAAAGVSGVIKVMALRHGVVPRTLHVDEPSRHVDWSAGAVRLATESQPWPD TGHPR
 RAGVSSFGISGTNAHVILEGVPEEPADTGEPSGLVPLLLSAKTPAALTHLEDRLRAYLTTEPNLPVASTLAQTRSL
 FDHRAVLLDGDGVVRGVAEPDRRVVVFVSGQGSQRAGMGDDLAAAFPVFAKIRQQVWDLLDIPDLVPDETGHQAQPALF
 30 ALQVALFGLLDSWGVRPDALVGHSIGELAAGYVAGIWSLEDACALVSARARLMQALPPGGVMVAVSVSEEQARAVLT
 DGVEIAAVNGPASVVLSGEETAVLQAAAALGGRSKRLATSHAFHSARMEPMLDEFMRVAERLSYGSQPQIPMAVGDGP
 DYWVRQVRET VRFGEQVAAHGGIFVELGPDGSLARLVDGIAVLDRDDEPRAALTALARLHVQGVKVDWPIAAGRRV
 LDLPTYPFQHQRWATRPAARRAPTDLTLVRDTAATVLGYPDSSAVPATTAFKDLGVDSL TAVELRNNLATSTGLR
 LPATLVFDYPTPATLAARLD
 35

SEQ ID NO: 16

EPLAIVGMACRLPGGVSTPEDLWQLVESGTDALISGF PADRGWDDYPYQGGFLTTAADFDAAFFGISPREALAMPQQ
 RLILEASWEAFERAGINPADARGSDTGVFMGAFSAGYGDDRDDSPATAGAVSVLSGRVSYFFGLEGPAMTVDTACSS
 SLVALHQAGYSLRHGECSMALVGGVTVMATPRTFVEFARQQGLAEDGRCKAFADTADGTGWAEGVGVLVERLSDAE
 40 RNGHRVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIIQAALDNAGLAPQDQVDVVEAHGTGTTLGDPPIEAQAVIATYQQ
 NRQQPLLLGSIKSNVGHTQAAAGVSGIIKMIMALRHGVVPRTLHVDEPSRHVDWTAGAVRLVTENQPWPDADRPRRA
 GVSSFGISGTNAHIIIEGVPEEPAQPDSELP TPLVISAKTAPALTQFEARLRSYLTTEPALS AVASTLAQTRSLFD

HRAVLLGGDTITGVAEPSRVVVFVFSGQGSQRAGMGDELAAPVFAKIRQQVWDLDDIPDLPVDETGHAPALFAL
 QVALFGLLDSSWGVPRDALIGHSIGELAAGYVSGIWSLEDACALVSARARLMQASPPGGAMVAVPVSEQQARAVLTDG
 VELAAVNGPSSVVLSGDETAVLQAAAAALGGRSKRLATSHAFHSARMEPMLDEFRAVAEQLSYRSPQIPMAVGDGPEY
 WVRQVRDTRVFGEQVAHDGAIFVELGPDGSLVRLIDGIPMLDRDDEPRAALTALARLHVRGVNVAWPIAADRRELD
 5 LPTYPFQRRERYWSTASLSALAPAEREQALRKVVSDSSAMVLGYAEGRAVAPTAAFKDLGVDLSLTAVELRNSLTAKATG
 LRLPATIVFDYPTPGALAVRL

SEQ ID NO: 17

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALSRFPADRGWDVDGLYDPDPDAPGKSYSVEGGFLDAVADFDAFF
 10 GISPREALAMDPQQRLLILEASWEAFERAGIEPGSLRGS DTGVFMGAYSSGYGIGADIPGLGVTAGAVSVVSGRVSYF
 FGLEGPAVTVDTACSSSLVALHQAGHALRRRECSLALVGGVTVMATPFGFVEFSRQRGLASDGRCKAFADTADGTSW
 SEGAGVLVVERLSDAERHGHQVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIIQAALANAGLTPHEVDVVEAHGTGTR
 LGDPIEAQAVIATYQGARGEPLLLGSIKSNVGHGTQAAAGVSGVIKVMALRHGVVPRTLHVDEPTRHVDWTTGAVRL
 ATENQPPETERPRRAGVSSFGVSGTNAHIIIEGVAAEPAQPGESPELTPLLLSAKTPAALTHLEDRLRAYLTTEPN
 15 LPAVASTLAQTRSLFDHRAVLLGGETVTGVAEPDPRVVFVFSGQGSQRAGMGDDLAAPFAFAKIRQQVWDQLDIPN
 LPVDETGHAPALFALQVALFGLLDSSWGVPRDALVGHSIGELAAGYVAGIWSLEDACALVSARARLMQALPPGGVMV
 AVSVSEEQARAVLTDGVEIAAVNGPASVVLSGEETAVLQAAAAALGGRSKRLATSHAFHSARMEPMLDEFRAVAEQLS
 YGSPRIPMAVGDGPDYVWRQVRDTRVFGEQVAHDGAIFVELGPDGSLARLIDGIAVLDRDDEPRAALTALARLHVR
 GVKVDWPIAAGRRELDLPTYPFQHORYWIDSRPTARRAPTDLTLVRD TTATV LGYPDNTAVTPTTAFTDLGIDSLT
 20 AIELRNMMATTTGLRLPATLVFDYPTPATLAARLD

SEQ ID NO: 18

EPLAIIIGMACRLPGGVTTPEDLWQLVETGTDALISALPTDRGWADHPYQGGFLTTAADFDAAFFGISPREALAMDPQQ
 RLILETSWEAFERAGINPADAHGSDTGVMGAYSSGYGIGADLAGFGATAGATSVLSGRVSYFFGLEGPATVDTAC
 25 SSSLVALHQAGHALRHGECSLALVGGVTVMATPDIFVEFARQRGLAADGRCKAFADTADGTGWAEGVGVLVVERLSD
 AERNGHRVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIIQAALDNAGLTPADIDVVEAHGTGTTLGDP IEAQAL IATY
 GQNREQPLLLGSLKSNIGHTQAAAGVSGVIKVMALQHGVVPRTLHVDPKPSRHVDWTAGAVRLLTESQPWPDTRPR
 RAGVSSFGVSGTNAHVIIIEGSPTPSPVADPSGDVLPVLISAKTPAALAAAYEDRLRTYLNATPEIDTRAVASTLAVTR
 SLFEHRAVLLGEDTVSGTALTEPRVVFVFPQGQWQWLGMAALMESVVF AERMTECATALSEFVDWNLTITVLNDPAV
 30 IDQVDVVQPACWAVMVSLAAVWQAAGVRPDAVIGHSGQEIAAACVAGAI SLRDAARIVALRSRLISERLGKGAMASI
 TLPADQITLAEGAWIAAYNGPTSTVVAGTPQAIEQM HGERVRR IAVDYASHTPHVEQIRAE LLDLT TDVSSQTPTLP
 WYSTVDGTWIDSPLDGDYWRNLRQPVG FHPAVQTLQALGETVFVEVSASPVLLPAMDDAVTIATLRRDEGTLTRMH
 TALAEAHVLGVTIDWPHVLGDTGERMLDLPTYAFQHRYWTTAARLTGRTTAAQHRLMLDFVLGNVAAVLGHGSAGD
 VAPDKPFKELGMDSLTSVELRNSLAKATGQRLPATIVFDHPTADALATYL
 35

SEQ ID NO: 19

EPIAIVSMACRVPGGVTSPEGLWRLVESGTDALISEFPGDRGWDVANLYSPDPDAPGKSYS LQGGFLDGAAAFDASFF
 GISPREALGMDPQQRLLLETSWEAVERARIDPKSLRGRDVG VYVGAAQGYGLGAAEAQRDNLTGGISLLSGRLS
 YALGLEGPGLTVDTACSSSLVALHLAAQALRQGECSLALVSGVSMPTPDVVFVEFSRQRGLAADGRCKSF AAAADGT
 40 SWSEGVGVLLLERLS DARRLGHEILAVVRGTAVNSDGASNGLTAPNGASQQRVIRQALASAGLGPADVDAVEAHGTG
 TKLGDP IEAEAILATY GKRPTPVWL GSLKSNIGHTMAASGV LGVIKMVESMRHGVLPRTLHVDEPSPHVDWAAGDV
 ALLTSNQPPAGRKPRRAGVSSFGLSGTNAHVLEQYRMPAAPVTTKEAGPLPWVLSAQTP EALRERAGQLATALAG

DPAWHPLDVGYSLAATRSTFAHRAVVVGDDREFVRTLGKLADGAGWPGLTTGVAKSRRIAFMFDDGQGTQRLAMGQGL
 YARFPAFTRTWDTVSAEFAKHLDTLTDVYLGGGGTAAAEVDDPLYAQAGIFAVEVALVELLAEWGVRPDVVTGHS
 IGEAAAAYTAGMFSLADVTALITARGAALRSAPPGAMLALRAGEPEVRDFLDRTGAALDVAAVNGPAAVVVSGAPDA
 VAGFASAWTASGRECRQLKVRRAFHSRHVEGVLDGDFRTVLKSLTFRTPALPIVSTVTGRLIDPAEMGTPEYWLSQVR
 5 QPVRFQDAVGELAGQGVSAFLEVGPSGTLASAGMECLDASFHALLRPRPAEDIGVLTALAELYAGGTAVDWATVLAG
 GRPVDLPVYPFQHQSYWLSAPDEPRTVLEMVHLEVASILGITDPDAVQDDSSFLELGFDSLGSVRLRNRLTQVTGL
 TLPATLLFDHDTPSALATELD

SEQ ID NO: 20

10 EPLAVVGMACRLPGGITSPEELWELVEDGGDAVGDFPTDRGWDVAALHAAAESATSRAGALMGAADFDAFFFGISPR
 EATALDPQQRILLEIAWEAIERAGIKADVLRGTDGTFVVGFFYYGYGAGADLGGFGAYSTQPAVLAGRLSYFFGLEG
 PAVTVDTACSSSLVALHQAGQALRAGECSLALVGGVTVMAS PQSFVEFSRQGGVAPDGRCKAFADAADGTGFAEGAG
 VLVVERLSDAERNGHTVLAVVRGSAVNQDGASNGISAPNGPAQQRVIRQALGSAGLAPADVDDVEAHGTGTVLGDPI
 EAQAVLATYGGQREVPLLLGSLKSNIGHAQAAAGVAGVIKMMAMRRGVVPRTLHVDEPSSHVDWTTGAVELLTEAR
 15 WPESDRPRRAGVSAFGVSGTNAHVILEEVAESSVRSGSSGLVPLPVSARTESSLAVQVERLGAYVRSGADLSAVA
 DGLVRERVVFGHRAVLLGESTVAGVAEGELRTVFVFPQGQSQWVGMRGELMGASEVFAARMRECAAALEPHTGWDLL
 DVLGEAVVADRVEVLQPASWAVAVSLAALWQAHGGTPDAVIGHSQGEIAAASVAGALSLEDAARIVALRSQTIAARL
 GRGAMASIAIPSAEVEVMEGVWVAARNGPSSTVIAGDPAAVEQVLARYEAEGVRVRRIAVDYASHTPHVEAIQDELA
 EVLEGVTAQVPTIPWWSTVDSWVTEPVDDDYWRNLRQPVAMDTAIGELDGSLEFIECSAHPVLLPALDQERTVASL
 20 RTDDGGWERFLTALAEAWTQGADVDTILVEPAPHRLDLPTYPFDHKRYWLLERLGAMTGADRDAALLTLVRDCAA
 VLGHVDAAGVPADAFAKDLGVDSLTAVELRNRLAAATGVRLPATLAFDHPTPRAIASRLD

SEQ ID NO: 21

EPLAIVGMACRLPGGVASPGDLWQMLDSGGDAVTGFPVDRGWDPGLTGGPDADRGGFLSDAADFDAFFFGISPREA
 25 LAMDPQQRILLETTWEAFENAGIVPGTLRGSDTGVMGAFSYGYGVGADLGGFGSIGVQPSVLTGRISYFYGLQGPA
 FTVDTACSSSLVALHQAGHALRHGEC SLALVGGVTVMANPDG FVEFEQQGGLSPDGRCCRAFADAANGTGWAEAGAGVL
 VVERLSDAERNGHTVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANAGLGAADVDDVEAHGTGTVLGDPIEA
 QAVLATYGGQREVPLLLGSLKSNIGHAQAAAGVAGVIKMMAMRRGVVPRTLHVDEPSSHVDWTAGAVEVVTEARPW
 PESGRVPRAGVSSFGVSGTNAHVVLEGAPEPSSGAEASSGGGLVPLPVSARTESSLAVQVERLGAYVRGGADLGAVA
 30 DGLVRGRAVFDRAVLLGESTVAGVAVEGARTVFVFPQGQSQWVGMRGELMGVSEVFAARMRECAAALEPYTGWDVL
 DVLGEAVVADRVEVLQPASWAVAVSLAALWQAHGVVDPDAVVGHSQGEIAAACVAGALSLEDAARVVALRSQTIAARL
 GHGAMASIALPASAVEVMEGVWIAARNGPSTVAGDPAAVERVLARYEAAGVRVRRIAVDYASHTPHVEAIQDELA
 DVLGGITSSAPDISWWSTVDSGWVTEAVGDDYWRNLRQPVAMDTAVSEL DGSLEFIECSAHPVLLPALDQERTVASL
 RTDDGGWDRFLTALAQAWTQGADVDTTLEIPAQHRLDLPTYPFDHKRYWLQPAGARNEVARHTDLLTLVRQKAAAL
 35 LGHAGPEDVPEDAAFRQLGVDSLIAVQLRNGLNEATGLRLSATLVFDYPTPRALAGRI

SEQ ID NO: 22

EPVAIVGMACRLPGGVTSPEDLWRLVASGTD AITEFPADRGWDVDALFDPDPDAVGRSTTRHGGFLTEATGFDAAFF
 GISPNEALAMDPQQRVLVLETSWEAFEHAGIVPDTLRESDTGVFMGAFHQGYGAGRDLGGLGVTATQTSVLSGRLSYF
 40 YGLQGPVAVTVDTACSSSLVALHQAAQALRSGECSLALAGGVTVMATPGSFVEFSRQRGLSPDGRCKAFADSADGTGF
 AEGVGVLVVERLSDAERNGHTVLAVVRGSAVNQDGASNGLSAPNGVAQQRVIRQALANAGLNGTDVDAVEAHGTGTV
 LGDPIEAQAVLATYGGQEREVPLLLGSVKSNVGHTQAAAGVAGVIKMMAMRRGVVPRTLHVDESSSHVDWSAGAVEV

VTEARPWPESGGARRAGVSSFGVSGTNAHVILEGVAESSVRSGGSAGLVPLPVSARTESSLALQVERLGEYVRGGA
 DLGAVADGLVRGRAVFGRRAVLLGESTVAGVAVEGARTVFVFPQGGSQWVGMGRELMGVSEVFAARMRECAAALEPH
 TGWDVLDVLGEAVVADRVEVLQPASWAVAVSLAALWQAHGVVPAVVGHSQGEIAAACVAGALSLEDAARVVALRSQ
 TIAARLGHGAMASIALPASAVEVMEGVWIAARNGPESTVVAGDPAEVERVLARYEAAGVRVRRIAVDYASHTPHVEA
 5 IQDELADVLGGITSSAPSVPPWSTVDSGWVTEPVDDDYWRNLRQPVAMDTATGELDGSLEFIECSAHPVLLPALDQE
 RTVASLRTDDGGWERFLTALAEAWTQGADVDWTTLIEPAQHRVDLPTYPFDHKRYWLQPARRTVRTGEDSGRDLLAV
 VCGATAAVLGHADASEIGPATAFKDLGIDSLSGIRLRLNSLAETTGVRLSATAVFDHPTPDALAARL

SEQ ID NO: 23

10 EPLAIVAMACRMPGGVDTPEDLWRLVESGGDAITEFPTDRGWDLAALYDPDPDAIGKVSVRHGGFLAGAADFDAEFF
 GISPREALAMDPQQRLLILEVSWEAFERAGILPASVRGSDAGVFMGAFTQGYGAGVDLGGFGATGTPTSVLSGRLSYY
 FGLEGPSVTVDTACSSSLVALHQAARSLRSGECSLALVGGVTVMATTTGTFVEFSRQRGLAPDGRAKAFADTADGTSE
 AEGAGVLIVERLSDATRLGHPVLAVVRGSAVNSDGASNGLSAPNGPAQRRVIERALDDAGLVPGDIDAVEAHGTGTR
 LGDPIEAQALEAAYGLDRVHPLLLIGSLKSNLGHGTQAAAGVAGVIKMLAMRHGVLPRTLHVDEPSRHDWGGGVRL
 15 RRNEPWPVTGRVRRAGVSSFGISGANAHVIEAGPPAAPATLPATEPVPEGVWPVSARTPDGVRDVAGRLVALTAP
 AAAIGHSLATTRTAMRHRVVPARDAEAFARGEVPGVVRGTADVTDARAVFVFPQGGSQWDGMAELLATEPVFAR
 RLGECAEALAPYTGWDLDDVIARRPGAPALDRVDVVPVSFAMMVALAELWRSRGVAPAAVVGHSQGEVAAACVAGV
 LTLDDAAKVVALRSRLVATELGHGGMVSVPPADFDAAAAGRLEVAAVNGPASIVVAGAADAEEELLAATPHARRIA
 VDYASHTAHVETIRDALLDALADLTPGAPEVPPFFSTVDEAWLDRPADAAAYWDNVRPVRFGAATARLAELGYRVFV
 20 EASPHPLVLTALADTLAGHPNTAVTGTLLRGDGGARRFTSSLAELWVRGVPVSWPSGESRRVPLPTYPFRRDRYWD
 AEAAPTAARDMLELVRTSAALVLGHRDAHAIEPTRAFKEVGFDLSLTGVELRNRLADATGLTLPATLVFDHPTAQALA
 AHL

SEQ ID NO: 24

25 EPLAIVGMACRLPGGVASPEDLWRLLESGGDGITTFPGDRGWDVEALYDPDPEHPGTSTVRHGGFLSGAGDFDAGFF
 GISPREAVAMDPQQRVVMETSWEALEYAGIDPHTLRGSDTGVMGGYFYGYGSGADRGGFGATSTQTSVLSGRLSYF
 YGLEGPAVTVDTACSSSLVALHQAGQSLRTGECSLALVGGVTVMASPSGFVDFSQQRGLAPDGRCKAFEAADGTGF
 AEGSGVLVVERLSDAERHGHRLVAVVRGSAVNQDGASNGLSAPNGPSQERVIRQALANAGLQPSDVEDAVEAHGTGTR
 LGDPIEATALLATYQDRATPLLLGSLKSNIGHTQAAAGVAGIIMVLAMHHDTPSTLHVDTTPSSHVDWTAGTVEL
 30 LTDARPWPETSRPHRAAVSSFGVSGTNAHVILESHPRPTPAPDTGSSTHPVPLLLISARTPRALSEHTTRVSAFLDAG
 GGDERAVASALLTRTAFTHRAALIGTDLITGTAVPDRLVWLFSGQGSQRPGMGDELAAYDVFASTRRDVLDALQV
 PAGLDIHDGTGYAQPAVFALQVALSAQLDAWGVRPDALVGHSIGELAAAYVAGVWSLDDACALVSARARLMQALPPGG
 AMAAVIASERDALPLLREGVEIAAVNGPASIVLSGDEDAVLVDVAARLGRFTRLRTSHAFHSARMEPMLDEFRDVAQR
 LTYHEPKLPMAAGADCATPEYWVRQVRDTRVFGEQVAAVDGAALLEIGPDRNLARLVDGIPVLHGDDEARSAMTALA
 35 RLHTGGVAVDWPEVIGAAPTHLNLPTYPFERTRYWLGSRDRIAGLTAADA EKAALAVVRECAAAVLGHEGPARI EAT
 ATFKELGVDSLTAVRLRNAFTEATGVRLPATAVFDFTPTQAVAAKL

SEQ ID NO: 25

EPLAIVGMACRLPGGVASPEDLWRLLESGGDGITAFPADRGWDVEALYDPDPEHPGTSTVRHGGFLSGAGDFDAGFF
 40 GISPREAIAMDPQQRVVLETSWEALEQAGIVPGTLRGSDTGVMGAFSDGYGLGTDLGGFGATGTQTSVLSGRLSYF
 YGLEGPAVTVDTACSSSLVALHQAGQSLRTGECSLALVGGVTVMASPGGFVEFSQQRGLAPDGRCKAFEAADGTAF
 AEGSGVLVVERLSDAERRGHRLVAVVRGSAVNQDGASNGLSAPNGPSQERVIRQALANAGLRPSDVEDAVEAHGTGTR

LGDPIEATALLATYQGDRATPLLLGSLKSNIGHTQAAAGVAGI IKMVLAMRHGSLPRTLYVDTPSSHVDWTAGGVEL
 LTDARPWPATTGPRRAAVSSFGVSGTNAHVILEAHAAPEPPALDSPVVEPSASLFATELTPLPVSARTSEAVDGGVQ
 RLREHLATHPGDDPRAVAAALLATRTDFPHRAVLLGDGVVTGTALTAPRTVFVFPQGGSQWLGMGRKLM AESPVFAA
 RMRQCADALAEHTGRDLIAMLDDPAVKSRVDVHPVCWAVMVSLAAVWEAAGVRPDAVIGHSQGEIAAACVAGAI SL
 5 EDGARLVALRSALLVELAGRGAMGSI AFAAADVEAAAAARIDGVWVAGRNGTATTIVSGRPDAVETLIADYETRGVWV
 TRLVVDCPTHPTFVVDPLYDELQRIVAATTSRAPEIPWFSTADERWIDAPLDDEYWFNRNMRNPVGF AAAVAAAAREPGD
 TVFIEVSAHPVLLPAINGTTVGTLRRGGGADRLLDSLAKAHTVGVAVDWAAHDAATGTADLPTYAFHHERYWIEPAE
 RLPDL SRKEQE QVLLDVVRDTAATLLGHADARAVTATAAFKDLGVDSL TALGLRDR LAEALGIPLPATLVFDHPAAG
 TLSRHL

10

SEQ ID NO: 26

EPLAIVGMACRLPGGVASPDLLWRLLESGGDGIGAFPGDRGWETGADGRGGFLSGAAGFDAAFFGVSPREALAMDPQ
 QRVVLETSWEALEHAGIDPHTLNGSDTG VFLGAFFQGYGIGADFDGYGTTSIHTSVLSGRLSYFYGLEGPAVTVDTA
 CSSSLVALHQAGQSLRTGEC SLALVGGVTVMASPA GFADFSEQGGLAPDGRCKAF AEAADGTAFSEGSGLVVERLS
 15 DAERHGHRI LAVVRGSAVNQDGASNGLSAPNGPSQERVIRQALANAGLQPSD VDAVEAHGTGTRLGDPIEATALLAT
 YGQHRTTPLLLGSLKSNIGHTQAAAGVAGI IKMVLAMHHDTLPP TLHVDTPSSHVDWTGGVELLTDARPWP TTTGP
 RRAGISSFGVSGTNAHVILESP TPVPSPGAEPGARPVLP I SARTPEALDEHTIRIRAFLDDNPGADHVAVAQTLAR
 RTPFEHRAVLLGDTLITADPNAGSGPVVFVYSGQSTLHPHTGRQLAATYPVFADAWGEVLGHLDADQGPATHF AHQI
 ALTALLRSWGIAPHAVIGHSLGEI SAACAAGVLSLGDASALLAARSRLMDELPA GGAMVTVLTSEENALRALRPGVE
 20 IAAVNGPHSVVLSGDEGPVLAVAQQLGIHHR LPTRHAGHSARMDPLVAPLLEAASGLTYHQPRIAIPGDPTTAAYWA
 RQVRDQVR FQAHAE RYPGATFLEIGPNQDLSPVVDGIPTQTGTPDEVQALHTALARLHTRGGVVDWPTVLGGDRAPV
 ALPTYPFQHKDYWL RATELAVLPDDERADALLAFVRNSTATVLGHLGAEDIPATATFKELGIDSLTAVQLRNALTTA
 TGVRLNATAVFDFTPRALAARL

25 SEQ ID NO: 27

EPLAIVGMACRLPGGVASPEGLWRLVASGTD AITEFPADR GWDVDALYDPDPAIGKTFVRHGGFLDGATGFDAGFFG
 ISPREALAMDPQQRVLLETSWEAFESAGITPDSARGSDTG VFIGAFSYGYGTGADTNGFGATGSQTSVLSGRLSYFY
 GLEGPSVTVDTACSSSLVALHQAGQSLRSGECSLALVGGVTVMASPGGFVEFSRQRGLAPDGRAKAFGAGADGTSFA
 EGAGALVVERLSDAERHGH TVLAVVRGSAVNSDGASNGLSAPNGPSQERVIRQALANAKLTPADVDAVEAHGTGTRL
 30 GDPIEAQALLATYQGDRATPLLLGSLKSNIGHAQAASGVAGI IKMVQAIRHGELPPTLHADEPSPHVDWTAGAVELL
 TSARPWPGTGRPRRAAVSSFGVSGTNAHI ILEAGPVKAGPVEAGPVPAAPPSAPGEDLPLLVSARSPEALDEQIGRL
 RTYLDTRPGVDRAAVAQTLARRTHFAHRAVLLGDTVIT TSPSHQADELVFVYSGQGTQHPAMGEQLAAAFPVFAETW
 HDALRRLDDPDHPDPTRSQHTLFAHQAAALTALLRSWDITPHAVIGHSLGEITAAAYAAGILSLDDACTLITTRARLMH
 TLPPPGAMVTVLTGEEEARQALRPGVEIAAVNGAHSVVLSGDEDAVL DVAQRLGIHHR LPAPHAGHSAHMEPVAAEL
 35 LATTRRLRYDRPHTAIPNDPTTA EYWAEQVRNPVLFHAHTQQYPDAVFVEIGPGQDLSPLVDGIALQNGPANEHAHL
 RTALARLFSRGATLDWPLVLGGASRHDPDVPSYAFQQRPYWIESARLAELPDADRDTALSTLVMDATAAVLGHADAS
 EIGPTTTFKDLGIDSLTAIELRNRLAEATGLRLSATMVFDHPTPRVLA AKL

SEQ ID NO: 28

40 EPLAIVGMACRLPGGVTSPEDLWRLVASGTD AITEFPTDRGWDIDRMFDPDPDAPGKTYVRHGGFLSEAAGFDAAFF
 GISPREAWAMPDQQRVILETVWEAFENAGIVPDTLRGSDTG VFMGAFSHGYGAGVDLGGFGATATQNSVLSGRLSYF
 FGMEGPAVTIDTACSSSMVALHQAAQSLRDGECSLALAGGVTVMP TPLGYVEFCRQRGLAPNGRAKAF AEGADGTSF

SEGAGVLVVERLSDAERNGHTVLALVRSSAVNQDGASNGISAPNGPSQQRVIRQALDKAGLTPADVDVVEAHGTGTP
LGDPIEAQAI IATYQDRDTPLYLGSVKSNIGHTQTTAGLAGVIKVMAMRHGLLPKTLHVDEPSSHVDWSAGAVEL
LTEARPWPDSDRPRRAGVSSLGISGTNAHVILEGVAESSVRSGGSSGLVPLPVSARTESSLALQVERVGEYVRGGAD
LGAVADGLVRGRAVFDRAVLLGESTVAGVAVEGARTVFVFPQGGSQWVGMGRELMGASEVFAARMRECAAALEPHT
5 GWDVLDVLGEAVVADRVEVLQPASWAVAVSLAALWQAHGVVPDVAVIGHSQGEIAAACVAGALSLEDAARVVALRSQT
IAARLGHGAMASIALPASAVEVAEGVWIAARNGPESTTVVAGDPGAVERVLARYEAAGVRVRRIAVDYASHTPHVEAI
EEQLADVLGGITSSAPDISWWSTVDSGWVTEPVGDDYWYRNLRQPVAMDTAISELDGSLFIECSAHPVLLPALDQEH
TVASLRTDGDDWRFLTLALAQAWTQGAPVDWTTLIEPAPHRDLPTYPFDHKRYWIEAAARLAGHTAAEQRRVMQEV
VLRQAAAVLAYGLGEQVAADRPFRDLGFDSLTAVDLRNRLAAETGLRLPTTVVFSHPTAEALATHL

10

SEQ ID NO: 29

EPIAIVAMACRLPGGVTSPEELWRLVESGTDAITMAPGDRGWDLDALYDPDPDAVGKAYNLRGGFLEGAAEFDAAFF
DISPRESLGMDPQQRLLETAWEAIERGRINPASLHGREIGVYVGAAAQGYGLGAEDTEGNAITGGSTSLLSGRLAY
VLGLEGPSVTVDTACSSSLVALHLACQGLRLGECALALAGGVSVLSSPAAFVEFSRQRGLAADGRCKSFSGSGADGTT
15 WAEGVGVLVLERLSDAERLGHTVLAVVRGSAVTS DGASNGLTAPNGLAQQRVIRKALAAAGLTAADV DLVEGHGTGT
RLGDPVEADALLATYQNRQEPVWLGLSKSNIGHATAAAGVAGVIKTVQAIGAGTMPRTLHADEPSPAVDWTAGRVS
LLTG NR PWPDDERARRAAVS AFGLSGTNAHVILEQHRPEPVAPRPPREPRPLPWVLSARTPAALRAQAARLRDHLA
AVPDADPLDIGYALATSRARFTHRAAVVATSSDEFRAGLDSVADGVEAPGVVGGTARERRVAF LFDGQGAQRVGMGR
ELHGRFPVF AA WDEVSDAFGKHLEHSPTDVFHGEHGD LAHDTLYAQVGLFTLEVALLRLLHEHWGVRPDVLVGH SVG
20 EVTAAYAAGVLT LADATALIVARGRALRALPPGAMTAVD GSPA EVGAFTGLDIAAVNGPSAVVLTGSPDDVTAFERE
WAAAGRRAKRLDVGHAFHSRHVDGALDDFRTVLESLSFGAARLPVVSTTTGRDAAGDLATPEHWLRHARRPVLFADA
VRELADLGVNMFVAVGPSGALASAASENTGGSAGTYHAVLRARTGEENAALTAVAE LHAHGAPVDLAAVLAGGRPVD
LPVYPFQHRSYWLAPDDLTVAEIVRRRAAALLGIADPGD V DADTTFFALGFDSLAVQRLRNQLTAATGLDLPTAVLF
DHDTPSALTAYL

25

SEQ ID NO: 30

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISGFPTDRGWDVENLYDPDPDASGKSYCVQGGFLDAAAAGFDAGFF
GISPREALAMDPQQRLVLEVSWEAFERAGIEPGSVRGTD TG VFI GAYPGGYGAGAGTELEGYGTTSGPSVLSGRVSY
FFGLEGPAITVD TACSSSLVALHQAGYALRQGECSLALVGGVTVMATPDVFTEFARQRGLAADGRSKAFSDSADGAG
30 FSEGIGVLLVERLSDAQAKGHQVLAVVRSSAVNQDGASNGLTAPNGPSQQRV IQAALGNAGLT TAEVDVVEGHGTGT
TLGDPIEAQALLATYQDRERPLLLGSLKSNIGHTQAAAGVSGVIKVMALRHGLVPRTLHVDEPSSHVDWTAGAVE
LVTANQPWPADADRPRRAGVSSFGVSGTNAHVILE SAPSTQAVDDVRPVETPVVGSELVPLVLSAKTLPALSGYEDRL
RAYLAGSPGVDLRAVASTLAVTRSVFEHRAVLLGDDTVTGTA VTDPRVVFI FPGQGSQRAGMGEE LAAAFPVFARIH
QQVWDL LDVPDLEVNETGYAQPALFALQVALFGLLESWGVGPDAVVGH SVGELAAAYVSGVWSLEDACTLV SARARL
35 MQALPPGGVMVAVPPEDEARAVLGEGVEIAAVNGPSSVVL SGDEAAVLR AAATLGKWMRLATSHAFHSARMEPMLD
EFRAVAERLTYQTPHLTMAAGEQVTT PDYWVRQVRDVVRFG EQVASFEDAVFVELGADRSLARLVDGVAMLHGAHEA
QAAISALAHLYVNGVTVDWPAVLGDVPGRVLDLPTYAFQHQR YWLEGWLAALTPAEREKALLKLVSDGAATVLGHAD
TSTIPVTGAFKDLGINSLTAVELRNSLAKATELRLPATLVFDYPTPATLAARLD

40 SEQ ID NO: 31

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISGFPTDRGWDVENLDSAGKSYRAEGGFLDAAAAGFDASFFGISPR
EALAMDPQQRLVLEVSWEAFERAGIEPGSLRGSDTGVMGAYPGGYGIGADLGGFGATAGATSVLSGRVSYFFGLEG

PAFTVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPQTFVEFSRQGLASDGRCKAFADAADGTGWAEGVG
 VLLVERLSDAQAKGHQVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIIQAALSNAGLTAEVDVVEAHGTGTTLGDP
 EAQALLATYGQDRERPLLLGSVKSNLGHTQAAAGVSGVIKMMALRHGLVPRTLHVDEPSRHVDWSTGAVELVTENQ
 PWPETGRPRRAGVSSFGISGTNAHVILESAPSAQVVENTVVESAPEWVPLVVSARTQSALADYEDRLRAYLAGSPGV
 5 DLRAVASTLAVTRSVFEHRAVLVGDDTVTGSVSDPRVVVFVFPQGQSQRAGMGEEELAAAFPLFAQIHQQVWDLLDVP
 DLEVNETGYAQPALFALQVALFGLLESWGVRPDAVIGHSVGELAAGYVCGVWSLEDAC TMVSARARLMQALPAGGVM
 VAVPVSEDEARAVLGEGVEIAAVNGPLSVILSGDEAAVLRAAATLGKWTRLATSHAFHSARMEPMEKFRVAEGLT
 YRTPRLTMAAGDQVATAEYWVRQVRDVVRFGQVASFEDAVFLELGADRSLARLVDGIAMLHGDHEAQAAI SALAHL
 YVNGMAVDWPAVLGDVRGRVLDLPTYAFEHQRYWLEGWLAVLAPAEREKALLKLVRDSAALVLGHADASTIPVAAA
 10 KDLGIDSLTAVELRNSLAKATGLRLPNTTVFDYPTPAILAARL

SEQ ID NO: 32

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISGFPTDRGWD AESLYDADPDAPGKSYCVEGGFLDNASSFDAGFF
 GISPREALAMDPQQLVLEVSWEAFERAGIEPGSIRGTD TGVMGAYAGGYGAGADLGGFAATASATSVLSGRVSYF
 15 FGLEGPAITVDTACSSSLVALHQAGYALRQGECSLALVGGATVMATPQS FVEFSRQRGLASDGRCKAFADAADGTGW
 AEGVGVLVLLVERLSDARRNGHQVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIIQAALSNAGLAAHEVDVVEAHGTGTT
 LGDPIEAQAVLATYQDRERPLLLGSLKSNIGHTQAASGVSGVIKMMALQHHTVPRTLHVNEPSRHVDWSAGAVEL
 VRENQSWPEGDRPRRAGVSSFGVSGTNAHI ILESAPAQSAEEVQPV EVPVVASDVLPLVVS AKTHSALTEAEDRLRA
 YLTASPEADMPAVASTLAVTRSVFEHRAVLVGDDTVTGTGTAMSDPRVVVFVFPQGWQWLGMSALRESSVVF AERM
 20 AECAAALSDFVDWDLFTVLDDPAVVDRVDVVPASWAVMVSLAAVWQAAGVRPDAVIGHSQGEIAAACIAGAVSLRD
 AARIVTLRSQAIARGPAGRGAMASIALPAQEIELADGAWIAAHNGPASTVIAGTPEAVDLVLTAEHAQGTRVRRITV
 DYASHTPHVELIRDELLHITAGIGSQVPVPWLSTVDGTWVEGPLDAEYWYRNLREPVG FAPAI SQLQAQGETVFIE
 VSASPVLLQAMDDDAVTATLRDDGDATRILTALAQAYTHGVTVDWPAILGTTTTRALDLPTYAFQHORYWLNRL
 TGRTSVEQHRVMLELVLGEAASVLGHGSPDAIATDTSFKDLGMDSLTAIELRNRLMAETGLQLPATMVFDYPTANAL
 25 ATHL

SEQ ID NO: 33

EPIAIVAMACRVPGGVSSPEGLWRLVESGTDVISGFPTDRGWDVEGLFDPDPDAPGKSYCVQGGFLDTAADFDAPFF
 GISPREALGMDPQQLLLETTWEAIERARIDPKSLRGRDVG VYVGAAQGYGVGVDDQQRDNGITGSSVSLLSGRVSY
 30 ALGLEPGVTVDTACSSSLVALHLASQALRQRECSLALVSGVSVMS SPAMFVEFSRQRGLSSDGRCKSF AASADGTI
 WSEGVLVLLVERLSDARRLGHRFLAVVRGSAVNSDGASNGLTAPNGASQQRVIRQALAGAGLTASD VDVVEAHGTGT
 KLGDP IEAEAILATYGQERSTPAWLGLSKSNIGHTMAASGVLGVIKMVEAMRHGSLPRTLHVDDPSPHVDWTS GSVA
 LLTEHQWPDDAKPRRAGVSSFGLSGTNAHVVLEQYQAPAPSVTPVTPVTPVTPNEPRPLAWVLSAQSPKALRE
 QAGRLYASLAEAPEWNSLDIGYSLATTRSDFAHRAVAVGSGREFLRALSKLADGASWPGLTTATAKARRVAF LFDGQ
 35 GAQRLGMGKELYDSSPVFARAWDTVSAGFDKHLHDHSLTDVYFGE GGGSTTAELVDDTL YAQAGIFAMEVALFGLLEDW
 GVRPDFVAGHSIGEATAAYASGMLSLEHVTTLIVARGRALRATPPGAMVALRAGEEEVRAFLDQTGAALDLAAVNSP
 EAVVVAGEPDAVAGFEAAWAASGREARKLRVRHAFHSRHVEAVLDEFRTTLES LKFSAPALPVVSTVTGQLIEPDEM
 GTPEYWLQRQVRQPVRFQDAVRELAEGVGTVEIGPSGALASAGMECLG GDASFHAVLRPRSPEDVCLMTAIAELYA
 GGT AIDWAKVLSGGRAVDLPVYPFQHQS YWLAPAEP SYADEPRTMLELVHMEVASVLGMTDPGVILDDSSFL ELGFD
 40 SLSAVRLRNRLSKATGLDLPSTLLFEHPTSAELASHLD

SEQ ID NO: 34

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALSGFPTDRGWDVEGLFDPDPDASGKSYCVRGGFLDSVGGFDASFF
 GISPREALAMDPQQRLLEVSWEAFERAGIEPGSVRGSDTGVMGGFPGGYGAGADLEGFGATAGAASVLSGRVSYF
 FGLEGPAITVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPQSFVEFSRQRLASDGRCKAFADAADGTGW
 5 AEGVGVLLVERLSDAQAKGHQVLGVVRGSAVNQDGASNGLSAPNGPSQQRVIRAALSNAGLTTAEVDVVEAHGTGTT
 LGDPIEAQAVLATYQDREQPLLLGSLKSNIGHAQAAAGVSGVIKVMALRHGLVPRTLHVDEPSRHVDWSAGAVEL
 VTENQSWPVTGRPRRAGVSAFGVSGTNAHVILESAPAQASEEAQPVVTPVVTPVASELVPLVVSAKTESALAEVEG
 RLRAYLAVSPGVDLRAVGSTLAVARSVFEHRAVLLGDDTGTGTGTAVSDPRVVFVFPQGQWQWLGMSALRGASV
 VFAERMAECAALGEFVDWDLFAVLDDPAVVDRVDVVQPASWAVMVSLAAVWEAAGVRPDAVVGHSSQGEIAAACVAG
 10 AVSLRDAARIVTLRSQVIAGLAGRGAMASVALPAHEIELVEGAWIAACNGPASTVIAGEPDAVDRVLAVHEARGVRV
 RRITVDYASHTPHVELIRDELLNITAGIGSQAPVVPWLSTVDGTWVEGPLDAEYWYRNLREPVGFDASVAGELRAQGD
 TVFVEVSASPVLLQAMDDDVSVATLRRDDGGAARMLTALAQAQFVEGVTVDWPAVLGNAPGRVLDLPTYAFEHQRYW
 LKSRWLARLAPVEREKALLKVVCDDGAATVLGHADASTIPAAGAFRDLGVDSLTAVELRNRLAKATGLRLPATLVFDY
 PTPTALAARL

15

SEQ ID NO:35

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDVSGFPTDRGWDVEDLFGPAAGDSYRLRGGFLDAAGGFDASFFGIS
 PREALAMDPQQRLVLEVSWEAFERAGIEPGSVRGTDGTGVMGAYPGGYGIGADLGGFGTTAGAVSVLSGRVSYFFGL
 EGPAFTVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPQTFAEFARQGGLAGDGRSKAFADSADGAGFSEG
 20 VGVLLVERLSDARRNGHQVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIQAALNNAGLTTAEVDVVEAHGTGTTLGD
 PIEAQALLAAYGQDRERPLLLGSLKSNLGHQAAAAGVSGVIKVMALRHGLVPRTLHVDEPSRHVDWSEGAVELVTE
 NQSWPDTGRPGRAGVSSFGISGTNAHVILESAPSAQTVENTVVESAPEWVPLVMSARTQSALADYEGRLRAYLAGSP
 GVDLRAVASTLAVTRSVFEHRAVLMGDDTGTGSAVSDPRVVFVFPQGQSQRAGMGEEELAAAFVFAQIHQQVWDLLD
 VPDLDVNETGYAQPALFALQVALFGLLESWGVGPDAVVGHSSVGELAAAYASGVWSLEDACTLV SARARLMQALPAGG
 25 VMVAVPVSEDEARAVLGEGVEIAAVNGPSSVVLSGDEAAVLRAAAGLGKWTRLATSHAFHSARMEPMLLEEFRAVAER
 LTYQTPHLMTMAAGEQVTTDPDYWVRQVRDVRVFGEQVASFEDAVFLELGADRSLARLVDGIAMLHGDHEAQDAISAMA
 HLYVSGVAVDWPAVLGDVGRVLDLPTYAFQHERYWLEGRWLAALAPAEREKALLKLVSDGAATVLGHADASTVPVS
 AVFRDLGVDSLTAVELRNRLAKATGLRLPATLVFDYPTPTALAARL

30 SEQ ID NO: 36

EPLAIVGMACRLPGGVASPEDLWRLVESGTDVSGFPTDRGWDVEDFDSAGKSYRAEGGFLDAAAGFDASFFGISPR
 EALAMDPQQRLLEVSWEAFERAGIEPGSVRGTDGTGVMGAYPGGYGIGADLGGFGATAGATSVLSGRVSYFFGLEG
 PAFTVDTACSSSLVALHQAGYALRQGECSMALVGGATVMATPELFTFESRQGLASDGRCKAFADSADGTGWAEVVG
 VLLVERLSDAQAKGHQVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIQAALSNAGLAAYEVDVVEAHGTGTTLGDPI
 35 EAQAVLATYQDREERPLLLGSLKSNIGHTQAASGVSGVIKVMALRQGLVPRTLHVDEPSRHVDWSAGAVELVRENQ
 SWPDTEGPRRAGVSSFGVSGTNAHVILESAPAQPAEEAQPVVTPVASELVPLVVSAKSQSALTEAEGRLRAYLAAS
 PGVDTRAVGATLAVARSVFEHRAVLLGDDTGTGTGTAMSDPRVVFVFPQGQWQWLGMSALRDSVVF AERMAECA
 ALSDFVDWDLFTVLDDPAVVDRVDVVQPASWAVMVSLAAVWEAAGVRPDAVIGHSSQGEIAAACIAGALSRLDAARIV
 SLRSQVIAGLAGRGAMASIALPAQDVELAEGAWIAAHNGPASTVIAGAPEAVDRVLAVHEARGVRVRRITVDYASHT
 40 PHVELIRDELLHITAGIGSQAPVVPWLSTVDGTWVEGPLDAEYWYRNLREPVGFAPAIRQLQDQGETVFIEVSASP
 VLLQAMDDDVSVATLRRDDGGAARMVTALAQAAYVQGVTVDWPAVLGNVPGRVLDLPTYAFEHQRYWLKSWLAALAPA
 EREKALLKVVCDSAAVVLGHADARSIPAAGAFKDLGVDSLMAVELRNRLVKATGLRLPATLVFDYPTPAALAARL

SEQ ID NO: 37

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDVAVSGFPTDRGWDLEDLDFDPDPEAAGKSYCVQGGFLDAAAAGFDAGFF
 GISPREALAMDPQQRLLEVSWEAFERAGIEPGSVRGSDTGVI GAFPVGYGVGF DREGYGATSGPSVLSGRVSYFF
 5 GLEGPATMDTACSSSLVALHLAAQALRNGECSMALAGGVTVMATPEVFTFARQRLASDGRCKAFADSADGAGFS
 EGAGLLLVERLS DARRNGHQVLAVVRGS AVNQDGASNGLTAPNGPSQQRV IQAALINAGLT TAEVDVVEAHGTGTTL
 GDPIEAQAVLATYGGQGRERPLLLGSLKSNIGHTQAASGVSGVIKVMALQ RGLVPRTLHVDEPSRHVDWSAGAVELV
 RENQSWPDSEGP RRAGVSSFGVSGTNAHVILESAPAQPAEEAQP VVTPVVASELVPLVVS AKTESALTEVEGRLRVY
 LAASPGVDTRAVASTLAVTRSVFEHRAVLLGDDTVTGTGTAVSDPRVVFVFPQGQWQWLGMSALRDSSVVF AERMA
 10 ECAAALSEFVDWDLFAVLDDPAVVDRVDVVQPASWAVMVSLAAVWQAAGVRPDAVIGH SQGEIAAACVAGVLSLRDA
 ARIVTLRSQAIAGLAGRGAMASIALPAQDVELVEGAWVAAHNGPASTVIAGAPEAVDRVLAVHEARGVRVRRIAVDY
 ASHTPHVELIRDELDDITAGIGSQAPVVPWLSTVDGTWVEGPLDAEYWYRNLREPVGFAPAVSQLQVQGETVFVEVS
 ASPVLLQAMDDDVSVATLRRDDGGAARMLTALAQA YTTQGVAVDWP AVLGTTTAQVLDLPTYAFQHRRYWVEWLAAL
 APEEREKALLRVVCDGAATVLGHADVGSIPVTA AFKDLGVDSL TAVELRNRLAKATGLRLPATLAFDYPTPTALAAR
 15 L

SEQ ID NO: 38

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISGFPTDRGWDVEHLYDPDPDAPGKAYCVQGGFLDSAGGF DASFF
 GISPREALAMDPQQRLVLEASWEAFERAGIEPGSLRGTD TGVMGAYPGGYGIGADLGGFGATAGAVSVLSGRVSYF
 20 FGLEGPAVTVD TACSSSLVALHQAGYALRQGECSLALVGGVTVMATPQSFVEFSRQRGLAGDGRCKAFADAADGTGW
 AEGVGVL LVERLSDAQAKGHQVLAVVRGS AVNQDGASNGLSAPNGPSQQRVIRQALANAGLT TAEVDVVEAHGTGTT
 LGDPIEAQALLATYQDRERPLLLGSVKS NLGHTQAAAGVSGVIKVMALRHGLVPRTLHVDEPSRHVDWSEGA VEL
 VTENQPWP DADRPRRAGVSSFGISGTNAHVILESAPSTQAVDDVRPVEAPVVASEWVPLVVSARTLPALVEYEGRLR
 AYLAGSPGVD MRAVGSTLAVTRSVFEHRAVLMGDDTVTGS AVSGPRVVFVFPQGGSQRAGMGEELAAAFPVFARIHQ
 25 QVWDL LDVPDLEVNETGYAQPALFALQVALFGLLESWGVGPDAVIGH SVGELAAGYVSGLWSLED ACTLVSARARLM
 QALPPGGVMVAVPVSEEEAKAVLCEGVEIAAVNGPSSVVLSGDETAVLR AAAALGKSTRLATSHAFHSARMEPMLDE
 FRAVAERLTYQTPRLPMAAGEQVTT PDYWVRQVREPVRFG EQAASCDAVFVELGADRSLARLVDGVAMLHG DHEAQ
 AAISALAHLYVNGVTVDWPAVLGDVPGRVLDLPTYAFQHQR YWLEGWLAALAPEERAKALLKVVCDTAATVLGHADA
 RTIPMTGAFRDLGIDSLTAVELRNGLAKATGLRLPATLVFDYPTPTVLAARL
 30

SEQ ID NO: 39

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISGFPTDRGWD AESLYDPDPDAPGKSYC VEGGFLDNAASF DAGFF
 GISPREALAMDPQQRLVLEVSWEAFERAGIEPGSVRGTD TGVMGAYAGGYGAGADLGGFAATASATSVLSGRVSYF
 FGLEGPAFTVD TACSSSLVALHQAGYALRQGECSLALVGGATVMATPQSFVEFSRQRGLASDGRCKAFADAADGTGW
 35 AEGVGVL LVERLS DARRNGHQVLAVVRSS AVNQDGASNGLSAPNGPSQQGVIRQALANAGLT PAEVDVVEAHGTGTT
 LGDPIEAQAVLATYQDRERPLLLGSLKSNIGHTQAASGVSGVIKVMALQHHTVPRTLHVNEPSRHVDWSAGAVQL
 VRENQSWPEGDRPRRAGVSSFGVSGTNAHI ILESAPAQSAEEVQPVEVPVVASDVLPLVVS AKTHSALTEAEDRLRA
 YLTASPEADMPAVASTLAVTRSVFEHRAVLLGDDTVTGTGTAVSDPRVVFVFPQGQWQWLGMSALRDSSVVF AERM
 AECAAALSDFVDWDLFAVLDDPAVVDRVDVVQPASWAVMVSLAAVWQAAGVRPDAVIGH SQGEIAAACIAGALS LR D
 40 AARIVTLRSQVIAGLAGRGAMASIALPAQEVELAEGAWIAAHNGPASTVIAGTPEAVDLVLTAEHAQGTRVRRIAVD
 YASHTPHVELIRDELDDITAGIGSQAPVVPWLSTVDGTWVEGPLDAEYWYRNLREPVGFAPAVRQLQDQGETVFIEV
 SASPVLLQAMGDDAVTVATLRCDDGGAARMLTALAQA YTTQGVAVDWP AVLGTTTARVLDLPTYAFQQR YWVEWLAG

LAPEERAKALLKVVCDTAATVLGHADARTIPLTGAFKDLGVDSLTAVELRNSLTAKATGLRLPATLVFDYPTPTALAV
RL

SEQ ID NO: 40

5 EPLAIVGMACRLPGGVSSPEDLWRLVESGTDVSGFPTDRGWDLEDLDFDPDPEAAGKSYCAEGGFLDAAAGFDAGFF
GISPREALAMDPQQRLLEVSWEAFERAGIEPGSVRGSDTGTFVIGAFPVGYAAGAAREGYGATAAPNVLSGRLSYFF
GLEGPAITMDTACSSSLVALHLAAQAVRNGECMALAGGVTVMATPEVFTTEFARQRLASDGRCKAFADSADGAGFS
EGAGLLLVERLSDARRNGHQVLAVVRGSAVNQDGASNGFTAPNGPSQQRVIQQALANAGLTAEVDVVEAHGTGTTL
GDPIEAQAVLATYGDREQPLLLGTLKSNIGHTQAAAGVSGVIKMMALQHDTPRTLHVNEPSRHVDWTAGAVELV
10 TENQSWPVTDPRRAGVSAFGVSGTNAHVILESAAPSVNNAQPVETPVVASELVPLVISAKTLPALTEHEDRLRAY
LAASPEADMPAVASTLAVTRSVFEHRAVLLGDDTVTGTGAAVSDPRVVFVFPQGQWQWLMGSGSLRGSSVVFARMA
ECAAALREFVDWDLFAVLDDPAVVDRVDVVQPASWAVMVSLAAVWEAAGVRPDAVVGHSQGEIAAACVAGAVSLRDA
ARIVTLRSQVIAGLAGRGGMASVALPAHEIELVEGAWIAARNGPAAATVIAGEPDAVDRVLAIHEAQGVRVRRIAVDY
ASHTPHVELIHDELLGVIAGVDSQAPVVPWLSTVDGTWVEGPLDAEYWYRNLREQVGFDPVAVSQLRAEGDTVFEVS
15 ASPVLLQAMDDDAATVATLRRDDGDAARMLTALAQAQFVEGVTVDWPAILGTATPGVLDLPTYAFQHQRFWAERWLAR
LAPVEREKALLKVVCDGAATVLGHADASTIPATAAFKDLGIDSLTAVELRNLAKATGLRLPATLVFDYPTPTALAA
RL

SEQ ID NO: 41

20 EPLAIVGMACRLPGGVSSPEDLWRLVESGTDVSGFPTDRGWDVGDVDFGPAAGDSYRLRGGFDAAGGFDASFFGIS
PREALAMDPQQRLVLEVSWEAFERAGIEPGSVRGTDGTGVMGAYPPGGYGIGADLGGFGATASATSVLSGRVSYFFGL
EGPAITVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPQTFVEFARQGGLAGDGRSKAFADSADGAGFSEG
VGVLLVERLSDAQAKGHQVLAMLRSSAVNQDGASNGLTAPNGPSQQRVIAAALSNAGLAPHEVDVVEAHGTGTTLGD
PIEAQALLATYGDREPRLLLSVKSNLGHTQAAAGVSGVIKMMALRNLVPRTLHVDEPSRHVDWSVGAVELVTE
25 NQSWPDSGRPRRAGVSSFGISGTNAHVILESEPPAQVVENTVVEPAPEWVPLVMSARTQSALADYEDRLRAYLAGSP
GVDLRAVGSTLAVTRSVFEHRAVLLGDDTVTGTAVSDPRVVFVFPQGQSQRAGMGEEELAAAFPVFARIHQQVWDLDD
VPDLEVNETGYAQPALFALQVALFGLLESWGVGPDAVIGHSVGELAAGYVSGLWSLEDAC TLVSARARLMQALPAGG
VMVAVPVSEEEAEAVLCEGVEIAAVNGPSSVVLSGDEAAVLRAAATLGKWTRLATSHAFHSARMEPMLLEEFRAVAEG
LTYRTPRLTMAAGDQIATAEYWVRQVRDVRVFGEQAASCGDAVFVELGADRSLARLVDGVAMLHGDHEAQAAI SALA
30 HLYVSGVAVDWPAVLGDVPGRVLDLPTYAFQHQRWLEGRWLAALTPEERAKALVKVVCDSAATVLGHADASTIPVT
AAFRDLGVDSLTAVELRNSLTAKATGLRLPATLVFDYPTAGALAARL

SEQ ID NO: 42

EPLAIVGMACRLPGGVFSPEDLWRLVESGTDVSGFPTDRGWDVDAENLDFDPDPAAGKSYCLEGGFLETAANFDASFF
35 EISPREALAMDPQQRLVLEASWEAFERAGIEPGSVRGSDTGTFVGMGAFPPGGYGIGADLEGYGATSGNLVLSGRLSYFF
GLEGPAVTVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPHTFVEFSRQRLASDGRCKAFADSADGTGWS
EGVGVLLVERLSDAQAKGHQVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANAGLTIAEVDVVEAHGTGTTL
GDPIEAQALLATYGDREQPLLLGSVKSNGHTQAAAGVSGVIKMMALRNLVPRTLHVDEPSRHVDWSEGA VELV
TENQPWPETGRPRRAGVSSFGVSGTNAHVILESAAPQVVDNTVVESAPEWVPLVMSARTQSALADYEDRLRAYLAG
40 SPGVDLRAVASTLAVTRSVFEHRAVLMGDDTVTGTAVSDPRVVFVFPQGQSQRAGMGEEELAAAFPVFARIHQQVWDL
LDVPDLEVNETGYAQPALFALQVALFGLLESWGVRPDAVVGHSAGELAAAYVSGVWSLEDACALVSARARLMQALPA
GGVMVAVPVSEEEAEAVLCEGVEIAAVNGPSSVVLSGDEAAVLRAAAGLGKWTRLATSHAFHSARMEPMLLEEFRAVA

EGLTYRTPRLTMAAGDQVATAEYWVRQVRDVVRFGQVASFEDAVFLELGADRSLARLVDGVAMLHGDHEAQAAISA
LAHLYVNGVTIDWPAVLGGVPGRVLDLPTYAFQHERYWAEAWLAALAPAEREKALLKLVS DGAATVLGHADASTIPV
TAAFKDLGIDSLTAVELRNSLAKATGLRLPATLVFDYPTPTALAARLD

5 SEQ ID NO: 43

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISGFPTDRGWDVENLYDPDPDAPGKSYSVRGGFLDAAANFDASFF
GISPREALAMDPQQRMLLEVSWEA FERAGIEPRSVRGS DTGVFIGAYPGGYGIGVDFEGFGATAGAASVLSGRVSYF
FGLEGPAFTVDTACSSSLVALHQAGYALRQGD CSLALVGGVTVMATPQTFVEF'SRQRGLSADGRCKAFADSADGTGW
AEGVGVLLVERLSDAQAKGHQVLGVVRGSAVNQDGASNGLSAPNGPSQQRVIRAALSNAGLAPHEVDVVEAHGTGTT
10 LGDPIEAQALLATYQGGRGEPLLLGSLKSNIGHTQAAAGVSGVIKVMALQYGLVPRTLHVDEPSRHVDWTAGAVEL
VGENQPWPETGRPHRAGVSSFGISGTNAHVILESAPAQPAEEAQPVVTPVVASELVPLVVS AKTESALTEVEGRLRA
YLAASPGVDTRAVASTLAVTRSVFEHRAVLLGDDTVTGTGTAMSDPRVVVFVFPQGQWQLMG SALRDSSVVF AERM
AECAAALSDFVDWDLFTVLDDPAVVDRVDVVQPASWAVMVSLAAVWQAAGVRPDAVIGHSQGEIAAACIAGAVSLRD
AARIVTLRSQAIAGLAGRGAMASIALPAQEIGLADGAWIAAHNGPASTVIAGAPEAVDRVLTAEHAQGARVRRIAVD
15 YASHTPHVELIRDELDDITAGIGSQAPVVPWLSTVDGTWVEGPLDAEYWYRNLREPVGFAPAVRQLQAQGETVFVEV
SASPVLLQAMDDDAVTATLRRDDGDATRLMTALAQA YTHGVTVDWPAILGTTTTRALDLPTYAFQHERYWAEAWLV
GLAPEERAKALLKLVS DSAAAVLGHADARGIPATGAFKDLGVDSL TAVELRNTLT KATGLRLPATMVFDYPTPADLA
ARL

20 SEQ ID NO: 44

EPLAIVGMACRLPGGVSSPEELWQLVESGGDAISPFP TDRGWDLETPYRGGFLTD PAGFDAGFFGISPREAVAMDPQ
QRVLLEASWEA FERAGIKPDSLGRS DTGVFVGGFSQGYGTGADLGGFGATSTQTSVLSGRLSYFYGFEGPAVTVDTA
CSSLVALHQAASSLHIGEC SLAVVGGVTVVATPGGFVEFARQGG LALDGRCKAFADAADGIGLAEGVGVLLVERLS
DAQRNGHTVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANARLTPADIDVVEAHGTGTRLGDPIEAQAILAT
25 YGQDRDQPLLLGSLKSNIGHTQAAAGVAGVIKTVMAMRHGTAPRTLHADEPSRHVDWSAGAVELLSENRLWPETDRP
RRAAVSSFGVSGTNAHVILESAPAESAEGPAGMGSESMGSESGPVVVVLSAKSASALAGQEERL RAYLASGADVRVAV
AAGLARRSVFEHRSVILGDSTVSGVAAGVPRVVFLFPQGQTQWAGMGADLLESSPVFAARMRQCAAELSKYTDWDLF
TALSDPALLDRVDVVQPVSWALMVSLAALWQHCGVQPD AVIGHSQGEIAAACVAGALTQD GARLITGRSALIAHLS
GRGTMASIALPADDLTLPDDVCIAAVNGPATTIIAGPTPAIEHLLATYEASNIHTRRIPVDYPSHTPHVEDLHDPLL
30 AITHTLTPHTPTTPWLSTVDNTWIHTPPHPDYWRNL RHPVQLAPAITTLTHPHPTH LIEISTHPVLLPAIDTTTTL
TTTATLRRNHGTPHQLLTSLAHAHTHGATINWPALLGNPPTATTADLPTYPFQHKRYWLQDTERVAGLPAAEREQVV
VKAVCETA AVVLGH AHADDILATTLFKDLGVDSL IAVELRNRLAADAGLRLPATLVFDYPTPHALATWL

SEQ ID NO: 45

EPLAIVGMACRFPGGVSSPEDLWRLVESGGDAISDIPADRGWDLETPYRGGFLADAGGFDAGFFGISPREALAMDPQ
QRVLLET SWEALERA EIEPGSLRGSDTG VFIGGFSQSYGIGADLGGFGTTGTIQT SVLSGRLSYFFGFEGPAFTVDTA
CSSLVALHQASSALRQGECSLALVGGVTVLADPSGFVEFARQGG LADGRCKAFADTADGTSLAEGVGVLLVERLS
DAQRNGHTVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANARLTPADIDVVEAHGTGTRLGDPIEAQAILAT
YGQDRDQPV LIGSLKSNIGHTQAAAGVAGVIKVMAMRQGTVPRTLHVDEPSHHVDWTAGSVQPI TQNQEWPQAGRV
40 RRAGVSSFGISGTNAHVIIIEGVPVAEPVVVADSGVVPLVLSARTPGALLEQEERL RAYLACGADVRAVAAGLARRSV
FEHRSVLVGDTVVS GTAADARLVLVFSGQGSQRAGMGEELAARFPVFAEIHQRVWDLLDVGPGLDVDDTGYAQPALF
ALQVALFGLLESWGV RPDVLIGHSIGELAAACVSGVWSLQD ACALVSARARLMQALPAGGVMAAVPVSEAEAEAVLR

EGVEIAAVNGPASIVLSGDEDAVLQAAASLGRFTRLSTSHAFHCARMDPMLDEFQRQVAESIAYQPPRIAMAAGDQVI
 TPDYVWRQVREPVRFGDQVAHADAVFLEIGPDRHLARLIDGIPTLSVDEVQSAMTALGELHVRGIDVDWATLLGTT
 PATPTDIPTYPFQHKHYWIDNRRISGLEPAERGQALLEIVREAAAVVLGHTDAREIAPTTAFRDLGIDSLTAIELRN
 RVATETGLRLPATLVFDHPTPTTLATWI

5

SEQ ID NO: 46

EPLAIVGMACRLPGGISSPEDLWQLVQSGGDAITDLPTDRGWDLTHLYDNDAPPVYRGGFLTDAGDFDAAFFGISPR
 EALAMDPQQRILLETSSWEAFERGGINPEAIRGSNTGVFIGGFSYGYGTGADLGGFGATSTQTSVLSGRLSYFYGFEG
 PAVTVDTACSSSLVALHQASSALRQGECSLALAGGVTVMATPAGFEEFARQGGLAADGRCKAFSDTADGTGWAEVGV
 10 VLLVERLSDAQNRNGHTVLAVVRGSANQDGASNGLSAPNGPSQQRVIRQALANARLTPADIDVVEAHGTGTRLGDP
 EAQAILATYGQDRDQPLLLGSLKSNIGHTQAAAGVAGIIKTMAMRHGTAPRTLHADEPSRHVDWSAGAVELLSEN
 LWPETDRPRRAAVSSFGVSGTNAHVIIIEQPPHTPAPEAERTTGLDVVPWLLSARTPAGLRAQAEQVSSLNEDFANIG
 FSLATTRTPMEHRAVVVADVSGISEAAVFAGGGSPDTPVSSGLANVRGKTVFVFPQGAQWAGMGAELEFATSPVFAER
 MTECAAAFAALVDWSLIDVLQOREGAPSPDRVDVQPLSFVAVMVSLAALWKSHGVVPDAVTGHSQGEVAAACVSGAL
 15 SLSDAATVVALRSRVIAQLAGHGMVALPATEFAAEYWAGRLELAAVNGPASVVVAGEPEALEELLAENPNARRIPV
 DYASHTSRVERIREELTGLLSGLAPRQPIVPFYSTVDNQWLDKPLDAEYWRNLRLQTVRFADAVHGLADAGFRAFVE
 VSPHPVLTSSMRDILDERETTAVVTGTLRRDAHGVREFVRSLARLWVSGFSVDWSGLFGNGPRRIPLPTYPFQRNRY
 WLQAELDLVRTHAAAVLGHAGPEAVAADHPFRDLGVDSLIAVELRNRLAAETGLRLPATLVFDYPTPRALAAWLD

20 SEQ ID NO: 47

EPLAIVGMACRFPGGVSSPEDLWRLVETSGDAISDIPADRGWDLETPYRGGFLIGAAGFDAGFFGISPREALAMDPQ
 QRLLEISWEALERAGINPESVRGSDTGTVFVGSSYGYGVGADLGGFGATSTHISVLSGRVSYFFGFEGPAFTVDTA
 CSSSLVALHQASSALRQGECSLALVGGVTVMATPAGFEEFARQGGLAADGRCKAFSDTADGTSLAEGVGVLLVERLS
 DAQRNGHTVLAVVRGSANQDGASNGLSAPNGPSQQRVIRQALANARLTPADIDVVEAHGTGTTLGDP
 25 YGQDRDQPVLLIGSLKSNIGHTQAAAGVAGVIKVMAMRQGTVPRTLHVDEPSHHVDWTAGSVQPIQONQEWPPQAGRV
 RRAGVSSFGISGTNAHVIIIEGVPVAEPVVVADSGVPLVLSARTPGALLEQEERLRLAYLACGADVRAVAGLARRSV
 FEHRSVLVGDTVVSCTAARLVLVFSGQGSQRAGMGEEAARFPVFAEIHQRVWDLDDVGPGLDVEDDTGYAQPALF
 ALQVALFGLLESWGVRPDVLIGHSIGELAAACVSGVWSLQDACALVSARARLMQALPAGGVMAAVPVSEAEAEAVLR
 EGVEIAAVNGPASIVLSGDEDAVLQAAASLGRFTRLSTSHAFHCARMDPMLDEFQRQVAESIAYQPPRIAMAAGDQVI
 30 TPDYVWRQVREPVRFGDQVAHADAVFLEIGPDRHLARLIDGIPTLSVDEVQSAMTALGELHVRGIDVDWATLLGTT
 PATPTDIPTYPFQHKHYWIDNTRISGLEPAERGQALLEIVREAAAVVLGHTDAREIAPTTAFRDLGIDSLTAIELRN
 RVATETGLRLPATLVFDHPTPTTLATWI

SEQ ID NO: 48

EPLAIVGMACRLPGGISSPEDLWQLVQSGGDAITDLPTDRGWDLETPYRGGFLTDPAGFDAGFFGISPREALAMDPQ
 QRVLLASWEAFERAGIKPDSLGRSDTGTVFVGGSQGYGTGADLGGFGATSTQTSVLSGRLSYFYGFEGPAFTVDTA
 CSSSLVALHQASSALRQGECSLALAGGVTVMATPAGFEEFARQGGLAADGRCKAFADTADGTGWAEVGVLLVERLS
 DAQRNGHTVLAVVRGSANQDGASNGLSAPNGPSQQRVIRQALANARLTPADIDVVEAHGTGTRLGDP
 35 YGQDRDQPLLLGSLKSNIGHTQAAAGVAGIIKTMAMRHGTAPRTLHADEPSRHVDWSAGAVELLSENRLWPETDRP
 RRAAVSSFGVSGTNAHVILESAPEAEAGPAGMGSESMGSESGPVVVVLSAKSASALAGQEERLRLAYLASGADVRAV
 AAGLARRSVFEHRSVILGDSTVSGVAAGVPRVVFLFPQGQTQWAGMGADLLESSPVFAARMRQCAAELSKYTDWDLF
 40 TALSDPALLDRVDVVPVSWALMVSLAALWQHCGVQPDVIGHSQGEIAAACVAGALTLQD GARLITGRSALIAHLS

GRGTMASIALPADDLTLPDDVCIAAVNGPATTIIAGPTPAIEHLLATYEASNIHTRRIPVDYPSHTPHVEDLHDPLL
AITHTLTPHTPTTPWLSTVDNTWIHTPPHPDYWYRNLRHPVQLAPAITTLTHPHPTHLEISTHPVLLPAIDTTTTL
TTTATLRRNHGTPHQLLTS LAHAH THGATINWPALLGNPPTATTADLPTYPFQHKRYWLQDTERLTTQSSVEQHRLM
LDLVTSHAAAVLGHSSAAAITTDTPFRLDGLFDSL TAVELRNRVAADTGLRLPATLVFNHPNADALTQYL

5

SEQ ID NO: 49

DPIAIVAMACRLPGGVSSPEDLWRLVETGTDAIGPFPTDRGWDTELYPVPDAPGKTYCVEGGFLTGAEEFDAAFFDI
SPREALAMDPQQRLLLETSWEAVERARINPKSLCGKDVGVYVGAQAQGYGLGAGDQTEGTAITGGSTSLSSGRVSYA
LGLEGPAVTVD TACSSSLVAMHLAQALRQGECSLALVGGVSVMASPALFVEFSRQRLAADGRCKSFSDAADGTNW
10 AEGVGVLILERLSDAQRNGHPVLAVIRGSAINSDGASNGLTAPNGLSQQRVIRQALTAAGLRPEDVD AVEAHGTGR
LGDPVEAEAILATYQONREQPLLLGSLKSNIGHAAAAAGSVAGVIKVMQAMRNGVLPRTLHIDE PSSQVDWTS GVAL
LTESRPWPDEDKPRRAGVSSFGISGTNAHIVLEQYRAAEPEDRPGDGPGERRPVAVVLSGKSPA AVRAQAGRLRAHL
VGTQGW RPVDVGYALATTRADF AHRVAVGSGPEFLHALEKLAEGASWPRLTTNRASARRVAF LFDGQGTQRLGMGR
ELHQRFPAFAEAWDTVDAEFAPYLD RSLTEVFFSDGGSGLMDDTLYAQAGLFAVETALFRLLAGWGV RPDFVAGHSA
15 GEITAAHVAGVLSVTD AVRLIVARGQALRLAPPGAMASVRSSAQEVRDFIAQSGLPVDLAA INSPG SVVAGSPETI
AEFEGAWTASGRQAKRLAVRHAFHSRHVDGVLDEFRAALGGCRFGVAELPLVSTATGELASPD ELGTPEHWRHARQ
TVRFQDAIRALTEQGVDTFVEIGPSGTLASAGMECGGGTAAFHAVMRARQPEEVSLMTAVAE LYAGGTPVEWSRVLD
GRSVVDLPVYPFQRQPYWLAPADELSQPEQQKALLELVKAEAAVLLGITDATAIEDDARFLELGFDSL SATRLRNQL
AKATGLALEQTLLFDFPTPAALAAHL

20

SEQ ID NO: 50

EPLAIVGMACRLPGGVSSPEDLWRLVESGGDVISDFPTDRGWDTTGEDSSFIRGGFLT D AGGF DAGFFGISPREAVA
MDPQQRLVLETSWEVLERAGIEPGSLRGSDTG VFI GGF SQGYGAGADLGGFGATGTQTSVLSGRVSY YLGLEGPAVT
VDTACSSSLVALHQ AASALRQGECSLALVGGVTVMATTHSFVEFARQGG LSSDGRCRSFADSADGTGWAEGVGVLLV
25 ERLSDARRSGHPVLALVRGS AVNQDGASNGLSAPNGLSQQRVIRQALATAGLDAADVDVVEAHGTGTVLGDPIEAQA
ILATYQGQREEPLLLGSLKSNVGH TQAAAGVAGVIKVMAMRQGTVPRTLHVDEP SHHVDWTAGRVELLTENRPWPQ
AGRVR RAGVSSFGISGTNAHVII EGV PVAEPVVVADSGVVPLVLSARTPGALLEQEERL RAYLACGADVRAVAAGLA
RRSVFEHRSVLVGD TVVSGTAADARLVLVFSGQGSQRAGMGEELARFPVFAEIHQRVWDLLDVGPGLDVDDTGYAQ
PALFALQVALFGLLESWGVRPDVLIGH SIGELAAACVSGVWSLQDACALVSARARLMQALPAGGVMAAVPVSEAEAE
30 AVLREGVEIAAVNGPASIVLSGDEDAVLQAAASLGRFTRLSTSHAFHCARMDPMLDEF RQVAESIAYQPPIAMAAG
DQVITPDYWVRQVREPVRFGDQVA AHADAVFLEIGPDRHLARLIDGIPTLSVDEVQSAMTALGELHVRGIDVDWATL
LGTTPATPTDIPTYPFQHKHYWIDNTRISGLEPAERGQALLEIVREAAAVVLGHTDAREIAPTTAFRDLGIDSLTAI
ELRNRVATETGLRLPATLVFDHPTPTTLATWI

35 SEQ ID NO: 51

EPLAIVGMACRLPGGISSPEDLWQLVQSGGDAITDLPTDRGWDLET PYRGGFLTDPAGFDAGFFGISPREALAMDPQ
QRVLLEASWEAFERAGIKPDSL RGS DTGVFVGGFSQGYGTGADLGGFGATSTQTSVLSGRLSYFYGFEGPAVTVDTA
CSSSLVALHQASSALRQGECSLALAGGV T MATPAGFEEFARQGG LAADGRCKAFADTADGTGWAEGVGVLLVERLS
DAQRNGHTVLAVVRGS AVNQDGASNGLSAPNGPSQQRVIRQALANARLTPADIDVVEAHGTGTRLGDPIEAQA I LAT
40 YGQDRDQPLLLGSLKSNIGHTQAAAGVAGIIKTMAMRHGTAPRTLHADEPSRHVDWSAGAVELLSENRLWPETDRP
RRAAVSSFGVSGTNAHVILESAPAESAEGPAGMGSESMGSESGPVVVVLSAKSASALAGQEERLRAYLASGADVRAV
AAGLARRSVFEHRSVILGDSTVSGVAAGVPRVVFLFPQGQTQWAGMGADLLESSPVFAARMRQCAAELSKYTDWDLF

TALSDPALLDRVDVVQPVSWALMVSLAALWQHCGVQPDVIGHGSQGEIAAACVAGALTLDQD GARLITGRSALIAHLS
GRGTMASIALPADDLTLPDDVCIAAVNGPATTIIAGPTPAIEHLLATYEASNIHTRRIPVDYPSHTPHVEDLHDPLL
AITHTLTPHTPTTPWLSTVDNTWIHTPPHPDYWYRNLRHPVQLAPAITTLTHPHPTHLEISTHPVLLPAIDTTTLT
TTTATLRRNHGTPHQLLTSLAHAHHTHGATINWPALLGNPPTATTADLPTYPFQRRRFWAERISGLEPAERGOALLEI
5 VREAAAVVLGHTDAREIAPTTAFRDLGIDSLTAIELRNRVATETGLRLPATLVFDHPTPTTLATWI

SEQ ID NO: 52

EPLAIVGMACRLPGGISSPEDLWQLVQSGGDAISDFPTDRGWDLTHLYDNDAPPVYRGGFLTDAGDFDAAFFGISPR
EALAMDPQQRILETSWEVLERAGIEPGTLRGSETGVFVGGFTQGYGTGADLGGFGMTSGHSSVLSGRVSYFFGFEG
10 PAVTVDTACSSSLVALHQASSALRQGECSLALVGGVTVMASPGGFTEFSRQGGGLSPDGRCKAFADAADGTGWAEGVG
VLLVERLSDAQRNGHTVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANARLTPADIDVVEAHGTGTTLGDPI
EAQAILATYQGDRDQPLLLGSLKSNIGHTQAAAGVAGVIKMMAMRHGTAPRTLHIDEPSRHIDWTTGSVALSTENQ
PWPETGHPRRAGVSFAFVSGTNAHVVLGVPVAGPPEEDVEPGVPLLISAKSRPALMEQEQRRLRTYLDGSQTDIRA
VAATLAHARSVFEHRSVLVGDTVVSAGTAADARLVLFVSGQGSQRAGMGEELAAARFPVFAEIHQRVWDLDDVGPGLDV
15 DDTGYAQPALFALQVALFGLLESWGVRPDVLIGHSIGELAAACVSGVWSLQDACALVSARARLMQALPAGGVMAAVP
VSEAEAEAVLREGVEIAAVNGPASIVLSGDEDAVLQAAASLGRFTRLSTSHAFHCARMDPMLDEFQRVAESIAYQPP
RIAMAAGDQVITPDYWVRQVREPVRFGDQVAHADAVFLEIGPDRTLARLIDGVPLLSKEDEVQAALVALAELHVRG
VPLEWSTVIGGMTSIVDLPTYPFRRKRYWIESAERLTTQSSVEQHRLMLDLVTSHAAAVLGHSSAAAITTDTPFRLD
GFDSLTAVELRNRVAADTGLRLPATLVFNHPNAGDLARHL

SEQ ID NO: 53

EPLAIVGMACRLPGGISSPEDLWQLVQSGGDAITDLPTDRGWDLTHLYDNDAPPVYRGGFLTDAGDFDAAFFGISPR
EALAMDPQQRILLETSWEAFAFERGGINPEAIRGSNTGVFIGGFSYGYGTGADLGGFGATSTQTSVLSGRLSYFYGFEG
PAVTVDTACSSSLVALHQASSALRQGECSLALAGGVTVMATPAGFEEFARQGGLAADGRCKAFADTADGTGWAEGVG
25 VLLVERLSDAQRNGHTVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANARLTPADIDVVEAHGTGTTLGDPI
EAQAILATYQGDRDQPLLLGSLKSNIGHTQAAAGVAGIIKTMAMRHGTAPRTLHADEPSRHVDWSAGAVELLSEN
LWPETDRPRRAAVSSFGVSGTNAHVILESAPAESAEGPAGMGSESMGSESGPVVVVLSAKSASALAGQEERLRAYLA
SGADVRAVAAGLARRSVFEHRSVILGDSTVSGVAAGVPRVVFLEPGQGTQWAGMGADLLESPVFAARMRQCAAELS
KYTDWDLFTALSDPALLDRVDVVQPVSWALMVSLAALWQHCGVQPDVIGHGSQGEIAAACVAGALTLDQD GARLITGR
30 SALIAHLSGRGTMASIALPADDLTLPDDVCIAAVNGPATTIIAGPTPAIEHLLATYEASNIHTRRIPVDYPSHTPHV
EDLHDPLLAITHTLTPHTPTTPWLSTVDNTWIHTPPHPDYWYRNLRHPVQLAPAITTLTHPHPTHLEISTHPVLLP
AIDTTTTLTTTATLRRNHGTPHQLLTSLAHAHHTHGATINWPALLGNPPTATTADLPTYPFQHKRYWLQDTRLALAP
AEREQALVKAVCETAAMVLGHADTREIAATTAFKELGLDSLTAVALRDLAAETGRKLPATLVFDYPSQPALAAWL

SEQ ID NO: 54

EPLAIVGMACRLPGGVASPEDLWRLVESGTDAITDFPTDRGWDLDEVADQSYCLQGGFLDNAAGFDAAFFGISPREA
LAMDPQQRVLVEASWEAFAERAGIKPGSLRGSDTGVMGAYPGGYGTGADLGGFGATAGAVSVLSGRISYFFGFEGPA
MTVDTACSSSLVALHQAGYALRQGECSIALVGGVTVMATPQSFIEFSRQRGLAADGRCKTFADAADGTGWAEGVGVL
LVERLS DARAKGHQILAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALVNAGLSPADVDVVEAHGTGTTLGDPIEA
40 QALLTTYGQGRSVPLLLGSVKSNLGHTQAAAGVTGVIKMMALRHGVVPRTLHVDEPSRHVDWSAGAVELVTSNREW
PVVDRPGRAGVSSFGISGTNAHVILEAVPSDTPASTSTDAVLPLVVSARTAPAAEDLTARLRLAYLSAAPETDQRAAA
ATLALTRSVFEHRAVVLGDELVSGQAVRDPRVVVFVSGQGSQRAGMGEQLAAVFPVFAEIHERVWALLDVPDGLDVD

DTGHAQPALFALQVALSGLLESWGVRPAAVIGHSIGELAAAYVSGVWSLEDACALVSARARLMQALPPGGVMVAVPV
 PEAEARAVLRDGEVIAAVNGPSSVVLSGDEDAVLQAVSGFAKWTRLKTSFAFHSAHMDPMLDEFRAVAERLTYRRPS
 VEMAAGDRVTTAEYWVRQVREAVRFGDQTTAYEDAVFVEIGPGRTLARLIDGITMLHGDTEREAALTGLSQLFVRGV
 DVDWATVIEDTTARILDPTYAFQHENYWLHWLSGLTPAEREQALLTAVRENAAVLGHADARTVPVNSAFRDLGFD
 5 SLTAIELRNSLAKATGLSLPATMAFDYPTPAVLATRL

SEQ ID NO: 55

EPLAIVGMACRLPGGVSSPEELWRLVESGVDAISGFPVDRGWDVENLFDPDPAAGKSYCVQGGFLDSAAEFDAAFF
 GISPREALAMDPQQRVLVLETSWEAFERAGIEPGSIKGS DTGVFMGAYQGGYGSGADLGGFGATAGATSVLSGRVSYF
 10 FGFEGPAMTVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPQSFVEFSRQRGLAVDGRSKAFADAADGTGW
 AEGVGVLLVERLSDAQAKGHQILAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANASLTPADVVDVVEAHGTGTT
 LGDPIEAQAVIATYQDRSTPLLLGSLKSNIGHTQAAAGVSGVIKVMALRHGVVPQTLHVDQPSRHVDWSAGAVEL
 LTSNQWPWSSERARRAGVSAFGVSGTNAHVILESAPAEPVVAEAGPVPVSDVPLVLSAKSAPALRALEQRLRAYD
 GAAGRALATARATFDHRAVLIGDDTVTGVAVPDPRVVFVFPQGQWQLGMGRELRGSSVFAERMAECAALSEFVD
 15 WDLFAALDDPAVVDRVDVVQPVCWAVMVSLAAVWQAAGVNPDAVVGHSSQGEIAAAVVAGSLSLRDGARVVALRSQLI
 KGLAGRGAMASIALPADQIGLVEGAWIAALNGPSSSTVIAGTPEAVEQVLAAQDARVRRIAVDYASHTPQVEAIRDEL
 LETAGVSSQPPTVPWLSTVDNTWVEGPLPADYWFRLREQVGFAQAVVTLGDAVFVEVSGSPVLMQSMGDAVTVAS
 FRRDDGSATRMVTSLAEAYVQGVNVNWA AVL GAGTERALDLPTYPFQRQHYWISLAALPPAERERALLKVVRDSAAV
 VLGHADGRTVPATAAFKDLGLDSL TAVELRNSLRKATGLQLPATLVFDYPSPVALAARL

SEQ ID NO: 56

EPLAIVGMSCLPGLGVSSPEDLWRLVESGVDAISGFPVDRGWDAEGLFDPDPDAAGKTYCVQGGFLEAAGEFDTAFF
 GISPREALAMDPQQRVLLEASWEAFERAGIGADTVRGTD TG VFI GAYPVAYGAGVDREGYGATAAPNVLSGRLSYFF
 GLEGPAITVDTACSSSLVALHLAASALRNGECSLALAGGVTVMATPEVFTTEFARQRGLAFDGRSKSFADAADGAGFS
 25 EGAGLLVLRLSDARRNGHQVLAVIRGSAVNQDGASNGFTAPNGPSQQRVIEAALGNAGLTAEVDVVEAHGTGTKL
 GDPIEAQAVLATYQDRDLPLLLGSLKSNIGHTQAASGVAGVIKVMALRHGVVPQTLHVDEPSRHVDWSAGAVELV
 TSNQWPWSSERPRRAGVSAFGVSGTNAHVILESAPVEPVVAEAGPVPVVGDVLPVLSAKSAPALTVELEQRLRAYEA
 ADEKAVAATLAAARATFGHRAVLLGGDTVTGVAVPDPRVVFVFPQGQWQLGMGRELRGSSVFAERMAECAALSE
 FVDWDLFTALDDPAVVDRVDVVQPVCWAVMVSLAAVWQASGVNPDAVVGHSSQGEIAAAVVAGSLSLRDGARVVALRS
 30 QLIKLAGRGAMASIALPAAEIDLVEGSWIAALNGPSSSTVIAGTPEAVEQVLAVQDARVRRIAVDYASHTPQVEAIR
 DELLELTGEVVSRRKPDVPWLSTVDNAWIEGPLGADYWFRLREQVGFAQAVVTLGDAVFVEVSASPVLQM SMGDAVC
 VPSLRDDGTATRMVTSLAEAYVQGVQVNWA AVL GAGTERALDLPTYPFQRQHYWALHWLRLSPAEREQALLKLC
 ESASVVLGHADAGAI PVTA AFKDLGVDSL TAVELRNSLATATGQRLPATAVFDYPTPAVLAARL

SEQ ID NO: 57

EPLAIVGMACRLPGGVSSPEGLWRLVVSGSDVISGFPADRGWGVEGLRGGLPGAADF'DAGFF'GISPREALAMDPQQ
 RLVLEASWEVLERAGIAPGSLRGS DTGVFMGAYPGGYGIGADLGGFGATAGAVSVLSGRVSYFFGFEGPAMTVDTAC
 SSSLVALHQAGHALRNSECSLALVGGVTVMASPTFVEFERQGGLAADGRSKAFSDGADGAGFSEGVGVLLVERLSD
 ARAKGHQILALVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANASLTVADVVDVVEAHGTGTT LGDPIEAQALLATY
 40 QQDRDRPLLLGSVKS NLGHTQAAAGVTGVIKVMALRHGVVPRTLHVDEPSRHVDWSAGAVELVTSNREWPVTD R PG
 RAGVSSFGISGTNAHVILEAVPVVSAVSTGGEVQPLVVSARTAPAAEDLTARLRTYLADTPD TDQRAAATTLALTRS
 VFEHRAVLLGGDTITGA AVPDPRVVFVFSGQGSQRAGMGEQLAAAFPVFAEIERVWALLDVPDGLDVDDTGHAQPA

LFALQVALSGLLESWGVRPAAVIGHSIGELAAAYVSGVWSLEDACVLVSARARLMQALPPGGVMVAVPVPEAEARAV
LRDGVEIAAVNGPSSVVLSGDEDAVLQAVAGFAKWTRLKTSFAFHSAHMDPMLDEFRAVAERLTYRRPSVEMAAGHG
VTTAEYWVRQVREAVRFGDQTTAYEDAVFVEIGPGRTLARLIDGITMLHGDTEREAALTGLSQLFVRGVDVDWPAVI
EDTTARILDLPITYPFQRQRYWLTTPRWLAGMSPEDRRQALLRVVRDSAAVVLGHAEAGTIPPNAAFKDLGIDSLTAVE
5 LRNSLATATGLRLPATLVFDYPAPETLAARLD

SEQ ID NO: 58

EPLAIVGMACRLPGGVASPEDLWRLVASGTDASISGFPTDRGWDVEGLFDPDPDVAGKTYCVQGGFLDTAARFDAAFF
GISPREALAMDPQQRVLVEASWEAFERAGIEPGSLRGSDTGVMGAFPPGGYGLGADLEGYGVTTGGPNAVSGRLSYFF
10 GLEGPAFTVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMGTPQTFVEFSRQRGLAVDGRSKSFSDQADGTGWS
EGVGVLVVERLS DARAKGHQILAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANASLTVADVDDVEAHGTGTTL
GDPIEAQALLATYGQDRDRPLLLGSLVKSNGHTQAAAGVAGVIKMMALQHGIVPQTLHVSEPSRHVDWTAGAVELV
TSNQWPWPSSGRPGRAGVSAFGVSGTNAHVILEGVPSNTPVSTAAGDVLPLVVSARTAPAVEDLTARLRITYLADTPGT
DQRAAATTLALTRSVEHRAVLLGEDTITGVAVPDSRVVVFVFSGQGSQRAGMGELAAAFPVFAAIHERVWALLDVP
15 DGLDVEDDTGHAQPALFALQVALSGLLESWGVRPDAVIGHSIGELAAAYVSGVWSLEDACALVSARARLMQALPSGGV
MVAVPVPEAEARAVLRDGVEIAAVNGPSSVVLSGDEDAVLQAVAGFAKWTRLKTSFAFHSAHMDPMLDEFRAVAERL
TYRRPSVEMAAGDRVTTAEYWVRQVREAVRFGDQTTAYEDAVFVEIGPGRTLARLIDGITMLHGENEGHAALAALSH
LFVQGVVRVDWPAVLGTTAERVDLPITYPFQHEHYWARAEHWLAGLPADEREKALLKIVRDSAAAVLGHADGRTVASGA
VFKELGLDSLTAVELRNSLGKATGLRLPSTAAFDYPTPAALATRL

SEQ ID NO: 59

EPLAIVGMACRLPGGVSSPEDLWRLVESGSDAISGFPTDRGWDVDGLFDPDPDAAGKSYCVQGGFLDSAAEFDAAFF
GISPREALAMDPQQRLLLETSWEAFERAGIDPGSVRGSDTGVMVGAFPPGGYGAGADIEGYGATAGPSVLSGRLSYFF
GLEGPAFTVDTACSSSLVALHQAGHALRQGECSLALVGGVTVMASPVTFVEFSRQRGLAADGRCKAFGDGADGTGWS
25 EGVGVLLVERLSDAQAKGHQILAVVRGSAVNQDGASNGLTAPNGPSQQRVIQAAALASAGLVTSDDVDVVEAHGTGTTL
GDPIEAQAVLATYGQDRSTPLLLGSLKSNIGHTQAAAGVSGVIKMMALRHGVVPQTLHADEPSRHVDWSAGAVELL
TSNRSWPSSERARRAGVSAFGVSGTNAHVILESAPVEPVVAVAGVPVVSVDVLPVLVLSAKSAPALTALEQRLRVYDG
AAGRALATARATFDHRAVLIGDDTVTGAVPDPVVVFVFPQGQWQWLGMGREL RDSSVVFASRMAECAAAALSEFVDW
DLFTALDDPAVVDRVDVVQVPCWAVMVSLAAVWQASGVNPDVAVGHSSQGEIAAAVVAGSLSLRDGARVVALRSQLIK
30 GLAGRGAMASIALPAAEIDLVEGSWIAALNGPSSSTVIAGTPEAVEQVLAVQDARVRRIAVDYASHTPQVEAIRDELL
ELTAEVESRRPDVPWLSTVDNTWVEGPLSADYWFRLREQVGFQAQAVVTLGDAVFVEVSASPVLQMGMGDAVTVA TL
RRDDGSALRMVTS LAEAYVQGVNVNWA AVLGAGTERALDLPITYPFQRQHYWVTAQSLAGLPAEDREKALLKIVRDSA
AQVLGHPDGRAVPAGAAFIELGVDSL TGVMERNRLGGITGLRLPATMVFDYPTPAALAGRL

SEQ ID NO: 60

EPLAIVGMACRLPGGVSSPEELWRLVESGVDAISGFVDRGWDVENLFDPPDAAGKSYCVQGGFLDTAAEFDAAFF
GISPREALAMDPQQRVLLETSWEAFERAGIEPGSLKGS DTGVYMGAFSGGYAADLEGFGATAGATSVLSGRVSYFFG
FEGPAMTVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMAS PQSFVEFSRQRGLAADGRSKAFADAADGTGWAE
GVGVLLVERLSDAQAKGHQILAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANASLTPADIDVVEAHGTGTTLG
40 DPIEAQAVIATYGQDRSTPLLLGSLKSNIGHTQAAAGVSGVIKMMALRHGVVPQTLHVDQPSRHVDWSAGAVELVT
SNQWPWPSSERPRRAGVSAFGVSGTNAHVILESAPAEPVVAEVGLVPVVSVDVLPVLVLSAKSAPALTVLEQRLRAYEAA
DERTVAATLATARATFDHRAVLIGTETVTGPLMTDPRVVVFVFPQGQWQWLGMGREL RGSSVVF AERMAECAAAALSDF

VDWDLFTALDDPAVVDRVDVVQPCWAVMVSLAAVWQAAGVNPDAVVGHSSQGEIAAAVVAGSLSLRDGALVVALRSQ
LIKGLAGRGAMASIALPADQIGLVEGAWIAALNGPSSTVIAGSPEAVEQVLAAQDARVRRIAVDYASHTPQVEAIRD
ELLELTAGVSSQPPTVPWLSTVDNTWVEGPLPADYWFRNLREQVGFAQAVVTLGDADFVEVSASPVLQMGMGDAVCV
PSLRDDGSATRMVTSLAEAYVQGVNVNWAAVLGAGTERALDLPTYPFQQRQYWAGHWLARLAPGERETALLKLVSE
5 SAAAVLGHADARSIPATAVFRDLGMDSLTAVEVRNSLAKTTGLRLPATLAFDYPTPAVLAARL

SEQ ID NO: 61

EPLAIVGMACRLPGGVSSPEGLWRLVVS GSDVISGFPADRGWGV EGLRGGFLPGAADF'DAGFF'GISPREALAMDPQQ
RLVLEASWEVLERAGIAPGSLRGS DTVGMGAYPGGYGIGADLGGFGATAGAVSVLSGRVSYFFGFEGPAMTVDTAC
10 SSSLVALHQAGHALRNSECSLALVGGVTVMASPTFVEFERQGG LAADGRSKAFSDGADGAGFSEGVGVLLVERLS
ARAKGHQILALVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANASLTVADVDVVEAHGTGTTLGDP IEAQALLATY
GQGRSVPLLLG SVKSNLGH TQAAAGVTGV IKVMALRHGVVPRTLHVDEPSRHVDWSAGAVELVTSNREWPVVD RRG
RAGVSSFGISGTNAHVILEGIPSNTPVSTAAGAVLPLVVSARTAPAAEDLTARL RAYLSAAPETDQRAAAATLALTR
SVFEHRTVLLGDDTITGAAMPDPRVVFVFSGQGSQRAGMGEQLAAVFPVFAEIERVWALLDVPDGLDIDDTGHAQP
15 ALFALQVALSGLLESWGV RPDVAVIGHSIGELAAAYVSGVWSLEDACALVSARARLMQALPPGGVMVAVPVSEAEART
VLRDGVEIAAVNGPSSVVLSGDEDAVLQAVSGFAKWTRLKTS HAFHSAHMDPMLDEFRAVAERLTYRRPSVEMAAGH
GVTTAEYWVRQVREAVRFGDQTTAYEDAVFVEIGPGRNLARLIDGITMLHGDTEREAALTGLS QLFVRGVDVDWATV
IEDTTARILDLP TYPFQHERYWLSWLVLPPAERAKALLKTVRDSAAVVLGHQGT RAIPVDGAFRELGMDSL TAVEL
RNSLAKATGLSLSATLVFDYPTPKVLADHLD

SEQ ID NO: 62

EPLAIVGMACRLPGGVSSPEELWRLVESGSDAISGFPVDRGWDADGLFDPDPDAAGKSYCVQGGFLDTAAEFDAAFF
GISPREALAMDPQQRLVLETSWEAFERAGIEPGSIKGS DTVGFIGAYPGGYGSGVELGGFGATSGAGSVLSGRVSYF
FGFEGPAMTVDTACSSSLVALHQAGYALRQGD CSMALVGGVTVMSTPHIFVEFSRQRGLAADGRCKAFGDGADGTGW
25 SEGVGVLLVERLS DARAKGHQVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIIHAALASAGLVTS DVDVVEAHGTGTT
LGDP IEAQAVIATY GQDRSTPLLLGSLKSNIGHTQAAAGVSGVIKVMALRHGVVPQT LHVDPQSRHVDWSAGAVEL
LTSNQWPWPSSERARRAGVSAFGVSGTNAHVILESAPVEPVVAEAGPVPVSDVLPLVLSAKSAPALRALEQRLRVYD
GAAGRALATARATFDHRAVLIGDDTVTGVAVPDPRVVFVFPQGQWQLGMGRELRGSSVFAERMAECAAAALSEFVD
WDLFAALDDPAVVDRVDVVQPCWAVMVSLAAVWQAAGVNPDAVVGHSSQGEIAAAVVAGSLSLRDGALVVALRSQ LI
30 KGLAGRGAMASIALPATEISLVEGAWIAALNGPSSTVIAGSPEAVEQVLAVQDARVRRIAVDYASHTPQVEAIRDEL
LELTAGVSSQLPTVPWLSTVDNTWVEGPLPADYWFRNLREQVGFAAAVQELGESV FVEVSGSPVLIQSMGDAVT VAT
LRRDDGSATRMVTSLAEAYVQGVQVNWAAVLGAGSERALDLPTYPFQRDHFWVLSLAALPSAEREKALVKIVCESAA
AVLGHTDTSAPVAAAFAKELGLDSLTAVDLRNRLRRATGLQLPATLVFDYPTPTAMAARL

SEQ ID NO: 63

EPLAIVGMSCRLPGGVSSPEDLWRLVESGSDAISGFPTDRGWDVDGLFDPDPDAAGKTYCVQGGFLEAAGEFDAAFF
GISPREALTMDPQQRVLLEASWEAFERAGIAPTSVRGTDTVGFIGAFPVGYGAGADHEGYTATAGVGSVLSGRLSYF
FGLEGPAMTMDTACSSSLVALHLAASALRN GECSLALAGGVTMATPEVFTEFARQRGLAADGRCKPFADAADGAGF
SEGAGLLVLERLS DARRNGHQVLAVIRGSAVNQDGASNGLTAPNGPAQQRVIRQALANAGLNSSD DVLVLEAHGTGTT
40 LGDP IEQAVLATY GQDRSTPLLLGSLKSNIGHTQAASGVAGVIKVMALRNLVPRSLHLDEPSRHVDWSAGAVEL
LTSNQWPWPSSDRPRRAGVSAFGVSGTNVHVILESAPAEPVGAEAGPLPVVG DVLPLVVS AKSAPALTALEQRLRAHV
AADERAAAATL ATARATFDHRAVLIGAETVTGVAAVDP RVVFVFPQGQWQLGMGRELRGSSVFAERMAECAAAALS

EFVDWDLFTALDDPAVVDRVDVVQPVCWAVMVSLAAVWQAVGVNPDVAVGHSSQGEIAAAVVAGSLSLRDLGALVVALR
 SQLIAGLAGRGAMASIALPADQISLVEGAWIAALNGPSSTVIAGTPEAVEQVLAQDARVRRIAVDYASHTPQVEAI
 RDELLELTGEVVSRRKPDVPWLSTVDNAWIEGPLGADYWFRLNREQVGFQAQAVVTLGDVAFVEVSASPVLQMGMGDAV
 CVPSLRRDDGTATRMVTSLAEAYVQGVQVNWAAVLGAGTERALDLPTYPFQRRERFWVLWLAGLAPQERETALLKLCV
 5 DSAAVVLGHGDGQAIPDTTAFKDLGVDSLTAVEVRNRLAAATGLRLPATMVFDDYPTPTALAARL

SEQ ID NO: 64

EPLAIVGMACRLPGGVASPEDLWRLVESGTDAITGFPADRGWTTTEPGQGGFLADAAGFDAAFFGISPREALAMDPQQ
 RLLLETSWEAFERAGIAPLSLRGSDTGVIYIGAYPDGYGIGADLGGFGTTAGSPSVLSGRVSYFFGLEGPATVDTAC
 10 SSSLVALHQAGYALRNNECSLALVGGVTVMATPEVFSAFALQDGLAADGRSKAFSDGADGAGFSEGVGVLLVERLS
 AQANGHQILALVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANASLTPADVVDIEAHGTGTTLGDPIDEAQALLATY
 GQGRSVPLLLGSGVKSNGHTQAAAGVTGVIMVMALRHGVVPRTLHVDEPSRHVDWSAGAVELVTSNREWPVTDPRG
 RAGVSSFGISGTNAHVILEAVPVVSAVSTGGEVQPLVVSARTAPAAEDLASRLRITYLADTPDQDQRAAAATLALTRS
 VFEHRTVLLGDDTITGAAMPDPRVVFVSGQGSQRAGMGEQLAAVFPVFAEIERVWALLDVPDGLDIDDTGHAQPA
 15 LFALQVALSGLLESWGVRPDAVIGHSIGELAAAYVSGVWSLEDACALVSARARLMQALPPGGVMVAVPVPEAEARTV
 LRDGVEIAAVNGPSSSVLSGDEDAVLQAVSGFAKWTRLKTSFAFHSAHMDPMLDEFRTVAERLTYRRPSVEMAAGHG
 VTTAEYWVRQVREAVRFGDQTTAYEDAVFVEIGPRTLARLIDGITMLHGDTEREAALTGLSQLFVRGVDVDWATVI
 EDTTARILDLPITYPFQHERYWAGRWLAGLAPDKRDAALLTMVRDSAARVLGHADGSAISPTATFRDLGVDSLTADEL
 RNRLARTAGLRLATTIVFDYPTPTALAAHL

SEQ ID NO: 65

EPLAIVGMACRLPGGVASPEDLWRLVESGTDAITGLPTDRGWDLGAVAAESYCVEGGFLDGVAGFDAAFFGISPREA
 LAMDPQQRLLETSWESLERAGIAPLSLRGSDTGVMGAYPGGYGAGADLGGFGTTSGAASVLSGRISYFFGLEGPA
 MTVDTACSSSLVALHLAGQALRNNECSLALVGGVTVMAAPDIFPEFARQRLASDGRSKAFADSADGTGWSEGVGV
 25 LVERLSDAQANGHQILALVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANASLTPADVVDIEAHGTGTTLGDPIDEA
 QALLATYGGQGRSVPLLLGSGVKSNGHTQAAAGVTGVIMVMALRHGVVPQTLHVDEPSRHVDWSAGAVELVTSNRE
 PVTDRPGRAGVSSFGISGTNAHVILEAVPSDTPAPTTTTDAVLPLVVSTRTAPAAEDLTARLRLAYLSAAPETDQRAAA
 ATLALTRSVEHRAVVLGEDTITGVAVPDPRVVFVSGQGSQRAGMGEQLAAAYPVFAAIEHVRWALLDVPDGLD
 DTGHAQPALFALQVALSGLLESWGVRPAAVIGHSIGELAAAYVSGVWSLEDACVLVSARARLMQALPPGGVMVAVPV
 30 PEAEARAVLRDGVEIAAVNGPSSSVLSGDEDAVLQAVSGFAKWTRLKTSFAFHSAHMDPMLDEFRAVAERLTYRRPS
 VEMAAGDRVTTAEYWVRQVREAVRFGDQTTAYEDAVFVEIGPRTLARLIDGITMLHGDTEREAALTGLSQLFVRGV
 DVDWATVIEDTTARILDLPITYPFQHEHYWLRAARTPAERAQELLKLVRDNAAAVLGHADGRTVPAFAAFRDLGVDS
 LIAVELRNALATGLQLPTTIVFDYPTASSLAERL

SEQ ID NO: 66

EPLAIVGMACRLPGGVESPEDLWRLVESGADAI SGFPTDRGWDADGLFDPDLAVGKTYCVQGGFLQTAAEFDPAFFG
 ISPREALAMDPQQRLVLETSWEAFERAGIEPGSLKGSDDTGVMGAYPGGYGMGADLGGFAATAGAGSVLSGRVSYFF
 GFEGPAMTVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMASPSQSFVEFSRQRLAADGRSKAFADAADGTGWA
 EGVGVLLVERLSDAQAKGHRILAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANASLTPADIDVVEAHGTGTTL
 40 GDPIDEAQAVIATYQDRSTPLLLGSLKSNIGHTQAAAGVSGVIMVMALRHGVVPQTLHVDPQSRHVDWSAGAVELV
 TSNQWPWSSERPRRAGVSAGVSGTNAHVILESAPVEPVGAEEAGLVPVADVLPWVVSASAPALRALEQRLRAYEA
 ADERTVVATLATARATFDHRAVLIGTETVTGPLMTDPRVVFVFPQGQWQWLGMGRELRGSSVFAERMAECAAALSE

FVDWDLFTALDDPAVVDVRVDVVQPVCWAVMVSLAAVWQAAGVNPDAVVGHSSQGEIAAAVVAGSLSLRDGALVVALRS
 QLIKGLAGRGAMASIALPADQIDLVEGAWIAALNGPSSSTVIAGTPEAVEQVLAAQDARVRRIAVDYASHTPQVEAIR
 DELLELTAEVLSRKPDVPWLSTVDNTWVEGPLPADYWFRNLREQVGFAQAVVTLGDAVFVEVSASPVLIIQSMGDAVT
 VATLRRDDGSATRMVTSALAEAYVQGVQVNWGAVLGAGTERALDLPTYPFQRQHYWALERLGERAGTERHRLMLEVVL
 5 GHAASVLGHSSAAALEPDRPFKDLGMDSLTAIELRNHLVAETGLRLPATMVFDFPTADALAGHL

SEQ ID NO: 67

EPIAVVSMACRLPGGVDTPEGLWRLVESGTDALSGFPTDRGWDLTDFYSADPQGGFLTGAEEFDAGFFGISPREALG
 MDPQQRLLLETTWEAIERAQLDPRSLRGRDVGYYVGGAAQGYGVGFAGEPRDNAITASSISLLSGRVSYALGLQGPG
 10 VTVDTACSSSLVALHLACQALRQRECSLALVGGVSVIATPDVFAEFSRQNGLAADGRCKSFGAAADGTGWSEGVGML
 VLRLSEATHRHGRILAVVRGSAVNSDGASNGLTAPNGQSQQRVIRQALSNAAGLAASDVVDVVEAHGTGTRLGDPIEA
 EAILATYGQDRAAPAWLGSLKSNIGHTMAASGVLGVIKMVEAMRHGTVPRTLHVDEPSPHVDWSAGRVALLTENQPW
 PDGAKPRRAGVSSFGLSGTNAHVVLEQHPEPASVPVARETGPVPWVLSAQSPKALQEQAQRLHAALVSDPRWHPLDV
 AFSLATTRSAFTHRTAVVASGRDLLEALSTLATSATATSTTARTRRVAFLLFDGQGTQRAGMGRELYERHPAFARAWD
 15 EVSAAFDKHEHPLHAVYFGAGALDELVDGTGYAQAAIFTFEVALFELLHEWGVPRDFVAGHSIGEVAAYVSGLS
 LADAAQLIVARGRALRSAPPGAMAALRAGETETREFLARTGTALDVAAVNSPEAVVSGSPEAVAEFTAATASGRR
 ARRLNVNRAFHSRHVDGLLDDFRAVLES LTCRTDTVLPVSTVTGRLIDPAELRTPQYWLSQVRDTRVFQDAVAELA
 ANGVGVFVEVGPSSSLASAGTETLGDEAHFQALQHSRTPADPALLTALAGLHSGGVGDWEKVLVGGRAVELPVYPF
 QHRAYWLAPASTQEPATMLELVRFEVAAVLGMPDPAVFEETSFLELGFDSL SAVRLRNRLTRSTGVLEPATLLFDH
 20 PTPAELAAHL

SEQ ID NO: 68

EPLAIVGMACRLPGGVASPEDLWRLVESGTDALISDFPTDRGWDVEGLYDPDPDVPGKSYAVKGGFLDAAGFDAAFFG
 ISPREAAAMDPQQRVLVEASWEAFERAGIEPGSVRGSDTGVMGAMANGYGAGADLGGFGATAGAGSVLSGRISYFF
 25 GLEGPAMTVDTACSSSLVALHQASFALRQGECSLALVGGVTVMPTPQLFVEFARQRLAVDGRSKAFADAADGSGFS
 EGVGVLVVERLSDAQAKGHRVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALDNAGLSSMDVDVVEGHGTGTRL
 GDPIEAQAVISTYQDRERPLLLGSLKSNIGHAQAAAGVSGVIKMMALRHGVVPQTLHVDEPSRHVDWAAGAVELV
 TENQPWPVAERARRAGVSSFGISGTNAHVILESAPEAAASASEPVTTPSEVSVPVVASDVVPLVVSATPGALTDIE
 ERLRGYLAAPADMQAVASTLAATRSVFEHRAVLLGDDTITGIATPDPRVVFVFPQGQWLGMGSVLRETSPVFA
 30 GRMAECAALREFVDWDLFSVLDDPTVVDVRVDVLQPACWAVMVSLAAVWQEAGVSPDAVIGHSSQGEIAAACVAGAVS
 LRDAARIVALRSKLIGARLGHGAMASIALPADAITLTDGAWIAAHNGPASTVIAGTPQAVDTVLAAYEAQGIRVRR
 TVDYASHSPQVEEIHTELLDATATVGSQTPAVPWLSTVDGAWVEGPLDADYWYRNLRREPVRFDQAVTHLQTQGETVF
 IEVSASPALTPAMNDDAITVATLRRDDSDPTRILTALAEAFVQGVGVDPVAVTGATTARVLDLPTYAFQRQRYWTL
 GLAAAERRQALAKLVRESAAVVLGHADPDSVPAFAAFKDLGVDSLTAVELRNSLGRSTGLRLPATMVFDPYTPDALA
 35 ARLD

SEQ ID NO: 69

EPLAIVGMACRLPGGVDSPEDLWRLVESGTDALISGFPTDRGWDLDLSLYDPILGASGEFYSAQGGFLDRAADFASFF
 GISPREALAMDPQQRVLVEVSWEALERAGIEASSVRGSDTGVMGAMANGYGIGADFGAFGMTASAGSVLSGRVSYF
 40 FGLEGPAMTVDTACSSSLVALHQAGYSLRQGECSMALVGGVTVMPTPQTFVEFARERGLAVDGRSKAFADAADGSGF
 SEGVGVLVVERLS DARARGHRVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANAGLASADVDVVEAHGTGTR
 LGDPIEAQAVIETYGQDRERPLLLGSLKSNIGHTQAAAGVSAVIKMMALRHGVVPQTLHVDEPSRHVDWTAGAVEL

ATEKLPWPASDRVRRAGVSSFGISGTNAHVILESVPAEVSVPSESSGPNLASDVVPLVVS AKTSGALVDIEERLRGY
 LAAVPGVDLGAVASVLAGRSVFGHRGVLVGGELVSGVALSGPRVVFVFSGQGSQCVMGERLAGVFPVFAEVYGRV
 WDL LDVPGSGLGVDDTG FVQPALFALQVGLFGLLESWGVRPEVVVGHSVGEVAAGYVAGLWSLEDACVLVSARARLM
 QGLPGGGVMVSVSVSEERARAALVEGVEIAAVNGPSSVVL SGDEAAVVGVAEGLGGRWRRLATSHAFHSARMDPMLD
 5 EFRVVAEGLEYREPRIVMAGGAGVVSPEYWVRQVRD TVRFGDQVAAYQGD AVFVEVGP GGS LARLIDGVA VGDGEDE
 VRAAVMAVAELFVRGVDVDWPAVVGTTATPVDLPTYPFQRQRYWTASWLVALEPEERGGQALLRMVREGASVVLGHAD
 ARAVEVDRAFRDLGVDSLTA VQVRNNLAKATGLRLPATMVFDYPTPAAL AARLD

SEQ ID NO: 70

10 EPLAIVGMACRLPGGVASPEDLWRLVESGTDVISGFPTDRGWDLDNLYDPDPAVGKSYCVQGYFLDDVADFDASFFG
 ISPREALAMDPQQRLILEASWEAFERAGIEPGSVRGSDTGVMGAFSSGYGIGADHSGFGMTAGAGSVLSGRISYLE
 GLEGPAMTVDTACSSSLVALHQASSALRQGECSLALVGGVTVMPTPQTFLFARQRGLAADGRSKAFSDAADGSGFS
 EGVGVLVVERLS DARARGHRVLAVVRGS AVNQDGASNGLSAPNGPSQQRVIRQALANAGLSSADVDVVEAHGTGTRL
 GDPIEAQAVIETYGQDRERPLLLGSLKSNIGHTQAAAGVSAVIKMMALRHGVVPQTLHVDEPSRHVDWTAGAVQLA
 15 TEKQPWPASDRARRAGVSSFGISGTNAHVILESAPVHSVETDETAPMALASDVVPLVVS AKTSGALVDIEERLRGYL
 AVAGSEVDLGAVASVLAGRSVFGHRGVLVGGELVSGVALSGPRVVFVFSGQGSQCVMGERLAGVFPVFAEVYGRV
 WDL LDVPGSGLGVDDTG FVQPALFALQVGLFGLLESWGVRPEVVVGHSVGEVAAGYVAGLWSLEDACVLVSARARLM
 QGLPGGGVMVSVSVSEERARAALVEGVEIAAVNGPSSVVL SGDEAAVVGVAEGLGGRWRRLATSHAFHSARMDPMLD
 EFRVVAEGLEYREPRIVMAGGAGVVSPEYWVRQVRD TVRFGDQVAAYQGD AVFVEVGP GGS LARLIDGVA VGDGEDE
 20 VRAAVMAVAELFVRGVDVDWPAVVGTTAAPVDLPTYPFQRQRYWTQTWLTGLASEDRRQALLKVV RDSAA TVLGHAD
 AGMIPATAAFKDLGLDSLTA VELRNSLGKSTGLSLPATMVFDYPTPDALADRLD

SEQ ID NO: 71

EPLAIVGMACRLPGGVASPEDLWRLVESGTD AISDFPTDRGWDVEGLYDPDPDAPGKSYAVKGGFLDAA GFDAAFFG
 25 ISPREALAMDPQQRLVLEASWEAFERAGIEPGSVRGSDTGVMGAFAPAGYGGDREGFGATAGAGSVLSGRVSYFFGL
 EGPAITVDTACSSSLVALHQAGYSLRQGECSLALVGGATVMATPQTFVEFSRQRGLSVDGRSKAFADAADGTGWAEG
 VGV LVVERLS DAQAKGHQVLAVVRGS AVNQDGASNGLTAPNGPSQQRVIRQALANAGLSSADVDVVEAHGTGTLGD
 PIEAQAVIATY GQDRERPLLLGSLKSNIGHTQAAAGVSGVIKMMALQHGVVPQTLHVDEPSRHVDWAAGAVELVTE
 NQPWPVAERARRAGVSSFGISGTNAHVILESAPAEAAASASEPVTTPSEASVPVVASDVVPLVVS AKTPGALTDIEER
 30 LRGYLAAASD VDMAAASTLAATRSVFEHRAVLLGDDTITGIATPDPRVVFVFPQGQWLGMGSVLRETSPVFAGR
 MAECAALGEFVDWDLFSVLDDPTVVDRVDVLQPACWAVMVSLAAVWQEAGVSPDAVIGHSSQGEIAAACVAGAVSLR
 DAARIVALRSKLIGARLGHGAMASIALPAGDVALVDGAWIAAHNGPASTVIAGTPQAVDTVLAAHEAQGIRVRRITV
 DYASHSPQVEEIHAE LLDATAAVGSQAPAVPWLSTVDGAWVEGPLDADYWRNLREPVRFDQAVTLLQTQGETVFIE
 VSASPALTPAMNDDAITVATLRDDDDSPARILTA LAEAFVQGVGVWDWPAVTGATTSRVDLPTYPFQHQR YWAWLAGL
 35 APEARQALLKVVRESAAVVLGHTGADTVPVTA AFKDLGLDSLTA VELRNSLGRSTGLRLPVTA VFDYPTPAAL AAR
 LD

SEQ ID NO: 72

EPLAIVGMACRLPGGVASPEDLWRLVESGRDVISDFPVDRGWDLDNLYDPDPAVGKTYCKRGGFLDAAA EFDAAFFG
 40 ISPREAAAMPDQQRLVLEASWEAFERAGIEPGSVRGSDTGVMGAFANEYGAGADF GAFGMTAGAGSVLSGRVSYLF
 GLEGPAMTVDTACSSSLVALHQAGSALRQGECSLALVGGVTVMPTPQLFVG FARERGLAVDGRSKAFSDAADGAGWA
 EGVGVLVVERLS DAQARGHRVLAVVRGS AVNQDGASNGLTAPNGPSQQRVIRQALANAGLSSADVDVVEAHGTGTRL

GDPIEAQAVIATYQDRERPLLLGSLKSNIGHTQAAAGVSGVIKMMALRHSVVPRTLHVDEPSRHVDWAAGAVELV
 TEKQPWPTSDRARRAGVSSFGISGTNAHVILESAQAQPLETDEALVPVVASDVMPPLVVS AKTPDALTDIEDRLRAHL
 AAAP EADMQAVASTLAATRSVF EHRVLLGDDTITGVAASGPRVVVFVFPQGQWLGMGSVLRETSPVFAGRMAECA
 AALREFVDWDLF SVLDDPTTVVDRVDVLQPACWAVMVSLAAVWQEAGVSPDAVIGHSQGEIAAACVAGAVSLQDAARI
 5 VALRSKLI AHLAGHGAMASIALPADAITLTDGAWIAAHNGTASTVIAGTPQAVDTVLATHEAQGIRVRRITVDYASH
 SPQVEEIHTELLDATATVGSQTPAVPWLSTVDNTWISRPLD TDYWRNLREPVRFDQAVTLLQTQGETVFIEVSASP
 ALTPAMND DAVTVATLRRDDDSPTRILTALAEAFVQGVGVDPVAVTGATTTPVDLPTYPFQRQRYWTASWLAGLAPE
 ARGQALLKV VRESTAVVLGHVDTETVPATAPFKDLGLDSLTA VQVRNGLAKATGLRLPATMVF DYPTPAALAAARLD

10 SEQ ID NO: 73

EPLAIVGMACRLPGGVASPEDLWRLVESGTD AISDFPTDRGWDVEGLYDPDPDVP GKSYAVKGGFLDAAGFDAAFFG
 ISPREALAMDPQQRLVLEASWEAFERAGIEPGSVRGSDIGVFMGAMANEY GAGADFGAFGMTAGAGSVLSGRVSYFF
 GLEGPAMTVDTACSSSLVALHQAGSALRQGECSMALVGGVTVMPTPQT FVEFARQ RGLATDGRSKAFADAADGSGFS
 EGVGVLVVERLS DARARGHRVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANAGLASADVDVVEAHGTGTRL
 15 GDPIEAQAVIETYQDRERPLLLGSLKSNIGHTQAAAGVSAVIKMMALRHGVVPQTLHVDEPSRHVDWTAGAVQLA
 TEKQPWPASDRARRAGVSSFGISGTNAHVILESA PVHSVETDETAPMALASDVVPLVVS AKTSGALVDIEERLRGYL
 AAVPGVDLGAVASVLAGSRVSFGHRGVLVG GELVSGVALSGPRVVVFVSGQGSQCVGMGERLAGVFPVFAEVYGRVW
 DLLDVP GSGLVDDTG FVQPALFALQVGLFGLLESWGV RPEVVVGH SVGEVAAGYVAGLWSLEDACVLVSARARLMQ
 GLPGGGVMVSVSVSEDRARAALVEGVEIAAVNGPSSVVL SGDEAAVVGVAEGLGGRWRRLATSHAFHSARMDPMLDE
 20 FRVVAEGLEYREPRIVMAGGAGVVSPEYWVRQVRD TVRFGDQVAAYQGD AVFVEVGP GGS LARLIDGVAVGDGEDEV
 RAAVMVAELFVRGVDVDWPAVVGTTATPVDLPTYPFQRQRYWAWLTGLASEDRRQALLKVVRDSAATVVLGHADARA
 VEVDRAFRDLGVDSLTA VQVRNNLAKATGLRLPATMVF DYPTPAALAAARLD

SEQ ID NO: 74

25 EPLAIVGMACRLPGGVASPEDLWRLVESGTD AISDFPTDRGWDVEGLYDPDPDVP GKSYAVKGGFLDAAGFDAAFFG
 ISPREALAMDPQQRLVLEASWEAFERAGIEPGSVRGSDTG VFMGAFTNGYGTGADLDGFGATAGTGSVLSGRVSYFF
 GLEGPAMTVDTACSSSLVALHQAGYSLRQGECSMALVGGVTVMPTPQT FVEFARQ RGLATDGRSKAFADAADGSGFS
 EGVGVLVVERLSAQARGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLASADVDVVEAHGTGTRL
 GDPIEAQAVIETYQDRERPLLLGSLKSNIGHTQAAAGVSGVIKMMALQHGVVPQTLHVDEPSRHVDWSAGAVELA
 30 RERQPWPVAGRARRAGVSSFGISGTNAHVILESA PVHSVETDETAPMALASDVVPLVVS AKTSGALVDIEERLRGYL
 AAVPGVDLGAVASVLAGSRVSFGHRGVLVG GELVSGVALSGPRVVVFVSGQGSQCVGMGERLAGVFPVFAEVYGRVW
 DLLDVP GSGLVDDTG FVQPALFALQVGLFGLLESWGV RPEVVVGH SVGEVAAGYVAGLWSLEDACVLVSARARLMQ
 GLPGGGVMVSVSVSEERARAALVEGVEIAAVNGPSSVVL SGDEAAVVGVAEGLGGRWRRLATSHAFHSARMDPMLDE
 FRVVAEGLEYREPRIVMAGGAGVVSPEYWVRQVRD TVRFGDQVAAYQGD AVFVEVGP GGS LARLIDGVAVGDGEDEV
 35 RAAVMVAELFVRGVDVDWPAVVGTTATPVDLPTYPFQRQRYWAWLTGLASEDRRQALLKVVRDSAATVVLGHADARA
 VEVDRAFRDLGVDSLTA VQVRNNLAKATGLRLPATMVF DYPTPAALAAARLD

SEQ ID NO: 75

EPLAIVGMACRLPGGVDSPEDLWRLVESGTDVISGFPTDRGWDLDNLYDPDPAVGKSYCVQGYFLDDVADF DASFFG
 40 ISPREALAMDPQQRLILEASWEAFERAGIEPGSVRGSDTG VFMGAFSSGYGIGADHSGFGMTAGAGSVLSGRVSYLF
 GLEGPAMTVDTACSSSLVALHQAGSALRQGECSLALVGGATVMPTPQT FVEFARQ RGLATDGRSKAFADAADGAGWA
 EGVGVLVVERLSAQARGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLSSADVDVVEAHGTGTRL

GDPIEAQAVIETYGQDRERPLLLGSLKSNIGHTQAAAGVSAVIKMMALRHGVVPQTLHVDEPSRHVDWTAGAVQLA
TEKQWPWASDRARRAGVSSFGISGTNAHVILESAPAQPLETDEPSAPIVASDVVPLVVSAKTLDALTDIEDRLRGYL
AAASDVDMAAVASTLAATRSIFEHRAVLLGDDTITGIATPGPRVVVFVFPQGQWQLGMGSTLRETSPVFAARMAECA
AALREFVDWDLFSILDDPTVVDRVDVLQPACWAVMVSLAAVWQEAGVSPDAVIGHSQGEIAAACVAGAVSLQDAARI
5 VALRSKLIAGHLAGHAGAMASIALPADAITLTDGAWIAAHNGTASTVIAGTPQAVDTVLATHEAQGIRVRRITVDYASH
SPQVEEIHTELLDATTTINPRTPAVPWLSTVDNTWISRPLDIDYWRNLREPVRFDQAVTLLQTRGETVFIEVSASP
ALTPAMNDDAITVATLRRDDDDSPARILTALAEAFVQGVGVDPVAVTGATTARVLDLPTYAFQHQRWATAWLAGLAP
AERGEALLKVVSDTVARVLGHADGRTIPATAAFKELGVDSLTAVELRNRLSAATGLRLPATMVFDPSPGALAGWL

10 SEQ ID NO: 76

QPLAIVGMACRLPGGVASPEDLWRLVESGTDIAISDFPVDRGWDLEGLYDPASDEPGVLYCDQGGFLDAAAGFDAAFF
GISPREAAAMDPQQLVLEASWEAFERAGIEPGSVRGSDTGVMGAMANGYGAGADLGGFGATAGAGSVLSGRISYF
FGLEGPAMTVDTACSSSLVALHQASFALRQGECSLALVGGVTVMPTPQTFVEFARQRLAADGRSKAFSDAADGAGW
AEGVGVLVVERLSDAQAKGHQVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLSSAEVDVVEAHGTGTR
15 LGDPIEAQAVIATYQDRERPLLLGSLKSNIGHAQAAAGVSGVIKMMALRHGVVPQTLHVDEPSRHVDWAAGAVEL
VTENQWPVPAERARRAGVSSFGISGTNAHVILESAPAEASASEPVTTPSEASVPVVASDVVPLVVSAKTPGALTDI
EERLRGYLAAPEADMQAVASTLAATRSVFEHRAVLLGDDTITGVAASGPRVVVFVFPQGQWQLGMGSVLRETSPVF
AGQMAECAALREFVDWDLFSVLDDPTVVDRVDVLQPACWAVMVSLAAVWQEAGVSPDAVIGHSQGEIAAACVAGAV
SLRDAARIVALRSKLIGARLGHGAMASIALPAGAITLTDGAWIAAHNGPASTVIAGTPQAVDAVLAAYEAQGIRVRR
20 ITVDYASHSPQVEEIRAELLDATATVGSQAPVVPWLSTVDGAWVEGPLDADYWRNLREPVRFDQAVTLLQTQGETV
FIEVSASPALTPAMNDDAITVATLRRDDDDSPARILTALAEAFVQGVGVDPVAVTGATTSRVDLPTYPFQHQRWWTQT
LSGLAAAERRQALAKLVRESAAVVLGHADPDSVPAFAAFKDLGVDSLTAVELRNLSLGRSTGLRLPATMVFDPYTPDA
LAARLD

25 SEQ ID NO: 77

EPLAIVGMACRLPGGVDSPEDLWRLVESGTDIAISGFPTDRGWDLDLSLYDPILGASGEFYSAQGGFLDRAADFASFF
GISPREALAMDPQQLVLEVSWEALERAGIEASSVRGSDTGVMGAFSSGYGTGSDFGAFGATSSAGSVLSGRISYF
FGLEGPAMTVDTACSSSLVALHQAGSALRQGECSLALVGGITVMSTPLTFAEFARQRLAPDGRSKAFSDAADGAGF
SEGVGVLVVERLSDAQAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLASADVDVVEAHGTGTR
30 LGDPIEAQAVIATYQDRERPLLLGSLKSNIGHAQAAAGVSGVIKMMALRHGVVPQTLHVDEPSRHVDWAAGAVEL
VTENQWPVPAERARRAGVSSFGISGTNAHVILESAPAEASASEPVTTPSEASVPVVASDVVPLVVSAKTPGALTDI
EERLRGYLAAASDVDMAVASTLAATRSVFEHRAVLLGDDTITGVAASGPRVVVFVFPQGQWQLGMGSVLRETSPVF
AGRMAECAALREFVDWDLFSVLDDPTVVDRVDVLQPACWAVMVSLAAVWQEAGVSPDAVIGHSQGEIAAACVAGAV
SLRDAARIVALRSKLIGARLGHGAMASIALPADAITLTDGAWIAAHNGPASTVIAGTPQAVDAVLAAYEAQGIRVRR
35 ITVDYASHSPQVEEIRAELLDATATVGSQAPVVPWLSTVDNTWISRPLDIDYWRNLREPVRFDQAVTLLQTQGETV
FIEVSASPALTPAMNDDAITVATLRRDDDDSPARILTALAEAFVQGVGVDPVAVTGATTARVLDLPTYPFQQRWAW
LTGLASEDRRQALLKVVDRSAATVLGHADARAVEVDRAFRDLGVDSLTAVQVRNNLAKATGLRLPATMVFDPYTPAA
LAARLD

40 SEQ ID NO: 78

EPLAIVGMACRLPGGVASPEDLWRLVESGTDVISGFPTDRGWDLDNLYDPDPAVGKSYCVQGYFLDDVADFASFFG
ISPREALAMDPQQLILEASWEAFERAGIEPGSVRGSDTGVMGAFSSGYGIGADHSGFGMTAGAGSVLSGRISYLF

GLEGPAMTVDTACSSSLVALHQASSALRQGECSLALVGGATVLTATPYGFVEISRQRGLAADGRSKAFSDAADGMSFS
EGAGVLVLERLSDAQAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLASADVDVVEAHGTGTRL
GDPIEAQAVIATYQONRERPVLGSLKSNIGHTHAAAGVSGVIKVMALQHGVPRTLHVDAPSRRHVDWAAGAVELV
TENQPWPVAERARRAGVSSFGISGTNAHVILESAPAQPLETGEPSAPIVASDVVPLVVSAPKTPDALTDIEDRLRAHL
5 AAPEADMQAVASTLAATRSIFEHRAVLLGDDTITGIATPDPRVVFVFPQGQWQLGMGSTLRETSPVFAARMAECA
TALREFVDWDLFSILDDPTVVDRVDVLQPACWAVMVSLAAVWQEAGVSPDAVIGHSSQGEIAAACVAGAVSLQDAARI
VALRSKLIHLAGHGAMASIALPADAITLTDGAWIAAHNGPASTVIAGTPQAVDTVLTATHEAQGIRVRRITVDYASH
SPQVEEIRAELLDATATVGSQAPVVPWLSTVDGAWVEGPLDADYWYRNLREPVRFDQAVTLLQTQGETVFIEVSASP
ALTPAMNDDAVTVATLRRDDSPTRILTALAEAFVQGVGDWPAVTGATTTPVDLPTYPFQRQRYWTASDRLSGRTS
10 GDQHRIMLELVLGHAAVSLGHGAADAVAADKPFKDLGMDSLTAIELRNHLVAETGLRLPATTAFDHPTADDLARRL

SEQ ID NO: 79

EPIAIVSMACRAPGGVDSPLDGLWRLVESGTDALISGFPTDRGWDVADLYSPDPAGYKSYCVQGGFLDTAADFDAAFFG
ISPREALGMDPQQRLLEASWEAIERARIDPRSLRGRSVGVFVGGASQGYGAGADDQQQSNAITGGSISLLSGRVSY
15 ALGLEPGVTVDTACSSSLVALHLASQALRQRECSLALVSGVSVMATPDVFVEFSRQRGLAPDGRCKSFSAADGTG
WSEGVLVLERLSEATRLGHRVHAVVRGSAVNQDGASNGLTAPNGASQQKVIQALANAGLAASEVDVAEHGTGT
KLGDPIDAEAILATYQDRAAPVWLGLSKSNIGHTMAASGVLGVKVMESMRHGLLPRTLNVDEPSPHVDWASGDVA
LLTENQPWPADVGPRAAVSSFGISGSNAHVLEQYGEPAAGPDLSDLTNTRAVNAADAPDRRQFPVPLMLSARSQAL
REQAGRLHAALAGAPDWRPLDIGYSLATTRSHFTHRAVAVGSGRELLRALSCLADGADWPALTTRIAKSRRVAFLED
20 GQGTQRLGMSGGLYAGFPVFAVWQVSAAFDKHLDHSLTDVFLGRDDRPAAELVDDTLYAQAGLFTLEVALFRLL
EEWGVPRPDFLAGHSIGEAAAAYAGGMFSLDVDLIVARGEALRLAPPGAMLALRASEEEVREFLGRTGAELDLAAV
NGPASVVVSGASEAVADFRARWTAAGRKARELNVSRFHSRHEAGLGRFREVLESFTFGTPVLPVIVSTVTGQLVDP
VEMSTPEYWLRLQVRQPVLFQDALRELSGQGVNTFVEIGPSGTLASAGLECLGGDASFHAVQQPRSPQDVGLMTAVAE
LHAGGTAVDWAKALAGGRATDLPVYPFQHESYWLAPADYAYPEEPGTMLLELVRLEAAKVLGITEPDTILEETSFLDL
25 GFDSLGTMLRLNRLSEVTELDLPATLLFDNPSPAELAAAYLD

SEQ ID NO: 80

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALISHFPTDRGWDLDNLYDPDPDAPGKGYRVQGGFLDAAAGFDAAFFG
ISPREAQAMDPQQRLVLEASWEAFERAGIDPGAMRGSHTGVMFGAMANGYAGADLGGFGATAGAVSVLSGRVSYLF
30 GLEGPAMTVDTACSSSLVALHQAAAYSLRQGECSLALVGGVTVMPTPQMFEVFEARQRGLAADGRSKAFADAADGAGFS
EGVGVLVVERLSDAQAKGHRVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANARLAPNEIDVVEAHGTGTTL
GDPIEAQALIAAYGQDREQPVLLGSLKSNIGHTQAAAGVSGVIKVMALQHGVPRTLHVDSPSRHVDWTAGAVELV
TENQPWPAIDRARRAGVSSFGISGTNAHVILESAPAQPVVETEEVPAPPVVASDMMPLVISAKTPSALVEFEGRLLRA
YLTSTPGVDMRAVASTLAGTRSVFEHRAVLLGDETVTGPGTGAGSGVAVSDPRVVFVFPQGSQRAGMGEQLAAVFP
35 VFAEIHQQVWDLDPDPGLDDETGYPALQVALFGLLESWGVPRQAVIGHSSVGEIAAGYVAGLWSLRDACT
LV SARARLMQTLPTGGAMVAVPVSEKQAQAALTDGVEIAAVNGPSSVVLTDGDETAVLETAALGRSTRLTTSFAHFS
ARMEPVLDDEFRTVAETLDYRTPHIPMAAGDAVVTPPEYVVRQIRDTVRFGDQVAAHENAVFVEIGPDRTLRLTDGIA
MLHGDNETQTAITALATLHTGVNIHWPVIGATTARVLDMPYAFQHQRYYTTLWLAGLAPEERKQALLKVVRDSAA
AVLGHAGADTVPVTAAFKDLGLDSLTAVELRNSLGKSTGLRLPATMVFDPNPTTLAARLD
40

SEQ ID NO: 81

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALISGFPTDRGWDIENLYDPDPDAPGKGYRVQGGFLDRAAEFDASFF

GISPREAQAMDPQQRLVLETSWEAFERAGIEPGAMRGSDTGVMGAMANGYGTGADLGAFGMTSAAVSVLSGRVSYL
 FGLEGPAMTVDTACSSSLVALHQAAYSRLRQGECSLALVGGVTVMPTPQTFVEFARQRLAADGRSKAFSDAADGAGF
 SEGVGVLVVERLS DARAKGHHVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALDNAGLSSTDVDVVEAHGTGTT
 LGDPIEAQALIAAYGQDREQPVLLGSLKSNIGHTQAAAGVSGVIKMMALQHGVPRTLHVDSRSHVDWTAGAVEL
 5 VTENQPWPAIDRARRAGVSSFGISGTNAHVILESAPDLPVVETEETPAPVVTSDMMPLVISAKTPAALADMEGRRLRS
 YLTSMPGVDMRAVASTLAGTRSVFEHRAVLLGDETVTGPGTGAVSGPRVVFVFPQGGSQRAGMGEQLAAVFPVFAE
 IHQQVWDLDDVPDPGLGADETGFAQPSLFALQVALFGLLESWGVRPQAVIGHSVGEIAAGYVAGLWSLRDACTLVSA
 RARLMQTLPTGGAMVAVPVSEKQAQAALTDGVEIAAVNGPSSVVLTDGETAVLETAALGKSTRLLTSHAFHSARME
 PVLDQFRTVAETLDYRTPHIPMAAGDAVVTPPEYWVRQIRDTVRFQDQVAAHENAVFVEIGPDRTLSRLTDGIAMLHG
 10 DNETQTAITALATLHTHG VNIHWP AVIGATTTTPVDLPTYAFERQRYWAWLAGLAPEERKQALLKTVRDNAAKVLGHA
 DARDIAVNTAFRDLGLDSLTA VQVRNSLAKATGLRLPTTTTFDYPNP TALATHLD

SEQ ID NO: 82

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISDFPTDRGWDLDNLYDPDPDTPGKAHNQGGFLDAAGFDASFFG
 15 ISPREAQAMDPQQRLVLETSWEAFERAGIDPASVRGSDTGVMGAFGSGYGTGADLGGFGATAGAVSVLSGRVSYLF
 GLEGPAMTVDTACSSSLVALHQAASSALRQDECSLALVGGVTVMPTPQTFVEFARQRLAADGRSKAFADAADGAGFS
 EGVGVLVVERLS DARAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIQSALDNAGLSSTDVDVVEAHGTGTTL
 GDPIEAQALIAAYGQDREQPVLLGSLKSNIGHTQAAAGVSGVIKMMALQHGVPRTLHVDEPSRSHVDWTAGAVELV
 TENQPWPAIDRARRAGVSSFGISGTNAHVILESAPDLPVVETEEVPAPPVVASDL MPLVISAKTPSALADMEGRRLRA
 20 YLTATPGVDMRAVASTLAGTRSVFEHRAVLLGDDTVTGPGVAVSGPRVVFVFPQGGSQRAGMGEQLAAVFPVFAEIH
 QQVWDLDDVPDPGLDDETGFAQPSLFALQVALFGLLESWGVRPQAVIGHSVGEIAAGYVAGLWSLRDACTLVSARA
 RLMQTLPTGGAMVAVPVSEKQAQAALTDGVEIAAVNGPSSVVLTDGETAVLETAALGKSTRLLTSHAFHSARMEPV
 LDQFRTVAETLDYRTPHIPMAAGDAVVTPPEYWVRQIRDTVRFQDQVAAHENAVFVEIGPDRTLSRLTDGIAMLHGDN
 ETQTAITALATLHTHG VNIHWPTIVGTTTPVLDLPTYAFQHQRYSWLAGLAPEERKQALLKTVRDSAAAVLGHVG
 25 TDTVPATAAFKDLGLDSLTA VELRNSLGKSTGLRLPATMVFDPNP TALAAARLD

SEQ ID NO: 83

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISGFPTDRGWDIENLYDPDPALGRITYTVQGGFLDIAGFDAAFFGI
 SPREAQAMDPQQRLVLEASWEAFERAGIEPGSMRGSDTGVMGAFSSGYGAEHEGFGATAGAVSVLSGRVSYFFGLE
 30 GPALTVDACSSSLVALHQAAGYSLRQGECSLALVGGVTVMPTPQTFVEFSRQRGMVAVDGRSKAFADAADGAGWAEGV
 GVLVVERLS DARAKGHRVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANAGLSADVDVVEAHGTGTRLGDP
 IEAQAVLATYQDREQPLLLGSLKSNIGHAQAAAGVSGVIKMMALQHGVPRTLHVDSRSHVDWTAGSVELVTEN
 QPWPVLERARRVGVSSFGISGTNAHVILESAPDPDVAVVEVEETPAPPVVVISAKTPSALADMEGRRLRAYLAARPGV
 DVRVAVASTLAGTRSVFEHRAVLLGDDTVTGSTGTGSGAAVSGVAVSGPRVVFVFPQGWQWLMGCGLRETSAVFA
 35 GRLEACAAALSEFVDWDLTLVDDPSVVDVLDLPACWAVMVSLAAVWQEAGVVPDAVIGHSQGEIAAACVAGALT
 LRDAARIVALRSRLIARLAGQGAMASIALPAHEIVLGDGAVVAARNGPAATVVAGTARAVERVLAVHEKEGARVRRRI
 TVDYASHSPQVEEIRTELLDILATTGSRTPVVPWLSTVDGTWTEQPLDPDYWYRNLREPVGHFPAVDTLRSMGETVF
 IEISASPTLTAMDDATTVATLRRDNDTPQQILTALAEATHGVNIHWP AII GTTTTPARVDLPTYAFQHQRYSW
 LAGLAPEERKQALLKMVRDSAAAVLGHAGADTVPVTA AFKDLGLDSLTA VELRNSLGKSTGLRLPVTMVFDPNP TT
 40 LAARLD

SEQ ID NO: 84

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALSGFPTDRGWDIENLYDPDPDAAGTTYTVQGGFLDIAGFDASFFG
 ISPREALAMDPQQRLVLETSWEAFERAGIEPSSMRGSDTGVMGAFNGYAGVDFGAFGGASAAVSVLSGRVSYFF
 GLEGPATVDTACSSSLVALHQAGSALRQGECSLALVGGVTVMPTPQLFVDFSRQRGLAADGRSKAFADPADGAGFS
 5 EGVGVLVVERLSDAQAKGHRVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALDNARLAPNEIDAVEAHGTGTTL
 GDPIEAQALIATYQDREQPVLLGSLKSNIGHTQAAAGVSGVIKMMALQHGVPRTLHVDSPSRHDVWTAGAVELV
 TENQPWPAIDRARRAGVSSFGISGTNAHVILESAPDPDVPVETEETPPPPVVVISAKTPSALADMEGRRLRAYLAAT
 PGVDVRAVASTLAGTRSVFEHRAVLLGDDTVTGTSTGTGSGAAVSGVVVSGPRVVFVFPQGQWQLGMGCGLRETS
 VFAGRLAECAAALSEFVDWDLTLVLDLDDPSVVDVRLQVQACWAVMVSLAAVWQVQAGVVPDAVIGHSQGEIAAACVAG
 10 ALTLRDAARIVALRSRLIARLAGQGAMASIALPAHEIALGDGAVVAARNGRAATVIAGTARAVDRVLAVHEKEGARV
 RRITVDYASHSPQVEEIRTELDDILATTGSRTPVVPWLSTVDGTWTEQPLDPDYWRNLREPVGFGHPAVDTLRSMGE
 TVFIEISASPTLTPAMDDATTVATLRRDNDTPQQILTALAEAHTHGVNIHWPVIMGATTTTRVDLPTYAFQHORYWTS
 WLAGLAPEERKQALLKVVRDSAAKVLGHAGADTVPVTAAFKDLGLDSLTAVELRNSLGKSTGLRLPATMVFDPNPNT
 TLAARLD

15

SEQ ID NO: 85

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALSGFPTDRGWDIENLYDPDPDAPGKGYRVQGGFLDRAAEFDASFF
 GISPREAQAMDPQQRLVLETSWEAFERAGIEPGAMRGSDTGVMGAMANGYGTGADLGAFGMTSAAVSVLSGRVSYL
 FGLEGPAMTVDTACSSSLVALHQAAYSLRQGECSLALVGGVTVMPTPQTFVEFARQRGLAADGRSKAFSDAADGAGF
 20 SEGVGVLVVERLSDARAKGHVLAHVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALDNAGLSSTDVDVVEAHGTGTT
 LGDPIEAQALIATYQDREQPVLLGSLKSNIGHTQAAAGVSGVIKMMALQHGVPRTLHVDSPSRHDVWTAGAVEL
 VTENQPWPAIDRARRVGVSSFGISGTNAHVILESAPAQPVVETEVEVPAPPVVASDMMPLVISAKTPSALVEFEGRRLR
 AYLSTSTPGVDMRAVASTLAGTRSVFEHRAVLLGDDTVTGPDTGTGAGSGVAVSDPRVVFVFPQGQSQRAGMGEQLAA
 VFPVFAEIHQQVWDLDDVPDPGLGADETGFAPSLFALQVALFGLLESWGVRPQAVIGHSVGEIAAGYVAGLWSLRD
 25 ACTLV SARARLMQTLPTGGAMVAVPVSEKQAQAALTDGVEIAAVNGPSSVVLTGDETAVLETAALGKSTRLTTS
 FHSARMEPVLDDQFRTVAETLDYRTPHIPMAAGDAVVTPEYWVRQIRDTVRFQDQVAAHENAVFVEIGPDRTLSRLTD
 GIAMLHGDNETQTAITALATLHTHGVNIHWPVAVIGATTARVLDLPTYAFERQRYWAWLAGLAPEERKQALLKVVRDS
 AAVALGHADARDIAVNTAFRDLGLDSLTAVALQVRNSLAKATGLRLPTTTVFDPNPNTALATHLD

30 SEQ ID NO: 86

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALSDFPDTRGWDLDNLYDPDPDTPGKAHNVQGGFLDAAGFDAAFFG
 ISPREALAMDPQQRLVLETSWEAFERAGIDPASVRGSDTGVMGAFGSGYGTGADLGGFGATAGAVSVLSGRVSYLF
 GLEGPAMTVDTACSSSLVALHQASSALRQDECSLALVGGVTVMPTPQTFVEFARQRGLAADGRSKAFADAADGAGFS
 EGVGVLVVERLSDARAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALDNAGLSSTDVDVVEAHGTGTRL
 35 GDPIEAQALIATYQDREQPVLLGSLKSNIGHTQAAAGVSGVIKMMALQHGVPRTLHVDSPSRHDVWTAGAVELV
 TENQPWPAIDRARRVGVSSFGISGTNAHVILESAPAQPVVETEVEVPAPPVVASDMMPLVISAKTPSALVEFEGRRLRA
 YLTSTSTPGVDMRAVASTLAGTRSVFEHRAVLLGDDTVTGPDTGTGAGSGVAVSDPRVVFVFPQGQSQRAGMGEQLAAV
 FFPVFAEIHQQVWDLDDVPDPGLDDETGYAQPSLFALQVALFGLLESWGVRPQAVIGHSVGEIAAGYVAGLWSLRDA
 CTLVSARARLMQTLPTGGAMVAVPVSEKQAQAALTDGVEIAAVNGPSSVVLTGDETAVLETAALGRSTRLTTS
 40 HSARMEPVLDEFRTVAETLDYRTPHIPMAAGDAVVTPEYWVRQIRDTVRFQDQVAAHENAVFVEIGPDRTLSRLTDG
 IAMLHGDNETQTAITALATLHTHGVNIHWPVAVIGATTARVLDMPPTYAFQHORYWTTWLAGLTPEERKQALLKTVRDS
 AAVALGHADARDIAVNTAFRDLGLDSLTAVALQVRNSLAKATGLRLPTTTVFDPNPNTALATHLD

SEQ ID NO: 87

EPLAIVGMACRLPGGVSSPEDLWQLVESGTD AISHFPTDRGWDIDNLYDPDPDTPGKTYCVQGYFLD GIAEFDASFF
 GTSPREALAMDPQQRLVLETSWEAFERAGIDPASVRGSDTGVMGAFSSGYGTGADLGGFGATAGAGSVLSGRVSYL
 5 FGLEGPAMTVDTACSSSLVALHQAGYSLRQGECSLALVGGVTVMPTPQAFVEFSRQRGLAADGRSKAFADAADGAGW
 AEGVGVLVVERLS DARAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRSALDNARLAPNEIDVVEAHGTGTR
 LGDPIEAQALIAAYGQDREQPVLLGSLKSNIGHTQAAAGVSGVIKVMALQHGVPVPTLHVDEPSRHVDWTAGAVEL
 VTENQWPWPAIDRARRAGVSSFGISGTNAHVILESPPAQPVVETEEVPAPPVVASDMMPVLVISAKTPSALADMEGRLR
 AYLAARPGVDVRAVASTLAGTRSVFEHRAVLLGDDTVTGTSTGTGSGAAVSGVVVSGPRVVFVFPQGQWQWLGMC CG
 10 LRETSAVFAGRLAECAAALSEFVDWDLTLVLD DPSVVD RVDVLQPACWAVMVSLAAVWQEAGVVPDAVIGH SQQGEIA
 AACVAGALT LRDAARIVALRSRLIARLAGQGAMASIALPAHEIALGDGAVVAARNGPAATVIAGTPRAVDRVLAVHE
 KEGARVRRITVDYASHSPQVEEIRTEL LDILATTGSRTPVVPWLSTVDGTWTEQPLDPDYWYRNLREPVGFHPAVDT
 LRSMGETVFIEISASPTLTPAMDDATTVATLRDNDTPRQILTALAEAHTHGVNIHWPTVIGTTTTTPARVDLPTYAF
 QHQRYWTSWLAGLAPAERDEALLKMVRDSAALVLGHAGGRTIPVAAAFKDLGVDSL TAVELRNRLSAATGLRLPATL
 15 VFDYPNPAALAGWL

SEQ ID NO: 88

QPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISDFPTDRGWDIENLYDPDSPGEGEAYS AQGGFLDAAGFDAAFFG
 ISPREAQAMDPQQRLVLEASWEAFERAGIDPGAMRGSH TGVMGAMANGYGAGADLGGFGATAGAGSVLSGRISYLF
 20 GLEGPAMTVDTACSSSLVALHQAA YSLRQGECSLALVGGVTVMATTQT FVEFARQRGLAADGRSKAFADAADGAGWA
 EGVGVLVVERLS DARAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRSALDNAGLSSADVDVVEAHGTG TTL
 GDPIEAQALIAAYGQDREQPVLLGSLKSNIGHAQAAAGVSGVIKVMALQHGVPVPTLHVDSPSRHVDWTAGAVELV
 TENQWPWPVTERARRAGVSSFGISGTNAHVILESAPDPDPV PVVETEKVPAPPVVVISAKTPSALVEFEGRLRAYLAAR
 PGVDVRAVASTLAGTRSVFGHRAVLLGDDTVTGTSTGTGSGAAVSGVVVSGPRVVFVFPQGQWQWLGMCGLRETS A
 25 VFAGRLAECAAALSEFVDWDLTLVLD DPSVVD RVDVLQPACWAVMVSLAAVWQEAGVVPDAVIGH SQQGEIAAACVAG
 ALT LRDAARIVALRSRLIARLAGQGAMASIALPAHEIALGDGAVVAALNGPAATVIAGTPRAVDRVLAVHEKEGARV
 RRITVDYASHSPQVEEIRTEL LDILATTGSRTPVVPWLSTVDGTWTEQPLDPDYWYRNLREPVGFHPAVDTLRSMGE
 TVFIEISASPTLTPAMDDATTVATLRDNDTPQQILTALAEAHTHGVNIHWPTVMGATTTRVDLPTYAFQHQRYWTS
 WLAGLAPEERKQALLKVVRDSAAAVLGHAGTDTVPVTA AFKDLGLDSL TAVELRNLSLGKSTGLRLPATLVFDYPNPT
 30 TLAARLD

SEQ ID NO: 89

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISGFPTDRGWDIENLYDPDPDAPGTGYRVQGGFLDRAAEFDASFF
 GISPREALAMDPQQRLVLETSWEAFERAGIEPGSVRGSDTGVMGAFSSGYGTGADFGAFGATSAAVSVLSGRVSYF
 35 FGLEGPAMTVDTACSSSLVALHQAGSALRQGECSLALVGGVTVMSTPLTFAEFARQRGLAADGRSKAFADAADGAGF
 SEGVGVLVVERLS DARAKGHRVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRQALANAGLSSADVDVVEAHGTGTR
 LGDPIEAQAVLATY GQDREQPLLLGSLKSNIGHTQAAAGVSGVIKVMALQHGVPVPTLHVDSPSRHVDWTAGAVEL
 VTENQWPWPVLERARRAGVSSFGISGTNAHVILESAPDPDLPVVEVEETPAPVVAVISAKTPSALVEFEGRLR TYLTA
 RPGVDVRAVASTLAGTRSVFGHRAVLLGDDTVTGTGP GAAVSGVVVSGPRVVFVFPQGQWQWLGMCGLRETS AVFA
 40 GR LAEC AAALSEFVDWDLTLVLD DPSVVD RVDVLQPACWAVMVSLAAVWQEAGVVPDAVIGH SQQGEIAAACVAGALT
 LRDAARIVALRSRLIARLAGQGAMASIALPAHEIALGDGAVVAARNGPAATVIAGTPRAVDRVLAVHEKQGARVRR I
 TVDYASHSPQVEEIRTEL LDILATTGSRTPVVPWLSTVDGTWTEQPLDPDYWYRNLREPVGFHPAVDTLRSMGETVF

IEISASPTLTPAMDDATTVATLRRDNDTPQQILTALAEAHTHGVNIHWPTVMGATTTTPVRVDLPTYAFERQRYWAWL
AGLTPEERKQALLKTVRDSAAAVLGHTDARDIAMNTAFRDLGLDSLTA VQVRNSLAKATGLRLPTTTTFDYPNP TAL
ATHLD

5 SEQ ID NO: 90

EPLAIVGMACRLPGGVSSPEDLWQLVESGTD AISHFPTDRGWDIDNLYDPDPDTPGKTYCVQGYFLDGIAEFDASFF
GISPREAQAMDPQQRVLVLETSWEAFERAGIDPASVRGSDTG VFMGAFGSGYGTGADLGGFGMTAGAGSVLSGRVSYF
FGLEGPA M TVDTACSSSLVALHQASSALRQGECSLALVGGTTVLATPYGLVEISRQRGLAADGRSKAFSDAADGMGF
SEGVGVLVVERLS DARAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRSALDNAGLSSADV DVEAHGTGTR
10 LGDPIEAQAVLATYQDREQPLLLGSLKSNIGHTHAAAGVSGVIKVMALQHGVVPRTLHVDEPSRHVDWTAGAVEL
VTENQPWPAIDRARRAGVSSFGISGTNAHVILESPPAQPVVEVEETPAPPVVASDMMLVISAKTPSALADMEGR LR
AYLAARPGVDVRAVASTLAGTRSVFEHRAVLLGDDTVTGTSTGTGSGAAVSGVVVSGPRVVFVFPQGQWQLGMGCC
LRETSAVFAGRLAECAALSEFVDWDLTLVLD DPSVVD RVDVLQ PACWAVMVSLAAVWQEAGVVPDAVIGHSQGEIA
AACVAGALT LRDAARIVALRSRLIARLAGQGAMASIALPAHEIALGDGAVVAARNGRAATVIAGTARA VDRVLAVHE
15 KEGARVRIA VDYASHSPQVEEIRTEL LDILATTGSRTPVVPWLSTVDGTWTEQPLDPDYWRNLREPVG FHPAVDT
LRSMGETVFIEISASPTLTPAMDDATTVATLRRDNDTPRQILTALAEAHTHGVNIHWPTVIGTTTTTPARVDLPTYAF
QHORYWTS DRLNGRTGLEQHRVMLELVLGHAASVLGHSAPDAIAADRPFKDLGMDSLTAIELRNHLVAETGLRLPAT
TAFDHPTADDLAKRL

20 SEQ ID NO: 91

EPIAIVSMACRVPGGVDSPEGLWHLVESGTD AISDFPTNRGWDVANLYSPDPAGYTSYCVQGGFLDSAADF DATFFG
ISPREALGMDPQQRVLVLEASWEAIERAQIDPRSLRGSNVGVFVGGASQGYGASANEQQQSNAITGGSSSLLSGRVTY
ALGLEGPA V TVDTACSSSLVALHLASQSLRQRECSLALVSGVSVMATPDVFVEFSRQRGLAPDGRCKSF SASADGTG
WSEGVGVLVLERLSEATRLGHRVLAVVRGSAVN SDGASNGLTAPNGASQQRVIRQALANAGLTASQVDAVEAHGTGT
25 TLGDPIEA EALLATYQD RSTPAWLGSLKSNIGHTMAASGVLGVIKMVEAMRHG LLPRTLHVDEPSPHVDWASGDIA
LLESERPWPDGSTPRRAGVSSFGISGTNAHV VLEQYRDPAGPDTPTGSDTQTGPETTTEHG PLPLMLSARS PKALRE
QAGRLHAALVEAPWRPLDIGYSLATTRSSF AHRAVAVGSDRELLRALSQ LADGGTSPALVTATAKAGRVAF LFDGQ
GTQRLGMGSGLYERFP AFARTWDLVSAAFDKHLNHS L TDVFLGRSGSVTAE LVDDTLYAQAGIFTMEVALFELLDEW
GIRPDFLTGHSIG EAAAAAYGAGMLSLEDVT TLIVARGQALRLSPPGAMVALRASEEEVREFLDRTGAALDLAAVNSP
30 TSVVVSGAPDAVSD FRTAWTESGREARALNVRHAFHSRHEAGLGRFREVLD SLTFRAPVLPVVSTVTGR LVEPAEM
STPEYWL RQVRQTVRFHDALRELSGRGVGT FVEIGPSGTLASAGLECLGGDAAFHAVQRPRSAEDVCLMTAVAE LHA
GGTAVDWT KVLAGGRRTDLPVYPFQHEAYWLTPAEP SYAEPLTTLELVCSEANVLGITEPGILLEDSSFLDLGLD
SLGAMRLRNRLSELTELDLPATLLFDNPNPTDLAAYLD

35 SEQ ID NO: 92

EPLAIVGMAARFPGGVASADDLWRLVVSGGDAIGGFPTDRGWDLD ELYDPDPAATGRSYVREGGFLSDATTFDASFF
RIGPREAKAMDPQQRLLLETSWEAFEHAGIRPETLRGTATAVFAGISLQDYGVLAGSDPELEGYAGTGNAPSVLSGR
LSYFYGLEGPAVTIDTACSSSLVALHLAGQSLRRDECTLAVVGGVTVMPSPNVFVEFSRQRGLAPDGRCKPF AAAAD
GTGWSEGA GLVVERLS DARRNNRRILAVVRGSAVNQDGASSGLTAPHG PSQQRVIRAALAAAAGL TAGDVDVVEAHG
40 TGTTLGDP IEAQGLVATY GDRKGAPVRLGSVKS NLGHTQAAAGVAGVIKMVQALRHGVMPRSLHIDEPSPHVDWTAG
RVELLTSNLPWPASERPRRAAVSSFGISGTNAHVILEQAFFPATEPEPPFTPVVSGPELPLIFS AKDPDALAAQTRVT
DGPGVAYALATSRSMFDHRTVRLGDMTVTGI AVTDPEVVVFVFPQGQTQWAGMGRDLMEASPVFAERMNECAAALEPY

LDLWAAIDAPDHVETLQPASWAMMVSLAAVWQAAGVQPAAVIGHSQGEIAAACVSGAISLQDAAAVVALRSKAI AAS
 LGKGAMASIPLPADAIELTGEVWVAALNGPSSTVVAGVPEAVELVRARYEGRRIAVDYASHTPHVEALRGQVVSVP
 QAPVIPWFSTVDSGWVEGPLDDDYWFRNL RQPVQFGPAAAGFDNAV FIEVSARPV LIPALEASVTVPSLRDDGGPE
 RMLASLAQAFVAGVPVDWTTIVAPAPFVELPTYFPQGERYWIDPRTLDEV LAVVRDSAATVLGHTDPTAITPDRSEK
 5 DLGFDSLAAVQLRNHLLTATGVRLSATAVFDFTPVVLAGEV

SEQ ID NO: 93

EPIAVVGMACRLPGDVSSPEDLWRLVSEGRDAVGPF PADRGWEPGDAAYARVGGFVTGATGFDAGFFGISPREAQAM
 DPQQRLLLEVAWEAFERAGIAPDEL RGS DTGVFVGT YQG YGELAVDGD AEGYVGIGNSGSVVSGRVSYFFGLEGPA
 10 VTVDTACSSSLVALHQAAQALRQGECSLALVGGVTVMSSPLIFQEFARQGG LAADGRCKAFADGADGTGWGEGVGV L
 VVERLSEAQRRTHTVLAVVRGSAVNQDGASNL TAPNGPSQQRVIRAALASAGLGVGDVDLVEAHGTGTALGDP IEA
 QALLATYGS DGSPVWLGLSKNIGHTQAAAGVAGVIKAVEAMRHGVMPRTLHVDQPS SHVDWSAGAVELLTANRPWD
 SGGPRPRAAVSSFGISGTNAHVILEESPSAPVPPEPGTAPLLSARSPAALAQFESRTAGLRPSRDLASTLSRRALF
 DHRVVLPDGD TVRGGVGDAPLVFVFAGQGSQRADMGSR LAEEFPVF AAAYERVWSLLDVDESLEVDHTGFAQPALF
 15 AFEVALAELLGVRPDAVIGH SVGELAAAYVAGAMSLEDACRLVSARARLMQALPSSGGVMVSVRVSEEAARTVLRDGV
 EIAAVNGPQAVVLSGDEDAVLAAAELGEFKRLRTSHAFHSARMEPML EEFRAVASTVAFDEPQIALSFVPSAEYFV
 RQVRETVRFGEQVA AFAPGTLFVEVGPDGSLSR LTGGVSAAEPMKALAYLWVRGVGV DWTPIYIGDGR LDDAPTYPFQ
 PERYWPEQRRRARHGD FLALVTATAAVVLGHPEGTDIPADTPFQSLGLDSL SAVDLRNQLAQATGVRLSPTAVFDYP
 TPRALAERL

SEQ ID NO: 94

DPIAIVGMACRYPGGVATADDLWDLVAEGGDAVGPF PVDRGWDLAALYDPDPEAAGKSYVREGGFLGGAADF DAAFF
 GISPREALAMDPQQRLLLETAWEAFEHAGIDPLDLRRSDTGVFVGTMAQEYGG LVTD SAHGLEGWIGTGN SQSVMSG
 RLSYFFGLQGP AVTVDTACSSSLVALHQAAQALRN GECALAVVGGVTVMSSPRTFQEF SRQRGMAPDGRCKPF AAAAA
 25 DGTGWSEGVGV LVVERLSEARRNGHAVLAVVRGTAVNQDGT SNGLTAPNGPAQQQVIRAALERAGLGVGDVDVVEAH
 GTGTALGDP IEAQAILD TYGSRTGGEPVRLGSVKS NLGHTQAAAGVAGVIKMVQAMRHATMPRSLHIDEPSPHVDWA
 SGAVELLTAERGWPATDRPRAAVSSFGISGTNAHVIVEGVADPELSREASPGGPLPFVLSAPTAEALSAQETR LRR
 FRVERPDVDERDIAITLAGRTGFAHRTVLIGDLTVSGVAVADRRVVFVFPQGTQWAGMGRDLMAASPVF AERMNEC
 AAAL EPYLDLWEAIDSPDRVETLQPASWAMMVSLAAVWQAAGVQPAAVIGH SQGEIAAACVSGAISLQDAAAVVALR
 30 SKAIGASLGKGAMASIPLPADAIELIDEVWIAALNGPSSTVVAGAPEAVEQIRARYDGRRIAVDYASHTPHVEALRG
 QVVSVP SRAPAIPWFSTVDSAWVEDPLDEDYWFRNL RQPVQFGPAAAGFDNAV FIEVSARPV LIPALEDAVTVPTLR
 RDDGGIDRLHASVAQAWTAGADVDWAALLPAGGRRIALPPYAF THERFWPRRPTAAGQDLLTVVRTAAATVLGHRDA
 ARVPADRAFKE LGFDSL SAVQLRNELLTATGVRLSATAVFDHPTAAALAEAL

SEQ ID NO: 95

EPIAIVGMACRLPGDVSSPDELWELVEAGRDAVGPF PADRGWNLSTLFD PDPDAPGKSYVREGGFLT GAGLFDADFF
 GISPREALAMDPQQRLLLEMAWEAFERAGIAPDEL RGS DTGVYVGT YAQGYGELAAATAGEGFVVGIGNSGSVVSGRV
 SYFLGLEGPAVTVDTACSSSLVALHQAAQALRLGEC SMALVGGVTVMASPLMFQEF SRQRGLSPDGRCKAFAESADG
 TGWGEVGVGV LVVERLSEARRRHTVLAVVRGSAVNQDGASNL TAPNGPSQQRVIRAALASAGLTVGDVDLVEGHGT
 40 GTALGDP IEAQALLATYGSAGSPVWLGLSKNIGHTQAAAGVAGVIKMVQAMRHGVMPRTLHVDQPS SHVDWSAGAV
 ELLTANRPWDSGGPRPRAAVSSFGISGTNAHVILEGVPAPEPAAGDAETAPLVLSARTAPALTDLEARVSARPSSPD
 LAATLAGRASFDHRAVVLPDGEVVRGRAGAAPVVLV FAGQGSQRADMASRLAGEFPVF AAAYERVWSLLDVDEALDT

DQTGFAQPALFAYEVALAELLNVRPDAVIGHSIGELAAAYTSGSLSLEDACRLVSARARLMQALPPGGAMVSVRVSE
EVAREVLRDGEVIAAVNGPQAVVLSGDEDAVLA AAAKLGEFKRLRTSHAFHSARMEPMLEEFRAVALTVEFREPEVA
LSFVPSAEYFVGQVRETVRFGEQVASFEPGTLFVEVGPDGSLSRLTGGVSA AEPLTALAYLWVHGVAIDWVPYLG
RLDLGAPTYPFQHERYWPARALAQLPARRGRALLDLVQNRVAKTLGLVRPADPGRAFTDLGFTSLTALELRNSIAE
5 ETGLPMPASLVFDHPNARSLAGYLD

SEQ ID NO: 96

EPLAIVGMACRLPGGISSAEELWRLVAEGGDAIGPFPDGRGWDVDALYDPDPAAGHTYTRSGGFLPGATDFDAAFFG
ISPREAQAMDPQHRQLLETSWEALEHAGIDPAGLRGRDVGVFAGFSGQDYIAEMGVGPAEAGGYQVTGRAASVLSGR
10 LSYFYGLEGPAVTVDTACSSSLVALHLAQSLRDGESSLALVGGVTVMSSPGLFVEFSRQRGLAPDGRCKAFSADAD
GTGWSEGVGVLVVERLSDARRNGHRILAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALAQSGLSVADVVDVVEAHG
TGTALGDPIEAQAVLATYGGRAGGEPVRLGSLKSNIGHTQAAAGIASLIKMVQAIRYGVMPTLHVSEPSPLVDWAS
GRVELLTSDIPWPDGVRRAAVSAFGISGTNAHVILEEAPAPAAVPSIRPVVSGPALPLVFSARDPSALAAQTRVTDG
PGVAYALATSRTMLDHRTVRLNDVTVTGIAVTDPEVVVFVFPQGGSQWAGMGRDLMGSSPVFAERMNECAALEPYLD
15 LWAAIDAPDRVETLQPASWAVMVSLAAVWQAAGVQPAAVIGHSSQGEIAAACVSGAISLQDAAAVVALRSRAIAASLG
KGAMASIPLPADAIELADEVWVAALNGPSSSTVVAGALEAVEQVRARYEGRRIAVDYASHTPHVEALRGQVVSVP
PAIPWFSTVDSGWIEGPLDDDYWFRNLRQQVQFGPAAAGFDNAVFI EVSARPVLI PALDASVTVP SLRRDDGGPERM
LASLAQAFVAGVPVDWTTIVPPAPHVELPSYFPQRQRHWIDMERLGQLPPGDRDRFLDLVRDAAA AVLGHGSRET
PASA AFKELGFDSLIAVQLRNAVSAATGVRLPATVTFDHPTPQALAA LL

SEQ ID NO: 97

EPLAIVGMACKFPGGVDSPERLWEMLEAGEDVIGPFPDDRGRGWDVDGGYDPDPEKAGSWYARAGGFLAGAADFDAAFF
GINPREALAMDPQQRLLLEVAWEAFERSGIAPDSL RGTDTGVFVGTFGGYGRLVSAGAPGLEAYSGTGNTG SVASG
RLSYVFGLEGPAVTVDTACSSSLVALHQASQSLHRGEC SLALVAGVTVMSTPDSFVEFSRQRGLSPDGRCKAF AAAAA
25 DGTGFSEGAGVLVVERLSDAQRNHGQILAVVRGSAVNQDGASNGLTAPHGPSQQRVINTALTDADLT TTDIDLVEAH
GTGTTLGDP IEAQAILATYGNRTTGNPVHLG SVKSNLGH TQAAAGVAGVIKVIQAMRHATMPKSLHIDQPSPHVDWT
AGRVELLTGNRPWPATDRPRRAAVSSFGVSGTNAHVILEERAAAEEQPPAVDGPVPLVLSARTPEALTAQEEAVRGL
STDDRHRVAPALALGRAALPHRAVLLGDSVIRGTASADDGRPVFLFPGQGAQWAGMGRELMAASPVFAERMRECAVA
LAGFVDWDLFAVLDDAEALRRT EIVQPASWAVMVSLAALWESWGVHPAAVVGHSQGETAAAVVAGAIGLRD GARLSA
30 TRSRVLALLAGHGALASIALPAAEVEVVDGVSVAAVNGPRATLISGDPAGVEAVTARYEASGVRVRRIPADVASHSP
HVERAEETLLTALAGIEARVPGVPWLSTATGDWITEPVDERYWYRNLRSPVLFHPAITTLRDRGHR LFL EISTHPQL
LPAMEDDLLTVGSLRRDDGGPDRMHTALAEAWAGGADVDP AVLGAGPVRALDLPTYPFQRRRFWPEAALPPVERDR
ALVEIVRDQAAAVLGHPDAGALTPGTAFRDLGFDSLTA VQLRNHLATATGLTLPATVIFDHPTPRALATFLD

SEQ ID NO: 98

EPLAVVGMACRLPGGVASPDQLWDLVVSGGDGIGPFPADRGWPTDDIFDPDPDAPGKTYVREGGFLDGAGEFDAAFF
GISPREALAMDPQQRLLLETSWEAFEHAGIDPAGLRGGDTGVFVGGFTQAYGVGTADLEGYAATGTVGSVLSGRLSY
FYGFEGPAVTVDTACSSSLVALHQAGQALRQGECTLAVVGGVTVMPTPVVFQEF SRQRGLATDGRCKAFAD EADGTG
FAEGAGVLLVCRLSDARRDGRRILAVVRGSAVNQDGASNGLTAPHGPSQQRVIRAALANARLGP GDVDLIEGHGTGT
40 TLGDPIEAQALLATHGSGASPVRLGSLKSNIGHTQAAAGVAGVIKVIQALRNGLMPRTLHAGTPSSRVDWSAGNVEL
LTSNLPWPAADRPRRAAVSSFGISGTNAHVILEEAPAAA AVPTISPVVSGPALPLVFSARDPSALAAQTRVTEGPGV
AFALATTRSMFEHRAVRIGDFS VSGAAVADRRVVVFVFPQGQTQWAGMGRDLMSASPVFAERMNECAALEPYLDLWE

AIDSPDRVETLQPASWAVMVSLAAVWQAAGVQPAAVIGHSQGEIAAACVAGSITLQDAAAVVALRSRAIAASLGKGA
MASIPLPAEQIELAGEVWVAALNGPSSTVVAGLPEAVEQVRARYEGRRIAVDYASHTPHVEALRGQVVSVPSPAPAI
PWFSTVDSGWIEGPLDEDYWFRNLRQPVQFGPAAGRFDDAVFIEVSARPVLI PALEDAATVPSLRRDDGGGDRMLAS
LAQAFVAGVPVDWTTTIIPPAPFVELPSYPFQHRRYWDSSDALRDLVREQAAAVLGYPDPSRITPGVAFRDLGFDS
5 LTAVQLRNALSAATGLRLSATVAFDHPTPAALAAAL

SEQ ID NO: 99

EPIAIVGMACRLPGGVSSPDELWELVESGRDAIGPFPADRGWNLDLYDPDPDAAGRSYVREGGFLTGAADF DAGFF
GINPREALAMDPQQRLLVLEVAWEAFERAGIAPDSLRTGTGVFLGAFAGGYLTLVNGAADLEGYAGTGNSVSVLSGR
10 LSYVLGLEGPAVTVDTACSSSLVALHQAQALRLGECSLAVVGGVTVMSTPDSHVEFSRQRALSPDGRCKAFADGAD
GTGWAEGAGVLVVERLSEARRRGHTVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRAALASAGLGVGDVDLVEGHG
TGTALGDPIEAQALLATYGS DGS PVWL GSLKSNIGHTQAAAGVAGVIKAVESMRRGVMPQTLHVGT PSSHVDWAAGA
VELLTANRAWDSVERPRRAAVSSFGISGTNAHVILEGVPAPPEPAAGSAESAPLLLSARSAAALA QFESLTSGLRPSR
DLASTLSRRAFFDHRAVVLPGGDVVRGRVGDAPVVLVFA GQGSQRADMASRLTAEFPVFAAAHERVWSLLDVDEGLG
15 IDQTGFAQPALFAYEVALAELLDVRPDAVIGHSIGELAAAYVAGAVSLEDACRLVSARARLMQALPPGGAMVSVRVS
EEAARAVLRDGEIAAVNGPQAVVLSGDEDAVLAAAELGEFKRLRTSHAFHSARMEPMLDEFRAVALTVEFREPEV
ALS FVPSAEYFVRQVRET VRFGEQVA AFPGTLFVEVGPDGSLSRLTGGVSAAEPLTALAYLWVRGVGVETPYVGG
GILDQGA TYPFQ RERYWVRPRLAGRTTDERDALLIDLVRDDVASVLGHSGRRRLETD RPLLELGFDSL TALRLNR
LAAATDVALPATLIFDYPNIQAI AVLH
20

SEQ ID NO: 100

EPLAVVGMACRYPGGVASADDLWRLVAAGGDAVGPFDDRGWELES LVDPDPEAVGRSTTGQGGFLADAAGFDAAFF
GISPREATAMDPQQRLLLEVSWAALEHAGLRADALRG SATGVFMGSNGQDYAGLLAGAPELEGWIGTGV SASVSGR
LSYFYGFEGPAVTVDTACSSSLVALHLAAQSLRTGESSLALVGGVTVMTSPTVFRSFSRQRGLAPDGRCKAFSAGAD
25 GTGWSEGVGVLVVERLSDAQRNHGQILAVVRGSAVNQDGASNGLTAPHGPSQQRVINTALTDADLT TTDIDLIEAHG
TGTTLGDP IEAQAILATYGNRTTG NPNVHLG SVKSNLGH TQAAAGIAGI IKA IQAMRHATMPRTLHIDEPSPHVDWTA
GRVELLTSNLWPATGRPRRAAVSSFGVSGTNAHVILEEAPAPAAVPSIRPVVSGPALPLVFS AKDPDALAEFQ SHT
PAGEGVAYALATSRSTLDHRSVRIGDVTVTGIAVTDPEVVVFVFPQGTQWAGMGRD LMSASPVFAERMNECAAALEP
YLDLWAAIDAPDRVETLQPASWAMMVSLAAVWQAAGVQPAAVIGHSQGEIAAACVSGAISLQDAAAVVALRSKAIAA
30 SLGKGAMASIPLPADAIELTGEVWVAALNGPSSTVVAGVPEAVELVRARYEGRRIAVDYASHTPHVEALRGQVVSVP
SQAPVIPWFSTVDSGWVEGPLDDDYWFRNLRQPVQFGPAAAGFDNAVFIEVSARPVLI PALDASVTVP SLRRDDGGP
ERMLASLAQAFVAGVPVDWTTIVPPAPHVDLPSYPFQHQRFWIEGRVTAAAGAERLRIMLEVVLAE TATV LGHGGAA
AIGPGRAFQDLGFDSL TAVELRNRLAAATGLTLPTTLVFNHPTPEALAAHL

SEQ ID NO: 101

EPVAIVGMACRLPGDVESPEDLWRLVAEGRDAVGPFPSDRGWNLTLD DPDAAGRSYVKEGGFLAGAAHFDPAFFGI
GPREALGMDPQQRILLEIAWEALEHARIA PGDLRGSETGVYVGAAAQGYGV DAPLEGNLLTGGSTSAMSGRVAYALG
LHGPAVTVDTACSSSLVALHLAAQALRHGECTLALAGGVAVMASPVLFTEFSRQRGLAPDGRCKAF AAAADGTGWSE
GAGLVVLERLSDAERHGHRLAVIRGSAVNSDGASNGLTAPNGTAQRRVIRSALRAAGLGAGD VDVVEAHGTGTTLG
40 DPVEADALIATYQ RSDTPPVRI GSLKSNIGHTVAAAGVAGVIKMVEAMRHGTMPRTLHVDRPTPHVDWSAGAAELL
TGELPWPRGDRPRRAAVSAFGLSGTNAHLILEDVAAA AEPPAGDDSGSGSETVPLLLSADDLP AVRDQAARLRAHLL
AHPELMRDVAYALATTRTARPHRAAVTATERELLRELALLAAGDQGP GTQLGEAVPHRRVAF LFDGQGTQRHGMGR

ALHQRHPVFFAAWDEVCAALDPLLDRGVAEVYFAEAGRDLDADDPLYTQAGLFALEVALYRLLTSWGVTPDAVAGHSV
 GEVAVAHVAGVLSLPDAASLLAARGAALRQLPPGAMAAIRASEDDTRGVLPDLDVAAVNGPEMTVVSGAEEAVDRF
 VAEQAGAGRQVRRLRVGRAYHSRHVDAVLAIEFGATLSALTTFHEPTLPVVSTVTGRPAGAGDLTTPEYWLRRHARRPVR
 FGAALASLSELGMDSFVEVGPSGSLSSMAGETVAGTFHPLLDRRVPDEIGAAAAAGELFTAGMALDWTAVLAGGRPI
 5 DLPVYPFRREFYWL GARYDLMAAAVRRDALLDLVRVQVALLLGRADAIGVRDNTSFLDVGLDSL GASRLRNRLAAAT
 GLTLPGGGLAFDHPTPARLADHLD

SEQ ID NO: 102

EPLAIVGMACRLPGGVWSPEDLWHLVASGTDALISDFPADRGWDVEKLFDPDPDAPGKTYCVQGGFLEATAAFDAAFF
 10 GISPREALAMDPQQRLMLEVSWEAERAGIEPGSVRGSDTGVLGAYPGGYGAGAGADLGGFGATGGAGSVLSGRVS
 YFFGLEGPAMTVDTACSSSLVALHQAAYSLRQRECSLALVGGVTVMGTPHMFVDFSRQRLSVDGRCKAFADAADGT
 SWSEGVGVLLVERLSDAQAKGHRVLAVVKGSAVNQDGASNGLTAPNGPSQQRVIRQALANADLAPHEVDVVEAHGTG
 TTLGDPIEAQALLATYQGDRQPLLLGSLKSNVGHTQAAAGVSGVIKVMALQHGVVPRTLHVDEPSRHVDWSAGAV
 ELVTQNQPWPSSDRPRRAGVSSFGVSGTNAHIILESAPAQPLAPSTPITGLVPLVISAKTAPALTAFEARLRSYVTA
 15 DADLTAAIATLATRSTFEHRAVLLGDDTVTGIA TPDPVVVFVFPQGQWLGMGSALRETSVVFKERMAECAAALS
 EFVDWDLFSVLDDPAVVDRVDVVQPACWAVMVSLAATWQAAGARPDVVGHSGQEIAAACVAGAI SLQDGARVVALR
 SQLIARLAGHGAMASIALPADQITLTDGVWIAARNGPAATVIAGAPEAVDSVLAHQDARVRRITVDYASHTPHVEK
 IRDELLPMLADIDSQTPLVPWLSTVDGLFIEGPKLADYWRNLREQVGFD TAVNQLPDSIFIEVSASPVLLPGMGDA
 LTVATLRRDEGGQERLVTALAEAYVQGVAVDWAAVIYNTTALVDLPTYPFQHEHYWLDSTRLMGLAAEERDKALVAV
 20 VRESAAVVLGHADARAIPATAAFRELGVDSLTA VQLRNSLAKATGLRLPTTLAFDYPTPAVLAARL

SEQ ID NO: 103

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALISDLPTDRGWDLNLYDPDPGAPGKSYCVQGGFLDTVADFDPAFF
 GISPREALAMDPQQRLVLEVSWEAERAGIEPGSVRGSDTGVMGAFGNGYGIDTDGGGFGATAGTGSVLSGRVSYF
 25 LGLEGPAMTVDTGCSSALVALHQAARYALRQGDCLALVGGVTVMASPYTFVEFSRQRGMAADGRCKAFADAADGTGW
 AEGVGVLVLLVERLSDAEAKGHQVLAVVRGSALNQDGASNGLTAPNGPSQQRV IQAALANAGLVSADVDVVEAHGTGTT
 LGDPIEAQAVLATYQDRERPLLLGSLKSNIGHTQAAAGVSGVIKVMALQHGVVPRTLHVDEPSRHVDWSAGAVEL
 VTENQPWPSVDRPPRAGVSAFGISGTNAHVILEAVPAPPFEPPTPVTGPVPLVISAKTRPALTAFEARLRAYVTADA
 DLTAIASTLATTRSIFEHRAVLLGDDTVTGIAVPDPVVVFVFPQGQWLGMGSALRESSVVF AERMAECAAALSDY
 30 VDWDLF SVLDDPAVVDRVDVVQPACWAVMVSLAATWQAAGVRPD AVIGHSGQEIAAACVAGAI SLRDAAQIVALRSQ
 LIAGLAGHGAMASIALSADQITLTDGAWIAARNGPAATVIAGAPAAVDSVLA AHEDARTRRITVDYASHTPHVEQIR
 TELLDLTDDLDSRAPVIPWLSTVDVTWVEGPLDADYWRNLREPVGFD TAVENLPDSVFIEVSANPVLLPAMGDALT
 VATLRRDAGGQTRLLTALAEAYVQGVAVDWVTVIGATPARVDLPTYAFQHQR YWVADRLHDRPSAEQHRLMREL VQR
 HAATVLGHASPD TIAADRPFKDLGLDSL TAVELRNHLVAETGLRLSATTAFDHPTADDLAGHL
 35

SEQ ID NO: 104

EPIAIVAMACRAPGGVSSPEGLWRLVESGTDATSGFPTDRGWVDNLFDPDPDAAGKTYSVRGGFLETAADF DAAFF
 GISPREALGMDPQQRLLETSWEAIERAQIDPKSLRGRDVG VYVGAAQGYGIGATDQQQENLITGSSISLLSGRVS
 YALGLEPGVTVDTACSSSLVALHLASQALRQRECSLALVSGVSMATPDVVFVEFSRQRLAADGRCKSFSAADGT
 40 TWSEGVGVLLVQLRSEAVREGHRVLAVVRGS AVNSDGASNGLTAPNGVSQQRVIRQALAGAGLTASEVDVVEAHGTG
 TKLGDPIEAAILATYQDRDAPAWLGSLKSNIGHTMAASGV LGVIKVMQAMRHGLLPRTLHVDEPSPHVDWARGDI
 ALLTENQPWPDGTRPRRAGVSSFGLSGTNAHVLEEYPAPVAAAPPVTPARGGPLPWVLSAQSPNALREQAARLYAA

LAEDPDWHPLDIGYSLATTRPGFPHRAVAVGSDREDFQRALSKLADGAGWPGLITATAAKDRRVAFLFDGQGTQRLG
MGRGLHRRFPVFARAWDAVSAFAKHLHDHSLTDIYLGESSPTNTDLADDTLYAQAGIFTLEVALVELLQDWGVRPDF
VTGHSIGAAAAAYVAGVLSLEDVTALIVARGKALRLTPPGDMVALRAGEADVRFNLRTGAALDLAAVNSPEAVVVS
GTPDAVADFRAAWTASGGQARNLTVRHAFHSRHVESALDEFRTTLETLTFRAPKVPLVSTATGRLVGPAELGAPEYW
5 LRQVRQTVRFEDALRDLSGRGVGTVEIGPSGSLATAGLECLGDDASFHAVQRPRSPEDVCLMTAVAEHLHAGGTTVD
WAKVLAGGRTVDLPVYPFQHRPYWIAPASYPDEPRTMRELVRLEVAGILGLSDPSVILDDSSFLELGFDSLSSRLRG
NRLATVTGLDLPSTLLFDYATPAALATHLD

SEQ ID NO: 105

10 EPLAIVGMACRLPGGVSTPEDLWRLVESGVDAISDFPADRGWDVANLFDPDPAAGKTYSVRGGFLDTAADFDAAFF
GISPREALAMPQQRLVLEASWEAFERAGIEPSSVRGSDTGVMGAFSAGYGTELEGFGATAGAVSVLSGRVSYFFG
LEGPAMTVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMASQSFVEFSRQRGLSVDGRCKAFADAADGTGWAE
GVGVLVVERLSDAEAKGHRIQAMVRSSAVNQDGASNGLSAPNGPSQQRVIRQALANAGLTGADVVDVEAHGTGTTLG
DPIEAQALLATYQGDRQPLLLGSLKSNVGHTQAAAGVSGVIKVMALQHGVVPRTLHVDEPSRHVDWSAGAVELVT
15 QNQPWPSFDRPRRAGVSSFGVSGTNAHIILESAPAQPLAPSTPIPLVPLVISAKTAPALTAFEARLRDYLTAADL
TAIAATLATRSTFEHRAVLLGDDTGTGIAAPDPRVVFVFPQGQWQLGMGSALCESSVVFASRMAECAAAALSEFVD
WDLFSVLDDPAVVDRVDVVQPACWAVMVSLAATWQAAGVHPDAVVGHSSQGEIAAACLAGAISLQD GARVVALRSQLI
ARLAGHGAMASIALPADQIALTDGAWIAARNGPAATVIAGAPEAVDSVLAHGDARVRRITVDYASHTPHVEQIRAE
LLAILADIDSRPPSIPWLSTVDDALVEGPLKADYWRNLREPVGFDTAVSALQDAVFIEVSANPVLLPAMGDAATVA
20 TLRRDDGGQDRLLTAVAEAYVQGVAVDWAAVIGATGARVLDLPTYAFQHQRFWARAASAAGLAPEALLKVQDSAAQ
VLGYADPGAIAVTAFAKDLGIDSLTAVEMRNTLAKKTGLRLPATLVFDYPTPGVLAGRL

SEQ ID NO: 106

EPLAIVGMACRLPGGVSSPEDLWRLVESGGDAISDFPADRGWDIENLFDPDPAAGKTYSVRGGFLDAAAGFDAAFF
25 GISPREALAMPQQRLVLEASWEAFERAGIEPGSVRGSDTGVMGAYPGGYGTGADVGGFGATAGAVSVLSGRVSYF
FGLEGPAMTVDTACSSSLVALHQAGHALRQGECSLALVGGVTVMATPHTFIEFSRQQGLASDGRSKAFADAADGAGF
SEGVGVLVVERLSDARAKGHRVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLTGADVVDVEAHGTGTT
LGDPIEAQAVLATYQDRQKPLLLGSLKSNVGHTQAAAGVSGVIKVMALQHGVVPRTLHVDEPSRHVDWSAGAVEL
VTENQPWPSVDRPRRAGVSSFGISGTNAHVILESVPVQLPVPSAGLAPLMISAKTAPALGDAEARLRGYLTADADLP
30 AIASTLATRSMFEHRAVLLGDTTITGTAAADPKVVFVFSQGSQRAGMGEQLAFPVFADIHRRVWDLDDVPDLDVD
QTGYAQPALFALQVALAGLLESWGVRPQAVIGHSVGELAAGYVAGLWSLEDAC TLVSARARLMQALPPGGVMVAVPV
SEDQARAALLEGVEIAAVNGPSSVVLSGDETAVLQVAAGLGKWTRLSTSHAFHSARMEPMLEEFRAVAEQVYTRTPV
ITMAAGAATPDYWVRQVRDTRVFGDQVAAFEGATFVEIGPDRTLARLVDGIAMLHGDDVEEAALTGLARLFVQGV PV
AWDNGARVLDLPTYPFQHQRWLDARRAASAGGDLK MVRDNAALILGHTNPGAISSETTAFRDLGVDSLTA VQLRNS
35 LAKATGLRLPATLVFDYPTPSVLAGRL

SEQ ID NO: 107

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISDFPADRGWDVDNLFDPDPAAGKTYSVQGGFLDAAAEFDAAFF
GISPREALAMPQQRLVLEASWEAFERAGIEPGSVRGSDTGVMGAYSGGYGAGADLDGFGATAGAGSVLSGRISYF
40 FGLEGPAMTVDTACSSSLVALHQAGSALRQGECSLAVVGGVTVMATPDLFVEFSRQRGLAADGRCKAFGDAADGTGW
AEGGVLLVERLSDAEAKGHRVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIRSALATAGLAPQDQDVVDVEAHGTGTT
LGDPIEAQAVLATYQDRERPLLLGSLKSNVGHTQAAAGVSGVIKVMALQHGVVPRTLHVDEPSRHVDWSAGAVEL

VTQNQPWPDVDRPRRAAVSAFGVSGTNAHVILESVPAVPPVPSAGPAPLMISAKTAPALGDAEARLRDYLTDADLT
 AIASTLATRSTFEHRSVIFENHTITGTAAAPDRVVVFVSGQGSQRAGMGEQLAATFPVFAEIHRRVWDLDDGPDLD
 VDQTYAQPALFALQVALVGLLESWGVRPEAVIGHSVGELAAGYVSGLWSLEDAC TLVSARARLMQALPPGGVMVAV
 PVPEDQARAALVEGVEIAAVNGPSSVVLSGDEAAVLQVAAGLGKWT RLATSHAFHSARMEPML EEFRGVAEQLYRT
 5 PVISMAAEVATPDYWVRQVRD TVRFGDQVAEFEGATFVEIGPDRTLARLIDGIAM LHGDDEVEAALNGLARLFVQGV
 PVAWDNGGRVLDLPTYPFQRQRYWAVSPEALLKAVRDSAAMILGHADPSAIS ETAAFRDLGVDSL TAVELRNSLAKA
 TGLRLPATLVFDYPTPAVLAARL

SEQ ID NO: 108

10 EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISDFPTDRGWD AASLFD PDPDAAGK TYSVQGGFLDAAADF DAAFF
 GISPREALAMDPQQRLVLEASWEAFERAGIEPGSVRGTD TGVMGAFSAGYGARLEGF GATAGAVSVLSGRVSYLFG
 LEGPAMTVDTACSSSLVALHQAAYSLRQGECSLALVGGVTVMATPQIFVDFSRQRGLAPDGRCKAF GDNADGTGWAE
 GVGVLVVERLSDAQAKGHRVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLT SADVDVVEAHGTGTTLG
 DPIEAQAVLATYQDRDRPLLLGSLKSNIGHTQAASGVSGVIKMVMALQHGVVPPTLHADQPSQHVDWSTGAVELVT
 15 QSQPWPSVDRPRRAGVSSFGISGTNAHVILESVPAQPPVPSAGPAPLMISAKTAPALGEAEARLRDYLTDADLPAI
 ASTLATRSTIFEHRAVLLGDDTTITGTAAADPKVVVFVSGQGSQRAGMGEQLAFPFVADIHRRVWDLDDVPDLVDQDT
 GYAQPALFALQVALAGLLESWGVRPQAVIGHSVGELAAGYVAGLWSLEDAC TLVSARARLMQALPPGGVMVAVPVSE
 EQARAALTEGVEIAAVNGPSSVVLSGDETAVLQVAAGLGKWT RLSTSHAFHSARMEPML EEFRAVAEQLYRTPTIT
 MTEEVTTTPDYWVRQVRD TVRFGDQVAEFEGATFVEIGPDRTLARLIDGIAM LHGDDETEAALTGLARLFVQGVPTW
 20 DNKARVLDLPTYPFQRQRYWAGWLAGLAAEERDKALVTVVRDSVA AVLGYADSRKIPVSAAFKDLGVDSL TAVELRN
 SLAKTTGLRLPATLVFDHPTLATLAARL

SEQ ID NO: 109

EPLAIVGVACRLPGGVSSPEALWQLVESGTD AISGFPADRGWDV DNLFD PDPPEASGK TYCVQGGFLDAVAEFDASFF
 25 GISPREALAMDPQQRLILEVSWEAFERAGIEPGSVRG SNTGVFMGAFSGSYGSDLEGF SATAGAGSVLSGRISYFFG
 LEGPAMTIDTACSSSLVALHQAAGYALRQGECSLALVGGATVMATPQT FIEFSRQRGLAADGRCKSF GDNADGTGWSE
 GVGALLVERLSDAQAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLAPQDVDVVEAHGTGTRLG
 DPIEAQAVLATYQDRERPLLLGSLKSNVGHTQAAGGVSGVIKMVMALQHGVVPQT LHVDEPSRQVDWSAGAVELVT
 ENQPWPDVDRPRRAAVSAFGVSGTNAHVILESAPAQP VAPSTPATGLT PLVISAKTAPALTASEARLRDYLTDADL
 30 TAI AATLAATRS AFEHRAVLLGDDTVTGIAAPDRVVVFVFPQGQWQLGMGSALRDSSVVF AERMAECAAAALS DYVD
 WDLFSVLDDPAVVDRVDVVPACWAVMVSLAATWQAAGVRPDAVVGH SQGEIAAACVAGAI SLRDGAKIVALRSQLI
 ARLAGHGAMASIALPADQITLTDGVWIAARNGPAATVIAGAPEAVDSVLSAHGDARVRR IAVDYASHTPHVEQIRTE
 LLPILADIDSQTPRIPWLSTVDDTWIEGPLGADYWYRNLREQVGFD TAVEHLQDSVFIEVSASP VLLPAMGDAITVA
 TLRRDEGGQDRLVTALAEAYVQGV PVDWAAVIDNTTARVLDLPT YAFQHQRFWVANLTPEALLKAVRDSAATVLGHA
 35 DPGTIPETAAFKDLGIDSLTAVELRNSLAKTTGLRLPATLVFDYPTPGVLAARL

SEQ ID NO: 110

EPLAIVGMACRLPGGVSSPEDLWRLVESGGDAISDFPVDRGWDV DNLFD PDPDAAGK TYSVQGGFLDTAAEF DAAFF
 GISPREALAMDPQQRLVLEASWEVFERAGIEPGSVRGSD TGVMGAYPGYYGIGADLDGFGATAGAGSVLSGRVSYF
 40 FGLEGPAMTVDTACSSSLVALHQAAGSALRQGECSLALVGGVTVMATPQTYVEFSRQRGLASDGRSKAFADAADGAGF
 SEGVGVLLVERLSDARRHGHRVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIRQALATAGLSPHEVDVVEAHGTGTT
 LGDPIEAQAVLATYQDRDRPLLLGSVKSNIGHTQAAGGVSGVIKMVMALQHGVVPPTLHVDEPSRHVDWSAGAVDL

VTENRPWPDLDPRRRAGVSSFGISGTNAHVILESVPVAVPPVPSAGPAPLMISAKTAPALGEAEARLRDYLTDADL
 AIASTLASTRSTFEHRAVIFQNHTITGTAAADPRVVFVFSGQGSQRAGMGEQLAATFPVFKDIHRRVWDLDDVDPDL
 VDQTYGAQPALFALQVALFGLLESWGVRPEAVIGHSVGELAAGYVAGLWSLEDAC TLVSARARLMQALPPGGVMVAV
 AVSEEHAQAALIKGVEIAAVNGPSSSVLSGDETA VLQVAAGLGKWT RLSTSHAFHSARMEPMLEKFRAVAEQ LTYRT
 5 PVITMAAEVTT PDYWVRQVRD TVRF GDQVA AFEGATFVEIGPDRTLARLVDGIAM LHGDDEVEAALTGLARLFVQGV
 PVTW DNGGRVLDLPT YAFQRQRYWATSTRWLAGLTPQERENALLKVVRD NAAVVLGHAGAGAIPATAAFRDLGVDSL
 TAVELRNSLATTTGLRLPATMVFDYPTPAAVAARL

SEQ ID NO: 111

10 EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISDFPTDRGWDVESLFD PDPDAAGK TYSVRGGFLDAAA SFDAAFF
 GISPREALAMPQQRLVLEASWEAFERAGIEPSSVRGSDTG VFMGAFSAGYGTELEGFGVTAGAVSVLSGRVSYFFG
 LEGPAMTVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMAS PQSFVEFSRQRGLSVDGRCKAFADAADGTGWAE
 GVGVLVVERLSDAQAKGHRVLAVVRSSAVNQDGASNGLSAPNGPSQQRVIRQALANAGLTGADVDVVEAHGTGTTLG
 DPIEAQAVLATYQGDR EQPLLLGSLKSNIGHTQAAAGVSGVIKVMALQHGVVPRTLHIDEPSQHVDWSAGAVELVT
 15 QNQWPWGNDRPRRAGVSSFGVSGTNAHVILESAPTQ PALPSVTATGPVPLVISAKTAPALTAFEARLRDYLTDADL
 TAI AATLATTRATFDHRAVLLGDDTVTGVA VPEPRVVFVFPQGQWQLGMGSALSESSVFAERMAECATALDEFVD
 WDLFSVLDDPAVVDRVDVVQPACWAVMVSLAATWQAAGVHPDAVIGH SQGEIAAACVAGAI SLRDGARIVALRSQLI
 ARLAGHGAMASIALPADQITLTDGVWIAARNGPAATVIAGDPAAVDSVLA AHQDARVRRITVDYASHTPHVEQIRAE
 LLAILSDIGSQTPVIPWLSTVDGEWVEGPLGNDYWRNLRET VGFDTAVGLLPDSVFIEVSASPVLLPAMGDAVTVA
 20 TLRRDDGGLTRLLTALAEAWVQGVAVDWAIGATTARVLDLPT YAFQH QHYWAVTGTGLTPEALLKVVDSTAQVLGY
 TDAAAI AVTAAFKDLGIDSLTAVEMRNTLAKATGLRLPATLVFDYPTPSLLAGRL

SEQ ID NO: 112

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISDFPTDRGWDVESLFD PDPDAAGK TYSVRGGFLDAAAGFDAAFF
 25 GISPREALAMPQQRLVLEASWEAFERAGIEPGSVRGSDTG VFMGAYPGGYGTGVDVGGFGATAGAVSVLSGRVSYF
 FGLEGPAMTVDTACSSSLVALHQAGHALRQGECSLALVGGVTVMATPHTFIEFSRQQGLASDGRSKAFADAADGAGF
 SEGVGVLVVERLSDAQAKGHRVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIRQALADAGLV SADVDVVEAHGTGTT
 LGDPIEAQAVLATYQGDRHPLLLGSLKSNVGHTQAAAGVSGVIKVMALQHGVVPRTLHVDEPSRHVDWSAGAVNL
 VTENLPWPSLDRPRRAGVSSFGISGTNAHVILESVP AQPPVSSTGPAPLVISAKTGPALTAFEARLR TYLAAA SEVD
 30 LGAVAATLATTRSVFEHRAVLLGEETIAGTA AVDPRVVFVFSGQGSQRAGMGEQLADAF PVFADIHRRVWDLDDVDP
 LDVNQTYGAQPALFALQVALFGLLESWGVRPAAVIGH SIGELAAGYVSGLWSLEDAC TLVSARARLMQALPPGGVMV
 AVPVSEEQARGVLVEGVEIAAVNGPSSSVLSGDEAVVLQV ASGLGKWT RLSTSHAFHSARMEPMLEEFQAVAEQ LTY
 RTPAIEMAAGEEVTT PDYWVRQVRD TVRFGEQVA AFSDAVFVEVGPDRTLARLIDGVAM LHGDDEPSAALTGLATLF
 VQGV PVDWSAVVSGTEARVLDLPT YAFQH QRYWLD RKAARRAASAGGDL LKMVRGNAALILGHADPSAIAATTAFRE
 35 LGVDSL TAVQLRNSLAKATGLRLPATLVFDYPTPAVLAGRL

SEQ ID NO: 113

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISDFPPDRGWDVENLFD PDPDAPGKTYSIHGGFLDTAAEFDAAFF
 GISPREALAMPQQRLVLEASWEAFERAGIEPGSVQGS DTGVFMGAYSAGYGAGADLDGFGATAGAGSVLSGRISYF
 40 FGLEGPAMTVDTACSSSLVALHQAGSALRQGECSLALVGGVTVMATPDLFVEFSRQRGLATDGRCKAFADTADGTW
 AEGVGVLVVERLSDAQAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLAPHEVDVVEAHGTGTT
 LGDPIEAQAVLATYQGDR EQPLLLGSLKSNVGHTQAAAGVSGVIKVMALQQGVVPQTLHVDEPSQHVDWSAGAVNL

VTQNQPWPDIDRPRRAAVSAFGVSGTNAHVILESVPASPPVPSTGPAPLVISAKTVPALTAFEARLRITYLAAPVEVD
 LGAVAATLATTRATFEHRAVLLGEETIAGTAAVDPRVVFVFSGQGSQRAGMGELAAAFFPVFADIHHRVWELLDIPD
 LDVDQGTGYAQPALFALQVALFGLLESWGVRPAAVIGHSVGELAAGYVSGLWSLEDAC TLVSARARLMQALPPGGVMV
 AVPVSEEQARAVLVEGVEIAAVNGPSSVVLSGDEAVVLQVASGLGKWT RLSTSHAFHSARVEPMLEEF RVIAGQLTY
 5 RTPVIEMAAGEQVTS PDYWVRQVRD TVRFGEQVAAFSDAVFVEIGPDRTLARLIDGVALLHGDD ETEAAMAGLARLF
 VQGV PVDWSAVLGGTEARVLDLPTYAFQH QRYWAALTPEALLKVVRDS AAMVLGHADPSAISGTA AFRDLGLDSLTA
 VELRNSLAKATGLRLPATLVFDYPTPSVLAGR L

SEQ ID NO: 114

10 EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISDFPPDRGWD TASLFD PDPDAAGK TYSVQGGFLDAVAEFDAGFF
 GISPREALAMPQQRLVLEASWEAFERAGIEPGSVRGTD TGVMGAFSAGYGAHLEGF GATAGAVSVLSGRVSYLFG
 LEGPAMTVDTACSSSLVALHQAA YSLRQGECSLALVGGVTVMATPQIFVDFSRQRGLAADGRCKAFADDADGTGWAE
 GVGVLVVERLSDAQAKGHRVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLT SADVDVVEAHGTGTTLG
 DPIEAQAVLATYQGDR EQPLLLGSLKSNLGHTQAAAGVSGVIKVMALQHGIVPRTLHVDQPSQHVDWSAGAVELVT
 15 ENQPWPSLDRPRRAGVSSFGISGTNAHVILESVPASPPVPSTGPAPLVISAKTGPALTA FEARLRITYLAATPDADLP
 TIASTLATTRS VFHRAVLLGEETIAGTAAVDPRVVFVFSGQGSQRAGMGELADAF PVFADIHRRVWDLDDVPDLD
 VNQTGYAQPALFALQVALFGLLESWGVRPAAVIGHSIGELAAGYVSGLWSLEDAC TLVSARARLMQALPPGGVMVAV
 PVSEEQARGVLVEGVEIAAVNGPSSVVLSGDEAVVLQVASGLGKWT RLSTSHAFHSARMEPMLEEFQAVAEQLTYRT
 PAIEMAAGEEVTTPDYWVRQVRD TVRFGEQVAAFSDAVFVEVGPDR TLARLIDGVAMLHGDD EPSAAGTALARLHVQ
 20 GVPVDWSAVLGGTGARVLDLPTYAFQRQRYWAGWLAGLAAEERDKALVTTVVRDSVA AVLGYADSRKIALSASFKELG
 VDSLTAVELRNNLAKTTGLRLPATLVFDHPTLAAMAARL

SEQ ID NO: 115

EPLAIVGVACRLPGGVSSPEALWRLVESGTD AISGFPADRGWDVDNLFDPDPEASGKTYCVQGGFLDTVADF DASFF
 25 GISPREALAMPQQRLILEVCWEAFERAGIEPGSVRGSDTGVMGAFGSGYGS DLEGF SATAGAGSVLSGRISYFFG
 LEGPAMTVDTACSSSLVALHQAGYALRQGECSLALVGGATVMATPQT FIEFSRQRGLAADGRCKSF GDNADGTGWSE
 GVGALLVERLSDAQAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLAPHEVDVVEAHGTGTRLG
 DPIEAQAVLATYQDRERPLLLGSLKSNVGHTQAAAGVSGVIKVMALQHSVVPQTLHVDAPS RQVDWSAGAVELVT
 QNQWPWPETGRARRAAVSAFGVSGTNAHVILESAPAQPPAPSTPVTGPVPLVISAKTASALGQAEARLRITYLADKPDA
 30 DLAAIAATLATTRSTFEHRAVLLGDETIRGVAVDPDRVVFVFPQGQWQLGMGSALRESSVVF AERMAECAALSDY
 VDWDLF SVLDDLA VDRVDVVPACWAVMVSLAATWQAAGVRPDAVIGH SQGEIAAACVAGAI SLRDAAQIVALRSQ
 LIAGLAGQGAMASIALPADQITLTDGVWIAARNGLAATVIAGDPAAVDGVLA AHQDARVRRITVDYASHTPHVEQIR
 TELLDLTDDISSRTPAIPWLSTVDSTWIEGPLD TDYWRNLREPVGFD TAVNLLPDSVFIEVSASP VLLPAMGDAAT
 VATLRRDDGSQTRLLTALAEAYVQGV AIDWTIGATTARVLDLPTYAFQHQRFWANALTPEALLKVVRDS AATVLGH
 35 ADPGTIPETA AFKDLGVDSLTAVELRNSLAKATGLRLPATLVFDYPTPSVLAGR L

SEQ ID NO: 116

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISDFPADRGWDIENLFD PDPDAPGK TYSVQGGFLDTAAEFDAGFF
 GISPREALAMPQQRLVLEASWEVFERAGIEPGSVRGSDTGVMGAYPGYYGIGADLDGFGATAGAGSVLSGRVSYF
 40 FGLEGPAMTIDTACSSSLVALHQAGSALRQGECSLALVGGVTVMATPQTYVEFSRQRGLASDGRSKAFADAADGAGF
 SEGVGVLLVERLSDARRHGHRVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIGSALANAGLAPHDVDVVEAHGTGTA
 LGDPIEAQAVLATYQGDR EQPLLLGSVKSNLGHTQAAAGVSGVIKVMALQHGIVPRTLHVD EPSRHVDWSAGAVEL

VTENQPWPEHDRPRRAGVSSFGISGTNAHVILESVPAPQPPVSSSTGPAPLVISAKTASALGQAEARLRITYLTVDADLP
AIAATLATTRAFFEHRAVLLGDDTTITGVAADPRVVVFVFSGQGSQRAGMGEQLAAAFPVFADTHRRVWDLDDVPDLDDV
DQTGYAQPALFALQVALFGLLESWGVPRPEAVIGHSVGELAAGYVSVGLWSLEDACALVSARARLMQALPPGGVMVAVA
VSEEQARTALVEGVEIAAVNGPGSVVLSGDEAVVLQVASGLGKWTRLATSHAFHSARMEPMLLEEFRAVAEQLTYYRTP
5 AIEMAAGEQVTTTPDYWVRQVRDTRVRFGEQVAAFGDAVFVEIGPDRTLARLIDGVAMLHGDDDETEAAMAGLAKLFVEG
IPVDWSAVLGGNAARVDLPTYAFQQRQRYWAASLLAGLTPEERGNALLKVVRDNAAVILGHAGAAAIPATAAFRDLGV
DSLTAVELRNSLATSTGLRLPATMVFDPYTPAAMAARLD

SEQ ID NO: 117

10 EPLAIVGMACRLPGGVFSPEDLWHLVESGTDATSGFPADRGWDVEKLFDPDPDAPGKTYCVQGGFLEATAAFDAAFF
GISPREALAMDPQQRMLLEVSWEAFERAGIEPGSVRGSDTGVLGAYPGGYGAGAGTDLGGFGATGGAGSVLSGRVS
YFFGLEGPAMTVDTACSSSLVALHQAAYSLRQRECSLALVGGVTVMGTPHMFVDFSRQRGLSVDGRCKAFADAADGT
SWSEGVGVLLVERLSDAQAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLAPHEVDVVEAHGTG
TTLGDPIEAQAVLATYQDREQPLLLGSLKSNVGHTQAAAGVSGVIKVMALQRGVVPQTLHVDQPSRHVDWSAGAV
15 DLTENRPPWPDTRPRRAGVSSFGVSGTNAHVILESAPAPQPTPSTPVTGPVPLVISAKTASALGQAEARLRDYLTA
DADLTAAATLAIITRSTFEHRAVLLGDDTTITGVATPDPRVVVFVFPQGWQWLGMSALRESSVFAERMAECAALD
EFVDWDLFSVLDDPAVVDRVDVVQPACWAVMVSLAATWQAAGVRPDAVIGHSSQGEIAAACVAGAIISLRDGAKIVALR
SQLIAGLAGQGAMASIALPADQITLTDGVWIAARNGPAATVIAGTPSAVDSVLAHQDARVRRITVDYASHTPHVEQ
IRTELLGILADIDSQTPLIPWLSTMEGTWVEGPLHSDYWRNLREPVGFDTAVSLLPDSVFIEVSASPVLLPAMGDA
20 LTVATLRRDEGGQNRMTALAEAYVQGVAVDWAAVIGATTARVLDLPTYAFQHEDYWLDSTRMLGLAAEERDKALVT
VVRESAAVVLGHADARAIPVTAAFRELGVDSLTAVALRNSLAKATGLRLPTTLAFDYPTPAVLAARL

SEQ ID NO: 118

EPLAIVGMACRLPGGVLSPELWRLVESGTDATISGLPTDRGWDIDNLYDPEPGAPGKSYCVQGGFLDTVADFDPAFF
25 GISPREALAMDPQQRLLVLEVSWEAFERAGIEPGSVRGSDTGVMGAFNGYGIDTDGGGFATAGTGSVLSGRVSFY
LGLEGPAMTVDTGCSSALVALHQAARYALRQGDCLALVGGVTVMASPYTFVEFSRQRGMAANGRCKAFADAADGTGW
AEGVGVLVVERLSDAEAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQAALANAGLVSADVDVVEAHGTGT
LGDPIEAQAVLATYQDREHPLLLGSLKSNIGHTQAAAGVSGLIKVMALQHGVPQTLHVDDEPSRHVDWSAGAVEL
VTENRPPWPSVDRPRRAGVSAFGISGTNAHVILESAPPSPAPSTPVTGLVPLVISAKTAPALGQAEARLRDYLTAADV
30 LTAIAATLVTTTRSTFEHRAVLLGDDTVTGVAVPDPRVVVFVFPQGWQWLGMSALRESSVFAERMAECASALSDYV
DWDLFTVLDDPAVVDRVDVVQPACWAVMVSLAATWQAAGVRPDAVIGHSSQGEIAAACVAGAIISLRDAAQIVALRSQ
IAGLAGHGAMASIALPADQITLTDGVWIAARNGPTATVIAGNPQAVDSVLAHQDARVRRITVDYASHTPHVEQIRT
ELLDLTDDVGSRTAIPWLSTVDGEWVEGPLDTDYWRNLREPVGFDTAVGMLPDSVFIEVSASPVLLPAMGDAATV
ATLRRDDGGQTRLLTALAEAYVQGVAVDWAVGATTARVLDLPTYAFQHQRYYWADRLHDPGVEQHRLMRELVLRHA
35 ATVLGHDSFDAIAADHPFKDLGLDSLTAVELRNHLVAETGLRLSATTAFDHPTADDLARHL

SEQ ID NO: 119

EPIAIVSMACRAPGGVSSPEGLWRLVESGTDATSGFPTDRGWDVENLFDPPDAAGKTYSMRGGFLETAADFDPAFF
GISPREALGMDPQQRLLLETAWEAIERAQIDPKSLRGQDVGVYVGGAAGYGGIGATDQQQENLITGSSISLLSGRVS
40 YALGLEGPVTVDTACSSSLVALHLAGQALRQRECSLALVSGVSMATPDVFVEFSRQRGLAADGRCKSFAASADGT
TWSEGVGVLLVQLRSEAVRQGHRVLAVVRGSAVNQDGASNGLTAPNGVSQRRVIRQALASAGLAASEVDVVEAHGTG
TKLGDPIEAAILATYQDRAAPAWLGSLKSNIGHTMAASGVLGVIKMEAMRHGLLPRTLHVDDEPSSHVDWERGDV

ALLTENQWPDPSTRPRRAGVSSFGLSGTNAHVLEEYPAPAAADPPVTPAGGGPLPWVLSAQSPNALREQAARLYAA
 LAEDPDWRPLDIGYSLATTRAGFPHRAVAVGSDREEFQRALSKLADGTGWPGITATAAKDRRMAFLFDGQGTQRLG
 MGKGLHRRFPVFARAWDAVSAFAKHLHDHSLTDIYLGPSPPASAEALDDTLYAQAGIFTMEVALVELLEDWGVPRDF
 VAGHSIGEAAAAAYTAGMFSLEDVTALIVARGRALRLTPPGEMVALRGGEADVRELLQRTGAALDLAAVNSPEAVVVS
 5 GAPDAVAEFRAAWTASGRRARDLTVRHAFHSRHVESVLDEFRAATLAALTFRAPALPVVSTMTGRLADPAEMGTPEYW
 LRQVRQTVRFEEAVRELSGQGVGTVEIGPSGALATAGLECLGGDATFHAVQRPRAPEDVCLMTAVAEHAGGTAVD
 WTKILAGGRPVDLPVYPFQHRPYWIAPAPSYDPDEPRTMRELVRLLEVAGILGLSDPSVILDDSSFLELGFDSLSSMRL
 GNRLATVTGLDLPSTLLFEYATPAALATHLD

10 SEQ ID NO: 120

EPLAIVGMACRLPGGVESPDLLWRLVASGTDASISGFPRDRGWDVDNLYDPDPDAPGKTYTVLGGFLDSVAGFDASFF
 GISPREALAMPQQRLVLEVAWEAFEHAGIAPRSVRGTDGTGVFMGAFSSGYDAELEEFGMTGDAVSVLSGRVSYFFG
 LEGPAMTVDTACSSSLVALHQASSALRQGECSLALVGGVTVLATPKTFVEFSRQRGLAGDGRSKAFADAADGAGWSE
 GVGVLVVERLS DARAKGHVVLGVVRS AVNQDGASNGLSAPNGPSQQRVIRQALAGAGLSPHEVDVVEAHGTGTKLG
 15 DPIEAQAVIATYQDRDQPALLGSLKSNVGHTQAAAGVAGVIKVMALQHATVPATLHVDPATRHVDWTAGAVELVT
 ENRPWPETGRARRAAVSSFGISGTNAHVILESAPAAPEETEPVAPVVASDRVPLVISAKTPAALTSTEDRLRAYLA
 AHPGTDPRAVASTLATRSVFEHRAVLLGENTVTGVSAGADPRVVVFVFGQGWQRLGMGRELLAASPVFAGRMAECA
 TVLREFVDWDLFTMLDDPAVVDRIEVVQPVCAIMVSLAAVWQAAGVRPDAVIGHSGQGEIAAACVSGAVSLRDAARI
 VTFRSDMIARMTGHGVMAVSVALHADDIPLVEGAWVAARNGPAATVVAGTPEAVDQVLAACEERGAVRRITAGVASH
 20 TPLAEHVRGELLDATGGLPSRVPDIPWLSTVDGTWVEKPLDPAYWFRNMREPVGFAVAVDLLRAQGDHVFLEISASP
 VLLPSMDDAVTVATLRRDDGSADRMLAALAEAHTHGVVVDWPRVLGTAGRVRLPTYAFQHQRYWAVSRPAVLTPDA
 LLKVVRDSAATVLGYTDADSITVTTAFRDLGVDSLSAVELRNNLAKSTGLRLPATLVFDYPTPADLATHL

SEQ ID NO: 121

EPLAIIIGMACRLPGGITSPEDLWRLVASGSDAISDFPDDRWDVGNLYDPDPDAPGRSTTVRGGFLEDEVAGFDASFF
 GISPREALAMPQQRLVLEVSWEAFERAGIEPGTVRGS DGTGVFMGAYPGGYGVGADLGGFGTTAVSGSVLSGRVSYF
 FGLEGPAMTVDTACSSSLVALHQAGHALRQGECSLALVGGVTVMPTPNIFVEFSRQRGLAADGRCKPFADAADGTGF
 SEGAGVLLVERLSDAQTNHHILAVVRASAVNQDGASNGLTAPNGPSQQRVIRSALANAGLTTADVVDVVEAHGTGTT
 LGDPIEAQAVIATYQDRAQPVLLGSLKSNIGHTQAAAGVSGVIKVMALRNNGTVPRTLHVDEPSRHIDWTAGAVEL
 30 ATENRPWPETERPRRAGVSSFGISGTNAHVILESTPTQPVPEPSTPAAHPLPLPISAKTPPALAALEARLRLAYLTSET
 DLAAVASTLASTRAVFEHRAVLLGDETIVGVAALDPRVVVFVFSGQGSQRAGMGELAAVFPVFAQIHREVLDDLDDIP
 DLDIDQTGHAQPALFAFQVALAGLLDSWGVPRDAVIGHSIGELAAAYIAGLWSLEDAC TLVSARARLMQALPSSGGAM
 VAVQATEEQARAVLIDGVEIAAVNGPSSVVLSGDETAVLQVAAELGGKSARLKTSHAFHSARMEPMLDQFRQVAEQL
 TYRSPVIEMAAGTTSYWVRQVRD TVRFGDQVRVHQGSVLVEIGPDRTLARLIDGIATSHGDDEVRAVMTALAEHLV
 35 RGVAVDWPGTTSARVLDLPTYAFQHERYWLANTAAELTAADLLKAVRDSAAVVLGHADADSIPATTAFKDLGFDSL
 AIELRNRLAKDIGLRLPATMAFDYPTPAALAARL

SEQ ID NO: 122

EPLAIVGMACRLPGGVTSPEDLWRLVASGTDAITFPTDRGWDVGNLYDPDPDAPGKSTTVHGGFLEGVAGFDASFF
 40 GISPREALAMPQQRLVLEVSWEAFERAGIEPGAVRGS DGTGVFMGAYPGGYGVGADLGGFGTTAGAGSVLSGRVSYF
 FGLEGPAMTVDTACSSSLVALHQAGHALRQGECSLALVGGVTVMPTPNIFVEFSRQRGLSADGRCKPFADAADGTGW
 SEGVGVLLVERLS DARANGHRILAVVRGS AVNQDGASNGLTAPNGPSQQRVIRQALANAGLTTADVVDVIEAHGTGTT

LGDPIEAQAVIATYQDRTQPVLLGSLKSNIGHTQAAAGVSGVIKVMALQHDTVPA SLHVDEPSRHVDWTAGAVEL
 ATERPWPKTGRAHRAGVSSFGVSGTNAHVILESAPTQPEEPSTPAPHPLPLPVSAKTSAAALTDLEDRI RAYLTPET
 DLA AVASTLA STRAMFEHRAVLLGDETITGVAAPDPRLVFVFSGQGSQRAGMGEQLAAVFPVFAQI HREVLDLLDVP
 DLDIDQTGHAQPALFAFQVALAGLLDSWGV RPD AVIGH SIGELAAAYVAGLWSLQDACALVSARARLMQALPPGGAM
 5 VAVAVPEEQARAVLIDGVEIAAVNGPSSVVLSGDETAVLQVAAELGGKSTRLRTSHAFHSARMEPMLDQFRQVAEQ L
 TYRSPVIEMAAGTTPDYWVRQVRD TVRFGDQVRVHQGSVLVEIGPDRTLARLIDGIATSHGDDEVRAAMTALAE LHV
 RGVAVDWPGTTSARVLDLPTYAFQHRRY WVAPARRAAGRPADLTPEGLLT TVRDSAAVVLGHADASAIPATAAFQAL
 GVDSLIAVELRN NLAKNTGLRLPATLIFDYPTPVDLATHL

10 SEQ ID NO: 123

EPLAIIGMSCRLPGGVTSPE DLWRLVASGTDAITGF PADRGWDLENLYDPDPDAPGR TTTVQGGFLDDVAGFDASFF
 GISPREAVAMDPQQR LALEASWEAFERAGIEPGSVRGSDTG VFMGAYPGGYGIGADLGAFMLTGRAGSVLSGRLSYF
 FGLEGPAMTVDTACSSSLVALHQASYALRQGECSMALVGGVTVMPTPVMFVEFSRQRNLADDGRCKAFADGADGTGW
 SEGVGVLLVERLS DALAKGHRIMAVVRGSAVNQDGASNGLTAPNGPSQQRV IQSALDSAGLTTADV D VIEAHGTGTT
 15 LGDPIEAQAVIATYQDRAQPVLLGSLKSNIGHTQAAAGVSGVIKVMALQNGVVPRTLHVDEPSRHVDWTAGAVEL
 ATENRPWPEVGRARRAAVSSFGFSGTNAHVILESAPAQPATPSAPVAHLLPLPISAKTPPALADLEARLRAYLTPEA
 DLP AVASTLA STRAVFEHRAVLLGDETIVGIAALDPRVVVFVFSGQGAQRAGMGEQLAAVFPVFAQI HREVLDLLDIP
 DLDIDQTGHAQPALFAFQVALAGLLESWGV RPD AVIGH SIGELAAAYIAGLWSLEDACALVSARARLMQALPSGGAM
 VAVQATEDQARAVLIDGVEIAAVNGPSSVVLSGDETAVLQVAAGLGGKSTRLRTSHAFHSARMEPMLDQFRQVAEQ L
 20 TYRSPVIEMAAGVT PDYWVRQVRD TVRFGDQVRVHQGSVLVEIGPDRTLARLIDGIATSHGDDEVQAAMTALAE LHV
 RGVAVDWPGTTSARVLDLPTYAFQHQR YWTVSWLAGLTPEEREGALVKVVRDSAAVVLGHADAGTIPVTA AFKDLGL
 DSLTAVELRN SLARSTGLRLPATMVFDYPTLGALAARLD

SEQ ID NO: 124

EPLAIVGMACRLPGGVTSPE DLWRLVESGTD AVSAFPADRGWDADALYDPDPDPAAGKTYCVRGGFLDGVAGFDASFF
 GISPREALAMDPQQR LILEASWEAFERAGIEPGSVRGSDTG VFMGAFPGSYGV DADLGGFGMTGGAASVLSGRVSYF
 FGLEGPAMTVDTVCSSSLVALHQAGHALRQGECSLALVGGVTVMSTPDTFVEFSRQRGLAADGRCKAFGDGADGTGW
 AEGAGVLLVERLS DAQAKGHRILAVVRGSAVNQDGASNGLTAPNGPSQQRVIRSALANAGLSSADV D VVEAHGTGTK
 LGDPIEAQAVLD TYGQDRERPLLGLSLKSNVGHTQAAAGVSGVIKVMALQHGVVPRTLHADVP SRQVDWTAGAVEL
 30 VTENRSWPEADRPRAAVSSFGLSGTNAHVILESPPDQPTTASAPTTPGVPLPISAKTPAALADLETRLRAYLTPET
 DLP AVAATLAVNRSLFEHRAVLIGDDTITGTASTEPRVVVFVFPQGQWHWLGMG SALLASSAVFADRMAECNAALSEF
 VDWDLF TALDDPAVFD RVDV VQPTCWAVMVSLAAVWQHAGVRPDAVLGHSQGEIAAACFAGAI SLQDAARIVALRSR
 LIGRLAGRGAMASVSLPPDEIPLIDGVTAVLNGPSAVIAGAPDAVDAVLADCEARGARVRKINVDYASHTPHVEQI
 RTEL LDITAGITAETPTVPWLSTVDGTWIDRPLDTEYWYRNLREPVGFGATIELLQAQGD TIFIEVSASP VLLQAID
 35 DSIAIPTLRDDGTPTRLLTALAEAHVHGVTIDWAKLLGSTASPVNLP TYAFQRQRYWAASAAA GRPAELTPEHLLK
 VVRDSAAVVLGHTDAGAIPATAAFQALGVD SLIAVELRN NLAKSTGLRLPATLIFDYPTPADLATHL

SEQ ID NO: 125

EPLAIIGMACRLPGGITSPE DLWRLVESGSDAISDFPDDR GWDVDR LFD PDPDAAGKTYTTQGGFLSEVAGFDASFF
 40 GISPREAVAMDPQQR LVLEVAWEAFERAGIEPGTVRGSDTG VFMGAYPDGYGSGTDLAGFGVTAGAGSVLSGRVSYF
 FGLEGPAMTVDTACSSSLVALHQAGSALRQGECSLALVGGVTVMPTPRTFVEFSRQRGLAADGRCKPFADAADGTGF
 SEGAGMLVVERLS DAQTNGHHILAVVRASAVNQDGASNGLTAPNGPSQQRV IQSALAGAGLV SADIDVIEAHGTGTT

LGDPIEAQAVIATYQGDRSQPVLLGSLKSNIGHTQAAAGVSGVIKVMALQHDTVATLHVDEPSRHVDWTAGAV
 VTENQWPWRNGHARRAGVSSFGVSGTNAHVIEEAPAEPPVEPVPAADVVPVPLVVSARDAIPLGDQAAARLAALVEAP
 DGPVLPALADALLTRRTTFAQRAVVVAGSRDDAAAGLRALATGTAHPALVTGAAGTSGRVVLMFPGQSQWDGMSQ
 LIGASPVFAARIADCAALQPWIDWDLQDVLGRNAPTDLLEVRDQVQASFAVMVGLAAVWESVGVVRPDAVLGHSQG
 5 EIAAAYVAGALTADAAKVVAVRSRLIAARLGRGGMASVALSPQDAAARRGRAELAAVNSPASVVLGASEALDETL
 AALEADGVRVRRVAVDYASHTGHVEELEQDLAEALADVRSQAPLVGFRSTVTGEWVTEAGALDGGYWRNLRQQVRF
 GPAVAALAEDGYSVFVEASAHPLVQPVTTETLDRDADAVVTGSLRRQDGGLSRLLTSVAEVFVGGVPVDWAGLLPAGA
 GRSWVDLPTYAFDHQHYWLPAGGTRGRSEAELELVRGAAAVLGHTDAGSIPATAAFKDLGLDSLTAVELRNSLAK
 STGLRLPATMVFDPYPTPAVAARL

10

SEQ ID NO: 126

EPLAIVGMACRLPGGITSPEDLWRLVASGSDAISDLPVDRGWTVDGHFQGGFLDEVAGFDASFFGISPREAVAMDPQ
 QRLVLEVAWEAFERAGIEPGSVRGTDAGVFMGAYADGYGMGTDLGGFGMTSVAVSVLAGRISYFFGLEGPAMTVDTA
 CSSSLVALHQAGHALRQGECSLALVGGVTVMPTPQTFVEFDRQRGLAADGRCKAFADAADGTSFSEGAGMLVVERLS
 15 DALANGHHILAVVRGSAVNQDGASNGLTAPNGPSQQRVIRSALANAGLTADVDVVEAHGTGTTLGDPPIEAQAVIAT
 YGQNRQRPLLLGSLKSNIGHTQAAAGVSGVIKVMALRNNGTVPATLHVDEPSRHIDWTAGAVALTENQWPETERP
 RRAGVSSFGISGTNAHVILESTPTPPATLSAQVAHPLPLPISAKTPPALADLEARLRAYLTPEADLAAVASTLASTR
 AVFEHRAVLLGDETIVGVAALDPRVVFVFSGQGSQRAGMGEQLAAVFPVFAQIHREVLDDLDDIPDLDDIDQTGHAQPA
 LFAFQVALAGLLDSWGVVRPDAVIGHSIGELAAAYVAGLWSLQDACALVSARARLMQALPSGGAMVAVAVPEDEARAV
 20 LIDGVEIAAVNGPSSVVLSGDETAVLQVAESLGGKSARLKTSHAFHSARMEPMLDQFRQVAEQLYRSPVIEMTAGV
 TPDYVWRQVRDTRVFRGQVRVHQGSVLVEIGPDRTLARLIDGIATSHGDDEVQAVMTALAEHLHVRGVAVDWPGTSSA
 RVLDLPTYAFQHDHYWAHPVDRTPEALLALVRDSAAVALGHAGAATVPATAAFQSLGMDSLIAVELRNNLARSTGLR
 LPATLVFDYPTPAALATRL

25 SEQ ID NO: 127

EPLAIVGMACRLPGGVTSPEDLWRLVASGSDAITGLPTDRGWEEDDRFRGGFLAGVAGFDASFFGISPREAVAMDPQ
 QRLVLEASWEAFERAGIEPGSVRGSDTGVMGAYPGGYGFGADLGGFALTSGSGSVLSGRVSYFFGLEGPAMTVDTA
 CSSSLVALHQAGYALRQGECSLALVGGVTIMPTPQTFIEFERQRGLAADGRSKAFADSADGTGWSEGVGLVVERLS
 DAQANGHHILAVVRGSAINQDGASNGLTAPNGPSQQRVIRSALANAGLTADIDVIEAHGTGTTLGDPPIEAQAVIAT
 30 YGQDRSQPVLLGSLKSNIGHTQAAAGVSGVIKVMALQHDTVATLHVDRPSRHVDWAAGAVELVTENRWPENGRV
 RRAGVSFAFGVSGTNAHVILESPPDQPVKPSAPAAGPVPLPISAKTPAALAALENRLRAYLTPETDLPAVASTLATTR
 AMFEHRAVLLGDDTITGTASTEPVVVFVFPQGQWHWLGMSALLASSAVFADRMAECNAALHEFVDWDLFTALDDPA
 VFDRVDIVQPTCWAVMMSLAALWQHAGVRPDAVLGHSQGEIAAACFAGAI SLQDAARIVALRSQ LIGRLAGRGAMAS
 VSLPPDEIPLIDGVTVAVLNGPSAVIAGSPEAVDAVLADCEARGARVRKINVDYASHTPHVEQIRTELLHITAATA
 35 ETPTVPWLSTVDGTWIDHPLDTEYWYRNLRREPVGFGATIELLQQTQGDITFIEVSASPVLLQAIDDSIAIPTLRDDG
 TPTRLLTALAEAHVNGVTIDWATVLGATGSPVDLPTYAFQHQRFWVGDRHLHGRTSAEQHRIMLDLVLGHATSVLGHQ
 TPDAVASDRAFKDLGMDSLTAVELRNHLVAETGLRLPATTAFDHPTADDLARRL

SEQ ID NO: 128

40 EPIAIVSMACRAPGCVTSPEGLWRLVESGSDAIADFPADRGWDLATLYSPDPIGYTSYCLQGGFLDAAADFDAFFG
 ISPREALGMDPQQRLLETSWEAIERARIDPRSLRGRDVGYYVGATQGYGVGAVDQQRDNVITGSSISLLSGRLSY
 ALGLEPGVTVDTACSSSLVALHLASQALRQRECSMALVSGVSVIPTPDVFVEFSRQRGLASDGRCKSFSAADGTI

WAEGVGVVLVLERLSEATRLGHEVLAVIRGS AVNSD GASNGLTAPNGASQQRVIRQALASAGLNAADVDTVEAHGTGT
 KLGDPIEAEAILATYQGDRSSPVWLGLSKSNIGH SMAASGVLGVIKMVEAMRHARLPRTLHVDEPSPHVDWASGDVA
 LLTENQWPWDGARPRRAGVSSFGLSGTNAHVVLEQHRAPAVPVAAETVADDVPLPLLLSARHPKALRDQAARLHAAL
 AEAPGWRPLDVGYSLATTRSFAHRAVAVGSGRELLRALAKLAEGAAPALVTGTAKAGRVAF LFDGQGTQRLGMGR
 5 VLHDRFPVFARAWDTV SARFDQHLDHSLTDVYLGRDTSAAALADDTLYAQAGIFTMEVALFELLAEWGVRPDLVSGH
 SIGEVAAAYAAGLFSLEDAATLIVARGRALRQMPPGAMLALRASEDQVRELLDRTGADLDVAAVNSPVSVVSGDPD
 AVAAFRAEWEASERDARALNVHHAFFHSRRVDAVLDEFRAVLGTLTFRTPALPVVSTVTGRLAGPAEMSTPEYWLRI
 RRTVRFQDAVRELSGQGAGTFVEIGPSGALAAAGLECVDA SFHAVQRPRSPEDACLLTAVAE LHAGGTAVDWAKVLA
 GGRATDLPVYPFQHETYWIPPASPPADTRTMEVVHEEAALVLGVTDPRVILDDSSFLDLGFDLSAMRLGNQLSAV
 10 TGLDLPPSLLFEHPTVGELAAHLD

SEQ ID NO: 129

EPLAIVGMAARFPGGVASADDLWRLVVS GGDAIGGF PADRGWDLEELYDPDPAATGRSYVREGGF LNDATTFDASFF
 RIGPREAKAMDPQQRLLLETSWEAFEHAGIRPETLRGTATGVFAGISLQDYGVLAGSDPELEGYAGTGNAPSVLSGR
 15 LSYFYGLEGP AVTIDTACSSSLVALHLAGQSLRRDECTLAVVGGVTVMPSPNVFVEFSRQRGLAPDGRCKPFAAAAD
 GTGWSEGVGVLVVERLS DARRNKRILAVVRGS AVNQDGASSGLTAPNGPSQQRVIRSALAAAGL TAGDVDVVEAHG
 TGTTLGDPIEAQGV LATYGDRSGAPVRLGSVKS NLGHTQAAAGIAGVIKMVQALRHGVMPRSLHIDEPSPHVDWTAG
 RVELLTSNLPWPTSERPRRAAVSSFGISGTNAHVILEQAFFPATEPEPSFTPVVSGPALPLVFSARD SGALATRTHLS
 DGPGVAYALATSRSMFDHRSVRIGDMT VTGVATTDEPVVFVFPQGQTQWAGMGRALMDASPVFAERMNECAALEPY
 20 LDLWEAIDTPDQVETLQPASWAVMVSLAAVWQAAGVRPAAVIGH S QGEIAAACVAGSLSLADAAA VVALRSKAIAAS
 LGKGAMASIPLPVEEIELIDEVWVAALNGPSSTVVAGAPDAVEQVRARYDGRRIAVDYASHTPHVEALRGQVVS VPS
 QAPDIPWFSTVDSEWVEGPLDDDYWFRNL RQPVQFGPAAARFDDAVFEVSARPV LIPALDASVTVP SLRRDDGGPE
 RMLASLAQAFVAGVAVDWTTIVPPAPFVDLPTYPFQGERFWIDLDDVLAVVRDCAATVLGHTDPAAIAPDRPFKDLG
 FDSLAAVQLRNHLLTVTGVRLSATAVDFDPTPAVLAGEV
 25

SEQ ID NO: 130

EPIAVVGMACRLPGDVASPEDLWRLVAEGRDAVGPF PADRGWELGEAAAYARVGGFVTGATGFDAGFFGISPREALAM
 DPQQRLLLEVAWEAFERAGIAPDALRGSDTG VFGTYGQGYGELAVDGD AEGYVGIGNSG SVVSGRVSYFFGLEGPA
 VTVDTACSSSLVALHQAAQALRQGECSLALVGGVTVMSSPLIFQEFARQGG LAADGRCKAFADGADGTGWGEGVGV L
 30 VVERLSEAQRGHTVLAVVRGS AVNQDGASNGLTAPNGPSQQRVIRSALASAGLGFGDVDLVEAHGTGTALGDPIEA
 QALLATYGSAGTPVWLGLSKSNIGHTQAAAGVAGVIKAVEAMRHGVLPQTLHADQPSSHVDWTAGAVELLTANRPWD
 SAGRPRRAAVSSFGISGTNAHVILEEFSSAPVSPEPGAGAAPLLLSARSAAALAEFESRVAALRPSRDLAATLAGRV
 FFDHRAVVLPGGEVVRGRVGDAPVVVFV FAGQGSQRSDMASRLAGEFPLFAAAHERVWSLLDVDES LDVDQTFGAQPA
 LFAYEVALAE LLGVRPDAVIGH SVGELAAAYVAGALSLEDACRLVSARARLMQALPPGGVMVSVRVSEEAARAVLRD
 35 GVELAAVNGPRAVVLSGDEGAVLAAA AELGEFRRLRTSHAFHSALMEPML EEFRAVASSVEFGEPEIALSFVPSADY
 FVRQVRET VRFGEQVA AFEPGTLFVEVGPDGSLSR LTGGVNAAEPLTALAHLWAHGAVVDWTPYTS DGRLD TAPTYP
 FQPERYWPEQRRRRARRGDSLALVIATAA AVLGHPEGTDIPADTPFQSLGFDLSAVDLRNQLAHATGVRLSPTAVF
 DHPTPRALAERL

40 SEQ ID NO: 131

DPIAIVGMACRYPGGVATADDLWDLVAEGGDAVGPF PADRGWDLAGLYDPDPEAAGKSYVREGGF LGGAADFDAAFF
 GISPREALAMDPQQRLLLETAWEAFEHAGIDPLDLRRSDTG VFGTMAQEYGG LVTD SAHGLEGWIGTGN SQSVMSG

RLSYFFGLQGPVAVTIDTACSSSLVALHQAAQALRSGECSLAVVGGVTVMSSPRTFQEF SRQRGMADGRCKPFAAAA
 DGTGWSEGVGVLVVERLSEARRNGHAVLAVVRGTAVNQDGT SNGLTAPNGPAQQQVIRAALERAGLGVGDVDVVEAH
 GTGTALGDPIEAQAAILDTYGSRTTGEPVRLGSKVSNLGHGTQAAAGVAGVIKMQAMRHATMPRSLHIDEPSPHVDWA
 SGAVELLTAERGWPATDRPRRAAVSSFGISGTNAHVIVEGVTEPEPSREAAAPSGPLPLMLSAPTAEALAEQETRLRR
 5 FRADRPDADERDIAVTLAGRTGFAHRTVLIGELSVSGVAVADRRVVFVFPQGTQWAGMGRDLMDASPVFAERMNEC
 AAALPYLDLWEAIDTPDRVETLQPASWAVMVSLAAVWQAAGVRPAAVIGHSQGEIAAACVAGSLSLADAAAVALR
 SKAIAASLGKGAMASIPLPAEEIELIDEVWVAALNGPSSSTVVAGAPDAVEQVRARYDGRRIAVDYASHTPHVEALRG
 QVVSVPSPQAPDIPWFSTVDSGWVEGPLDDDYWFRNLRQPVQFGPAAARFDDAVFIEVSARPVLIPVLEDAVTVPTRLR
 RDDGGIGRLHASVAQAWTAGADVDWAALLPAGGRRIALPPYAFTHERFWPRRPAAAGQDLLTVVRTAAATVTLGHRDA
 10 ARVPADRAFKELGFDLSAVQLRNELLTATGVRLSATAVFDHPTAAALAEAL

SEQ ID NO: 132

EPIAIVGMACRLPGDVSSPDELWELVESGRDAIGPFPADRGWNLSTLFDPPDPAPGKSYVREGGFLT GAGLFDADFF
 GISPREALAMDPQQRLLLEVAWEAFERAGIAPDALRGSDTG VYVGTYAQGYGELAAATAGEGFVGIGNSGSVVSGRV
 15 SYFLGLEGPAVTVDTACSSSLVALHQAAQALRLGEC SLALVGGVTVMASPLMFQEF SRQRGLSPDGRCKAFAEGADG
 TGWGEAGVGVLVVERLSEARRRGHTVLAVVRGS AVNQDGASNGLTAPNGPSQQRVIRAALAAAAGLTFGDVDVVEGHGT
 GTALGDPIEAQALLATYGAAGSPVRLGSLKSNIGHTQAAAGVAGVIKMQAMRHGVMPRTLHVDQPSSHVDWSAGAV
 ELLTANRTWEAPGRPRRAAVSSFGISGTNAHVILEGVPAPPEPAAGSAETAPLLLSARTVPALNDFEARVSARPSSPD
 LAATLSRRVFFDHRAVVLPGGEVVRGRVGDAPVVVFV FAGQGSQRADMASRLAGEFPVFAAAHERVWSLLDVDEGLAV
 20 DQTGLAQPALFAYEVALAELLGVRPDAVIGH SVGELAAAYVAGALSLEDACRLVSARARLMQALPPGGVMVSVRVSE
 EAARAVLRDGEVIAAVNGPRAVVLSGDEDAVLA AAAELGEFRRLRTSHAFHSARMEPMLEEFRAVASSVVFGEPEIA
 MSFVPSADYFVRQVRETVRFGEQVASFDPGSLFVEV GPDGSLSRLTGGVSAAEPMKALAYLWVRGVGDWAPYVGGG
 RLDLGAPTYPFQREGFWPTREALAQLPARRGRALLDL VQNRVAKTLGLVRPADPGRAFTDLGFTSLTALELRNSIA
 EETGLPLPASLVFDHPNARALAAAYLD

25

SEQ ID NO: 133

EPLAIVGMACRLPGGISSAEELWRLVAEGGDAIGPFP GDRGWDIDALYDPDPDAAGRITYTRSGGFLPGAGDFDAAFF
 GISPREAQAMDPQHRQLLETSWEALEHAGIDPAGLRGRDVG VFAGFSGQDYIAEMGVGPAAEAGGYQVTGRAASVLSG
 RLSYFYGFEGPAVTVDTACSSSLVALHLAGQSLRDGE SSSLALVGGVTVMSSPGLFVEFSRQRGLAPDGRCKAFSVDA
 30 DGTGWSEGVGVLVVERLSDARRNNHQILAVVRGS AVNQDGASNGLTAPNGPSQQRVIRQALAQSGLSVGDVDVVEAH
 GTGTALGDPIEAQAVLATYGSRTGGEVRLGSLKSNIGHTQAAAGIASLIKMQSIRYGVMPRTLHVSEPSPLVDWA
 AGRVELLTSDVPWPEGVRRAAVSAFGISGTNAHVILEE APAPAEAVPSIRPVVSGPELPLVFSARDADALAAQSRLT
 DGPVGAHALVTARTVFDHRSVRMGDVTVTGVATPDPE VVFVFPQGTQWPGMGRDLMAASPVFADRMNECALALSPY
 LDLWAAIDAPDRVETLQPASWAVMVSLAAVWQAAGV QPAAVIGHSQGEIAAACVAGSLSLADAAAVALRSRAIASL
 35 AGKGAMASIPLPAEEIELVDEVWVAALNGPSSSTV VAGTPDAVEQIRSRDGRRIAVDYASHTPHVEALRGQVVSVP
 QSPAVPWFSTVDSAWVEGPLDEDYWFRNLRQPVQFGP AAAGFDNAVFEVSARPVLIPALDASVTVP SLRRDDGGPE
 RMLASLAQAFVAGVAVDWTTIVPPAPHVDLPTYPFR RQRHWIDMERLGQLPPGDRDRFLDLVRDAAAVALGHGSR
 TVPASAAFKELGFDLSIAVQLRNAVAATGVSLPATVTF DHPTPQALAVLL

40 SEQ ID NO: 134

EPLAIVGMACKFPGGVDSPERLWEMVEAGEDVIGPFP DDRGWDVDGGYDPDPEKAGSWYARAGGFLAGAADFDAAFF
 GINPREALAMDPQQRLLLEVAWEAFERSGIAPDSL RGTDTGVFVGTFTGQGYGRLVAAGAPGLEAYSGTGNTG SVASG

RLSYVFGLEGPAVTVDTCSSSLVALHQAGRSLQSGECSLALVAGVTVMSTPDSFVEFSRQRGLSPDGRCKAFAAAA
 DGTGFSEAGAGVLVVERLSDARRNNHQILALVRGSAVNQDGASNGLTAPNGPSQQRVITAALTDARLTTTDDIDLVEAH
 GTGTTLGDPIDAEQAILATYGNRTTGNPVHLGVSXSNLGHGTQAAAGIAGVIKAIQAIRHTTMPKSLHIDQPSPHVDWT
 SGRVELLTSNQWPATDRPRRAAVSSFGVSGTNAHVILEEQTPVEEPPPASAGPVPLALSARTPEALTAQEKAVERGL
 5 PDGDRRRAAPALALGRAALPHRAVLLGDSVIRGTASADDGRPVFLFPGQGAQWAGMGRELMAASPVFAERMRECAVA
 LAGFVDWDLFAVLDDAEALRRTIIVQPASWAMMVSLAALWESWGVPRPAAVVGHSSQGETAAAVVAGAIGLRDGRARLSA
 TRSRVLALLAGHGALASIALPAGEVEEVVDGVSVAAVNGPRATLISGDPAGVEAVTARYEASGVRVRRIPADVASHSP
 HVERAEETLLAALAGIEARVPGVPWLSTATGDWITEPVDERYWYRNLRSPLFHPAITTLRDRGHRFLFLEISTHPQL
 LPAMEDDLLTVGSLRRDDGDLDRMHAALAEAWAAGADVWRAFLGSGPVRALDLPTYPFQRRRFWPEAGALPPAERE
 10 RALVEIVRDQAAAVLGDPDAGALTPGTAFRDLGFDLSLTAVQLRNHLATATGLTLPATVIFDHPTPRALATFLD

SEQ ID NO: 135

EPLAVVGMACRLPGGVSSPDQLWDLVVSGGDGIGPFPGRGWATDEIYDPDPDASGKTYVREGGF LDSAGDFDAAFF
 GISPREALAMDPQQRLLETSWEAFEHAGIDPAGLRGGDTGVFVGGFTQAYGVGTADLEGYAATGTVGSVLSGRLSY
 15 FYGFEGPAVTIDTACSSSLVALHQAGQALRQGECTLAVVGGVTVMPTPVVFQEF SRQRGLAADGRCKAFAD EADGTG
 FAEGAGVLLVCR LSDARRDGRRILAVVRGSAVNQDGASNGLTAPNGPSQQRVIRAALASARLGPDVDLIEGHGTGT
 TLGDPIEAQALLATHGSGASPVRLGSLKSNIGHTQAAAGVAGVIKVIQALRHGLMPRTLHVGT PSSHVQVDSAGNVEL
 LTSNLPWPATDRPRRAAVSSFGISGTNAHVILEEAPAPA AVPSITPVVSGPALPLVFSARD SGALAARTRLTDGPGV
 AFALATSRSMFDHRAVRIGDLSVSGVAVADRRVVFVFPQGTQWAGMGRALMDASPVFAERMNECAAALSPYLDLWE
 20 AIDAPDRVETLQPASWAVMVSLAAVWQAVGVEPAAVIGHSSQGEIAAACVAGSISLPDAAAVALRSKAIASLAGKGA
 MASIPLPDQIDLIDQVWIAALNGPSSTVVAGSPEAVEQVRARYDGRRIAVDYASHTPHVEALRGQVVSVP SQAPDI
 PWFSTVDSAWVEKPLDGDYWFRNL RQPVQFGPAAARFDDAVFIEVSARPVLI PALDTSVTVPSLRDDGGPERMLAS
 LAQAFVAGVAVDWT TIVPPAPFVELPTYPFQRRRYWIDSSEEALRDLVREQAAAVLGYPDPSRITPGVAFRDLGFD S
 LTAVQLRNALSAATGLRLSATVAFDHPTPAALAAAL

25

SEQ ID NO: 136

EPIAIVGMACRLPGDVSSPDELWDLVESGRDAIGPFPADRGWNLD ELYDPDPDATGRSYVREGGF LAGAADF DAEFF
 GINPREALAMDPQQRLVLEVAWEAFERAGIAPDSL RGTDTGVFLGAFAGGYLTLVNGAADLEGYAGTGNSVSVLSGR
 LSYVLGLEGPAVTVDTCSSSLVALHQAQALRQGECSLAVAGGVTVMSTPDSHVEFSRQRALSPDGR CRAFTADGAD
 30 GTGWSEAGAGVLVVERLSEARRRGHTVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIRSALASAGLFGD VDLVEGHG
 TGTALGDPIEAQALLATYGSAGTPVWL GSLKSNIGHTQAAAGVAGVIKAVEAMRRGVMPRTLHVDAPSSHVEWSSGS
 VELLTANRPWDGVGRPRRAAVSSFGISGTNAHVILEGVPAP EPAGTGQAPLLLSARSVSALAEFESRIAGLVP SRDL
 AATLAGRAFFDHRAVILPDGDVVRGRAGGAPLVFVFAGQGSQRADMASRLAEFFPAFAAAHERVWSLLDVDEGLDVD
 QTGLAQPALFAYEVALAELLGVRPDAVIGH SIGELAAAYVAGALSLEDACRLVSARARLMQDLPSGGAMVSVRVSEE
 35 AARAVLRDGEIAAVNGPQAI VLSGDEDAVLAAAAELGEFRRLRTSHAFHSGRMEPML EEFRLVASSVVFREPEIAM
 SFVPSADYFVRQVRET VRFGEQVASFDAGAVFVEVGPDGSLSR LTGGVSAAEPLTALAYLWVRGVGV DWAPYVGGGR
 LDLGAPTYPFQRRERYWVRPRLAGRTTDERDALLISLVRDDVASVLGHPDRRRRLATDRPLLELGFDSL TALRLRNRLA
 AATDIALPATLIFDYPNIQAI AVHL

40 SEQ ID NO: 137

EPLAVVGMACRYPGGVASADDLWRLVTAGGDAIGPFPDDR GWELES LVDPDPEAVGRSTTGQGGFLADAAGFDAAFF
 GISPREATAMDPQQRLLEVSWAAL EHA GLRADALRG SATGVFMGSNGQDYAGLLAGAPELEGWIGTVGSASVVSGR

LSYFYGFEGPAVTVDACSSSLVALHLAAQSLRTGESSLALVGGVTVMTSPTVFRSFSRQRLAPDGRCKAFSAGAD
GTGWSEGVGVLVVERLSDARRNNHQILALVRGSAVNQDGASNGLTAPNGPSQQRVITAALTDARLTTTDLVEAHG
TGTTLGDPIDAEQAILATYGNRTTGNPVHLGSKSNLGHGTQAAAGIAGVIKAIQAIRHTTMPKSLHIDQPSPHVDWTS
GRVELLTSNQWPATDRPRRAAVSSFGVSGTNAHVILEEAPAPAEAVPPIRPVVS GPALPLVFSARDSGALATRHL
5 SDGPGVAYALATSRSMFDHRSVRIGDMTIVTG VATTDPEVVVFVFPQGGTQWAGMGRALMDASPVFAERMNECAAALEP
YLDLWAAIDAPDQVETLQPASWAVMVSLAAVWQAAGVRPAAVIGHSQGEIAAACVAGSITLQDAAAVALRSKAI
SLGKGAMASIPLPVEEIELIDEVWVAALNGPSSTVVAGAPDAVEQVRARYDGRRIAVDYASHTPHVEALRGQVVSVP
SQTPAVPWFSTVDSEWVEGQLDDDYWFRNLRQPVQFGPAAARFDDAVFIEVSARPVLPALDASVTVP SLRRDDGGP
ERMLASLAQAFVAGVAVDWTIVPPAPFVDLPTYPFQHERFWIEGRVAAATGAERPRILLEVVLAEATATVLGHGGAA
10 AIGPDRAFQDLGFDLSLTAVELRNRLAAATALTLP TTLVFNHPTPEALAAHL

SEQ ID NO: 138

ELVAIVGMACRLPGDVASPEDLWRLVAEGRDAVGPF PADRGWNLTLD DPDAAGRSYVKEGGFLAGAAHFDPGFFGI
GPREALGMDPQQRILLEIAWESLERARIAPGSLRGSETGVYVGAAGQYGV DAPLEGNLLTGGSTSAMSGRVAYS LG
15 LHGPAVTIDTACSSSLVALHLAAQALRN GECTLALAGGVAVMASPVLFTEFSRQRLAPDGRCKAFAAAAADGTGWSE
GAGLVVLERLSDAERHGHFVLAVIRGSAVN SDGASNGLTAPNGTAQRRVIRSALRAAGLGAGD VDVVEAHGTGTTLG
DPVEADALIATYQRP GMPVVRIGSLKSNIGHTVAAAGVAGVIKMEAMRHDTMPRTLHVDRPTPHVDWSAGAAELL
TGEQPWPRGDRPRRAAVSAFGLSGTNAHLILEDVAPGAASGAEP PGADETVP LLLSADDLPAVRDQAARL RAYLLA
RPELRMRDVAYALATTRTARPHRAAVAATEREFLRELALLAAGDQGP GTQLGEAVPHRRVAF LFDDGQGTQRHGMGRA
20 LHQRHPVFAAAWDEVCAALDPLLGRGVADVYFAEAGRD LADDPLYTQAGLFALEVALYRLLTSWGVTPDAVAGHSVG
EVAAAHVAGVLSLPDAAALLAARGAALRRLPAGAMAAIRASEADTRAVLPDLDVA AVNGPEMTVVSGAPDAVDRFI
AEQAGAGRQVRRLRVGRAYHSRHVDAVLA EFGATLSALTFHEPVLPVVS TVTGRPAGAGDLTTPEYWL RHARRPVRF
GAALAALSELGMDSFVEVGP SGSLSSMAGETVAGTFHPMLDRRVPDEI AVAAGELFTAGMVL DWAAVLAGGRTIDL P
VYPFRREFYWL GARRYDLMAAAERRDALLDLVRVQVALLLGRADAIGVRDNTSFLDVGLDSL GASRLRNRLAAATGL
25 TLPGGVAFDHPTPARLADHL

SEQ ID NO: 139

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDATSGFPVDRGWADSSMRGGFLDAAADF DAAFFGISPREALAMPQQ
RLVLEASWEAFERAGIEPGSVRGSDTG VFMGAFSGGYGAGADLGGFGVTAGAVSVLSGRVSYFFGLEGP AVTVDTAC
30 SSSLVALHQAGHALRQGECSLALVGGVTVMSTPDIFAEFSRQGGLASDGRCKAFADTADGT SWSEGVGVLVVERLSD
ARAKGHQVLAVVRGSAVNQDGASNGLTAPNGPSQQRV IQAALTHAGLTAEVDVVEAHGTGTTLGDPIDAEQAVIATY
GRDRERPVLGSLKSNVGHTQAAAGVSGVIKVMALQHGVVPRTLHVDEPSRHVDW TAGAVRLATESQPWPDTGRPR
RAAVSSFGVSGTNAHVILEGVAEEPAQSEESSELVPLVISAKTPAALTRLEERLRAYLTAESNLSAVASTLAETRSL
FEHRAVLLGDDTIKGTAPNPRVVFVFSGQGSQRAGMGDELA AAFPVFAKIRQQVWDLLDVPDLEVNDTGHAQPALF
35 ALQVALFGLLESWGV RPQALIGHSIGELAAGYVSGIWSLEDACTLV SARARLMQSLPPGGAMVAVPVSEQQARAVLT
DGVEIAAVNGPSSVVL SGDEEAVLRAAAALDGRSKRLVTSHAFHSARMEPMLDEFRAVAEQ LTYRAPRIPMAVGEGP
EYVWRQVRET VRFGEQVA AHGAVFVELGPEGTLARLIDGVAVLDREDEPRAALTALGKLHVRGVRVDWPLTSGRRV
DLPTYAFQ RERYWATALTPAEREQALLKLVRDSAAVVLGYTDAVPVSGSFKDLGIDSLTAVELRN SLATTTGLRLPA
TLVFDYPTPATLAARL

40

SEQ ID NO: 140

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AIAFPPTDRGWDVEALFDPDPDAAGKSYCVRGGFLDG VADF DASFF

GISPREALAMDPQQRLILEASWEAFERAGIDPADARGSDTGVFMGAFTSGYGADLEGFGGTAGALSVLSGRVSYFFG
LEGPAVTVDTACSSSLFALHQAGYALRQGECSMALVGGVTVMATPRTFVEFSRQRGLASDGRCKAFGDTADGTGWAE
GVGVLVVERLSDAQAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIQAALHNARLTPADVVEAHGTGTTLG
DPIEAQAVIAAYGQGRDEPLLLGSLKSNVGHTQAAAGVSGVIKVMALQHGVVPRTLHVDEPSRHVDWTAGEVRLVT
5 ENQSWPDTGRPRRAGVSAFGVSGTNAHVILEGPPTQPPATAPQEPAPLVISAKTPAALADYEGRLRAYLAATPGTDA
RALAVTRSLFEHRAVLLGDDTISGAAVTDPRVVVFVFPQGQWQLGMGVALRDSVVFVAERMTECAAALSEFVDWDLF
AVLDDPAVVDRVDVVQPACWAVMVSLAAVWQAAGVHPDAVVGHSSQGEIAAACVAGAISLRDAARVVALRSRLIGERL
GQGAMASVTLPADEISLVDGVWIAAYNGPASTVIAGSPDAVDQMGDRVRRIAVDYASHSPQVEQIKDELDDITADV
GSRTPTVPWFSTVDGSGWIEGPLDADYWRNLQRPVGFHPAVEALRALGETVFVEVSASPVLLPAMDDALTVATLRRD
10 DGTIARMHTALAEAHVHGVNVDWAAVLGVAARHVDLPTYAFQRQRFWADERELASLGAEREQALRKLVS DTAAGVL
GYADPGAVPIKAAFRELGVDSL TAVELRNGLAKATGMRLPATMVFDYPTPHALAAARL

SEQ ID NO: 141

EPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISEFPADRGWDVENLYDPDPDAAGKSYCVRGGFLDAAA EFDASFF
15 GISPREALAMDPQQRLILEASWEAFERAGIEPGSVRGSDTGVFMGA FSGGYGADVEGFGATAGAGSVLSGRVSYFFG
LEGPAITVDTACSSSLVALHQAGYSLRQGECSLALVGGATVMAKPQSFVEFSRQRGLAADGRCKAFADAADGTGWAE
GVGVLLVERLSDAERNHQLAVVRSSAVNQDGASNGLSAPNGPSQQRVIRQALANARLTAADVVEAHGTGTTLG
DPIEAQALIAAYGQDREWPLLLGSLKSNVGHTQAAAGVSGVIKVMALQHGVVPRTLHVDEPSRHVDWTAGAVRLVT
DNQPWPETGRPRRAGVSSFGVSGTNAHVILESPPTQPSGTFKKPAHEPQPLIISAKTPAALADYEDRLGAYLTAAPG
20 VDVPAVAATLAVTRSLFEHRAVLLGDNTVTGTAITDPRVVVFVFPQGQWQLGMGAALRGSSVVFVAERMTECAAALSE
FVDWDLFAVLDDPAVVDRVDVVQPACWAVMVSLAAVWQAAGVHPDAVVGHSSQGEIAAACVASAVSLRDAARVVALRS
RLISERLGQGAMASVALPADQIVLADGVWIAAHNGPTSTVVAGSPDAVEQMLGDRVRKIAVDYASHTPHVEQIKTEL
LGITAGIGSRTPTVPWFSTVDGSGWIEGPLDADYWRNLQRPVGFDAAVGRRLRALGATVFVEVSASPVLLPAMDDAVT
IATLRRDEGSITRMHTALAEAHVLGVNVDWPTLLGDTDRRALDLPTYAFQRQRYWGDAAGLAPAEREQALLKLVRDS
25 AALVLGYAGGDAVPATDAFKDLGIDSLTAVELRNGLAKATGLRLPATLVFDYPTPQVLAARL

SEQ ID NO: 142

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDATSGFPVDRGWADSSMRGGFLDAAADF DAAFFGISPREALAMDPQQ
RLVLEASWEAFERAGIEPGSVRGSDIGVFMGAYPGGYGIGADLAGFGATAGAGSVLSGRVSYFFGLEGP AVTVDTAC
30 SSSLVALHQAGHALRLGEC SLALVGGVTVMATPDTFVEFARQGGGLASDGRSKAFADSADGAGFSEGVGVLLVERLS D
AQRHGHRLAVVRGSAVNQDGASNGLTAPNGPSQQRVIQAALAHAGLAPHEVDVVEAHGTGTTLGDP IEAQAVIATY
GQDRDEPLLLGSVKS NVGHTQAAAGVAGVIKVMALRHGVVPQTLHVDEPSRHVDWTAGAVRLLTEKQPWPSTDRPR
RAGVSSFGISGTNAHVILEGVAE EPAQSEDSSSELVPLVISAKTPAALTQVEERLRAYLTAESNLSAVASTLAETRSL
FEHRAVLLDGHAVRGVAESNPRVVVFVSGQGSQRAGMGDELA AAFPVFAKIRGQVWDLLDVPDLVDNDTGH AQPALF
35 ALQVALFGLLESWGVRPHALIGHSIGELAAGYVSGIWSLEDACALVSARARLMQALPPGGAMVAVPVSEQQARAVLT
DGVEVA AVNGPSSAVLSGDEEAVLRAAAALGGRWKRLATSHAFHSARMEPMLDEFRAVAEQ LTYRAPRIPMAVGEGP
EYVWRQVRET VRFGEQVA AHGAVFVELGPDGSLARLIDGIATLDRDDEPRVALTALAE LHVGRVDVDWPLTSGRRV
DLPTYAFQRQRYWIDRAGRTPAEREQALLKVVRDSAATVLGHADGGSVGAAAAFKDLGVDSL TAVELRN SLAKATGL
RLPATLVFDYPTPA AVAVRL

40

SEQ ID NO: 143

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDATSGFPTDRGWADSSMRGGFLVAAADF DAAFFGISPREALAMDPQQ

RLVLEASWEAFERAGIEPGTVRGSDTGVFMGAYPGGYGIGADLAGFGATAGAGSVLSGRVSYFFGLEGPAVTVDTAC
 SSSLVALHQAGHALRQGECSLALVGGVTVMATPDLFVEFARQGGGLASDGRCKAFGDTADGTGWAEGVGVLVERLSD
 AQAKGHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIIQAALHNARLTPADVDDVEAHGTGTTLGDPIDIAQAVIAAY
 GQGRDEPLLLGSLKSNVGHTQAAAGVSGVIKVMALQHGVVPRTLHVDEPSSHVDWTAGAVRLVTENQSWPDTGRPR
 5 RAAVSAFGVSGTNAHVILESSAAPSPTIPQPPSAEPMPLVISAKTPAALADYEGRLRAYLTAPGVDVPAVAATLAVT
 RSLFEHRAVLLGGNTVTGTAVADPRVVVFVFPQGQWQLGMGAALRGSSVVF AERMTECAAALSEFVDWDLFAVLDDP
 AVVDRVDVVQPACWAVMVSLAAVWQAAGVHPDAVVGHSSQGEIAAACVAGALSRLDAARVVALRSRLIGERLGRGAMA
 SVSLPADQIVLADGVWIAAHNGPASTVIAGGAGAVDQMVGERVRRIAVDYASHTPDVEQIQTELLDITADVGSQAPV
 VPWFSTVDGVWVDGPLDRDYWRNLQRPVGFHPAVEALQALGETVFVEVSASPVLLPAMDDAVTVATLRRDEGSITR
 10 MHTALAEAHVLGVNVWDPTVVGDTDRRTLDLPTYAFQHRYWISAAARLDGLTAAEKHSLLLDIVLANAATVLGHHT
 VDTIAPDKPFKDLGIDSLTAVELRNGLAKATGLRLPATLVFDYPTPDMAAARL

SEQ ID NO: 144

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALISDFPTDRGWADAAGAPYSPQGGFVDAAADFDAFFGISPREALA
 15 MDPQQRLVLEASWEAFERAGIEPGTVRGSDTGVFMGAYPGGYGIGADQAGFGTTAGAGSVLSGRVSYFFGLEGPAVT
 VDTACSSSLVALHQAGHALRQGECSLALVGGVTVMGTPDIFAEFSRQGGGLASDGRCKAFGDDADGTGWGEGVGILLV
 ERLSDAQHRSHRVLAVVRGSAVNQDGASNGLTAPNGPSQQRVIIQAALAHAGLAPHEVDVVEAHGTGTTLGDPIDIAQA
 VIATYGQDRDEPLLLGSLKSNVGHTQAAAGVAGVIKVMALRHGVVPRTLHADQPSRHVDWTAGAVRLATENQWPWA
 IDRPRRAGVSSFGISGTNAHVILEGVAEPAQSEESSPLMPLVISAKTPAALTRLEERLRLAYLAAPETSLGAVAST
 20 LAETRSLFEHRAVLLNGDVVRGVAEPNPRVVVFVFSGQGSQRAGMGDEVAFAFPVFAKIRRVWDLLDVPDLVDNDTG
 HAQPALFALQVALFGLLESWGVPRDALIGHSIGELAAGYVSGIWSLEDACALVSARARLMQALPAGGAMVAVPVSEQ
 QARAVLTDGVEIAAVNGPSSVVLSGDEEAVLRAAAGLGSRWKRLATSHAFHSARMEPMLDEFVRVVAEQLSYKTPRIP
 VAVGEGPEYWVRQVRETVRFGEHVAHDGAVFVELGPDGSLARLIDGIATLDRDDEPRAALTALAEHLVRGVDVDWP
 LTSGRRVDLPTYAFQRQRYWTTAGLTRAEREQALLKLVRDTAAVVLGYGDGNAVPTAAFKDLGVDSLTAVELRNGL
 25 AEAIGLRLPATLVFDYPTPATLAVRL

SEQ ID NO: 145

EPLAIVGMACRLPGGVESPDDLWRLVESGTDALITGFPTDRGWPDVTGTSHSQHGGLHTAADFDAFFGISPREALA
 MDPQQRLILEASWEAFERAGINPADAHGTDGTGVFMGAFSAGYDADRDDSPATAGAVSVLSGRISYFFGLEGPAMTVD
 30 TACSSSLVALHQAGYSLRQGECSMALVGGVTVMATPRTFVEFSRQGGGLASDGRCKAFGDTADGTGWSEGVGLVVER
 LSDARAKGHRVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIIQAALHNAHLTPADVDDVEAHGTGTTLGDPIDIAQAVI
 AAYGQDRDEPLLLGSLKSNVGHTQAAAGVSGVIKVMALRHGVVPRTLHADQPSRHVDWNAGAVQLVTENQSWPETG
 RPRRAAVSSFGISGTNAHVILEGVPEQPAQPEPPSERVPLMISAKSTSALSQLEDRLRAYLAARPEASLGAVASTLA
 TRSLFEHRAVLLDGQVVKGVAEPNPRVVVFVFSGQGSQRAGMGDELAFAFPVFAKIRGQVWDLLDVPDLVDNDTGHAQ
 35 PALFALQVALFGLLESWGVPRDALIGHSIGELAAGYVSGIWSLEDACTLV SARARLMQALPAGGAMVAVPVSEQQAR
 AVLTGGVEIAAVNGPSSVVLSGDEGAVLRAAAALGGRWKRLATSHAFHSARMEPMLDEFRAAAEQLTQYQTPRIPMVV
 GDGPDYVWRQVRETVRFGEQVAHDGAVFVELGPDRLARLIDGIATLDRDDEPRAALTALAEHLVRGVDVDWPHDG
 QLVLDLPTYAFQRERYWATALAALPLAEREQALLAVVSDNAAVVLGYAEGRDVTQTAAFKDLGVDSLTAVELRNTLAK
 ATGLRLPATIVFDYPTPDTLAARL
 40

SEQ ID NO: 146

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALISRFPPDRGWDVEGLFDPDPDAPGKSYVEGGFLDAVADFDAFF

GISPREALAMDPQQRLVLEASWEAFERAGIEPGSVRGSDTGVMGAYSGGYGIGADLPGLGVTAGAVSVVSGRVSYF
 FGLEGPAVTVDTACSSSLVALHQAGHALRRRECSLALVGGVTVMATPFGFVEFSRQRGLAADGRCKAFADTADGTSW
 SEGVGVLVVERLS DARANGHQVLAVVRGSAVNQDGASNGLTAPNGPSQQRV IQAALAHADLAPHEVDVVEAHGTGTR
 LGDPIEAQAVIATYQGGRDEPLLLGSIKSNVGHTQAAAGVSGVIKMMALRHGVVPQTLHVDEPTQHVDWTAGAVRL
 5 ATENQWPWDTGRPRRAGVSSFGVSGTNAHVILEGVAEPAQSEESSELVPLVISAKTPAALTRLEERLRAYLSAESN
 LSAVASTLAETRSLFEHRAVLLGDDTIKGTAPNPRVVFVFSGQGSQRAGMGDELA AAFV FARIRRQVWDL LDVPD
 VSVDDTGFAQPALFALQVALFGLLESWGVRPDALIGHSIGELAAGYVSGIWSLEDACTLV SARARLMQALPAGGAMV
 AVPVSEQQARAVLTGGVEIAAVNGPSSVVLSGDEEAVLRAAAALGGRSKRLVTSHAFHSARMEPMLDEFQAVAEQLT
 YQAPRIPMAVGDPDYWVRQVRETVRFGDQVAAQDGA VFVELGPDRSLARLIDGIATLDRDDEPRAALTALAEHLVR
 10 GVDVDWPLTSGRRVDLPTYAFQRQRYWIDSALT PAEREQALLKVVRDSAAVVLGYTDAVPVSGSFKDLGIDSLTAVE
 LRNSLAKVTGLRLPATLVFDYPTPATLAARL

SEQ ID NO: 147

EPLAIVGMACRLPGGVESPDDLWRLVESGTDAITGFPTDRGWP DVTGTSHSQHGGFLHTAADFDAAFFGISPREALA
 15 MDPQQRLILEASWEAFERAGINPADAHGTD TGVMGAYSGGYGIGADLAGFGATSGATSVLSGRVSYFFGLEGPAIT
 VDTACSSSLVALHQAGHALRQGE CAMALVGGVTVMATPDIFVEFSRQRGLAADGRCKAFADAADGTGWAEGVGVLLV
 ERLSDAERNGHRVLAVVRSSAVNQDGASNGLTAPNGPSQQRV IQAALHNARLTPADV DVEAHGTGTT LGDPIEAQA
 VIAAYGQGRDEPLLLGSLKSNVGHTQAAAGVSGVIKMMALQHGVVPRTLHVDEPSSHVDWTAGAVRLATENQSWPD
 TGRPRRAAVSAFGVSGTNAHVILESSAAPSPTIPQPPSAEPMPLVISAKTPAALADYEDRLRAYLTNP GVDVPAVAA
 20 TLAMTRSLFEHRAVLLGGNTVTGTAVADPRVVFVFPQGQWQLMGGAALRGSSSVFAERMTECAAAALSEFVDWDLFA
 VLDDPAVVD RVDVVQPACWAVMVSLAAVWQAAGVHPDAVLGHSQGEIAAACVAGAI SLQDAARVVALRSQAISGLSG
 KGAMASIALPADQIALPDGAWIAAHNGPASTVVAGSPDAVEQMLGDRVRKIAVDYASHTPHVEQIQTELLDITAGIG
 SRTPTIPWFSTVDGMWVDGPLDRDYWYRNLRQPVG FHPAVEALQALGETV FVEVSASP VLLPAMDDAVTVATLRRDE
 GSITRMHTALAEAHVLGVNVDWPTLLGDTGRRTLDLPTYAFQH HRYWINGSRLIGRTTAEQHRMLLAFVLGNVASVL
 25 GHGSADAIAADKPFKDLGMSLTSVELRNSLAKATELRLPATIVFDHPTADALAAHL

SEQ ID NO: 148

EPIAIVSMACRVPGGVTSPEGLWRLVESGTD AISAFPGDRGWDIANLYSPDPDAPGKSYSVQGGFLDGAAAFDASFF
 GISPREALGMDPQQRVLLETAWEAVERARIDPRSLRGRDVG VYVGAAQGYGLGAAEAHRDNLITGGSISLLSGRLS
 30 YALGLEGPGLTVDTACSSSLVALHLAAQALRQGECSLALVSGVSVMP TPDVFVEFSRQRGLASDGRCKSFAASADGT
 SWSEGVGVLVLERLSEARRLGHQVLAVVRGTAVNSDGASNGLTAPNGAAQQRVIRQALANAGLSTADVDAVEAHGTG
 TTLGDPIEA EAILATY GKDRSTPVWLGSLKSNIGHTMAASGV LGVIKMEAMRHGVLPRTLHVDEPSPHVDWAAGEV
 ALLTENQTPGDVRPRRAGVSSFGLSGTNAHVLEQDEAPAAPVTTKESG PLPWVLSAQSPKALRQRAGQLATALAE
 DSTWHPLDVAYSLATTRSDFAHRAVVVGADRELLRTL GKVADGAGWPGLTTGTAKARRVAF LFDGQGTQRLTMGQGL
 35 YGSFPAFARAWDTVSAEFGKHL DHPLADVYFDGSGGAATADLVDDPLYAQAGIFAVEVALVELLAEWGVRPDVVTGH
 SIGEAAAAYTAGMLSLSDVTTLIVARGAALRSAPPGAMLALRAGEQEVRNFDGTGAALDLAAVNGPAAVVVSGAPD
 AVTDFASAWTASGREARRLKVRRAFHSRHVEGVLD DFRTALESLSFRTPLLPVVSTVTGRLIDPAEMGTPEYWLDQV
 RQPVRFQEAVQELAGQGVGT FVEVGPSGTLASAGMECLDGDASFHALLRPSAEDVGVL TALAE LHAGGTAIDWPTV
 LAGGRPMDLPVYPFQHQS YWL VSTDEPRTTLELVHLEVARVLGITDPD TVLDDASFLELGFDSLGGVRLRNRLAQVT
 40 GLTLPPTLLFDHVT PAALAAELD

SEQ ID NO: 149

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALSGFPADRGWDVENLYDPDPEAAGKSYCVQGGFLDSAGGFDASFF
 GISPREALAMDPQQRLVLEASWEAFERAGIEPGSLRGSDTGVMGAYPPGGYGVGADLGGFGATAGAVSVLSGRVSYF
 FGLEGPAVTVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPQSFVEFSRQRGLASDGRCKAFADAADGTGW
 5 AEGAGVLLVERLSDAQAKGHQVLAVVRGSAVNQDGASNGLSAPNGPSQQRVIRAALSNAGLSTAEDDVVEAHGTGTT
 LGDPIEAQALLATYQGDRQPLLLGSVKSNLGHGTQAAAGVSGVIKVMALQHGLVPRTLHVDEPSRHVDWTDGAVAL
 VTENQWPWMDGRPRRAGVSSFGISGTNAHVILESAPPTQAVDDVPPAEAPVVASELVPLVISARTLPALVEYEDRLR
 AYLAASPGVDVRGVASTLAVTRSVFEHRAVLLGDDTVTGTTVSDPRVVFVFPQGGSQRAGMGEELAAAFPVFARIHQ
 QVWGLLDVPDLEVNETGYAQPALFALQVALFGLLESWGVPRDAVVGHSGVGLAAGYVSGLSLEDACTLV SARARLM
 10 QALPPGGVMVAVPVSEDEARAVLGEGVEIAAVNGPSSVVLSGDETAVLQAAAAALGKSTRLATSHAFHSARMEPMLEE
 FRTVAERLTYQTPRLAMAAGDRVTTAEYWVRQVRDTRVRFGEQVASYEDAVFIELGADRSLARLVDGVAMLHTDHEAQ
 AAISALAHLYVNGVTVDWTALLGDAPATRVLDPTYAFQHQRWLEGWLAALAPEERAKALLKVVRDTAATVLGHADA
 RTIPVTGAFRDLGIDSLTAVELRNGLAKVTGLRLPATLVFDYPTPAVLAARL

15 SEQ ID NO: 150

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDALSGFPADRGWDAESLFDPDPAVGKSYCVEGGFLDSAASF DAGFFG
 ISPREALAMDPQQRLIMEVSWEAFERAGIEPGSVRGSDTGVMGAYAGGYGAGADLGGFAATASATSVLSGRVSYFF
 GLEGPATITVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPQSFVEFSRQRGLASDGRCKAFADSADGTGWA
 EGVGVLLVERLSDAQAKGHQVLAVVRSSAVNQDGASNGLSAPNGPSQQGVIQAALSNAGLAAHEVDVVEAHGTGTTL
 20 GDPIEAQAVIATYQGDRERPLLLGSLKSNIGHAQAAAGVSGVIKVMALQHNTVPRTLHVDEPSRHVDWAAGAVELV
 RENQWPWPGTDRPRRAGVSSFGVSGTNAHVILESAPPAQPAEEAQPVETPVVASDVLPLVISAKTQPALTEHEDRLRA
 YLAASPGVDTRAVASTLAVTRSVFEHRAVLLGDDTVTGTAVSDPRVVFVFPQGWQWLGMSALRDSSIVFAERMAE
 CAAALREFVDWDLFTVLDDPAVVDRVDVVQPASWAMMVSLAAVWQAAGVRPDAVIGHSQGEIAAACVAGAVSMRDAA
 RIVTLRSQAIAGLAGRGAMASVALPAQDVELVDGAWIAAHNGPASTVIAGTPEAVDHVLTAEHARGVRVRRITVDYA
 25 SHTPHVELIRDELDDITSDSSSQAPVVPWLSTVDGSWVDSPLDVEYWYRNLRPVGFGHPAVGQLQAQGDTVFVEVSA
 SPVLLQAMDDDDVVTATLRRDDGDATRLMTALAQAAYVHGVTVDWPAILGTTTTTRVDLPTYAFQHQRWVEWLAALAP
 AEREKALLKVVCDSAAVVLGHADARTIPVTGAFKDLGVDSL TAVELRNSLVKATGLRLPATMVF DYPTPTALAAARLD

SEQ ID NO: 151

EPLAIVGMACRLPGGVSSPEDLWRLVESGTDVAVSGFPTDRGWDVEGLFDPDPDAAGKSYRAEGGFLDTAAGFDAGFF
 GISPREALAMDPQQRLLEVSWEAFERAGIEPGSVRGSDTGVF IGAFPVGYGAGAAAREGYGATAAPNVLSGRLSYFF
 GLEGPATITMDTACSSSLVALHLAAQALRNGECSMALAGGVTVMATPEVFTTEFARQRGLASDGRCKAFADSADGAGFS
 EGAGLLLVERLSDARRNGHQVLAVVRGSAVNQDGASNGFTAPNGPAQQRVIRQALANAGLTAEVDVVEAHGTGTTL
 GDPIEAQAVIATYQGDRQPLLLGTLKSNVGHGTQAAAGVSGVIKVMALQHSTVPRTLHVNEPSRHVDWSAGAVELV
 35 TENQSWPVTGRPRRAGVSFAFVSGTNAHVVLLESAPPAQSVNNAQPVATPVVASELVPLVISAKTLPALTEHEDRLRA
 YLAASPGADMRAVGSTLALTRSVFEHRAVLLGHDTVTVTGTGTAVSNPRVVFVFPQGWQWLGMSALRGSSVVF AE
 RMAECAALSEFVDWDLFAVLDDPAVVDRVDVVQPASWAVMVSLAAVWQAAGVRPDAVIGHSQGEIAAACVAGAVSL
 RDAARIVTLRSQAIAGLAGRGAMASVALPAHEIELVDGAWIAAHNGPASTVVAGAPEAVDRVLAVHEARGVRVRRIA
 VDYASHTPHVELIRDELDDITAGIGSQAPVVPWLSTVDGTWVEGPLDVEYWYRNLRPVGFD SAVGQLRAEGD TVFV
 40 EVSASPVLLQAMDDDDVVTATLRRDDGDATRLMTALAQA FVEGVTVDWPAILGTATTRVDLPTYAFQHQRFWAEGWL
 ARLAPVEREKALLKLVCDGAATVLGHADASTIPATAAFKDLGIDSLTAVELRNSLT KATGLRLPATLVFDYPTPTAL
 AARL

SEQ ID NO: 152

EPLAIVGMACRLPGGVSSPEDLWRLLESGETDAVSGFPTDRGWDVENLFGPAAGDSYRLQGGFLDAAAGFDASFFGIS
PREALAMPQQRLVLEVSWEAFERAGIEPGSVRGTDGTFMGAYPGGYGIGADLGGFGATASAVSVLSGRVSYFFGL
5 EGPATVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPQTFVEFARQGGLAGDGRSKAFADSADGAGFSEG
VGVLVERLSDAQAKGHQVLAMLRSSAVNQDGASNGLTAPNGPSQQRVIAAALSNAAGLAAHEVDVVEAHGTGTTLGD
PIEAQALLATYGQDREQPLLLGSVKSNLGHTQAAAGVSGVIKMMALQRGFVPRTLHVDEPSRHVDWSAGAVALVTE
NQWPDMGRARRAGVSSFGISGTNAHVILESAPPTQPADNAVIERAPEWLPMVISARTQSALTEHEGRLRAYLAASP
GVDMAVASTLAMTRSVFEHRAVLLGDDTVTGTAATDPRVVFVFPQGGSQRAGMGEEELAAAFPVFARIHQQVWDLDD
10 VPDLEVNETGYAQPALFALQVALFGLLESWGVRPDAVVGHSVGELAAAGYVSGLWSLEDAC TLVSARARLMQALPAGG
VMVAVPVSEDEARAVLGEGVEIAAVNGPSSVVLSGDEAAVLQAAEGLGKWTRLATSHAFHSARMEPMLLEEFRTVAEG
LTYRTPQVSMAAGDQVTTTEYWVRQVRD TVRFGEQVASYEDAVFVELGADRSLARLVDGVAMLHGDHEAQAAVSALA
HLYVNGVTVDWPALLGDAPATRVLDPTYAFQHORYWLEGRWLAALAPEERAKALVKVVCDSAATVLGHADVDSIPVT
AAFRDLGVDSLTAVELRNSLT KATGLRLPATLVFDYPTPGALAARL

15

SEQ ID NO: 153

EPLAIVGMACRLPGGVSSPEDLWRLVESGETDAISGFPTDRGWDVENLSDPDAAGKSYCVEGGFLATAANFDASFFGI
SPREALAMPQQRLVLEVSWEAFERAGIEPGSVRGSDTGTFMGAFPGGYGIGADLEGYGATAGLNVLSGRLSYFFGL
EGPAVTVDACSSSLVALHQAGYALRQGECSLALIGGVTVMATPHTFVEFSRQRLASDGRCKAFADSADGTGWSEG
20 VGVLVERLSDAQAKGHQVLAVVRSSAVNQDGASNGLSAPNGPSQQRVIRQALANAGLTAEVDVVEAHGTGTTLGD
PIEAQAVIATYGQDRDQPVLLGSVKS NVGHTQAAAGVSGVIKMMALQHGLVPRTLHVDEPSRHVDWTDGAVELVTE
NQSWPEAGRPRRAGVSSFGVSGTNAHVILESAPPTQAVDDVRPADAPVVASVMASELVPLVISAKTQSALAEYEGRL
RAYLAASPGVDMRAVASTLAMTRSVFEHRAVIVGDDTVSGTAATDPRVVFVFPQGGSQRAGMGAEELAAAFPVFARIH
QQVWDLDDVPDLEVNETGYAQPALFALQVALFGLLESWGVRPDAVIGH SVGELAAAYVSGLWSLEDAC TLVSARARL
25 MQALPAGGVMVAVPVSEDEARAVLGEGVEIAAVNGPSSVVLSGDEAAVLQAAEGLGKWTRLATSHAFHSARMEPMLE
EFRAVAQGLTYHAPGVMAAGDRVMTAEYWVRQVRD TVRFGEQVASYEDAVFVELGADRSLARLVDGVAMLHGDHET
QAAIGALAHLYVNGVTVDWTALLGDVPVTRVDLPTYAFQQQRYWAERWLAALAPAEREKALLKLVS DGAA TVLGHAD
TSTIPATTAFKDLGIDSLTAVELRNSLAKATELRLPATLVFDYPTPTALAAARLD

30 SEQ ID NO: 154

EPLAIVGMACRLPGGVSSPEDLWRLVESGETDAISGFPTDRGWDVENLYDPDPDAPGKSYSVQGGFLDAAAGFDASFF
GISPREALAMPQQRLMLEVSWEAFERAGIEPGSVRGSDTGTFVIGAYPGGYGIGADLGGFGTTAGAASVLSGRVSYF
FGLEGPAFTVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPQTFVEFSRQRLSADGRCKAFADAADGTGW
AEGVGVLVERLSDAQANGHQILAVVRSSAVNQDGASNGLSAPNGPSQQRVIRAALSNA GLAPHEVDVVEAHGTGTT
35 LGDPIEAQAVIATYGQGRGEPLLLGSLKSNVGHTQAAAGVSGVIKMMALQHSMVPRTLHVDEPSRHVDWSAGAVEL
VAENQWPETGRPRRAGVSSFGISGTNAHVILESAPAQSVGDTAGSTPVLVSELVPLVISAKTQPALTEHEDRLRAY
LAASPGVDIRAVASTLAVTRSVFEHRAVLLGDET VTGTAVSDPRIVFVFPQGWQWLGMGSALRDSVVFAERMAEC
AAALSEFVDWDLFAVLDDPAVVD RVDVVQPASWAVMVSLAAVWQAAGVRPDAVIGH SQGEIAAACVAGAVSMRDAAR
IVTLRSQAIAGLAGRGAMASVALPAQDVELVDGAWIAAHNGPASTVIAGTPEAVDHLVTAHEAQGVRVRRITVDYAS
40 HTPHVELIRDELDDITS DSSSQTPLVPWLSTVDGTWVDSPLDGEYWYRNLREPVG FHPAVSQLQAQGD TVFVEVSAS
PVLLQAMDDDVTVATLRRDDGDATRMLTALAQA YVHGVTVDWRAVLGDVPATRVLDLPTYAFQHORYWAEAWLVGLA
PEERAKALLKVVRDSAATVLGHADARSIPATGAFKDLGVDSLTAVELRNSLT KATGLRLPATMVFDYPTPADLAARL

SEQ ID NO: 155

EPLAIVGMACRLPGGVSSPEDLWRLLESGETDAVSGFPTDRGWDVENLYDMAGKSHRAEGGF LDAAAAGFDAGFFGISP
 REALAMDPQQRLVLEVSWEAFERAGIEPGSVRGSDTGVMGAYPGGYGAGADLGGFAATASATSVLSGRVSYFFGLE
 5 GPAFTVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPELFTFESRQRGLASDGRCKAFADSADGTGWAEGV
 GVLLVERLSDAQAKGHQVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIIQAALSNAGLAAHEVDVVEAHGTGTTLGDP
 IEAQAVIATYQDRERPLLLGSLKSNIGHAQAAAGVSGVIKVMALQHNTVPRTLHVDEPSRHVDWAAGAVELVREN
 QPWPGTDRPRRAGVSSFGVSGTNAHVILESAPPAQPAEEAQPVETPVVASDVLPLVISAKTQPALTEHEDRLRAYLA
 ASPGVDTRAVASTLAVTRSVFEHRAVLLGDDTVTGTAVSDPRVVFVFPQGQWQWLGMSALRDSSVFAERMAECAA
 10 ALSEFVDWDLFTVLDDPAVVDRVDVVQPASWAVMVSLAAVWQAAGVRPDAVIGHSQGEIAAACVAGAVSLRDAARIV
 TLRQAIAGLAGRGAMASVALPAQDVELVDGAWIAAHNGPASTVIAGTPEAVDHLTAHEARGVRVRRITVDYASHT
 PHVELIRDELDDITSDSSSQAPLVPWLSTVDGWSVDSPLDGEYWYRNLREPVGFHPAVGQLQAEGDTVFEVSASP
 VLLQAMDDDDVTVATLRRDDGDATRMLTALAQAAYVHGVTVDWPAILGTATTRVDLPTYAFQHQRVWLSWLAALAPAE
 REKALLKLVCD SAAMVLGHADARSI PAAGAFKDLGVDSLMAVELRNLVKATGLRLPATLVFDYPTPTVLAARLD

15

SEQ ID NO: 156

EPLAIVGMACRLPGGVSSPEDLWRLVESGETDAVSGFPTDRGWDVENLYDSDPEAAGKSYCVQGGFLDTAAGFDAGFF
 GISPREALAMDPQQRLLEVSWEAFERAGIEPGSVRGSDTGVMFIGAFPVGYGAGFDREGYGATSGPSVLSGRVSYVF
 GLEGPAITMDTACSSSLVALHLAAQALRNCECSMALAGGVTVMATPEVFTFARQRGLASDGRCKAFADSADGAGFS
 20 EGAGLLLVERLSDARRNGHQVLAVVRGS AVNQDGASNGLTAPNGPSQQRVIRAALSNAGLSTADV DVEAHGTGTTL
 GDPIEAQALLATYQDREQPLLLGSLKSNIGHTQAASGVSGVIKVMALRHGFVPRTLHVDEPSRHVDWAAGAVELV
 RENQPWPGTDRPRRAGVSSFGVSGTNAHVLESAPPAQPAEEEQPVETPVVASDVLPLVISAKTQPALTEHEDRLRA
 YLAASPGADTRAVASTLAVTRSVFEHRAVLLGDDAVTGTAVTDPRVVFVFPQGQWQWLGMSALRDSSVFAERMAE
 CAAALSEFVDWDLFAVLDDPAVVDRVDVVQPASWAVMVSLAAVWQAAGVRPDAVIGHSQGEIAAACVAGAVSLRDAA
 25 RIVTLRSQAIAGLAGRGAMASVALPAHEIELVDGAWIAAHNGPASTVIAGTPEAVDHLTAHEARGVRVRRITVDYA
 SHTPHVELIRDELLGITAGIGSQPPVVPWLSTVDGWSVDSPLDGEYWYRNLREPVGFHPAVSQLQAQGDVFEVSA
 SPVLLQAMDDDDVTVATLRRDDGDATRMLTALAQAAYVHGVTVDWPAILGTTTARVLDLPTYAFQHQRVWKS WLAAL
 APEERAKALLRVVCD SAATVLGHADIDSIPVTA AFKDLGVDSL TAVELRNSLAKATGLRLPATLVFDYPTPTALAAR
 LD

30

SEQ ID NO: 157

EPLAIVGMACRLPGGVSSPEDLWRLVESGETDAISDFPADRGWDVENLYDPDPDASGKSYCVQGGFLDSAGGF DASFF
 GISPREALAMDPQQRLVLEVSWEAFERAGIEPGSLRGSDTGVMFIGAYPGGYGAGAGADLEGYGTSGPSVLSGRVSY
 FFGLEGPAITVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPDVFTFARQRGLATDGRSKAFADSADGAG
 35 FSEGIGVLLVERLSDAEAKGHQVLAVVRSSAVNQDGASNGLTAPNGPSQQRVIIQTALSNAGLTAEVDVVEGHGTGT
 TLGDPIEAQAVIATYQDREQPLLLGSLKSNIGHTQAAAGVSGVIKVMALRHALLVPRTLHVDEPSRHVDW TAGAVE
 LV TENQPWPEIGRPRRAGVSSFGVSGTNAHVILESAPPTQAEDAAQPV EAPVMGSEPVPLVISAKTLPALNAHEDRL
 RAYLAASPGVDMRAVASTLAMTRSMFEHRGVLLGDGT VSGTAVSDPRVVFVFPQGQSQRAGMGEEELAAAFPVFARIH
 QQVWDLLDVPDLVDVNETGYAQPALFALQVALFGLLESWGV RPDVIGHSVGELAAAYVSGVWSLEDACTLV SARARL
 40 MQALPAGGVMVAVPVSEDEARAVLGEGVEIAAVNGPSSVVL SGDEAAVLQAAEGLGKWTRLATSHAFHSARMEPMLE
 EFRAVAEGLTYRTPQVAMAAGDQVMTAEYWVRQVRD TVRFGEQVASFEDAVFVELGADRSLARLVDGVAMLHGDHEA

QAAVGALAHLYVNGVSVEWSAVLGDVPVTRVDLPTYAFQHORYWLEGRWLAALAPAEREKALLKLVSDGAATVLGHA
DTSTIPATTAFKDLGINSLTAVELRNSLAKATELRRLPATLVFDYPTPAALAARLD

SEQ ID NO: 158

5 EPLAIVGMACRLPGGVSSPEDLWRLLESGETDAVSGFPTDRGWDVENLYDMAGKSHRAEGGF LDAAAGFDAGFFGISP
REALAMDPQQRLLVLEVSWEAFERAGIEPGSVRGSDTGVMGAYPGGYGIGADLGGFGATASSVSVLSGRVSYFFGLE
GPAFTVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPQTFVEFSRQGLASDGRCKAFADAADGTGWAEGV
GVLLVERLS DARRNGHQVLAVVRGS AVNQDGASNGLTAPNGPSQQRVIRAALSNAGLSTAEVDVVEAHGTGTTLGDP
IEAQALIATYQQRDQPVLLGSVKSNGHTQAAAGVSGVIKVMALQHGLVPRTLHVDEPSRHVDWSAGAVQLVTEN
10 QPWPDMGRARRAGVSSFGISGTNAHVILESAPPTQPADNAVIERAPEWVPLVISARTQSALTEHEGRLRAYLAASPG
VDMRAVASTLAMTRSVFEHRAVLLGDDTVTGTAVSDPRAVFVFPGQGSQRAGMGEELAAAFPVFARIHQQVWDLDDV
PDLEVNETGYAQPALFAMQVALFGLLESWGVRPDAVIGHSVGELAAAYVSGVWSLEDACTLV SARARLMQALPAGGV
MVAVPVSEDEARAVLGEGVEIAAVNGPSSVVLSGDEAAVLQAAEGLGKWTRLATSHAFHSARMEPMLEEFRAVAEGL
TYRTPQVSMAGVDQVTTAEYWVRQVRDTRVFEQVASYEDAVFVELGADRSLARLVDGVAMLHGDHEIQAAIGALAH
15 LYVNGVTVDWPALLGDAPATRVLDLPTYAFQHORYWLEGRWLAALAPAEREDALLKLVDRD SAALVLGHADASTIPAAA
AFKDLGIDSLTAVELRNSLAKATGLRLPNTTVFDYPTPAILATRL

SEQ ID NO: 159

EPLAVVGMACRLPGGVSSPEDLWRLVESGETDAISGF PADRGWDAESLFDPDPAVGKSYCVEGGFLD SAASF DAGFFG
20 ISPREALAMDPQQRLLIMEVSWEAFERAGIEPGSVRGSDTGVMGAYAGGYGAGADLGGFAATASATSVLSGRVSYFF
GLEGPAITVDTACSSSLVALHQAGYALRQGECSLALVGGVTVMATPQS FVEFSRQGLASDGRCKAFADSADGTGWA
EGVGVLVERLSDAQAKGHQVLAVVRSSAVNQDGASNGLTAPNGPSQQRV IQAALSNAGLAAHEVDVVEAHGTGTTL
GDPIEAQALIATYQQRDRPLLLGSLKSNIGHAQASGVSGVIKVMALQHNTVPRTLHVDEPSRHVDWAAGAVELV
RENQPWPGTDRPRRAGVSSFGVSGTNAHVILESAPPAQPAEEAQPVETPVVASDVLPLVISAKTQPALTEHEDRLRA
25 YLAASPGADIRAVASTLAVTRSVFEHRAVLLGDDTVTGTAVTDPRIVFVFPGQGWQLGMGSALRDSSV VFAERMAE
CAAALREFVDWDLFTVLDDPAVVDRVDVVPASWAMMVSLAAVWQAAGVRPDAVIGHSQGEIAAACVAGAVSLRDAA
RIVTLRSQAIAGLAGRGAMASVALPAQDVELVDGAWIAAHNGPASTVIAGTPEAVDHVLTAEHAQGVRRRITVDYA
SHTPHVELIRDELDDITS DSSSQTPLVPWLSTVDGTWVDSPLDGEYWRNLREPVG FHPAVSQLQAQGD TVFVEVSA
SPVLLQAMDDDVVTATLRRDDGDATRLTALAQA YVHGVTVDWPA ILGTTTTRVDLPTYAFQHORYWLKSRLTGRT
30 SVEQHRIMLELVLGEAASVLGHSSADAIATDTSFKDLGMDSLTAIELNRNLVAETGLQLPATMVF DYPTANALAAHL

SEQ ID NO: 160

EPIAIVAMACRLPGGVSSPEGLWHLVESGETDAISGFPTDRGWDVEGLFDPDPDAAGKSYCVQGGFLDTAADFDAPFF
GISPREALGMDPQQRLLLETTWEAIERAQIDPKSLRGRDVG VYVGAAQGYGVGVDDQHDNGITGSSVSLLSGRVSY
35 ALGLEPGVTVDACSSSLVALHLASQALRQRECSLALVSGVSVMS SPAMFVEFSRQRLSSDGRCKSFAASADGTI
WSEGVGVLVERLS DARRLGHRVLATVRGS AVNSDGASNGLTAPNGTSQQRVIRQALANAGLTASDVDVVEAHGTGT
KLGDP IEAEAILATYQ QERSAPAWLGS LKSNIGHAMAASGVLSVIKMEAMHGSLPRTLHV DAPSPHVDWTS GSVA
LLTEHQWPDDTKLRRAGVSSFGLSGTNAHV VLEQYQAPAPPVTPVTPAPPTPVTPVTPNEPGPLPWLSAQSPKAL
REQAGRLYASLAGDSEWNSLDIGYSLATTRSDFAHRAVAVGSGREFLRALSKLADGAPWPGLTTATATAKARRVAF L
40 FDGQGTQRLGMGKELYDSYPA FARAWDTVSAGFDKHL DLSLTDVCFGEGGSTTAGLVDDTL YAQAGIFAMEAALFGL
LEDWGV RPDFVAGHSIGEATAAYASGMLSLENTT LIVARGRALRTTPPGAMVALRAGEEEVREFLSRTGAALDLAA
VNSPEAVVVS GEPEPVADFEAAWTASGREARKLKV RHAFHSRHVEAVLDEFRTALES LKFRAPALPVVSTVTGR LID

QDEMGTPEYWLRQVRRPVRFQDAVRELAEQGVGTFVEVGP SGALASAGVECLGGDASFHAVLRPRSPEDVCLMTAIA
ELHAGGTAIDWAKVLSGGRAVDLPVYPFQHQS YWLAPAEP SYADEPRTMLELVHMEVASLLGMADPGVILDDSSFLE
LGFDLS SAVRLRNRLSKATGLDLPSTLLFEHPTSAE
LAAHLD

5

SEQ ID NO: 170

MSRAELVRPIYDLLRANAERLGDKMAYVDSRLALTHAELAARTGRIAGHLVDMGVDRGDRVAILLGNRVENIESYLA
IARASAVAVPLNPDATGA EVAHFLSDSGAVVITDSAHLDDVRR TAPAVTIVLVGEERIPPGVRSFAELATAEPQQS
ARDDLGLDEAAWMLYTS GTTGT PKGV LSTQGSGLWSAAYCDIPAWELTENDVLLWPAPLFHSLALHLCVLATTAVGA
10 TARIMNGFVASEVLEELTEHPCTVLVGVPTMFRYLLGAADTFEPRTSS LKMGLVAGSVAPASLIEGFEDVFGVPLLD
TYGCTETSGSLTVNWL SGERVPGSCGLPVPGLSLRFVDPISGADVADGEEGELWASGPSIMIGYHEQPEATAEVLSD
GWYRTGDLARRSETGHVTITGRIKELIIRSGENIHPHEIEAVALDVPGVKDAAAAGKRHPVLGEVPVLYVVPETGGV
DADMVLAVCRERLSYFKVP EEIYRVDAIPRTASGKVKRSSLTEEPAELLAGASGGETLHRLEWIPLELPEQAAPDGH
VVVRVDSLASDDSD LADADLADAVRDLARSWLADKRRADSTLVFVTRRAVHTGPSDIPSPEHAAVWD AIRREQTENP
15 GV FVVVIDVDDDDDDDDVNDREDDDTLLPALAGLGEPQVALRDGNPLVPRLAHANTPDSGSLTIPEDRAWLLEHSRS
GTLRDLALVPADAAERPLHPGEVRISVRAAGLNFRDVLIALGTYPGEGLMGGEAAGVVLEV GSEVSD LAPGDRVFG L
VGS AFGTVAIADRLLGAIPDTWSFATAASIPIVFATAYYGLVDLAGLSAGESVLIHAAAGGVGMAATQIARHLGAR
IFATASVGKQHILSEAGLEDTRIAGSRTLAFREAF LN TTDGQGV DVVLNSLSGDFVDASLDLLPRGGRFLEMGKTDI
RDADRITADRP GTTYQAFD LLDAGPDRLREIIAELLPLFAQGVLRPLPLLTWDIRKARDAFSWMSRARHTGKITFTTI
20 PRQLDPGGTVLIADGSGVLTGTVARHLVAEQGV RHLLLLSRSTPDEALINELIESGARVDTAVRDVSDRAGLEQAL A
GISPEHPLTAVIHTGGPAVAHESHQLHGLTKRLDLAA FVVF SQDAPASVDALARRRRAEGLPTTTI AWGIPEAEAVV
VQGPLLGRAMASADSAHIVTRLNTVGLRALAAADTL PPLLRNLVGAQTDNTQQQAWSRQLLAAEEAAREQALRDLVRS
CVM DILGLS AADRYAPDKTFREM GIDSLTAVELRN SLAKATDLRLPATMVFDYPTPAMLVVRLGE

25 SEQ ID NO: 171

MSREEFIQPIHDLRLVNAERLGDKIAYADSRRELTHAELRTRTGRIAGHLVDLAVERGDRVAILLGNRVETIESYLA
IARAGAI AVPLNPDATGA EVAHFLADSGAVLVITDSAHLDDVRR AAAAVTVVLVDEGPLPAGTRSF AELATAEPPTP
ARDDLGLDEAAWMLYTS GTTGT PKGV VSTQGSGLWSAANC DVP AWELTENDVLLWPAPLFHSLAHHLCLLATTAVGA
TARIMSGFVAGEVLHELEEHACTVLVGVPTMYHYLLGAVGEAGPRLPS LKMGLVAGAVSPPALIEGFERVFGVPLLD
30 TYGCTETTGS LTVNRLSGPRMPGSCGQAVPGISLRFVDPHTGA EVAEGEEGELWASGPSLMIGYHGRPDATREVLSD
GWYRTGDLARRSETGHVTITGRVKELIIRGGENIHPRDIEA VALELPGVRDAAAAGKQHPVLGEIPALYLV PDADGV
DAEAVLAACREKLSYFKVP EEIYRVDAIPRTL SGKVKRAALTEAPAELL SAASGNGSLYRLEWVPAETPPAGTGGPV
AVHVTRRAVATGPADLPDQEQAATWDALRGEQTGPGGPVLIDLDGADID DARLSALASLGEPQIVVRDDTPLVARLA
REKSPALTIPGERAWVLEPDHSGVLQELALVAADTDVRPLRPGEVRIEVRAAGLNFRDVLVALGTDLDG DGVFGAEGA
35 GVVLETGSDVRDLRPGDRVFG LLEGGHGSIAIADRRMLAVIPEGWSFATAASVPEVFVIAYYGLVDLAGLRAGESVL
IHAATGGVGMAATQIARHLGAQVYATAGVGKQHILRDAGLGDDRIADSRTTDFREAFRDSTQGRGVDVVLNSLKGDF
VDASLDLLADGGRFLELGQTDIRDAGEIAAERP GTTYHSFTRM NAGPDRLREIIAELLALFEQGVLRPSPVHTWDIR
HAREAFSWMSGGRHTGKMVLTMPQRIDPGGTVLIAGDSEALARIAARHLGV RHLLLD RGVADAAPDAVVC DVSDHDA
LERVLADLSPEHPLTAVIHTGGA AVTDEIRRLHDLTESLDLTDFVVF SQDAPAAVEAFARSRRAGLPVRTIAWGIP
40 EADPVVADEHLLGRALASAEQAQIVARVNTAGLRALTAANALPTLLRN LIRAEPEETGQSAWPHRFEAAGADREEAL
LDLIRANVVDILSLPTADRYAPDRTFREM GIDSLTAVGLRN SLAKATGLPLPTTMVFDYPTP

SEQ ID NO: 172

MSHAKLIQPIYDLLRVNAERLGDKIAYADSRHALTYTELEARTGRLAGHLADLGVERGDRVAILLDNRVETIESYLA
 IARASAIIVPLNPAAAGDELAHFLSDSGSVLVITDSAHLLDDVRLVAPAVTVVRVDEDPVPPGVRSAELVAVEPRTQ
 ARDDLGLDEAAWMLYTS GTTGT PKGVVSTQGSGLWSA AFCDPVAPWELTEEDVLLWPAPLFHSLAHHLCLLATITAGA
 5 TARIMNGFVASEVLNELEKHACTVLVCVPTMYHYLLGAVGEGESRTFSLKLGVVAGSVSPPALIEGF EKAFGAPLLD
 TYGCTETTGS LTVNWLNGPRVPGSCGTAVPGVTLRFVDPSTGADVADGEEGELWASAPSVMTGYHGQPEATREVLTD
 GWYHTGDLARRSETGHVTITGRIKELIIRGGENIHPQEIEAAVLGLPGVRDAAAAGRPHPV LGDVPALYIVPDADGV
 DADAVLAACRERLSYFKVPEEIIYRVDAIPRTMSGKVKRTSLTEAPAELLAGASGSDALYRLKWVPAETPGPAATGGH
 VIVRVASLRADGTELAGAARDLARSWLS DERRAGATLVFVTGRAVSAGPSDVPVPEHAAVWDAIRDEQ TENPGAFVL
 10 IDLEEAETEPEESA APEAGDPQADTPGADDTRLSTLVALGEPQIALRDSTPLVPRLAPESSTALTTPAARAWVLEPA
 RSGTLRELSLVAADTDARPLRPGEVRVDVRAAGLNFRDVLIALGTYPGDGVMMGGEAAGVVLEVGP EVNDLSVGDRVF
 GLVTDGFGPVTITDRLLAAMPQDWSFTTAASAAMAFATAHYGLVELAGLKAGESVLIHAATGGVGM AATQIAHHLG
 AHIYATASSGKQHLLRAAGIDDDRIANSRTTGFRDAFLDSTGGRGVDVVLNLSLGEFVDSSDLLAHGGRFIEMSTD
 IRDAGRIAAERP GTTYQAFHLVDADPDRLEILTELLALFDQ GILDPLPVQAWDIRQAREAFSWMSRARHTGKLVLT
 15 IPQHIDPDGTVLITGGSGGLAGVVARHLVADKGARRLLLLSCDTLDATLAAELTESGARVDTAVCDVSDRAALAQVL
 AGVSPEHPLTAIVHAGGA AVADESRQLHHLTKNRDLAA FVVF SQDAPAAATEAFAGIRQA EGLPVT TIANGIPEAE PV
 VVGQHLLDRAMASADRAHVAA RVNTAGLRALAAANALPPVLKNLVGAETDGTGHQDWSRRFMVAEAAARQQELLDLIR
 TTVMEILSLPTTARYFPDRTFRENGIDSLTAVELVNSLAKTTGLRLSATMVFDYPTPTALAGRMREL

20 SEQ ID NO: 173

MSRLDLIRPLSESLCASAASFGDKVAYTDSRRSVTYAELQIRTGRLAGHLAEHGVARGDRVAILLGNRVEIIESYLA
 VARASAVAVPLNPDAMDAELAHFLRD SGAVVVITDLAHLEQTRGVAPAMTVVLIGDGRTPGTSSFAELADTEPASP
 ARDDLRLDEPAWMLYTS GTTGT PKGVVSAQRSGLWSAASCDVPAPWDLSD EDDLWWPAPLFHSLGHHLCLLAVVAVGA
 SARIMSGFAADEVLDALREHPCTVLVGVPTMYRYLLAAVGESGADAPALKMALIAGSVTPASLVEAFERSFGVPLID
 25 TYGCTETTGSVTANRLHGERVPGSCGVPVPGVEIRLVDPVTGADVPLGAEGELWAKTPSVMIGYHGQPEATGEVLVD
 GWYRSGDLARRQESGHITITGRVKELIIRSGENIHPREIETVALEVPGVEDAAAAGKPHRVLG EVPVLYVVP AEAGV
 DVTAVFAACREQLSYFKVPEEIIYQVESIPRTPSGKVKRGLLTEQPAELLAAADGGGSLYRVEWRPAVPPGAGDTGGD
 SPVVVRVDSLPADEQELLGAVRDLIHDRIADPRRTTAPLVFVTRHAVLSRNP AHAHA AVWDLVSRAQADNPGLFVLV
 DADGDDAPLPSAVGLGEPRVAWRDGGLLVPRLAHPGTEALIAPESGSWLLAETGGGTLRDLALVGTDTADRVLLPGE
 30 VRIAVRAAGLNFRDVTVALGVVSDDRLMGGEAAGVVLVDVAPDVTDL EPGDRVFG LVEGAFASVAVTD RRLGRIPAG
 WSFATAASVPVVFSTAYHALTDLVDLRPGEAILIHAAAGGVGM AATQLARHIGAKIYATASPAKQHALLGVDQVANS
 RTTEFRGTGFLEATGGRGVDVVLNLSLAGEHIDASLELLPRGGRFLELGKTDLRDPRHLPAGVSYQVLNRLDSSPDRVR
 EILAELLVMFERGVLRLPVRTWDIREAPEAFSWMSRGRHLGKIVLTIPRDLPDGTVLVTGASADHMARYLSAERG
 HAHVLVSDDPAAVPATHPLTAVVHTGGDEVVSESTRLHQLTRELDLAAFAVFSQSAPASVEALVRHRRTEGLPATAV
 35 SWGLPEAEPA PVQGALLDRTIASVEPAHV VTRVNSAGLRALANS GELPSVLRDVT PALS AKWPRPGT PRPGT PRPGT
 PHPAALDQAALLDLVRESLTTVLGLPGVESCAPDRPFRETGLDSLTTIGLANTLSARLGRKLPATMIFDHPTPRTLA
 TRLA EEL

SEQ ID NO: 174

40 MTPAYDVRPLPELLIANAERLGDKPAYTGLHRTVGHAE LADRTRRAAGHLAGFAARGDRIALLGNRVEMVEGYLAV
 ARAAAVAVPLNPQASDAELAHFLTDSEAVAVLADAEHTEQVRRVAPGLRLVPIGEWETLATTEPDRPARDDLGLDEV
 AWMMLYTS GTTGTAPKGVLS TQGSGLWSAYHCDVPALGLTDADVLLWPAPLFHSLAHHLAVLAATVSGATVRLLSGFAA

DEVLRELREEGCTLLAGVPTMYHYLLGAAGPDDEV RAPALRGAVVAGAVTPAALITAFGERFGAPLLDTYGCTETT
 SLTINRLDGPVRPGSCGVAVPGVRLRLVDPRTGDDAPEGGEI WASGPSLMRGYHRRPDATAEVLADGWYRTGD LA
 VRAATGHITITGRVKELIIRGGENIHPREIEAVLA E VPGVADVAVAGRSHAVLGDL PVAYLVTEAGLDPAALFAACR
 ERLSSFKVP EEEVYRVA AVPRTPSGKIKREL VAGPAELLATAGGAETLLRTRWTA VDVLDPASLDGWRV VHADQEVD
 5 LGGR LDDDGP AIVVTTTRAVRTSAD ERPSASAAAAWDLVTAAQARRPGRYLLVD TDGVP GGLGAALATGEPQVAVRED
 VVLVPRLEAAGETGAPVRLDGTVLVTGEHTERVARHLRARGVTVTDDPAARPLHAVVHVGGTSGLAELAE LTGCPER
 AAFVVCTEDSRATADALVRAIPGGVAVGVGLPGIEPAALLPELLDRLTADGPYVVARPGSTGLRALATAGRLPAGLG
 ALVDTGAAPDPDAAVRRDLVRRLIALPRRARDQALVELVWD AVRATLGAGATPGGPGQAFSEVGFDSLTSVQLRNRL
 VAATGVRLSATAVFDFPTPRALADELGRVLI

10

In another aspect, the disclosure provides a chimeric polyketide synthase, wherein at least one module of the chimeric polyketide synthase has been modified as compared to a polyketide synthase having the sequence of SEQ ID NO: 175-176.

15 SEQ ID NO: 175

MSREEFIQPIHDLLRVNAERLGDKIAYADSRRELTHAELRTRTGRIAGHLVDLAVERGDRVA ILLGNRVETIESYLA
 IARAGAI AVPLNP DATGA EVAHFLADSGAVLVITDSAHLDDVRRAAA AVTVVLVDEG PLPAGTRSF AELATAEPPTP
 ARDDLGLDEAAWMLYTS GTTGT PKGVVSTQGSGLWSAANCDVPAWELTENDVLLWPAPLFHSLAHLCLLATTAVGA
 TARIMSGFVAGEVLHELEEHACTVLVGVPTMYHYLLGAVGEAGPRLPSLKMGLVAGAVSPPALIEGFERVFGVPLLD
 20 TYGCTETTGS LTVNRLSGPRMPGSCGQAVPGISLRFVDPHTGA EVAE GEEGELWASGPSLMIGYHGRPDATREVLSD
 GWYRTGDLARRSETGHVTITGRVKELIIRGGENIHPRIEAVALELPGV RDAAAAGKQH PVLGEIPALYLVPDADGV
 DAEAVLAACREKLSYFKVP E E IYRVDAIPRTL SGKVKRAALTEAPAELL SAASNGSLYRLEWVPAETPPAGTGGPV
 AVHVTRRAVATGPADLPDQEQAATWDALRGEQTGPGGPVLIDLDGADID DARLSALASLGEPQIVVRDDTPLVARLA
 REKSPALTIPGERAWVLEPDHSGVLQELALVAADTDVRPLRPGEVRIEVRAAGLNFRDVLVALGTDLDG DGVFGAEGA
 25 GVVLETGSDVRDLRPGDRVFG LLEGGHGSI AIAIDRRMLAVIPEGWSFATAASVPEVFVIAYYGLVDLAGLRAGESVL
 IHAATGGVGMAATQIARHLGAQVYATAGVGKQHILRDAGLGDDRIADSR TTDFREAFRDSTQGRGVDVVLNSLKGDF
 VDASLDLLADGGRFLELGQTDIRDAGEIAAERP GTTYHSFTRMNAGPDRLREIIAELLALFEQGVL RPSPVHTWDIR
 HAREAFSWMSSGRHTGKMVLTMPQRIDPGGT VLIAGDSEALARIAARHLGVRHLLLDRGVADAAPDAVVC DVS DHDA
 LERVLADLSPEHPLTAVIHTGGA AVTDEIRRLHDLTESLDLTD FVVFSQDAPAAVEAFARSRRAGLPVRTIAWGIP
 30 EADPVVADEHLLGRALASAEQAQIVARVNTAGLRALTAANALPTLLRNLI RAEP EETGQSAWPHRFEAAGADREEAL
 LDLIRANVVDILSLPTADRYAPDRTFREM GIDSLTAVGLRNSLAKATGLPLPTTMVFDYPTPAVL TARMRELLAGES
 PAPARTAARAVAQDEPLAIVGMACRLPGGVSSPDDLWRLVAAGTDAISEFPADRGWDVDNLYDPDPDAPGKTYTVLG
 GFLDG VAGFDASFFGISPREALAMPQQRLMLEVSWEAFEHAGIPPRSVRGSDAGVFMGAFPSGYDAGLEEFGMTGD
 AVSVLSGRVSYFFGLEGP AITVDTACSSSLVALHQASSALRQGECSLALVGGVTVLATPQT FVEFSRQRGLALDGRS
 35 KAFADAADGAGWAEGVGVLVVERLS DARAKGHQIWGVIRGS AVNQDGASNGLSAPNGPSQQRVIRQALANAGLAPHE
 VDVVEAHGTGTTLGDP IEAQAVIATYQDREQ PLLLGSLKSNVGHTQAAAGVSGVIKMVMALQHDTV PATLHVDAPS
 RHVDWTAGAVELVTENRPWPETGRVRRAGVSSFGISGTNAHVILESAP EQPVSPPEAVAPVVASDRVPLVISAKTPA
 ALAEMENRLRAYLAAAPGADPRAVASTLATARSVF EHRVLLGENTITGT VAGADPRVVFVFPQGQWQQLGMGRALR
 ESSPVFAARMAECAALSEFVDWDLFTMLDDPAVIDRIDVLQPACWAVMMSLA AVWQAAGVRPDAVIGHSQGEIAAA
 40 CVAGALSLRDAARIVALRSQLLAREMVGHGVMAAVALPADDIPLVDGVWIGACNGPSSTVISGTPEAVEVVAACEE
 RGARVRRITA AVASHSPLGEKIRTELLGISASIPSRTPVVPWLSTADGIWIEAPLDPAYWWRNLREPVGFGPAVDLL
 QARGENVFLEMSASPVLLPAMNDAVTVATLRRDDGT PDRMLTALAEAHAGVIVDWPRVFGSTTRVLDLPTYAFEHQ

RYWAVSADRPSDAGHPMVETVVPLPASGGVALTGRVSLATHAWLADHAVRG TALLPGTAFVELVTRAATEVDCPVID
ELVIEAPLPLTQTGAVQLSTTVGEADESGRRPVTVFSQADGTDWTRHVTATIGRAASLPDPVAVPPAQAEPVDTV
FYDELAAGYEGPAFQGLRAAWSGDVTYAEVVLAAEQAEVDRYAVHPALLDAALQAGMVNTAGTGQGVRLPFWS
NGIQVHSTGATTLRVAATPLADGWSVRAAADNGRPVATIGSLVTRPVTTDMLGSTTDDLFAVVWTEITAPEPGDPSD
5 VGVFTALPEAGGDPLTQTRALTAQVLQTVQQWLAGEDRPLVVRTGTDLASAAVSGLVRSAQSEHPGRLLILVESDDEL
TPEQLAGTAGLDEPRIRIDGGHYEVPRLAREDA SLTPEDRAWLLELPGSGTLRDLRVIPTD TAERPLRWGEVRVGV
RAGGLNFRDVVVALGMVTDPRPAGGEAAGVVLETGPGVEDLSPGDRVFGILDGGFGSVAIADRRL LAVIPDGWSFTT
AASIPVVFATAYYGLVDLAGLRAGESVLIHAATGGVGMAATQIARHLGAEIYGTAGIAKQHVLRDAGLGDDRIADSR
TTGFRETFRDSTQGRGVDVVLNSLSGDFVDASLDVLAEGGRFIEMGKTDIRD AEQITHATYRAFDLMDAGPDRVREI
10 IAELLGLFEQGVLRPLPVQAWDIRQARDAFTWMSRARHIGKIVLTIPQQLDPDGTVLISGGSGVLGILARHLVAER
GVRHLLLVRSAPSEALISELTALGAQVETVACDVSDRVALEQVLDGVPLTAVFHTAAALDDGVVESLTPQRVDTVL
RPKADAAWYLHELTRDADLAAFV MYSSVAGIMGAAGQGN YAAANAFDLALAAHRRREGLPALSLAWGLWEDASGLSA
GLTETDHDRIIRGGLEAIAAEHGMRLFD TATRQGEVLLASPLNLTRQGEVPALLRTLHRPVARRAATANGRPADLT
PEALLKLVCGRAAAVLGHVDADAVPVAVAFRDLGVDSL TAVELRNLSLAKATGLRLPATLVFDYPTPTVLAGRLGELL
15 AGGTAPVRAAVVRRAAASDEPLAIVGMACRLPGGVLS PEDLWRLVESGGDAISGFPVDRGWDVENLFD PDPDAAGRT
YAVRGGFLDGAAGFDASFFGISPREAQAMD PQQRLVLEVSWEAFERAGIEPGSVRGSDTG VFMGAYPGGYGVGTDLG
GFGMTSVAVSVLAGRVSYFFGLEGPAMTVDTACSSSLVALHQAGSALRQGECSLALVGGVTVMPTPQT FVEFSRQRG
LAADGRCKAFADAADGTGFSEG VGVLVERLSDAQARGHNILAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQALAN
AGLAGAEVDVVEAHGTGTTLGDP IEAQAVIATYGQDRDQP VLLGSLKSNLGHTQAAAGVSGVIKVMALRHDTVPAT
20 LHIDEPSRHIDWTAGAVELVTENQSWPETGRARRAAVSSFGISGTNAHVILE SAPAQPVPLVDTPVSAV TAGVVPLP
ISARTVPALADLEDRLRAYLT TTPETDLPAVASTLAVTRS VFHRAVLLGEETVTGIAVSDPRVVFV FSGQGSQRVG
MGEELAAAFPLFARLHRQVWDL LDVPDLEVDDTGYVQPALFALQVALFGLLESWGVRPEAVIGH SVGEVAAGYVAGV
WSLEDACTLV SARARLMQALPAGGAMVAVPVSEERARAVLVDGVEIAAVNGPASVVLSGDES AVLRVAEGLGRWTRL
SASHAFHSVRMEPMLEEF RQVASELTYREPRIVMAAGEQVTTPEYWVRQVRD TVRFQVAAFGDAVFLEIGPDRTL
25 SRLIDGIPTLHGDD EQHAVVAALAEHLVQGVPI DWSSILGANPARVLDLPTYAFQH ERYWMVSTGRVGGEGHPLL GW
GVPVAEAGGRLYTGRVARQDGPVLSVA AFVEMAFAAAGGRPIRELSVDALLYIPDDGTAE LQTWVSEHRLTIHARYR
DTEPWTRLATAALDTTAPATHTTPHPGLITTALTLTGDEAPAIWHDLT LHTSNATELH THITPGDDGTLTITATDTT
GQPVLTATHTATPTTIPVHTPTTPADDLLTLTWTQIPTPGPDPTDIAVCTALPD PDGDPLAQTRTLTAQVLQSIQTT
LTGEDRPLVVHTGTGLASAAVSGLV RSAQSEHPDRFILVESDDSLPQAQLAAVAGLDEPWLRITGSCYEVPRLTKTT
30 TATATAVSEPWNPDGTVLITGGSGALAGILARHLVTERGVRHLL LISRSTPSTTLTDELRELGAHV DVAACDVSDR
DALARVLDGVDLTAVFHTAGALDDGVVESLTPQR LDTVLT PKADGAHWLHELTRDRDLTAFV MYSSAAGVMGAAGQG
NYAAANAFDLALAEHRHADGLPALSLAWGMWDDTDGMTASLSGTDHRRIRRSQRAITAEHGMRL LDKASGRSEPVL
VATAMNPIPD TDLPALRLSLYPKTARKSQPIQELSPEALLKIVRDSAALMLGHPNTDAIAATTA FRDLGVDSLIAVE
LRNSLAKATGLRLPATLVFDYPTPTVLAGRLGELLAGVTPQRHATVRTGTASDEPLAIVGMACRLPGGVSSPEDLWR
35 LVESGTDAITDFPTDRGWD TDDLFD PDPDTAGKTYTVHGGFLDDVAGFDASFFGISPREAQAMD PQQRLVLEAAWEA
FERAGIEPGSVRGSDTG VFMGAYPGGYGIGADLGGFGATAGAGSVLSGRLSYFFGLEGPAMTVDTACSSSLVALHQ A
GSALRQGECSLALVGGVTVIANPQIFVEFSRQRGLAADGRCKAFADSADGTGWSEGVGVLVERLSDAQARGHNILA
VVRGSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLAGAEVDVVEAHGTGTTLGDP IEAQAVIATYGQDRDQSVLL
GSLKSNLGHTQAAAGVSGVIKVMALQNGVVPRTLHADQPSRHIDWTAGAVELVTENQPWPELDRPRRAAVSAFGVS
40 GTNAHVILESAPDQPVPLVDTPVSAV TAGVVPLPISARTVPALADLEDQLRAYLT TAPETDLPAVASTLATTRS VFE
HRAVLLGEDTVTGTAIPDPRIVFV FSGQGSQRAGMGEELAAAFPLFARLHRQVWDL LDVPDLVDVDDTGYVQPALFAL
QVALFGLLESWGVRPRAVIGH SVGEVAAGYVAGVWSLEDACALVSARARLMQALPAGGAMVAVPVSEERARAVLVDG

VEIAAVNGPASVVLSGDEAAVLRVAEGLGRWTRLSASHAFHSVRMEPMLLEEFQVVSRLTYREPRIVMAAGEQVTTT
EYWVRQVRETVRFGDQVAAFQDAVFLEIGPDRTLSRLIDGIAMLDGDDEVRAAVAALAMMHVQGVGDWPAILGTTT
GRVLDLPTYAFQHERRYWMANADEGHPLLKGVEHPLLGSVMALPNSDGVVLTGRISLATHAWLADHVVRGTVLLPGTG
FVEMVARAAAEEVGCVIDELLIEAPLLLPEHGGVHLSVSVGEADGAGRRPVTVFAQADDAEVWVRQVTATISPAGPA
5 VSLPELEVWPPVQAEPVDVSTFYERLARADWQWGAFAQGLRAAWRDGDTIYAEIVLADEEAREADQFLVHPALLDAA
LQTSVLKTPDDLRLPFSWNQIEFHATGAAILRVAVTPVADRWIVHAADSTGRPVATIGALVSRPVTAETLGSNTDDL
FALTWTEIPTPGPGDPADVAVCTALPEPDSPLTQTRTLTAQVLQSIQTSLTGEDRPLVVHTGTGLASAAVSGLVRS
AQSEHPDRFILVECDDETLPDQLAATAGLDEPWLRITGGHYEVPRLTKTTTAAATTVSEPVDPDGTVLITGGSGA
LAGILARHLVTERSVRHLLLISRSTPSTTLINELRELGAHIETAACDVSDRDALARVLDGVDLTAVFHTAGALDDGV
10 VESLTPQRLD TVLMPKADAAWHLHELTRDRDLAAFMYSAAAGVMGAAGQGNAAAAFLDALAEHRRADGLPALSL
AWGMWDDADGMTASLSGTDHRRIRRSQRAITAEHGMRLLDKASGRSEPVLVATAMNPAGEGEVPALLRLTHRPVAR
RAATTNGRPADLTPEALLKVVRDSA AAVVLGHASADTVPAATAFQELGLDSLIAVELRNSLAKATGLRLPATMVFDP
TPAALAGRLGELLAGETTPATAAVVRRATASDEPLAIVGMACRLPGGVSSPEDLWRLVESGFDAITGFPTDRGWDVD
NLYDPDPDAPGKSTTLHGGLDDVAGFDASFFGISPREAVAMPQQRLAMEVSWEAFERAGIEPGSVRGSDTGVMG
15 AYPGGYGIGAEELGGFMTGRAGSVLAGRVSYFFGLEGPAMTVDTACSSSLVALHQAAYALRQGECSLALVGGVTVM
TPVMFVEFSQQQNLADDGRCKAFADSADGTGWSEGVGVLLVERLSDAQARGHNLAVVRGSAVNQDGASNGLTAPNG
PSQQRVIRSALTSAGLTTADVVDVEAHGTGTTLGDP IEAQAVLATYQDRDQPVLLGSLKSNLGHGTQAAAAGVSGVIK
MVMALQNGVVPRTLHVEEPSRHVDWTAGAVELVTENQSWPETGRARRAAVSSFGSGTNAHVILESAPAQPVPMDT
PAPAVTTGVVPLPISAKSLPALADLEDQLRAYLTATPETDLPVASTLAMTRSVFEHRAVLLGEETVTGTAI DPRI
20 VFVFSGQGSQVRGMGEELAAAFPLFARLHRQVWDLDDVPDLVDVDTGYVQPALFALQVALFGLLESWVRPRAVIGH
SVGEVAAGYVAGVWSLEDACALVSARARLMQALPAGGAMVAVPVSEERARVALVDGVEIAAVNGPASVVLSGDEAAV
LQIAEGLGRWTRLSASHAFHSVRMEPMLLEEFQVASELTYQEPRIVMAAGEQVTTPEYWVRQVRD TVRFGDQVAAFQ
DAVFLEIGPDRTLSRLIDGIAMLDGDDEVRAAVAALAEHLHVQGVPI DWPAVLGTTTGRVLDLPTYAFQHORYWAAS
DRPAGDGHPLLDTVVALPGADGVVLTGRISLATHAWLADHAVRGTVLLPGTG FVEMVARAAAEEVGC AAVDELVIEAP
25 LLLPASGGVQLSVSVGEADDAHRPVTVHSQLADETEAWVRHVTATISPSPGPIVSPPEFEVWPPAQAEPEVARFYDE
LAAAGYEYGA AFQGLRAAWRAGETIYAEVVLAEQDTLEAARFTVHPALLDAALQANILNASGDLRLPFSWGQVQFHT
TGAATLRVAVTPVADGWTIQATDDAGRPVATVGSVVARPVAGLGATAEDLFALTWNEIPAPGQGGRTVGRFEDLADD
GPVPELVVFTALPDVDADPLVRTRALTARVLEAIQRWLGEPRFADSTLVVRTGTDLASAAVSGLVRS AQSEHPDRFI
LVEGDSSPVEIGLDEPWLRVDGGRYEVPRLIRLSAEPVQEAAWNPDGMVLITGGTGALAGILARHLVAENKARRLLL
30 VSRVPPDALISELTELGAEVGTAVCDVSDRAALARVLAGVPSLTAVIHTAGVLDGVMESLTPQRLD TVLRKADG
AWHLHELTRDRDLAAFMYSAAAGLMGSPGQGNAAAAFLDALAVERRAEGLPALSLAWGFWEETTGLTANLTGAD
RDRIRRGGLQTITAERGMRMFDATQHGEPVLLAAPISPVRDGEVPALLRSLHRRGTRRGTTADASAQWLAGLAPEE
REGALIKVVRDTAAVVLGHADAGTIPVTA AFKDLGLDSL TAVELRNSLAKSTGLRLPATMVFDPYTPASLAARLDDL
MNPRVSSTALLAELDRIEGMFD SVTFDEKQASLVKDRLS AALGKWQQISRSADVATVALANADAGEILDFIDREFGN
35 PTI

SEQ ID NO: 176

MPDHDKLVEYLRWATAELHTTRAKLQAATEAGTQPLAIVGMACRLPGGVSSPEDLWRLVESGTD AISGFVPDRGWDV
DGLYDPDPDVP GKSYTVEGGFLDAVTGFDAPFFGISPREALAMPQQRLVLEASWEAFERAGIEPGSVRGSDTGVM
40 GAFPGGYGTGADLGGFGMTGGAASVLSGRVSYFFGLEGPAMTVDTVCSSSLVALHQAAGYALRHGECSLALVGGVTVM
STPQTFVEFSRQRLAADGRCKAFADNADGTGWSEGVGVLLVERLSDAQARGHNLAVVRGSAVNQDGASNGLTAPN
GPSQQRVIRQALANAGLTGADVVDVEAHGTGTTLGDP IEAQAVIATYGRDRDQPVLLGSLKSNLGHGTQAAAAGVSGVI

KMVMALQNGVVPRTLHIEEPSRHVDWTAGAVQLVTENRPWPELGRARRAAVSSFGLSGTNAHVILESAPDQPPAPTT
DTPVSAVTAGVVPLPISAKTVPALADLEDRLRTYLTTTPDTPDAVASTLATTRSLFEHRAVLLGEDTVTGTAI PDP
RVVFVFPQGQWQWQGMGSALLTSSTVFAERMAECAAALSEFVDWDLTVLDDPSVVDVQVQACWAVMISLAAVW
QAAGIHPDIVLGHSQGEIAAACLAGAISLPDAARIVAQRSQLIAHQLTGHGAMASISLPADDIPTTDKVVIAAHNGT
5 STVIAGDPQAVEAVLATCETRGARVRKINVYASHTPHVEQIRTELLEDITTGIEAHTPAVPWLSTTDNTWIDQPLDP
TYWYRNLREPVRFGPAIDLLQTQDNNLFIEISASPVLLQTMDNAATVATLRRDEDTTQRLLTAFABEAHVHGATIDWP
TVLDTTTTTPVLDLPTYPFQRQRYWATSNRSTGQGHPLLETVVALPGTDGVALTGRISLATHPWLT DHTVRGTVLLP
GTAFVELVTRAATEVNCQIIDELIIEAPLPLPQTDGVQLSVTVGEADEAGHRPVTVYSQTDSEDDWIQHVTATIGPG
ASLPETAAWPPAHAEPVNVGTGLYDNLAAAGYEYGPAPFQGLQAAWRAGDTVYAEVTLAEQQAQETARFTMHPALLDAA
10 LHTIALHDTGDLHLPFSWTRVQFHGTGAATLRVAVTPAADGWNIRATDDTGRAVATIGSLVTRPMAAETDDDLLALT
WTEIPAPEFPVDPTDVVVFTALPDTVEDVPAQTRALTTRVLHTIQEWLADDDRTLIVRTGTDLASAAVSGLVRSAQSE
HPGRFILVESADEALTQEQLAATAGLDEPRLRITGGRYEVPRLTREDTALAVPTDRAWLLEQPRSGSLEDLALLPTD
AAERPLQAGEVRIGVRAAGMNF RDVVVALGMVTDTRLAGGEAAGVVLEVGTVDVNDFRPGDRVFGILEGGFGSVAICD
HRTLAVIPDGWSFTTAASVPIAFATAYYGLVDLAGLRAGESVLIHAATGGVGIAATQIARHLGAEIYG TASVGKQHV
15 LRDAGLADDRIADSRTTDFRDTFRDGTQGRGVDVVLNSLRGEFIDASLDLLVDGGRFIEMGKTDIRDAAQIPDATYH
AFDLMDAGHDLREIMTELLALFEQGV LHPMPVHAFDIRQAREAFSWSMRARHIGKLVLTIPQPIDPDGTVLITGGS
GVLAGIVARYLV TENRARHLLLLSRSAPSASLIDELTALGAHV DVAACDVADRAALAEILDGVDLTAVIHTAGALDD
GVVESLTPQRLD TVLTPKADGAWHLHELTRDRDLAAFIVYSSAAGVLGAAGQGNYAAANAFDLALAVHRRLEGLPGL
SLAWGLWEDASGLTADLTADRDRIRRSGQRAITAAYGMRMLDAATRQSEAILLAAPISPIQDGDVPAILRSLHRRV
20 GRRASVAHGHPADLTPEALLKVVRDSAAMVLGHTNADTVPTATAFQELGLDSL TAVELRNSLT KATGLRLPATMAFD
YPTPDALAARLGELLAGEAAPKAAAARRATASDEPLAIVGMACRLPGGVSSPEDLWRLVESGTDAITDFPTDRGWD
TDTLFDPDPTPGKTYTVHGGFLNDVAGFDAPFFGISPREAVAMPQQRLVLESSWEAFERAGIQPDSIRGSDTG VF
MGAYPDGYGIGADLAGFGVTAGAGSVLSGRVSYFFGLEGPAMTVDTACSSSLVALHQAAYALRQGECSLALVGGVTV
MPSPRTFIEFSRQRLAADGRSKAFADAADGTGFSEGVGVLLVERLSDAQAKGHNILALVRSSAVNQDGASNGLTAP
25 NGPSQQRVIQSALAGAGLTSADVDVVEAHGTGTTLGDP IEAQAVLATYQGQDRDQPVLLGSLKSNLGH TQAAAGVSGV
IKMVMALQHNTVPATLHVDAPS RHVDWTAGAVRLATENQFPWPETNRPRRAGVSSFGVSGTNAHVILEQAPAA SPVEP
VDTTDVVIPLVVSARSSGSLSDQADRLAALVGSPDAPALTSLADALLTRRTVFSQRAVVVAGSHEQAAAGLRALASG
DSHPALVTGAAGPARGVVLVFPQGSQWAGMGAELLDTS PVFAARIAECAEALRPWVDWSLDEVLRGDASADVLGRV
DVVQPASFVVMVGLAAVWESAGVRPDAVLGHSQGEIAAAYVAGALS LTDAKIVAVRSRLIAARLAGRGMASVALA
30 PDEAAAKLGRTELA AVNGPASVVIAGDAEALDET LAMLEGEAVRVRRAVDYASHTPHVEELEQSM AEALADVRSRQ
PRVGFLSTVTGDWVTEAGALDGGYWRNL RQPVRFGPAVASLAEAGYTVFVEASHPVLVQPV AETLDRTDAVVTGT
LRRQDGG LPRLLTSMAELFVGGVPVNWVLLPAGAVRGWVDLPTYAFDHQRYWLENRVATDAAA LGLAGADHPLLGA
IVAVPQSGGVAMTSRLSPRNHPWLA EHTLGGVPTVPTSVLVELAVRAGDEVGCGVVEELTVDAPLLLPERGGVRVQV
IVGATDANGQRGLDIFSAPEDTGQEA WTRHATGTLAPGGDIAADV DLSAWPPANAQPV DVTG DYDLLERAGYGYGPA
35 FQGVRAIWRERGEELFAEVALEPELTDTAARFGLHPALLDAAWHPELRDEVAETSPDGRRWWSQPSRWAGLR LHTAGA
TVLRVRLAPVDADSM SLQAADETGDPVLTVD SLSLCAVSADQLTTAESSDDALFRLEWTPLSKAPTAA RSWVPVETG
ADVAALDGQAVVDAMVLEAAGTGDAL ELTCRVLEV VQAWLTLPGWDESRLVVVTRGAVGAVGDPAGSA VWGLVRAAQ
AENPDRIALLDL DGGRPVEPLLAESE PQLAIRGAELVPRLIRAAAATDAPALFDESQTVLITGGT GSLGGLLARHL
VGRYGLRRLVLVSRRGPDAPGAYELAAELAAHGAEAAALVACDLTDRDAVARLLTEHHPTAVVHAAGVSDDGVIGTLT
40 SDRLAYVFGPKATAARHLDEL TRELLPD LAAFVTYSSISAVFLGAGSGGYAAANAYLDGLMARRHAEGLPGLSLAWG
LWDQEADGGGMAAGLQDITRNMRRRRGGVLSFTPAEGMALF DAAMATDEALVVPVR LDPALRAEAVAEGRSAPVLL
RGLVRPGRRLARTVSGGTGV LADLTPEALLKLVRGAAAVLGHVDADAVPVAAAFKDLGVDSL TAVELRNSLAKATG

LRLPATLVFDYPTPTVLAGRLGELLAGGTAPVRAAVVRRRAASDEPLAIVGMACRLPGGVLSPEDLWRLVESGGDAI
SGFPVDRGWDVENLFDPDPAAGRTYAVRGGFLDGAAGFDASFFGISPREAQAMPDQQRLVLEVSWEAFERAGIEPG
SVRGSDTGVMGAYPGGYGMGTDLGGFGMTSVAVSVLAGRVSYFFGLEGPAMTVDTACSSSLVALHQAGSALRQGE
SLALVGGVTVMPTPQTTFVEFSRQRGLAADGRCKAFADAADGTGFSEGVGVLLVERLSDAQARGHNILAVVRGSAVNQ
5 DGASNGLTAPNGPAQQRVIQSALAGAGLASADVVDVEAHGTGTTLGDP IEAQAVIATYGQDRDQPVLLGSLKSNLGH
TQAAAGVSGVIKMMALQNGVVPRTLHIDEPSRHIDWTAGAVELVTENQSWPETGRARRAAVSSFGISGTNAHVILE
SAPAQPVPVLVDTPVSDVTAGVVPLPISARTVPALADLEDQLRAYLTTAPETDLPVASTLAMTRSVFEHRAVLLGEE
TVTGIASVSDPRVVFVSGQGSQRVGMGEELAAAFPLFARLHRQVWDLLDVPDLEVD DDTGYVQPALFALQVALFGLLE
SWGVRPRAVIGHSVGEVAAGYVAGVWSLEDACTLV SARARLMQALPAGGAMVAVPVSEERARAVLVDGVEIAAVNGP
10 ASVVLSGDESAVLRVAEGLGRWTRLSASHAFHSVRMEPMLEEF RQVASELTYREPRIVMAAGEQVTTPEYWVRQVRD
TVRFGDQVAAFGDAVFLEIGPDRTL SRLIDGIAMLDGDDEVRAAVAALAMMHVQGVGVDPVAVLGT TTTGRVLDLPTY
AFQHERYWMVSTGRPGGEGHPLL GWGVPVAEADGRLYTGRVARQDGPVLPVAAFVEMAF AAAGGRPIRELSVDALLY
IPDDGTAELQ TWVSEHRLTIHARYRDT EPWTRLATATLDTTEPATHTPHPLITTALT LTGDEAPAIWHDLT LHTS
NATELHTHITPGDDGLTITATDATGQPVLT AHAATPTTIPVHTPTTPADDLLTLTWTQIPTPGPGDGADIAVCTAL
15 PDPDS DPLAQTRTLTAQVLHSIQASLTGEDRPLVVHTGTGLASAAVSGLVRSAQSEHPDRFILVESDETLTPDQLAA
VAGLDEPWLRI TDGRYEVPRLT KTTTTATATAVSEP VWDPDGTVLITGGSGALAGILARHLVTERGVRHLLLVS RST
PSTTLIDELRELGAHV DVAACDVSDRAALARVLDGVDLTAVFHTAGALDDGVVESLTPQRVDAVLRPKADGAWHLHE
LTRDRDLTAFVMYSSAAGVMGAAGQGN YAAANAFLDALAEHRRADGLPALSLAWGMWDDADGMTASLSGTDHRRIRR
SGQRAITA EHGMRLLDKASGRSEPV LVATAMNPIPD TDLPALRLSLYPKTARKSQPIQELSPEALLKIVRDS AAMVL
20 GHANADTVPTATALQELGLDSL TAVELRNSLT KATGLRLPATMAFDYPTPAALAGRLGELLAGD TTPATAAVRRAT
ASDEPLAIVGMACRLPGGVSTPEDLWRLVESG TDAITDFPTDRGWDTD DLFDPDPDTPGKTYTVHGGFLDDVAGFDA
SFFGISPREALAMDSQQRLVLEAAWEAFERAGIEPGSVRGSDTGVMGAYPDGYGIGADLGGFGATAGAGSVLSGRL
SYFFGLEGPAMTVDTACSSSLVALHQAGSALRQGECSLALVGGVTVIANPQIFVEFSRQRGLAADGRCKAFADNADG
TGFSEGVGVLLVERLSDAQAKGHN ILALVRSSAVNQDGASNGLTAPNGPSQQRVIRQALANAGLTGA EVDVVEAHGT
25 GTTLGDP IEAQAVLATY GQDRDQPVLLGSLKSNLGH TQAAAGVSGVIKMMALRHDTVPATLHIDEPSRHIDWTAGA
VELVTENQWPVVLGRPRRAAVSAFGVSGTNAHVILE SAPDQPPAPATDTPAPAATAGVVPLPISAKTVPALADLEDR
LRTYLTTTTPETDLPVASTLAT TTRSLFEHRAVLLGEDTVTGT TTPDPRIVFVFPQGQWQWQMGGSALLTSSTVFAER
MAECAAAALSEFVDWDL TLVDDPSIVDRVDVVQPACWAVMISLA AVWQAAGIHPDIVLGHSQGEIAAACLAGAISLP
DAARIVAQRSQLIAHQLTGHGAMASISLPADDIPTTDKVWIAAHNGTSTVIAGDPQALD TVLATCETHGARVRKINV
30 DYASHTPHVEQIRTELLDITTDIEAHTPTVPWLSTTDNTWIDQPLDPTYWYRNLREPVRFGPAIDLLQTQDNNLFIE
ISASPVLLQTM DNATTVATLRRDEDTTQRLLTAF AEAHVHGATIDWPTVLD TTTTTPVLDLPTYPFQRQRYWATSNGR
PTSQGHPLLETVVALPGTHGVALTGRISLATHPWLTDHTVRGTVLLPGTAFVELVTHAATEVNCQVIDELIIEAPLP
LPQNGGVQLSVTVGEADEAGHRPVTVYSQTDESDDWVQHVTATIAPGVSSSESAAWPPAQAEPVNVTGLYDNLAAAG
YEYGPAFQGLQTAWRDGSTVYAEVTLAE EQA QETARFTMHPALLDAALHTIALHDTADLQLPFSWRQVQFHGSGAAT
35 LRVAVTPAADGWNIRATDDTGQTVATIGSLVTRPMAAETTNDLLALTWTEIPAPEPVDPADV VVFTALPEPGSDPLA
QTRALTTRVLHTIQEWLADDDRTLIVRTGTD LASAAVSGLVRSAQSEHPGRFILVESDDETLTHEQLAATAGLDEPR
LRITDGRYEVPRLTREDTALAVPEGGA WMLDQPSRSGTLQDLRLVPTDAAERPLRPGEVRVGVRAAGLNFRDVAVAL
GMVTDTRLIGGEGAGVVLEAGPGVEDLRPGDRVFG LLEGGFGPVAVADRRALALIPDGWSFTTAASVP IAFATAYYG
LLDLAGLRAGESVLIHAATGGVGMAATQIARHLGADVYATASTGKQHVLRDAGLSDDRIADSRTTGFRETF RDSTDG
40 RGVDVVLNSLKGDFVDASL DLLVDGGRFIEMGKTDIRDAAQIPDATYRAFDLMDAGPERLREIITELLALFEQGVLR
PLPVHAFDIRQARDAFGWMSRARHIGKLVLTIPQPIDPDGTVLITGGSGVLAGIVARHLVIAEGLRNLLLLSR SAPS
EALIGELTALGAQVETAACDIADRAALARVLDGVPLTAVIHTAGALDDGVVESLDPQR LDSVLTPKADGAWHLHELT

RDRDLAAFIMYSSAAGVLGAAGQGNIAAANAFVDALAVHRRFMGLPALSLAWGLWDDTSALTAGLTDSHDRIIRRS
 ARTITAHEGMRMFDAATRQSEAVLLAAPMGPIRGEDVPALLRGLATVRQPRTRAKRDMGPERLRDLRLNGRTSVEQHR
 IMVELVLAHATSVLGHESPDIAIPDRAFKDLGMDSLTAIELRNHLVAETGVRLPATTAFDHPHTADDLAKRLLAEVGL
 TPAPQRTEADIREEVVVREPAGDDSWTSEPIAIVSMSCRAPGGVDSPELWRLVESGTDAITDFPGDRGWDVAGLYS
 5 PDPDTGYKTYCVQGGFLDAAADFDAFFGISPREALGMDPQQRLLETSWEAIERARIDPRSLRGRNVGVYVGAAQ
 GYGVGAIDQQRDNVITGSSISLLSGRLSYALGLEGPVTVDTACSSSLVALHLACQALRQRECSMALVSGVSVIPTP
 DVFVEFSRQRLAADGRCKSFSSASADGTIWAEGVGVLVLERLSEATRLGHRVLAVVRGSAVNSDGASNGLTAPNGVS
 QQRVIRQALTGAGLTAADVVDVEAHGTGTLGDPIEAAILATYQDRSTPVCLGSLKSNIGHAMAASGVLAVIKMV
 EAMRHGLIPRTLHVEEPSPHVDWASGDVALLTENQPPWDDAKLRAGVSSFGLSGTNAHVVLEQYRAPAAPDITTE
 10 HEPLAWTLSARDPKALREQAGRLHAALTESPQWRPLDIGYSLATTRSNFAHRAVAVGSDREDLLRALS KLADGSAWP
 ALVTATAKDRRVAYLFDGQGSQRPDMGSGLYERFPFAFARAWDRISAIEFGKHLHDHSLTDVYLGGRDAATADLVDDTLY
 AQAGLFTMEIALFELLAEWGVRPDFVSGHSIGETAAAYAAGVLSLEDVTTLIVARGRALRQVPPGAMVALRAGEDEA
 REFLGRTGAALDLAAVNSPTSVVVS GASEAVAGFRARWTESGREARTLNVRHAFHSRHVEAVLGEFREVLES LTFRT
 PALPVVSTVTGR LIEPTELSTSEYWL RQVRQTVRFHDVARELSGQGVGT FVEIGPSGALASAGLECLGDEASFHAVQ
 15 RPGSPGDVCLMTAVALHAGGTTVDWATVLAGGRATDLPVYPFQHGSYWLAPVTRAADGAPSAGVPAPGEYARPSAP
 EEPRTMLELVRLEAAIALSITDPGLIADDSSFLDLGFDSSISALRLSNRLAAVTGLDLPPSLLFDHPTPAELAAARLDE
 LSAADLDGAGVYALLEEIDELDDEDLDMTEEEQT AISSELLTKLSAKWSR

In some embodiments, the disclosure provides a chimeric polyketide synthase where at least one
 20 module includes a portion having at least 90% identity to any one of SEQ ID NO: 1-174.

In another aspect, the disclosure provides a nucleic acid encoding any one of the above
 described polyketide synthases.

In some embodiments of any of the above described aspects, the nucleic acid encoding any one
 of the above described polyketide synthases further encodes an LAL in which the sequence encoding the
 25 LAL is operatively linked to the sequence encoding the polyketide synthase.

In some embodiments, the LAL may be a heterologous LAL.

In some embodiments, the LAL may include a portion having at least 80% (e.g., at least 80%, at
 least 85%, at least 90%, at least 95%, at least 99%) sequence identity to SEQ ID NO: 177. In some
 embodiments, the LAL may include a portion having the sequence of SEQ ID NO: 177. In some
 30 embodiments, the disclosure provides a nucleic in which the LAL has the sequence of SEQ ID NO: 177.
 In some embodiments, the LAL lacks a TTA inhibitory codon in an open reading frame.

SEQ ID NO 177:

MPAVESYELDARDDELRRLEEAVGQAGNGRGVVVTITGPIACGKTELLDAAAASDAITLRAVCSEEEERALPYALIG
 QLIDNPAVASQLPDPVSMALPGEHLSPEAENRLRGDLTRTLLALAAERPVLIGIDDMHHADTASLNCLLHLARRVGP
 35 ARIAMVLTELRLTPAHSQFHAELLSLGHREIALRPLGPKHIAELARAGLGPVDVDEDVLTGLYRATGGNLLNLGHGL
 IKDVREAWATGGTGINAGRAYRLAYLGS LYRCGPVPLRVARVA AVL GQSANTTLVRWISGLNADAVGEATEILTEGG
 LLHDLRFPHPAARSVVLNDLSARERRRLHRSALVLDVDPVEVVAHHQAGAGFIHGPKAAEIFAKAGQELHVRGELD
 AASDYLQLAHHASDDAVTRAALRVEAVAIERRRNPLASSRHDELTVAAARAGLLSLEHAALMIRWLALGGRSGEAAE
 VLAAQRPRAVTDQDRAHLRAAEVSLALVSPGASGVSPGASGPDRRRPRPLPPDELANLPKAARLCAIADNAVISALHG
 40 RPELASAEAEENVLKQADSAADGATASALTALLYAENTDTAQLWADKLVSETGASNEEEGAGYAGPRAETALRRGDL
 AA AVEAGSAILDHRRGSLLGITAALPLSSAVAAAIRLGETERAEKWLAEPLEAIRDSLFGLHLLSARGQYCLATGR
 HESAYTAFRTCGERMNRNWGDVPGLSLWRVDAAEALLHGRDRDEGRRLIDEQLTHAMGPRSRALT LRVQAAYSPQAQ

RVDLLEEAADLLLSCNDQYERARVLADLSEAFSALRHHSRARGLLRQARHLAAQCGATPLLRRRLGAKPGGPGWLEES
GLPQRIKSLTDAERRVASLAAGGQTNRVIADQLFVTASTVEQHLTNVFRKLGVKGRQHLPAELANAE

In some embodiments of any of the foregoing nucleic acids, the nucleic acid includes an LAL binding site, in which the sequence encoding the LAL binding site is operatively linked to the sequence encoding the polyketide synthase.

In some embodiments, the LAL binding site includes a portion having at least 80% (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 99%) sequence identity to the sequence of SEQ ID NO: 178 (CTAGGGGGTTGC). In some embodiments, the LAL binding site includes a portion having the sequence of SEQ ID NO: 178 (CTAGGGGGTTGC). In some embodiments, the LAL binding site has of the sequence of SEQ ID NO: 178 (CTAGGGGGTTGC). In some embodiments of the above described aspect, the LAL binding site has the sequence of SEQ ID NO: 179 (GGGGGT).

In some embodiments of any of the foregoing nucleic acids, the binding of an LAL to the LAL binding site promotes expression of the polyketide synthase.

In some embodiments of any of the foregoing nucleic acids, the nucleic acid encoding any one of the above described polyketide synthases, further encodes a nonribosomal peptide synthase.

In some embodiments of any of the foregoing nucleic acids, the nucleic acid encoding any one of the above described polyketide synthases further encodes a P450 enzyme.

In some embodiments of any of the foregoing nucleic acids, the nucleic acid encoding any one of the above described polyketides and a first P450 enzyme, further encodes a second P450 enzyme.

In another aspect, the disclosure provides an expression vector including any of the foregoing nucleic acids. In some embodiments, the expression vector may be an artificial chromosome, e.g., a bacterial artificial chromosome.

In another aspect, the disclosure provides a host cell including any of the above described expression vectors.

In another aspect, the disclosure provides a host cell including any of the foregoing polyketide synthases, in which the polyketide synthase is heterologous to the host cell.

In some embodiments of any of the foregoing host cells, the host cell naturally lacks an LAL and/or an LAL binding site.

In some embodiments of any of the foregoing host cells, the host cell includes an LAL capable of binding to an LAL binding site and regulating expression of a polyketide synthase. In some embodiments, the LAL and/or LAL binding site may be heterologous to the cell. In some embodiments, the host cell includes an LAL with a portion having at least 80% (e.g., at least 80%, at least 85%, at least 90%, at least 95%, at least 99%) sequence identity to the sequence of SEQ ID NO: 177.

In some embodiments of any of the foregoing host cells, the host cell is a bacterium, e.g., an actinobacterium, such as an actinobacterium selected from the group consisting of *Streptomyces ambofaciens*, *Streptomyces hygroscopicus*, or *Streptomyces malayensis*. In some embodiments in which the host cells is an actinobacterium, the actinobacterium is S1391, S1496, or S2441.

In some embodiments of any of the foregoing host cells, the host cell has been modified to enhance expression of a polyketide synthase. For example, the host cell has been modified to enhance expression of a compound-producing protein by (i) deletion of an endogenous gene cluster which expresses a compound-producing protein; (ii) insertion of a heterologous gene cluster which expresses a

compound-producing protein; (iii) exposure of the host cell to an antibiotic challenge; and/or (iv) introduction of a heterologous promoter that results in an at least 2-fold increase in expression of a compound compared to the homologous promoter.

5 In another aspect, the disclosure provides a method of producing a polyketide by culturing any of the foregoing host cells under suitable conditions.

In another aspect, the disclosure provides a method of producing a polyketide by culturing a host cell engineered to express any of the foregoing polyketide synthases under conditions suitable for the polyketide synthase to produce a polyketide.

10 In another aspect, the disclosure provides a method of producing a compound, the method including: (a) providing a parent polyketide synthase sequence capable of producing a compound; (b) determining the compatibility of at least one module of a second polyketide synthase with at least two modules of the parent polyketide synthase; (c) producing a nucleic acid encoding a modified polyketide synthase, wherein the modified polyketide synthase includes at least one module of a second polyketide synthase which has been determined to be compatible with the at least two modules of the parent
15 polyketide synthase.

In another aspect, the disclosure provides a method of producing a compound, the method including: (a) providing a parent nucleic acid encoding a parent polyketide synthase; (b) modifying the parent nucleic acid to create a modified nucleic acid encoding a modified polyketide synthase capable of producing a compound, wherein the modification produces a modified polyketide synthase including at
20 least one heterologous module.

In another aspect, the disclosure provides a method of producing a compound, the method including: (a) providing a parent polynucleotide sequence capable of producing a compound; (b) identifying one or more heterologous modules suitable for replacement of one or more modules in the parent polynucleotide sequence; (c) producing a nucleic acid encoding a modified polyketide synthase,
25 wherein the modified polyketide synthase includes at least one heterologous module identified in step (b).

In another aspect, the disclosure provides a method of producing a plurality of engineered polyketide synthases, wherein each of the plurality of polynucleotides corresponds to an engineered polyketide synthase, and wherein each of the plurality of polynucleotides includes one or more heterologous modules with altered enzymatic activity relative to a reference polyketide. The method
30 includes the steps of: (a) providing a parent polynucleotide sequence encoding a polyketide synthase; (b) identifying one or more modules for replacement in the parent polynucleotide sequence; (c) identifying two or more heterologous modules suitable for replacement for each of the modules identified in step (b); (d) generating a plurality of polynucleotides, wherein each of the plurality of polynucleotides corresponds to an engineered polyketide synthase, and wherein each of the plurality of polynucleotides includes a
35 heterologous module selected from the two or more heterologous modules identified in step (c) in replacement of each of the one or more modules to be replaced identified in step (b)..

Definitions

A "polyketide synthase" refers to an enzyme belonging to the family of multi-domain enzymes
40 capable of producing a polyketide. A polyketide synthase may be expressed naturally in bacteria, fungi, plants, or animals.

As used herein, the term “engineered polyketide synthase” is used to describe a non-natural polyketide synthase whose design and/or production involves action of the hand of man. For example, in some embodiments, an “engineered” polyketide synthase is prepared by production of a non-natural polynucleotide which encodes the polyketide synthase.

5 A cell that is “engineered to contain” and/or “engineered to express” refers to a cell that has been modified to contain and/or express a protein that does not naturally occur in the cell. A cell may be engineered to contain a protein, e.g., by introducing a nucleic acid encoding the protein by introduction of a vector including the nucleic acid.

10 The term “gene cluster that produces a small molecule” or “gene cluster that produces a compound,” as used herein, refers to a cluster of genes which encodes one or more compound-producing proteins.

The term “heterologous,” as used herein, refers to a relationship between two or more proteins, nucleic acids, compounds, and/or cell that is not present in nature. For example, the LAL having the sequence of SEQ ID NO: 177 is naturally occurring in the S18 *Streptomyces* strain and is thus
15 homologous to that strain and would thus be heterologous to the S12 *Streptomyces* strain.

The terms “homologous” or “native,” as used interchangeably herein, refer to a relationship between two or more proteins, nucleic acids, compounds, and/or cells that is present naturally. For example, the LAL having the sequence of SEQ ID NO: 177 is naturally occurring in the S18 *Streptomyces* strain and is thus homologous to that strain.

20 The term “recombinant,” as used herein, refers to a protein that is produced using synthetic methods.

As used herein, the term “reference polyketide synthase” refers to a polyketide synthase that has a sequence having at least 80% identity (e.g., at least 80% identity, at least 85% identity, at least 90% identity, at least 95% identity, at least 99% identity, or 100% identity) to the sequence of an engineered
25 polyketide synthase except to the sequence of the one or more modules which are modified.

As used here, the term “compatibility” refers to a measure of the likelihood of two adjacent modules to form a competent module-module junction, in which polyketide translocation is not substantially inhibited. A heterologous module may be considered compatible if it meets at least one of the following criteria: 1) the module is present in the same module clade as one or more adjacent
30 modules of the reference PKS, as determined by the module-level phylogeny classification described in the detailed description of the invention; 2) the module is assigned a score of greater than or equal to 0.90 in the inter-module covariation analysis algorithm described in the detailed description of the invention; or 3) the module belongs to the same functional clade or sub-clade as one or more adjacent modules of the reference PKS, as determined by the evolutionary trace methodology outlined in the
35 detailed description of the invention.

As used here, the term “linking sequence” refers to a sequence directly upstream or downstream of an inter-modular junction. For example, in a single module swap, the ACP for the upstream homologous module, the ACP and KS-AT didomain of the inserted heterologous module, and the KS of the downstream homologous module may all be considered linking sequences.

40 As used herein, the term “module” refers to a region of a polyketide synthase that includes multiple domains. Modules present in a polyketide synthase may include i) a loading module; ii)

extending modules; and iii) releasing and/or cyclization modules, depending on whether the final polyketide is linear or cyclic. The domains which may be included in a given module include, but are not limited to, acyltransferase (AT), acyl carrier protein (ACP), keto-synthase (KS), ketoreductase (KR), dehydratase (DH), enoylreductase (ER), methyltransferase (MT), sulfhydrylase (SH), and thioesterase (TE).

As used here, the term “acceptor module” refers to a homologous module within a PKS cluster subject to engineering by module swapping. In the resulting engineered PKS cluster, the acceptor module is absent.

As used here, the term “donor module” refers to a heterologous module that is introduced into an engineered PKS cluster.

As used here, the term “module swapping” refers to the exchange of one or more heterologous donor modules for one or more homologous acceptor modules.

As used here, the term “does not substantially inhibit polyketide translocation” refers to the ability of a heterologous PKS module to function in a biosynthetic assembly line. For example, a heterologous loading module does not substantially inhibit polyketide translocation if the loading module is able to load a starter unit onto its ACP domain and pass the starter unit to the KS domain of the adjacent (n+1) extender module. A heterologous extender module does not substantially inhibit polyketide translocation if the extender module is able to receive a starter unit or polyketide chain from the previous (n-1) module, catalyze the addition of an extender unit, and pass the elongated polyketide chain to the adjacent (n+1) module. In some embodiments, a heterologous module does not substantially inhibit polyketide translocation if the engineered PKS that includes the heterologous module produces a compound in levels that are detectable by a highly sensitive detection method, e.g., LC-TOF mass spectrometry.

An extender unit, e.g., a malonyl-CoA, is loaded onto the acyl carrier protein domain of the current module catalyzed by another acyltransferase domain. The polyketide chain is then elongated by subsequent extender modules after being passed from the acyl carrier protein domain of module n to the ketosynthase domain of the n+1 module. The acyl carrier protein bound extender unit reacts with the polyketide chain bound to the ketosynthase domain with expulsion of CO₂ to produce an extended polyketide chain bound to the acyl carrier protein. Each added extender unit may then be modified by β -ketoprocessing domains, i.e., ketoreductase (which reduces the carbonyl of the elongation group to a hydroxy), dehydratase (which expels H₂O to produce an alkene), and enoylreductase (which reduces alkenes to produce saturated hydrocarbons).

Brief Description of the Figures

FIGS. 1A and 1B are schematics illustrating the mechanisms by which PKS biosynthesis proceed. FIG. 1A depicts polyketide chain elongation and β -carbonyl processing within a module. FIG. 1B depicts translation between modules.

FIG. 2A is a diagram depicting complementary bioinformatics approaches to the prediction of functional protein-protein interactions at the module-module junction.

FIG. 2B is a phylogenetic tree resulting from multiple sequence alignments of complete FK-family modules.

FIGS. 2C-2E depict how inter-module residue covariation is used to generate an algorithm that ranks module-module junction compatibility. FIG. 2C is a diagram that illustrates the upstream and downstream module-module junctions used to determine the compatibility of a given heterologous module. FIG. 2D is a correlation map that depicts the alignment of the ACP domain of a given module and the KS-AT didomain of a second module. FIG. 2E depicts the compatibility score resulting from inter-domain residue covariation analysis for a series of heterologous modules. Scores are normalized to the homologous module for the polyketide synthase in question, which is given a score of 1.00.

FIGS. 2F and 2G depict how evolutionary trace analysis is used to predict module-module junction compatibility. FIG. 2F is a phylogenetic tree generated by multiple sequence alignments of FK-family KS and ACP domains, in which group-specific residues have been concatenated into functional clades or sub-clades. The distance between modules can be used to predict module-module junction compatibility. FIG. 2G is a schematic depicting the compatibility relationships predicted by evolutionary trace analysis between KS and ACP domains for the FK-family.

FIG. 3A is a schematic depicting a single module swap in which a donor module replaces either module 3 or module 4 of the PKS gene cluster that produces Compound 1.

FIG. 3B is an image of the engineered PKS that includes the heterologous module 3 from the S17 *Streptomyces* strain in place of the homologous module 3 in the PKS that produces Compound 1. The engineered PKS module 3 now includes an ER domain, and thus, the resulting compound produced by the engineered PKS, Compound 2, is reduced relative to Compound 1.

FIG. 3C is an image depicting compounds, e.g., Compound 2, Compound 3, Compound 4, and Compound 5, produced by single module swaps of either module 3 or module 4 in the PKS that produces Compound 1 with compatible heterologous modules.

FIG. 4A is a schematic depicting combinatorial swapping of a dimodule unit.

FIG. 4B is a schematic depicting the synthesis of dimodule units from exogenous donor modules by a first round of Gibson assembly. The dimodule product is shown as analyzed by DNA gel electrophoresis.

FIG. 4C is a schematic depicting dimodule capture, amplification, and enrichment in a shuttle vector. Dimodule units resulting from a first round of Gibson assembly are captured in a shuttle vector by a second round of Gibson assembly. This allows for the dimodule assembly to be amplified, enriched, and ligated into the intended PKS.

FIG. 4D is a schematic depicting the construction of dimodule libraries by combinatorial synthesis.

FIG. 4E is an image depicting the possible resulting compounds that may be generated by an exemplary dimodule library swapped into module 3 and module 4 of the PKS that produces Compound 1.

FIG. 4F depicts oversampling required for sufficient coverage of a large combinatorial dimodule library. FIG. 4F is a graphical representation of the oversampling required to achieve 90% or greater coverage of a 225 member dimodule combinatorial library. 18% of the 650 sampled clones were found to have produced polyketide compounds resulting from the engineered PKS cluster, as determined by LC-TOF mass spectrometry analysis.

FIG. 4G is a schematic depicting a method of preparing combinatorial dimodule libraries and characterizing the resulting libraries using NanoPore sequencing.

FIG. 4H is a schematic depicting the core informatics workflow for deconvoluting the sequences of combinatorial dimodule libraries by NanoPore sequencing.

FIGS. 5A and 5B depict the construction of trimodule libraries by combinatorial synthesis. **FIG. 5A** is a schematic illustrating a trimodule swap of modules 4, 5, and 6 of the PKS cluster that produces Compound 7, to produce a theoretical library size of 2,197 engineered polyketide synthases. **FIG. 5b** is an image of high efficiency trimodule assembly by Gibson assembly as analyzed by DNA gel electrophoresis.

FIG. 6A is a schematic illustrating a module swap that results in ring expansion by exchanging a single module acceptor for a dimodule donor. The resulting expanded ring compound produced by the engineered PKS, Compound 8, is also depicted.

FIG. 6B is a spectrogram that shows the production of an expanded ring compound, Compound 8, as analyzed by LC-TOF mass spectrometry.

FIG. 7A is schematic depicting the enzymatic domains of five PKS loading modules, including Rapamycin and novel PKS cluster, X23. Also shown is the starter unit associated with each loading module.

FIG. 7B depicts the compounds produced by engineered PKS clusters resulting from single module swaps in the X23 PKS cluster. The products include Compound 11 and 12, which are produced by an engineered PKS that contains a heterologous loading module.

Detailed Description of the Invention

The present invention describes compositions and methods for the production of polyketide compounds by an engineered polyketide synthase that includes one or more heterologous modules. The present invention also describes methods for predicting the compatibility of linking sequences of heterologous module-module junctions to produce an engineered polyketide synthase that does not substantially inhibit translocation during polyketide biosynthesis.

Compounds

Compounds that may be produced with the methods of the invention include, but are not limited to, polyketides and polyketide macrolide antibiotics such as erythromycin; hybrid polyketides/non-ribosomal peptides such as rapamycin and FK506; carbohydrates including aminoglycoside antibiotics such as gentamicin, kanamycin, neomycin, tobramycin; benzofuranoids; benzopyranoids; flavonoids; glycopeptides including vancomycin; lipopeptides including daptomycin; tannins; lignans; polycyclic aromatic natural products, terpenoids, steroids, sterols, oxazolidinones including linezolid; amino acids, peptides and peptide antibiotics including polymyxins, non-ribosomal peptides, β -lactams antibiotics including carbapenems, cephalosporins, and penicillin; purines, pteridines, polypyrroles, tetracyclines, quinolones and fluoroquinolones; and sulfonamides.

Proteins

Polyketide Synthases

Polyketide synthases (PKSs) are a family of multi-domain enzymes that produce polyketides. Type I polyketide synthases are large, modular proteins which include several domains organized into

modules. The modules generally present in a polyketide synthase include i) a loading module; ii) extending modules; and iii) releasing and/or cyclization modules depending on whether the final polyketide is linear or cyclic. The domains which generally are found in the modules are acyltransferase (AT), acyl carrier protein (ACP), keto-synthase (KS), ketoreductase (KR), dehydratase (DH), enoylreductase (ER), methyltransferase (MT), sulfhydrolase (SH), and thioesterase (TE).

A polyketide chain and the starter groups are generally bound to the thiol groups of the active site cysteines in the ketosynthase domain (the polyketide chain) and acyltransferase domain (the loading group and malonyl extender units) through a thioester linkage. Binding to acyl carrier protein (ACP) is mediated by the thiol of the phosphopantetheinyl group, which is bound to a serine hydroxyl of ACP, to form a thioester linkage to the growing polyketide chain. The growing polyketide chain is handed over from one thiol group to another by trans-acylations and is released after synthesis by hydrolysis or cyclization.

The synthesis of a polyketide begins by a starter unit, being loaded onto the acyl carrier protein domain of the PKS catalyzed by the acyltransferase in the loading module. An extender unit, e.g., a malonyl-CoA, is loaded onto the acyl carrier protein domain of the current module catalyzed by another acyltransferase domain. The polyketide chain is then elongated by subsequent extender modules after being passed from the acyl carrier protein domain of module *n* to the ketosynthase domain of the *n*+1 module. The acyl carrier protein bound extender unit reacts with the polyketide chain bound to the ketosynthase domain with expulsion of CO₂ to produce an extended polyketide chain bound to the acyl carrier protein. Each added extender unit may then be modified by β -ketoprocessing domains, i.e., ketoreductase (which reduces the carbonyl of the elongation group to a hydroxy), dehydratase (which expels H₂O to produce an alkene), and enoylreductase (which reduces alkenes to produce saturated hydrocarbons). Once the synthesis of the polyketide is complete, a thioesterase domain in the releasing modules hydrolyzes the completed polyketide chain from the acyl carrier protein of the last extending module. The compound released from the PKS may then be further modified by other proteins, e.g., nonribosomal peptide synthase. In some cases, the biosynthetic cluster harbors polyketide megasynthases and a non-ribosomal peptide synthase (NRPS). This hybrid architecture is referred to as hybrid PKS/NRPS.

Polyketide synthase extender modules

PKS biosynthesis proceeds by two key mechanisms: polyketide chain *elongation* within a module and *translocation* between modules (FIGS. 1A and 1B). The basic functional unit of polyketide synthase clusters is the extender module, which encodes a 2-carbon extender unit derived from malonyl-CoA. Within the extender module, the minimal domain architecture required for polyketide chain elongation includes the ketosynthase (KS), acyl-transferase (AT) and the ACP (acyl-carrier protein) domains, and the specific chemistry of each module is encoded by the AT domain and by the presence of the beta-carbonyl processing domains: ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) domains. Productive chain elongation depends on the concerted function of numerous domains

β-ketone processing domains

β-ketone processing domains are the domains in a PKS which result in modification of the elongation groups added during the synthesis of a polyketide. Each β-ketone processing domain is capable of changing the oxidation state of an elongation group. The β-ketone processing domains include ketoreductase (which reduces the carbonyl of the elongation group to a hydroxy), dehydratase (which expels H₂O to produce an alkene), and enoylreductase (which reduces alkenes to produce saturated hydrocarbons).

Module swapping to produce engineered polyketide synthases

The present disclosure provides methods and compositions related to engineered polyketide synthases produced by swapping modules between related PKS clusters. Polyketide translocation is controlled by protein-protein interactions at the inter-modular junctions. In some embodiments, module swapping is guided by bioinformatic predictions to determine which modules have the highest probability of functioning in assembly-line polyketide biosynthesis. Multiple bioinformatics methods are used to determine the structural information in PKS sequence alignments to predict protein-protein interactions that mediate polyketide translocation at the inter-modular junction. The present disclosure includes a DNA assembly strategy to swap one or more heterologous donor modules for one or more acceptor modules to generate hybrid PKS clusters.

In some embodiments, module swapping is achieved by single, di- or tri-, or multi-module capture. In some embodiments, module swapping may be performed by exchange of the loading module. In some embodiments, module swapping may be performed by exchange of one or more extender modules. In some embodiments, module swapping may be performed by exchange of one or more releasing or cyclization modules. In some embodiments, two or more heterologous donor modules may replace a single acceptor module which may result in the production of a ring-expanded compound. In some embodiments, a single heterologous donor module may replace two or more acceptor modules which may result in a contracted ring compound. In some embodiments, the engineered polyketide synthases may produce novel compounds.

Combinatorial libraries of engineered polyketide synthases

In some embodiments, the pooled capture and transfer of single, di- or tri-, or multi-module units enables the production of combinatorial libraries of engineered polyketide synthases. A dimodule unit, for example, consists of two heterologous modules, each of which may be independently selected from a pool of heterologous modules. A trimodule unit, example, consists of three heterologous modules, each of which may be independently selected from a pool of heterologous modules. One or more modules of a polyketide synthase may be replaced with a single, di-, tri-, or multi-module unit, where the single, di-, tri- or multi-module unit is selected from a pool of single- di-, tri- or multi-module units produced by combinatorial synthesis. Exemplary methods for the production of combinatorial libraries of engineered polyketide synthases (e.g., dimodule and trimodule combinatorial libraries) are provided in Examples 2 and 4.

Characterization of engineered PKS libraries by single-molecule long-read sequencing

In some embodiments of the invention, single-molecule long-read sequencing technology (e.g., Nanopore sequencing or SMRT sequencing) may be used to characterize libraries of engineered polyketide synthases which are produced by any of the methods described herein. In particular, single-molecule long-read sequencing (e.g., Nanopore sequencing or SMRT sequencing) may be used to characterize (e.g., deconvolute) combinatorial libraries of engineered polyketide synthases (e.g., combinatorial libraries of engineered polyketide synthases which are produced by pooled capture and transfer of single, di- or tri-, or multi-module units). Single-molecule long-read sequencing enables the identification of the module or modules which are incorporated into the combinatorial library. This further enables the prediction of the chemistry of the resulting plurality of engineered polyketide synthases. The predicted enzymatic chemistry can therefore be connected to the compounds produced by the engineered polyketide synthases. The resulting compounds may be identified by chemical methods of analysis known to one of skill in the art (e.g., mass spectrometry or high performance liquid chromatography). Furthermore, the predicted enzymatic chemistry can be connected to the function of the resulting compounds (e.g., binding to a target protein or inducing a phenotype, such as a cell based phenotype). Accordingly, long-read sequencing of a genetically encoded molecule may allow for genotypic-phenotypic linkage.

Single-molecule long-read sequencing technologies may be considered to include any sequencing technology which enables the sequencing of a single molecule of a biopolymer (e.g., a polynucleotide such as DNA or RNA), and which enables read lengths of greater than 2 kilobases (e.g., greater than 5 kilobases, greater than 10 kilobases, greater than 20 kilobases, greater than greater than 50 kilobases, or greater 100 kilobases). Single-molecule long-read sequencing technologies may enable the sequencing of multiple single molecules of DNA or RNA in parallel. Single-molecule long-read sequencing technologies may include sequencing technologies that rely on individual compartmentalization of each molecule of DNA or RNA being sequenced.

Nanopore sequencing is an exemplary single-molecule long-read sequencing technology that may be used to characterize libraries of engineered polyketide synthases that are prepared by any of the methods described herein. Nanopore sequencing enables the long-read sequencing of single molecules of biopolymers (e.g., polynucleotides such as DNA or RNA). Nanopore sequencing relies on protein nanopores set in an electrically resistant polymer membrane. An ionic current is passed through the nanopores by setting a voltage across this membrane. If an analyte (e.g., a biopolymer such as DNA or RNA) passes through the pore or near its aperture, this event creates a characteristic disruption in current. The magnitude of the electric current density across a nanopore surface depends on the composition of DNA or RNA (e.g., the specific base) that is occupying the nanopore. Therefore, measurement of the current makes it possible to identify the sequence of the molecule in question. Exemplary methods for the use of Nanopore sequencing to characterize combinatorial libraries of engineered polyketide synthases are provided in Example 3.

Single molecule real-time (SMRT) sequencing (PacBio) is an exemplary single-molecule long-read sequencing technology that may be used to characterize libraries of engineered polyketide synthases that are prepared by any of the methods described herein. SMRT is a parallelized single molecule DNA sequencing method. SMRT utilizes a zero-mode waveguide (ZMW). A single DNA

polymerase enzyme is affixed at the bottom of a ZMW with a single molecule of DNA as a template. The ZMW is a structure that creates an illuminated observation volume that is small enough to observe only a single nucleotide of DNA being incorporated by DNA polymerase. Each of the four DNA bases is attached to one of four different fluorescent dyes. When a nucleotide is incorporated by the DNA
 5 polymerase, the fluorescent tag is cleaved off and diffuses out of the observation area of the ZMW where its fluorescence is no longer observable. A detector detects the fluorescent signal of the nucleotide incorporation, and the base call is made according to the corresponding fluorescence of the dye.

Computational approaches for the prediction of functional inter-modular junctions

10 The present disclosure provides complementary bioinformatic approaches for the prediction of functional protein-protein interactions at the module-module junction (FIG. 2A). In some embodiments, these bioinformatic approaches serve as the predictive basis for the design of chimeric PKS proteins by module swapping.

Module-level phylogeny

15 Sequence divergence between polyketide modules and inter-module linkers suggests importance in module-module compatibility. In some embodiments, a module-level phylogenetic map may be constructed by multiple sequence alignment of PKS modules. For example, a module-level phylogenetic map was generated by multiple sequence alignments of complete FK-family modules (FIG. 2B). This
 20 enabled the identification of 10 module clades including 8 elongation, 1 loading, and 1 off-loading. In some embodiments, a heterologous module is compatible if it is present in the same module clade as the adjacent modules.

Inter-module residue covariation

25 Inter-module residue covariation across the intermodule junction was computed to generate an algorithm to rank order intermodule compatibility (FIGS. 2C-2E). Type I polyketide synthase protein sequences were extracted from Genbank and an internal database using Hidden Markov Models trained on the ketosynthase (KS) and acyl carrier protein (ACP) domains. Shorter peptide sequences, starting with the ACP of a module and extending through the KS and acyl transferase (AT) of the following
 30 module, were extracted to generate a multiple alignment. Positions not aligning to an amino acid from PDB entry 2JU1 (for the ACP) or 2HG4 (for KS and AT and associated linkers) were removed to compress the multiple alignment. Evolutionary couplings were then calculated using the package FreeContact. These couplings take the form of a score matrix with two indices: the first amino acid position in the multiple alignment (I) and the second amino acid position in the multiple alignment (J,
 35 which is always greater than I) and the amino acid at position J. I,J pairs with a score above a specified cutoff and in which I is within the ACP and J within the KS-AT didomain are saved.

To generate a score for a potential single module substitution, the following alignments are retrieved from the original multiple alignment: the ACP for the upstream domain, the ACP and KS-AT didomain for the inserted module, and the KS for the downstream module. These are used to synthesize
 40 two rows compatible with the original multiple alignment: one with the ACP of the upstream module and KS-AT of the inserted module and a second with the ACP of the inserted module and KS-AT of the

downstream module. For each I,J pair in the saved coupling matrix, the amino acids at position I and J in the synthesized alignment are retrieved (aal, aaj). The mutual information for this amino acid pair within the alignment is multiplied by the coupling score to generate a raw score. The raw scores are computed for each I,J pair in the saved coupling matrix and for each of the two synthesized alignments. The sum of the raw scores for the heterologous donor domain is divided by the sum of the raw scores for the homologous native domain to generate a normalized percentage score. Candidate swaps with the same chemistry are ranked by this score. In the case of multiple module swaps, the process is expanded, e.g., if N donor domains are to be swapped in, then one synthetic alignment is generated for the preceding module's ACP domain and the first donor module's KS-AT didomain, another for the first donor modules' ACP domain and the second donor module's KS-AT didomain and so forth, concluding with the final donor domain's ACP and the first module of the recipient synthase downstream of the breakpoint. Scores are computed and normalized in the same manner: the scores for the swapped modules are normalized for the score computed for the native modules. In some embodiments, a heterologous module is compatible if the module is assigned a score of greater than or equal to 0.90 in the inter-module covariation analysis algorithm described herein.

Evolutionary trace analysis to identify modules within functional clades or sub-clades

As an additional test of module compatibility, evolutionary trace analysis may be used to identify modules that belong to the same functional clade or sub-clade (FIGS. 2F-2G). For example, phylogenetic trees with uniform branch lengths were constructed based on multiple sequence alignments of FK-family KSs and ACPs. For every non-terminal node in a tree, a vertical cutoff was applied by which terminal nodes were partitioned into groups based on shared parental nodes at the cutoff. Residues globally conserved across all groups and residues locally conserved within groups, but specific to a given group, were identified as functional residues. Globally conserved residues suggest rules that likely must be observed for all members of the FK-family. Group-specific residues suggest guidelines that may provide predictive power for engineering within the FK class. For each tree, the earliest cutoff at which the number of group-specific residues exceeded the number of globally conserved residues was selected for further analysis. Group-specific residues were concatenated into functional clades and unrooted phylogenetic trees of the clades were constructed. Distances between terminal nodes in the phylogenetic tree were used to create an evolutionary distance score (EDS). The KS and ACP EDSs between a homologous acceptor module and a proposed heterologous donor module were calculated and used to predict engineering compatibility. KS and ACP clade classifications were then used to create network maps of neighboring KSs and ACPs weighted by the frequency a given KS-ACP or ACP-KS pair was observed in FK-family polyketides. Superimposing a proposed module swap onto the network map was used to predict engineering compatibility with upstream ACPs and downstream KSs. In some embodiments, a heterologous module is compatible if the module belongs to the same functional evolutionary clade or sub-clade as one or more adjacent modules in the reference PKS.

Regulation of polyketide synthase expression

The Large ATP-binding regulators of the LuxR family of transcriptional activators (LALs) are known transcriptional regulators of polyketides such as FK506 or rapamycin. The LAL family has been

found to have an active role in the induction of expression of some types of natural product gene clusters, for example PikD for pikromycin production and RapH for rapamycin production. Binding of the LAL or multiple LALs in a complex to specific sites in the promoters of genes within a gene cluster that produces a small molecule (e.g., a polyketide synthase gene cluster) potentiates expression of the gene cluster and hence promotes production of the compound (e.g., a polyketide). In some embodiments, LALs may be used for the regulation of the expression of engineered PKS clusters.

LALs

LALs include three domains, a nucleotide-binding domain, an inducer-binding domain, and a DNA-binding domain. A defining characteristic of the structural class of regulatory proteins that include the LALs is the presence of the AAA+ ATPase domain. Nucleotide hydrolysis is coupled to large conformational changes in the proteins and/or multimerization, and nucleotide binding and hydrolysis represents a “molecular timer” that controls the activity of the LAL (e.g., the duration of the activity of the LAL). The LAL is activated by binding of a small-molecule ligand to the inducer binding site. In most cases the allosteric inducer of the LAL is unknown. In the case of the related protein MalT, the allosteric inducer is maltotriose. Possible inducers for LAL proteins include small molecules found in the environment that trigger compound (e.g., polyketide) biosynthesis. The regulation of the LAL controls production of compound-producing proteins (e.g., polyketide synthases) resulting in activation of compound (e.g., polyketide) production in the presence of external environmental stimuli. Therefore, there are gene clusters that produce small molecules (e.g., PKS gene clusters) which, while present in a strain, do not produce compound either because (i) the LAL has not been activated, (ii) the strain has LAL binding sites that differ from consensus, (iii) the strain lacks an LAL regulator, or (iv) the LAL regulator may be poorly expressed or not expressed under laboratory conditions. Since the DNA binding region of the LALs of the known PKS LALs are highly conserved, the known LALs may be used interchangeably to activate PKS gene clusters other than those which they naturally regulate. In some embodiments, the LAL is a fusion protein.

In some embodiments, an LAL may be modified to include a non-LAL DNA-binding domain, thereby forming a fusion protein including an LAL nucleotide-binding domain and a non-LAL DNA-binding domain. In certain embodiments, the non-LAL DNA-binding domain is capable of binding to a promoter including a protein-binding site positioned such that binding of the DNA-binding domain to the protein-binding site of the promoter promotes expression of a gene of interest (e.g., a gene encoding a compound-producing protein, as described herein). The non-LAL DNA binding domain may include any DNA binding domain known in the art. In some instances, the non-LAL DNA binding domain is a transcription factor DNA binding domain. Examples of non-LAL DNA binding domains include, without limitation, a basic helix-loop-helix (bHLH) domain, leucine zipper domain (e.g., a basic leucine zipper domain), GCC box domain, helix-turn-helix domain, homeodomain, srf-like domain, paired box domain, winged helix domain, zinc finger domain, HMG-box domain, Wor3 domain, OB-fold domain, immunoglobulin domain, B3 domain, TAL effector domain, Cas9 DNA binding domain, GAL4 DNA binding domain, and any other DNA binding domain known in the art. In some instances, the promoter is positioned upstream to the gene of interest, such that the fusion protein may bind to the promoter and induce or inhibit expression of the gene of interest. In certain instances, the promoter is a heterologous

promoter introduced to the nucleic acid (e.g., a chromosome, plasmid, fosmid, or any other nucleic acid construct known in the art) containing the gene of interest. In other instances, the promoter is a pre-existing promoter positioned upstream to the gene of interest. The protein-binding site within the promoter may, for example, be a non-LAL protein-binding site. In certain embodiments, the protein-binding site binds to the non-LAL DNA binding domain, thereby forming a cognate DNA binding domain/protein-binding site pair.

In some embodiments, the LAL is encoded by a nucleic acid having at least 70% (e.g., at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%) sequence identity to any one of SEQ ID Nos: 180-212 or has a sequences with at least 70% (e.g., at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%) sequence identity to any one of SEQ ID Nos: 180-212.

SEQ ID NO: 180

```
ATGCCTGCCGTGGAGTGCTATGAACTGGACGCCCCGCGATGACGAGCTCAGAAAACTGGAGGAGGTTGTGACCGGGCG
GGCCAACGGCCGGGGTGTGGTGGTCACCATCACCGGACCGATCGCCTGCGGCAAGACCGAACTGCTCGACGCAGCCG
15 CCGCGAAGGCCGACGCCATCACGTTACGAGCGGTCTGCTCCGCGGAGGAACAGGCACTCCCGTACGCCCTGATCGGG
CAGCTCATCGACAACCCGGCGCTCGCCTCCACGCGCTGGAGCCGGCCTGCCCGACCCCTCCCGGGCGAGCACCTGTC
GCCGGAGGCCGAGAACCGGTGCGCAGCGACCTACCCGTACCCCTGCTGGCGCTCGCCGCCGAACGGCCGGTGTCTGA
TCGGCATCGACGAGTCACACGGAACGCTTTGTGTCTGCTCCACCTGGCCCGAAGGGTCGGCTCGGCCCGGATCGCC
ATGGTCCTCACCGAGTTGCGCCGGCTCACCCCGGCCACTCACAGTTCCAGGCCGAGCTGCTCAGCCTGGGGCACCA
20 CCGCGAGATCGCGCTGCGCCCGCTCAGCCCCAAGCACACCGCCGAGCTGGTCCGCGCCGGTCTCGGTCCCGACGTG
ACGAGGACGTGCTCACGGGGTTGTACCGGGCGACCGGCGGCAACCTGAACCTCACCCGCGGACTGATCAACGATGTG
CGGGAGGCCTGGGAGACGGGAGGGACGGGCATCAGCGCGGGCCGCGCGTACCGGCTGGCATACCTCGGTTCCCTCTA
CCGCTGCGGCCCGGTCCCGTTGCGGGTCGACGGGTGGCCGCGTGTGGGCCAGAGCGCCAACACCACCTGGTGC
GCTGGATCAGCGGGCTCAACGCGGACGCGGTGGGCGAGGCAACCGAGATCCTCACCGAAGGCGGCCTGCTGCACGAC
25 CTGCGGTTCCCGCACCCGGCGGCCCGTTCCGGTGGTACTCAACGACATGTCCGCCAGGAACGACGCCCGCTGCACCG
GTCCGCTCTGGAAGTGCTGGACGACGTGCCCCGTGGAAGTGGTTCGCGCACCACCAGGTGCGCGCCGGTCTCCTGCACG
GCCCCAAGGCCCGCGAGATATTCGCCAAGGCCGGCCAGGAGCTGCATGTGCGCGCGAGTTGGACACCGCGTCCGAC
TATCTGCAACTGGCCCACCAGGCCTCCGACGACGCCGTACCCGGGATGCGGGCCGAGGCCGTGGCGATCGAGCGCCG
CCGCAACCCGCTGGCCTCGAGCCGGCACCTCGACGAGCTGACCGTCGTCGCCCCGTGCCGGGTGCTCTTCCCCGAGC
30 ACACGGCGCTGATGATCCGCTGGCTGGGCGTGGCGGGCGGTCCGGCGAGGCAGCCGGGTGCTGGCCTCGCAGCGC
CCCCGTGCGGTACCGACCAGGACAGGGCCCATATGCGGGCCGCGAGGTATCGCTCGCGCTGGTCAGCCCCGGCAC
GTCCGGCCCGGACCGGCGGCCGCGTCCGCTCACGCCGATGAGCTCGCGAACCTGCCGAAGGCGGCCCGGCTCTGCG
CGATCGCCGACAATGCCGTATGTGCGCCCTGCGCGGTGCTCCCGAGCTCGCCGCGGCCGAGGCGGAGAACGTCCTG
CAGCACGCCGACTCGGCGGCGGCCGGCACACCAGCCCTCGCCGCGCTGACCGCCTTGCTGTACGCGGAGAACACCGA
35 CACCGCTCAGCTCTGGGCCGACAAGCTGGTCTCCGAGACCGGGGCGTGAACAGGAGGAGGCGGGCTACGCGGGGC
CGCGCGCCGAAGCCGCGTTGCGTTCGCGGCGACCTGGCCGCGGCGGTGAGGCAGGCAGCACCGTTCTGGACCACCGG
CGGCTCTCGACGCTCGGCATCACCGCCGCGCTACCGCTGAGCAGCGCGGTGGCCGCCGCCATCCGGCTGGGCGAGAC
CGAGCGGGCGGAGAAGTGGCTCGCCCAGCCGCTGCCGAGGCCATCCAGGACGGCCTGTTCCGGCTGCACCTGCTCT
CGGCGCGCGGCCAGTACAGCCTCGCCACGGGCCAGCACGAGTCGGCGTACACGGCGTTTCGCACCTGCGGGGAACGT
40 ATGCGGAACCTGGGGCGTTGACGTGCCGGTCTGTCCCTGTGGCGCGTCGACGCCGCCGAGGCGCTGCTGCACGGCCG
CGACCGGGACGAGGGCCGACGGCTCGTCGACGAGCAACTACCCGTGCGATGGGACCCCGTTCCCGCGCCTTGACGC
TGCGGGTGCAGGCGGCGTACAGCCCGCCGGCGAAGCGGGTTCGACCTGCTCGATGAAGCGGCCGACCTGCTGCTCTCC
```

TGCAACGACCAGTACGAGCGGGCACGGGTGCTCGCCGACCTGAGCGAGACGTTTCAGCGCGCTCCGGCACCACAGCCG
GGCGCGGGGACTGCTTCGGCAGGCCCCGGCACCTGGCCGCCCAGCGCGGCGCGATAACCGCTGCTGCGCCGACTCGGGG
CCAAGCCCCGAGGCCCCGGCTGGCTGGAGGAATCCGGCCTGCCGAGCGGATCAAGTCGCTGACCGACGCGGAGCGG
CGGGTGGCGTCGCTGGCCGCCGGCGGACAGACCAACCGCGTGATCGCCGACCAGCTCTTCGTCACGGCCAGCACGGT
5 GGAGCAGCACCTCACGGACGTCTCCACTGGGTCAAGGCCGCCAGCACCTGCCGCCGAACCTCGTCTAG

SEQ ID NO: 181

ATGCCTGCCGTGGAGTGCTATGAACTGGACGCCCCGCGATGACGAGCTCAGAAAACTGGAGGAGGTTGTGACCGGGCG
GGCCAACGGCCGGGGTGTGGTGGTACCATACCGGACCGATCGCCTGCGGCAAGACCGAACTGCTCGACGCAGCCG
10 CCGCGAAGGCCGACGCCATCACGCTGCGAGCGGTCTGCTCCGCGGAGGAACAGGCACTCCCGTACGCCCTGATCGGG
CAGCTCATCGACAACCCGGCGCTCGCCTCCACGCGCTGGAGCCGGCCTGCCCGACCTCCCGGGCGAGCACCTGTC
GCCGGAGGCCGAGAACCGGCTGCGCAGCGACCTACCCGTACCCTGCTGGCGCTCGCCGCCGAACGGCCGGTGCTGA
TCGGCATCGACGAGTCACACGCGAACGCTTTGTGTCTGCTCCACCTGGCCCCGAAGGGTCGGCTCGGCCCGGATCGCC
ATGGTCCTCACCGAGTTGCGCCGGCTCACCCCGGCCACTCACAGTTCCAGGCCGAGCTGCTCAGCCTGGGGCACCA
15 CCGCGAGATCGCGCTGCGCCCGCTCAGCCCCAAGCACACCGCCGAGCTGGTCCGCGCCGGTCTCGGTCCCCGACGTCG
ACGAGGACGTGCTCACGGGGTTGTACCGGGCGACCGGCGGCAACCTGAACCTCACCCGCGGACTGATCAACGATGTG
CGGGAGGCCTGGGAGACGGGAGGGACGGGCATCAGCGCGGGCCGCGCGTACCGGCTGGCATACTCGGTTCCCTCTA
CCGCTGCGGCCCGGTCCCCTTGCGGGTCGCACGGGTGGCCGCCGTGCTGGGCCAGAGCGCCAACACCACCTGGTGC
GCTGGATCAGCGGGCTCAACGCGGACGCGGTGGGCGAGGCAACCGAGATCCTCACCGAAGCGGCCCTGCTGCACGAC
20 CTGCGGTTCCCGCACCCGGCGGCCCGTTCCGGTGGTACTCAACGACATGTCCGCCAGGAACGACGCCCGCTGCACCG
GTCCGCTCTGGAAGTGCTGGACGACGTGCCCGTGAAGTGGTCGCGCACCAACAGGTGCGCGCCGGTCTCCTGCACG
GCCCCAAGGCCCGCCGAGATATTCGCCAAGGCCGGCCAGGAGCTGCATGTGCGCGGCGAGTTGGACACCGCGTCCGAC
TATCTGCAACTGGCCCCACAGGCCTCCGACGACGCGCTACCGGGATGCGGGCCGAGGCCGTGGCGATCGAGCGCCG
CCGCAACCCGCTGGCCTCGAGCCGGCACCTCGACGAGCTGACCGTCGTCGCCCGTGCCGGGCTGCTCTTCCCCGAGC
25 ACACGGCGCTGATGATCCGCTGGCTGGGCGTCGGCGGGCGGTCCGGCGAGGCAGCCGGGCTGCTGGCCTCGCAGCGC
CCCCGTGCGGTACCGACAGGACAGGGCCCATATGCGGGCCCGGAGGTATCGCTCGCGCTGGTCAGCCCCGGCAC
GTCCGGCCCGGACCGGCGGCCGCTCCGCTCACGCCGATGAGCTCGCGAACCTGCCGAAGCGGCCCGGCTCTGCG
CGATCGCCGACAATGCCGTATGTGCGCCCTGCGCGGTCTGCCGAGCTCGCCGCGGCCGAGGCGGAGAACGTCCTG
CAGCACGCCGACTCGGCGGCGGCCGGCACCAACCGCCCTCGCCGCGCTGACCGCCTTGCTGTACGCGGAGAACACCGA
30 CACCGCTCAGCTCTGGGCCGACAAGCTGGTCTCCGAGACCGGGGCGTCGAACGAGGAGGAGGCGGGCTACGCGGGG
CGCGCGCCGAAGCCGCGTTGCGTCGCGGCGACCTGGCCGCGGCGGTGCGAGGCAGGCAGCACCGTTCTGGACCACCGG
CGGCTCTCGACGCTCGGCATCACCGCCGCGCTACCGCTGAGCAGCGCGGTGGCCGCCGCCATCCGGCTGGGCGAGAC
CGAGCGGGCGGAGAAGTGCTCGCCCAGCCGTGCCGAGGCCATCCAGGACGGCCTGTTCCGGCCTGCACCTGCTCT
CGGCGCGCGGCCAGTACAGCCTCGCCACGGGCCAGCACGAGTCGGCGTACACGGCGTTTCGCACCTGCGGGGAACGT
35 ATGCGGAACCTGGGGCGTTGACGTGCCGGGTCTGTCCCTGTGGCGCGTCGACGCCGCCGAGGCGCTGCTGCACGGCCG
CGACCGGGACGAGGGCCGACGGCTCGTCGACGAGCAACTACCCGTGCGATGGGACCCCGTTCCCGCGCCTTGACGC
TGCGGGTGACGGCGGCGTACAGCCCGCCGGCGAAGCGGGTCGACCTGCTCGATGAAGCGGCCGACCTGCTGCTCTCC
TGCAACGACCAGTACGAGCGGGCACGGGTGCTCGCCGACCTGAGCGAGACGTTTCAGCGCGCTCCGGCACCACAGCCG
GGCGCGGGGACTGCTTCGGCAGGCCCCGGCACCTGGCCGCCAGCGCGGCGCGATAACCGCTGCTGCGCCGACTCGGGG
40 CCAAGCCCCGAGGCCCCGGCTGGCTGGAGGAATCCGGCCTGCCGAGCGGATCAAGTCGCTGACCGACGCGGAGCGG
CGGGTGGCGTCGCTGGCCGCCGGCGGACAGACCAACCGCGTGATCGCCGACCAGCTCTTCGTCACGGCCAGCACGGT
GGAGCAGCACCTCACGGACGTCTCCACTGGGTCAAGGCCGCCAGCACCTGCCGCCGAACCTCGTCTAG

SEQ ID NO: 182

GTGGTTCTGAAGTGCGAGCAGCCCCGACGAACTGATCGCCCGCGATGACGAGCTGAGCCGCCTCCAACGGGGCACT
CACCAGGGCGGGGAGCGGAAGGGGCGGGCTCGTCGCCATCACCAGGGCCCATCGCCAGCGGAAAGACGGCGCTGCTCG
5 ACGCCGGAGCGGCCAAGTCCGGCTTCGTGCGACTCCGTGCGGTGTGCTCCTGGGAAGAGCGCACTCTGCCGTACGGG
ATGCTGGGCCAGCTCTTCGACCATCCCGAACTGGCCGCCAGGCGCCGACCTTGCCCACTTCACGGCTTCGTGCGA
GAGCCCTCAGGCCGGTACCGACAACCGCCTGCGGGCCGAGTTACCCGCAACCCTGCTGGCGCTCGCCGCGGACTGGC
CCGTCTGATCGGCATCGACGACGTGCACCACGCCGACGCGGAATCACTGCGCTGTCTGCTCCACCTCGCCCGCCGC
ATCGGCCCGGGCCGCATCGCGGTCTGACTGACCGAGCTGCGCAGACCGACGCCCGCGACTCCCCTTCCAGGCGGA
10 ACTGCTGAGCCTGCGCTCCTACCAGGAGATCGCGCTCAGACCGCTCACCAGGGCGCAGACCGGCGAACTCGTACGTC
GGCACCTCGGCGCGGAGACCCACGAGGACGTCTCCGCCGATACGTTCCGGGCGACCGGCGGGAACCTGCTCCTCGGG
CACGGTTTGATCAATGACATCCGGGAGGCGCGGACAGCGGGACGGCCGGGGGTCTGTCGGGGGCGGGCGTACCGGCT
CGCGTACCTCAGCTCGCTCTACCGCTGCGGCCCGAGCGCGCTGCGTGTGCCCCGGGCGTCCGCCGTGCTCGGCGCGA
GCGCCGAAGCCGTGCTCGTCCAGCGGATGACCGGACTGAACAAGGACGCGGTGGAACAGGTCTATGAGCAGCTGAAC
15 GAGGGACGGCTGCTGCAGGGCGAGCGGTTTCCGCACCCGGCGGCCCGCTCCATCGTCTTGACGACCTGTGCGGCCCT
GGAACGCAGAAACCTGCACGAGTCGGCGCTGGAGCTGCTGCGGGACACGGCGTGGCCGGCAACGTGCTCGCCCGCC
ACCAGATCGGCGCCGGCCGGGTGCACGGCGAGGAGGCCGTGAGCTGTTACCCGGGCGCGACGGGAGCACCACTG
CGCGGTGAACCTGGACGACGCGGCCGGATACCTGGAACCTGCCCCACCGTGCTCCGACGACCCCGTCACGCGCGCCGC
ACTACGCGTCGGCGCCCGCGGATCGAGCGCCTCTGCAATCCGGTACGGGCAGGCCGGCATCTGCCCGAGCTGCTCA
20 CCGCGTCGCGCGCGGGACTGCTCTCCAGCGAGCACGCCGTGTGCTCGCTCGCCGACTGGCTGGCGATGGGCGGGCGCCCCG
GGCGAGGCGGGCGAGGTCTCGCGACGACGCTCCCGCGGCCGACAGCGAGCAGCACCGCGCACTCCTGCGCAGCGG
CGAGTTGTCCCTCGCGCTGGTCCACCCCGGCGCGTGGGATCCGTTGCGCCGACCGATCGGTTCCCGCGGGGCGGGC
TCGGCTCGCTTCCCGGACCCGCCCGGACCCGCGCGGTGCGCGACCAAGCCGTCATCGCGGCGCTGCGTGGACGTCTC
GACCGGGCGGACGCCAACGCGGAGAGCGTTCTCCAGCACACCGACGCCACGGCGGACCGGACCACGGCCATCATGGC
25 GTTGCTGGCCCTGCTCTACGCGGAGAACACCGATGCTGTCCAGTTCTGGGTGACAAACTGGCCGGTGACGAGGGCA
CCAGGACACCGGCCGACGAGGCGGTCCACGCGGGGTTCAACGCCGAGATCGCGCTGCGCCGCGGCGACTTGATGAGA
GCCGTGAGTACGGCGAGGACGCGCTCGGCCACCGGCACCTGCCCACCTGGGGAATGGCCGCCGCTCTGCCGCTGAG
CAGCACCGTGTTGCCGCGATCCGGCTCGGCGACCTCGACAGGGCCGAGCGGTGGCTCGCCGAGCCGCTGCCGCGAGC
AGACGCCGGAGAGCCTCTTCGGGCTGCACCTGCTCTGGGCCCCGCGGGCAGCACCACTCGCGACCGGGCGGCACGGG
30 GCGGCGTACACGGCGTTACAGGAATGCGGCGAGCGGATGCGGCGGTGGGCCGTGACGTGCCGGGCTGGCCCTGTG
GCGGGTTCGACGCCGCCGAATCGCTGCTGCTGCTCGGCCGTGACCGTGCCGAAGGACTGCGGCTCGTCTCCGAGCAGC
TGTCCCGGCCGATGCGCCCTCGCGCGCGCGTGCAGACGTTACGGGTACAGGCGGCCTACAGTCCGCCGCCCAACGG
ATCGACCTGCTCGAAGAGGCCGCCGACCTGCTGGTCACTGCAACGACAGTACGAACCTGGCAAACGTACTCAGCGA
CTTGGCAGAGGCCTCCAGCATGGTCCGGCAGCACAGCAGGGCGCGGGGTCTGCTCCGCCGGGCACGGCACCTCGCCA
35 CCCAGTGGCGCGCCGTGCCGCTCCTGCGGCGGCTCGGCGCGGAACCCCTCGGACATCGGCGGAGCCTGGGACGCGACG
CTGGGACAGCGGATCGCGTCACTGACGGAGTCGGAGCGGCGGGTGGCCGCGCTCGCCGCGGTGGGGCTACGAACAG
GGAGATCGCCGAGCAGCTGTTTCGTACAGGCCAGCACGGTGGAACAGCACCTCACGAACGTGTTCCGCAAACCTGGCGG
TGAAGGGCCGCCAGCAGCTTCCGAAGGAACCTGGCCGACGTGCGCGAGCCGGCGGACCGCGACCGCCGGTGCGGGTAG

40 SEQ ID NO: 183

ATGGTTCTGAAGTGCGAGCAGCCCCGACGAACTGATCGCCCGCGATGACGAGCTGAGCCGCCTCCAACGGGGCACT
CACCAGGGCGGGGAGCGGAAGGGGCGGGCTCGTCGCCATCACCAGGGCCCATCGCCAGCGGAAAGACGGCGCTGCTCG

ACGCCGGAGCGGCCAAGTCCGGCTTCGTGCGACTCCGTGCGGTGTGCTCCTGGGAAGAGCGCACTCTGCCGTACGGG
ATGCTGGGCCAGCTCTTCGACCATCCCGAACTGGCCGCCAGGCGCCGACCTTGCCCACTTCACGGCTTCGTGCGA
GAGCCCTCAGGCCGGTACCGACAACCGCCTGCGGGCCGAGTTACCCGACCCCTGCTGGCGCTCGCCGCGGACTGGC
CCGTCTGATCGGCATCGACGACGTGCACCACGCCGACGCGGAATCACTGCGCTGTCTGCTCCACCTCGCCCCGCCG
5 ATCGGCCCCGGCCGCATCGCGGTCTGACTGACCGAGCTGCGCAGACCGACGCCCGCCGACTCCCGCTTCAGGCGGA
ACTGCTGAGCCTGCGCTCCTACCAGGAGATCGCGCTCAGACCGCTCACCAGGGCGCAGACCGGCGAACTCGTACGTC
GGCACCTCGGCGCGGAGACCCACGAGGACGTCTCCGCCGATACGTTCCGGGCGACCGGCGGGAACCTGCTCCTCGGG
CACGGTTTGATCAATGACATCCGGGAGGCGCGGACAGCGGGACGGCCGGGGTTCGTGCGGGGCGGGCGTACCGGCT
CGCGTACCTCAGCTCGCTCTACCGCTGCGGCCGAGCGCGCTGCGTGTGCGCCGGGCGTCCGCCGTGCTCGGCGCGA
10 GCGCCGAAGCCGTGCTCGTCCAGCGGATGACCGGACTGAACAAGGACGCGGTGGAACAGGTCTATGAGCAGCTGAAC
GAGGGACGGCTGCTGCGAGGGCGAGCGGTTTTCCGCACCCGGCGGCCCGCTCCATCGTCTTGACGACCTGTGCGCCCT
GGAACGCAGAAACCTGCACGAGTCGGCGCTGGAGCTGCTGCGGGACCACGGCGTGCCCGGCAACGTGCTCGCCCGCC
ACCAGATCGGCGCCGGCCGGGTGCACGGCGAGGAGGCCGTGAGCTGTTACCCGGGGCCGCACGGGAGCACCACTG
CGCGGTGAACTGGACGACGCGGCCGGATACCTGGAACCTGCCCCACCGTGCTCCGACGACCCCGTCACGCGCGCCG
15 ACTACGCTCGGCGCCGCCGCGATCGAGCGCCTCTGCAATCCGGTACGGGCAGGCCGGCATCTGCCCGAGCTGCTCA
CCGCGTCGCGCGCGGGACTGCTCTCCAGCGAGCACGCCGTGTGCTCGCTCGCCGACTGGCTGGCGATGGGCGGGCGCCG
GGCGAGGCGGCCGAGGTCTTCGCGACGCGAGCGTCCCGCGGCCGACAGCGAGCAGCACCGCGCACTCCTGCGCAGCGG
CGAGTTGTCCCTCGCGCTGGTCCACCCCGGCGCGTGGGATCCGTTGCGCCGGACCGATCGGTTGCGCGGGGCGGGC
TCGGCTCGCTTCCCGGACCCGCCCGGCACCGCGCGGTGCGCGACCAAGCCGTATCGCGGCGCTGCGTGGACGTCTC
20 GACCGGGCGGACGCCAACGCGGAGAGCGTTCTCCAGCACACCGACGCCACGGCGGACCGGACCACGGCCATCATGGC
GTTGCTGGCCCTGCTCTACGCGGAGAACACCGATGCTGTCCAGTTCTGGGTGACAAACTGGCCGGTGACGAGGGCA
CCAGGACACCGGCCGACGAGGCGGTCCACGCGGGTTCAACGCCGAGATCGCGCTGCGCCGCGGCGACTTGATGAGA
GCCGTGAGTACGGCGAGGCGAGCGCTCGGCCACCGGCACCTGCCCACCTGGGGAATGGCCGCCGCTCTGCCGCTGAG
CAGCACCGTGTTGCCGCGATCCGGCTCGGCGACCTCGACAGGGCCGAGCGGTGGCTCGCCGAGCCGCTGCCGCAGC
25 AGACGCCGAGAGCCTCTTCGGGTGCACCTGCTCTGGGCCCCGCGGGCAGCACCACTCGCGACCGGGCGGCACGGG
GCGGCGTACACGGCGTTACAGGAATGCGGCGAGCGGATGCGGCGGTGGGCCGTGACGTGCCGGGCTGGCCCTGTG
GCGGGTCGACGCGCCGAATCGCTGCTGCTGCTCGGCCGTGACCGTGCCGAAGGACTGCGGCTCGTCTCCGAGCAGC
TGTCCCGGCCGATGCGCCCTCGCGCGCGCTGCAGACGCTGCGGGTACAGGCGGCCTACAGTCCGCCGCCCAACGG
ATCGACCTGCTCGAAGAGGCCGCCGACCTGCTGGTACCTGCAACGACCAGTACGAACTGGCAAACGTACTCAGCGA
30 CTTGGCAGAGGCCTCCAGCATGGTCCGGCAGCACAGCAGGGCGCGGGGTCTGCTCCGCCGGGCACGGCACCTCGCCA
CCCAGTGCGGCGCCGTCGCGCTCCTGCGGCGGCTCGGCGCGGAACCTCGGACATCGGCGGAGCCTGGGACGCGACG
CTGGGACAGCGGATCGCGTCACTGACGGAGTCGGAGCGGCGGGTGGCCGCGCTCGCCGCGGTGGGGCTACGAACAG
GGAGATCGCCGAGCAGCTGTTCTGTCACGGCCAGCACGGTGGAACAGCACCTCACGAACGTGTTCCGCAAACTGGCGG
TGAAGGGCCGCCAGCAGCTTCCGAAGGAACTGGCCGACGTGCGCGAGCCGGCGGACCGCGACCGCCGGTGCGGGTAG
35

SEQ ID NO: 184

GTGATAGCGCGCTTATCTCCCCAGACCTGATCGCCCGGATGACGAGTTCGGTTCCCTCCACCGGGCGCTCACCCG
AGCGGGGGGCGGGCGGGGCGTCTGTCGCCCGCTACCGGGCCGATCGCCTGCGGCAAGACCGAACTCCTCGACGCCG
CCGCGGCCAAGGCCGGCTTCGTACCCCTTCGCGCGGTGTGCTCCATGGAGGAGCGGGCCCTGCCGTACGGCATGCTC
40 GGCCAGCTCCTCGACCAGCCGAGCTGGCCGCCCGGACACCGGAGCTGGTCCGGCTGACGGCATCGTGCGAAAACCT
GCCGGCCGACGTGACAACCGCCTGGGGACCGAACTACCCGCACGGTGTGACGCTCGCCGCGGAGCGGGCCGTAC
TGATCGGCATCGACGACGTGCACCACGCCGACGCGCGTTCGTGCGCTGCCTGCTCCACCTCGCGCGCCGCATCAGC

CGGGCCCGTGTGCGCATCGTGCTGACCGAGCTGCTCCGGCCGACGCCCCGCCCCTCCCAATTCCGGGCGGCACTGCT
GAGTCTGCGCCACTACCAGGAGATCGCGCTGCGCCCGCTCACCGAGGCGCAGACCACCGAACTCGTGCGCCGGCACC
TCGGCCAGGACGCGCACGACGACGTGGTGGCCAGGCGTTCCGGGCGACCGGCGGCAACCTGCTCCTCGGCCACGGC
CTGATCGACGACATCCGGGAGGCACGGACACGGACCTCAGGGTGCTTGAAGTGGTCGCGGGGCGGGCGTACCGGCT
5 CGCCTACCTCGGGTCGCTCTATCGTTGCGGCCCCGCGCGCTGAGCGTCGCCCCGAGCTTCCGCCGTGCTCGGCGAGA
GTGTGCAACTCACCTCGTCCAGCGGATGACCGGCCCTCGACACCGAGGCGGTGAGCAGGCCCCAGAACAGCTGGTC
GAGGGGCGGCTGCTGCGGGAAGGGCGGTTCCCGCACCCCGCGGCCCGCTCCGTCTACTCGACGACCTCTCCGCCGC
CGAGCGGCGTGGCCTGCACGAGCTGGCGCTGGAAGTGTGCGGGACCGCGCGTGGCCAGCAAGGTGCTCGCCCCGCC
ACCAGATGGGTACCGGCCGGGTGCACGGCGCCGAGGTGCGCGGGCTGTTACCGACGCCGCGCGAGCACCACCTG
10 CGCGGCGAGCTCGACGAGGCGGTACCTACCTGGAGTTCGCTACCGGGCCTCCGACGACCCCGCCGTCCACGCCGC
ACTGCGCGTCGACACCGCCGCCATCGAGCGGCTCTGCGATCCCGCCAGATCCGGCCGGCATGTGCCCGAGCTGCTCA
CCGCGTCGCGGGAACGGCTCCTCTCCAGCGAGCACGCCGTGTGCTCGCTCGCCTGCTGGCTGGCGATGGACGGGCGGCCG
GGCGAGGCCCGCGAGGTCTGGCGGCCAGCGCTCCGCCGCCCGAGCGAGCAGGGCCGGGCGACCTGCGCGTTCGC
GGACCTGTCCCTCGCGCTGATCTATCCCGCGCGGCCGATCCGCCGCGTCCGGCCGATCCGCCGGCCGAGGACGAGG
15 TCGCCTCGTTTTCCGGAGCCGTCCGGCACCGCGCCGTGCGCGACAAGGCCCTGAGCAACGCGCTGCGCGGCTGGTCC
GAACAGGCCGAGGCCAAAGCCGAGTACGTGCTCCAGCACTCCCGGGTCACGACGGACCGGACCACGACCATGATGGC
GTTGCTGGCCCTGCTCTACGCCGAGGACACCGATGCCGTCCAGTCTTGGGTGACAAGCTGGCCGGTGACGACAACA
TGCGGACCCCGGCCGACGAGGCGGTCCACGCGGGGTTCCGCGCCGAGGCCGCGCTGCGCCGCGGCGACCTGACCGCC
GCCGTGCAATGCGGCGAGGCCGCGCTCGCCCCCGGGTCTGCCCCCTCTGGGGGATGGCCGCCGCATTGCCGCTGAG
20 CAGCACCGTGGCCGCCGCGATCCGACTGGGCGACCTGGACCGGGCGGAGCGGTGGCTCGCCGAGCCGTTGCCGGAGG
AGACCTCCGACAGCCTCTTCGGAAGTGCACATGGTCTGGGCCCGTGGGCAACACCATCTCGCGGCCGGGCGGTACCGG
GCGGCGTACAACGCGTTCCGGGACTGCGGGGAGCGGATGCGACGCTGGTCCGTGACGTGCCGGGCCTGGCCCTGTG
GCGGGTCGACGCCGCCGAAGCGCTTCTGCTGCTCGGCCGCGGCCGTGACGAGGGGCTGAGGCTCATCTCCGAGCAGC
TGTCCCGGCCGATGGGGTCCCGGGCGCGGGTGATGACGCTGCGGGTGCAGGCGGCCCTACAGTCCGCCGGCCAAGCGG
25 ATCGAACTGCTCGACGAGGCCGCCGATCTGCTCATCATGTGCCGCGACCAAGTACGAGCTGGCCCCGCTCCTCGCCGA
CATGGGCGAAGCGTGCGGCATGCTCCGGCGGCACAGCCGTGCGCGGGGACTGTTCCGCCGCGCACGGCACCTCGCGA
CCCAGTGGGAGCCGTGCCGCTCCTCCGGCGGCTCGGTGGGGAGTCTCGGACGCGGACGGCACCCAGGACGTGACG
CCGGCGCAGCGGATCACATCGCTGACCGAGGCGGAGCGGCGGGTGGCGTCGCACGCCGCGGTGGGGCGACCAACAA
GGAGATCGCCAGCCAGCTGTTTCGTACCTCCAGCACGGTGGAACAGCACCTCACCAACGTGTTCCGCAAGCTGGGGG
30 TGAAGGGCCGTGAGCAACTGCCCAAGGAAGTGTCCGACGCCGGCTGA

SEQ ID NO: 185

ATGATAGCGCGCCTGTCTCCCCAGACCTGATCGCCGCGATGACGAGTTCGGTTCCTCCACCGGGCGCTCACCCG
AGCGGGGGGCGGGCGGGGCGTCTGCGCCCGCTACCGGGCCGATCGCCTGCGGCAAGACCGAACTCCTCGACGCCG
35 CCGCGGCCAAGGCCGGCTTCGTACCTTTCGCGCGGTGTGCTCCATGGAGGAGCGGGCCCTGCCGTACGGCATGCTC
GGCCAGCTCCTCGACAGCCCGAGCTGGCCGCCCGGACACCGGAGCTGGTCCGGCTGACGGCATCGTGCGAAAACCT
GCCGGCCGACGTGACAACCGCCTGGGGACCGAACTCACCCGCACGGTGTGACGCTCGCCGCGGAGCGGCCGTAC
TGATCGGCATCGACGACGTGCACCACGCCGACGCGCGTCTGCTGCGCTGCCTGCTCCACCTCGCGCGCCGCATCAGC
CGGGCCCGTGTGCGCATCGTGCTGACCGAGCTGCTCCGGCCGACGCCCGCCCCTCCCAATTCCGGGCGGCACTGCT
40 GAGTCTGCGCCACTACCAGGAGATCGCGCTGCGCCCGCTCACCGAGGCGCAGACCACCGAACTCGTGCGCCGGCACC
TCGGCCAGGACGCGCACGACGACGTGGTGGCCAGGCGTTCCGGGCGACCGGCGGCAACCTGCTCCTCGGCCACGGC
CTGATCGACGACATCCGGGAGGCACGGACACGGACCTCAGGGTGCTTGAAGTGGTCGCGGGGCGGGCGTACCGGCT

CGCCTACCTCGGGTCGCTCTATCGTTGCGGCCCGGCCGCGCTGAGCGTCGCCCCGAGCTTCCGCCGTGCTCGGCGAGA
GTGTGCAACTCACCTCGTCCAGCGGATGACCGGCCCTCGACACCGAGGCGGTGAGCAGGCCCCACGAACAGCTGGTC
GAGGGGCGGCTGCTGCGGGAAGGGCGGTTCCCGCACCCCGCGGCCCGCTCCGTCTGTAATCGACGACCTCTCCGCCGC
CGAGCGGCGTGGCCTGCACGAGCTGGCGCTGGAAGTCTGTCGGGACCGCGGCGTGGCCAGCAAGGTGCTCGCCCCGC
5 ACCAGATGGGTACCGGCCGGGTGCACGGCGCCGAGGTGCGCGGGCTGTTACCGACGCCGCGCGGAGCACCACCTG
CGCGGCGAGCTCGACGAGGCGGTACCTACCTGGAGTTCGCTACCGGGCCCTCCGACGACCCCGCCGTCCACGCCGC
ACTGCGCGTCGACACCGCCGCCATCGAGCGGCTCTGCGATCCCGCCAGATCCGGCCGGCATGTGCCCAGCTGCTCA
CCGCGTCGCGGGAACGGCTCCTCTCCAGCGAGCACGCCGTGTCGCTCGCCTGCTGGCTGGCGATGGACGGGCGGCCG
GGCGAGGCCCGGAGGTCTGGCGGCCAGCGCTCCGCCGCCCGAGCGAGCAGGGCCGGGCGACCTGCGCGTCCG
10 GGACCTGTCCCTCGCGCTGATCTATCCCGGCGCGGCCGATCCGCCGCGTCCGGCCGATCCGCCGGCCGAGGACGAGG
TCGCCTCGTTTTCCGGAGCCGTCCGGCACCGCGCCGTGCGCGACAAGGCCCTGAGCAACGCGCTGCGCGGCTGGTCC
GAACAGGCCGAGGCCAAAGCCGAGTACGTGCTCCAGCACTCCCGGGTCACGACGGACCGACCACGACCATGATGGC
GTTGCTGGCCCTGCTCTACGCCGAGGACACCGATGCCGTCCAGTCTGGGTGCGACAAGCTGGCCGGTGACGACAACA
TGCGGACCCCGGCCGACGAGGCGGTCCACGCGGGTTCCGCGCCGAGGCCGCGCTGCGCCGCGGCGACCTGACCGCC
15 GCCGTGCAATGCGGCGAGGCCGCGCTCGCCCCCGGGTCGTGCCCTCCTGGGGGATGGCCGCCGCATTGCCGCTGAG
CAGCACCGTGGCCGCCGCGATCCGACTGGGCGACCTGGACCGGGCGGAGCGGTGGCTCGCCGAGCCGTTGCCGGAGG
AGACCTCCGACAGCCTCTTCGACTGCACATGGTCTGGGCCCCGTGGGCAACACCATCTCGCGGCCGGGCGGTACCGG
GCGGCGTACAACGCGTTCCGGGACTGCGGGGAGCGGATGCGACGCTGGTCCGTGACGTGCCGGGCCCTGGCCCTGTG
GCGGGTCGACGCCGCCGAAGCGCTTCTGCTGCTCGGCCGCGGCCGTGACGAGGGGCTGAGGCTCATCTCCGAGCAGC
20 TGTCCCGGCCGATGGGGTCCCGGGCGCGGTGATGACGCTGCGGGTGCAGGCGGCCCTACAGTCCGCCGGCCAAGCGG
ATCGAACTGCTCGACGAGGCCGCCGATCTGCTCATCATGTGCCGCGACCAGTACGAGCTGGCCCGCGTCTCCGCCGA
CATGGGCGAAGCGTGCGGCATGCTCCGGCGGCACAGCCGTGCGCGGGGACTGTTCCGCCGCGCACGGCACCTCGCGA
CCCAGTGGGAGCCGTGCCGCTCCTCCGGCGGCTCGGTGGGGAGTCTCGGACGCGGACGGCACCCAGGACGTGACG
CCGGCGCAGCGGATCACATCGCTGACCGAGGCGGAGCGGCCGGTGGCGTGCACGCCGCGGTGGGGCGACCAACAA
25 GGAGATCGCCAGCCAGCTGTTTCGTACCTCCAGCACGGTGGAAACAGCACCTCACCAACGTGTTCCGCAAGCTGGGGG
TGAAGGGCCGTGAGCAACTGCCCAAGGAAGTGTCCGACGCCGGCTGA

SEQ ID NO: 186

GTGGAGTTTTACGACCTGGTCGCCCCGCGATGACGAGCTCAGAAGGTTGGACCAGGCCCTCGGCCGCGCCGCCGGCGG
30 ACGGGGTGTCGTGGTCACCGTCACCGGACCGGTGCGGCTGCGGCAAGACCGAACTGCTGGACGCGGCCGCGGCCGAGG
AGGAATTTCATCACGTTGCGTGCGGTCTGCTCGGCCGAGGAGCGGGCCCTGCCGTACGCCGTGATCGGCCAACTCCTC
GACCATCCCGTACTCTCCGCACGCGCGCCCGACCTGGCCTGCGTGACGGCTCCGGGCCGGACGCTGCCGGCCGACAC
CGAGAACC GCCTGCGCCGCGACCTCACCCGGGCCCTGCTGGCCCTGGCCTCCGAACGACCGGTTCTGATCTGCATCG
ACGACGTGCACCAGGCCGACACCGCCTCGCTGAACTGCCTGCTGCACCTGGCCCGGCGGGTCGCCTCGGCCCGGATC
35 GCCATGATCCTACCGAGTTGCGCCGGCTCACCCGGCTCACTCCCGGTTGAGGCGGAACTGCTCAGCCTGCGGCA
CCGCCACGAGATCGCGCTGCGTCCCTCGGCCCGGCCGACACCGCCGAAGTGGCCCGCGCCCGGCTCGGGCGCCGGCG
TCACCGCCGACGAGCTGGCCAGGTCCACGAGGCCACAGCGGGAACCCCAACCTGGTTCGGAGGCGCTGGTCAACGAC
GTGCGAGAGGCTGGGCGGCCGGTGGCACGGGCATTGCGGCGGGGCGGGCGTACCGGTGGCGTACCTCAGTCCGT
GTACCGCTGTGGTCCGGTCCCGTTGCGGATCGCCCAGGCGGCGGGTGTGGGTCCCAGCGCCACCGTCACGCTGG
40 TCGCGCCGATCAGCGGGCTCGACGCCGAGACGGTGGACGAGGCGACCGCGATCCTCACCGAGGGCGGCCCTGCTCCGG
GACCACCGGTTCCCGCATCCGGCGGCCCGCTCGGTGCTACTCGACGACATGTCCGCGCAGGAACGCCGCCGCTGCA
CCGGTCCACGCTGGACGTGCTGGACGGCGTACCGTTCGACGTGCTCGCGCACCAACAGGCCGGCGCCGGTCTGCTGC

ACGGCCCGCAGGCGGCCGAGATGTTTCGCCCCGGGCCAGCCAGGAGCTGCGGGTACGCGGCGAGCTGGACGCCGCGACC
GAGTACCTGCAACTGGCCTACCGGGCCTCCGACGACGCCGGCGCCCCGGGCCGCCCTGCAGGTGGAGACCGTGGCCGG
CGAGCGCCGCCGCAACCCGCTGGCCGCCAGCCGGCACCTGGACGAGCTGGCCGCCGCCCGGGGCCGGCCTGCTGT
CGGCCGAGCACGCCGCCCTGGTCTGTGCACTGGCTGGCCGACGCCGGACGACCCGGCGAGGCCGCCGAGGTGCTGGCG
5 CTGCAGCGGGCGCTGGCCGTACCGACCACGACCGGGCCCCGCTGCGGGCGGCCGAGGTGTCGCTCGCGCTGTTCCA
CCCCGGCGTCCCCGGTTTCGGACCCGCGGGCCCCCTCGCGCCGGAGGAGCTCGCGAGCCTGTCCCTGTGCGGCCGGCACG
GTGTGACCGCCGACAACGCGGTGCTGGCGGGCGCTGCGCGGCCGTCCCGAGTCGGCCGCCGCCGAGGCGGAGAACGTG
CTGCGCAACGCCGACGCCGCCCGCTCCGGCCCCGACCGCCCTGGCCGCGCTGACGGCCCTGCTCTACGCCGAGAACAC
CGACGCCGCCAGCTCTGGGCGGACAAGCTGGCCGCGGGCATCGGGGCGGGGGAGGGGGAGGCCGGCTACGCGGGGC
10 CGCGGACCGTGGCCGCCCTGCGTTCGCGGGCGACCTGACCACCGCGGTCCAGGCGGCCGGCGCGGTCTTGACCGCGGC
CGGCCGTGCTCGCTCGGCATCACCGCCGTGTTGCCGTTGAGCGGCGCGGTGCGCCGCCGCGATCCGGCTGGGCGAGCT
CGAGCGGGCCGAGAAGTGGCTGGCCGAGCCGCTGCCCCAAGCCGTCCACGACAGCCTGTTCCGGCTGCACCTGCTGA
TGGCGCGGGGCCGCTACAGCCTCGCGGTGGGCCGGCACGAGGCGGCGTACGCCCGCTTCCGGGACTGCGGTGAACGG
ATGCGCCGGTGGGACGTGACGTGCCCGGGCTGGCCCTGTGGCGGGTGGACGCGGCCGAGGCGCTGCTGCCCGGGCGA
15 TGACCGGGCGGAGGGCCGGCGGGCTGATCGACGAGCAGCTACCCGGCCGATGGGGCCCCGGTACGAGCCCTGACCC
TGCGGGTACGAGCGGCCTACGCCCCGCCGGCGAAACGGATCGACCTGCTCGACGAAGCGGCCGACCTGCTGCTCTCC
AGCAACGACCAGTACGAGCGGGCACGGGTGCTGGCCGACCTGAGCGAGGCGTTCAGCGCGCTCCGGCAGAACGGCCG
GGCGCGCGGCATCCTGCGGCAGGCCCGGCACCTGGCCGCCAGTGCGGGGCGGTCCCCCTGCTGCGCCGGCTGGGCG
TCAAGGCCGGCCGGTCCGGTTCGGCTCGGCCGGCCGCCGAGGGAATCCGCTCCCTGACCGAGGCCGAGCGCCGGGTG
20 GCCACGCTGGCCGCCGCCGGGACACCAACCGGGAGATCGCCGACCAGCTCTTCGTCACCGCCAGCACGGTCGAGCA
GCACCTACCAACGTGTTCCGCAAGCTCGGCGTGAAGGGCCGCCAGCAATTGCCGGCCGAGCTGGCCGACCTGCGGC
CGCCGGGCTGA

SEQ ID NO: 187

ATGGAGTTTTACGACCTGGTTCGCCCCGCGATGACGAGCTCAGAAGGTTGGACCAGGCCCTCGGCCGCGCCGCCGGCGG
ACGGGGTGTGCTGGTACCGTACCGGACCGGTTCGGCTGCGGCAAGACCGAACTGCTGGACGCGGCCGCCGGCCGAGG
AGGAATTCATCACGTTGCGTTCGGTCTGCTCGGCCGAGGAGCGGGCCCTGCCGTACGCCGTGATCGGCCAACTCCTC
GACCATCCCGTACTCTCCGCACGCGCGCCCCGACCTGGCCTGCGTGACGGCTCCGGGCCGGACGCTGCCGGCCGACAC
CGAGAACC GCCTGCGCCGCGACCTACCCGGGCCCTGCTGGCCCTGGCCTCCGAACGACCGGTTCTGATCTGCATCG
30 ACGACGTGCACCAGGCCGACACCGCCTCGCTGAACTGCCTGCTGCACCTGGCCCCGGCGGGTTCGCTCGGCCCGGATC
GCCATGATCCTACCGAGTTGCGCCGGCTCACCCCGGCTCACTCCCGGTTTCGAGGCGGAACTGCTCAGCCTGCGGCA
CCGCCACGAGATCGCGCTGCGTCCCTCGGCCCGGCCGACACCGCCGAACTGGCCCCGCGCCCGGCTCGGCGCCGGCG
TCACCGCCGACGAGCTGGCCAGGTCCACGAGGCCACGAGCGGGAACCCCAACCTGGTTCGAGGCTTGGTCAACGAC
GTGCGAGAGGCTGGGCGGCCGGTGGCACGGGCATTGCGGCGGGGCGGGCGTACCGGCTGGCGTACCTCAGCTCCGT
35 GTACCGCTGTGGTCCGGTCCCGTTGCGGATCGCCAGGCGGGCGGGTGTGGGTCCCAGCGCCACCGTCACGCTGG
TGCGCCGGATCAGCGGGCTCGACGCCGAGACGGTGGACGAGGCGACCGCGATCCTACCGAGGGCGGCCTGCTCCGG
GACCACCGGTTCCCGCATCCGGCGGCCCGCTCGGTTCGTACTCGACGACATGTCCGCGCAGGAACGCCGCCCGCTGCA
CCGGTCCACGCTGGACGTGCTGGACGGCGTACCGTTCGACGTGCTCGCGCACCAACAGGCCGGCGCCGGTCTGCTGC
ACGGCCCGCAGGCGGCCGAGATGTTTCGCCCCGGGCCAGCCAGGAGCTGCGGGTACGCGGCGAGCTGGACGCCGCGACC
40 GAGTACCTGCAACTGGCCTACCGGGCCTCCGACGACGCCGGCGCCCCGGGCCGCCCTGCAGGTGGAGACCGTGGCCGG
CGAGCGCCGCCGCAACCCGCTGGCCGCCAGCCGGCACCTGGACGAGCTGGCCGCCGCCCGGGGCCGGCCTGCTGT
CGGCCGAGCACGCCGCCCTGGTCTGTGCACTGGCTGGCCGACGCCGGACGACCCGGCGAGGCCGCCGAGGTGCTGGCG

CTGCAGCGGGCGCTGGCCGTCACCGACCACGACCGGGCCCCGCTGCGGGCGGGCGAGGTGTCGCTCGCGCTGTTCCA
CCCCGGCGTCCCCGGTTTCGGACCCGCGGGCCCCCTCGCGCCGGAGGAGCTCGCGAGCCTGTCCCTGTCGGCCCCGGCACG
GTGTGACCGCCGACAACGCGGTGCTGGCGGGCGCTGCGCGGCCGTCCCGAGTCGGCCGCCGCCGAGGCGGAGAACGTG
CTGCGCAACGCCGACGCCGCCGCGTCCGGCCCCGACCGCCCTGGCCGCGCTGACGGCCCTGCTCTACGCCGAGAACAC
5 CGACGCCGCCAGCTCTGGGCGGACAAGCTGGCCGCGGGCATCGGGGCGGGGGAGGGGGAGGCCGGCTACGCGGGGC
CGCGGACCGTGGCCGCCCTGCGTTCGCGGGCGACCTGACCACCGCGGTCCAGGCGGGCGGCGCGGTCTTGACCGCGGC
CGGCCGTGCTCGCTCGGCATCACCGCCGTGTTGCCGTTGAGCGGCGCGGTGCGCCGCCGCGATCCGGCTGGGCGAGCT
CGAGCGGGCCGAGAAGTGGCTGGCCGAGCCGCTGCCCCAAGCCGTCCACGACAGCCTGTTCCGGCCTGCACCTGCTGA
TGGCGCGGGGCCGCTACAGCCTCGCGGTGGGCCGGCACGAGGCGGCGTACGCCGCGTTCCGGGACTGCGGTGAACGG
10 ATGCGCCGGTGGGACGTCGACGTGCCCCGGGCTGGCCCTGTGGCGGGTGGACGCGGCCGAGGCGCTGCTGCCCCGGCGA
TGACCGGGCGGAGGGCCGGCGGGCTGATCGACGAGCAGCTCACCCGGCCGATGGGGCCCCGGTACAGAGCCCTGACCC
TGCGGGTACGAGCGGCCTACGCCCCGCCGGCGAAACGGATCGACCTGCTCGACGAAGCGGCCGACCTGCTGCTCTCC
AGCAACGACAGTACGAGCGGGCACGGGTGCTGGCCGACCTGAGCGAGGCGTTACGCGCGCTCCGGCAGAACGGCCG
GGCGCGCGGCATCTGCGGCAGGCCCGGCACCTGGCCGCCAGTGGGGGCGGTCCCCCTGCTGCGCCGGCTGGGGC
15 TCAAGGCCGGCCGGTCCGGTTCGGCTCGGCCGGCCGCCGAGGGAATCCGCTCCCTGACCGAGGCCGAGCGCCGGGTG
GCCACGCTGGCCGCCGCCGGGCGAGACCAACCGGGAGATCGCCGACCAGCTCTTCGTACCGCCAGCACGGTCGAGCA
GCACCTCACCAACGTGTTCCGCAAGCTCGGCGTGAAGGGCCGCCAGCAATTGCCGGCCGAGCTGGCCGACCTGCGGC
CGCCGGGCTGA

20 SEQ ID NO: 188

GTGGTCACCGTCACCGGCCCAATCGCCTGCGGCAAGACAGAACTGCTTGACGCGGCTGCCGGAAGGTGAGGCCAT
CATTCTGCGCGCGGTCTGCGCGCCAGAAGAGCGGGCTATGCCGTACGCCATGATCGGGCAGCTCATCGACGACCCGG
CGCTCGCGCATCGGGCGCCGGGGCTGGCTGATCGGATAGCCCAGGGCGGGCAGCTGTGCTGAGGGCCGAGAACCGA
CTGCGCAGGGATCTCACCCGTGCCCTGCTGGCGCTTGCCGTGACCGGCCCTGTGCTGATCGGCGTCGACGATGTGCA
25 TCACGCCGACACCGCCTCTTTGAACTGTCTGCTGCATTTGGCGCGCCGGGTCCGTCCGGCCCCGATATCCATGATCT
TCACCGAGTTGCGCAGCCTCACCCCTACTCAGTCACGGTTCAAGGCGGAGCTGCTCAGCCTGCCGTACCACCACGAG
ATCGCGCTGCGTCCGTTTCGGACCGGAGCAATCGGCGGAGCTGGCCCCGCGCCGCTTCGGCCCCGGCCCTCGCCGAGGA
TGTGCTCGTGGGGTTGTATAAAACGACCAGGGGCAATCTGAGTCTCAGCCGTGGACTGATCAGCGATGTGCGGGAGG
CCCTGGCCAACGGAGAGAGCGCCTTCGAGGCGGGCCGCGCGTTCCGGCTGGCGTACCTCGGCTCGCTCTACCGCTGT
30 GGCCCGGTGCGCTGCGGGTTCGCGGAGTGGCTGCCGTGCTGGGCCCCGAGCGCCACCACCGCTGGTGCGCCGTCT
AAGCGGGCTCAGCGCGGAGACGATAGACCGGGCAACCAAGATCCTCACCGAGGGCGGGCTGCTGCTCGACAGCAGT
TCCCGCACCCGGCCGCCGCTCGGTGGTGGTGTGATGACATGTCCGCCCAGGAACGACGCGGCCTGCACACTCTCGCC
CTGGAAGTGTGACGAGGCGCCGGTTGAAGTGTGCTGCGCACCAACAGGTGCGCGCCGGTCTCATAACGGGGCCCAA
GGCTGCGGAGATGTTCCGCAAGGCCGGCAAGGCTCTGGTTCGTACGCAACGAGTTGGGCGACGCGGCAGAATACCTGC
35 AACTGGCTCACCGGGCCTCCGACGATGTCTCCACCCGGGCCGCTTACGGGTGAGGCGGTGGCGATCGAGCGCCGC
CGCAATCCGCTGGCCTCCAGTCGGCACATGGACGAGCTGAGCGCCGCCGGCCGCGCCGGTCTGCTTTCCCCAAGCA
TGCGGCGCTGGCCGTCTTCTGGCTGGCCGACGGCGGGCGATCCGGCGAGGCAGCCGAGGTGCTGGCGTCCGAACGCC
CGCTAGCGACCAACGATCAGAACCAGGGCCCACTTGCGATTTGTGAGGTGACTCTCGCGCTGTTCTCTCCCGGCGCC
TTCGGATCGGACCGGCGCCACCTCCGCTGACGCGGACGAACTCGCCAGCCTGCCGAAGGCGGCCTGGCAATGCGC
40 GGTGCGCGACAACGCGGCCATGACCGCCTTGACGGTCATCCAGAACTTGCCACCGCTCAGGCGGAAACAGTTCTGC
GGCAGGCTGATTCCGCGAGCCGACGCGATCCCCGCCGCGCTGATCGCCCTGTTGTACGCGGAGAACACCGAGTCCGCT
CATATCTGGGCCGACAAGCTGGGCAGCACGAATGGCGGGGTATCGAACGAGGCGGAAGCGGGCTACGCCGGCCCCGTG

CGCCGAGATCGCCCTGCGGCGGGCGACCTGGCCACGGCGTTTCGAGGCTGGTAGCACCGTCTCTGGACGACCGGTTCG
TGCCGTGCTCGGCATCACCGCCGCATTGCTGTTGAGCAGCAAGACGGCCGCCGCTGTCCGGCTGGGCGAACTCGAG
CGTGCGGAGAAGCTGCTCGCCGAGCCGCTTCCGAACGGCGTCCAGGACAGCCTTTTCGGTCTGCACCTGCTCTCGGC
ATACGGCCAGTACAGCCTCGCGATGGGCCGATATGAATCGGCTCTCCGGGCGTTTCACACCTGCGGAGAACGTATGC
5 GCAGCTGGGATGTTGACGTGCCTGGTCTGGCCCTGTGGCGTGTGACGCCGCCGAGGCGCTGCTCAGCCTCGACCGG
AACGAGGGCCAGCGGCTCATCGACGAACAACCTACCCGTCCGATGGGGCCCTCGTTCCCGCGCGTTAACGCTGCGGAT
CAAGGCGGCATACCTCCCGCGGACGAAGCGGATCCCCCTGCTCCATGAGGCGGCCGAGCTGCTGCTCCCCTGCCCCG
ACCCGTACGAGCAAGCGCGGGTGTCTGCGCGATCTGGGCGACACGCTCAGCGCGCTCAGACGCTATAGCCGGGCGCGG
GGAGTTCTCCGGCAGGCTCGTCACCTGGCCGCCAGTGCGGTGCTGTCCCGCTGCTGCGCAGGCTCGGGGGCGAGCC
10 CGGCCGGATCGACGACGCCGGCCTGCCGCAGCGGAGCACATCGTTGACCGATGCGGAGCGGCGGGTGGCGGCGCTGG
CCGCGGCCCGGACAGACCAACCGGGAGATCGCCAAACAGCTGTTTCGTACGGCCAGCACAGTGAACAGCACCTCACA
AGCGTCTTCCGCAAACTGGGGGTCAAGGGTCGCAAGCAGCTGCCGACCGCGCTGGCCGACGTGGAACAGACCTGA

SEQ ID NO: 189

15 ATGTATAGCGGTACCTGCCGTGAAGGATACGAACTCGTCGCACGCGAGGACGAACTCGGCATTCTACAGAGGTCTCT
GGAACAAGCGAGCAGCGGCCAGGGCGTCTGGTCAACCGTCAACGGCCCAATCGCCTGCGGCAAGACAGAACTGCTTG
ACGCGGCTGCCGCGAAGGCTGAGGCCATCATTCTGCGCGCGGTCTGCGCGCCAGAAGAGCGGGCTATGCCGTACGCC
ATGATCGGGCAGCTCATCGACGACCCGGCGCTCGCGCATCGGGCGCCGGGGCTGGCTGATCGGATAGCCAGGGCGG
GCAGCTGTGCTGAGGGCCGAGAACCGACTGCGCAGGGATCTCACCCGTGCCCTGCTGGCGCTTGCCGTGACCCGGC
20 CTGTGCTGATCGGCGTCGACGATGTGCATCACGCCGACACCGCCTCTTTGAACTGTCTGCTGCATTTGGCGCGCCGG
GTCCGTCCGGCCCGGATATCCATGATCTTACCGAGTTGCGCAGCCTCACCCCTACTCAGTCACGGTTCAAGGCGGA
GCTGCTCAGCCTGCCGTACCACCAGAGATCGCGCTGCGTCCGTTCGGACCGGAGCAATCGGCGGAGCTGGCCCCGG
CCGCCTTCGGCCCGGGCCTCGCCGAGGATGTGCTCGTGGGGTTGTATAAAACGACCAGGGGCAATCTGAGTCTCAGC
CGTGGACTGATCAGCGATGTGCGGGAGGCCCTGGCCAACGGAGAGAGCGCCTTCGAGGCGGGCCGCGCTTCCGGCT
25 GCGGTACCTCGGCTCGCTCTACCGCTGTGGCCCGGTGCGCTGCGGGTTCGCCCGAGTGGCTGCCGTGCTGGGCCCGA
GCGCCACCACCAGCTGGTGCGCCGTCTAAGCGGGCTCAGCGCGGAGACGATAGACCGGGCAACCAAGATCCTCACC
GAGGGCGGGCTGCTGCTCGACCAGCAGTTCCCGCACCCGGCCGCCCGCTCGGTGGTGCTTGATGACATGTCCGCCCCA
GGAACGACGCGGCCTGCACACTCTCGCCCTGGAACCTGCTGGACGAGGCGCCGGTTGAAGTGCTCGCGCACCACCAGG
TCGGCGCCGGTCTCATAACGGGCCCAAGGCTGCGGAGATGTTTCGCCAAGGCCGGCAAGGCTCTGGTCTGACGCAAC
30 GAGTTGGGCGACGCGGCAGAATACCTGCAACTGGCTCACCGGGCCTCCGACGATGTCTCCACCCGGGCGCCCTGCG
GGTTCGAGGCGGTGGCGATCGAGCGCCGCCGCAATCCGCTGGCCTCCAGTCGGCACATGGACGAGCTGAGCGCCGCCG
GCCGCGCCGGTCTGCTTTCCCCCAAGCATGCGGCGCTGGCCGTCTTCTGGCTGGCCGACGGCGGGCGATCCGGCGAG
GCAGCCGAGGTGCTGGCGTCGGAACGCCCGCTAGCGACACCGATCAGAACCAGGGCCCACTTGCGATTTGTGAGGT
GACTCTCGCGCTGTTCTCTCCCGGCGCCTTCGGATCGGACCGGCGCCACCTCCGCTGACGCCGGACGAACTCGCCA
35 GCCTGCCGAAGGCGGCCTGGCAATGCGCGGTGCGCGACAACGCGGCCATGACCGCCTTGACGGTCTATCCAGAACTT
GCCACCGCTCAGGCGGAAACAGTTCTGCGGCAGGCTGATTCGGCAGCCGACGCGATCCCCGCCGCGCTGATCGCCCT
GTTGTACGCGGAGAACACCGAGTCCGCTCATATCTGGGCCGACAAGCTGGGCAGCACGAATGGCGGGGTATCGAACG
AGGCGGAAGCGGGCTACGCCGGCCCGTGCGCCGAGATCGCCCTGCGGCGCGGCGACCTGGCCACGGCGTTGAGGCT
GGTAGCACCGTCTTGACGACCGGTGCTGCGCTCGCTCGGCATCACCGCCGATGCTGTTGAGCAGCAAGACGGC
40 CGCCGCTGTCCGGCTGGGCGAACTCGAGCGTGCAGAGAAGCTGCTCGCCGAGCCGCTTCCGAACGGCGTCCAGGACA
GCCTTTTCGGTCTGCACCTGCTCTCGGCATACGGCCAGTACAGCCTCGCGATGGGCCGATATGAATCGGCTCTCCGG
GCGTTTCACACCTGCGGAGAACGTATGCGCAGCTGGGATGTTGACGTGCCTGGTCTGGCCCTGTGGCGTGTGACGCG

CGCCGAGGCGCTGCTCAGCCTCGACCGGAACGAGGGCCAGCGGCTCATCGACGAACAACCTACCCGTCGGATGGGGC
CTCGTTCCCGCGCGCTGACGCTGCGGATCAAGGCGGCATACTCCCGCGGACGAAGCGGATCCCCCTGCTCCATGAG
GCGGCCGAGCTGCTGCTCCCCTGCCCCGACCCGTACGAGCAAGCGCGGGTGTCTGCCGATCTGGGCGACACGCTCAG
CGCGCTCAGACGCTATAGCCGGGCGCGGGGAGTTCTCCGGCAGGCTCGTCACCTGGCCGCCAGTGCGGTGCTGTCC
5 CGCTGCTGCGCAGGCTCGGGGGCGAGCCCGGCCGGATCGACGACGCCGGCCTGCCGACGCGGAGCACATCGTTGACC
GATGCGGAGCGGCGGGTGGCGGCGCTGGCCGCGGCGGACAGACCAACCGGGAGATCGCCAAACAGCTGTTCTGTCAC
GGCCAGCACAGTGAACAGCACCTCACAAGCGTCTTCCGCAAACTGGGGGTCAAGGGTCGCAAGCAGCTGCCGACCG
CGCTGGCCGACGTGGAACAGACCTGA

10 SEQ ID NO: 190

ATGCCTGCCGTGGAGAGCTATGAACTGGACGCCCCGCGATGACGAGCTCAGAAGACTGGAGGAGGCGGTAGGCCAGGC
GGGCAACGGCCGGGGTGTGGTGGTCACCATCACCGGGCCGATCGCCTGCGGCAAGACCGAACTGCTCGACGCGGCCG
CCGCGAAGAGCGACGCCATCACATTACGTGCGGTCTGCTCCGAGGAGGAACGGGGCCCTCCCGTACGCCCTGATCGGG
CAGCTCATCGACAACCCGGCGGTGCGCTCCAGCTGCCGGATCCGGTCTCCATGGCCCTCCCGGGCGAGCACCTGTC
15 GCCGGAGGCCGAGAACCGGTGCGCGGCGACCTACCCGTACCTGCTGGCGCTCGCCGCCGAACGGCCGGTGTCTGA
TCGGCATCGACGACATGCACCACGCGGACACCGCCTCTTTGAACTGCCTGCTCCACCTGGCCCGGAGGGTTCGGCCCCG
GCCCGGATCGCCATGGTCCTCACCGAGCTGCGCCGGCTCACCCCGGCCCACTCCAGTTCCACGCCGAGCTGCTCAG
CCTGGGGCACCAACGCGAGATCGCGCTGCGCCCGCTCGGCCCGAAGCACATCGCCGAGCTGGCCCGCGCCGGCCTCG
GTCCCGATGTCTGACGAGGACGTGCTCACGGGGTTGTACCGGGCGACCGGCGGCAACCTGAACCTCGGCCACGGACTG
20 ATCAAGGATGTGCGGGAGGCCTGGGCGACGGGCGGGACGGGCATCAACGCGGGCCGCGCGTACCGGCTGGCGTACCT
CGGTTCCCTCTACCGCTGCGGCCCGGTCCCGTTGCGGGTCGCACGGGTGGCCGCCGTGCTGGGCCAGAGCGCCAACA
CCACCCTGGTGCCTGGATCAGCGGGCTCAACGCGGACGCGGTGGGCGAGGCGACCGAGATCCTCACCGAGGGCGGGC
CTGCTGCACGACCTGCGGTTCCCGCATCCGGCGGCGCCGTTGCGTCTGTAACGACCTGTCCGCCCGGGAACGCCG
CCGACTGCACCGGTCCGCTCTGGAAGTGTGATGACGTACCCGTTGAAGTGGTCGCGCACCACCAGGCCGGTGGCG
25 GTTTCATCCACGGTCCCAAGGCCGCCGAGATCTTCGCCAAGGCCCGGCCAGGAGCTGCATGTGCGCGGCGAGCTGGAC
GCCGCGTCCGACTATCTGCAACTGGCCCAACACGCTCCGACGACGCGCTACCCGGGGCCGCGCTGCGGGTCGAGGC
CGTGGCGATCGAGCGCCGCCGAACCCGCTGGCCTCCAGCCGCCACCTCGACGAGCTGACCGTCGCCGCCCGTGCCG
GTCTGCTCTCCCTCGAGCACGCCGCGCTGATGATCCGCTGGCTGGCTCTCGGCGGGCGGTCCGGCGAGGCGGCCGAG
GTGCTGGCCGCGCAGCGCCCGCTGCGGTACCGACCAAGGACAGGGCCCACTGCGGGCCGCCGAGGTATCGCTGGC
30 GCTGGTCAGCCCGGGCGCGTCCGGCGTCAGCCCGGTGCGTCCGGCCCCGATCGGCGGCCGCGTCCGCTCCCGCCGG
ATGAGCTCGCGAACCTGCCGAAGGCGGCCCGGCTTTGTGCGATCGCCGACAACGCCGTCATATCGGCCCTGCACGGT
CGTCCCGAGCTTGCTCGGCCGAGGCGGAGAACGTCTGAAGCAGGCTGACTCGGCGGCGGACGGCGCCACCGCCCT
CTCCGCGCTGACGGCCTTGCTGTACGCGGAGAACACCGACACCGCTCAGCTCTGGGCCGACAAGCTCGTCTCCGAGA
CCGGGGCGTCGAACGAGGAGGAAGGCGCGGGCTACGCGGGCCGCGCGCCGAGACCGCGTTGCGCCGCGGCGACCTG
35 GCCGCGGCGGTGAGGCGGGCAGCGCCATTCTGGACCACCGGCGGGGGTGGTTGCTCGGCATCACCGCCGCGCTACC
GCTGAGCAGCGCGGTAGCCGCCGCCATCCGGCTGGGCGAGACCGAGCGGGCGGAGAAGTGGCTCGCCGAGCCGCTGC
CGGAGGCCATTGGGACAGCCTGTTGCGGCTGCACCTGCTCTCGGCGCGCGGCCAGTACTGCCTCGCGACGGGCGCG
CACGAGTCGGCGTACACGGCGTTCCGCACCTGCGGGGAACGGATGCGGAACTGGGGCGTCGACGTGCCGGGTCTGTC
CCTGTGGCGCGTCGACGCCGCCGAGGCGCTGCTGCACGGCCGCGACCGGGACGAGGGCCGACGGCTCATCGACGAGC
40 AGCTCACCCATGCGATGGGACCCCGTTCCCGCGCTTTGACGCTGCGGGTGAGGCGGCGTACAGCCCGCAGGCGCAG
CGGGTCGACCTGCTCGAAGAGGCGGCCGACCTGCTGCTCTCTTGCAACGACCAGTACGAGCGGGCGCGGGTGTCTGCG
CGATCTGAGCGAGGCGTTAGCGCGCTCAGGCACACAGCCGGGCGCGGGGACTGCTCCGGCAGGCCCGGCACCTGG

CCGCCCAGTGC GGCGCGACCCCGCTGCTGCGCCGGCTCGGGGCCAAGCCCGGAGGCCCCGGCTGGCTGGAGGAATCC
GGCCTGCCGCAGCGGATCAAGTCGCTGACCGACGCGGAGCGGGCGGGTGGCGTCGCTGGCCGCCGGCGGCCAGACCAA
CCGCGTGATCGCCGACCAGCTCTTCGTACAGGCCAGCACGGTGGAGCAGCACCTCACGAACGTCTTCCGCAAGCTGG
GCGTCAAGGGCCGCCAGCACCTGCCGGCCGAACCTCGCCAACGCGGAATAG

5

SEQ ID NO: 191

ATGCCTGCCGTGGAGAGCTATGAACTGGACGCCCCGCGATGACGAGCTCAGAAGACTGGAGGAGGCGGTAGGCCAGGC
GGGCAACGGCCGGGGTGTGGTGGTCACCATCACCGGGCCGATCGCCTGCGGCAAGACCGAACTGCTCGACGCGGCCG
CCGCGAAGAGCGACGCCATCACACTGCGTGCGGTCTGCTCCGAGGAGGAACGGGCCCTCCCGTACGCCCTGATCGGG
10 CAGCTCATCGACAACCCGGCGGTGCGCTCCAGCTGCCGGATCCGGTCTCCATGGCCCTCCCGGGCGAGCACCTGTC
GCCGGAGGCCGAGAACCGGTGCGCGGGCGACCTACCCGTACCCTGCTGGCGCTCGCCGCCGAACGGCCGGTGTGA
TCGGCATCGACGACATGCACCACGCCGACACCGCCTCTTTGAACTGCCTGCTCCACCTGGCCCGGAGGGTCGGCCCCG
GCCCGGATCGCCATGGTCCTCACCGAGCTGCGCCGGCTCACCCCGGGCCACTCCAGTTCCACGCCGAGCTGCTCAG
CCTGGGGCACCACCGCGAGATCGCGCTGCGCCCGCTCGGCCCCGAAGCACATCGCCGAGCTGGCCCGCGCCGGCCTCG
15 GTCCCGATGTGCGACGAGGACGTGCTCACGGGGTTGTACCGGGCGACCGGGCGCAACCTGAACCTCGGCCACGGACTG
ATCAAGGATGTGCGGGAGGCCTGGGCGACGGGCGGGACGGGCATCAACGCGGGCCGCGCTACCGGCTGGCGTACCT
CGGTTCCCTCTACCGCTGCGGCCCGGTCCCGTTGCGGGTCGCACGGGTGGCCGCCGTGCTGGGCCAGAGCGCCAACA
CCACCCTGGTGGCTGGATCAGCGGGCTCAACGCGGACGCGGTGGGCGAGGCGACCGAGATCCTCACCGAGGGCGGC
CTGCTGCACGACCTGCGGTTCCCGCATCCGGCGGGCCCGTTGCGTCTACTCAACGACCTGTCCGCCCGGGAACGCCG
20 CCGACTGCACCGGTCCGCTCTGGAAGTGCTGGATGACGTACCCGTTGAAGTGGTCGCGCACCAACAGGCCGGTGCCG
GTTTCATCCACGGTCCCAAGGCCGCCGAGATCTTCGCCAAGGCCGGCCAGGAGCTGCATGTGCGCGGCGAGCTGGAC
GCCGCGTCCGACTATCTGCAACTGGCCCCACCACGCCTCCGACGACGCGGTACCCGGGCCGCGCTGCGGGTTCGAGGC
CGTGGCGATCGAGCGCCGCCGAACCCGCTGGCCTCCAGCCGCCACCTCGACGAGCTGACCGTCGCCGCCCGTGCCG
GTCTGCTCTCCCTCGAGCACGCCGCGCTGATGATCCGCTGGCTGGCTCTCGGCGGGCGGTCCGGCGAGGCGGCCGAG
25 GTGCTGGCCGCGCAGCGCCCGCTGCGGTACCGACAGGACAGGGCCCACTGCGGGCCCGCGAGGTATCGCTGGC
GCTGGTCAGCCCGGGCGCGTCCGGCGTCAGCCCGGTGCGTCCGGCCCGGATCGGCGGCCGCGTCCGCTCCCGCCGG
ATGAGCTCGCGAACCTGCCGAAGGCGGCCCGGCTTTGTGCGATCGCCGACAACGCCGTATATCGGCCCTGCACGGT
CGTCCCAGCTTGCTCGGCCGAGGCGGAGAACGTCTGAAGCAGGCTGACTCGGCGGCGGACGGCGCCACCGCCCT
CTCCGCGCTGACGGCCTTGCTGTACGCGGAGAACACCGACACCGCTCAGCTCTGGGCCGACAAGCTCGTCTCCGAGA
30 CCGGGGCGTCAACGAGGAGGAAGGCGCGGGCTACGCGGGGCCGCGCGCCGAGACCGCGTTGCGCCGCGGCGACCTG
GCCGCGGCGGTGAGGCGGGCAGCGCCATTCTGGACCACCGGCGGGGGTGGTTGCTCGGCATCACCGCCGCGCTACC
GCTGAGCAGCGCGGTAGCCGCCGCCATCCGGCTGGGCGAGACCGAGCGGGCGGAGAAGTGGCTCGCCGAGCCGCTGC
CGGAGGCCATTGCGGACAGCCTGTTGCGGTGACCTGCTCTCGGCGCGCGGCCAGTACTGCCTCGCGACGGGCCGG
CACGAGTCGGCGTACACGGCCTTCCGCACCTGCGGGGAACGGATGCGGAACTGGGGCGTCGACGTGCCGGGTCTGTC
35 CCTGTGGCGCGTCGACGCCGCCGAGGCGCTGCTGCACGGCCGCGACCGGGACGAGGGCCGACGGCTCATCGACGAGC
AGCTCACCCATGCGATGGGACCCCGTTCCCGCGCTTTGACGCTGCGGGTGCAGGCGGCGTACAGCCCGCAGGCGCAG
CGGGTGCACCTGCTCGAAGAGGCGGCCGACCTGCTGCTCTCTGCAACGACCACTACGAGCGGGCGCGGGTGTCTCGC
CGATCTGAGCGAGGCGTTAGCGCGCTCAGGCACCACAGCCGGGCGCGGGGACTGCTCCGGCAGGCCCGGCACCTGG
CCGCCCAGTGC GGCGCGACCCCGCTGCTGCGCCGGCTCGGGGCCAAGCCCGGAGGCCCCGGCTGGCTGGAGGAATCC
40 GGCCTGCCGCAGCGGATCAAGTCGCTGACCGACGCGGAGCGGGCGGGTGGCGTCGCTGGCCGCCGGCGGCCAGACCAA
CCGCGTGATCGCCGACCAGCTCTTCGTACAGGCCAGCACGGTGGAGCAGCACCTCACGAACGTCTTCCGCAAGCTGG
GCGTCAAGGGCCGCCAGCACCTGCCGGCCGAACCTCGCCAACGCGGAATAG

SEQ ID NO: 192

GTGAAGCGCAACGATCTGGTTGCCC GCGATGGCGAGCTCAGGTGGATGCAAGAGATTCTCAGTCAGGCGAGCGAGGG
CCGGGGGGCCGTGGTCACCATCACGGGGGCGATCGCCTGTGGCAAGACGGTGCTGCTGGACGCCGCGGCAGCCAGTC
5 AAGACGTGATCCAAC T GCGTGCGGTCTGCTCGGCGGAGGAGCAGGAGCTGCCGTACGCGATGGTCGGACAAC TACTC
GACAATCCGGTGCTCGCCGCGCGAGTGCCGGCCCTGGGCAACCTGGCTGCGGCGGGCGAGCGGTGCTGCCGGGCAC
CGAGAACAGGATCCGGCGGGAGCTCACCCGCACCCTGCTGGCTCTCGCCGACGAACGACCGGTGCTGATCGGCGTCG
ACGACATGCACCATGCGGACCCCGCTCGCTGGACTGCCTGCTGCACCTGGCCCCGGCGGGTCGGCCCCGGCCCCGCATC
GCGATCGTTCTGACCGAGTTGCGCCGGCTCACCCCGGCTCACTCGCGCTTCCAGTCCGAGCTGCTCAGCCTGCGGTA
10 CCACCACGAGATCGGGTTGCAGCCGCTCACCGCGGAGCACACCGCCGACCTGGCCCCGCTCGGCCTCGGTGCCGAGG
TCGACGACGACGTGCTCACCGAGCTCTACGAGGCGACCGGCGGCAACCCGAGTCTGTGCTGCGGCCCTGATCAGGGAC
GTGCGGCAGGACTGGGAGGCCGGGGTCACCGGTATCCACGTGCGCCGGGCGTACCGGCTGGCCTATCTCAGTTCGCT
CTACCGCTGCGGCCCGGCGGCGCTGCGGACCGCCCGCGCGGCCGCGGTGCTGGGCGACAGCGCCGACGCCTGCCTGA
TCCGCCGGGTGACGCGCCTCGGTACGGAGGCCGTGGGCCAGGCGATCCAGCAGCTCACCGAGGGCGGCCTGCTGCGT
15 GACCAGCAGTTCCCGCACCCGGCGGCCCGCTCGGTGCTGCTCGACGACATGTCCGCGCAGGAACGCCACGCGATGTA
TCGACGCGCCCGGAGGCGAGCCGCCGAAGGTACGGCCGACCCGGCACCCGGGCGAGCCGCGGGCGGCTACGGCGT
ACGCCGGGTGTGGTGAGCAAGCCGGTGACTACCCGGAGCCGGCCGGCCGGCCCTGCGTGACGGTGCCGGTCCGGCC
GAGTACTGCGGCGACCCGCACGGCGCCGACGACGACCCGGACGAGCTGGTCGCCGCGCTGGGCGGGCTGCTGCCGAG
CCGGCTCGTGCGATGAAGATCCGGCGCCTGGCGGTGGCCGGGCGCCCCGGGGCGGCTGCCGAGCTGCTGACCTCGC
20 AGCGGTTGCACGCGGTGACCAGCGAGGACCGGGCCAGCCTGCGGGCCGCCGAGGTGGCGCTGCCACGCTGTGGCCG
GGTGCGACCGGCCCGGACCGGCATCCGCTCACGGAGCAGGAGGCGGCGAGCCTGCCGAGGGTCCGCGCCTGCTCGC
TGCCGCCGACGATGCCGTGCGGGCCGCCCTGCGCGGTGCGCGCGAGTACGCCGCGGCCGAGGCGGAGAACGTCCTGC
GGCACGCCGATCCGGCAGCCGGTGGTGACGCCTACGCCGCCATGATCGCCCTGCTGTACACGGAGCACCCCGAGAAC
GTGCTGTTCTGGGCCGACAAGCTCGACGCGGGCCGCCCGACGAGGAGACCAGTTATCCCGGGCTGCGGGCCGAGAC
25 CGCGGTGCGGCTCGGTGACCTGGAACGGCGATGGAGCTGGGCCGCGACGGTGCTGGACCAGCGGCGGTGCCGTCCC
TGGGTGTCGCCGCGGGCCTGCTCCTGGGCGGCGCGGTGACGGCCGCCATCCGGCTCGGCGACCTCGACCGGGCGGAG
AAGTGGCTCGCCGAGCCGATCCCCGACGCCATCCGTACCAGCCTCTACGGCCTGCACGTGCTGGCCGCGCGGGGCCG
GCTCGACCTGGCCGCGGGCCGCTACGAGGCGGCGTACACGGCGTTCCGGCTGTGTGGCGAGCGGATGGCAGGCTGGG
ATGCCGATGTCTCCGGGTGGCGCTGTGGCGCGTCGACGCCGCCGAGGCCCTGCTGTCCGCGGGCATCCGCCCGGAC
30 GAGGGCCGCAAGCTCATCGACGACCAGCTCACCCGTGAGATGGGGGCCCGCTCCCGGGCGCTGACGCTGCGGGCGCA
AGCGGCGTACAGCCTGCCGTTGACCGGGTGGGCCTGCTCGACGAGGCGGCCGGCCTGCTGCTCGCCTGCCATGACG
GGTACGAGCGGGCGCGGGTGCTCGCGGACCTGGGGGAGACCCTGCGCACGCTGCGGCACACCGACGCGGCCAGCGG
GTGCTCCGGCAGGCCGAGCAGGCGGCCGCGCGGTGCGGGTCCGTCCCGCTGCTGCGGCGGCTCGGGGCCGAACCCGT
ACGCATCGGCACCCGGCGTGGTGAACCCGGCCTGCCGACGCGGATCAGGCTGCTGACCGATGCCGAGCGGCGGGTTG
35 CCGCGATGGCCGCCGCCGGGCGAGACCAACCGGGAGATCGCCGGTCCGGCTCTTCGTACGCGCCAGCACGGTGGAGCAG
CACCTGACCAGCGTCTTCCGCAAGCTGGGCGTCAAGGGCCGCCGGTTCTGCCGACCGAGCTCGCCCAAGCCGTCTG
A

SEQ ID NO: 193

ATGCCTGCCGTGAAGCGCAACGATCTGGTTGCCC GCGATGGCGAGCTCAGGTGGATGCAAGAGATTCTCAGTCAGGC
GAGCGAGGGCCGGGGGGCCGTGGTCACCATCACGGGGGCGATCGCCTGTGGCAAGACGGTGCTGCTGGACGCCGCGG
CAGCCAGTCAAGACGTGATCCAAC T GCGTGCGGTCTGCTCGGCGGAGGAGCAGGAGCTGCCGTACGCGATGGTCGGA

CAACTACTCGACAATCCGGTGCTCGCCGCGCGAGTGCCGGCCCTGGGCAACCTGGCTGCGGCGGGCGAGCGGCTGCT
GCCGGGCACCGAGAACAGGATCCGGCGGGAGCTCACCCGCACCCTGCTGGCTCTCGCCGACGAACGACCGGTGCTGA
TCGGCGTCGACGACATGCACCATGCGGACCCCGCCTCGCTGGACTGCCTGCTGCACCTGGCCCGGCGGGTCGGCCCCG
GCCCGCATCGCGATCGTTCTGACCGAGTTGCGCCGGCTCACCCCGGCTCACTCGCGCTTCCAGTCCGAGCTGCTCAG
5 CCTGCGGTACCACCACGAGATCGGGTTGCAGCCGCTCACCGCGGAGCACACCGCCGACCTGGCCCGGCTCGGCCTCG
GTGCCGAGGTGACGACGACGTGCTCACCGAGCTCTACGAGGCGACCGGCGGCAACCCGAGTCTGTGCTGCGGCCTG
ATCAGGGACGTGCGGCAGGACTGGGAGGCCGGGGTCACCGGTATCCACGTGCGCCGGGCGTACCGGCTGGCCTATCT
CAGTTGCTCTACCGCTGCGGCCCCGGCGGCGCTGCGGACCGCCCCGCGCGGGCCGCGGTGCTGGGCGACAGCGCCGACG
CCTGCCTGATCCGCCGGGTACGCGGCCTCGGTACGGAGGCCGTGGGCCAGGCGATCCAGCAGCTCACCGAGGGCGGGC
10 CTGCTGCGTGACCAGCAGTTCCCGCACCCGGCGGGCCGCTCGGTGCTGCTCGACGACATGTCCGCGCAGGAACGCCA
CGCGATGTATCGCAGCGCCCCGGGAGGCAGCCGCCGAAGGTACGGCCGACCCCGGCACCCCGGGCGAGCCGCGGGCGG
CTACGGCGTACGCCGGGTGTGGTGAGCAAGCCGGTGAATAACCGGAGCCGGCCGGCCGGGCTGCGTGGACGGTGCC
GGTCCGGCCGAGTACTGCGGCGACCCGCACGGCGCCGACGACGACCCGGACGAGCTGGTGCCTGCGCGCTGGGCGGGCT
GCTGCCGAGCCGGCTCGTGGCGATGAAGATCCGGCGCTGGCGGTGGCCGGGCGCCCCGGGGCGGCTGCCGAGCTGC
15 TGACCTCGCAGCGGTTGCACGCGGTGACCAGCGAGGACCGGGCCAGCCTGCGGGCCGCGGAGGTGGCGCTCGCCACG
CTGTGGCCGGGTGCGACCGGGCCCGGACCGGCATCCGCTCACGGAGCAGGAGGCGGCGAGCCTGCCGGAGGGTCCCGG
CCTGCTCGCTGCCGCCGACGATGCCGTGCGGGCCGCCCTGCGCGGTGCGCGCGAGTACGCCGCGGCCGAGGCGGAGA
ACGTCTGCGGCACGCCGATCCGGCAGCCGGTGGTGACGCCTACGCCGCCATGATCGCCCTGCTGTACACGGAGCAC
CCCGAGAACGTGCTGTTCTGGGCCGACAAGCTCGACGCGGGCCGCCCGACGAGGAGACCAGTTATCCCGGGCTGCG
20 GGCCGAGACCGCGGTGCGGCTCGGTGACCTGGAACCGCGATGGAGCTGGGCCGACCGGTGCTGGACCAGCGGCGGC
TGCCGTCCCTGGGTGTGCGCGCGGGCCTGCTCCTGGGCGGCGCGGTGACGGCCGCCATCCGGCTCGGCGACCTCGAC
CGGGCGGAGAAGTGGCTGCGCGAGCCGATCCCCGACGCCATCCGTACCAGCCTCTACGGCCTGCACGTGCTGGCCGC
GCGGGGCCGGCTCGACCTGGCCGCGGGCCGCTACGAGGCGGCGTACACGGCGTTCGGCTGTGTGGCGAGCGGATGG
CAGGCTGGGATGCCGATGTCTCCGGGCTGGCGCTGTGGCGCGTGCACGCCGCCGAGGCCCTGCTGTCCGCGGGCATC
25 CGCCCGGACGAGGGCCGCAAGCTCATCGACGACAGCTCACCCGTGAGATGGGGGCCCGCTCCCGGGCGCTGACGCT
GCGGGCGCAAGCGGCGTACAGCCTGCCGGTGCACCGGTGGGCTGCTCGACGAGGCGGCCGGCCTGCTGCTCGCCT
GCCATGACGGGTACGAGCGGGCGCGGGTGTGCGGACCTGGGGGAGACCCTGCGCACGCTGCGGCACACCGACGCG
GCCAGCGGGTGCTCCGGCAGGCCGAGCAGGCGGCCGCGCGGTGCGGGTGGTCCCGCTGCTGCGGCGGCTCGGGGC
CGAACCCGTACGCATCGGCACCCGGCGTGGTGAACCCGGCCTGCCGACGCGGATCAGGCTGCTGACCGATGCCGAGC
30 GCGGGGTTGCCGCGATGGCCGCCGCGGGCAGACCAACCGGGAGATCGCCGGTGGCTCTTCGTACGCGCCAGCACG
GTGGAGCAGCAGCTGACCAGCGTCTTCCGCAAGCTGGGCGTCAAGGGCCGCCGGTTCCTGCCGACCGAGCTCGCCCA
AGCCGTCTGA

SEQ ID NO: 194

GTGGTCACCGTACCGGCCCAATCGCCTGCGGCAAGACAGAACTGCTTGACGCGGCTGCCGGAAGGCTGAGGCCAT
35 CATTCTGCGCGCGGTCTGCGCGCCAGAAGAGCGGGCTATGCCGTACGCCATGATCGGGCAGCTCATCGACGACCCGG
CGCTGCGCATCGGGCGCCGGGGCTGGCTGATCGGATAGCCCAGGGCGGGCAGCTGTGCTGAGGGCCGAGAACCGA
CTGCGCAGGGATCTCACCCGTGCCCTGCTGGCGCTTGCCGTGGACCGGCCTGTGCTGATCGGCGTCGACGATGTGCA
TCACGCCGACACCGCCTCTTTGAACTGTCTGCTGCATTTGGCCCCCGGGTCCGTCCGGCCCCGATATCCATGATCT
40 TCACCGAGTTGCGCAGCCTCACCCCTACTCAGTCACGGTTCAAGGCGGAGCTGCTCAGCCTGCCATACCACCACGAG
ATCGCGCTGCGTCCATTCCGACCGGAGCAATCGGCGGAGCTGGCTGCGCGCGCCTTCGGCCCCGGCCTCGCCGAGGA
TGTGCTGCGGGGTTGTATAAAACGACAGGGGCAATCTGAGTCTCAGCCGTGGACTGATCAGCGATGTGCGGGAGG

CCCTGGCCAACGGAGAGAGCGCTTTTCGAGGCGGGCCGCGCGTTCCGGCTGGCGTACCTCAGCTCGCTCTACCGCTGT
GGCCCGGTTCGCGCTGCGGGTTCGCCCAGTGGCTGCCGTGCTGGGCCCCAAGCGCCACCACCACGCTGGTTCGCGCCGGCT
AAGCGGGCTCAGCGCGGAGACGATAGACCGGGCAACCAAGATCCTCACTGAGGGCGGGCTGCTGCTCGACCAGCAGT
TCCCGCACCCGGCCGCCCGCTCGGTGGTGTCTGATGACATGTCCGCCCAGGAACGACGAGCCTGCACACTCTCGCC
5 CTGGAAGTGTGAGACGAGGCGCCGGTTGAAGTGTCTGCGCACACCACCAGGTTCGGCGCCGGTCTCATAACGGGCCCCAA
GGCTGCGGAGATGTTTCGCCAAGGCCGGCAAGGCTCTGGTTCGTACGCAACGAGTTGGGCGACGCGGCCGAATACCTGC
AACTGGCTCACCGGGCCTCCGACGATGTCTCCACCCGGGCGCCCTTACGGGTTCGAGGCCGTGGCCATCGAGCGCCGC
CGCAATCCGCTGGCCTCCAGTCGGCACATGGACGAACTGAGCGCCGCCGGCCGCGCCGGTCTGCTTTCCCCCAAGCA
TGCGGCGCTGGCCGTCTTCTGGCTAGCCGACGGCGGGCGATCCGGCGAGGCAGCCGAAGTGTGGCGTCGGAACGCC
10 CGCTCGCGACCACCGATCAGAACC GGCCCCACCTGCGATTTGTGAGGTGACTCTCGCGCTGTTCTCTCCCGGCGCC
TTCGGATCGGACCGGCGCCACCTCCGCTGACGCCGACGAACTCGCCAGCCTGCCGAAGGCGGCCCTGGCAATGCGC
GGTCGCCGACAACGCGGCCATGACCGCCTTGACGGCCATCCAGAACTTGCCACCGCTCAGGCGGAAACAGTTCTGC
GGCAGGCTGATTTCGGCAGCCGACGCGATCCCCGCGCGCTGATCGCCCTGTTGTACGCGGAGAACACCGAGTCCGCT
CATATCTGGGCGGACAAGCTGGGCAGCACGAATGCCGGGTATCGAACGAGGCGGAAGCGGGCTACGCCGGCCCCGTG
15 CGCCGAGATCGCCCTGCGGCGCGGCGACCTGGCCACGGCGTTTCGAGGCTGGTAGCGCCGTCTTGACGACCCGGTTCGC
TGCCGTGCTCGGCATCACCGCCGATTGCTGTTGAGCAGCAAGACGGCCGCCGCTGTCCGGCTGGGCGAACTCGAG
CGTGCGGAGAAGCTGCTCGCCGAGCCGCTTCCGAACGGCGTCCAGGACAGCCTTTTCGGTCTGCACCTGCTCTCGGC
GTACGGCCAGTACAGCCTCGCGATGGGCCGATATGAATCAGCTCACCGGGCGTTTCGCACCTGCGGAGAACGTATGC
GCAGCTGGGATGTTGACGTGCCTGGTCTGGCCCTGTGGCGTGTGACGCGCCGCGAGGCGCTGCTCAGCCTCGACCGG
20 AACGAGGGCCAGCGGCTCATCGACGAACAACCTACCCGTCCGATGGGGCCTCGTTCCACGCGTTAACGCTGCGGAT
CAAGGCGGCATACCTCCCGCGGACGAAGCGGATCCCCCTGCTCCATGAGGCGGCCGAGCTGCTGCTCCCCTGCCCCG
ACCCGTACGAGCAAGCGCGGGTGTCTCGCCGATCTGGGCGACACGCTCAGCGCGCTCAGACGCTATAGCCGGGCGCGG
GGAGTTCTCCGGCAGGCTCGTCACCTGGCCACCCAGTGCGGTGTGTCGCGTGTGCTGCGCAGGCTCGGGGGCGAGCC
CGGCCGGATCGACGACGCGGCCCTGCCGACGCGGAGCACATCGTTGACCGATGCGGAGCGGCGGGTGGCGGCGCTGG
25 CCGCGGCCCGGACAGACCAACCGGGAGATCGCCGAACAGCTGTTTCGTACAGGCCAGCACAGTGAACAGCACCTCACA
AGCGTCTTCCGCAAGCTGGGCGTCAAGGGCCGCAAGCAGCTGCCGACCGCGCTGGCCGACGTGGAACAGACCTGA

SEQ ID NO: 195

ATGTATAGCGGTACCTGCCGTGAAGGATACGAACTCGTCGCACGCGAGGACGAACTCGGTATTCTACAGAGGTCTCT
30 GGAACAAGCGAGCAGCGGCCAGGGCGTCTGGTACCGTACCGGCCCAATCGCCTGCGGCAAGACAGAACTGCTTG
ACGCGGCTGCCGCGAAGGCTGAGGCCATCATTCTGCGCGCGGTCTGCGCGCCAGAAGAGCGGGCTATGCCGTACGCC
ATGATCGGGCAGCTCATCGACGACCCGGCGCTCGCGCATCGGGCGCCGGGGCTGGCTGATCGGATAGCCCAGGGCGG
GCAGCTGTGCTGAGGGCCGAGAACCAGTTCGCGAGGGATCTCACCCGTGCCCTGCTGGCGCTTGCCGTGGACCGGC
CTGTGCTGATCGGCGTCGACGATGTGCATCACGCCGACACCGCCTCTTTGAACTGTCTGCTGCATTTGGCCCCGCGG
35 GTCCGTCCGGCCCGGATATCCATGATCTTACCGAGTTGCGCAGCCTCACCCCTACTCAGTCACGGTTCAAGGCGGA
GCTGCTCAGCCTGCCATACCACCACGAGATCGCGCTGCGTCCATTTCGACCGGAGCAATCGGCGGAGCTGGCTCGCG
CCGCCTTCGGCCCGGGCCTCGCCGAGGATGTGCTCGCGGGGTGTATAAAACGACCAGGGGCAATCTGAGTCTCAGC
CGTGACTGATCAGCGATGTGCGGGAGGCCCTGGCCAACGAGAGAGCGCTTTTCGAGGCGGGCCGCGCTTCCGGCT
GGCGTACCTCAGCTCGCTCTACCGCTGTGGCCCGGTTCGCGCTGCGGGTTCGCCGAGTGGCTGCCGTGCTGGGCCCCAA
40 GCGCCACCACCACGCTGGTGCGCCGGCTAAGCGGGCTCAGCGCGGAGACGATAGACCGGGCAACCAAGATCCTCACT
GAGGGCGGGCTGCTGCTCGACCAGCAGTTCCCGCACCCGGCCGCCCGCTCGGTGGTGTCTGATGACATGTCCGCCCA
GGAACGACGACGCTGCACACTCTCGCCCTGGAAGTGTGAGACGAGGCGCCGGTTGAAGTGTCTGCGCACACCACGAG

TCGGCGCCGGTCTCATACACGGGCCCAAGGCTGCGGAGATGTTTCGCCAAGGCCGGCAAGGCTCTGGTTCGTACGCAAC
GAGTTGGGCGACGCGGCCGAATACCTGCAACTGGCTCACCGGGCCTCCGACGATGTCTCCACCCGGGCGCCCTGCG
GGTCGAGGCCGTGGCCATCGAGCGCCGCCGAATCCGCTGGCCTCCAGTCGGCACATGGACGAACTGAGCGCCGCCG
GCCGCGCCGGTCTGCTTTCCCCAAGCATGCGGCGCTGGCCGTCTTCTGGCTAGCCGACGGCGGGCGATCCGGCGAG
5 GCAGCCGAAGTGCTGGCGTCGGAACGCCCCGCTCGCGACCACCGATCAGAACC GGCGCCACCTGCGATTTGTCGAGGT
GACTCTCGCGCTGTTCTCTCCCGGCGCCTTCGGATCGGACCGGCGCCACCTCCGCTGACCGCGGACGAACTCGCCA
GCCTGCCGAAGGCGGCCTGGCAATGCGCGGTGCGCGACAACGCGGCCATGACCGCCTTGACGCGCCATCCAGAACTT
GCCACCGCTCAGGCGGAAACAGTTCTGCGGCAGGCTGATTGCGCAGCCGACGCGATCCCCGCCGCGCTGATCGCCCT
GTTGTACGCGGAGAACACCGAGTCCGCTCATATCTGGGCCGACAAGCTGGGCAGCACGAATGCCGGGGTATCGAACG
10 AGGCGGAAGCGGGCTACGCCGGCCCGTGCGCCGAGATCGCCCTGCGGCGCGGCGACCTGGCCACGGCGTTTCGAGGCT
GGTAGCGCCGTCTTGACGACCGGTGCGTGCCGTGCTCGGCATCACCGCCGCATTGCTGTTGAGCAGCAAGACGGC
CGCCGCTGTCCGGCTGGGCGAACTCGAGCGTGCGGAGAAGCTGCTCGCCGAGCCGCTTCCGAACGGCGTCCAGGACA
GCCTTTTTCGGTCTGCACCTGCTCTCGGCGTACGGCCAGTACAGCCTCGCGATGGGCCGATATGAATCAGCTCACCGG
GCGTTTTCGCACCTGCGGAGAACGTATGCGCAGCTGGGATGTTGACGTGCTGGTCTGGCCCTGTGGCGTGTGACGCG
15 CGCCGAGGCGCTGCTCAGCCTCGACCGGAACGAGGGCCAGCGGCTCATCGACGAACAACCTCACCCGTCCGATGGGGC
CTCGTTCCACGCGCTGACGCTGCGGATCAAGGCGGCATACCTCCCGCGGACGAAGCGGATCCCCCTGCTCCATGAG
GCGGCCGAGCTGCTGCTCCCCTGCCCCGACCCGTACGAGCAAGCGCGGGTGTCTCGCCGATCTGGGCGACACGCTCAG
CGCGCTCAGACGCTATAGCCGGGCGCGGGGAGTTCTCCGGCAGGCTCGTCACCTGGCCACCCAGTGCGGTGCTGTCC
CGCTGCTGCGCAGGCTCGGGGGCGAGCCCGCCGGATCGACGACGCCGGCCTGCCGCAGCGGAGCACATCGTTGACC
20 GATGCGGAGCGGCGGGTGGCGGCGCTGGCCGCGGCCGACAGACCAACCGGGAGATCGCCGAACAGCTGTTTCGTAC
GGCCAGCACAGTGAACAGCACCTCACAAGCGTCTTCCGCAAGCTGGGCGTCAAGGGCCGCAAGCAGCTGCCGACCG
CGCTGGCCGACGTGGAACAGACCTGA

SEQ ID NO: 196

GTGTATAGCGGTACCTGCCGTGAAGGATACGAACTCGTCGCCCCGCGAGGACGAACTCGGCATTCTGCAGAGGTCTCT
GGAAGAAGCAGGCAGCGGCCAGGGCGCCGTGGTCACCGTCACCGGCCCGATCGCCTGCGGCAAGACAGAAGTGTCTG
ACGCGGCTGCCGCGAAGGCTGACGCCATCATTCTGCGCGCGGTCTGCGCGCCCGAAGAGCGCGCTATGCCGTACGCC
ATGATCGGGCAGCTCATCGACGACCCGGCGCTCGCGCATCGGGCGCCGGAGCTGGCTGATCGGATAGCCAGGGCGG
GCATCTGTCTGCTGAGGGCCGAGAACCGACTGCGCAGGGATCTCACCCGTGCCCTGCTGGCGCTTGCCGTGACCGGC
30 CTGTGCTGATCGGCGTCGACGATGTGCATCACGCCGACACCGCCTCTTTGAACTGTCTGCTGCATTTAGCCCGCCGG
GTCCGTCCGGCCCGGATATCCATGATCTTCACCGAGTTGCGCAGCCTCACCCCTACTCAGTCACGATTCAAGGCGGA
GCTGCTCAGCCTGCCGTACCACCACGAGATCGCGCTGCGTCCACTCGGACCGGAGCAATCGGCGGAGCTGGCCACG
CCGCCTTCGGCCCGGGCCTCGCCGAGGATGTGCTCGCGGGGTGTATGGGATGACCAGGGGCAACCTGAGTCTCAGC
CGTGGA CTGATCAGCGATGTGCGGGAGGCCCAGGCCAACGGAGAGAGCGCTTTTCGAGGTGGGCGCGCGTTCCGGCT
35 GCGGTACCTCAGCTCGCTCTACCGCTGTGGCCCGATCGCGCTGCGGGTTCGCCGAGTGGCTGCCGTGCTGGGCCCAA
GCGCCACCACCACGCTGGTGCGCCGTCTAAGCGGGCTCAGCGCGGAGACGATAGACCGGGCAACCAAGATCCTCACT
GAGGGCGGGCTGCTGCTCGACCAACAGTTCCCGCACCCGGCCGCCCGCTCGGTGGTGTCTCGATGACATGTCCGCCCA
GGAACGACGCAGCCTGCACACTCTCGCCCTGGAAGTGTGAGACGAGGCGCCGGTTGAAGTGTCTGCGCACCACCAGG
TCGGCGCCGGTCTCATACACGGGCCCAAGGCTGCGGAGATATTCGCCAGGGCTGGCCAGGCTCTGGTTGTACGCAAC
40 GAGTTGGGCGACGCGGCCGAATACCTGCAACTGGCTCACCGAGCCTCCGACGATGTCTCCACCCGGGCGCCCTTACG
GGTCGAGGCCGTGGCAATCGAGCGCCGCCGAATCCGCTGGCCTCCAGTCGTACATGGACGAGCTGAGCGCCGCCG
GCCGCGCCGGTCTGCTTTCCCCAAGCATGCGAGCGCTGGCTGTCTTCTGGCTGGCCGACGGCGGGCGATCCGGCGAG

GCAGCCGAGGTGCTGGCGTCGGAACACCCGCTCGCGACCACCGATCAGAACCGAGCACACCTGCGATTTGCCGAGGT
GACTCTCGCGCTGTTCTGTCCCGGCGCCTTCGGGTTCGGACCGGCGCCACCTCCGCTGGCGCCGGACGAGCTCGCCA
GCTTGCCGAAGGCGGCCTGGCAATGCGCGGTGCGCGACAACGCGGTTCATGACAGCGTTGCATGCTCATCCAGAACTT
GCCACCGCTCAGGCGGAAACAGTTCTGCGGCAGGCTGATTTCGGCAGCCGACGCAATCCCCGCCGCACTGATCGCCCT
5 GTTGTACGCAGAGAACACCGAGTCCGCTCAGATCTGGGCCGACAAGCTGGGCAGCACCAATGCCGGGGTATCGAACG
AGGCGGAAGCGGGCTACGCCGGCCCGTGCGCCGAGATCGCCCTGCGGCGCGGGGACCTGGCCACGGCGTTTCGAGGCT
GGTGGCACCGTCTTGACGACCGGCCGCTGCCGTGCTCGGCATCACCGCCGCATTGCTGTTGAGCAGCAAGACGGC
AGCCGCTGTCCGCTGGGCGAACTCGAGCGTGCAGGAGAAGCTGCTCGCTGAGCCGCTTCCGAACGGTGTCCAGGACA
GCCTTTTCGGTCTGCACCTGCTCTCGGCGCACGGCCAGTACAGCCTCGCGATGGGCCGATATGAATCGGCTCACCGG
10 GCGTTTCACACCTGCGGAGAACGTATGCGCAGCTGGGGTGTGACGTGCCTGGTCTAGCCCTGTGGCGTGTGACGC
CGCCGAGGCACTGCTCAGCCTCGACCGGAACGAGGGCCAGCGGCTCATCGACGAACAACCTCGCCCGTCCGATGGGAC
CTCGTTCCCGCGCATTAAACGCTGCGGATCAAGGCGGCATACCTCCCCGCGACGAAGCGGATCCCCCTGCTCCATGAG
GCAGCTGAGCTGCTGCTCTCCTGCCCCGACCCGTACGAGCAAGCGCGGGTGTCTGCCGATCTGGGCGACACGCTCAG
CGCGCTCAGACGCTATAGCCGGGCGCGGGGAGTTCTCCGGCAGGCTCGTCACCTGGCCACCCAGTGCGGTGTGTCC
15 CGCTGCTGCGCCGACTCGGGGGCGAGCCCGGCCGGATCGACGACGCCGGCCTGCCGAGCGGAGCACATCGTTGACC
GATGCGGAGCGGCGGGTGTGCGCCCTGGCCGCGGCCGGACAGACCAACCGGGAGATCGCCAAACAGCTATTTCGTCAC
GGCCAGCACCGTGGAAACAGCACCTCACAAGCGTCTTCCGCAAGCTGGGCGTTAAGGGCCGAGGCAGCTACCGACCG
CGCTGGCCGACGTGGAATAG

20 SEQ ID NO: 197

ATGTATAGCGGTACCTGCCGTGAAGGATACGAACTCGTCGCCCCGCGAGGACGAACTCGGCATTCTGCAGAGGTCTCT
GGAAGAAGCAGGCAGCGGCCAGGGCGCCGTGGTACCGTACCGGCCCGATCGCCTGCGGCAAGACAGAACTGCTTG
ACGCGGCTGCCGCGAAGGCTGACGCCATCATTCTGCGCGCGGTCTGCGCGCCCGAAGAGCGCGCTATGCCGTACGCC
ATGATCGGGCAGCTCATCGACGACCCGGCGCTCGCGCATCGGGCGCCGGAGCTGGCTGATCGGATAGCCCAGGGCGG
25 GCATCTGTGCTGAGGGCCGAGAACCAGCTGCGCAGGGATCTCACCCGTGCCCTGCTGGCGCTTGCCGTGACCGGC
CTGTGCTGATCGGCGTCGACGATGTGCATCACGCCGACACCGCCTCTTTGAACTGTCTGCTGCATCTGGCCCCGCCG
GTCCGTCCGGCCCGGATATCCATGATCTTACCGAGTTGCGCAGCCTCACCCCTACTCAGTCACGATTCAAGGCGGA
GCTGCTCAGCCTGCCGTACCACCACGAGATCGCGCTGCGTCCACTCGGACCGGAGCAATCGGCGGAGCTGGCCCACG
CCGCCTTCGGCCCGGGCCTCGCCGAGGATGTGCTCGCGGGGTTGTATGGGATGACCAGGGGCAACCTGAGTCTCAGC
30 CGTGGACTGATCAGCGATGTGCGGGAGGCCAGGCCAACGGAGAGAGCGCTTTCGAGGTGGGCGCGCGTTCGGGCT
GGCGTACCTCAGCTCGCTCTACCGCTGTGGCCCGATCGCGCTGCGGGTTCGCCGAGTGGCTGCCGTGCTGGGCCCAA
GCGCCACCACCACGCTGGTGCGCCGTCTAAGCGGGCTCAGCGCGGAGACGATAGACCGGGCAACCAAGATCCTCACT
GAGGGCGGGCTGCTGCTCGACCACCAGTTCCCGCACCCGGCCCGCCGCTCGGTGGTGTCTCGATGACATGTCCGCCCCA
GGAACGACGCAGCCTGCACACTCTCGCCCTGGAACCTGCTGGACGAGGCGCCGGTTGAAGTGTCTCGCGCACCACCAGG
35 TCGGCGCCGGTCTCATACAGGGCCCCAAGGCTGCGGAGATATTGCCAGGGCTGGCCAGGCTCTGGTTGTACGCAAC
GAGTTGGGCGACGCGGCCGAATACCTGCAACTGGCTCACCGAGCCTCCGACGATGTCTCCACCCGGGCGCCCTGCG
GGTTCGAGGCCGTGGCAATCGAGCGCCGCCGAATCCGCTGGCCTCCAGTCGTACATGGACGAGCTGAGCGCCGCCG
GCCGCGCCGGTCTGCTTTCCCCAAGCATGCAGCGCTGGCTGTCTTCTGGCTGGCCGACGGCGGGCGATCCGGCGAG
GCAGCCGAGGTGCTGGCGTCGGAACACCCGCTCGCGACCACCGATCAGAACCGAGCACACCTGCGATTTGCCGAGGT
40 GACTCTCGCGCTGTTCTGTCCCGGCGCCTTCGGGTTCGGACCGGCGCCACCTCCGCTGGCGCCGGACGAGCTCGCCA
GCTTGCCGAAGGCGGCCTGGCAATGCGCGGTGCGCGACAACGCGGTTCATGACAGCGTTGCATGCTCATCCAGAACTT
GCCACCGCTCAGGCGGAAACAGTTCTGCGGCAGGCTGATTTCGGCAGCCGACGCAATCCCCGCCGCACTGATCGCCCT

GTTGTACGCAGAGAACACCGAGTCCGCTCAGATCTGGGCCGACAAGCTGGGCAGCACCAATGCCGGGGTATCGAACG
AGGGCGAAGCGGGCTACGCCGGCCCGTGCGCCGAGATCGCCCTGCGGCGCGGCGACCTGGCCACGGCGTTTCGAGGGCT
GGTGGCACCGTCTTGACGACCGGCCGCTGCCGTCGCTCGGCATCACCGCCGCATTGCTGTTGAGCAGCAAGACGGC
AGCCGCTGTCCGCTGGGCGAACTCGAGCGTGC GGAGAAGCTGCTCGCTGAGCCGCTTCCGAACGGTGTCCAGGACA
5 GCCTTTTCGGTCTGCACCTGCTCTCGGCGCACGGCCAGTACAGCCTCGCGATGGGCCGATATGAATCGGCTCACCGG
GCGTTTCACACCTGCGGAGAAGCTATGCGCAGCTGGGGTGTGACGTGCCTGGTCTAGCCCTGTGGCGTGTGACGCG
CGCCGAGGCACTGCTCAGCCTCGACCGGAACGAGGGCCAGCGGCTCATCGACGAACAACCTGCCCCGTCCGATGGGAC
CTCGTTCCCGCGCACTGACGCTGCGGATCAAGGCGGCATACCTCCCCGCGGACGAAGCGGATCCCCCTGCTCCATGAG
GCAGCTGAGCTGCTGCTCTCCTGCCCCGACCCGTACGAGCAAGCGCGGGTGTCTGCGCGATCTGGGCGACACGCTCAG
10 CGCGCTCAGACGCTATAGCCGGGCGCGGGGAGTTCTCCGGCAGGCTCGTCACCTGGCCACCCAGTGCGGTGCTGTCC
CGCTGCTGCGCCGACTCGGGGGCGAGCCCGCCGGATCGACGACGCCGGCCTGCCGACGCGGAGCACATCGTTGACC
GATGCGGAGCGGCGGGTGTGCGCCCTGGCCGCGGCCGACAGACCAACCGGGAGATCGCCAAACAGCTATTCTGTAC
GGCCAGCACCGTGGAACAGCACCTCACAAGCGTCTTCCGCAAGCTGGGCGTTAAGGGCCGACGGCAGCTACCGACCG
CGCTGGCCGACGTGGAATAG

SEQ ID NO: 198

GTGTATAGCGGTACCTGCCGTGAAGGATACGAACTCGTCGCACGCGAGGACGAACTCGGCATTCTACAGAGGTCTCT
GGAACAAGCGAGCAGCGGCCAGGGCGTCTGTGGTACCGTACCGGCCCAATCGCCTGCGGCAAGACAGAACTGCTTG
ACGCGGCTGCCGCGAAGGCTGAGGCCATCATTCTGCGCGCGGTCTGCGCGCCCCGAAGAGCGGGCTATGCCGTACGCC
20 ATGATCGGGCAGCTCATCGACGACCCGGCGCTCGCGCATCGGGCGCCGGGGCTGGCTGATCGGATAGCCCAGGGCGG
GCAGCTGTGCTGAGGGCCGAGAACCAGTCTGCGCAGGGATCTCACCCGTGCCCTGCTGGCGCTTGCCGTGCACCCGGC
CTGTGCTGATCGGCGTCTGATGATGTGCATCACGCCGACACCGCCTCTTTGAACTGTCTGCTGCATTTGGCGCGCCGG
GTCCGTCCGGCCCGGATATCCATGATCTTACCGAGTTGCGCAGCCTCACCCCTACTCAGTCACGATTCAAGGCGGA
GCTGCTCAGCCTGCCGTACCACCACGAGATCGCGCTGCGTCCATTTCGACCGGAGCAATCGGCGGAGCTGGCTCGCG
25 CCGCCTTCGGCCCGGGCCTCGCCGAGGATGTGCTCGCGGGGTTGTATAAAACGACCAGGGGCAATCTGAGTCTCAGC
CGTGACTGATCAGCGATGTGCGGGAGGCCCTGGCCAACGAGAGAGCGCTTTTCGAGGCGGGCCGCGCGTTCCGGCT
GGCGTACCTCAGCTCGCTCTACCGCTGTGGCCCGGTGCGCTGCGGGTCTGCCCCGAGTGGCTGCCGTGCTGGGCCCCAA
GCGCCACCACCACGCTGGTGCGCCGGCTAAGCGGGCTCAGCGCGGAGACGATAGACCGGGCAACCAAGATCCTCACC
GAGGGCGGGCTGCTGCTCGACCAGCAGTTTCCGCAACCGGCCCGCCCGCTCGGTGGTGTCTGATGACATGTCCGCCCA
30 GGAACGACGCGGCCTGCACACTCTCGCCCTGGAAGTCTGAGACGAGGCGCCGGTTGAAGTGTCTCGCGCACCACCAGG
TCGGCGCCGGTCTCATACAGGGCCCCAAGGCTGCGGAGATGTTGCGCAAGGCCGGCAAGGCTCTGGTCTGTACGCAAC
GAGTTGGGCGACGCGGCCGAATACCTGCAACTGGCTCACCGGGCCTCCGACGATGTCTCCACCCGGGCCCGCTTACG
GGTCGAGGCCGTGGCGATCGAGCGCCGCCGAATCCGCTGGCCTCCAGTCGGCACATGGACGAGCTGAGCGCCGCCG
GCCGCGCCGGTCTGCTTTCCCCAAGCATGCGGCGCTGGCCGTCTTCTGGCTGGCCGACGGCGGGCGATCCGGCGAG
35 GCAGCCCAGGTGCTGGCGTGGAAACGCCCGCTCGCGACCACCGATCAGAACCGGGGCCACCTGCGATTGTGCGAGGT
GACTCTCGCGCTGTTCTCTCCGGCGCCTTCGGATCGGACCGGCGCCACCTCCGCTGACGCGGACGAACTCGCCA
GCCTGCCGAAGGCGGCCTGGCAATGCGCGGTGCGCGACAACGCGGCCATGACCGCCTTGACGCGCCATCCAGAACTT
GCCACCGCTCAGGCGGAAACAGTTCTGCGGCAGGCTGATTGCGCAGCCGACGCGATCCCCGCCGCGCTGATCGCCCT
GTTGTACGCGGAGAACACCGAGTCCGCTCATATCTGGGCCGACAAGCTGGGCAGCATGAATGCCGGGGTATCGAACG
40 AGGGCGAAGCGGGCTACGCCGGCCCGTGCGCCGAGATCGCCCTGCGGCGCGGCGACCTGGCCACGGCGTTTCGAGGGCT
GGTAGCACCGTCTTGACGACCGGTCACTGCCGTCGCTCGGCATCACCGCCGCATTGCTGTTGAGCAGCAAGACGGC
CGCCGCTGTCCGGCTGGGCGAACTCGAGCGTGC GGAGAAGCTGCTCGCCGAGCCGCTTCCGAACGGCGTCCAGGACA

GCCTTTTCGGTCTGCACCTGCTCTCGGCGTACGGCCAGTACAGCCTCGCGATGGGCCGATATGAATCGGCTCACCGG
GCGTTTCGCACCTGCGGAGAACGTATGCGCAGCTGGGATGTTGACGTGCCTGGTCTGGCCCTGTGGCGTGTGACGC
CGCCGAGGCGCTGCTCAGCCTCGACCGGAACGAGGGCCAGCGGCTCATCGACGAACAACCTACCCGTCCGATGGGAC
CTCGTTCCCGCGCGTTAACGCTGCGGATCAAGGCGGCATACCTCCCGCGGACGAAGCGGATCCCCCTGCTCCATGAG
5 GCGGCCGAGCTGCTGCTCCCCTGCCCCGACCCGTACGAGCAAGCGCGGGTGTCTGCCGATCTGGGCGACACGCTCAG
CGCGCTCAGACGCTATAGCCGGGCGCGGGGAGTTCTCCGGCAGGCTCGTCACCTGGCCACCCAGTGCGGTGCTGTCC
CGCTGCTGCGCCGACTCGGGGGCGAGCCCGGCCGGATCGACGACGCCGGCCTGCCGCAGCGGAGCACATCGTTGACC
GATGCGGAGCGGCGGGTGGCGGGCGCTGGCCGCGGCCGGACAGACCAACCGGGAGATCGCCGAACAGCTGTTTCGTAC
GGCCAGCACAGTGAACAGCACCTCACAAGCGTCTTCCGCAAGCTGGGCGTCAAGGGCCGCAAGCAGCTGCCGACCG
10 CGCTGGCCGACGTGGAACAGACCTGA

SEQ ID NO: 199

ATGTATAGCGGTACCTGCCGTGAAGGATACGAACTCGTCGCACGCGAGGACGAACTCGGCATTCTACAGAGGTCTCT
GGAACAAGCGAGCAGCGGCCAGGGCGTCGTGGTCACCGTCACCGGCCCAATCGCCTGCGGCAAGACAGAACTGCTTG
15 ACGCGGCTGCCGCGAAGGCTGAGGCCATCATTCTGCGCGCGGTCTGCGCGCCCCGAAGAGCGGGCTATGCCGTACGCC
ATGATCGGGCAGCTCATCGACGACCCGGCGCTCGCGCATCGGGCGCCGGGGCTGGCTGATCGGATAGCCAGGGCGG
GCAGCTGTGCTGAGGGCCGAGAACCGACTGCGCAGGGATCTCACCCGTGCCCTGCTGGCGCTTGCCGTGCACCGGC
CTGTGCTGATCGGCGTCGATGATGTGCATCACGCCGACACCGCCTCTTTGAACTGTCTGCTGCATTTGGCGCGCCGG
GTCCGTCCGGCCCGGATATCCATGATCTTCACCGAGTTGCGCAGCCTCACCCCTACTCAGTCACGATTCAAGGCGGA
20 GCTGCTCAGCCTGCCGTACCACCACGAGATCGCGCTGCGTCCATTTCGGACCGGAGCAATCGGCGGAGCTGGCTCGCG
CCGCCTTCGGCCCGGGCCTCGCCGAGGATGTGCTCGCGGGGTGTATAAAACGACCAGGGGCAATCTGAGTCTCAGC
CGTGGA CTGATCAGCGATGTGCGGGAGGCCCTGGCCAACGGAGAGAGCGCTTTCGAGGCGGGCCGCGCGTTCCGGCT
GGCGTACCTCAGCTCGCTCTACCGCTGTGGCCCGGTGCGCTGCGGGTTCGCCGAGTGGCTGCCGTGCTGGGCCCAA
GCGCCACCACCACGCTGGTGCGCCGGCTAAGCGGGCTCAGCGCGGAGACGATAGACCGGGCAACCAAGATCCTCACC
25 GAGGGCGGGCTGCTGCTCGACCAGCAGTTTCCGCACCCGGCCGCCCGCTCGGTGGTGTCTCGATGACATGTCCGCCCCA
GGAACGACGCGGCCTGCACACTCTCGCCCTGGAACCTGCTGGACGAGGCGCCGGTTGAAGTGTCTCGCGCACCACCAGG
TCGGCGCCGGTCTCATAACGGGCCCAAGGCTGCGGAGATGTTTCGCCAAGGCCGGCAAGGCTCTGGTTCGTACGCAAC
GAGTTGGGCGACGCGGCCGAATACCTGCAACTGGCTCACCGGGCCTCCGACGATGTCTCCACCCGGGCCGCCCTGCG
GGTCGAGGCCGTGGCGATCGAGCGCCGCCGCAATCCGCTGGCCTCCAGTCGGCACATGGACGAGCTGAGCGCCGCCG
30 GCCGCGCCGGTCTGCTTTCCCCAAGCATGCGGCGCTGGCCGTCTTCTGGCTGGCCGACGGCGGGCGATCCGGCGAG
GCAGCCAGGTGCTGGCGTCGGAACGCCCGCTCGCGACCAACGATCAGAACCAGGGCCACCTGCGATTTGTGAGGT
GACTCTCGCGCTGTTCTCTCCCGGCGCCTTCGGATCGGACCGGCGCCACCTCCGCTGACGCCGGACGAACTCGCCA
GCCTGCCGAAGGCGGCCTGGCAATGCGCGGTGCGCGACAACGCGGCCATGACCGCTTGACGCGCCATCCAGAACTT
GCCACCGCTCAGGCGGAAACAGTTCTGCGGCAGGCTGATTCGGCAGCCGACGCGATCCCCGCCGCGCTGATCGCCCT
35 GTTGTACGCGGAGAACACCGAGTCCGCTCATATCTGGGCCGACAAGCTGGGCAGCATGAATGCCGGGGTATCGAACG
AGGCGGAAGCGGGCTACGCCGGCCCGTGCGCCGAGATCGCCCTGCGGCGCGGCGACCTGGCCACGGCGTTTCGAGGCT
GGTAGCACCGTCTTGGACGACCGGTCACTGCCGTGCTCGGCATCACCGCCGATTGCTGTTGAGCAGCAAGACGGC
CGCCGCTGTCCGGCTGGGCGAACTCGAGCGTGCAGGAGAAGCTGTCTCGCCGAGCCGCTTCCGAACGGCGTCCAGGACA
GCCTTTTCGGTCTGCACCTGCTCTCGGCGTACGGCCAGTACAGCCTCGCGATGGGCCGATATGAATCGGCTCACCGG
40 GCGTTTCGCACCTGCGGAGAACGTATGCGCAGCTGGGATGTTGACGTGCCTGGTCTGGCCCTGTGGCGTGTGACGC
CGCCGAGGCGCTGCTCAGCCTCGACCGGAACGAGGGCCAGCGGCTCATCGACGAACAACCTACCCGTCCGATGGGAC
CTCGTTCCCGCGCGCTGACGCTGCGGATCAAGGCGGCATACCTCCCGCGGACGAAGCGGATCCCCCTGCTCCATGAG

GCGGCCGAGCTGCTGCTCCCCTGCCCCGACCCGTACGAGCAAGCGCGGGTGCTCGCCGATCTGGGCGACACGCTCAG
CGCGCTCAGACGCTATAGCCGGGCGCGGGGAGTTCTCCGGCAGGCTCGTCACCTGGCCACCCAGTGCGGTGCTGTCC
CGCTGCTGCGCCGACTCGGGGGCGAGCCCGGCCGATCGACGACGCCGGCCTGCCGAGCGGAGCACATCGTTGACC
GATGCGGAGCGGCGGGTGCGGGCGCTGGCCGCGGCCGACAGACCAACCGGGAGATCGCCGAACAGCTGTTCTGTCAC
5 GGCCAGCACAGTGGAAACAGCACCTCACAAGCGTCTTCCGCAAGCTGGGCGTCAAGGGCCGCAAGCAGCTGCCGACCG
CGCTGGCCGACGTGGAACAGACCTGA

SEQ ID NO: 200

GTGCGAGCTATTAATGCGTCCGACACCGGTCTGAACTGGTCGCCCCGGAAGACGAACTGGGACGTGTACGAAGTGC
10 CCTGAACCGAGCGAACGGCGGCCAAGGTGTCTGATCTCCATTACCGGTCCGATCGCCTGCGGCAAGACCGAACTGC
TTGAGGCTGCCGCTCGGAAGTTGACGCCATCACTCTGCGCGCGGTCTGTGCCGCCGAGGAACGGGCGATACCTTAT
GCCCTGATCGGGCAGCTTATCGACAACCCCCGCGCTCGGCATTCCGGTTCCGGATCCGGCCGGCCTGACCGCCAGGG
CGGACGACTGTATCGAGCGCCGAGAACCGACTGCGTCGCGACCTCACCCGTGCCCTGCTGACGCTCGCCACCGACC
GGCTGGTGCTGATCTGTGTGATGACGTGCAGCACGCCGACAACGCCTCGTTGAGCTGCCTTCTGTATCTGGCCCCGA
15 CGGCTTGTCGCCGCTCGAATCGCTCTGGTATTACCGAGTTGCGAGTCTCACCTCGTCTCAGTTACGGTTCAACGC
GGAGCTGCTCAGCTTGCGGAACCACTGCGAGATCGCGCTGCGCCCACTCGGCCCGGGGCATGCGGCCGAGCTGGCCC
GCGCCACCCTCGGCCCGGCCTCTCCGACGAAACACTCACGGAGCTGTACCGGGTGACCGGAGGCAACCTGAGTCTC
AGCCGCGGGCTGATCGACGATGTGCGGGACGCCTGGGCACGAGGGGAAACGGGCGTCCAGGTGGGCCGGGCGTTCCG
GCTGGCCTACCTCGGTTCCCTCCACCGCTGTGGTCCGCTGGCGTTGCGGGTCGCCCCGCTAGCCGCCGTACTGGGCC
20 CGAGCGCCACCAGCGTCTCGGTGCGCCGGATCAGTGGGCTCAGCGCGGAGGCCATGGCCCAGGCGACCGATATCCTC
GCTGACGGCGGCCTCTGCGCGACCAGCGGTTACACATCCAGCGGCCCGCTCGGTGGTGCTCGACGACATGTCCGC
CGAGGAACGACGCAGCGTGCACAGCCTCGCCCTGGAAGTCTGCTGGACGAGGCACCGGCCGAGATGCTCGCGCACCACC
GGGTGCGGCGCCGGTCTCGTGACGGGCCGAAGGCCGCGGAGACATTACCGGGGCGCGCCGGGCACTGGCCGTTCCG
GGCATGCTGGGCGAGGCGAGCCGACTACCTGCAACTGGCGTACCGGGCCTCCGGCGACGCCGCTACCAAGGCCGCGAT
25 ACGCGTCGAGTCCGTGGCGGTGCGAGCGCCGACGCAATCCGCTGGTTCGTCAGTCGCCATTGGGACGAGCTGAGCGTCG
CGGCCCGCGCCGGTCTGCTCTCCTGCGAGCACGTGTCCAGGACGGCCCCGCTGGCTGACCGTCGGTGGGCGGCCCGGC
GAGGCGGCCAGGGTGCTGGCGTCGCAACACCGACGGGTCTGTCACCGATCAGGACCGGGCCACCTGCGGGTCGCCGA
GTTCTCGCTCGCGCTGCTGTACCCCGGTACGTCCGGCTCGGACCGGCGCCCGCACCCGCTCACGTGCGACGAACTCG
CGGCCCTACCGACTGCGACCAGACACTGCGCGATCGCCGATAACGCTGTATGGCTGCCTTGCGTGGTTCATCCGGAG
30 CTTGCCACCGCCGAGGCGAGAAGCCGTTCTGCGAGCAAGCCGACGCGGCGGACGGCGCTGCTCTACCGCGCTGATGGC
CCTGCTGTACGCGGAGAGCATCGAGGTGCGTGAAGTCTGGGCGGACAAGCTGGCGGCAGAGGCCGAGCATCGAACG
GGCAGGACGCGGAGTACGCCGGTATACGCGCCGAAATCGCCCTGCGGCGCGGCGATCTGACCGCGGCCGTCGAGACC
GCCGGCATGGTCTTGACGGCGCGCGCTGCCGTGCTCGACATCACCGCCACGTTGCTGTTGGCCGGCAGGGCGTC
CGTCGCCGTCCGGCTGGGCGAACTCGACCACGCGGAGGAGCTGTTCCGCGCGCCCGCGGAGGACGCCTTCCAGGACA
35 GCCTCTTCGGTCTGCATCTGCTCTCGGCGCACGGCCAGTACAGCCTCGCGACAGGCCGGCCCCGAGTCGGCATAACGG
GCCTTTCTGTCCTGCGGCGAACGTATGCGCGATTGGGGCTTCGACGCGCCCCGGTGTGGCCCTGTGGCGGTCGCGGC
CGCCGAGGCGCTGCTCGGCCTCGACCGGAACGAGGGCCGACGGCTCATCGACGAACAGCTGAGCCGGACGATGGCCC
CCCGGTCCACGCGTTGACGCTGCGGATAAAAGCGGCGTACATGCCGAGCCGAAGCGGGTCGACCTGCTCTACGAA
GCGGCTGAGCTGCTGCTCTCCTGCCGGGACCAGTATGAGCGAGCGCGGGTGCTCGCCGATCTGGGCGAGGCGCTCAG
40 CGCGCTCGGGAACCTACCGGCAGGCGCGAGGTGTGCTCCGGCAGGCTCGGCATCTGGCCATGCGAACGGCGCGGACC
CGCTGCTGCGCCGGCTCGGAATCAGGCCCGGCCGCGAGGACGACCCCGACCCGAGCCGCGGAGCAGATCGCTGACC
AACGCTGAGCGGCGTGCGGCGTCGCTGGCCGCGACCGGACTGACCAACCGGGAGATCGCCGACCGGCTCTTCGTCAC

CGCCAGCACCGTGGAGCAGCACCTCACCAACGTCTTCCGCAAGCTGGGCGTCAAGGGCCGCAAGCAGCTGCCGGCCG
AGTTGGACGACATGGAATAG

SEQ ID NO: 201

5 ATGCGAGCTATTAATGCGTCCGACACCGGTCTGAACTGGTCGCCCCGGAAGACGAACCTGGGACGTGTACGAAGTGC
CCTGAACCGAGCGAACGGCGGCCAAGGTGTCTGATCTCCATTACCGGTCCGATCGCCTGCGGCAAGACCGAACTGC
TTGAGGCTGCCGCCTCGGAAGTTGACGCCATCACTCTGCGCGCGGTCTGTGCCGCCGAGGAACGGGCGATACCTTAT
GCCCTGATCGGGCAGCTTATCGACAACCCCGCGCTCGGCATTCCGGTTCCGGATCCGGCCGGCCTGACCGCCCAGGG
CGGACGACTGTATCGAGCGCCGAGAACCGACTGCGTCGCGACCTCACCCGTGCCCTGCTGACGCTCGCCACCGACC
10 GGCTGGTGCTGATCTGTGTGATGACGTGCAGCACGCCGACAACGCCTCGTTGAGCTGCCTTCTGTATCTGGCCCCGA
CGGCTTGTCCCGGCTCGAATCGCTCTGGTATTACCGAGTTGCGAGTCCTCACCTCGTCTCAGCTGCGGTTCAACGC
GGAGCTGCTCAGCTTGCGGAACCACTGCGAGATCGCGCTGCGCCCACTCGGCCCGGGGCATGCGGCCGAGCTGGCCC
GCGCCACCCTCGGCCCGGCCTCTCCGACGAAACACTCACGGAGCTGTACCGGGTGACCGGAGGCAACCTGAGTCTC
AGCCGCGGGCTGATCGACGATGTGCGGGACGCCTGGGCACGAGGGGAAACGGGCGTCCAGGTGGGCCGGGCGTTCCG
15 GCTGGCCTACCTCGGTTCCCTCCACCGCTGTGGTCCGCTGGCGTTGCGGGTGCCTCCGCTAGCCGCCGTACTGGGCC
CGAGCGCCACCAGCGTCTGGTGCCTCGGATCAGTGGGCTCAGCGCGGAGGCCATGGCCCAGGCGACCGATATCCTC
GCTGACGGCGGCCTCTGCGCGACCAGCGGTTACACATCCAGCGGCCCGCTCGGTGGTGCTCGACGACATGTCCGC
CGAGGAACGACGCGAGCGTGCACAGCCTCGCCCTGGAACCTGCTGGACGAGGCACCGGCCGAGATGCTCGCGCACCACC
GGGTGCGCGCCGGTCTCGTGACGGGCCGAAGGCCGCGGAGACATTCACCGGGGCCGCGCGGGCACTGGCCGTTTCGC
20 GGCATGCTGGGCGAGGCGAGCCGACTACCTGCAACTGGCGTACCGGGCTCCGGCGACGCCGCTACCAAGGCCGCGAT
ACGCGTCGAGTCCGTGGCGGTGAGCGCCGACGCAATCCGCTGGTTCGTGAGTCGCCATTGGGACGAGCTGAGCGTCG
CGGCCCGCGCCGGTCTGCTCTCCTGCGAGCACGTGTCCAGGACGGCCCGCTGGCTGACCGTCCGTGGGCGGCCCGGC
GAGGCGGCCAGGGTGTGGCGTCGCAACACCGACGGGTCTGTCACCGATCAGGACCGGGCCACCTGCGGGTCGCCGA
GTTCTCGCTCGCGCTGCTGTACCCCGGTACGTCCGGCTCGGACCGGCGCCCGCACCCGCTCACGTGCGACGAACTCG
25 CGGCCCTACCGACTGCGACCAGACACTGCGCGATCGCCGATAACGCTGTGATGGCTGCCTTGCGTGGTTCATCCGGAG
CTTGCCACCGCCGAGGCGAGAAGCCGTTCTGCGAGCAAGCCGACGCGGCGGACGGCGCTGCTCTACCCGCGCTGATGGC
CCTGCTGTACGCGGAGAGCATCGAGGTGCTGAACTCTGGGCGGACAAGCTGGCGGCAGAGGCCGAGCATCGAACG
GGCAGGACGCGGAGTACGCCGGTATACGCGCCGAAATCGCCCTGCGGCGCGGCGATCTGACCGCGGCCGTCGAGACC
GCCGGCATGGTCTTGACGGCGGCCGCTGCCGTGCTCGACATCACCGCCACGTTGCTGTTGGCCGGCAGGGCGTC
30 CGTCGCCGTCCGGCTGGGCGAACTCGACCACGCGGAGGAGCTGTTTCGCCGCGCCGCCGAGGACGCCTTCCAGGACA
GCCTCTTCGGTCTGCATCTGCTCTCGGCGCACGGCCAGTACAGCCTCGCGACAGGCCGGCCCCGAGTCGGCATACCGG
GCCTTTTCGTGCTGCGGCGAACGTATGCGCGATTGGGGCTTCGACGCGCCCGGTGTGGCCCTGTGGCGCGTCGGCGC
CGCCGAGGCGCTGCTCGGCCTCGACCGGAACGAGGGCCGACGGCTCATCGACGAACAGCTGAGCCGGACGATGGCCC
CCCGGTCCACGCGTTGACGCTGCGGATAAAAGCGGCGTACATGCCGAGCCGAAGCGGGTCGACCTGCTCTACGAA
35 GCGGCTGAGCTGCTGCTCTCCTGCCGGGACCAGTATGAGCGAGCGCGGGTGTGCTGCGCGATCTGGGCGAGGCGCTCAG
CGCGCTCGGGAACCTACCGGCAGGCGCGAGGTGTGCTCCGGCAGGCTCGGCATCTGGCCATGCGAACCGGCGCGGACC
CGCTGCTGCGCCGGCTCGGAATCAGGCCCCGGCCGGCAGGACGACCCCGACCCGACGCCGCGGAGCAGATCGCTGACC
AACGCTGAGCGGCGTGCAGGCGTGGTGGCGGACCGGACTGACCAACCGGGAGATCGCCGACCGGCTCTTCGTAC
CGCCAGCACCGTGGAGCAGCACCTCACCAACGTCTTCCGCAAGCTGGGCGTCAAGGGCCGCAAGCAGCTGCCGGCCG
40 AGTTGGACGACATGGAATAG

SEQ ID NO: 202

MPAVECYELDARDDELRLKEEVVTGRANGRGVVVTITGPIACGKTELLDAAAAKADAITLRVCSAEEQALPYALIG
 QLIDNPALASHALEPACPTLPGEHLSPEAENRLRSDLTRTLLALAAERPVLIGIDESHANALCLLHLARRVGSARIA
 MVLTELRLTPAHSQFQAELLSLGHREIALRPLSPKHTAELVRAGLGPDVDEDVLTGLYRATGGNLLNLTRGLINDV
 REAWETGGTGISAGRAYRLAYLGSLYRCGPVPLRVARVAAVLGGQSANTTLVRWISGLNADAVGEATEILTEGGLLHD
 5 LRFPHPAARSVVLNMSAQERRRLHRSALVLEDDVPVEVVAHHQVGAGLLHGPKAAEIFAKAGQELHVRGELDTASD
 YLQLAHQASDDAVTGMRAEAVAIEERRRNPLASSRHLDELTVVARAGLLFPEHTALMIRWLVGGRSGEAAAGLLASQR
 PRAVTDQDRAHMRAAEVSLALVSPGTS GPDRRPRPLTPDELANLPKAAARLCAIADNAVMSALRGRPELAAAAEAENVL
 QHADSAAAGTTALAALTALLYAENTDTAQLWADKLVSETGASNEEEAGYAGPRAEAAALRRGDLAAAVEAGSTVLDHR
 RLSTLGITAALPLSSAVAAAIRLGETERAEKWLAQPLPQAIQDGLFGLHLLSARGQYSLATGQHESAYTAFRTCGER
 10 MRNWGVDVPGLSLWRVDAAEALLHGRDRDEGRRLVDEQLTRAMGPRSRALTLRVQAAYSPPAKRVDLLDEAADLLLS
 CNDQYERARVLADLSETFSALRHHSRARGLLRQARHLAAQGAIPLLRRLGAKPGGPGWLEESGLPQRIKSLTDAER
 RVASLAAGGQTNRVIADQLFVTASTVEQHLTDVSTGSRPPAPAAELV

SEQ ID NO: 203

MVPEVRAAPDELIARDDELSRLQRALTRAGSGRGVVVAITGPIASGKTALLDAGAAKSGFVALRAVCSWEERTLPYG
 MLGQLFDHPELAAQAPDLAHTASCESPQAGTDNRLRAEFTRTLLALAADWPVLIGIDDVHHADAESLRCLLHLARR
 IGPARIAVVLTELRRPTPADSRFQAELLSLRSYQEIALRPLTEAQTGELVRRHLGAETHEDVSADTFRATGGNLLLG
 HGLINDIREARTAGRPGVVAGRAYRLAYLSSLYRCGPSALRVARASAVLGASAEAVLVQRMGTGLNKDAVEQVYEQLN
 EGRLLQGERFPHPAARSIVLDDLALERRNLHESALELLRDHGVAGNVLARHQIGAGRVHGEEAVELEFTGAAREHHL
 20 RGELDDAAGYLELAHRASDDPVTRAALRVGAAAIERLCNPVRAGRHLPELLTASRAGLLSSEHAVSLADWLAMGGRP
 GEAAEVLATQRPAADEQHRALLRSGELSLALVHPGAWDPLRRTDRFAAGGLGSLPGPARHRAVADQAVIAALRGRL
 DRADANAESVLQHTDATADRTTAIMALLALLYAENTDAVQFWVDKLAGDEGTRTPADEAVHAGFNAEIALRRGDLMR
 AVEYGEAALGHRHLPTWGMAAALPLSSTVVAAIRLGDLDRAERWLAEPPLPQQTPESLFGLHLLWARGQHHLATGRHG
 AAYTAFRECGERMRRWAVDVPGLALWRVDAAESLLLLGRDRAEGLRLVSEQLSRPMRPRARVQTLRVQAAYSPPPQR
 25 IDLLEEAADLLVTCNDQYELANVLSDLAEASSMVRQHRSRARGLLRRARHLATQCGAVPLLRRLGAEPDIDGAWDAT
 LGQRIASLTESERRVAALAAVGRTNREIAEQLFVTASTVEQHLTNVFRKLAVKGRQQLPKELADVGEPADRRRCG

SEQ ID NO: 204

MIARLSPPDLIARDDEFGSLHRALTRAGGGRGVVAAVTGPIACGKTELLDAAAAKAGFVTLRAVCSMEERALPYGML
 30 GQLLDQPELAARTPELVRLTASCENLPADVNDRLGTTELTRTVLTLAAERPVLIGIDDVHHADAPSLRCLLHLARRIS
 RARVAIVLTELLRPTPAHSQFRAALLSLRHYQEIALRPLTEAQTTELVRRLHGLQDAHDDVVAQAFRATGGNLLLGHG
 LIDDIREARTRTSGCLEVVAGRAYRLAYLGSLYRCGPAALSVARASAVLGESVELTLVQRMGTGLDTEAVEQAHEQLV
 EGRLLREGRFPHPAARSVVLDDLSAAERRGLHELALELLRDRGVASKVLARHQMGTGRVHGAEVAGLFTDAAREHHL
 RGELDEAVTYLEFAYRASDDPAVHAALRVDTAAIERLCDPARSGRHVPELLTASRERLLSSEHAVSLACWLAMDGRP
 35 GEAAEVLAAQRSAAPSEQGRAHLRVADLSLALIYPGAADPPRPADPPAEDEVASFSGAVRHRRAVADKALSNALRGWS
 EQAEAKAEYVLQHSRVTTDRTTMMALLALLYAEDTDAVQSWVDKLAGDDNMRTPADEAVHAGFRAEAAALRRGDLTA
 AVECGEAAALAPRVVPSWGMAAALPLSSTVVAAIRLGDLDRAERWLAEPPEETSDSLFGHLMVWARGQHHLAAGRYR
 AAYNAFRDCGERMRRWSVDVPGLALWRVDAAEALLLLGRGRDEGLRLISEQLSRPMGSRARVMTLRVQAAYSPPAKR
 IELLDEAADLLIMCRDQYELARVLADMGEACGMLRRHSRARGLFRRARHLATQCGAVPLLRRLGGESSDADGTQDVT
 40 PAQRITSLTEAERRVASHAAVGRTNKEIASQLFVTSSTVEQHLTNVFRKLGVKGRQQLPKELSDAG

SEQ ID NO: 205

MEFYDLVARDDDELRRRLDQALGRAAGGRGVVVTVTGPGCGKTELLDAAAAEEEFITLRAVCSAEERALPYAVIGQLL
 DHPVLSARAPDLACVTAPGRTLTPADTENRLRRDLTRALLALASERPVLICIDDVHQADTASLNCLLHLARRVASARI
 AMILTELRRRLTPAHSRFEAEELSLRHRHEIALRPLGPADTAELARARLGAGVTADELAQVHEATSGNPNLVGGLVND
 5 VREAWAAGGTGIAAGRAYRLAYLSSVYRCGPVPLRIAQAAAVLGPSATVTLVRRISGLDAETVDEATAILTEGGLLR
 DHRFPHPAARSVVLDDMSAQERRRLHRSTLDVLDGVPVDVLAHHQAGAGLLHGPQAAEMFARASQELRVRGELDAAT
 EYLQLAYRASDDAGARAALQVETVAGERRRNPLAASRHLDELAAAARAGLLSAEHAALVVHWLADAGRPGEEAAEVL
 LQRALAVTDHDRARLRAAEVSLALFHPGVPGPSDRPLAPEELASLSLSARHGVTADNAVLAALRGRPEASAAEAENV
 LRNADAAASGPTALAALTALLYAENTDAAQLWADKLAAGIGAGEGEAGYAGPRTVAALRRGDLTTAVQAAGAVLDGR
 10 RPSSLGITAVLPLSGAVAAAIRLGELERAEKWLAEPLEAVHDSLFGLHLLMARGRYSLAVGRHEAAYAAFRDCGER
 MRRWDVDVPGALWVRVDAEALLPGDDRAEGRRLIDEQLTRPMGPRSRALTLRVRAAYAPPKAKRIDLLDEAADLLS
 SNDQYERARVLADLSEAFSALRQNGRARGILRQARHLAAQCGAVPLLRRLGVKAGRSGRLLGRPPQGIRSLTEAERRV
 ATLAAAGQTNREIADQLFVTASTVEQHLTNVFRKLGVKGRQQLPAELADLRPPG

15 SEQ ID NO: 206

MYSGTCREGYELVAREDELGILQRSLEQASSGQGVVVTVTGPIACGKTELLDAAAAKAEAIILRAVCAPEERAMPYA
 MIGQLIDDPALAHRAPGLADRIAQGGQLSLRAENRLRRDLTRALLALAVDRPVLIGVDDVHHADTASLNCLLHLARR
 VRPARISMIFTELRLSTPTQSRFKAELLSLPYHHEIALRPFPGPEQSAELARAAFGPGLAEDVLVGLYKTRGNLSLS
 RGLISDVREALANGESAFEAGRAFRLAYLGSLYRCGPVALRVARVAAVLGPSATTTLVRRLSGLSAETIDRATKILT
 20 EGGLLLDQQFPHPAARSVVLDDMSAQERRGLHTLALELLDEAPVEVLAHHQVGAGLIHGPKAAEMFAKAGKALVVRN
 ELGDAAEYLQLAHRASDDVSTRAALRVEAVAIERRRNPLASSRHMDLSAAGRAGLLSPKHAALAVFWLADGGRSGE
 AAEVLASERPLATTDQNRAHLRFVEVTLALFSPGAFGSDRRPPPLTPDELASLPKAAWQCAVADNAAMTALHGHPEL
 ATAQAETVLRQADSAADAIPAALIALLYAENTESAHIWADKLGSTNGGVSNEAEAGYAGPCAELARRGDLATAFEA
 GSTVLDDRSLPSLGITAALLSSKTAAAVRLGELERAEKLLAEPLPNGVQDSLFGLHLLSAYGQYSLAMGRYESALR
 25 AFHTCGERMRSWDVDVPGALWVRVDAEALLSLDRNEGQRLIDEQLTRPMGPRSRALTLRIKAAYLPRTRKRIPLHE
 AAELLPLPCDPYEQARVLADGLTSLARRYSRARGVLRQARHLAAQCGAVPLLRRLGGEPRGRIDDAGLPQRSTSLT
 DAERRVAALAAAGQTNREIAKQLFVTASTVEQHLTSVFRKLGVKGRQLPTALADVEQT

SEQ ID NO: 207

MPAVESYELDARDDELRRLEEAVGQAGNGRGVVVTITGPIACGKTELLDAAAAKSDAITLRAVCSEEEERALPYALIG
 QLIDNPAVASQLPDPVSMALPGEHLSPEAENRLRGDLTRTLALAAERPVLIGIDDMHHADTASLNCLLHLARRVGP
 ARIAMVLTTELRLTPAHSQFHAELLSLGHHEIALRPLGPKHIAELARAGLGPVDVDEDLTGLYRATGGNLLNLGHGL
 IKDVREAWATGGTGINAGRAYRLAYLGSLYRCGPVPLRVARVAAVLGQSANTTLVRWISGLNADAVGEATEILTEGG
 LLHDLRFPHPAARSVVLNDLSARERRRLHRSALVLDVPEVVAHHQAGAGFIHGPKAAEIFAKAGQELHVRGELD
 35 AASDYLQLAHHASDDAVTRAALRVEAVAIERRRNPLASSRHLDELTVAAARAGLLSLEHAALMIRWLALGGRSGEAAE
 VLAAQRPRAVTDQDRAHLRAAEVSLALVSPGASGVSPGASGPDRRPRPLPPDELANLPKAARLCAIADNAVISALHG
 RPELASAEAEENVLKQADSADGATASALTALLYAENTDTAQLWADKLVSETGASNEEEGAGYAGPRAETALRRGDL
 AAAVEAGSAILDHRRGSLGITAALPLSSAVAAAIRLGETERAEKWLAEPLEAIRDSLFGLHLLSARGQYCLATGR
 HESAYTAFRTCGERMNRNWGDVPGLSLWVRVDAEALLHGRDRDEGRRLIDEQLTHAMGPRSRALTLRVQAAYSPQAQ
 40 RVDLLEEAADLLSCNDQYERARVLADLSEAFSALRHHSRARGLLRQARHLAAQCGATPLLRRLGAKPGPGWLEES
 GLPQRIKSLTDAERRVASLAAGQTNRVIAADQLFVTASTVEQHLTNVFRKLGVKGRQHLPAELANAE

SEQ ID NO: 208

MPAVKRNDLVARDGELRWMQEILSQASEGRGAVVTITGAIACGKTVLLDAAAASQDVIQLRAVCSAEEQELPYAMVG
 QLLDNPVLAARVPALGNLAAAGERLLPGTENRIRRELTRTLLALADERPVLIGVDDMHADPASLDCLLHLARRVGP
 ARIAIVLTELRRLTPAHSRFQSELLSLRYHHEIGLQPLTAHTADLARVGLGAEVDDDDVLTELYEATGGNPSLCCGL
 5 IRDVRQDWEAGVTGIHVGRAYRLAYLSSLYRCGPAALRTARAAAVLGDSADACLIRRVSGLGTEAVGQAIQQLTEGG
 LLRDQQFPHPAARSVVLDMSAQERHAMYRSAREAAAEGQADPGTPGEPRAATAYAGCGEQAGDYEPAGRACVDGA
 GPAEYCGDPHGADDDPDELVAALGGLLPSRLVAMKIRRLAVAGRPGAAAELLTSQRLHAVTSEDRA SLRAAEVALAT
 LWPGATGPD RHPLTEQEAAASLPEGPRLLAAADDAVGAALRGRAEYAAAAEAENVLRHADPAAGGDAYAAAMIALLYTEH
 PENVLFWADKLDAGRPDEETSYPLGRAETAVRLGDLETAMELGRTVLDQRRPLPSLGVAAGLLLGGA VTAAIRLGDL
 10 RAEKWLAEPIDAIRTSLYGLHVLAARGRLDLAAGRYEAA YTAFRLCGERMAGWDADV SGLALWRVDAAEALLSAGI
 RPDEGRKLIDDQLTREM GARSRALTLRAQAAYSLPVHRVGLLDEAAGLL LACHDGYERARVLADLGETLRTL RHTDA
 AQRVLRQAEQAAARCGSVPLLRRLGAEPVRIGTRRGEPGLPQRIRLLTDAERRVAAMAAAGQTNREIAGRLFVTAST
 VEQHLTSVFRKLGVKGRRLPTELAQAV

15 SEQ ID NO: 209

MYSGTCREGYELVAREDELGILQRSLEQASSGQGVVTVTGPIACGKTELLDAAAAKAEAIILRAVCAPEERAMPYA
 MIGQLIDDPALAHRA PGLADRIAQGGQLSLRAENRLRRDLTRALLALAVDRPVLIGVDDVHHADTASLNCLLHLARR
 VRPARISMIFTELRLSTPTQSRFKAELLSLPYHHEIALRPFPGPEQSAELARAAFGPGAEDVLAGLYKTTRGNLSLS
 RGLISDVREALANGESAFEAGRAFRLAYLSSLYRCGPVALRVARVAAVLGPSATTTLVRRLSGLSAETIDRATKILT
 20 EGGLLLDQQFPHPAARSVVLDMSAQERRSLHTLALELLDEAPVEVLAHHQVGAGLIHGPKAAEMFAKAGKALVVRN
 ELGDAAEYLQLAHRASDDVSTRAALRVEAVAIERRRNPLASSRHMDLSAAGRAGLLSPKHAALAVFWLADGGRSGE
 AAEVLASERPLATTDQNRHLRFVEVTLALFSPGAFGSDRRPPPLTPDELASLPKAAWQCAVADNAAMTALHGHPEL
 ATAQAETVLRQADSAADAIPAALIALLYAENTESAHIWADKLGSTNAGVSNEAEAGYAGPCA EIALRRGDLATAFEA
 GSAVLDDRSLPSLGITAALLSSKTAAAVRLGELERA EKLLAEPLNGVQDSLFGHLHLSAYGQYSLAMGRYESAHR
 25 AFRTCGERMRSWDVDVPGALALWRVDAAEALLSLDRNEGQRLIDEQLTRPMGPRSHALT LRIKAA YLPRTKRIPLLHE
 AAELLPLCPDPYEQARVLADLGD TLSALRRYSRARGVLRQARHLATQCGAVPLLRRLGGEPRIDDAGLPQRSTSLT
 DAERRVAALAAAGQTNREIAEQLFVTASTVEQHLTSVFRKLGVKGRKQLPTALADVEQT

SEQ ID NO: 210

MYSGTCREGYELVAREDELGILQRSLEEAGSGQGA VVTVTGPIACGKTELLDAAAAKADAIILRAVCAPEERAMPYA
 MIGQLIDDPALAHRAPELADRIAQGGHLSLRAENRLRRDLTRALLALAVDRPVLIGVDDVHHADTASLNCLLHLARR
 VRPARISMIFTELRLSTPTQSRFKAELLSLPYHHEIALRPLGPEQSAELAHAAFGPGAEDVLAGLYGMTRGNLSLS
 RGLISDVREAQANGESAFEVGRAFRLAYLSSLYRCGP IALRVARVAAVLGPSATTTLVRRLSGLSAETIDRATKILT
 EGGLLLDHQFPHPAARSVVLDMSAQERRSLHTLALELLDEAPVEVLAHHQVGAGLIHGPKAAE IFARAGQALVVRN
 35 ELGDAAEYLQLAHRASDDVSTRAALRVEAVAIERRRNPLASSRHMDLSAAGRAGLLSPKHAALAVFWLADGGRSGE
 AAEVLASEHPLATTDQNRHLRF AEVTLALFCPGA FGSDDRRPPPLAPDELASLPKAAWQCAVADNAVMTALHAHPEL
 ATAQAETVLRQADSAADAIPAALIALLYAENTESAQI WADKLGSTNAGVSNEAEAGYAGPCA EIALRRGDLATAFEA
 GGTVLDDRPLPSLGITAALLSSKTAAAVRLGELERA EKLLAEPLNGVQDSLFGHLHLSAHGQYSLAMGRYESAHR
 AFHTCGERMRSWGDVPGALALWRVDAAEALLSLDRNEGQRLIDEQLARPMGPRSRALT LRIKAA YLPRTKRIPLLHE
 40 AAELLSCPDYPYEQARVLADLGD TLSALRRYSRARGVLRQARHLATQCGAVPLLRRLGGEPRIDDAGLPQRSTSLT
 DAERRVSALAAAGQTNREIAKQLFVTASTVEQHLTSVFRKLGVKGRRLPTALADVE

SEQ ID NO: 211

MYSGTCREGYELVAREDELGILQRSLEQASSGQGVVVTGTPIACGKTELLDAAAAKAEAIILRAVCAPEERAMPYA
 MIGQLIDDPALAHRAPGLADRIAQGGQLSLRAENRLRRDLTRALLALAVHRPVLIGVDDVHHADTASLNCLLHLARR
 VRPARISMIFTELRLSTPTQSRFKAELLSLPYHHEIALRPFQPEQSAELARAAFGPGLAEDVLAGLYKTTRGNLSLS
 5 RGLISDVREALANGESAFEAGRAFRLAYLSSLYRCGPVALRVARVAAVLGPSATTTLVRRLSGLSAETIDRATKILT
 EGGLELDDQQFPHPAARSVVLDDMSAQERRGLHTLALELLDEAPVEVLAHHQVGAGLIHGPKAAEMFAKAGKALVVRN
 ELGDAAEYLQLAHRASDDVSTRAALRVEAVAIERRRNPLASSRHMDLSAAGRAGLLSPKHAALAVFWLADGGRSGE
 AAQVLASERPLATTDQNRHLRFVEVTLALFSPGAFGSDRRPPPLTPDELASLPKAAWQCAVADNAAMTALHGHPEL
 ATAQAETVLRQADSAADAIPAALIALLYAENTESAHIWADKLGSMNAGVSNEAEAGYAGPCAEIALRRGDLATAFEA
 10 GSTVLDDRSPLSLGITAALLSSKTAAAVRLGELERAEKLLAEPLPNGVQDSLFGHLHLSAYGQYSLAMGRYESAHR
 AFRTCGERMRSWDVDVPGALWVRVDAEALLSLDRNEGQRLIDEQLTRPMGPRSRALTTRIKAAYLPRTKRIPLLHE
 AAELLPLCPDPYEQARVLADLGDLSALRRYSRARGVLRQARHLATQCGAVPLLRRLGGEPGRIDDAGLPQRSTSLT
 DAERRVAALAAAGQTNREIAEQLFVTASTVEQHLTSVFRKLGVKGRKQLPTALADVEQT

SEQ ID NO: 212

MRAINASDTGPELVAREDELGRVRSALNRANGGQGVLSITGTPIACGKTELLEAAAASEVDAITLRAVCAAEERAIPY
 ALIGQLIDNPALGIPVPDPAGLTAQGGRLLSSAENRLRRDLTRALLTLATDRLVLICVDDVQHADNASLSCLLYLAR
 RLVPARIALVFTELRLVTSSQLRFNAELLSLRNHCEIALRPLGPGHAAELARATLGPGLSDETLTELYRVTGGNLSL
 SRGLIDDVRDAWARGETGVQVGRAFRLAYLGS LHRCGPLALRVARVAAVLGPSATSVLVRRISGLSAEAMAQATDIL
 20 ADGGLLRDQRFTHPAARSVVLDDMSAEERRSVHSLALELLDEAPAEMLAHHRVGAGLVHGPKAAETFTGAGRALAVR
 GMLGEAADYLQLAYRASGDAATKAAIRVESVAVERRRNPLVVSRLHWDELSVAARAGLLSCEHVSRTARWLTVGGRPG
 EAARVLASQHRRVVTDQDRAHLRVAEFSLALLYPGTSGSDRRPHPLTSDELAALPTATRHCAIADNAVMAALRGHPE
 LATAEAEAVLQQADAADGAALTALMALLYAESIEVAEVWADKLAAEAGASNGQDAEYAGIRAEIALRRGDLTAAVET
 AGMVLDGRPLPSLDITATLLLAGRASVAVRLGELDHAEELFAAPPEDAFQDSLFGHLHLSAHGQYSLATGRPESAYR
 25 AFRACGERMRDWGFDAPGVALWRVGAAEALLGLDRNEGRRIDEQLSRTMAPRSHALTTRIKAAYMPEPKRVDLLYE
 AAELLLSRCDQYERARVLADLGEALSALGNRYQARGVLRQARHLAMRTGADPLLRRLGIRPGRQDDPDPQPRSRSLT
 NAERRAASLAATGLTNREIADRLFTASTVEQHLTNVFRKLGVKGRKQLPAELDDME

LAL Binding Sites

In some embodiments, a gene cluster (e.g., a PKS gene cluster) includes one or more promoters
 that include one or more LAL binding sites. The LAL binding sites may include a polynucleotide
 consensus LAL binding site sequence (e.g., as described herein). In some instances, the LAL binding
 site includes a core AGGGGG (SEQ ID NO: 213) motif. In certain instances, the LAL binding site
 includes a sequence having at least 80% (e.g., 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100%)
 35 homology to SEQ ID NO: 213. The LAL binding site may include mutation sites that have been restored
 to match the sequence of a consensus or optimized LAL binding site. In some embodiments, the LAL
 binding site is a synthetic LAL binding site. In some embodiments, synthetic LAL binding sites may be
 identified by (a) providing a plurality of synthetic nucleic acids including at least eight nucleotides; (b)
 contacting one or more of the plurality of nucleotides including at least eight nucleotides with one or more
 40 LALs; (c) determining the binding affinity between a nucleic acid of step (a) and an LAL of step (b),
 wherein a synthetic nucleic acid is identified as a synthetic LAL binding site if the affinity between the

synthetic nucleic acid and an LAL is greater than X. The identified synthetic LAL binding sites may then be introduced into a host cell in a compound-producing cluster (e.g., a PKS cluster).

In some embodiments, a pair of LAL binding site and a heterologous LAL or a heterologous LAL binding site and an LAL that have increased expression compared to a natural pair may be identified by (a) providing one or more LAL binding sites; (b) contacting one or more of the LAL binding sites with one or more LALs; (c) determining the binding affinity between a LAL binding site and an LAL, wherein a pair having increased expression is identified if the affinity between the LAL binding site and the LAL is greater than the affinity between the LAL binding site and its homologous LAL and/or the LAL at its homologous LAL binding site. In some embodiments, the binding affinity between the LAL binding site and the LAL is determined by determining the expression of a protein or compound by a cell which includes both the LAL and the LAL binding site.

Constitutively active LALs

In some embodiments, the recombinant LAL is a constitutively active LAL. For example, the amino acid sequence of the LAL has been modified in such a way that it does not require the presence of an inducer compound for the altered LAL to engage its cognate binding site and activate transcription of a compound producing protein (e.g., polyketide synthase). Introduction of a constitutively active LAL to a host cell would likely result in increased expression of the compound-producing protein (e.g., polyketide synthase) and, in turn, increased production of the corresponding compound (e.g., polyketide).

Engineering Unidirectional LALs

FK gene clusters are arranged with a multicistronic architecture driven by multiple bidirectional promoter-operators that harbor conserved (in single or multiple, and inverted to each other and/or directly repeating) GGGGGT (SEQ ID NO: 179) motifs presumed to be LAL binding sites. Bidirectional LAL promoters may be converted to unidirectional ones (UniLALs) by strategically deleting one of the opposing promoters, but maintaining the tandem LAL binding sites (in case binding of LALs in the native promoter is cooperative, as was demonstrated for MalT). Functionally this is achieved by removal of all sequences 3' of the conserved GGGGGT (SEQ ID NO: 179) motif present on the antisense strand (likely containing the -35 and -10 promoter sequences), but leaving intact the entire sequence on the sense strand. As a consequence of this deletion, transcription would be activated in one direction only. The advantages of this feed-forward circuit architecture would be to tune and/or maximize LAL expression during the complex life cycle of *Streptomyces* vegetative and fermentation growth conditions.

Host cells

In some embodiments, the host cell is a bacteria such as an Actinobacterium. For example, in some embodiments, the host cell is a *Streptomyces* strain. In some embodiments, the host cell is *Streptomyces anulatus*, *Streptomyces antibioticus*, *Streptomyces coelicolor*, *Streptomyces peucetius*, *Streptomyces sp. ATCC 700974*, *Streptomyces canus*, *Streptomyces nodosus*, *Streptomyces (multiple sp.)*, *Streptoalloteicus hindustanus*, *Streptomyces hygroscopicus*, *Streptomyces avermitilis*, *Streptomyces viridochromogenes*, *Streptomyces verticillus*, *Streptomyces chartruensis*, *Streptomyces (multiple sp.)*, *Saccharothrix mutabilis*, *Streptomyces halstedii*, *Streptomyces clavuligerus*, *Streptomyces venezuelae*,

Streptomyces roseochromogenes, *Amycolatopsis orientalis*, *Streptomyces clavuligerus*, *Streptomyces rishiriensis*, *Streptomyces lavendulae*, *Streptomyces roseosporus*, *Nonomuraea* sp., *Streptomyces peucetius*, *Saccharopolyspora erythraea*, *Streptomyces filipinensis*, *Streptomyces hygroscopicus*, *Micromonospora purpurea*, *Streptomyces hygroscopicus*, *Streptomyces narbonensis*, *Streptomyces kanamyceticus*, *Streptomyces collinus*, *Streptomyces lasaliensis*, *Streptomyces lincolnensis*, *Dactyloporangium aurantiacum*, *Streptomyces toxitricini*, *Streptomyces hygroscopicus*, *Streptomyces plicatus*, *Streptomyces lavendulae*, *Streptomyces ghanaensis*, *Streptomyces cinnamomensis*, *Streptomyces aureofaciens*, *Streptomyces natalensis*, *Streptomyces chattanoogensis* L10, *Streptomyces lydicus* A02, *Streptomyces fradiae*, *Streptomyces ambofaciens*, *Streptomyces tendae*, *Streptomyces noursei*, *Streptomyces avermitilis*, *Streptomyces rimosus*, *Streptomyces wedmorensis*, *Streptomyces cacaoi*, *Streptomyces pristinaespiralis*, *Streptomyces pristinaespiralis*, *Actinoplanes* sp. ATCC 33076, *Streptomyces hygroscopicus*, *Lechevalieria aerocolonigenes*, *Amycolatopsis mediterranei*, *Amycolatopsis lurida*, *Streptomyces albus*, *Streptomyces griseolus*, *Streptomyces spectabilis*, *Saccharopolyspora spinosa*, *Streptomyces ambofaciens*, *Streptomyces staurosporeus*, *Streptomyces griseus*, *Streptomyces* (multiple species), *Streptomyces acromogenes*, *Streptomyces tsukubaensis*, *Actinoplanes teichomyceticus*, *Streptomyces glaucescens*, *Streptomyces rimosus*, *Streptomyces cattleya*, *Streptomyces azureus*, *Streptoalloteicus hindustanus*, *Streptomyces chartreusis*, *Streptomyces fradiae*, *Streptomyces coelicolor*, *Streptomyces hygroscopicus*, *Streptomyces* sp. 11861, *Streptomyces virginiae*, *Amycolatopsis japonicum*, *Amycolatopsis balhimycini*, *Streptomyces albus* J1074, *Streptomyces coelicolor* M1146, *Streptomyces lividans*, *Streptomyces incarnates*, *Streptomyces violaceoruber*, or *Streptomyces griseofuscus*. In some embodiments, the host cell is an *Escherichia* strain such as *Escherichia coli*. In some embodiments, the host cell is a *Bacillus* strain such as *Bacillus subtilis*. In some embodiments, the host cell is a *Pseudomonas* strain such as *Pseudomonas putida*. In some embodiments, the host cell is a *Myxococcus* strain such as *Myxococcus xanthus*.

Examples

Example 1. Single module swapping to produce an engineered PKS

Inter-module residue covariation analysis and evolutionary trace analysis were used to predict 10 heterologous donor modules that would successfully replace module 3 of the PKS that produces Compound 1 (FIG. 3A). Seven of the 10 predicted donor modules, ranging in length from 4-6kb, were selectively amplified in their entirety using a GC-rich long PCR method. In parallel, a bacterial artificial chromosome (BAC) that harbored the PKS that produces Compound 1 was converted to a module swap acceptor for heterologous donor modules by introducing the restriction sites *Afl*III and *Spe*I to the flanking intermodule sequence of module 3. The modified acceptor BAC was linearized by digestion with *Afl*III and *Spe*I, and the 7 donor modules were gel-purified and subcloned by Gibson cloning. The resulting constructs were subjected to Sanger sequencing of region of interest, PCR-based analysis to confirm cluster integrity, and Illumina NGS to sequence the entire BAC. The PCR-mediated error rate of the module amplification protocol was determined to be approximately 1bp per 5000bp, or approximately 1 mutation per module.

A single module was swapped to produce an engineered PKS by replacing module 3 of the PKS that produces Compound 1 with module 3 of *Streptomyces* strain S317. The donor S317 module 3 was

PCR amplified and Gibson cloned into position 3 of the PKS that produces Compound 1 (FIG. 3B). The resulting clone was conjugated into a *Streptomyces* expression host and fermented. Production of compound was analyzed by LC-TOF mass spectrometry analysis by co-injecting purified native FKBP12, the protein to which both compounds are expected to bind, with either the product of the native PKS, Compound 1, or the compound produced by the engineered PKS cluster, Compound 2. Comparative LC-TOF analysis of indicated that Compound 2 had the expected mass of 611.38, corresponding to the conversion of the module 3 alkene to a fully reduced module at that position. Compound 2 was re-fermented at large scale, purified to homogeneity and the structure was confirmed by NMR spectroscopy.

To replace module 4 in the PKS that produces Compound 1, module swapping prediction algorithms based on inter-module covariation were used to generate a list of 16 modules encoding 4 chemistries. Gibson-based subcloning into module 4 was not as efficient as module 3. Gibson cloning, which involves a ssDNA intermediate, is difficult in high GC-rich regions, and direct ligation of donor modules to restriction sites with 4bp overhangs may not be sensitive to local GC content. Therefore *Afl*I and *Spe*I sites were introduced at new positions in the inter-module flanking region to generate a direct ligation acceptor BAC. This direct ligation acceptor BAC was linearized by digestion with *Afl*I and *Spe*I, and 12 donor modules were gel-purified, digested with *Afl*I and *Xba*I and subcloned by ligation.

Single module swaps of either module 3 or module 4 in the PKS that produces Compound 1 generated novel Compounds 2-5 (FIG. 3C). Therefore, single module swapping was used to introduce a range of module encoded chemistries and generate novel compounds. LC-TOF mass spectrometry analysis indicated that of the module swaps at module 3 and module 4, the resulting hybrid clusters yielded a range of compound expression.

Example 2. Library construction by combinatorial dimodule swapping

Pooled transfer of dimodule libraries was used to simultaneously replace modules 3 and 4 in the PKS that produces Compound 1 and generate a plurality of engineered PKS clusters (FIG. 4A). A total of 31 modules were amplified for transfer to the module 3 position and 25 modules for the module 4 position. To optimize Gibson dimodule assembly cloning, phosphothiorate-modified DNA oligos were synthesized for PCR amplification of the donor modules. Phosphothiorate-capped module ends function by constraining the exonuclease step of the Gibson cloning protocol, which resulted in a dramatic increase in Gibson capture of GC-rich DNA (FIG. 4B). An intermediate plasmid-based dimodule capture protocol was developed to assemble, capture, amplify, and enrich the dimodule units (FIG. 4C). Pooled module 3 and module 4 amplicons were mixed with a linear backbone amplicon based on pBR322 for a 3-part Gibson assembly reaction. Shuttle vectors containing dimodule assemblies could be resolved from empty vector by fractionating on a preparative 0.4% agarose gel. After dimodule capture, the assembled dimodule fragments were released from the shuttle vector by digestion with *Afl*I and *Xba*I and subcloned by direct ligation to an expression vector containing the PKS that produces Compound 1, in which the PKS lacked the native module 3 and module 4.

Replicate BACs encoding single module and dimodule swaps were conjugated to optimized *Streptomyces* producer strain S2441 and solid-phase extracted samples were subjected to LC-TOF mass spectrometry with the expected protein binding partner, purified FKBP12 protein. Further analysis confirmed that dimodule library generation is capable of engineering PKS clusters that express novel

compounds in high yield (FIG. 4D). As a representative example, Compound 6 was generated by dimodule swapping of a module encoding mDEK chemistry at module 3 and K chemistry at module 4 of the PKS that produced Compound 1. The expected mass of Compound 6 was observed by LC-TOF analysis, confirming that the dimodule assembly protocol yields engineered derivatives Compound 1.

5 A 650-member combinatorial library of engineered derivatives of the PKS that produces Compound 1 was produced by dimodule swapping. A total of 31 modules were amplified for transfer the module 3 position and 25 modules for the module 4 position of the PKS that produces Compound 1 (FIG. 4E). Clusters were cloned onto BACs, and the cloned BACs were subsequently used as templates to PCR modules of diverse sources from multiple heterologous donors.

10 A subset of the library corresponding to 15 different donor modules at the module 3 position and 15 different donor modules at the module 4 position produced a potential combinatorial library of 225 novel PKS clusters and resulting novel compounds (the 15x15 dimodule library). Because the dimodule library was assembled as a pool, rarefaction analysis was performed to determine how many clones needed to be conjugated, fermented, and extracted to effectively sample >90% of the diversity of the
15 library. Rarefaction analysis indicated that 650 clones corresponded to a statistical sampling >90% of the dimodule library (FIG. 4F). 650 clones were prosecuted and subjected to LC-TOF mass spectrometry analysis. 115 of the 650 sampled clones expressed compounds with novel masses.

20 **Example 3. Characterization of a combinatorial dimodule library by single-molecule long-read sequencing**

A library corresponding to 15 different donor modules at the module 3 position and 15 different donor modules at the module 4 position (the 15x15 dimodule library), produced according to the methods of Example 2, was characterized by Nanopore sequencing (FIG. 4G). The dimodules present in the 15x15 dimodule library were excised from the PKS clusters using CRISPR/Cas9 (NEB). The resulting
25 excised dimodules each had a length of approximately 7-12 kilobases. The dimodules were purified by 96-well column purification, and well-specific adaptors were ligated to the dimodules. The resulting dimodules were normalized and pooled and prepared for sequencing according to the standard ligation preparation protocol for Nanopore sequencing of oligonucleotides. Nine 96-well plates (864 dimodule clones total) were sequenced by Nanopore and the resulting sequencing data was analyzed according to
30 the informatics workflow provided in FIG. 4H, with 73.1% of clones being called. The comparison of the resulting sequencing data against the table of input of the donor modules allows the deconvolution of the resulting combinatorial library by identification of the resulting dimodules. The results of Nanopore sequencing of the 15x15 dimodule library are provided in Table 1.

Table 1.

Library Plate IDs	NoCall	Ambiguous	Single Read	Called	Grand Total
163846	45	4	11	36	96
163848	14	10	14	58	96
163851		16		80	96
163896	5	8	5	78	96
163897		21	1	74	96
163898	3	10	11	72	96
163899	4	6	2	84	96
163900	1	26	3	66	96
50066321		12		84	96
Grand Total	72	113	47	632	864
%	8.3%	13.1%	5.4%	73.1%	

Example 4. Library construction by combinatorial trimodule swapping

The combinatorial module swap protocols were modified to generate trimodule assemblies in the PKS that produces Compound 7 (FIG. 5A). Increasing the number of module swaps increases the size, and therefore diversity, of a PKS library. For example, given a collection of 13 different module-encoded chemistries, increasing size and diversity is based on the number of modules that are swapped such that the maximal library size of a single mod swap is 13; with a dimodule swap the maximal library size is $13^2=169$; and for a trimodule swap, the maximal library is $13^3 = 2197$.

Trimodule assembly leverages the technical advances of the dimodule protocol with an additional “proof-reading” Gibson cloning step to insert the captured trimodule assembly into the PKS that produces Compound 7 (FIG. 5B). As before, phosphorothioate chemistry was used to constrain the ssDNA intermediate for the first round of Gibson cloning into a shuttle vector. Shuttle vector clones harboring trimodule assemblies were enriched by preparative gel fractionation and isolation. Finally, Gibson-mediated “error correction” was used to trim restriction sites for scarless cloning in the expression vector. First, flanking PmeI restriction sites were introduced within the linker regions between Module 3 and Module 4, as well as between Module 6 and Module 7. Sites with reduced GC content and secondary structure (as predicted by DNAfold; <8 kcal/ml) were selected for optimal Gibson homology arms. A Gibson Assembly Ultra Kit (SGI-DNA) was used to clone the trimodule assembly into the PKS that produces Compound 7 enabling the replacement of Modules 4, 5, and 6 and simultaneously removal of the additional extraneous PmeI sequence retained after digestion. This resulted in >95% correct assembly for the industrial scale production of compounds produced by trimodule swapped PKS clusters (>200 per week).

Example 5. Ring expansion by swapping a single module acceptor with a dimodule donor

A heterologous dimodule donor assembly encoding mDEK chemistry and K chemistry was swapped into module 3, a single module acceptor, of the PKS that produces Compound 1 by the methods described above (FIG. 6A). The compound produced by engineered PKS, Compound 8, was observed in

high yield and had a mass of 655.41, as determined by LC-TOF analysis (FIG. 6B). This corresponds to a ring-expanded compound product in which Compound 8 contains an additional 2-carbon extender unit. Thus reprogramming PKS biosynthesis via module swapping by insertion of a dimodule assembly to replace a single module may produce functional PKS expression.

5

Example 6. Module swapping of a PKS loading module

Rapamycin is a natural product synthesized by a mixed polyketide synthase (PKS)/nonribosomal peptide synthetase (NRPS) system. Rapamycin shares a common structural motif with related natural product FK506 which is responsible for binding to FK506-binding proteins (FKBPs). During biogenesis of Rapamycin, loading modules bind and load a 4,5-dihydroxycyclohexa-1,5-dienecarboxylic acid starter unit via a CaiC domain, which functions as a carboxylic acid ligase (CL) like domain (FIG. 7A). Loading modules may possess similar domain structure as conventional elongation PKS modules, including ketoreductase-like domains and an enoyl-reductase domain, which may or may not be catalytically active. The final chemistry of the starter unit depends on the presence and the sequence of the domains in the loading module, so the resulting "starter unit" can be engineered by swapping the loading module

15

The X23 PKS cluster produces Compound 9 and Compound 10 (FIG. 7B). The Rapamycin loading module from *Streptomyces* strain S303 was swapped into the X23 cluster by the methods described previously for a single module swap. The engineered PKS produced Compounds 11 and 12, in which the starter unit is replaced with the starter unit of Rapamycin. Additional single elongation module swaps of Module 2 and Module 7 of X23 produced Compounds 13 and 14, respectively.

20

Other Embodiments

It is to be understood that while the present disclosure has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the present disclosure, which is defined by the scope of the appended claims. Other aspects, advantages, and alterations are within the scope of the following claims.

25

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments in accordance with the invention described herein. The scope of the present invention is not intended to be limited to the above Description, but rather is as set forth in the appended claims.

30

In the claims, articles such as "a," "an," and "the" may mean one or more than one unless indicated to the contrary or otherwise evident from the context. Claims or descriptions that include "or" between one or more members of a group are considered satisfied if one, more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process unless indicated to the contrary or otherwise evident from the context. The invention includes embodiments in which exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The invention includes embodiments in which more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process.

35

It is also noted that the term "comprising" is intended to be open and permits but does not require the inclusion of additional elements or steps. When the term "comprising" is used herein, the term "consisting of" is thus also encompassed and disclosed.

40

Where ranges are given, endpoints are included. Furthermore, it is to be understood that unless otherwise indicated or otherwise evident from the context and understanding of one of ordinary skill in the art, values that are expressed as ranges can assume any specific value or subrange within the stated ranges in different embodiments of the invention, to the tenth of the unit of the lower limit of the range, unless the context clearly dictates otherwise.

In addition, it is to be understood that any particular embodiment of the present invention that falls within the prior art may be explicitly excluded from any one or more of the claims. Since such embodiments are deemed to be known to one of ordinary skill in the art, they may be excluded even if the exclusion is not set forth explicitly herein. Any particular embodiment of the compositions of the invention (e.g., any polynucleotide or protein encoded thereby; any method of production; any method of use) can be excluded from any one or more claims, for any reason, whether or not related to the existence of prior art.

What is claimed is:

CLAIMS

1. An engineered polyketide synthase comprising one or more heterologous modules with altered enzymatic activity relative to a reference polyketide, wherein the engineered polyketide synthase is capable of producing a polyketide when expressed under conditions suitable to allow expression of a compound by the engineered polyketide synthase and wherein the one or more heterologous modules do not substantially inhibit polyketide translocation during polyketide biosynthesis.
2. An engineered polyketide synthase comprising one or more heterologous modules with altered enzymatic activity relative to a reference polyketide, wherein the engineered polyketide synthase is capable of producing a polyketide when expressed under conditions suitable to allow expression of a compound by the engineered polyketide synthase and wherein the one or more heterologous modules comprise linking sequences which are compatible to the linking sequences of the modules adjacent thereto.
3. An engineered polyketide synthase comprising one or more heterologous modules with altered enzymatic activity relative to a reference polyketide, wherein the engineered polyketide synthase is capable of producing a polyketide when expressed under conditions suitable to allow expression of a compound by the engineered polyketide synthase and wherein the polyketide expression level of the engineered polyketide synthase is at least 1% of the polyketide expression level of the reference polyketide synthase.
4. The engineered polyketide synthase of any one of claims 1 to 3, wherein the one or more heterologous modules comprise native linking sequences.
5. The engineered polyketide synthase of any one of claims 1 to 4, wherein the engineered polyketide synthase comprises two or more heterologous modules.
6. The engineered polyketide synthase of claim 5, wherein the two or more heterologous modules are adjacent.
7. The engineered polyketide synthase of any one of claims 1 to 6, wherein the engineered polyketide synthase comprises three or more heterologous modules.
8. The engineered polyketide synthase of claim 7, wherein the three or more heterologous modules are adjacent.
9. The engineered polyketide synthase of any one of claims 1 to 8, wherein the heterologous module is an elongation module which modifies a β -carbonyl unit in the variable region of the polyketide.

10. The engineered polyketide synthase of any one of claims 1 to 9, wherein at least one of the one or more heterologous modules comprises a portion having at least 90% identity to any one of SEQ ID NO: 1-174.

11. The engineered polyketide synthase of any one of claims 1 to 10, wherein at least one of the one or more heterologous modules comprises a portion having the sequence of any one of SEQ ID NO: 1-174.

12. A chimeric polyketide synthase, wherein at least one module of the chimeric polyketide synthase has been modified as compared to a polyketide synthase having the sequence of SEQ ID NO: 175-176.

13. The chimeric polyketide synthase of claim 12, wherein the at least one module comprises a portion having at least 90% identity to any one of SEQ ID NO: 1-174.

14. A nucleic acid encoding a polyketide synthase of any one of claims 1 to 13.

15. The nucleic acid of claim 15, wherein the nucleic acid further encodes an LAL, wherein the sequence encoding the LAL is operatively linked to the sequence encoding the polyketide synthase.

16. The nucleic acid of claim 15, wherein the LAL is a heterologous LAL.

17. The nucleic acid of claim 15 or 16, wherein LAL comprises a portion having at least 80% identity to SEQ ID NO: 177.

18. The nucleic acid of claim 17, wherein the LAL comprises a portion having the sequence of SEQ ID NO: 177.

19. The nucleic acid of claim 18, wherein the LAL has the sequence of SEQ ID NO: 177.

20. The nucleic acid of any one of claims 14 to 19, wherein the nucleic acid encoding the LAL lacks a TTA inhibitory codon in an open reading frame.

21. The nucleic acid of any one of claims 14 to 20, wherein the nucleic acid further comprises an LAL binding site, wherein the sequence encoding the LAL binding site is operatively linked to the sequence encoding the polyketide synthase.

22. The nucleic acid of claim 21, wherein the LAL binding site comprises a portion having at least 80% sequence identity to the sequence of SEQ ID NO: 178.

23. The nucleic acid of claim 22, wherein the LAL binding site comprises a portion having the sequence of SEQ ID NO: 178.

24. The nucleic acid of claim 23, wherein the LAL binding site has of the sequence of SEQ ID NO: 178.

25. The nucleic acid of claim 21, wherein the LAL binding site has the sequence GGGGGT (SEQ ID NO: 179).

26. The nucleic acid of any one of claims 21 to 25, wherein the binding of an LAL to the LAL binding site promotes expression of the polyketide synthase.

27. The nucleic acid of any one of claims 14 to 26, wherein the nucleic acid further encodes a nonribosomal peptide synthase.

28. The nucleic acid of any one of claims 14 to 27, wherein the nucleic acid further encodes a first P450 enzyme.

29. The nucleic acid of claim 28, wherein the nucleic acid further encodes a second P450 enzyme.

30. An expression vector comprising a nucleic acid of any one of claims 14 to 29.

31. The expression vector of claim 30, wherein the expression vector is an artificial chromosome.

32. The expression vector of claim 31, wherein the artificial chromosome is a bacterial artificial chromosome.

33. A host cell comprising an expression vector of any one of claims 30 to 32.

34. A host cell comprising a polyketide synthase of any one of claims 1 to 13, wherein the polyketide is heterologous to the host cell.

35. The host cell of claim 33 or 34, wherein the host cell naturally lacks an LAL.

36. The host cell of any one of claims 33 to 35, wherein the host cell naturally lacks an LAL binding site.

37. The host cell of any one of claims 33 to 36, wherein the host cell comprises an LAL capable of binding to an LAL binding site and regulating expression of a polyketide synthase.

38. The host cell of claim 37, wherein the LAL is heterologous.
39. The host cell of claim 37 or 38, wherein the LAL comprises a portion having at least 80% identity to the sequence of SEQ ID NO: 177.
40. The host cell of any one of claims 33 to 39, wherein the host cell is a bacterium.
41. The host cell of claim 40, wherein the bacterium is an actinobacterium.
42. The host cell of claim 41, wherein the actinobacterium is *Streptomyces ambofaciens*, *Streptomyces hygroscopicus*, or *Streptomyces malayensis*.
43. The host cell of claim 42, wherein the actinobacterium is S1391, S1496, or S2441.
44. The host cell of any one of claims 33 to 43, wherein the host cell has been modified to enhance expression of a polyketide synthase.
45. The host cell of claim 44, wherein the host cell has been modified to enhance expression of a compound-producing protein by (i) deletion of an endogenous gene cluster which expresses a compound-producing protein; (ii) insertion of a heterologous gene cluster which expresses a compound-producing protein; (iii) exposure of the host cell to an antibiotic challenge; and/or (iv) introduction of a heterologous promoter that results in an at least 2-fold increase in expression of a compound compared to the homologous promoter.
46. A method of producing a polyketide, the method comprising culturing a host cell of any one of claims 33 to 45 under suitable conditions.
47. A method of producing a polyketide, the method comprising culturing a host cell engineered to express a polyketide synthase of any one of claims 1 to 13 under conditions suitable for polyketide synthase to produce a polyketide.
48. A method of producing a compound, the method comprising:
- a) providing a parent polyketide synthase sequence capable of producing a compound;
 - (b) determining the compatibility of at least one module of a second polyketide synthase with at least two modules of the parent polyketide synthase;
 - (c) producing a nucleic acid encoding a modified polyketide synthase, wherein the modified polyketide synthase comprises at least one module of a second polyketide synthase which has been determined to be compatible with the at least two modules of the parent polyketide synthase.
49. A method of producing a compound, the method comprising:
- (a) providing a parent nucleic acid encoding a parent polyketide synthase;

(b) modifying the parent nucleic acid to create a modified nucleic acid encoding a modified polyketide synthase capable of producing a compound, wherein the modification produces a modified polyketide synthase comprising at least one heterologous module.

50. A method of producing a compound, the method comprising:

- (a) providing a parent polynucleotide sequence capable of producing a compound;
- (b) identifying one or more heterologous modules suitable for replacement of one or more modules in the parent polynucleotide sequence;
- (c) producing a nucleic acid encoding a modified polyketide synthase, wherein the modified polyketide synthase comprises at least one heterologous module identified in step (b).

51. A method of producing a plurality of polynucleotides, wherein each of the plurality of polynucleotides corresponds to an engineered polyketide synthase, and wherein each of the plurality of polynucleotides comprises one or more heterologous modules with altered enzymatic activity relative to a reference polyketide, wherein the method comprises:

- (a) providing a parent polynucleotide sequence encoding a polyketide synthase;
- (b) identifying one or more modules for replacement in the parent polynucleotide sequence;
- (c) identifying two or more heterologous modules suitable for replacement for each of the modules identified in step (b);
- (d) generating a plurality of polynucleotides, wherein each of the plurality of polynucleotides corresponds to an engineered polyketide synthase, and wherein each of the plurality of polynucleotides comprises a heterologous module selected from the two or more heterologous modules identified in step (c) in replacement of each of the one or more modules to be replaced identified in step (b).

FIG. 1A

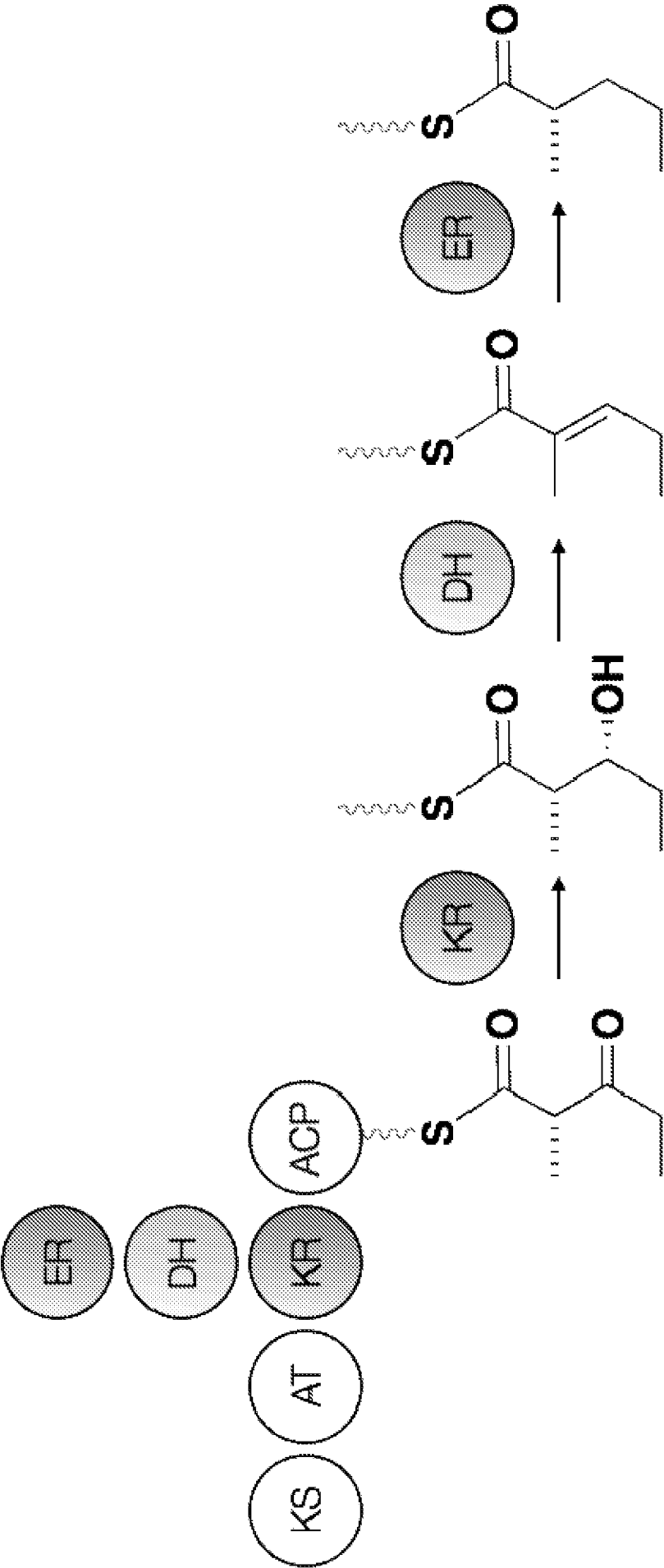


FIG. 1B

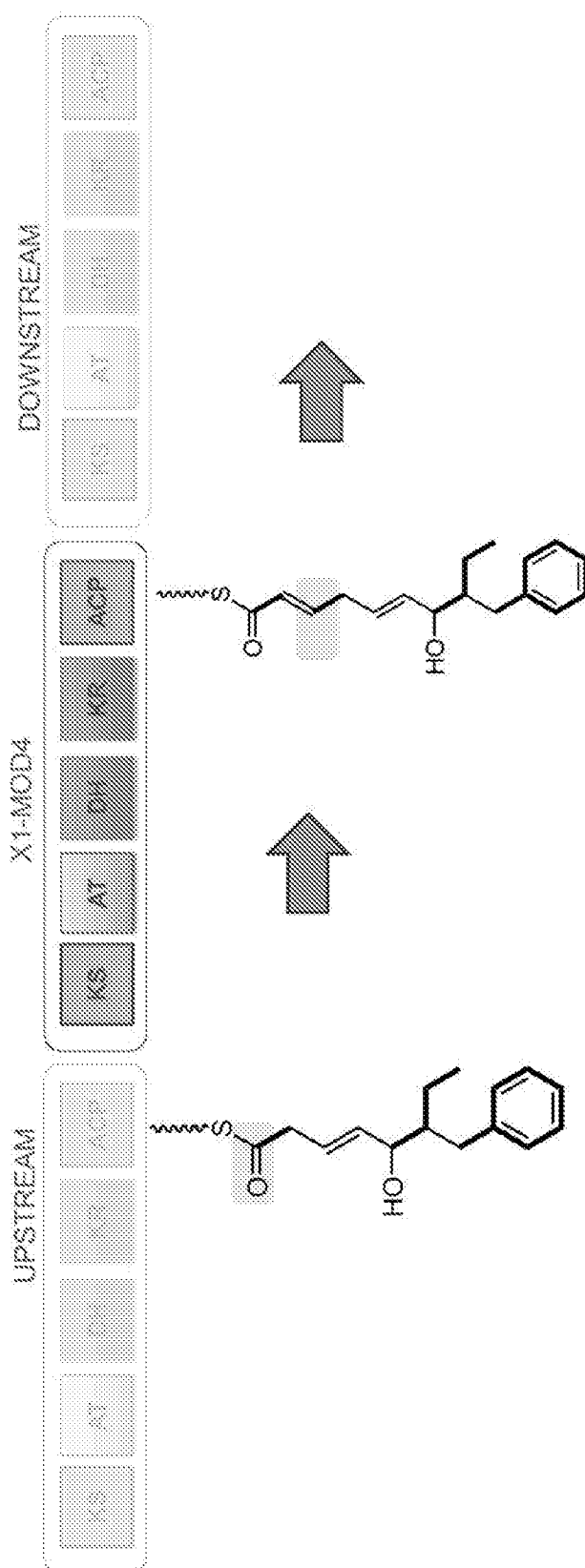


FIG. 2A

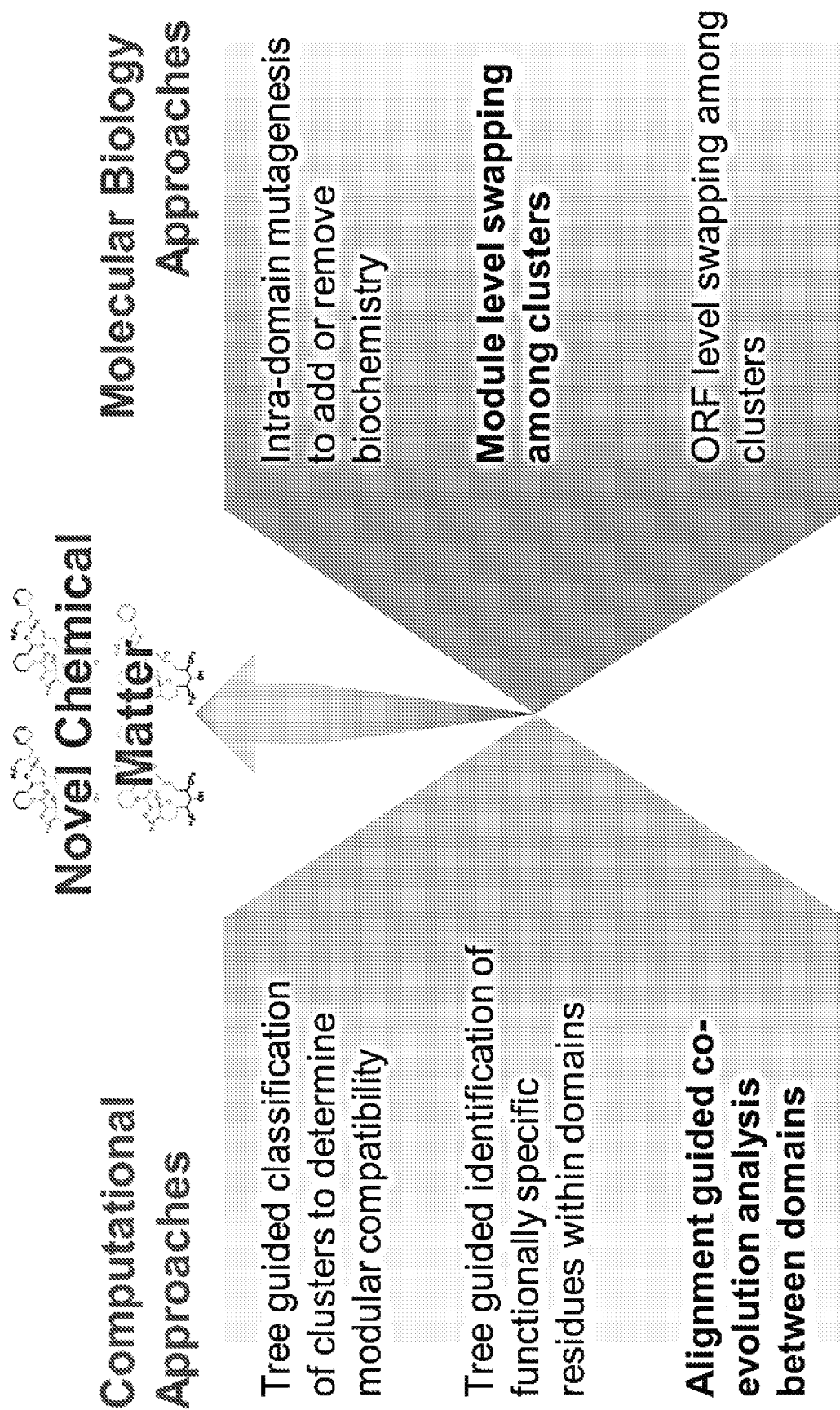


FIG. 2B

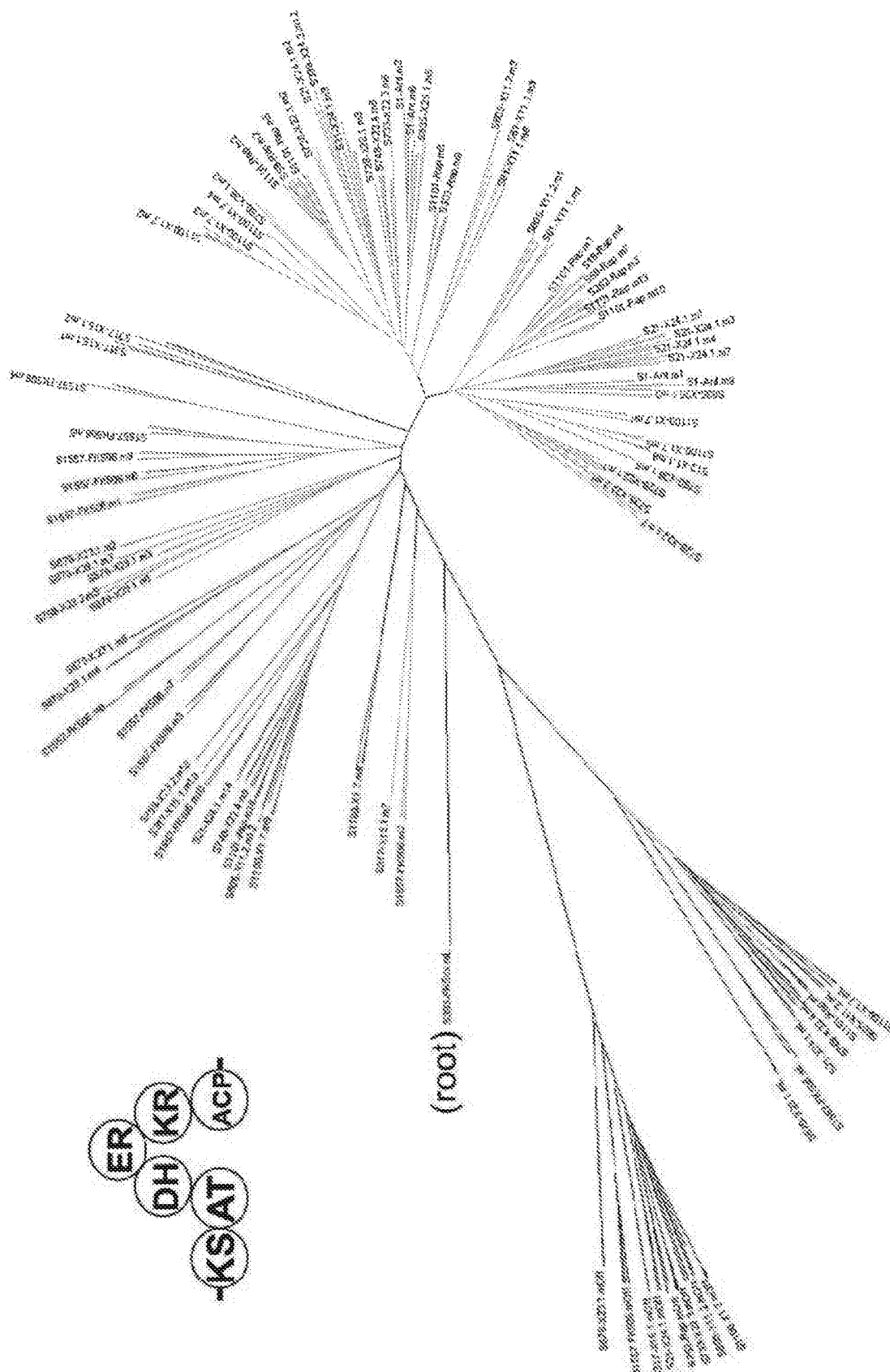
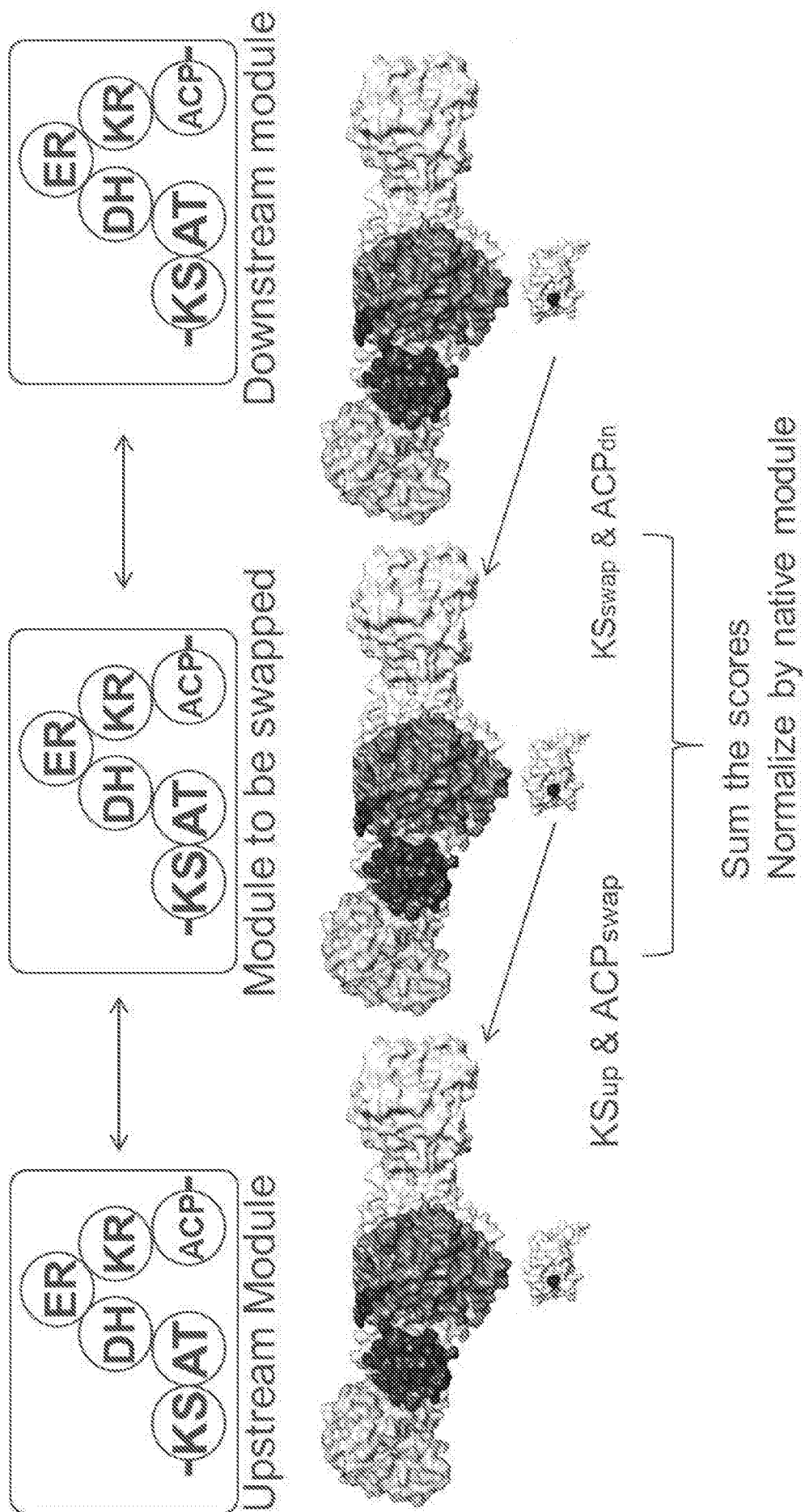


FIG. 2C



Correlation Score

X High
+ Medium
★ Low

FIG. 2D

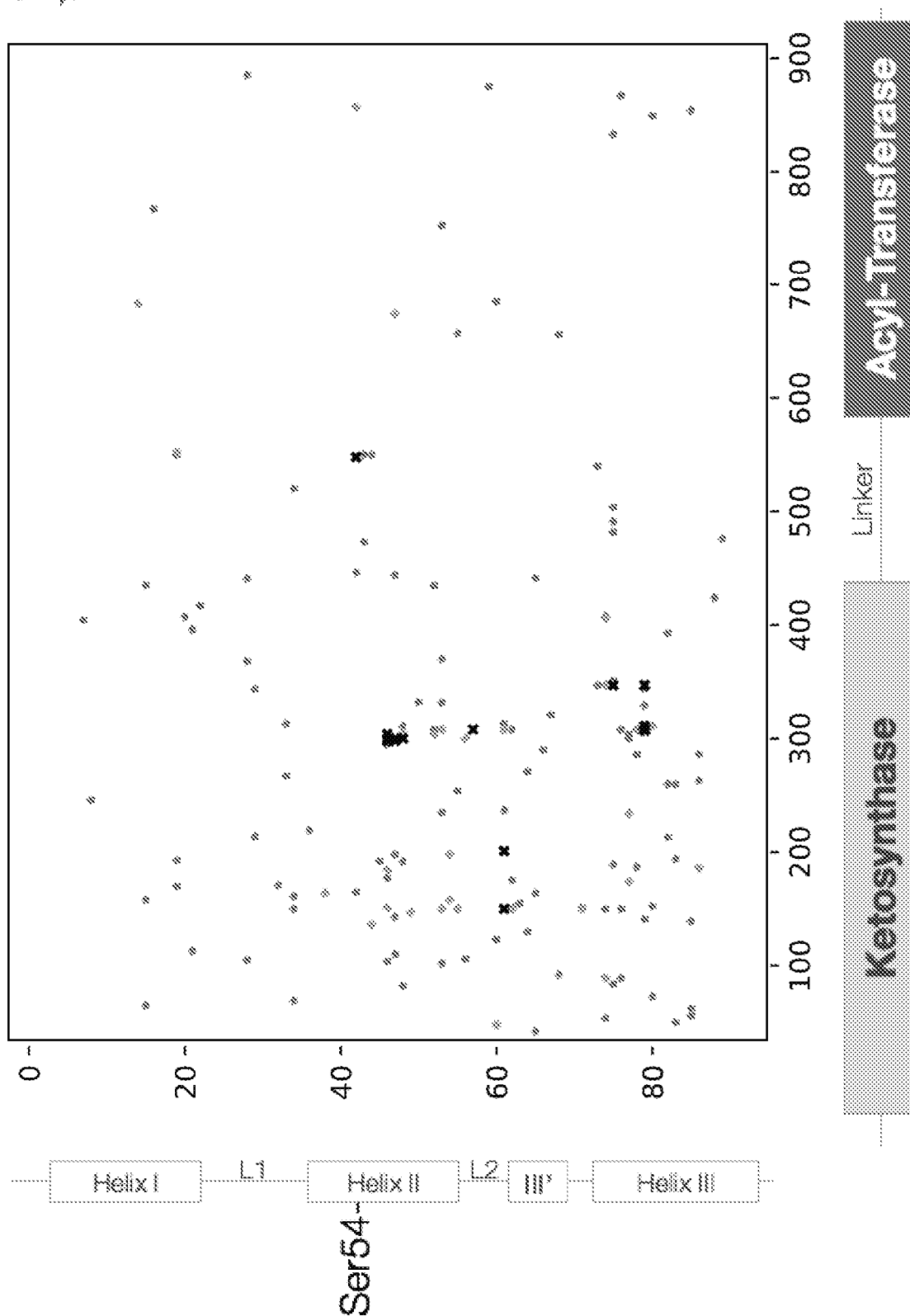
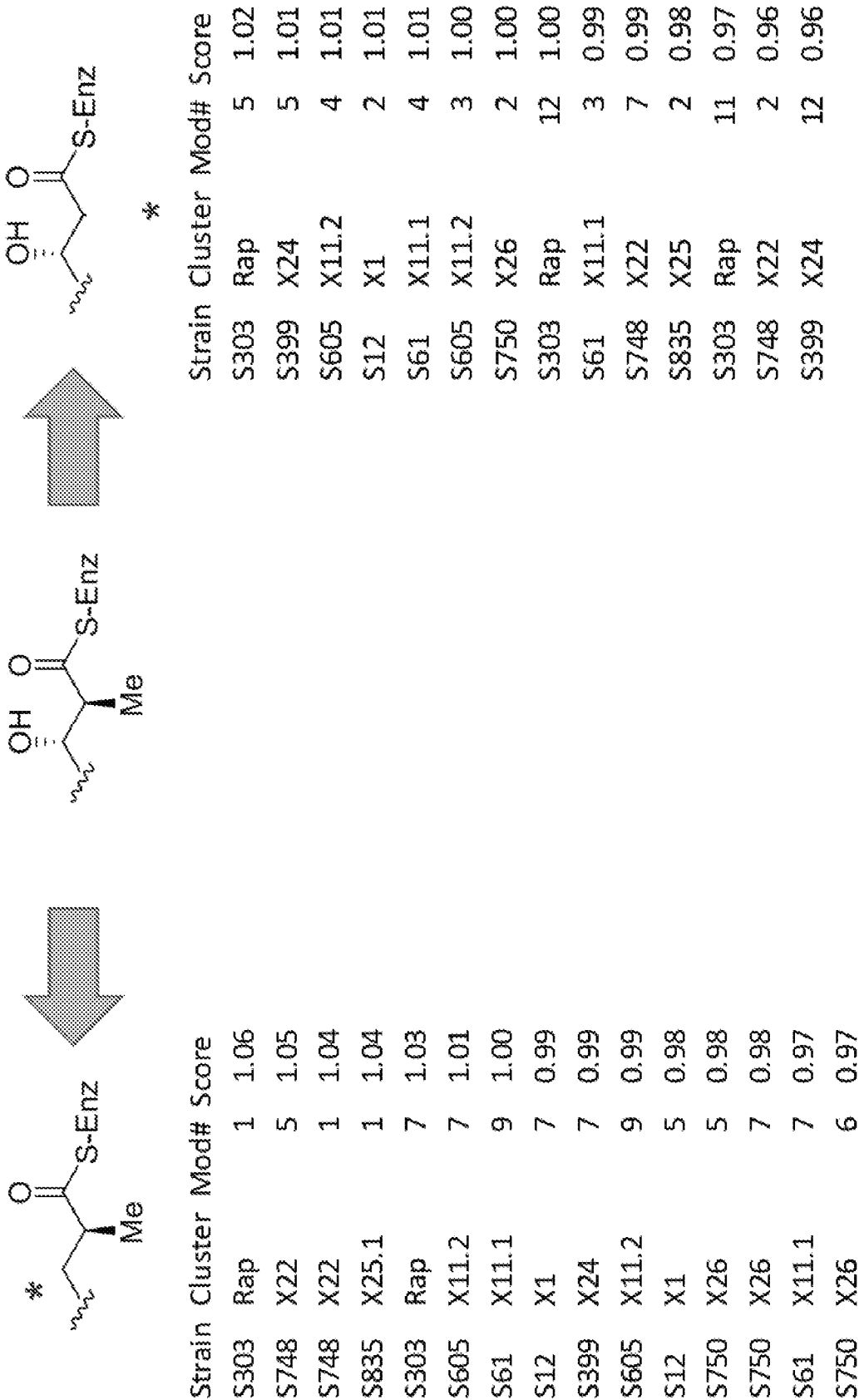


FIG. 2E



Score >0.95 from clusters available for plucking

FIG. 2F

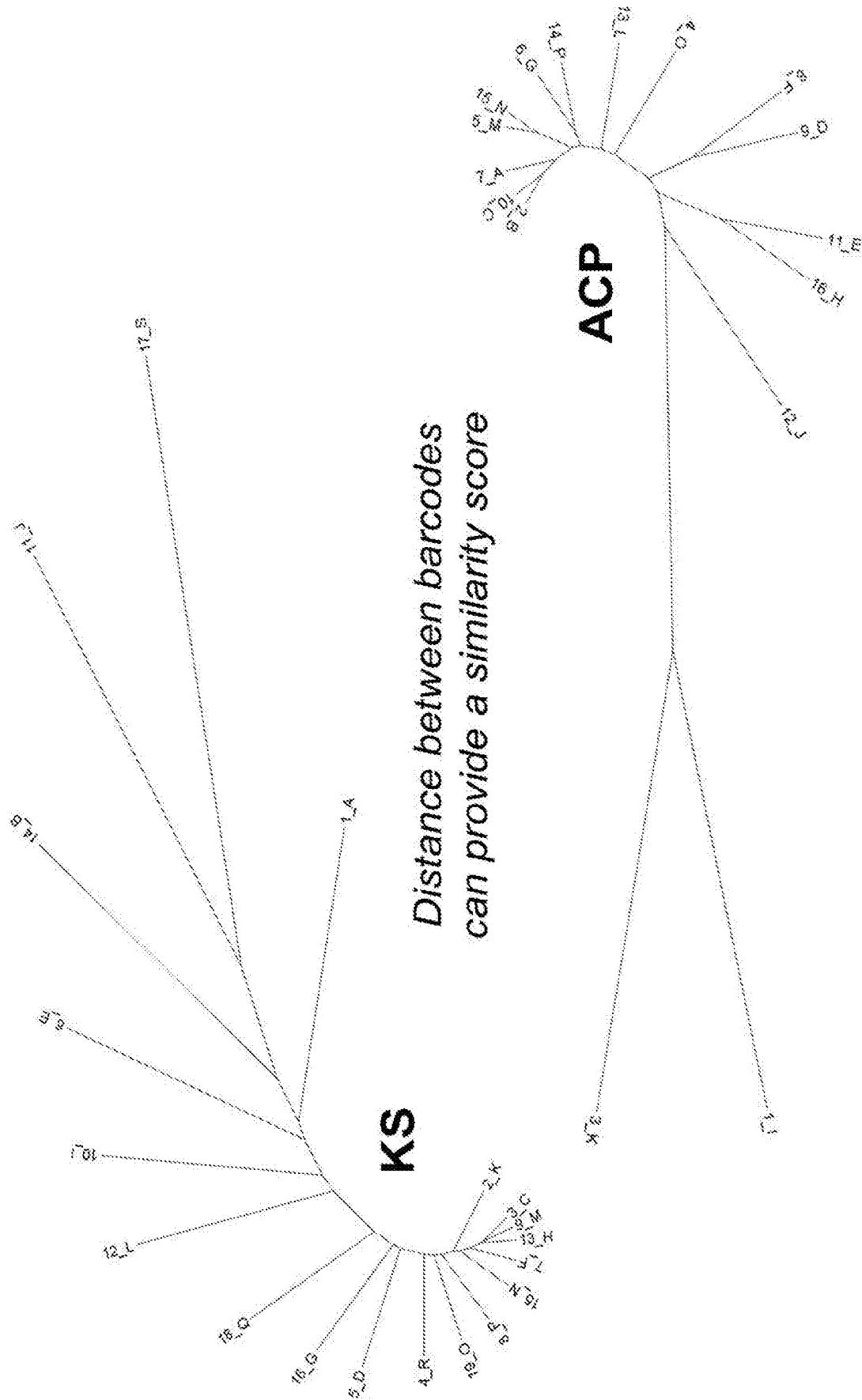


FIG. 3A

Introduce Unique Restriction Sites to BAC

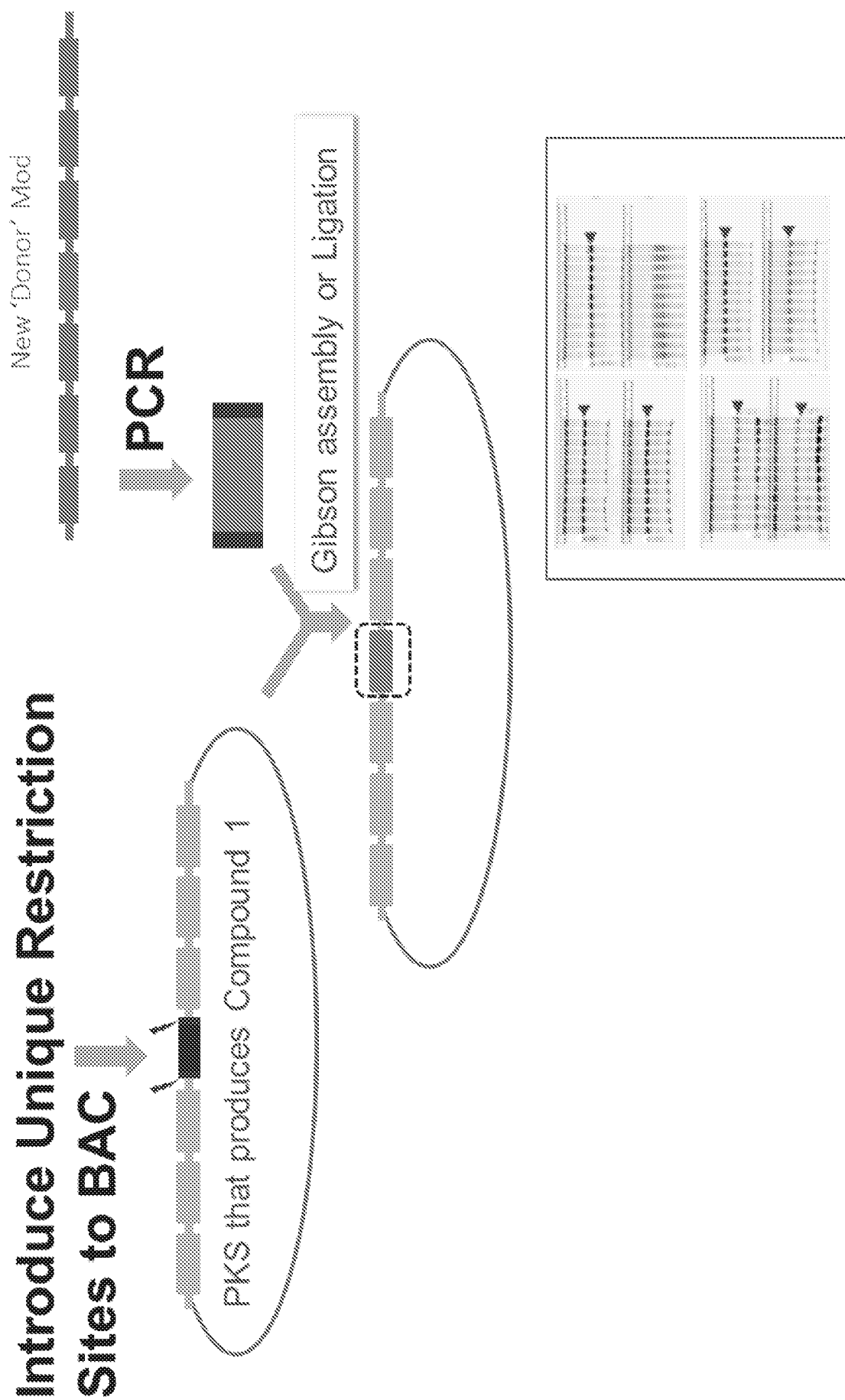
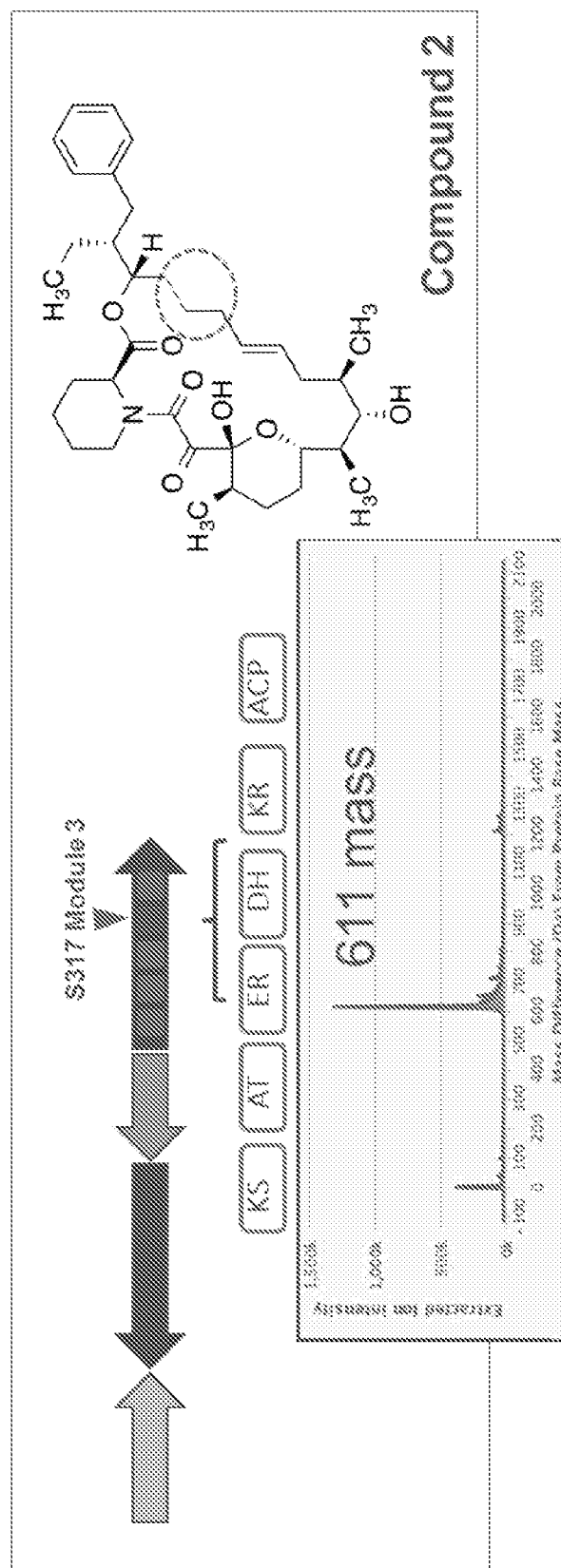
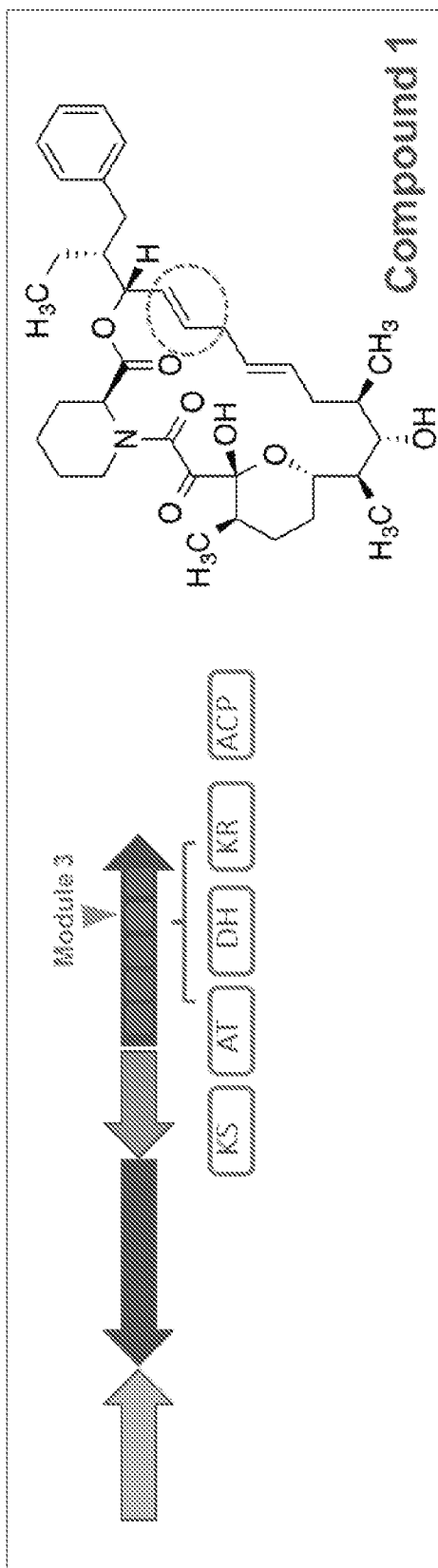


FIG. 3B



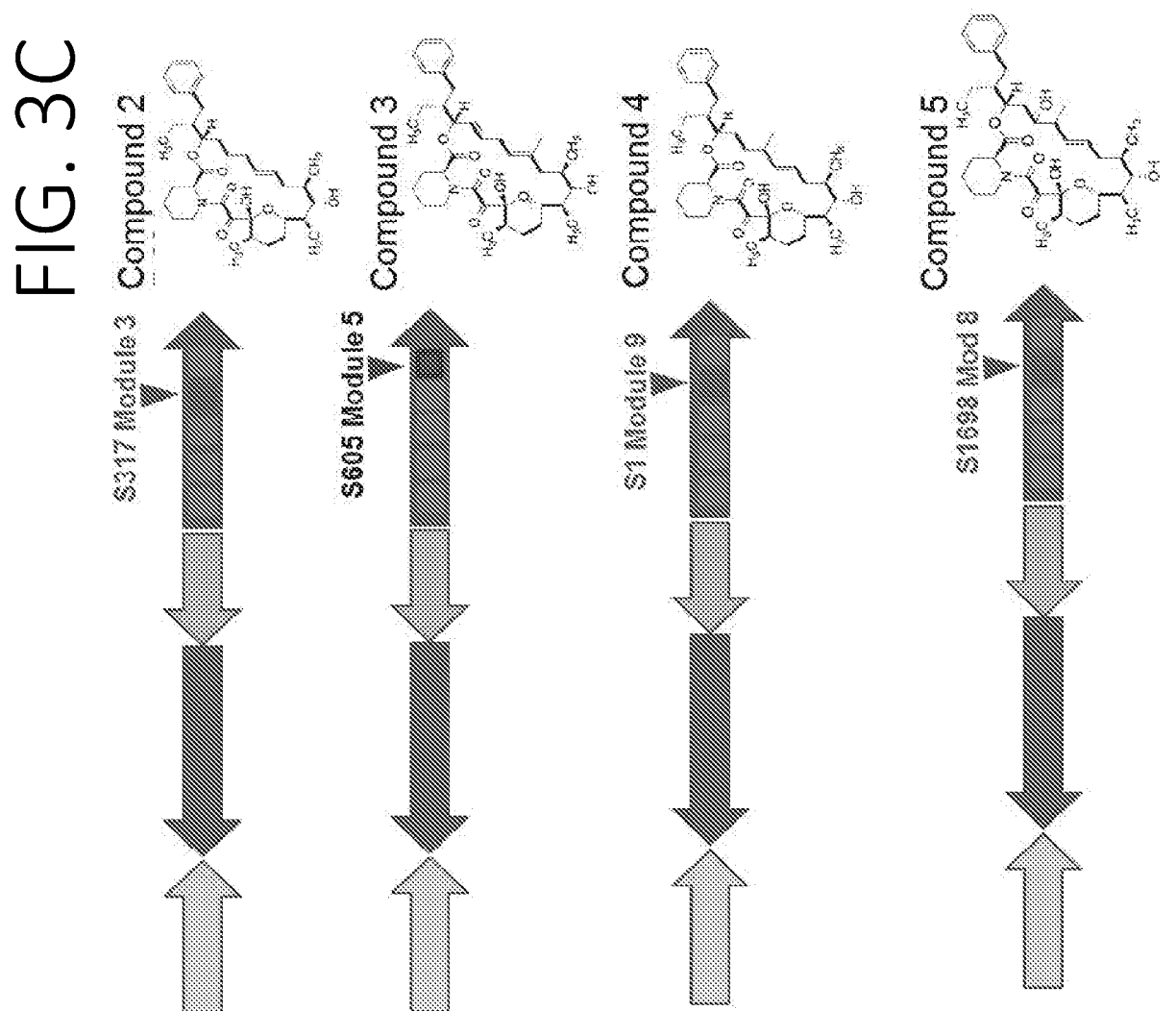


FIG. 4A

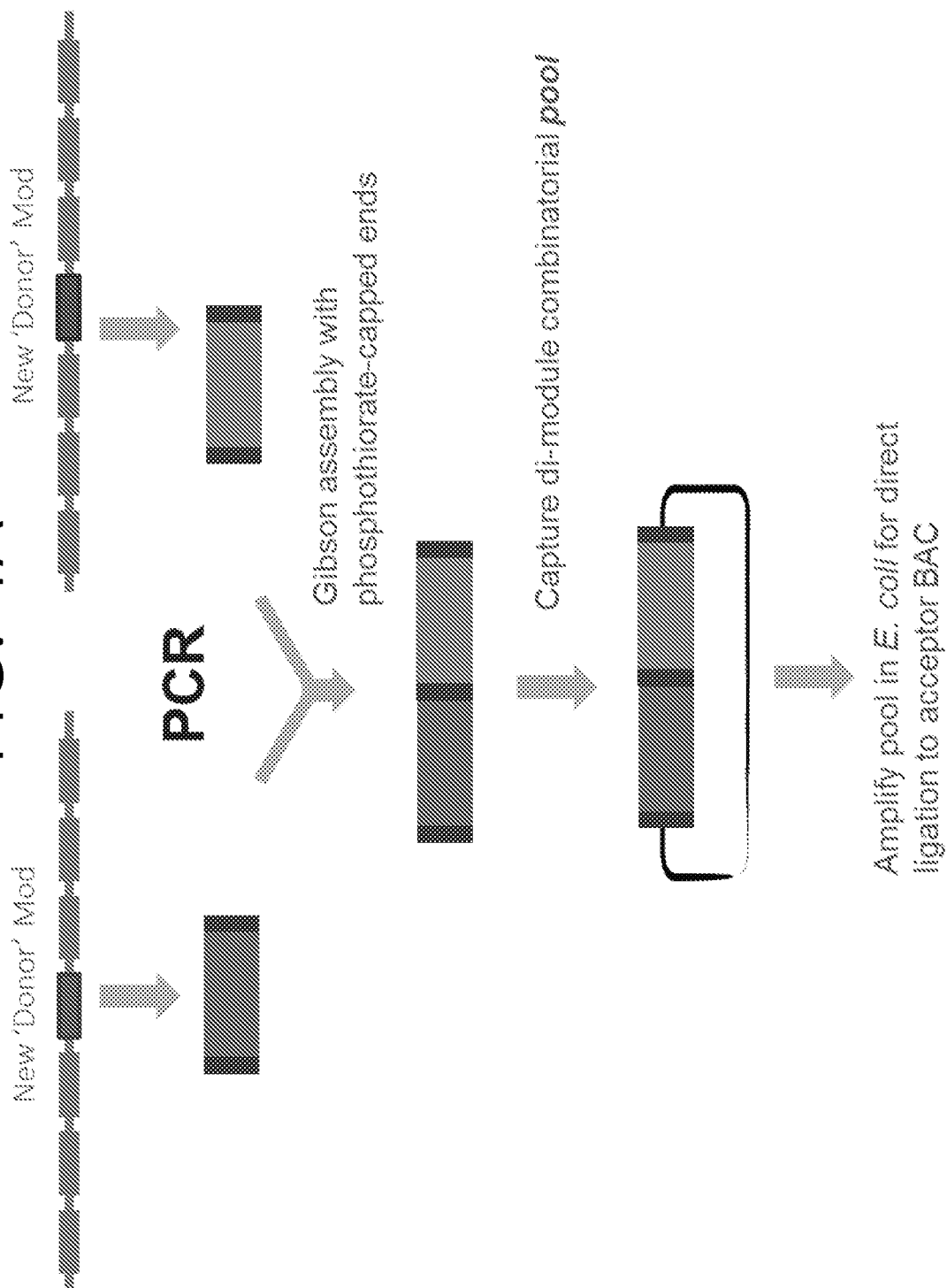


FIG. 4B

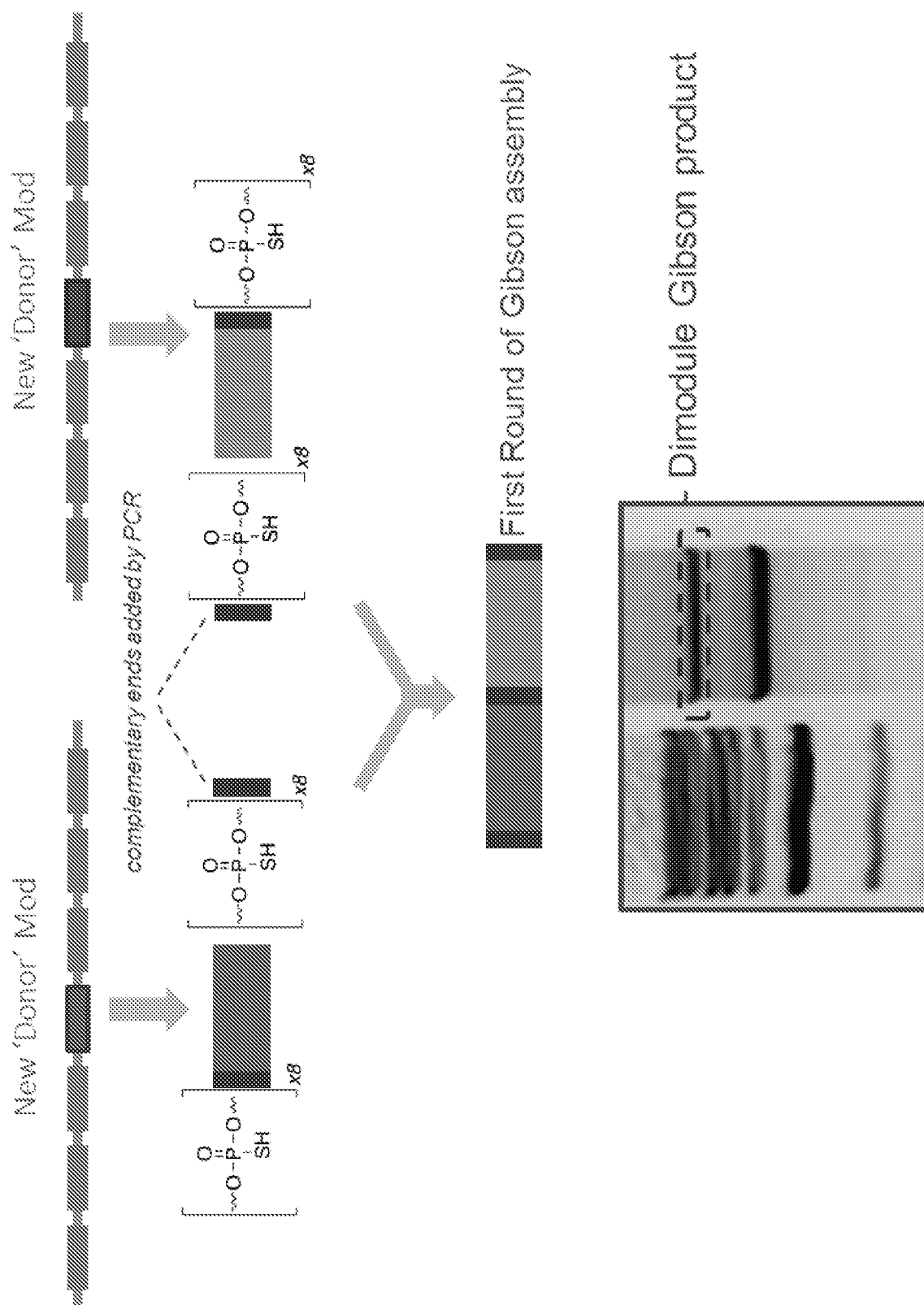


FIG. 4C

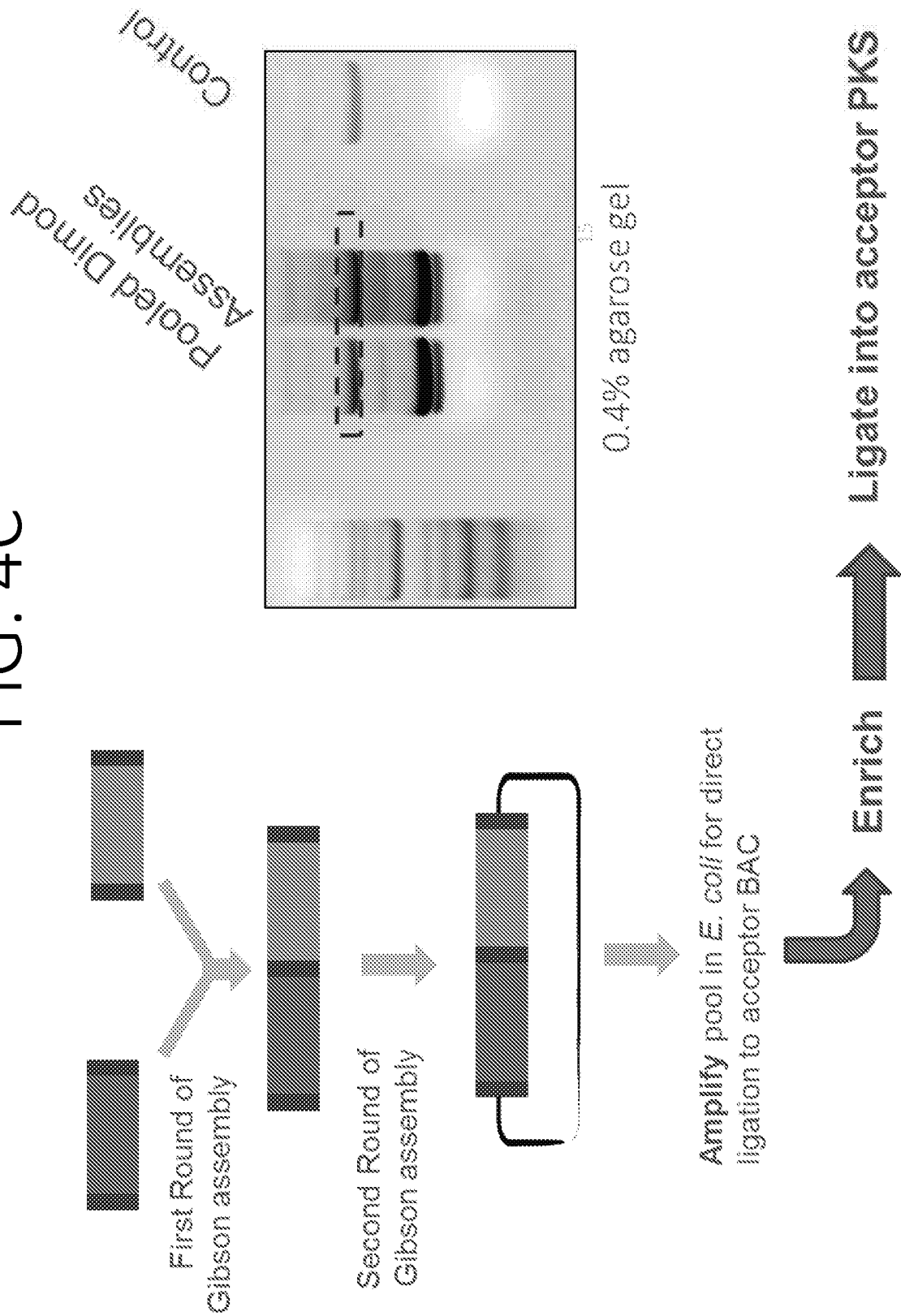


FIG. 4D

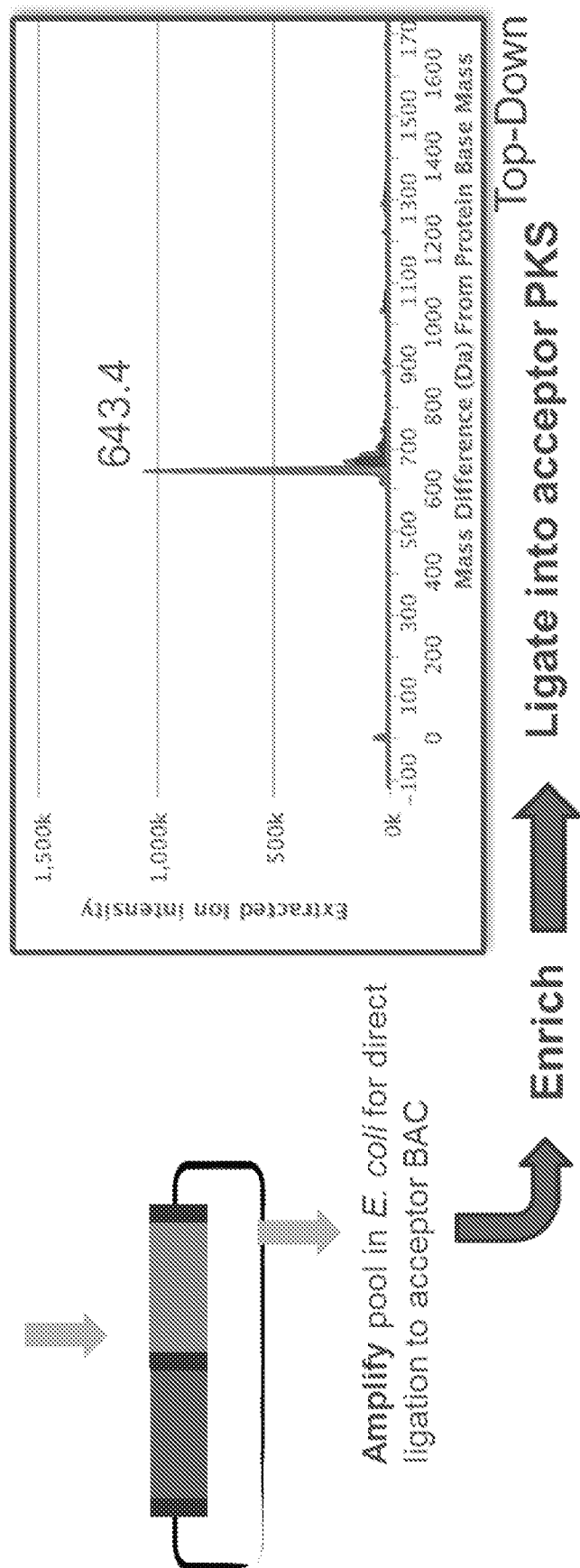
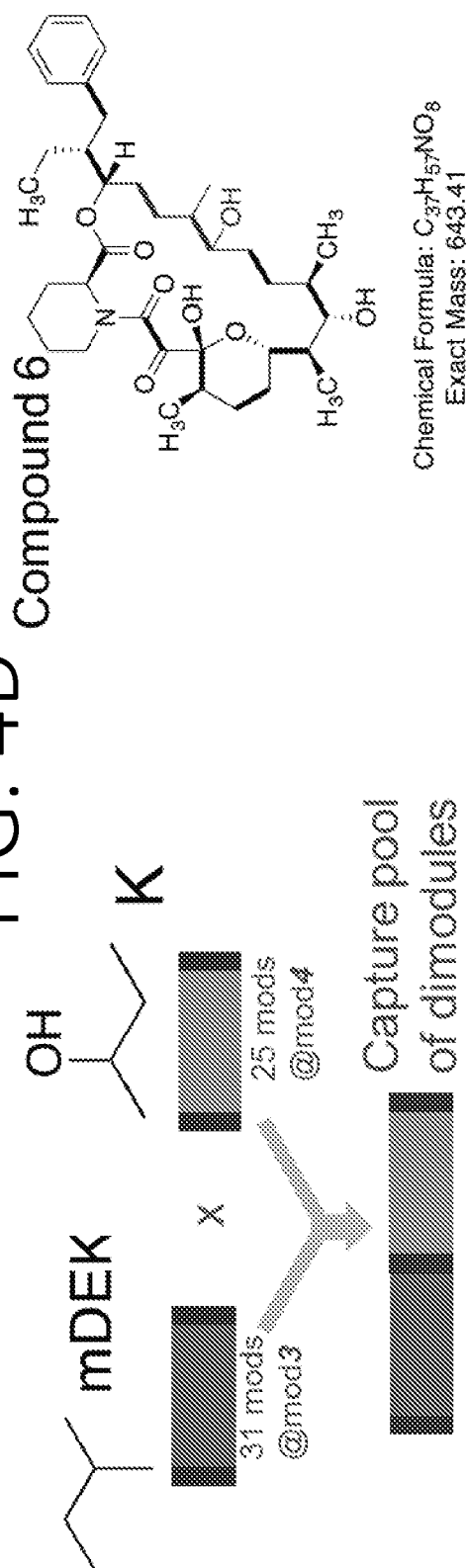
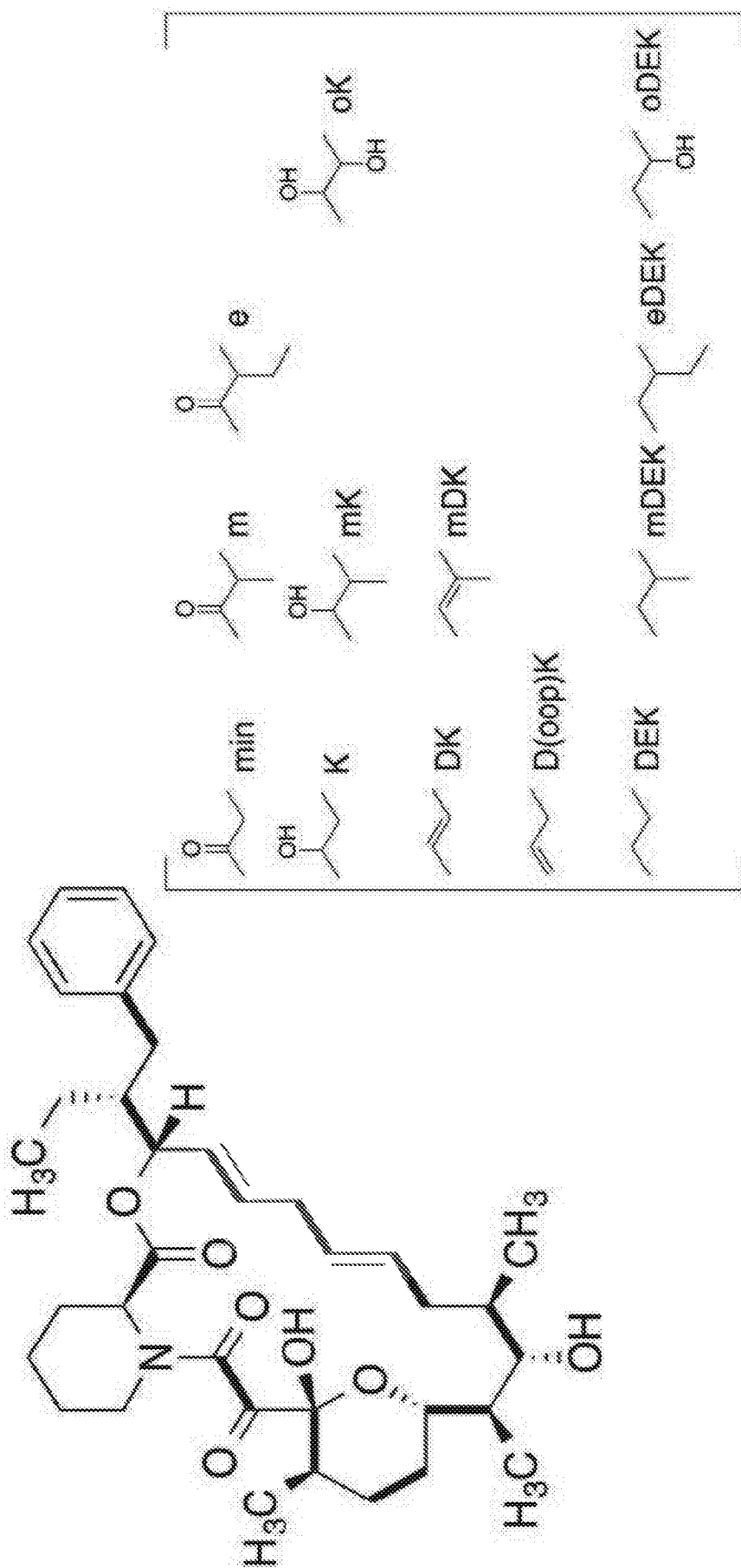


FIG. 4E



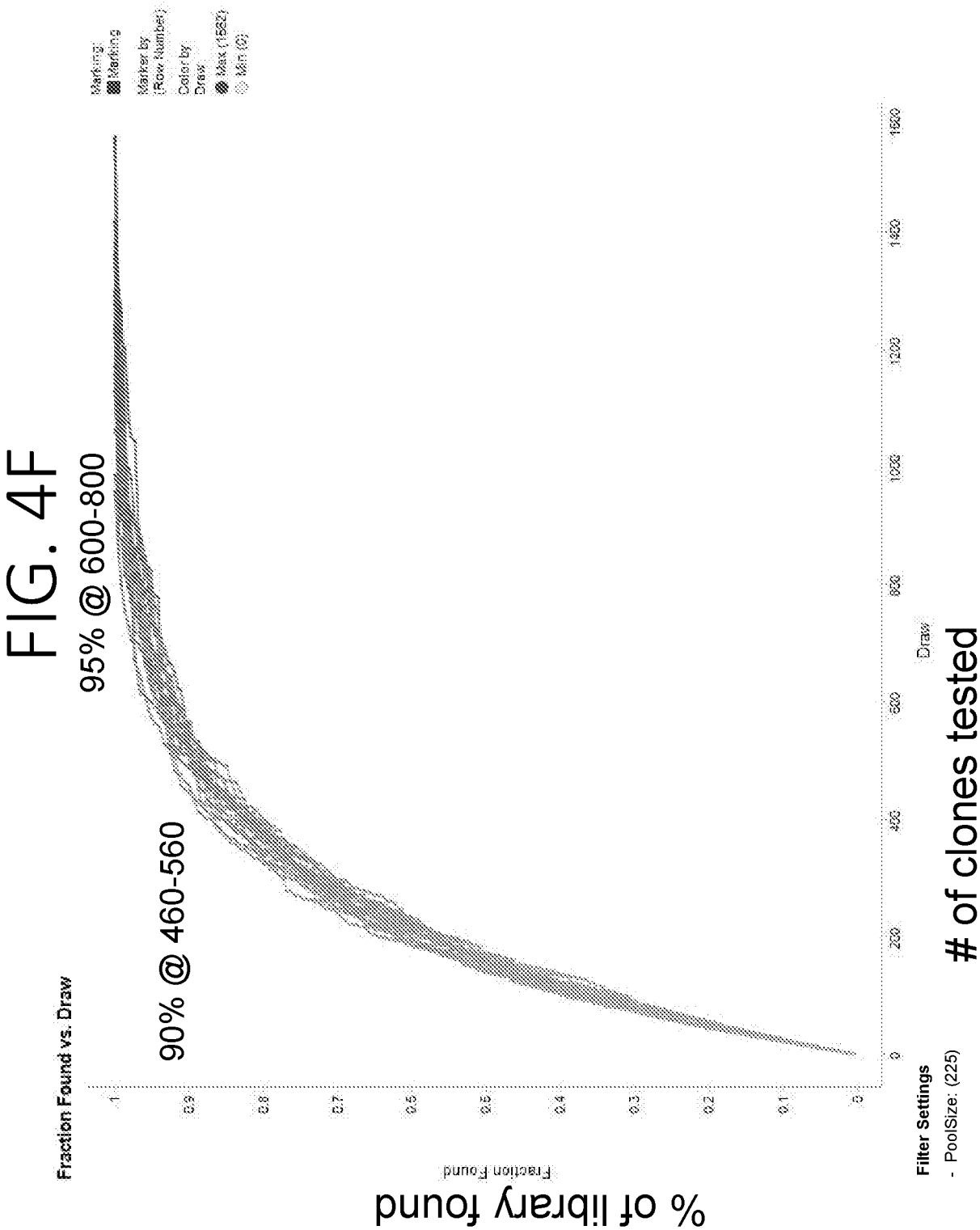
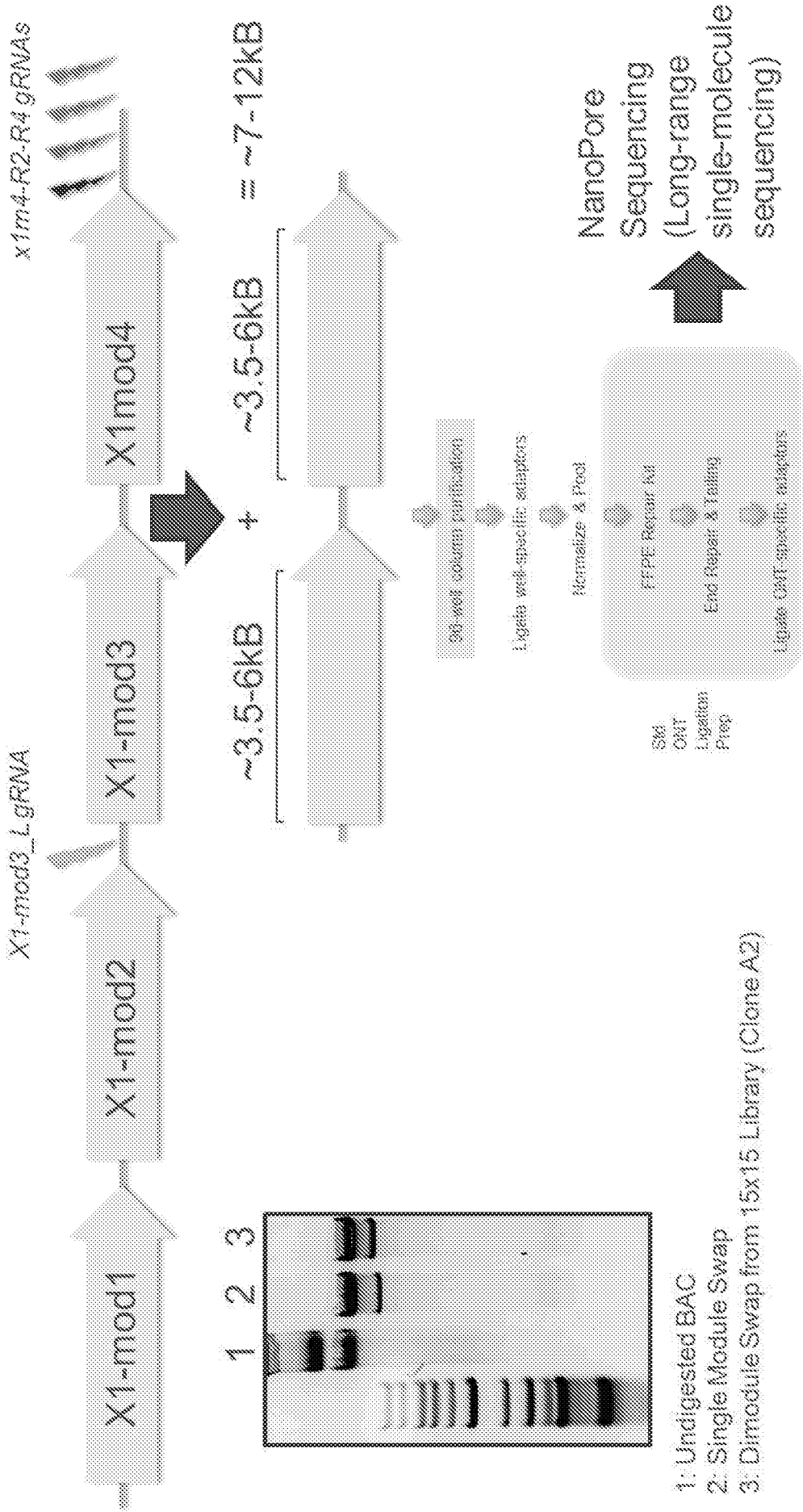


FIG. 4G

In vitro CRIPSR / Cas9 module release for NanoPore Library Preparation



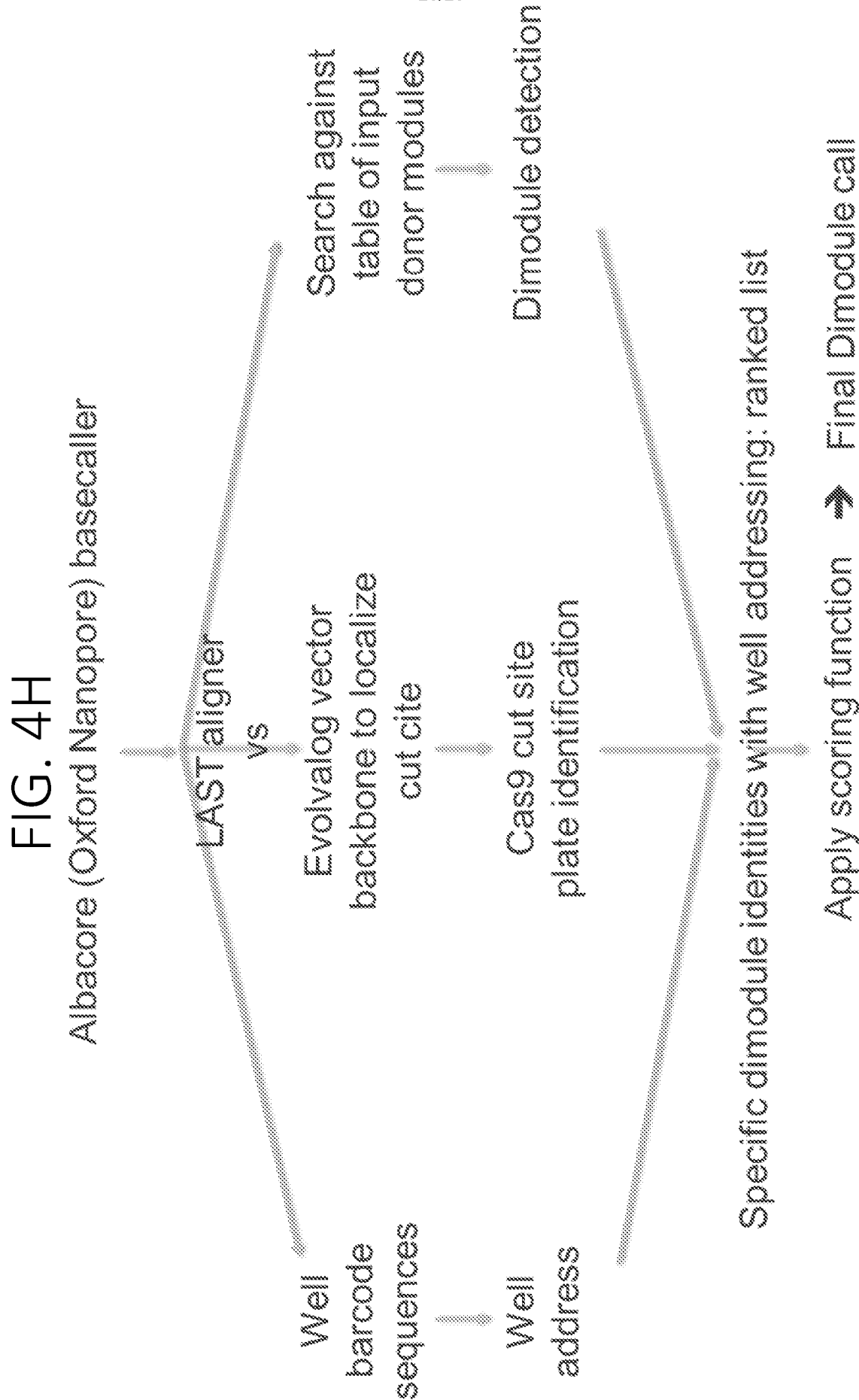


FIG. 5A

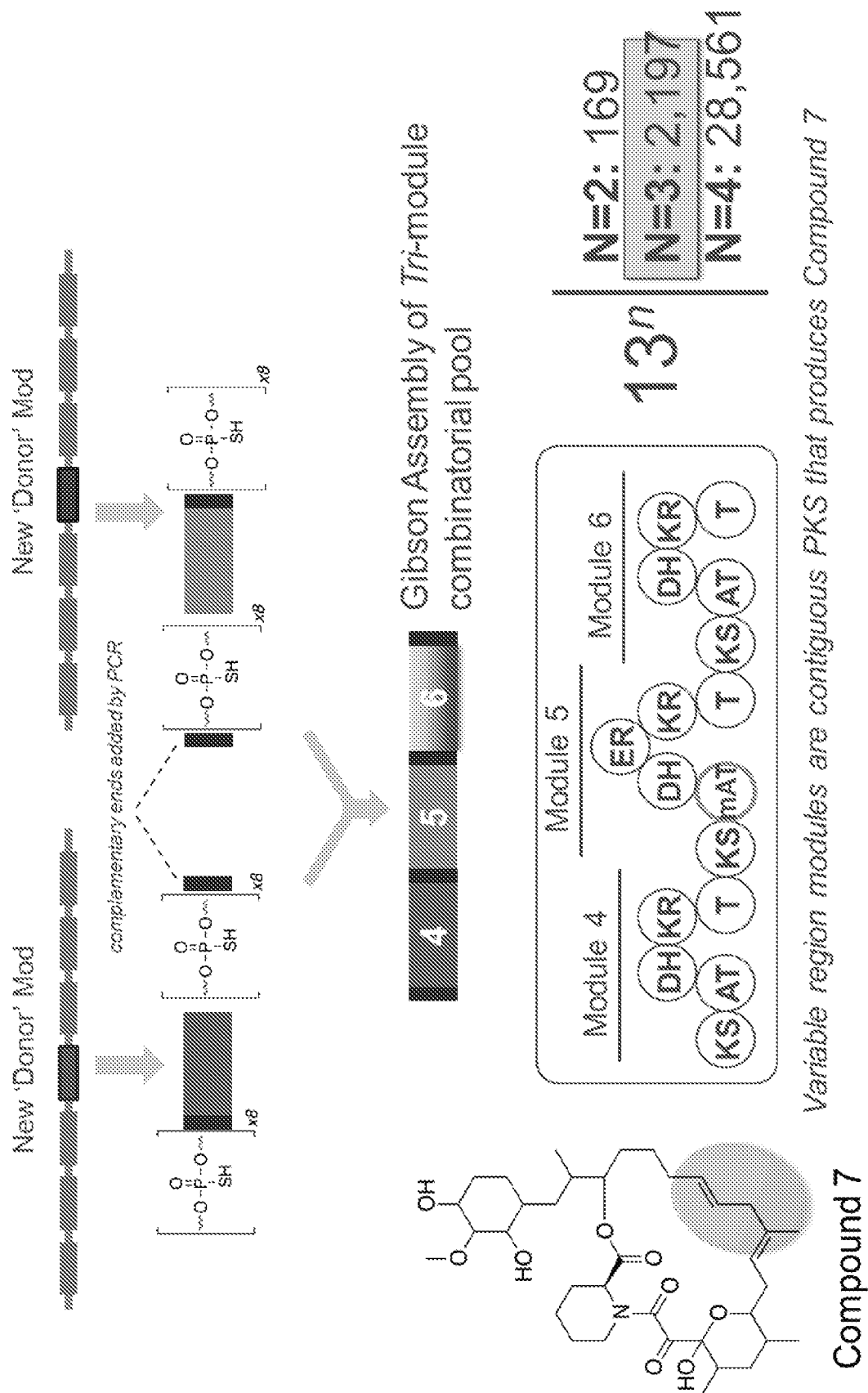


FIG. 5B

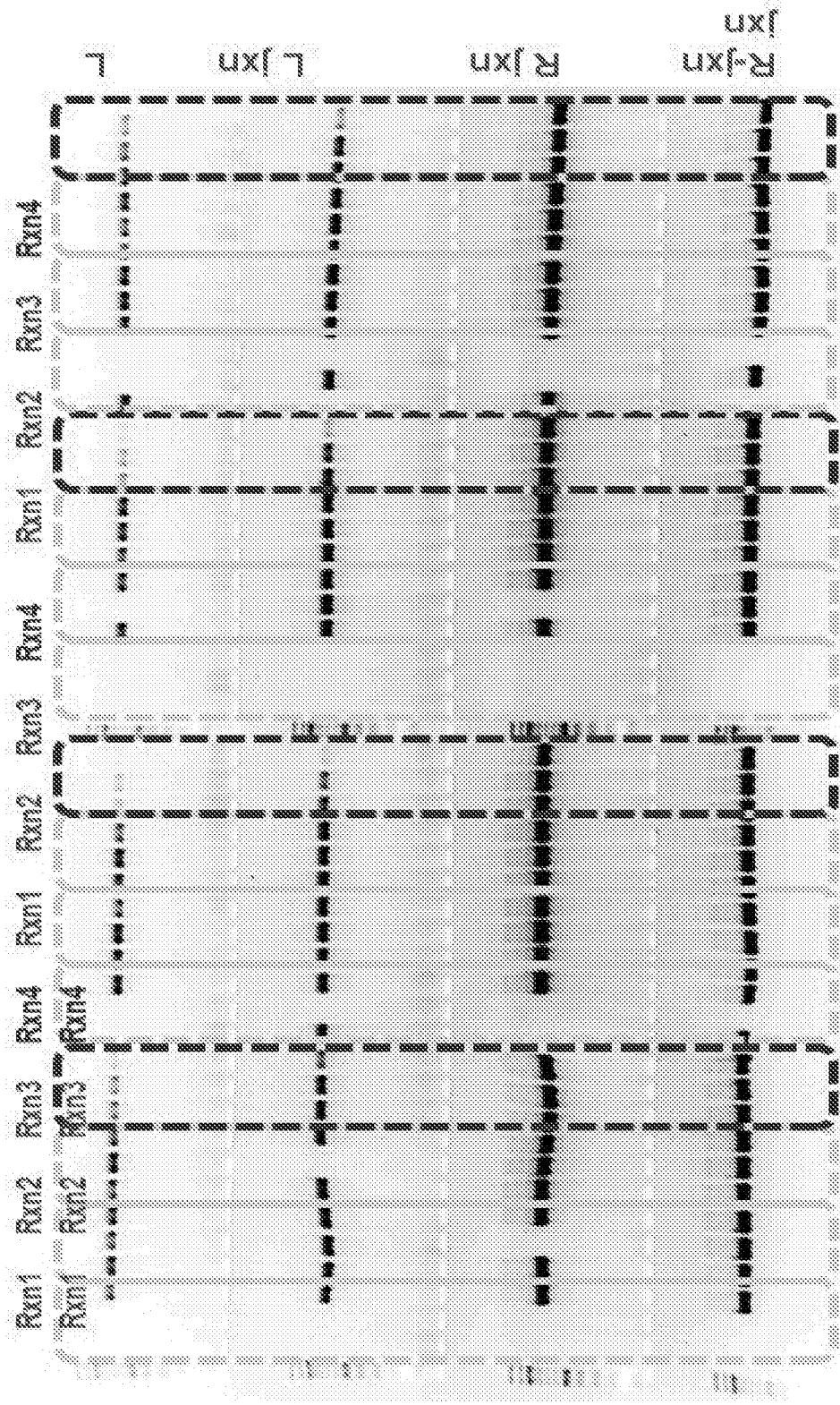
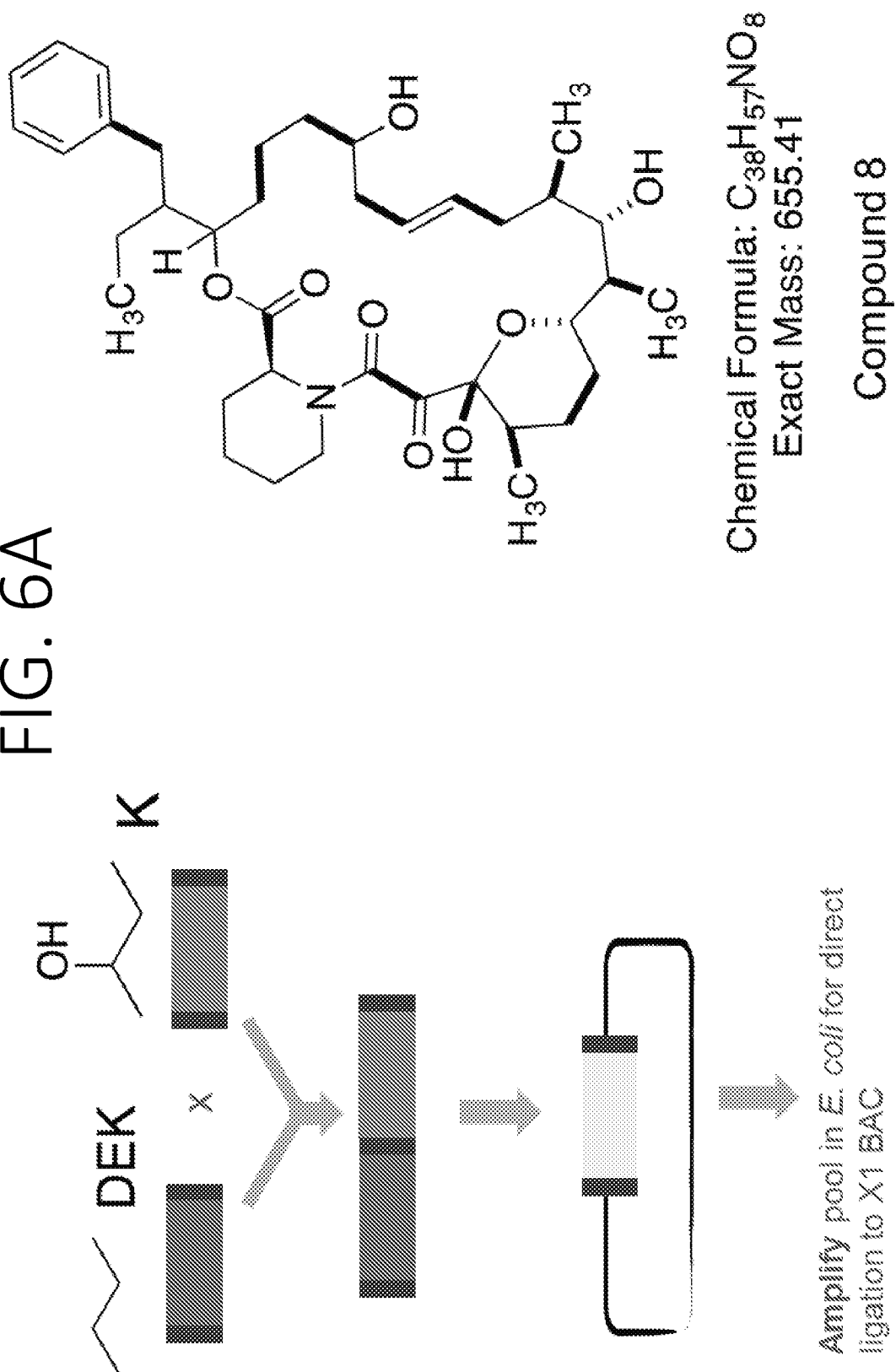


FIG. 6A



Conjugate → Ferment

FIG. 6B

WDB: 0011752 #557-701 RT: 8.97-10.44 AV: 13 NL: 4.58E-4
T: FTMS + p NSI Fullms [1975.00-2375.00]
1994.588

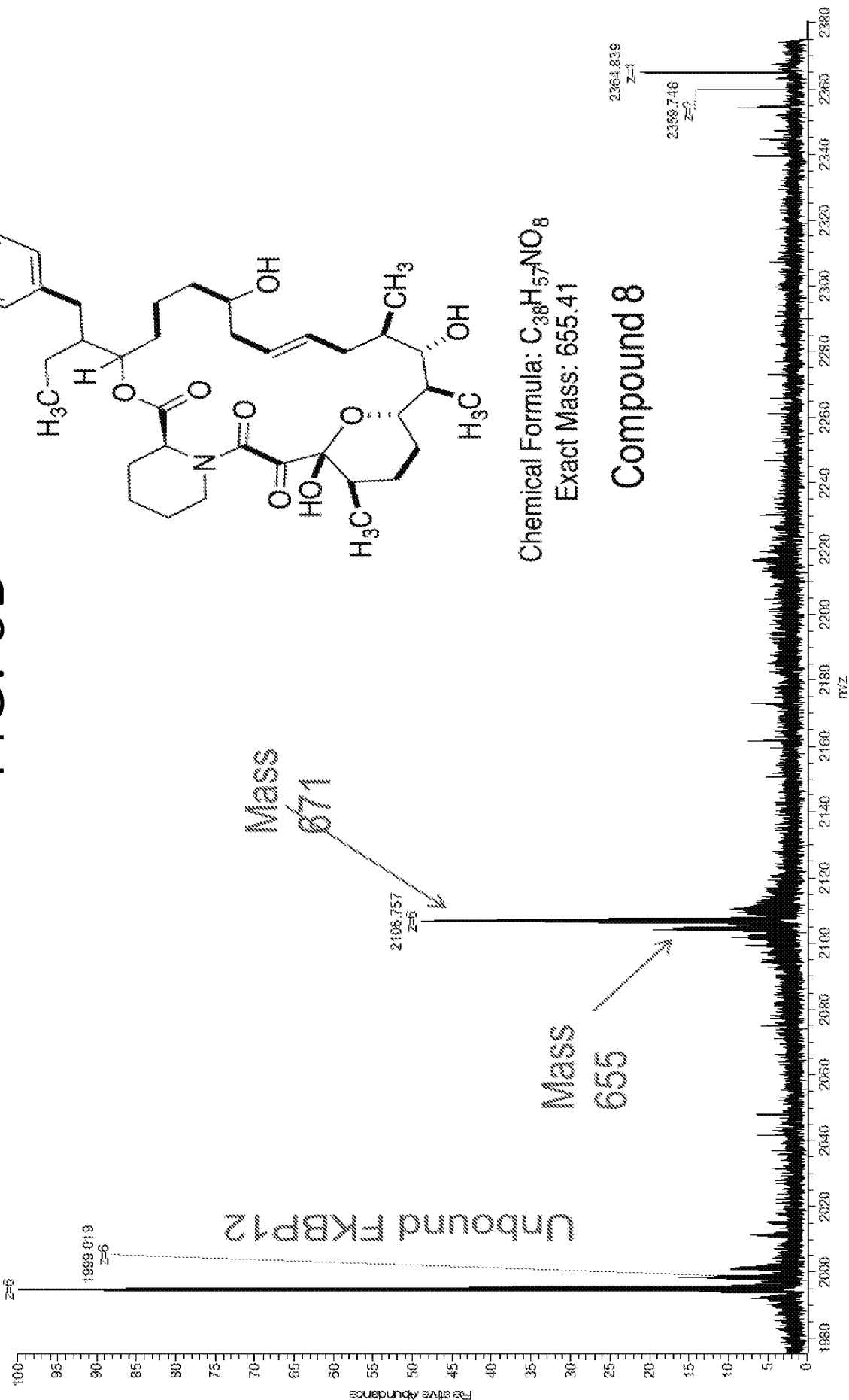


FIG. 7A

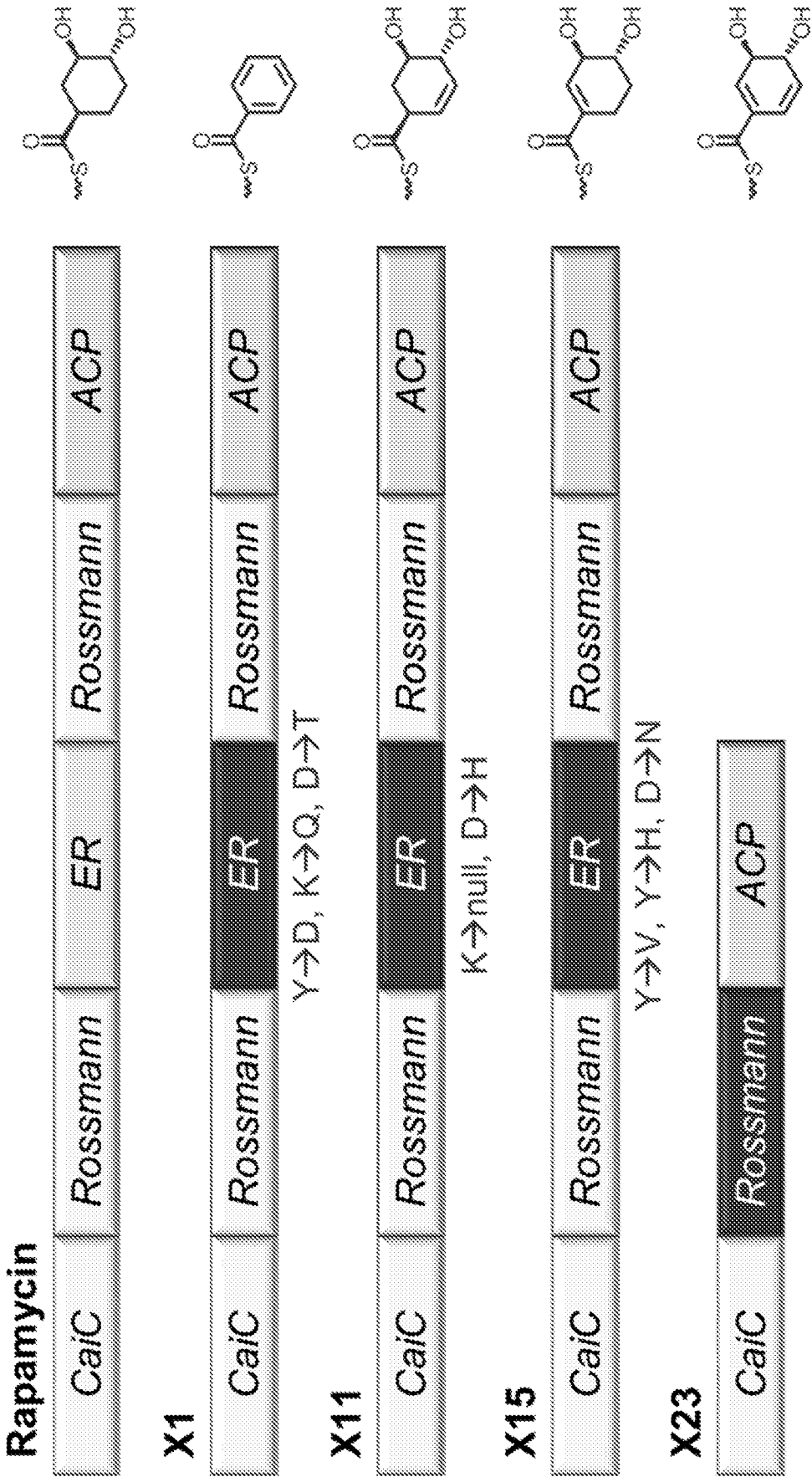
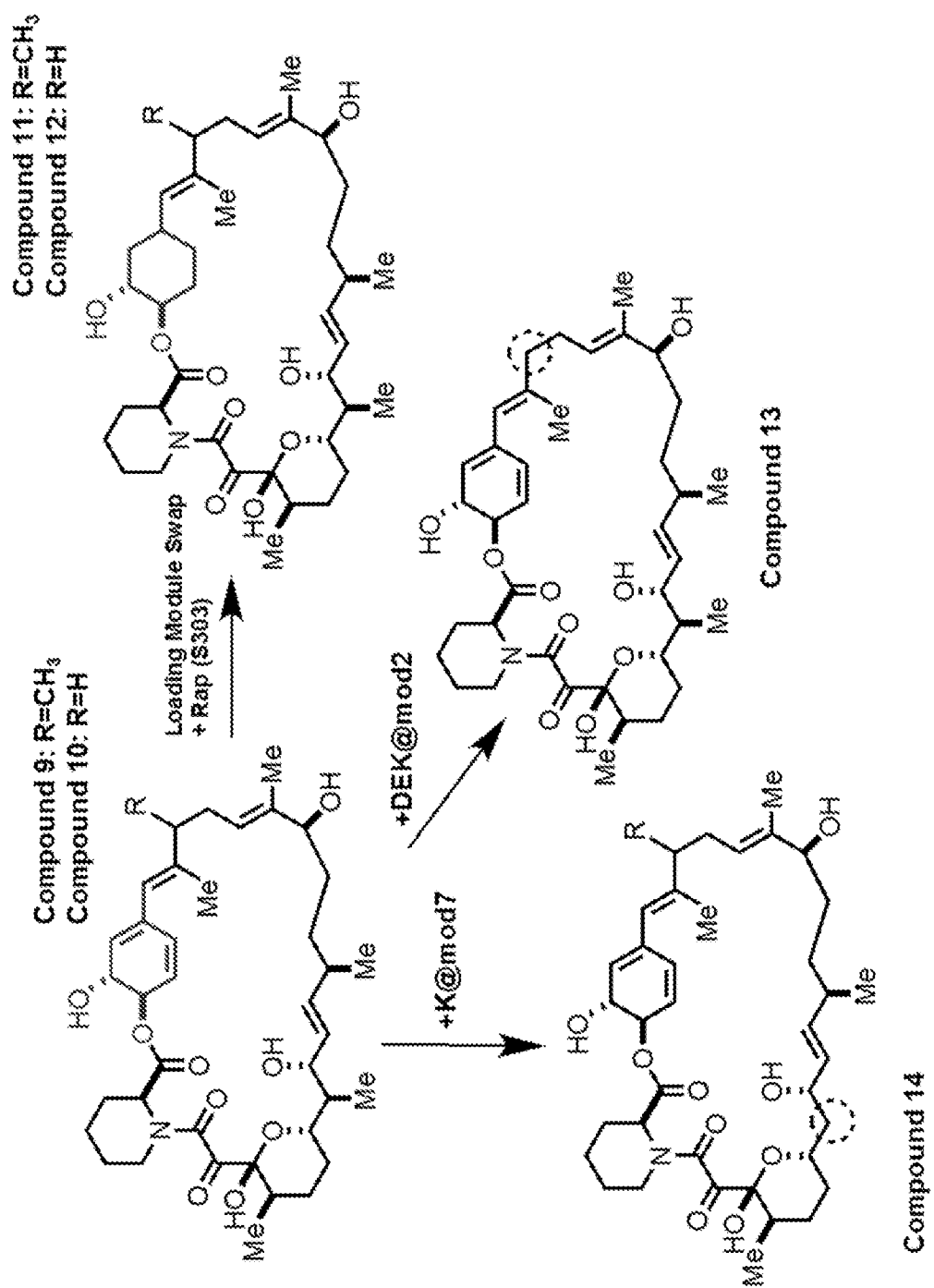


FIG. 7B



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 17/58800

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - A61K 31/436, A61P 17/00, A61P 17/06, A61P 25/00 (2018.01)
 CPC - C07D 498/18, C07K 2319/00, C12N 9/1029, C12N 9/1205

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History Document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Search History Document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History Document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	US 2015/0307855 A1 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 29 October 2015 (29.10.2015) Claim 1; Claim 2; Claim 3; Claim 14; para [0026]; para [0047]; para [0083]; para [0088]; para [0119]; para [0124-0125]	1, 2 --- 3, 4
X -- A	US 2003/0175901A1 (REEVES et al.,) 18 September 2003 (18.09.2003) Abstract; Claim 5; p98 Sequence listing	12 --- 13
Y	~ PFEIFER et al., Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of E. coli. Science. 2 March 2001, Volume 291, pg 1790-1792. p1791, col 2-3;	3, 4/3
Y	~ RANGANATHAN et al., Knowledge-based design of bimodular and trimodular polyketide synthases based on domain and module swaps: a route to simple statin analogues. Chemistry & Biology. October, 1999, Volume 6, Issue 10, pages 731-741. p735, Fig. 4 legend	4
A	~ UniProtKB Accession No. A0A061A6I8 "Streptomyces iranensis Modular polyketide synthase" 3 September 2014 [online]. [Retrieved on 2 February 2018]. Retrieved from the internet: <URL: http://www.uniprot.org/uniprot/A0A061A6I8 > Entire document	13
A	~ UniProtKB Accession No. Q54296 "Streptomyces hygroscopicus Polyketide synthase" 1 November 1996 [online]. [Retrieved on 2 February 2018]. Retrieved from the internet: <URL: http://www.uniprot.org/uniprot/Q54296 > Entire document	13



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

14 February 2018

Date of mailing of the international search report

03 APR 2018

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
 P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300
 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 17/58800

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 5-11, 14-47
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I+: Claims 1-4, 12 and 13, directed to an engineered polyketide synthase comprising one or more heterologous modules. The polyketide synthase will be searched to the extent that the module sequence encompasses SEQ ID NO: 1 (the first claimed module sequence). It is believed that claims 1-4, 12 and 13 encompass this first named invention, and thus these claims will be searched without fee to the extent that the module sequence encompasses SEQ ID NO: 1. Additional module sequence(s) will be searched upon the payment of additional fees. Applicants must specify the claims that encompass any additionally elected module sequence(s). Applicants must further indicate, if applicable, the claims which encompass the first named invention, if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched. An exemplary election would be module sequence SEQ ID NO: 2 (claims 1-4, 12 and 13).

-----Please see continuation in first extra sheet-----

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-4, 12 and 13 limited to SEQ ID NO: 1

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 17/58800

Continuation of Box No. III. Observations where unity of invention is lacking

Group II: Claims 48-51, directed to a method of producing a compound or a method of producing a plurality of polynucleotides corresponding to an engineered polyketide synthase.

The inventions listed as Groups I+ and II do not relate to a single special technical feature under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Special technical features:

Group I+ has the special technical feature of an engineered polyketide synthase, that is not required by Group II.

Group II has the special technical feature of producing a nucleic acid encoding a modified polyketide synthase, that is not required by Group I+.

No technical features are shared among the amino acid sequences of polyketide synthase modules of Group I+ and, accordingly, these groups lack unity a priori.

Additionally, even if the inventions listed as Groups I+ were considered to share technical features, these shared technical features are previously disclosed by the prior art, as further discussed below.

Common technical features:

Groups I+ and II share the common technical feature of an engineered polyketide synthase comprising one or more heterologous modules with altered enzymatic activity relative to a reference polyketide.

The inventions of Groups I+ share the technical features of an engineered polyketide synthase comprising one or more heterologous modules with altered enzymatic activity relative to a reference polyketide, wherein the engineered polyketide synthase is capable of producing a polyketide when expressed under conditions suitable to allow expression of a compound by the engineered polyketide synthase and wherein the one or more heterologous modules do not substantially inhibit polyketide translocation during polyketide biosynthesis, wherein the one or more heterologous modules comprise linking sequences which are compatible to the linking sequences of the modules adjacent thereto, and wherein the polyketide expression level of the engineered polyketide synthase is at least 1 % of the polyketide expression level of the reference polyketide synthase; and a chimeric polyketide synthase, wherein at least one module of the chimeric polyketide synthase has been modified as compared to a polyketide synthase.

However, these shared technical features are previously made obvious by US 2015/0307855 A1 to The Regents Of The University Of California (hereinafter 'Univ California'), in view of the article entitled 'Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of E. coli' by Pfeifer et al., (hereinafter 'Pfeifer').

Univ California teaches an engineered polyketide synthase comprising one or more heterologous modules with altered enzymatic activity relative to a reference polyketide (Claim 1 - 'A non-naturally occurring polyketide synthase (PKS), or functional variant thereof, capable of synthesizing a 3-hydroxycarboxylic acid or ketone.'; Claim 3 - 'The PKS of claim 1, wherein the PKS is a hybrid PKS comprising modules, domains, and/or portions thereof, or functional variant thereof, from two or more PKSs.'; para [0088] - 'We have uncovered broad substrate specificity of the loading didomain, which was not anticipated from the study of the PKS in its native host and led us to speculate that novel lipomycin analogues can be produced by increasing intracellular acyl-CoA'; para [0124] - 'Using this method, we created mutant 75 where the lipAT1 domain was exchanged with the one from module 1 of the borrelidin PKS, which is specific to malonyl-CoA. This mutant maintained the wild-type kinetics to produce the corresponding methyl ketones, which includes the industrially important MEK (ketone 8) and methyl isobutyl ketone, MIBK (ketone 11) when the corresponding starter units, propionyl-CoA and isovaleryl-CoA, respectively, were added to the reaction mixtures in vitro (FIG. 16B).'; para [0125] - 'Specifically, we mutated the active site serine residues of KR domain (LFSSIAG, SEQ ID NO:7) of both LipPks1+TE and mutant 75 to alanine residues to produce ketones even in the presence of NADPH. These engineered PKSs were designated as mutant 87 and 88, respectively. The mutant proteins were purified and we confirmed that they were able to produce ketones in vitro in the presence of NADPH', wherein the engineered polyketide synthase is capable of producing a polyketide when expressed under conditions suitable to allow expression of a compound by the engineered polyketide synthase (Claim 2 - 'The PKS of claim 1, wherein the 3-hydroxycarboxylic acid or ketone is not a compound synthesized by a naturally occurring PKS.'; Claim 14 - 'A method of producing a 3-hydroxycarboxylic acid or ketone, comprising: providing a host cell of claim 12, and culturing said host cell in a suitable culture medium such that the 3-hydroxycarboxylic acid or ketone is produced') and wherein the one or more heterologous modules do not substantially inhibit polyketide translocation during polyketide biosynthesis (para [0119] - 'The extended polyketide chain is translocated to the KS of the next module for another round of condensation and -carbon modification'; para [0026] - 'FIG. 15 shows structurally diverse ketone production by engineered PKSs. . The extended polyketide chains tethered to the ACP domain are translocated to the downstream TE to release -keto-carboxylic acids in the absence of NADPH or when the KR domain is inactivated.').

-----Please see continuation in next extra sheet-----

Continuation from last extra sheet

Univ California further teaches a chimeric polyketide synthase, wherein at least one module of the chimeric polyketide synthase has been modified as compared to a polyketide synthase wherein the one or more heterologous modules comprise linking sequences which are compatible to the linking sequences of the modules adjacent thereto (para [0047] - 'In some embodiments, the components of the hybrid PKS are arranged onto polypeptides having interpolypeptide linkers that direct the assembly of the polypeptides into the functional PKS protein, such that it is not required that the PKS have the same arrangement of modules in the polypeptides as observed in natural PKSs.'), and wherein the polyketide expression level of the engineered polyketide synthase is measured (para [0083] - 'LipPks1+TE was purified by Ni affinity chromatography followed by anion exchange chromatography to yield 2 mg of protein from 1 L of E. coli culture (FIG. 6C)'). Univ California does not expressly teach the polyketide expression level of the engineered polyketide synthase is at least 1 % of the polyketide expression level of the reference polyketide synthase.

Pfeifer teaches biosynthesis of complex polyketides in a metabolically engineered strain of E. coli is superior to wild-type S. erythraea (p. 1791, col 1, para 1 'Formation of triketide lactone under both feeding conditions confirmed the ability of BAP1 to produce polyketides and encouraged us to attempt 6dEB production in the same host'; p 1791, col 2-3, - 'the specific productivity of this cellular catalyst is 0.1 mmol of 6dEB per gram of cellular protein per day. This is superior to wild-type S. erythraea and compares well to an industrially relevant strain that overproduces erythromycin (0.2 mmol of erythromycin per gram of cellular protein per day)'). Since both Pfeifer and Univ California use E.coli for heterologous expression, it would have been obvious to one of ordinary skill in the art that the polyketide expression level of the engineered polyketide synthase of Univ California could be at least 1 % of the polyketide expression level of the reference polyketide synthase according to Pfeifer.

As the technical features were known in the art at the time of the invention, they cannot be considered special technical features that would otherwise unify the groups.

Therefore, Group I+ and II inventions lack unity under PCT Rule 13 because they do not share the same or corresponding special technical feature.