

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
24 December 2003 (24.12.2003)

PCT

(10) International Publication Number
WO 03/105755 A2

- (51) International Patent Classification⁷: **A61K** (74) Agents: **BAK, Mary, E.** et al.; Howson and Howson, Spring House Corporate Center, P.O. Box 457, Spring House, PA19477 (US).
- (21) International Application Number: PCT/US03/16216
- (22) International Filing Date: 16 June 2003 (16.06.2003) (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (25) Filing Language: English (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (26) Publication Language: English
- (30) Priority Data: 10/173,718 17 June 2002 (17.06.2002) US
- (63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application: US 10/173,718 (CON) Filed on 17 June 2002 (17.06.2002)
- (71) Applicant (*for all designated States except US*): **ISIS PHARMACEUTICALS INC.** [US/US]; 2292 Faraday Avenue, Carlsbad, CA 92008 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **ZHANG, Hong** [US/US]; 3339 Cadencia Street, Carlsbad, CA 92009 (US). **DOBIE, Kenneth, W.** [US/US]; 703 Stratford Court, #4, Del Mar, CA 92014 (US).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 03/105755 A2

(54) Title: ANTISENSE MODULATION OF VEGF-C EXPRESSION

(57) Abstract: Antisense compounds, compositions and methods are provided for modulating the expression of VEGF-C. The compositions comprise antisense compounds, particularly antisense oligonucleotides, targeted to nucleic acids encoding VEGF-C. Methods of using these compounds for modulation of VEGF-C expression and for treatment of diseases associated with expression of VEGF-C are provided.

ANTISENSE MODULATION OF VEGF-C EXPRESSION**FIELD OF THE INVENTION**

5 The present invention provides compositions and methods for modulating the expression of VEGF-C. In particular, this invention relates to compounds, particularly oligonucleotides, specifically hybridizable with nucleic acids encoding VEGF-C. Such compounds have been shown to
10 modulate the expression of VEGF-C.

BACKGROUND OF THE INVENTION

 All vessels of the circulatory system are lined with endothelial cells. This endothelial cell lining is formed by
15 two processes: vasculogenesis, the *de novo* formation of new endothelial channels from differentiating angioblasts, and angiogenesis, the sprouting or splitting of capillaries from pre-existing vessels. These two processes are regulated by polypeptide growth factors and their receptors. Adult
20 vasculature is normally quiescent, but it can become activated to form new capillaries as a part of wound healing or tumorigenesis. During tumorigenesis, the balance between angiogenesis inhibitors, such as endostatin and thrombospondin-1, and angiogenesis inducers, such as vascular
25 endothelial growth factor (VEGF), is shifted and rapid vessel ingrowth occurs, supporting tumor expansion (Olofsson et al., *Curr. Opin. Biotechnol.*, 1999, 10, 528-535).

 Aberrant regulation of endothelial cell growth and proliferation contributes to tumor formation, cardiovascular
30 disease and atherosclerosis, and diseases such as psoriasis and rheumatoid arthritis (Enholm et al., *Trends Cardiovasc. Med.*, 1998, 8, 292-297). During embryonic vasculogenesis, VEGF is an important regulator of endothelial cell

-2-

proliferation, chemotaxis, migration and vascular permeability, as well as of normal and pathological angiogenesis. A critical role of VEGF in embryogenesis is demonstrated by the unprecedented finding that inactivation of even a single VEGF allele results in embryonic lethality (Joukov et al., *J. Cell Physiol.*, 1997, 173, 211-215).

A family of VEGF-related molecules has recently been characterized, and consists of at least five members: VEGF/VEGF-A, VEGF-B, VEGF-2/VEGF-C, VEGF-D and placenta growth factor (PlGF). Within the VEGF family of growth factors, VEGF-C and its closest relative, VEGF-D constitute a subgroup characterized by the presence of unique amino- and carboxy-terminal extensions flanking the common VEGF-homology domain. VEGF family members transmit their signals by binding to the protein tyrosine kinase receptors VEGFR-1/FLT1, VEGFR-2/KDR/FLK1, and VEGFR-3/FLT4, which are structurally related to the PDGF family of class III transmembrane receptors. Upon ligand binding, the receptors auto- or trans-phosphorylate specific cytoplasmic tyrosine residues to initiate an intracellular cascade of signaling that ultimately reaches cytoskeletal proteins and nuclear transcription factor effectors (Olofsson et al., *Curr. Opin. Biotechnol.*, 1999, 10, 528-535; Wang et al., *Blood*, 1997, 90, 3507-3515).

The receptor tyrosine kinase VEGFR-3/(FLT4) is expressed mainly on lymphatic endothelia and originally was considered an orphan receptor, as its ligand was unknown and it was found not to bind VEGF. The ligand of VEGFR-3/FLT4 was purified by receptor-affinity chromatography from medium conditioned by PC-3 prostatic adenocarcinoma cells, a partial amino acid sequence was obtained, and the cDNA encoding the ligand was cloned using a degenerate PCR-based strategy. This ligand specific for the VEGFR-3/FLT4 receptor was vascular endothelial growth factor-C (VEGF-C; also known as VEGFC,

-3-

vascular endothelial growth factor c precursor, FLT4 ligand, VEGF-2, vascular endothelial growth factor related protein, and VRP) (Joukov et al., *EMBO J.*, 1996, 15, 290-298).

Independently, a sequence was identified in the EST database as homologous to VEGF, and using the EST clone as a probe, a full length cDNA encoding this VEGF-related protein (VRP/VEGF-C) was isolated from a cDNA library made from the human glioma cell line G61. Additionally, two VEGF-C cDNA clones containing 152 and 557 base pair deletions when compared with the full-length clone were also identified, presumably generated by alternative splicing. The predicted proteins encoded by these two deleted cDNAs contain either only part or none of the VEGF-homology domain (Lee et al., *Proc. Natl. Acad. Sci. U. S. A.*, 1996, 93, 1988-1992).

The genomic organization of the human and mouse VEGF-C genes was characterized, and both genes comprise over 40 kilobase pairs of genomic DNA and consist of seven coding exons. Alternative splicing leads to one mRNA transcript lacking exon 4 and another putative mRNA form lacking exons 2-4, corresponding to the 152 and 557 nucleotide deletions, respectively, described by Lee, et al. (Chilov et al., *J. Biol. Chem.*, 1997, 272, 25176-25183).

Antibodies recognizing two different regions of the VEGF-C precursor were generated and allowed the characterization of the VEGF-C polypeptide and its variously processed forms. VEGF-C is synthesized as a precursor protein that undergoes proteolytic processing in which the carboxy-terminal domain is cleaved upon secretion but remains bound to the amino-terminal domain by disulfide bonds. Thus, the major secreted form of VEGF-C is a homodimer formed through the disulfide linkage of the C-terminal propeptide of one polypeptide chain to the N-terminal part of the other chain. Further proteolytic processing of the N-terminal propeptide

-4-

then releases the mature, biologically active, non-disulfide bonded, soluble form of VEGF-C. This mature form consists essentially of a dimer of two polypeptide chains corresponding to the VEGF-homology domain and possesses VEGF-like effects on endothelial cells, stimulating their proliferation and migration, as well as increasing permeability of blood vessels *in vivo* (Enholm et al., *Trends Cardiovasc. Med.*, 1998, 8, 292-297; Joukov et al., *EMBO J.*, 1997, 16, 3898-3911). The mature form of VEGF-C also has receptor specificity; although both the full-length and mature forms of VEGF-C bind to VEGFR-3/FLT4, only mature VEGF-C can bind and activate VEGFR-2/KDR (Joukov et al., *EMBO J.*, 1997, 16, 3898-3911).

Northern blot analyses detect 2.0 and 2.4 kb RNA species from many embryonal and adult tissues. In adult humans, the VEGF-C mRNA is expressed most prominently in heart, placenta, ovary, small intestine, and the thyroid gland, and tumor cells express almost exclusively the 2.4 kb form (Joukov et al., *EMBO J.*, 1996, 15, 290-298).

By *in situ* hybridization, VEGF-C mRNA was found to be expressed in mesenchymal cells of postimplantation mouse embryos, particularly in regions where the lymphatic vessels undergo sprouting from embryonic veins, such as the perimetanephric, axillary and jugular regions, and in the developing mesenterium, which is rich in lymphatic vessels. In adult mice, the expression of VEGF-C decreases but mRNA can still be observed in the lung, heart, liver and kidney. The pattern of expression of VEGF-C in relation to its receptor VEGFR-3/FLT4 suggests a paracrine mode of action (Kukk et al., *Development*, 1996, 122, 3829-3837).

VEGF-C also has VEGF like properties, including stimulation of blood vascular endothelial cell proliferation and migration, as well as increasing vascular permeability

-5-

(Joukov et al., *EMBO J.*, 1996, 15, 290-298; Joukov et al., *EMBO J.*, 1997, 16, 3898-3911; Lee et al., *Proc. Natl. Acad. Sci. U. S. A.*, 1996, 93, 1988-1992). VEGF-C is also a potent inducer of lymphangiogenesis. Transgenic mice overexpressing VEGF-C directed to basal keratinocytes of the skin exhibited superficial lymphatic, but not vascular, endothelial proliferation and vessel enlargement. Thus, VEGF-C induces selective hyperplasia of the lymphatic vasculature, and VEGF-C is the first growth factor shown to specifically induce lymphangiogenesis (Jeltsch et al., *Science*, 1997, 276, 1423-1425).

VEGF-C is highly chemoattractive for lymphatic endothelial cells. The lymphangiogenic potency of VEGF-C was further demonstrated when exogenously added, recombinant, mature VEGF-C was found to induce proliferation of lymphatic endothelial cells and development of new lymphatic sinuses in differentiated avian chorioallantoic membrane (Oh et al., *Dev. Biol.*, 1997, 188, 96-109).

The lymphatic system is a low-flow, low-pressure system in intimate contact with extracellular matrix and lymph is actively exchanged with interstitial tissue fluid. Cellular trafficking in the lymphatic system is important for immunosurveillance and for pathological processes such as the metastatic spread of tumors. Aberrant function of the lymphatic system is implicated in many disease conditions such as edema, ascites, inflammation, infectious and immune diseases, fibrosis, and tumors such as Kaposi's sarcoma (KS) and lymphangioma/lymphangiomatosis (Enholm et al., *Trends Cardiovasc. Med.*, 1998, 8, 292-297).

A number of cytokines have been postulated to have a role in the pathogenesis of KS. The proliferative effects of basic fibroblast growth factor (bFGF) and oncostatin M (OSM) may occur via activation of the c-Jun N-terminal kinase (JNK)

-6-

signaling pathway in KS cells. Serum, several growth factors such as platelet derived growth factor (PDGF), epidermal growth factor (EGF), transforming growth factor-beta (TGF- β), and proinflammatory cytokines such as tumor necrosis factor-
5 alpha (TNF- α) and interleukin-1 alpha and beta (IL-1 α and IL-1 β) have also been demonstrated to increase the steady state of VEGF-C mRNA but not VEGF-B mRNA in human lung fibroblast IMR-90 cells (Enholm et al., *Oncogene*, 1997, 14, 2475-2483; Ristimaki et al., *J. Biol. Chem.*, 1998, 273, 8413-8418).

10 Furthermore, the VEGFR-3/FLT4 receptor is robustly expressed in KS cells, and stimulation of KS cells with VEGF or VEGF-C resulted in an increase in JNK activity (Liu et al., *J. Clin. Invest.*, 1997, 99, 1798-1804). Thus, VEGF-C may play a role in the pathogenesis of KS.

15 Co-expression of VEGF-C and its receptor VEGFR-3/FLT4 was also detected in most samples from human thyroid adenomas and adenocarcinomas. VEGF-C mRNA was found in different types of thyroid disorder, including benign as well as malignant tumors (Shushanov et al., *Int. J. Cancer*, 2000, 86, 47-52).

20 Furthermore, in one half of 36 human tumor tissues tested, VEGF-C mRNA was detected and, notably, all lymphomas contained low levels of VEGF-C mRNA, possibly reflecting the cell-specific pattern of expression of the gene in the corresponding normal cells. VEGF-C might also be involved in
25 metastasis, as upregulated expression of VEGF-C was detected in prostatic adenocarcinoma PC-3 cells and was correlated with tumor dissemination into lymph nodes. Thus, cancer cells that produce VEGF-C and metastasize to lymph nodes may have a growth advantage because of their capacity to stimulate both
30 hematic and lymphatic endothelia (Salven et al., *Am. J. Pathol.*, 1998, 153, 103-108).

-7-

Expression of VEGF-C and VEGFR-3/FLT4 are associated with angiogenesis and perhaps lymphangiogenesis in breast cancer. Affinity purified, polyclonal antibodies were produced against VEGF-C and used to stain and demonstrate the presence of VEGF-C in eight intraductal carcinoma and 5 invasive breast carcinoma cell lines. The presence of VEGF-C in intraductal carcinoma cells as well as the VEGFR-3 positive capillaries surrounding the affected ducts suggest that VEGF-C secreted by the cancer cells acts predominantly 10 as an angiogenic growth factor for blood vessels, is involved in paracrine signaling between cancer cells and the endothelium, and may be involved in modifying the permeabilities of both blood and lymphatic vessels during tumor metastasis into the axillary lymph nodes (Valtola et 15 al., *Am. J. Pathol.*, 1999, 154, 1381-1390).

Immunohistochemistry was used to assess VEGF-C expression in 228 primarily surgically treated cases of postmenopausal uterine endometrial carcinoma and evaluate the correlation with established histopathologic risk factors and 20 clinical outcome, such as vascular invasion, depth of invasion (myometrial vs. serosal-parametrial), lymphatic vessel invasion, lymph node metastasis, and 5- and 10-year disease-free survival rates. VEGF-C expression was highly correlated with all of these these histopathologic features, 25 which bore prognostic significance and are thus predictive of the clinical outcome of postmenopausal uterine endometrial carcinoma (Hirai et al., *Gynecol. Oncol.*, 2001, 80, 181-188).

Because angiogenesis is suggested to be a rate limiting step in tumor development, and because of the selective 30 nature of the VEGF-C ligand, which potently induces lymphangiogenesis as well as angiogenesis, VEGF-C is an ideal target for therapeutic modulation of growth factor signaling

-8-

in pathologic conditions such as tumor growth, metastasis, and diabetic retinopathy.

Currently, there are no known therapeutic agents which effectively inhibit the synthesis of VEGF-C. To date,
5 investigative strategies aimed at modulating VEGF-C function have involved the use of a VEGF-C mutant protein and a soluble VEGFR-3 fusion protein inhibitor.

A selective agonist mutant has been described, which binds and activates only VEGFR-3 but neither binds nor
10 activates VEGFR-2, and neither induces vascular permeability *in vivo* nor stimulates migration of endothelial cells in culture (Joukov et al., *J. Biol. Chem.*, 1998, 273, 6599-6602).

VEGF-C overexpression in MCF-7 mammary tumors strongly
15 and specifically induces the growth of tumor-associated lymphatic vessels and significantly increases tumor growth. These effects of VEGF-C are inhibited by a soluble VEGFR-3 fusion protein, suggesting that the VEGF-C facilitated tumor metastasis via the lymphatic vessels can be inhibited by
20 blocking the interaction between VEGF-C and its receptor (Karpanen et al., *Cancer Res.*, 2001, 61, 1786-1790).

Consequently, there remains a long felt need for agents capable of effectively inhibiting VEGF-C function.

Disclosed and claimed in US Patent 6,130,071 is a
25 purified and isolated mutant encoding a VEGF-C protein which is capable of binding to the VEGFR-3/FLT4 receptor but fails to bind to human KDR/VEGFR-2, a vector comprising a nucleic acid encoding said mutant, a host cell transformed or transfected with said nucleic acid, sequences complementary
30 to and which hybridize with VEGF-C, and a method of making a polypeptide that binds to VEGFR-3/FLT4 (Alitalo and Joukov, 2000).

Disclosed and claimed in US Patents 6,221,839 and 6,245,530 are a purified and isolated VEGF-C polypeptide, cDNAs and vectors encoding said polypeptide, a pharmaceutical composition comprising said polypeptide, a method of
5 modulating the activity of human VEGFR-3/FLT4 receptor tyrosine kinase comprising administering to a person in need said pharmaceutical composition, and generally disclosed are inhibitors, including antibodies, antisense oligonucleotides, and peptides which block the VEGFR-3 receptor, which may be
10 used to control endothelial cell proliferation and lymphangiomas, to arrest metastatic growth, or to control other aspects of endothelial cell expression and growth (Alitalo and Joukov, 2001; Alitalo and Joukov, 2001).

Disclosed and claimed in PCT Publication WO 99/08522 is
15 a method of stimulating angiogenesis in endothelial cells comprising co-administering to said cells at least two cytokines selected from the group consisting of VEGF, VEGF-B, VEGF-C, and FGF, a method of inhibiting endothelial cell permeation, invasion and/or metastasis in a patient
20 comprising administering to said patient an effective endothelial cell proliferation inhibiting amount of a VEGF-C antagonist and a method of modulating angiogenic activity of endothelial cells comprising transfecting or transforming the cells with a vector containing an antisense nucleic acid for
25 VEGF-C (Pepper et al., 1999).

Disclosed and claimed in PCT Publication WO 01/52904 is a composition comprising one or more antisense oligonucleotides directed against vascular endothelial growth factor (VEGF) wherein said antisense oligonucleotide inhibits
30 proliferation-of cells exhibiting autocrine VEGF activity, wherein said one or more antisense oligonucleotide is selected from a group of fragments encoding a portion of the VEGF protein. One antisense oligonucleotide resulting in

-10-

decreased VEGF-C mRNA levels is disclosed (Gill and Masood, 2001).

Disclosed and claimed in PCT Publication WO 00/45835 is a method for treating injury or degeneration of photoreceptors, comprising administering to a subject suffering from such photoreceptor injury or degeneration a therapeutically effective amount of VEGF-C, and a composition comprising an isolated antibody, wherein said antibody specially binds the VEGF-C polypeptide. Also generally disclosed as a potential VEGF-C antagonist is an antisense DNA oligonucleotide as well as small molecule inhibitors (Rosen et al., 2000).

Antisense technology is emerging as an effective means for reducing the expression of specific gene products and may therefore prove to be uniquely useful in a number of therapeutic, diagnostic, and research applications for the modulation of VEGF-C expression.

SUMMARY OF THE INVENTION

The present invention provides compositions and methods for modulating VEGF-C expression, including modulation of the alternatively spliced forms of VEGF-C. The present invention is directed to compounds, particularly antisense oligonucleotides, which are targeted to a nucleic acid encoding VEGF-C, and which modulate the expression of VEGF-C. Pharmaceutical and other compositions comprising the compounds of the invention are also provided. Further provided are methods of modulating the expression of VEGF-C in cells or tissues comprising contacting said cells or tissues with one or more of the antisense compounds or compositions of the invention. Further provided are methods of treating an animal, particularly a human, suspected of having or being prone to a disease or condition associated

-11-

with expression of VEGF-C by administering a therapeutically or prophylactically effective amount of one or more of the antisense compounds or compositions of the invention.

5 DETAILED DESCRIPTION OF THE INVENTION

The present invention employs oligomeric compounds, particularly antisense oligonucleotides, for use in modulating the function of nucleic acid molecules encoding VEGF-C, ultimately modulating the amount of VEGF-C produced.

10 This is accomplished by providing antisense compounds which specifically hybridize with one or more nucleic acids encoding VEGF-C. As used herein, the terms "target nucleic acid" and "nucleic acid encoding VEGF-C" encompass DNA encoding VEGF-C, RNA (including pre-mRNA and mRNA)

15 transcribed from such DNA, and also cDNA derived from such RNA. The specific hybridization of an oligomeric compound with its target nucleic acid interferes with the normal function of the nucleic acid. This modulation of function of a target nucleic acid by compounds which specifically

20 hybridize to it is generally referred to as "antisense". The functions of DNA to be interfered with include replication and transcription. The functions of RNA to be interfered with include all vital functions such as, for example, translocation of the RNA to the site of protein translation,

25 translocation of the RNA to sites within the cell which are distant from the site of RNA synthesis, translation of protein from the RNA, splicing of the RNA to yield one or more mRNA species, and catalytic activity which may be engaged in or facilitated by the RNA. The overall effect of

30 such interference with target nucleic acid function is modulation of the expression of VEGF-C. In the context of the present invention, "modulation" means either an increase (stimulation) or a decrease (inhibition) in the expression of

-12-

a gene. In the context of the present invention, inhibition is the preferred form of modulation of gene expression and mRNA is a preferred target.

It is preferred to target specific nucleic acids for antisense. "Targeting" an antisense compound to a particular nucleic acid, in the context of this invention, is a multistep process. The process usually begins with the identification of a nucleic acid sequence whose function is to be modulated. This may be, for example, a cellular gene (or mRNA transcribed from the gene) whose expression is associated with a particular disorder or disease state, or a nucleic acid molecule from an infectious agent. In the present invention, the target is a nucleic acid molecule encoding VEGF-C. The targeting process also includes determination of a site or sites within this gene for the antisense interaction to occur such that the desired effect, e.g., detection or modulation of expression of the protein, will result. Within the context of the present invention, a preferred intragenic site is the region encompassing the translation initiation or termination codon of the open reading frame (ORF) of the gene. Since, as is known in the art, the translation initiation codon is typically 5'-AUG (in transcribed mRNA molecules; 5'-ATG in the corresponding DNA molecule), the translation initiation codon is also referred to as the "AUG codon," the "start codon" or the "AUG start codon". A minority of genes has a translation initiation codon having the RNA sequence 5'-GUG, 5'-UUG or 5'-CUG, and 5'-AUA, 5'-ACG and 5'-CUG have been shown to function *in vivo*. Thus, the terms "translation initiation codon" and "start codon" can encompass many codon sequences, even though the initiator amino acid in each instance is typically methionine (in eukaryotes) or formylmethionine (in prokaryotes). It is also known in the art that eukaryotic

-13-

and prokaryotic genes may have two or more alternative start codons, any one of which may be preferentially utilized for translation initiation in a particular cell type or tissue, or under a particular set of conditions. In the context of the invention, "start codon" and "translation initiation codon" refer to the codon or codons that are used *in vivo* to initiate translation of an mRNA molecule transcribed from a gene encoding VEGF-C, regardless of the sequence(s) of such codons.

10 It is also known in the art that a translation termination codon (or "stop codon") of a gene may have one of three sequences, i.e., 5'-UAA, 5'-UAG and 5'-UGA (the corresponding DNA sequences are 5'-TAA, 5'-TAG and 5'-TGA, respectively). The terms "start codon region" and
15 "translation initiation codon region" refer to a portion of such an mRNA or gene that encompasses from about 25 to about 50 contiguous nucleotides in either direction (i.e., 5' or 3') from a translation initiation codon. Similarly, the terms "stop codon region" and "translation termination codon
20 region" refer to a portion of such an mRNA or gene that encompasses from about 25 to about 50 contiguous nucleotides in either direction (i.e., 5' or 3') from a translation termination codon.

The open reading frame (ORF) or "coding region," which
25 is known in the art to refer to the region between the translation initiation codon and the translation termination codon, is also a region which may be targeted effectively. Other target regions include the 5' untranslated region (5'UTR), known in the art to refer to the portion of an mRNA
30 in the 5' direction from the translation initiation codon, and thus including nucleotides between the 5' cap site and the translation initiation codon of an mRNA or corresponding nucleotides on the gene, and the 3' untranslated region

-14-

(3'UTR), known in the art to refer to the portion of an mRNA in the 3' direction from the translation termination codon, and thus including nucleotides between the translation termination codon and 3' end of an mRNA or corresponding nucleotides on the gene. The 5' cap of an mRNA comprises an N7-methylated guanosine residue joined to the 5'-most residue of the mRNA via a 5'-5' triphosphate linkage. The 5' cap region of an mRNA is considered to include the 5' cap structure itself as well as the first 50 nucleotides adjacent to the cap. The 5' cap region may also be a preferred target region.

Although some eukaryotic mRNA transcripts are directly translated, many contain one or more regions, known as "introns," which are excised from a transcript before it is translated. The remaining (and therefore translated) regions are known as "exons" and are spliced together to form a continuous mRNA sequence. mRNA splice sites, i.e., intron-exon junctions, may also be preferred target regions, and are particularly useful in situations where aberrant splicing is implicated in disease, or where an overproduction of a particular mRNA splice product is implicated in disease. Aberrant fusion junctions due to rearrangements or deletions are also preferred targets. mRNA transcripts produced via the process of splicing of two (or more) mRNAs from different gene sources are known as "fusion transcripts". It has also been found that introns can be effective, and therefore preferred, target regions for antisense compounds targeted, for example, to DNA or pre-mRNA.

It is also known in the art that alternative RNA transcripts can be produced from the same genomic region of DNA. These alternative transcripts are generally known as "variants". More specifically, "pre-mRNA variants" are transcripts produced from the same genomic DNA that differ

-15-

from other transcripts produced from the same genomic DNA in either their start or stop position and contain both intronic and extronic regions.

Upon excision of one or more exon or intron regions or portions thereof during splicing, pre-mRNA variants produce smaller "mRNA variants". Consequently, mRNA variants are processed pre-mRNA variants and each unique pre-mRNA variant must always produce a unique mRNA variant as a result of splicing. These mRNA variants are also known as "alternative splice variants". If no splicing of the pre-mRNA variant occurs then the pre-mRNA variant is identical to the mRNA variant.

It is also known in the art that variants can be produced through the use of alternative signals to start or stop transcription and that pre-mRNAs and mRNAs can possess more than one start codon or stop codon. Variants that originate from a pre-mRNA or mRNA that use alternative start codons are known as "alternative start variants" of that pre-mRNA or mRNA. Those transcripts that use an alternative stop codon are known as "alternative stop variants" of that pre-mRNA or mRNA. One specific type of alternative stop variant is the "polyA variant" in which the multiple transcripts produced result from the alternative selection of one of the "polyA stop signals" by the transcription machinery, thereby producing transcripts that terminate at unique polyA sites.

Once one or more target sites have been identified, oligonucleotides are chosen which are sufficiently complementary to the target, i.e., hybridize sufficiently well and with sufficient specificity, to give the desired effect.

In the context of this invention, "hybridization" means hydrogen bonding, which may be Watson-Crick, Hoogsteen or reversed Hoogsteen hydrogen bonding, between complementary

-16-

nucleoside or nucleotide bases. For example, adenine and thymine are complementary nucleobases which pair through the formation of hydrogen bonds. "Complementary," as used herein, refers to the capacity for precise pairing between two nucleotides. For example, if a nucleotide at a certain position of an oligonucleotide is capable of hydrogen bonding with a nucleotide at the same position of a DNA or RNA molecule, then the oligonucleotide and the DNA or RNA are considered to be complementary to each other at that position. The oligonucleotide and the DNA or RNA are complementary to each other when a sufficient number of corresponding positions in each molecule are occupied by nucleotides which can hydrogen bond with each other. Thus, "specifically hybridizable" and "complementary" are terms which are used to indicate a sufficient degree of complementarity or precise pairing such that stable and specific binding occurs between the oligonucleotide and the DNA or RNA target. It is understood in the art that the sequence of an antisense compound need not be 100% complementary to that of its target nucleic acid to be specifically hybridizable.

An antisense compound is specifically hybridizable when binding of the compound to the target DNA or RNA molecule interferes with the normal function of the target DNA or RNA to cause a loss of activity, and there is a sufficient degree of complementarity to avoid non-specific binding of the antisense compound to non-target sequences under conditions in which specific binding is desired, i.e., under physiological conditions in the case of in vivo assays or therapeutic treatment, and in the case of in vitro assays, under conditions in which the assays are performed. It is preferred that the antisense compounds of the present invention comprise at least 80% sequence complementarity to a

-17-

target region within the target nucleic acid, moreover that they comprise 90% sequence complementarity and even more comprise 95% sequence complementarity to the target region within the target nucleic acid sequence to which they are targeted. For example, an antisense compound in which 18 of 20 nucleobases of the antisense compound are complementary, and would therefore specifically hybridize, to a target region would represent 90 percent complementarity. Percent complementarity of an antisense compound with a region of a target nucleic acid can be determined routinely using basic local alignment search tools (BLAST programs) (Altschul et al., *J. Mol. Biol.*, 1990, 215, 403-410; Zhang and Madden, *Genome Res.*, 1997, 7, 649-656).

Antisense and other compounds of the invention, which hybridize to the target and inhibit expression of the target, are identified through experimentation, and representative sequences of these compounds are hereinbelow identified as preferred embodiments of the invention. The sites to which these preferred antisense compounds are specifically hybridizable are hereinbelow referred to as "preferred target regions" and are therefore preferred sites for targeting. As used herein the term "preferred target region" is defined as at least an 8-nucleobase portion of a target region to which an active antisense compound is targeted. While not wishing to be bound by theory, it is presently believed that these target regions represent regions of the target nucleic acid which are accessible for hybridization.

While the specific sequences of particular preferred target regions are set forth below, one of skill in the art will recognize that these serve to illustrate and describe particular embodiments within the scope of the present invention. Additional preferred target regions may be identified by one having ordinary skill.

-18-

Target regions 8-80 nucleobases in length comprising a stretch of at least eight (8) consecutive nucleobases selected from within the illustrative preferred target regions are considered to be suitable preferred target regions as well.

Exemplary good preferred target regions include DNA or RNA sequences that comprise at least the 8 consecutive nucleobases from the 5'-terminus of one of the illustrative preferred target regions (the remaining nucleobases being a consecutive stretch of the same DNA or RNA beginning immediately upstream of the 5'-terminus of the target region and continuing until the DNA or RNA contains about 8 to about 80 nucleobases). Similarly good preferred target regions are represented by DNA or RNA sequences that comprise at least the 8 consecutive nucleobases from the 3'-terminus of one of the illustrative preferred target regions (the remaining nucleobases being a consecutive stretch of the same DNA or RNA beginning immediately downstream of the 3'-terminus of the target region and continuing until the DNA or RNA contains about 8 to about 80 nucleobases). One having skill in the art, once armed with the empirically-derived preferred target regions illustrated herein will be able, without undue experimentation, to identify further preferred target regions. In addition, one having ordinary skill in the art will also be able to identify additional compounds, including oligonucleotide probes and primers that specifically hybridize to these preferred target regions using techniques available to the ordinary practitioner in the art.

Antisense compounds are commonly used as research reagents and diagnostics. For example, antisense oligonucleotides, which are able to inhibit gene expression with exquisite specificity, are often used by those of ordinary skill to elucidate the function of particular genes.

-19-

Antisense compounds are also used, for example, to distinguish between functions of various members of a biological pathway. Antisense modulation has, therefore, been harnessed for research use.

5 For use in kits and diagnostics, the antisense compounds of the present invention, either alone or in combination with other antisense compounds or therapeutics, can be used as tools in differential and/or combinatorial analyses to elucidate expression patterns of a portion or the entire
10 complement of genes expressed within cells and tissues.

Expression patterns within cells or tissues treated with one or more antisense compounds are compared to control cells or tissues not treated with antisense compounds and the patterns produced are analyzed for differential levels of
15 gene expression as they pertain, for example, to disease association, signaling pathway, cellular localization, expression level, size, structure or function of the genes examined. These analyses can be performed on stimulated or unstimulated cells and in the presence or absence of other
20 compounds which affect expression patterns.

Examples of methods of gene expression analysis known in the art include DNA arrays or microarrays (Brazma and Vilo, *FEBS Lett.*, 2000, 480, 17-24; Celis, et al., *FEBS Lett.*, 2000, 480, 2-16), SAGE (serial analysis of gene
25 expression) (Madden, et al., *Drug Discov. Today*, 2000, 5, 415-425), READS (restriction enzyme amplification of digested cDNAs) (Prashar and Weissman, *Methods Enzymol.*, 1999, 303, 258-72), TOGA (total gene expression analysis) (Sutcliffe, et al., *Proc. Natl. Acad. Sci. U. S. A.*, 2000, 97, 1976-81),
30 protein arrays and proteomics (Celis, et al., *FEBS Lett.*, 2000, 480, 2-16; Jungblut, et al., *Electrophoresis*, 1999, 20, 2100-10), expressed sequence tag (EST) sequencing (Celis, et al., *FEBS Lett.*, 2000, 480, 2-16; Larsson, et al., J.

-20-

Biotechnol., 2000, 80, 143-57), subtractive RNA fingerprinting (SuRF) (Fuchs, et al., *Anal. Biochem.*, 2000, 286, 91-98; Larson, et al., *Cytometry*, 2000, 41, 203-208), subtractive cloning, differential display (DD) (Jurecic and Belmont, *Curr. Opin. Microbiol.*, 2000, 3, 316-21), comparative genomic hybridization (Carulli, et al., *J. Cell Biochem. Suppl.*, 1998, 31, 286-96), FISH (fluorescent *in situ* hybridization) techniques (Going and Gusterson, *Eur. J. Cancer*, 1999, 35, 1895-904) and mass spectrometry methods (reviewed in To, *Comb. Chem. High Throughput Screen*, 2000, 3, 235-41).

The specificity and sensitivity of antisense is also harnessed by those of skill in the art for therapeutic uses. Antisense oligonucleotides have been employed as therapeutic moieties in the treatment of disease states in animals and man. Antisense oligonucleotide drugs, including ribozymes, have been safely and effectively administered to humans and numerous clinical trials are presently underway. It is thus established that oligonucleotides can be useful therapeutic modalities that can be configured to be useful in treatment regimes for treatment of cells, tissues and animals, especially humans.

In the context of this invention, the term "oligonucleotide" refers to an oligomer or polymer of ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) or mimetics thereof. This term includes oligonucleotides composed of naturally-occurring nucleobases, sugars and covalent internucleoside (backbone) linkages as well as oligonucleotides having non-naturally-occurring portions which function similarly. Such modified or substituted oligonucleotides are often preferred over native forms because of desirable properties such as, for example, enhanced cellular uptake, enhanced affinity for nucleic acid

-21-

target and increased stability in the presence of nucleases.

While antisense oligonucleotides are a preferred form of antisense compound, the present invention comprehends other oligomeric antisense compounds, including but not limited to oligonucleotide mimetics such as are described below. The antisense compounds in accordance with this invention preferably comprise from about 8 to about 80 nucleobases (i.e. from about 8 to about 80 linked nucleosides). Particularly preferred antisense compounds are antisense oligonucleotides from about 8 to about 50 nucleobases, even more preferably those comprising from about 12 to about 30 nucleobases. Antisense compounds include ribozymes, external guide sequence (EGS) oligonucleotides (oligozymes), and other short catalytic RNAs or catalytic oligonucleotides which hybridize to the target nucleic acid and modulate its expression.

Antisense compounds 8-80 nucleobases in length comprising a stretch of at least eight (8) consecutive nucleobases selected from within the illustrative antisense compounds are considered to be suitable antisense compounds as well.

Exemplary preferred antisense compounds include DNA or RNA sequences that comprise at least the 8 consecutive nucleobases from the 5'-terminus of one of the illustrative preferred antisense compounds (the remaining nucleobases being a consecutive stretch of the same DNA or RNA beginning immediately upstream of the 5'-terminus of the antisense compound which is specifically hybridizable to the target nucleic acid and continuing until the DNA or RNA contains about 8 to about 80 nucleobases). Similarly preferred antisense compounds are represented by DNA or RNA sequences that comprise at least the 8 consecutive nucleobases from the 3'-terminus of one of the illustrative preferred antisense

-22-

compounds (the remaining nucleobases being a consecutive stretch of the same DNA or RNA beginning immediately downstream of the 3'-terminus of the antisense compound which is specifically hybridizable to the target nucleic acid and continuing until the DNA or RNA contains about 8 to about 80 nucleobases). One having skill in the art, once armed with the empirically-derived preferred antisense compounds illustrated herein will be able, without undue experimentation, to identify further preferred antisense compounds.

Antisense and other compounds of the invention, which hybridize to the target and inhibit expression of the target, are identified through experimentation, and representative sequences of these compounds are herein identified as preferred embodiments of the invention. While specific sequences of the antisense compounds are set forth herein, one of skill in the art will recognize that these serve to illustrate and describe particular embodiments within the scope of the present invention. Additional preferred antisense compounds may be identified by one having ordinary skill.

As is known in the art, a nucleoside is a base-sugar combination. The base portion of the nucleoside is normally a heterocyclic base. The two most common classes of such heterocyclic bases are the purines and the pyrimidines. Nucleotides are nucleosides that further include a phosphate group covalently linked to the sugar portion of the nucleoside. For those nucleosides that include a pentofuranosyl sugar, the phosphate group can be linked to either the 2', 3' or 5' hydroxyl moiety of the sugar. In forming oligonucleotides, the phosphate groups covalently link adjacent nucleosides to one another to form a linear polymeric compound. In turn, the respective ends of this

-23-

linear polymeric structure can be further joined to form a circular structure, however, open linear structures are generally preferred. In addition, linear structures may also have internal nucleobase complementarity and may therefore
5 fold in a manner as to produce a double stranded structure. Within the oligonucleotide structure, the phosphate groups are commonly referred to as forming the internucleoside backbone of the oligonucleotide. The normal linkage or backbone of RNA and DNA is a 3' to 5' phosphodiester linkage.

10 Specific examples of preferred antisense compounds useful in this invention include oligonucleotides containing modified backbones or non-natural internucleoside linkages. As defined in this specification, oligonucleotides having modified backbones include those that retain a phosphorus
15 atom in the backbone and those that do not have a phosphorus atom in the backbone. For the purposes of this specification, and as sometimes referenced in the art, modified oligonucleotides that do not have a phosphorus atom in their internucleoside backbone can also be considered to
20 be oligonucleosides.

Preferred modified oligonucleotide backbones include, for example, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates including 3'-
25 alkylene phosphonates, 5'-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, selenophosphates and borano-
30 phosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity wherein one or more internucleotide linkages is a 3' to 3', 5' to 5' or 2' to 2' linkage. Preferred oligonucleotides having inverted

-24-

polarity comprise a single 3' to 3' linkage at the 3'-most internucleotide linkage i.e. a single inverted nucleoside residue which may be abasic (the nucleobase is missing or has a hydroxyl group in place thereof). Various salts, mixed salts and free acid forms are also included.

Representative United States patents that teach the preparation of the above phosphorus-containing linkages include, but are not limited to, U.S.: 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; 5,194,599; 5,565,555; 5,527,899; 5,721,218; 5,672,697 and 5,625,050, certain of which are commonly owned with this application, and each of which is herein incorporated by reference.

Preferred modified oligonucleotide backbones that do not include a phosphorus atom therein have backbones that are formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom and alkyl or cycloalkyl internucleoside linkages, or one or more short chain heteroatomic or heterocyclic internucleoside linkages. These include those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfoxide and sulfone backbones; formacetyl and thioformacetyl backbones; methylene formacetyl and thioformacetyl backbones; riboacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH₂ component parts.

Representative United States patents that teach the preparation of the above oligonucleosides include, but are

-25-

not limited to, U.S.: 5,034,506; 5,166,315; 5,185,444;
5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564;
5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677;
5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289;
5 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070;
5,663,312; 5,633,360; 5,677,437; 5,792,608; 5,646,269 and
5,677,439, certain of which are commonly owned with this
application, and each of which is herein incorporated by
reference.

10 In other preferred oligonucleotide mimetics, both the
sugar and the internucleoside linkage, i.e., the backbone, of
the nucleotide units are replaced with novel groups. The
base units are maintained for hybridization with an
appropriate nucleic acid target compound. One such
15 oligomeric compound, an oligonucleotide mimetic that has been
shown to have excellent hybridization properties, is referred
to as a peptide nucleic acid (PNA). In PNA compounds, the
sugar-backbone of an oligonucleotide is replaced with an
amide containing backbone, in particular an aminoethylglycine
20 backbone. The nucleobases are retained and are bound
directly or indirectly to aza nitrogen atoms of the amide
portion of the backbone. Representative United States
patents that teach the preparation of PNA compounds include,
but are not limited to, U.S.: 5,539,082; 5,714,331; and
25 5,719,262, each of which is herein incorporated by reference.
Further teaching of PNA compounds can be found in Nielsen et
al., *Science*, 1991, 254, 1497-1500.

Most preferred embodiments of the invention are
oligonucleotides with phosphorothioate backbones and
30 oligonucleosides with heteroatom backbones, and in particular
-CH₂-NH-O-CH₂-, -CH₂-N(CH₃)-O-CH₂- [known as a methylene
(methylimino) or MMI backbone], -CH₂-O-N(CH₃)-CH₂-, -CH₂-
N(CH₃)-N(CH₃)-CH₂- and -O-N(CH₃)-CH₂-CH₂- [wherein the native

-26-

phosphodiester backbone is represented as -O-P-O-CH₂-] of the above referenced U.S. patent 5,489,677, and the amide backbones of the above referenced U.S. patent 5,602,240. Also preferred are oligonucleotides having morpholino backbone structures of the above-referenced U.S. patent 5,034,506.

Modified oligonucleotides may also contain one or more substituted sugar moieties. Preferred oligonucleotides comprise one of the following at the 2' position: OH; F; O-, S-, or N-alkyl; O-, S-, or N-alkenyl; O-, S- or N-alkynyl; or O-alkyl-O-alkyl, wherein the alkyl, alkenyl and alkynyl may be substituted or unsubstituted C₁ to C₁₀ alkyl or C₂ to C₁₀ alkenyl and alkynyl. Particularly preferred are O[(CH₂)_nO]_mCH₃, O(CH₂)_nOCH₃, O(CH₂)_nNH₂, O(CH₂)_nCH₃, O(CH₂)_nONH₂, and O(CH₂)_nON[(CH₂)_nCH₃]₂, where n and m are from 1 to about 10. Other preferred oligonucleotides comprise one of the following at the 2' position: C₁ to C₁₀ lower alkyl, substituted lower alkyl, alkenyl, alkynyl, alkaryl, aralkyl, O-alkaryl or O-aralkyl, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, an RNA cleaving group, a reporter group, an intercalator, a group for improving the pharmacokinetic properties of an oligonucleotide, or a group for improving the pharmacodynamic properties of an oligonucleotide, and other substituents having similar properties. A preferred modification includes 2'-methoxyethoxy (2'-O-CH₂CH₂OCH₃, also known as 2'-O-(2-methoxyethyl) or 2'-MOE) (Martin *et al.*, *Helv. Chim. Acta*, 1995, 78, 486-504) i.e., an alkoxyalkoxy group. A further preferred modification includes 2'-dimethylaminooxyethoxy, i.e., a O(CH₂)₂ON(CH₃)₂ group, also known as 2'-DMAOE, as described in examples hereinbelow, and 2'-dimethylaminoethoxyethoxy (also known in the art as 2'-O-

-27-

dimethyl-amino-ethoxy-ethyl or 2'-DMAEOE), i.e., 2'-O-CH₂-O-CH₂-N(CH₃)₂, also described in examples hereinbelow.

Other preferred modifications include 2'-methoxy (2'-O-CH₃), 2'-aminopropoxy (2'-OCH₂CH₂CH₂NH₂), 2'-allyl (2'-CH₂-CH=CH₂), 2'-O-allyl (2'-O-CH₂-CH=CH₂) and 2'-fluoro (2'-F). The 2'-modification may be in the arabino (up) position or ribo (down) position. A preferred 2'-arabino modification is 2'-F. Similar modifications may also be made at other positions on the oligonucleotide, particularly the 3' position of the sugar on the 3' terminal nucleotide or in 2'-5' linked oligonucleotides and the 5' position of 5' terminal nucleotide. Oligonucleotides may also have sugar mimetics such as cyclobutyl moieties in place of the pentofuranosyl sugar. Representative United States patents that teach the preparation of such modified sugar structures include, but are not limited to, U.S.: 4,981,957; 5,118,800; 5,319,080; 5,359,044; 5,393,878; 5,446,137; 5,466,786; 5,514,785; 5,519,134; 5,567,811; 5,576,427; 5,591,722; 5,597,909; 5,610,300; 5,627,053; 5,639,873; 5,646,265; 5,658,873; 5,670,633; 5,792,747; and 5,700,920, certain of which are commonly owned with the instant application, and each of which is herein incorporated by reference in its entirety.

A further preferred modification includes Locked Nucleic Acids (LNAs) in which the 2'-hydroxyl group is linked to the 3' or 4' carbon atom of the sugar ring thereby forming a bicyclic sugar moiety. The linkage is preferably a methylene (-CH₂)_n group bridging the 2' oxygen atom and the 4' carbon atom wherein n is 1 or 2. LNAs and preparation thereof are described in WO 98/39352 and WO 99/14226.

Oligonucleotides may also include nucleobase (often referred to in the art simply as "base") modifications or substitutions. As used herein, "unmodified" or "natural" nucleobases include the purine bases adenine (A) and guanine

-28-

(G), and the pyrimidine bases thymine (T), cytosine (C) and uracil (U). Modified nucleobases include other synthetic and natural nucleobases such as 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-halouracil and cytosine, 5-propynyl (-C≡C-CH₃) uracil and cytosine and other alkynyl derivatives of pyrimidine bases, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 2-F-adenine, 2-amino-adenine, 8-azaguanine and 8-azaadenine, 7-deazaguanine and 7-deazaadenine and 3-deazaguanine and 3-deazaadenine. Further modified nucleobases include tricyclic pyrimidines such as phenoxazine cytidine (1H-pyrimido[5,4-b][1,4]benzoxazin-2(3H)-one), phenothiazine cytidine (1H-pyrimido[5,4-b][1,4]benzothiazin-2(3H)-one), G-clamps such as a substituted phenoxazine cytidine (e.g. 9-(2-aminoethoxy)-H-pyrimido[5,4-b][1,4]benzoxazin-2(3H)-one), carbazole cytidine (2H-pyrimido[4,5-b]indol-2-one), pyridoindole cytidine (H-pyrido[3',2':4,5]pyrrolo[2,3-d]pyrimidin-2-one). Modified nucleobases may also include those in which the purine or pyrimidine base is replaced with other heterocycles, for example 7-deaza-adenine, 7-deazaguanosine, 2-aminopyridine and 2-pyridone. Further nucleobases include those disclosed in United States Patent No. 3,687,808, those disclosed in *The Concise Encyclopedia Of Polymer Science And Engineering*, pages 858-859, Kroschwitz, J.I., ed. John Wiley & Sons, 1990, those disclosed by Englisch et al., *Angewandte Chemie*,

-29-

International Edition, 1991, 30, 613, and those disclosed by Sanghvi, Y.S., Chapter 15, *Antisense Research and Applications*, pages 289-302, Crooke, S.T. and Lebleu, B. , ed., CRC Press, 1993. Certain of these nucleobases are particularly useful for increasing the binding affinity of the oligomeric compounds of the invention. These include 5-substituted pyrimidines, 6-azapyrimidines and N-2, N-6 and O-6 substituted purines, including 2-aminopropyladenine, 5-propynyluracil and 5-propynylcytosine. 5-methylcytosine substitutions have been shown to increase nucleic acid duplex stability by 0.6-1.2°C (Sanghvi, Y.S., Crooke, S.T. and Lebleu, B., eds., *Antisense Research and Applications*, CRC Press, Boca Raton, 1993, pp. 276-278) and are presently preferred base substitutions, even more particularly when combined with 2'-O-methoxyethyl sugar modifications.

Representative United States patents that teach the preparation of certain of the above noted modified nucleobases as well as other modified nucleobases include, but are not limited to, the above noted U.S. 3,687,808, as well as U.S.: 4,845,205; 5,130,302; 5,134,066; 5,175,273; 5,367,066; 5,432,272; 5,457,187; 5,459,255; 5,484,908; 5,502,177; 5,525,711; 5,552,540; 5,587,469; 5,594,121, 5,596,091; 5,614,617; 5,645,985; 5,830,653; 5,763,588; 6,005,096; and 5,681,941, certain of which are commonly owned with the instant application, and each of which is herein incorporated by reference, and United States patent 5,750,692, which is commonly owned with the instant application and also herein incorporated by reference.

Another modification of the oligonucleotides of the invention involves chemically linking to the oligonucleotide one or more moieties or conjugates which enhance the activity, cellular distribution or cellular uptake of the oligonucleotide. The compounds of the invention can include

-30-

conjugate groups covalently bound to functional groups such as primary or secondary hydroxyl groups. Conjugate groups of the invention include intercalators, reporter molecules, polyamines, polyamides, polyethylene glycols, polyethers, 5 groups that enhance the pharmacodynamic properties of oligomers, and groups that enhance the pharmacokinetic properties of oligomers. Typical conjugate groups include cholesterols, lipids, phospholipids, biotin, phenazine, folate, phenanthridine, anthraquinone, acridine, fluores- 10 ceins, rhodamines, coumarins, and dyes. Groups that enhance the pharmacodynamic properties, in the context of this invention, include groups that improve oligomer uptake, enhance oligomer resistance to degradation, and/or strengthen sequence-specific hybridization with RNA. Groups that 15 enhance the pharmacokinetic properties, in the context of this invention, include groups that improve oligomer uptake, distribution, metabolism or excretion. Representative conjugate groups are disclosed in International Patent Application PCT/US92/09196, filed October 23, 1992 the entire 20 disclosure of which is incorporated herein by reference. Conjugate moieties include but are not limited to lipid moieties such as a cholesterol moiety (Letsinger *et al.*, *Proc. Natl. Acad. Sci. USA*, 1989, 86, 6553-6556), cholic acid (Manoharan *et al.*, *Bioorg. Med. Chem. Lett.*, 1994, 4, 1053- 25 1060), a thioether, e.g., hexyl-S-tritylthiol (Manoharan *et al.*, *Ann. N.Y. Acad. Sci.*, 1992, 660, 306-309; Manoharan *et al.*, *Bioorg. Med. Chem. Lett.*, 1993, 3, 2765-2770), a thiocholesterol (Oberhauser *et al.*, *Nucl. Acids Res.*, 1992, 20, 533-538), an aliphatic chain, e.g., dodecandiol or 30 undecyl residues (Saison-Behmoaras *et al.*, *EMBO J.*, 1991, 10, 1111-1118; Kabanov *et al.*, *FEBS Lett.*, 1990, 259, 327-330; Svinarchuk *et al.*, *Biochimie*, 1993, 75, 49-54), a phospholipid, e.g., di-hexadecyl-rac-glycerol or triethyl-

-31-

ammonium 1,2-di-O-hexadecyl-rac-glycero-3-H-phosphonate
(Manoharan et al., *Tetrahedron Lett.*, 1995, 36, 3651-3654;
Shea et al., *Nucl. Acids Res.*, 1990, 18, 3777-3783), a
polyamine or a polyethylene glycol chain (Manoharan et al.,
5 *Nucleosides & Nucleotides*, 1995, 14, 969-973), or adamantane
acetic acid (Manoharan et al., *Tetrahedron Lett.*, 1995, 36,
3651-3654), a palmityl moiety (Mishra et al., *Biochim.*
Biophys. Acta, 1995, 1264, 229-237), or an octadecylamine or
hexylamino-carbonyl-oxycholesterol moiety (Crooke et al., *J.*
10 *Pharmacol. Exp. Ther.*, 1996, 277, 923-937). Oligonucleotides
of the invention may also be conjugated to active drug
substances, for example, aspirin, warfarin, phenylbutazone,
ibuprofen, suprofen, fenbufen, ketoprofen, (S)-(+)-
pranoprofen, carprofen, dansylsarcosine, 2,3,5-triiodobenzoic
15 acid, flufenamic acid, folinic acid, a benzothiadiazide,
chlorothiazide, a diazepine, indomethicin, a barbiturate, a
cephalosporin, a sulfa drug, an antidiabetic, an
antibacterial or an antibiotic. Oligonucleotide-drug
conjugates and their preparation are described in United
20 States Patent Application 09/334,130 (filed June 15, 1999)
which is incorporated herein by reference in its entirety.

Representative United States patents that teach the
preparation of such oligonucleotide conjugates include, but
are not limited to, U.S.: 4,828,979; 4,948,882; 5,218,105;
25 5,525,465; 5,541,313; 5,545,730; 5,552,538; 5,578,717,
5,580,731; 5,580,731; 5,591,584; 5,109,124; 5,118,802;
5,138,045; 5,414,077; 5,486,603; 5,512,439; 5,578,718;
5,608,046; 4,587,044; 4,605,735; 4,667,025; 4,762,779;
4,789,737; 4,824,941; 4,835,263; 4,876,335; 4,904,582;
30 4,958,013; 5,082,830; 5,112,963; 5,214,136; 5,082,830;
5,112,963; 5,214,136; 5,245,022; 5,254,469; 5,258,506;
5,262,536; 5,272,250; 5,292,873; 5,317,098; 5,371,241,
5,391,723; 5,416,203, 5,451,463; 5,510,475; 5,512,667;

-32-

5,514,785; 5,565,552; 5,567,810; 5,574,142; 5,585,481;
5,587,371; 5,595,726; 5,597,696; 5,599,923; 5,599,928 and
5,688,941, certain of which are commonly owned with the
instant application, and each of which is herein incorporated
5 by reference.

It is not necessary for all positions in a given
compound to be uniformly modified, and in fact more than one
of the aforementioned modifications may be incorporated in a
single compound or even at a single nucleoside within an
10 oligonucleotide. The present invention also includes
antisense compounds which are chimeric compounds. "Chimeric"
antisense compounds or "chimeras," in the context of this
invention, are antisense compounds, particularly
oligonucleotides, which contain two or more chemically
15 distinct regions, each made up of at least one monomer unit,
i.e., a nucleotide in the case of an oligonucleotide
compound. These oligonucleotides typically contain at least
one region wherein the oligonucleotide is modified so as to
confer upon the oligonucleotide increased resistance to
20 nuclease degradation, increased cellular uptake, increased
stability and/or increased binding affinity for the target
nucleic acid. An additional region of the oligonucleotide
may serve as a substrate for enzymes capable of cleaving
RNA:DNA or RNA:RNA hybrids. By way of example, RNase H is a
25 cellular endonuclease which cleaves the RNA strand of an
RNA:DNA duplex. Activation of RNase H, therefore, results in
cleavage of the RNA target, thereby greatly enhancing the
efficiency of oligonucleotide inhibition of gene expression.
The cleavage of RNA:RNA hybrids can, in like fashion, be
30 accomplished through the actions of endoribonucleases, such
as interferon-induced RNaseL which cleaves both cellular and
viral RNA. Consequently, comparable results can often be
obtained with shorter oligonucleotides when chimeric

-33-

oligonucleotides are used, compared to phosphorothioate deoxyoligonucleotides hybridizing to the same target region. Cleavage of the RNA target can be routinely detected by gel electrophoresis and, if necessary, associated nucleic acid hybridization techniques known in the art.

Chimeric antisense compounds of the invention may be formed as composite structures of two or more oligonucleotides, modified oligonucleotides, oligonucleosides and/or oligonucleotide mimetics as described above. Such compounds have also been referred to in the art as hybrids or gapmers. Representative United States patents that teach the preparation of such hybrid structures include, but are not limited to, U.S.: 5,013,830; 5,149,797; 5,220,007; 5,256,775; 5,366,878; 5,403,711; 5,491,133; 5,565,350; 5,623,065; 5,652,355; 5,652,356; and 5,700,922, certain of which are commonly owned with the instant application, and each of which is herein incorporated by reference in its entirety.

The antisense compounds used in accordance with this invention may be conveniently and routinely made through the well-known technique of solid phase synthesis. Equipment for such synthesis is sold by several vendors including, for example, Applied Biosystems (Foster City, CA). Any other means for such synthesis known in the art may additionally or alternatively be employed. It is well known to use similar techniques to prepare oligonucleotides such as the phosphorothioates and alkylated derivatives.

The compounds of the invention may also be admixed, encapsulated, conjugated or otherwise associated with other molecules, molecule structures or mixtures of compounds, as for example, liposomes, receptor-targeted molecules, oral, rectal, topical or other formulations, for assisting in uptake, distribution and/or absorption. Representative United States patents that teach the preparation of such

uptake, distribution and/or absorption-assisting formulations include, but are not limited to, U.S.: 5,108,921; 5,354,844; 5,416,016; 5,459,127; 5,521,291; 5,543,158; 5,547,932; 5,583,020; 5,591,721; 4,426,330; 4,534,899; 5,013,556; 5 5,108,921; 5,213,804; 5,227,170; 5,264,221; 5,356,633; 5,395,619; 5,416,016; 5,417,978; 5,462,854; 5,469,854; 5,512,295; 5,527,528; 5,534,259; 5,543,152; 5,556,948; 5,580,575; and 5,595,756, each of which is herein incorporated by reference.

10 The antisense compounds of the invention encompass any pharmaceutically acceptable salts, esters, or salts of such esters, or any other compound which, upon administration to an animal, including a human, is capable of providing (directly or indirectly) the biologically active metabolite 15 or residue thereof. Accordingly, for example, the disclosure is also drawn to prodrugs and pharmaceutically acceptable salts of the compounds of the invention, pharmaceutically acceptable salts of such prodrugs, and other bioequivalents.

The term "prodrug" indicates a therapeutic agent that is 20 prepared in an inactive form that is converted to an active form (i.e., drug) within the body or cells thereof by the action of endogenous enzymes or other chemicals and/or conditions. In particular, prodrug versions of the oligonucleotides of the invention are prepared as SATE 25 [(S-acetyl-2-thioethyl) phosphate] derivatives according to the methods disclosed in WO 93/24510 to Gosselin et al., published December 9, 1993 or in WO 94/26764 and U.S. 5,770,713 to Imbach et al.

The term "pharmaceutically acceptable salts" refers to 30 physiologically and pharmaceutically acceptable salts of the compounds of the invention: i.e., salts that retain the desired biological activity of the parent compound and do not impart undesired toxicological effects thereto.

-35-

Pharmaceutically acceptable base addition salts are formed with metals or amines, such as alkali and alkaline earth metals or organic amines. Examples of metals used as cations are sodium, potassium, magnesium, calcium, and the like. Examples of suitable amines are

5 N,N'-dibenzylethylenediamine, chlorprocaine, choline, diethanolamine, dicyclohexylamine, ethylenediamine, N-methylglucamine, and procaine (see, for example, Berge et al., "Pharmaceutical Salts," *J. of Pharma Sci.*, 1977, 66, 1-

10 19). The base addition salts of said acidic compounds are prepared by contacting the free acid form with a sufficient amount of the desired base to produce the salt in the conventional manner. The free acid form may be regenerated by contacting the salt form with an acid and isolating the

15 free acid in the conventional manner. The free acid forms differ from their respective salt forms somewhat in certain physical properties such as solubility in polar solvents, but otherwise the salts are equivalent to their respective free acid for purposes of the present invention. As used herein,

20 a "pharmaceutical addition salt" includes a pharmaceutically acceptable salt of an acid form of one of the components of the compositions of the invention. These include organic or inorganic acid salts of the amines. Preferred acid salts are the hydrochlorides, acetates, salicylates, nitrates and

25 phosphates. Other suitable pharmaceutically acceptable salts are well known to those skilled in the art and include basic salts of a variety of inorganic and organic acids, such as, for example, with inorganic acids, such as for example hydrochloric acid, hydrobromic acid, sulfuric acid or

30 phosphoric acid; with organic carboxylic, sulfonic, sulfo or phospho acids or N-substituted sulfamic acids, for example acetic acid, propionic acid, glycolic acid, succinic acid, maleic acid, hydroxymaleic acid, methylmaleic acid, fumaric

-36-

acid, malic acid, tartaric acid, lactic acid, oxalic acid, gluconic acid, glucaric acid, glucuronic acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, salicylic acid, 4-aminosalicylic acid, 2-phenoxybenzoic acid,
5 2-acetoxybenzoic acid, embonic acid, nicotinic acid or isonicotinic acid; and with amino acids, such as the 20 alpha-amino acids involved in the synthesis of proteins in nature, for example glutamic acid or aspartic acid, and also with phenylacetic acid, methanesulfonic acid, ethanesulfonic
10 acid, 2-hydroxyethanesulfonic acid, ethane-1,2-disulfonic acid, benzenesulfonic acid, 4-methylbenzenesulfonic acid, naphthalene-2-sulfonic acid, naphthalene-1,5-disulfonic acid, 2- or 3-phosphoglycerate, glucose-6-phosphate, N-cyclohexylsulfamic acid (with the formation of cyclamates),
15 or with other acid organic compounds, such as ascorbic acid. Pharmaceutically acceptable salts of compounds may also be prepared with a pharmaceutically acceptable cation. Suitable pharmaceutically acceptable cations are well known to those skilled in the art and include alkaline, alkaline earth,
20 ammonium and quaternary ammonium cations. Carbonates or hydrogen carbonates are also possible.

For oligonucleotides, preferred examples of pharmaceutically acceptable salts include but are not limited to (a) salts formed with cations such as sodium, potassium,
25 ammonium, magnesium, calcium, polyamines such as spermine and spermidine, etc.; (b) acid addition salts formed with inorganic acids, for example hydrochloric acid, hydrobromic acid, sulfuric acid, phosphoric acid, nitric acid and the like; (c) salts formed with organic acids such as, for
30 example, acetic acid, oxalic acid, tartaric acid, succinic acid, maleic acid, fumaric acid, gluconic acid, citric acid, malic acid, ascorbic acid, benzoic acid, tannic acid, palmitic acid, alginic acid, polyglutamic acid,

naphthalenesulfonic acid, methanesulfonic acid, p-toluenesulfonic acid, naphthalenedisulfonic acid, polygalacturonic acid, and the like; and (d) salts formed from elemental anions such as chlorine, bromine, and iodine.

5 The antisense compounds of the present invention can be utilized for diagnostics, therapeutics, prophylaxis and as research reagents and kits. For therapeutics, an animal, preferably a human, suspected of having a disease or disorder which can be treated by modulating the expression of VEGF-C,
10 is treated by administering antisense compounds in accordance with this invention. The compounds of the invention can be utilized in pharmaceutical compositions by adding an effective amount of an antisense compound to a suitable pharmaceutically acceptable diluent or carrier. Use of the
15 antisense compounds and methods of the invention may also be useful prophylactically, e.g., to prevent or delay infection, inflammation or tumor formation, for example.

 The antisense compounds of the invention are useful for research and diagnostics, because these compounds hybridize
20 to nucleic acids encoding VEGF-C, enabling sandwich and other assays to easily be constructed to exploit this fact. Hybridization of the antisense oligonucleotides of the invention with a nucleic acid encoding VEGF-C can be detected by means known in the art. Such means may include
25 conjugation of an enzyme to the oligonucleotide, radiolabelling of the oligonucleotide or any other suitable detection means. Kits using such detection means for detecting the level of VEGF-C in a sample may also be prepared.

30 The present invention also includes pharmaceutical compositions and formulations which include the antisense compounds of the invention. The pharmaceutical compositions of the present invention may be administered in a number of

-38-

ways depending upon whether local or systemic treatment is desired and upon the area to be treated. Administration may be topical (including ophthalmic and to mucous membranes including vaginal and rectal delivery), pulmonary, e.g., by inhalation or insufflation of powders or aerosols, including 5 by nebulizer; intratracheal, intranasal, epidermal and transdermal), oral or parenteral. Parenteral administration includes intravenous, intraarterial, subcutaneous, intraperitoneal or intramuscular injection or infusion; or 10 intracranial, e.g., intrathecal or intraventricular, administration. Oligonucleotides with at least one 2'-O-methoxyethyl modification are believed to be particularly useful for oral administration.

Pharmaceutical compositions and formulations for topical 15 administration may include transdermal patches, ointments, lotions, creams, gels, drops, suppositories, sprays, liquids and powders. Conventional pharmaceutical carriers, aqueous, powder or oily bases, thickeners and the like may be necessary or desirable. Coated condoms, gloves and the like 20 may also be useful. Preferred topical formulations include those in which the oligonucleotides of the invention are in admixture with a topical delivery agent such as lipids, liposomes, fatty acids, fatty acid esters, steroids, chelating agents and surfactants. Preferred lipids and 25 liposomes include neutral (e.g. dioleoylphosphatidyl DOPE ethanolamine, dimyristoylphosphatidyl choline DMPC, distearoylphosphatidyl choline) negative (e.g. dimyristoylphosphatidyl glycerol DMPG) and cationic (e.g. dioleoyltetramethylaminopropyl DOTAP and dioleoylphosphatidyl 30 ethanolamine DOTMA). Oligonucleotides of the invention may be encapsulated within liposomes or may form complexes thereto, in particular to cationic liposomes. Alternatively, oligonucleotides may be complexed to lipids, in particular to

-39-

cationic lipids. Preferred fatty acids and esters include but are not limited arachidonic acid, oleic acid, eicosanoic acid, lauric acid, caprylic acid, capric acid, myristic acid, palmitic acid, stearic acid, linoleic acid, linolenic acid, 5 dicaprate, tricaprinate, monoolein, dilaurin, glyceryl 1-monocaprinate, 1-dodecylazacycloheptan-2-one, an acylcarnitine, an acylcholine, or a C₁₋₁₀ alkyl ester (e.g. isopropylmyristate IPM), monoglyceride, diglyceride or pharmaceutically acceptable salt thereof. Topical 10 formulations are described in detail in United States patent application 09/315,298 filed on May 20, 1999 which is incorporated herein by reference in its entirety.

Compositions and formulations for oral administration include powders or granules, microparticulates, 15 nanoparticulates, suspensions or solutions in water or non-aqueous media, capsules, gel capsules, sachets, tablets or minitables. Thickeners, flavoring agents, diluents, emulsifiers, dispersing aids or binders may be desirable. Preferred oral formulations are those in which 20 oligonucleotides of the invention are administered in conjunction with one or more penetration enhancers surfactants and chelators. Preferred surfactants include fatty acids and/or esters or salts thereof, bile acids and/or salts thereof. Preferred bile acids/salts include 25 chenodeoxycholic acid (CDCA) and ursodeoxychenodeoxycholic acid (UDCA), cholic acid, dehydrocholic acid, deoxycholic acid, glucholic acid, glycholic acid, glycodeoxycholic acid, taurocholic acid, taurodeoxycholic acid, sodium tauro-24,25-dihydro-fusidate and sodium glycodihydrofusidate. Preferred 30 fatty acids include arachidonic acid, undecanoic acid, oleic acid, lauric acid, caprylic acid, capric acid, myristic acid, palmitic acid, stearic acid, linoleic acid, linolenic acid, dicaprate, tricaprinate, monoolein, dilaurin, glyceryl 1-

-40-

monocaprato, 1-dodecylazacycloheptan-2-one, an acylcarnitine, an acylcholine, or a monoglyceride, a diglyceride or a pharmaceutically acceptable salt thereof (e.g. sodium). Also preferred are combinations of penetration enhancers, for example, fatty acids/salts in combination with bile acids/salts. A particularly preferred combination is the sodium salt of lauric acid, capric acid and UDCA. Further penetration enhancers include polyoxyethylene-9-lauryl ether, polyoxyethylene-20-cetyl ether. Oligonucleotides of the invention may be delivered orally, in granular form including sprayed dried particles, or complexed to form micro or nanoparticles. Oligonucleotide complexing agents include poly-amino acids; polyimines; polyacrylates; polyalkylacrylates, polyoxethanes, polyalkylcyanoacrylates; cationized gelatins, albumins, starches, acrylates, polyethyleneglycols (PEG) and starches; polyalkylcyanoacrylates; DEAE-derivatized polyimines, pullulans, celluloses and starches. Particularly preferred complexing agents include chitosan, N-trimethylchitosan, poly-L-lysine, polyhistidine, polyornithine, polyspermines, protamine, polyvinylpyridine, polythiodiethylamino-methylethylene P(TDAE), polyaminostyrene (e.g. p-amino), poly(methylcyanoacrylate), poly(ethylcyanoacrylate), poly(butylcyanoacrylate), poly(isobutylcyanoacrylate), poly(isohexylcyanoacrylate), DEAE-methacrylate, DEAE-hexylacrylate, DEAE-acrylamide, DEAE-albumin and DEAE-dextran, polymethylacrylate, polyhexylacrylate, poly(D,L-lactic acid), poly(DL-lactic-co-glycolic acid (PLGA), alginate, and polyethyleneglycol (PEG). Oral formulations for oligonucleotides and their preparation are described in detail in United States applications 08/886,829 (filed July 1, 1997), 09/108,673 (filed July 1, 1998), 09/256,515 (filed February 23, 1999), 09/082,624 (filed May 21, 1998) and

09/315,298 (filed May 20, 1999), each of which is incorporated herein by reference in their entirety.

Compositions and formulations for parenteral, intrathecal or intraventricular administration may include
5 sterile aqueous solutions which may also contain buffers, diluents and other suitable additives such as, but not limited to, penetration enhancers, carrier compounds and other pharmaceutically acceptable carriers or excipients.

Pharmaceutical compositions of the present invention
10 include, but are not limited to, solutions, emulsions, and liposome-containing formulations. These compositions may be generated from a variety of components that include, but are not limited to, preformed liquids, self-emulsifying solids and self-emulsifying semisolids.

15 The pharmaceutical formulations of the present invention, which may conveniently be presented in unit dosage form, may be prepared according to conventional techniques well known in the pharmaceutical industry. Such techniques include the step of bringing into association the active
20 ingredients with the pharmaceutical carrier(s) or excipient(s). In general, the formulations are prepared by uniformly and intimately bringing into association the active ingredients with liquid carriers or finely divided solid carriers or both, and then, if necessary, shaping the
25 product.

The compositions of the present invention may be formulated into any of many possible dosage forms such as, but not limited to, tablets, capsules, gel capsules, liquid syrups, soft gels, suppositories, and enemas. The
30 compositions of the present invention may also be formulated as suspensions in aqueous, non-aqueous or mixed media. Aqueous suspensions may further contain substances which increase the viscosity of the suspension including, for

-42-

example, sodium carboxymethylcellulose, sorbitol and/or dextran. The suspension may also contain stabilizers.

In one embodiment of the present invention the pharmaceutical compositions may be formulated and used as
5 foams. Pharmaceutical foams include formulations such as, but not limited to, emulsions, microemulsions, creams, jellies and liposomes. While basically similar in nature these formulations vary in the components and the consistency of the final product. The preparation of such compositions
10 and formulations is generally known to those skilled in the pharmaceutical and formulation arts and may be applied to the formulation of the compositions of the present invention.

Emulsions

15 The compositions of the present invention may be prepared and formulated as emulsions. Emulsions are typically heterogenous systems of one liquid dispersed in another in the form of droplets usually exceeding 0.1 μm in diameter (Idson, in *Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 199; Rosoff, in *Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., Volume 1, p. 245; Block
20 in *Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 2, p. 335; Higuchi et al., in *Remington's Pharmaceutical Sciences*, Mack Publishing Co., Easton, PA, 1985, p. 301). Emulsions are often biphasic systems comprising two immiscible liquid phases intimately mixed and dispersed with
30 each other. In general, emulsions may be of either the water-in-oil (w/o) or the oil-in-water (o/w) variety. When an aqueous phase is finely divided into and dispersed as minute droplets into a bulk oily phase, the resulting

-43-

composition is called a water-in-oil (w/o) emulsion. Alternatively, when an oily phase is finely divided into and dispersed as minute droplets into a bulk aqueous phase, the resulting composition is called an oil-in-water (o/w) emulsion. Emulsions may contain additional components in addition to the dispersed phases, and the active drug which may be present as a solution in either the aqueous phase, oily phase or itself as a separate phase. Pharmaceutical excipients such as emulsifiers, stabilizers, dyes, and anti-oxidants may also be present in emulsions as needed. Pharmaceutical emulsions may also be multiple emulsions that are comprised of more than two phases such as, for example, in the case of oil-in-water-in-oil (o/w/o) and water-in-oil-in-water (w/o/w) emulsions. Such complex formulations often provide certain advantages that simple binary emulsions do not. Multiple emulsions in which individual oil droplets of an o/w emulsion enclose small water droplets constitute a w/o/w emulsion. Likewise a system of oil droplets enclosed in globules of water stabilized in an oily continuous phase provides an o/w/o emulsion.

Emulsions are characterized by little or no thermodynamic stability. Often, the dispersed or discontinuous phase of the emulsion is well dispersed into the external or continuous phase and maintained in this form through the means of emulsifiers or the viscosity of the formulation. Either of the phases of the emulsion may be a semisolid or a solid, as is the case of emulsion-style ointment bases and creams. Other means of stabilizing emulsions entail the use of emulsifiers that may be incorporated into either phase of the emulsion. Emulsifiers may broadly be classified into four categories: synthetic surfactants, naturally occurring emulsifiers, absorption bases, and finely dispersed solids (Idson, in *Pharmaceutical*

-44-

Dosage Forms, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 199).

Synthetic surfactants, also known as surface active agents, have found wide applicability in the formulation of emulsions and have been reviewed in the literature (Rieger, in *Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 285; Idson, in *Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), Marcel Dekker, Inc., New York, N.Y., 1988, volume 1, p. 199). Surfactants are typically amphiphilic and comprise a hydrophilic and a hydrophobic portion. The ratio of the hydrophilic to the hydrophobic nature of the surfactant has been termed the hydrophile/lipophile balance (HLB) and is a valuable tool in categorizing and selecting surfactants in the preparation of formulations. Surfactants may be classified into different classes based on the nature of the hydrophilic group: nonionic, anionic, cationic and amphoteric (Rieger, in *Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 285).

Naturally occurring emulsifiers used in emulsion formulations include lanolin, beeswax, phosphatides, lecithin and acacia. Absorption bases possess hydrophilic properties such that they can soak up water to form w/o emulsions yet retain their semisolid consistencies, such as anhydrous lanolin and hydrophilic petrolatum. Finely divided solids have also been used as good emulsifiers especially in combination with surfactants and in viscous preparations. These include polar inorganic solids, such as heavy metal hydroxides, nonswelling clays such as bentonite, attapulgite, hectorite, kaolin, montmorillonite, colloidal aluminum

-45-

silicate and colloidal magnesium aluminum silicate, pigments and nonpolar solids such as carbon or glyceryl tristearate.

A large variety of non-emulsifying materials are also included in emulsion formulations and contribute to the properties of emulsions. These include fats, oils, waxes, fatty acids, fatty alcohols, fatty esters, humectants, hydrophilic colloids, preservatives and antioxidants (Block, in *Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 335; Idson, in *Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 199).

Hydrophilic colloids or hydrocolloids include naturally occurring gums and synthetic polymers such as polysaccharides (for example, acacia, agar, alginic acid, carrageenan, guar gum, karaya gum, and tragacanth), cellulose derivatives (for example, carboxymethylcellulose and carboxypropylcellulose), and synthetic polymers (for example, carbomers, cellulose ethers, and carboxyvinyl polymers). These disperse or swell in water to form colloidal solutions that stabilize emulsions by forming strong interfacial films around the dispersed-phase droplets and by increasing the viscosity of the external phase.

Since emulsions often contain a number of ingredients such as carbohydrates, proteins, sterols and phosphatides that may readily support the growth of microbes, these formulations often incorporate preservatives. Commonly used preservatives included in emulsion formulations include methyl paraben, propyl paraben, quaternary ammonium salts, benzalkonium chloride, esters of p-hydroxybenzoic acid, and boric acid. Antioxidants are also commonly added to emulsion formulations to prevent deterioration of the formulation. Antioxidants used may be free radical scavengers such as

-46-

tocopherols, alkyl gallates, butylated hydroxyanisole, butylated hydroxytoluene, or reducing agents such as ascorbic acid and sodium metabisulfite, and antioxidant synergists such as citric acid, tartaric acid, and lecithin.

5 The application of emulsion formulations via dermatological, oral and parenteral routes and methods for their manufacture have been reviewed in the literature (Idson, in *Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 199). Emulsion formulations for oral delivery
10 have been very widely used because of ease of formulation, as well as efficacy from an absorption and bioavailability standpoint (Rosoff, in *Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 245; Idson, in
15 *Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 199). Mineral-oil base laxatives, oil-soluble vitamins and high fat nutritive preparations are among the materials
20 that have commonly been administered orally as o/w emulsions.

In one embodiment of the present invention, the compositions of oligonucleotides and nucleic acids are formulated as microemulsions. A microemulsion may be defined as a system of water, oil, and amphiphile which is a single
25 optically isotropic and thermodynamically stable liquid solution (Rosoff, in *Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 245). Typically microemulsions are systems that are prepared by first dispersing an oil in an
30 aqueous surfactant solution and then adding a sufficient amount of a fourth component, generally an intermediate chain-length alcohol to form a transparent system. Therefore, microemulsions have also been described as

-47-

thermodynamically stable, isotropically clear dispersions of two immiscible liquids that are stabilized by interfacial films of surface-active molecules (Leung and Shah, *in: Controlled Release of Drugs: Polymers and Aggregate Systems*, Rosoff, M., Ed., 1989, VCH Publishers, New York, pages 185-215). Microemulsions commonly are prepared via a combination of three to five components that include oil, water, surfactant, cosurfactant and electrolyte. Whether the microemulsion is of the water-in-oil (w/o) or an oil-in-water (o/w) type is dependent on the properties of the oil and surfactant used and on the structure and geometric packing of the polar heads and hydrocarbon tails of the surfactant molecules (Schott, *in Remington's Pharmaceutical Sciences*, Mack Publishing Co., Easton, PA, 1985, p. 271).

The phenomenological approach utilizing phase diagrams has been extensively studied and has yielded a comprehensive knowledge, to one skilled in the art, of how to formulate microemulsions (Rosoff, *in Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 245; Block, *in Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 335). Compared to conventional emulsions, microemulsions offer the advantage of solubilizing water-insoluble drugs in a formulation of thermodynamically stable droplets that are formed spontaneously.

Surfactants used in the preparation of microemulsions include, but are not limited to, ionic surfactants, non-ionic surfactants, Brij 96, polyoxyethylene oleyl ethers, polyglycerol fatty acid esters, tetraglycerol monolaurate (ML310), tetraglycerol monooleate (MO310), hexaglycerol monooleate (PO310), hexaglycerol pentaoleate (PO500), decaglycerol monocaprates (MCA750), decaglycerol monooleate

-48-

(MO750), decaglycerol sequioleate (SO750), decaglycerol decaoleate (DAO750), alone or in combination with cosurfactants. The cosurfactant, usually a short-chain alcohol such as ethanol, 1-propanol, and 1-butanol, serves to increase the interfacial fluidity by penetrating into the surfactant film and consequently creating a disordered film because of the void space generated among surfactant molecules. Microemulsions may, however, be prepared without the use of cosurfactants and alcohol-free self-emulsifying microemulsion systems are known in the art. The aqueous phase may typically be, but is not limited to, water, an aqueous solution of the drug, glycerol, PEG300, PEG400, polyglycerols, propylene glycols, and derivatives of ethylene glycol. The oil phase may include, but is not limited to, materials such as Captex 300, Captex 355, Capmul MCM, fatty acid esters, medium chain (C8-C12) mono, di, and tri-glycerides, polyoxyethylated glyceryl fatty acid esters, fatty alcohols, polyglycolized glycerides, saturated polyglycolized C8-C10 glycerides, vegetable oils and silicone oil.

Microemulsions are particularly of interest from the standpoint of drug solubilization and the enhanced absorption of drugs. Lipid based microemulsions (both o/w and w/o) have been proposed to enhance the oral bioavailability of drugs, including peptides (Constantinides *et al.*, *Pharmaceutical Research*, 1994, 11, 1385-1390; Ritschel, *Meth. Find. Exp. Clin. Pharmacol.*, 1993, 13, 205). Microemulsions afford advantages of improved drug solubilization, protection of drug from enzymatic hydrolysis, possible enhancement of drug absorption due to surfactant-induced alterations in membrane fluidity and permeability, ease of preparation, ease of oral administration over solid dosage forms, improved clinical potency, and decreased toxicity (Constantinides *et al.*,

-49-

Pharmaceutical Research, 1994, 11, 1385; Ho et al., *J. Pharm. Sci.*, 1996, 85, 138-143). Often microemulsions may form spontaneously when their components are brought together at ambient temperature. This may be particularly advantageous when formulating thermolabile drugs, peptides or oligonucleotides. Microemulsions have also been effective in the transdermal delivery of active components in both cosmetic and pharmaceutical applications. It is expected that the microemulsion compositions and formulations of the present invention will facilitate the increased systemic absorption of oligonucleotides and nucleic acids from the gastrointestinal tract, as well as improve the local cellular uptake of oligonucleotides and nucleic acids within the gastrointestinal tract, vagina, buccal cavity and other areas of administration.

Microemulsions of the present invention may also contain additional components and additives such as sorbitan monostearate (Grill 3), Labrasol, and penetration enhancers to improve the properties of the formulation and to enhance the absorption of the oligonucleotides and nucleic acids of the present invention. Penetration enhancers used in the microemulsions of the present invention may be classified as belonging to one of five broad categories - surfactants, fatty acids, bile salts, chelating agents, and non-chelating non-surfactants (Lee et al., *Critical Reviews in Therapeutic Drug Carrier Systems*, 1991, p. 92). Each of these classes has been discussed above.

Liposomes

There are many organized surfactant structures besides microemulsions that have been studied and used for the formulation of drugs. These include monolayers, micelles, bilayers and vesicles. Vesicles, such as liposomes, have

-50-

attracted great interest because of their specificity and the duration of action they offer from the standpoint of drug delivery. As used in the present invention, the term "liposome" means a vesicle composed of amphiphilic lipids arranged in a spherical bilayer or bilayers.

Liposomes are unilamellar or multilamellar vesicles which have a membrane formed from a lipophilic material and an aqueous interior. The aqueous portion contains the composition to be delivered. Cationic liposomes possess the advantage of being able to fuse to the cell wall. Non-cationic liposomes, although not able to fuse as efficiently with the cell wall, are taken up by macrophages *in vivo*.

In order to cross intact mammalian skin, lipid vesicles must pass through a series of fine pores, each with a diameter less than 50 nm, under the influence of a suitable transdermal gradient. Therefore, it is desirable to use a liposome which is highly deformable and able to pass through such fine pores.

Further advantages of liposomes include; liposomes obtained from natural phospholipids are biocompatible and biodegradable; liposomes can incorporate a wide range of water and lipid soluble drugs; liposomes can protect encapsulated drugs in their internal compartments from metabolism and degradation (Rosoff, in *Pharmaceutical Dosage Forms*, Lieberman, Rieger and Banker (Eds.), 1988, Marcel Dekker, Inc., New York, N.Y., volume 1, p. 245). Important considerations in the preparation of liposome formulations are the lipid surface charge, vesicle size and the aqueous volume of the liposomes.

Liposomes are useful for the transfer and delivery of active ingredients to the site of action. Because the liposomal membrane is structurally similar to biological membranes, when liposomes are applied to a tissue, the

-51-

liposomes start to merge with the cellular membranes and as the merging of the liposome and cell progresses, the liposomal contents are emptied into the cell where the active agent may act.

5 Liposomal formulations have been the focus of extensive investigation as the mode of delivery for many drugs. There is growing evidence that for topical administration, liposomes present several advantages over other formulations. Such advantages include reduced side-effects related to high
10 systemic absorption of the administered drug, increased accumulation of the administered drug at the desired target, and the ability to administer a wide variety of drugs, both hydrophilic and hydrophobic, into the skin.

 Several reports have detailed the ability of liposomes
15 to deliver agents including high-molecular weight DNA into the skin. Compounds including analgesics, antibodies, hormones and high-molecular weight DNAs have been administered to the skin. The majority of applications resulted in the targeting of the upper epidermis.

20 Liposomes fall into two broad classes. Cationic liposomes are positively charged liposomes which interact with the negatively charged DNA molecules to form a stable complex. The positively charged DNA/liposome complex binds to the negatively charged cell surface and is internalized in
25 an endosome. Due to the acidic pH within the endosome, the liposomes are ruptured, releasing their contents into the cell cytoplasm (Wang et al., *Biochem. Biophys. Res. Commun.*, 1987, 147, 980-985).

 Liposomes which are pH-sensitive or negatively-charged,
30 entrap DNA rather than complex with it. Since both the DNA and the lipid are similarly charged, repulsion rather than complex formation occurs. Nevertheless, some DNA is entrapped within the aqueous interior of these liposomes. pH-sensitive

-52-

liposomes have been used to deliver DNA encoding the thymidine kinase gene to cell monolayers in culture. Expression of the exogenous gene was detected in the target cells (Zhou et al., *Journal of Controlled Release*, 1992, 19, 269-274).

One major type of liposomal composition includes phospholipids other than naturally-derived phosphatidylcholine. Neutral liposome compositions, for example, can be formed from dimyristoyl phosphatidylcholine (DMPC) or dipalmitoyl phosphatidylcholine (DPPC). Anionic liposome compositions generally are formed from dimyristoyl phosphatidylglycerol, while anionic fusogenic liposomes are formed primarily from dioleoyl phosphatidylethanolamine (DOPE). Another type of liposomal composition is formed from phosphatidylcholine (PC) such as, for example, soybean PC, and egg PC. Another type is formed from mixtures of phospholipid and/or phosphatidylcholine and/or cholesterol.

Several studies have assessed the topical delivery of liposomal drug formulations to the skin. Application of liposomes containing interferon to guinea pig skin resulted in a reduction of skin herpes sores while delivery of interferon via other means (e.g. as a solution or as an emulsion) were ineffective (Weiner et al., *Journal of Drug Targeting*, 1992, 2, 405-410). Further, an additional study tested the efficacy of interferon administered as part of a liposomal formulation to the administration of interferon using an aqueous system, and concluded that the liposomal formulation was superior to aqueous administration (du Plessis et al., *Antiviral Research*, 1992, 18, 259-265).

Non-ionic liposomal systems have also been examined to determine their utility in the delivery of drugs to the skin, in particular systems comprising non-ionic surfactant and cholesterol. Non-ionic liposomal formulations comprising

Novasome™ I (glyceryl dilaurate/cholesterol/polyoxyethylene-10-stearyl ether) and Novasome™ II (glyceryl distearate/cholesterol/polyoxyethylene-10-stearyl ether) were used to deliver cyclosporin-A into the dermis of mouse skin. Results
5 indicated that such non-ionic liposomal systems were effective in facilitating the deposition of cyclosporin-A into different layers of the skin (Hu et al. *S.T.P. Pharma. Sci.*, 1994, 4, 6, 466).

Liposomes also include "sterically stabilized"
10 liposomes, a term which, as used herein, refers to liposomes comprising one or more specialized lipids that, when incorporated into liposomes, result in enhanced circulation lifetimes relative to liposomes lacking such specialized lipids. Examples of sterically stabilized liposomes are
15 those in which part of the vesicle-forming lipid portion of the liposome (A) comprises one or more glycolipids, such as monosialoganglioside G_{M1}, or (B) is derivatized with one or more hydrophilic polymers, such as a polyethylene glycol (PEG) moiety. While not wishing to be bound by any
20 particular theory, it is thought in the art that, at least for sterically stabilized liposomes containing gangliosides, sphingomyelin, or PEG-derivatized lipids, the enhanced circulation half-life of these sterically stabilized liposomes derives from a reduced uptake into cells of the
25 reticuloendothelial system (RES) (Allen et al., *FEBS Letters*, 1987, 223, 42; Wu et al., *Cancer Research*, 1993, 53, 3765).

Various liposomes comprising one or more glycolipids are known in the art. Papahadjopoulos et al. (*Ann. N.Y. Acad. Sci.*, 1987, 507, 64) reported the ability of
30 monosialoganglioside G_{M1}, galactocerebroside sulfate and phosphatidylinositol to improve blood half-lives of liposomes. These findings were expounded upon by Gabizon et

-54-

al. (*Proc. Natl. Acad. Sci. U.S.A.*, 1988, 85, 6949). U.S. Patent No. 4,837,028 and WO 88/04924, both to Allen *et al.*, disclose liposomes comprising (1) sphingomyelin and (2) the ganglioside G_{M1} or a galactocerebroside sulfate ester. U.S. Patent No. 5,543,152 (Webb *et al.*) discloses liposomes comprising sphingomyelin. Liposomes comprising 1,2-*sn*-dimyristoylphosphatidylcholine are disclosed in WO 97/13499 (Lim *et al.*).

Many liposomes comprising lipids derivatized with one or more hydrophilic polymers, and methods of preparation thereof, are known in the art. Sunamoto *et al.* (*Bull. Chem. Soc. Jpn.*, 1980, 53, 2778) described liposomes comprising a nonionic detergent, 2C₁₂15G, that contains a PEG moiety. Illum *et al.* (*FEBS Lett.*, 1984, 167, 79) noted that hydrophilic coating of polystyrene particles with polymeric glycols results in significantly enhanced blood half-lives. Synthetic phospholipids modified by the attachment of carboxylic groups of polyalkylene glycols (e.g., PEG) are described by Sears (U.S. Patent Nos. 4,426,330 and 4,534,899). Klibanov *et al.* (*FEBS Lett.*, 1990, 268, 235) described experiments demonstrating that liposomes comprising phosphatidylethanolamine (PE) derivatized with PEG or PEG stearate have significant increases in blood circulation half-lives. Blume *et al.* (*Biochimica et Biophysica Acta*, 1990, 1029, 91) extended such observations to other PEG-derivatized phospholipids, e.g., DSPE-PEG, formed from the combination of distearoylphosphatidylethanolamine (DSPE) and PEG. Liposomes having covalently bound PEG moieties on their external surface are described in European Patent No. EP 0 445 131 B1 and WO 90/04384 to Fisher. Liposome compositions containing 1-20 mole percent of PE derivatized with PEG, and methods of use thereof, are described by Woodle *et al.* (U.S. Patent Nos. 5,013,556 and 5,356,633) and Martin *et al.* (U.S.

-55-

Patent No. 5,213,804 and European Patent No. EP 0 496 813 B1). Liposomes comprising a number of other lipid-polymer conjugates are disclosed in WO 91/05545 and U.S. Patent No. 5,225,212 (both to Martin *et al.*) and in WO 94/20073 (Zalipsky *et al.*) Liposomes comprising PEG-modified ceramide lipids are described in WO 96/10391 (Choi *et al.*). U.S. Patent Nos. 5,540,935 (Miyazaki *et al.*) and 5,556,948 (Tagawa *et al.*) describe PEG-containing liposomes that can be further derivatized with functional moieties on their surfaces.

10 A limited number of liposomes comprising nucleic acids are known in the art. WO 96/40062 to Thierry *et al.* discloses methods for encapsulating high molecular weight nucleic acids in liposomes. U.S. Patent No. 5,264,221 to Tagawa *et al.* discloses protein-bonded liposomes and asserts
15 that the contents of such liposomes may include an antisense RNA. U.S. Patent No. 5,665,710 to Rahman *et al.* describes certain methods of encapsulating oligodeoxynucleotides in liposomes. WO 97/04787 to Love *et al.* discloses liposomes comprising antisense oligonucleotides targeted to the raf
20 gene.

Transfersomes are yet another type of liposomes, and are highly deformable lipid aggregates which are attractive candidates for drug delivery vehicles. Transfersomes may be described as lipid droplets which are so highly deformable
25 that they are easily able to penetrate through pores which are smaller than the droplet. Transfersomes are adaptable to the environment in which they are used, e.g. they are self-optimizing (adaptive to the shape of pores in the skin), self-repairing, frequently reach their targets without
30 fragmenting, and often self-loading. To make transfersomes it is possible to add surface edge-activators, usually surfactants, to a standard liposomal composition. Transfersomes have been used to deliver serum albumin to the

skin. The transferrin-mediated delivery of serum albumin has been shown to be as effective as subcutaneous injection of a solution containing serum albumin.

Surfactants find wide application in formulations such as emulsions (including microemulsions) and liposomes. The most common way of classifying and ranking the properties of the many different types of surfactants, both natural and synthetic, is by the use of the hydrophile/lipophile balance (HLB). The nature of the hydrophilic group (also known as the "head") provides the most useful means for categorizing the different surfactants used in formulations (Rieger, in *Pharmaceutical Dosage Forms*, Marcel Dekker, Inc., New York, NY, 1988, p. 285).

If the surfactant molecule is not ionized, it is classified as a nonionic surfactant. Nonionic surfactants find wide application in pharmaceutical and cosmetic products and are usable over a wide range of pH values. In general their HLB values range from 2 to about 18 depending on their structure. Nonionic surfactants include nonionic esters such as ethylene glycol esters, propylene glycol esters, glyceryl esters, polyglyceryl esters, sorbitan esters, sucrose esters, and ethoxylated esters. Nonionic alkanolamides and ethers such as fatty alcohol ethoxylates, propoxylated alcohols, and ethoxylated/propoxylated block polymers are also included in this class. The polyoxyethylene surfactants are the most popular members of the nonionic surfactant class.

If the surfactant molecule carries a negative charge when it is dissolved or dispersed in water, the surfactant is classified as anionic. Anionic surfactants include carboxylates such as soaps, acyl lactylates, acyl amides of amino acids, esters of sulfuric acid such as alkyl sulfates and ethoxylated alkyl sulfates, sulfonates such as alkyl benzene sulfonates, acyl isethionates, acyl taurates and

-57-

sulfosuccinates, and phosphates. The most important members of the anionic surfactant class are the alkyl sulfates and the soaps.

5 If the surfactant molecule carries a positive charge when it is dissolved or dispersed in water, the surfactant is classified as cationic. Cationic surfactants include quaternary ammonium salts and ethoxylated amines. The quaternary ammonium salts are the most used members of this class.

10 If the surfactant molecule has the ability to carry either a positive or negative charge, the surfactant is classified as amphoteric. Amphoteric surfactants include acrylic acid derivatives, substituted alkylamides, N-alkylbetaines and phosphatides.

15 The use of surfactants in drug products, formulations and in emulsions has been reviewed (Rieger, in *Pharmaceutical Dosage Forms*, Marcel Dekker, Inc., New York, NY, 1988, p. 285).

20 Penetration Enhancers

In one embodiment, the present invention employs various penetration enhancers to effect the efficient delivery of nucleic acids, particularly oligonucleotides, to the skin of animals. Most drugs are present in solution in both ionized and nonionized forms. However, usually only lipid soluble or lipophilic drugs readily cross cell membranes. It has been discovered that even non-lipophilic drugs may cross cell membranes if the membrane to be crossed is treated with a penetration enhancer. In addition to aiding the diffusion of non-lipophilic drugs across cell membranes, penetration enhancers also enhance the permeability of lipophilic drugs.

30 Penetration enhancers may be classified as belonging to one of five broad categories, *i.e.*, surfactants, fatty acids,

-58-

bile salts, chelating agents, and non-chelating non-surfactants (Lee *et al.*, *Critical Reviews in Therapeutic Drug Carrier Systems*, 1991, p.92). Each of the above mentioned classes of penetration enhancers are described below in
5 greater detail.

Surfactants: In connection with the present invention, surfactants (or "surface-active agents") are chemical entities which, when dissolved in an aqueous solution, reduce
10 the surface tension of the solution or the interfacial tension between the aqueous solution and another liquid, with the result that absorption of oligonucleotides through the mucosa is enhanced. In addition to bile salts and fatty acids, these penetration enhancers include, for example,
15 sodium lauryl sulfate, polyoxyethylene-9-lauryl ether and polyoxyethylene-20-cetyl ether) (Lee *et al.*, *Critical Reviews in Therapeutic Drug Carrier Systems*, 1991, p.92); and perfluorochemical emulsions, such as FC-43. Takahashi *et al.*, *J. Pharm. Pharmacol.*, 1988, 40, 252).

20 Fatty acids: Various fatty acids and their derivatives which act as penetration enhancers include, for example, oleic acid, lauric acid, capric acid (n-decanoic acid), myristic acid, palmitic acid, stearic acid, linoleic acid, linolenic acid, dicaprate, tricaprate, monoolein (1-
25 monooleoyl-*rac*-glycerol), dilaurin, caprylic acid, arachidonic acid, glycerol 1-monocaprate, 1-dodecylazacycloheptan-2-one, acylcarnitines, acylcholines, C₁₋₁₀ alkyl esters thereof (e.g., methyl, isopropyl and *t*-butyl), and mono- and di-glycerides thereof (*i.e.*, oleate,
30 laurate, caprate, myristate, palmitate, stearate, linoleate, *etc.*) (Lee *et al.*, *Critical Reviews in Therapeutic Drug Carrier Systems*, 1991, p.92; Muranishi, *Critical Reviews in*

-59-

Therapeutic Drug Carrier Systems, 1990, 7, 1-33; El Hariri et al., *J. Pharm. Pharmacol.*, 1992, 44, 651-654).

Bile salts: The physiological role of bile includes the facilitation of dispersion and absorption of lipids and fat-soluble vitamins (Brunton, Chapter 38 in: Goodman & Gilman's *The Pharmacological Basis of Therapeutics*, 9th Ed., Hardman et al. Eds., McGraw-Hill, New York, 1996, pp. 934-935). Various natural bile salts, and their synthetic derivatives, act as penetration enhancers. Thus the term "bile salts" includes any of the naturally occurring components of bile as well as any of their synthetic derivatives. The bile salts of the invention include, for example, cholic acid (or its pharmaceutically acceptable sodium salt, sodium cholate), dehydrocholic acid (sodium dehydrocholate), deoxycholic acid (sodium deoxycholate), glucolic acid (sodium glucolate), glycholic acid (sodium glycocholate), glycodeoxycholic acid (sodium glycodeoxycholate), taurocholic acid (sodium taurocholate), taurodeoxycholic acid (sodium taurodeoxycholate), chenodeoxycholic acid (sodium chenodeoxycholate), ursodeoxycholic acid (UDCA), sodium tauro-24,25-dihydro-fusidate (STDHF), sodium glycodihydrofusidate and polyoxyethylene-9-lauryl ether (POE) (Lee et al., *Critical Reviews in Therapeutic Drug Carrier Systems*, 1991, page 92; Swinyard, Chapter 39 In: *Remington's Pharmaceutical Sciences*, 18th Ed., Gennaro, ed., Mack Publishing Co., Easton, PA, 1990, pages 782-783; Muranishi, *Critical Reviews in Therapeutic Drug Carrier Systems*, 1990, 7, 1-33; Yamamoto et al., *J. Pharm. Exp. Ther.*, 1992, 263, 25; Yamashita et al., *J. Pharm. Sci.*, 1990, 79, 579-583).

Chelating Agents: Chelating agents, as used in connection with the present invention, can be defined as

-60-

compounds that remove metallic ions from solution by forming complexes therewith, with the result that absorption of oligonucleotides through the mucosa is enhanced. With regards to their use as penetration enhancers in the present invention, chelating agents have the added advantage of also serving as DNase inhibitors, as most characterized DNA nucleases require a divalent metal ion for catalysis and are thus inhibited by chelating agents (Jarrett, *J. Chromatogr.*, 1993, 618, 315-339). Chelating agents of the invention include but are not limited to disodium ethylenediaminetetraacetate (EDTA), citric acid, salicylates (e.g., sodium salicylate, 5-methoxysalicylate and homovanilate), *N*-acyl derivatives of collagen, laureth-9 and *N*-amino acyl derivatives of beta-diketones (enamines) (Lee et al., *Critical Reviews in Therapeutic Drug Carrier Systems*, 1991, page 92; Muranishi, *Critical Reviews in Therapeutic Drug Carrier Systems*, 1990, 7, 1-33; Buur et al., *J. Control Rel.*, 1990, 14, 43-51).

Non-chelating non-surfactants: As used herein, non-chelating non-surfactant penetration enhancing compounds can be defined as compounds that demonstrate insignificant activity as chelating agents or as surfactants but that nonetheless enhance absorption of oligonucleotides through the alimentary mucosa (Muranishi, *Critical Reviews in Therapeutic Drug Carrier Systems*, 1990, 7, 1-33). This class of penetration enhancers include, for example, unsaturated cyclic ureas, 1-alkyl- and 1-alkenylazacyclo-alkanone derivatives (Lee et al., *Critical Reviews in Therapeutic Drug Carrier Systems*, 1991, page 92); and non-steroidal anti-inflammatory agents such as diclofenac sodium, indomethacin and phenylbutazone (Yamashita et al., *J. Pharm. Pharmacol.*, 1987, 39, 621-626).

-61-

Agents that enhance uptake of oligonucleotides at the cellular level may also be added to the pharmaceutical and other compositions of the present invention. For example, cationic lipids, such as lipofectin (Junichi *et al*, U.S. Patent No. 5,705,188), cationic glycerol derivatives, and polycationic molecules, such as polylysine (Lollo *et al.*, PCT Application WO 97/30731), are also known to enhance the cellular uptake of oligonucleotides.

Other agents may be utilized to enhance the penetration of the administered nucleic acids, including glycols such as ethylene glycol and propylene glycol, pyrrols such as 2-pyrrol, azones, and terpenes such as limonene and menthone.

Carriers

Certain compositions of the present invention also incorporate carrier compounds in the formulation. As used herein, "carrier compound" or "carrier" can refer to a nucleic acid, or analog thereof, which is inert (*i.e.*, does not possess biological activity *per se*) but is recognized as a nucleic acid by *in vivo* processes that reduce the bioavailability of a nucleic acid having biological activity by, for example, degrading the biologically active nucleic acid or promoting its removal from circulation. The coadministration of a nucleic acid and a carrier compound, typically with an excess of the latter substance, can result in a substantial reduction of the amount of nucleic acid recovered in the liver, kidney or other extracirculatory reservoirs, presumably due to competition between the carrier compound and the nucleic acid for a common receptor. For example, the recovery of a partially phosphorothioate oligonucleotide in hepatic tissue can be reduced when it is coadministered with polyinosinic acid, dextran sulfate, polycytidic acid or 4-acetamido-4'-isothiocyano-stilbene-2,2'-

disulfonic acid (Miyao *et al.*, *Antisense Res. Dev.*, 1995, 5, 115-121; Takakura *et al.*, *Antisense & Nucl. Acid Drug Dev.*, 1996, 6, 177-183).

5 Excipients

In contrast to a carrier compound, a "pharmaceutical carrier" or "excipient" is a pharmaceutically acceptable solvent, suspending agent or any other pharmacologically inert vehicle for delivering one or more nucleic acids to an animal. The excipient may be liquid or solid and is selected, with the planned manner of administration in mind, so as to provide for the desired bulk, consistency, *etc.*, when combined with a nucleic acid and the other components of a given pharmaceutical composition. Typical pharmaceutical carriers include, but are not limited to, binding agents (e.g., pregelatinized maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose, *etc.*); fillers (e.g., lactose and other sugars, microcrystalline cellulose, pectin, gelatin, calcium sulfate, ethyl cellulose, polyacrylates or calcium hydrogen phosphate, *etc.*); lubricants (e.g., magnesium stearate, talc, silica, colloidal silicon dioxide, stearic acid, metallic stearates, hydrogenated vegetable oils, corn starch, polyethylene glycols, sodium benzoate, sodium acetate, *etc.*); disintegrants (e.g., starch, sodium starch glycolate, *etc.*); and wetting agents (e.g., sodium lauryl sulphate, *etc.*).

Pharmaceutically acceptable organic or inorganic excipient suitable for non-parenteral administration, which do not deleteriously react with nucleic acids, can also be used to formulate the compositions of the present invention. Suitable pharmaceutically acceptable carriers include, but are not limited to, water, salt solutions, alcohols, polyethylene glycols, gelatin, lactose, amylose, magnesium

-63-

stearate, talc, silicic acid, viscous paraffin,
hydroxymethylcellulose, polyvinylpyrrolidone and the like.

Formulations for topical administration of nucleic acids may include sterile and non-sterile aqueous solutions, non-
5 aqueous solutions in common solvents such as alcohols, or solutions of the nucleic acids in liquid or solid oil bases. The solutions may also contain buffers, diluents and other suitable additives. Pharmaceutically acceptable organic or inorganic excipients suitable for non-parenteral
10 administration, which do not deleteriously react with nucleic acids, can be used.

Suitable pharmaceutically acceptable excipients include, but are not limited to, water, salt solutions, alcohol, polyethylene glycols, gelatin, lactose, amylose, magnesium
15 stearate, talc, silicic acid, viscous paraffin, hydroxymethylcellulose, polyvinylpyrrolidone and the like.

Other Components

The compositions of the present invention may
20 additionally contain other adjunct components conventionally found in pharmaceutical compositions, at their art-established usage levels. Thus, for example, the compositions may contain additional, compatible, pharmaceutically-active materials such as, for example,
25 antipruritics, astringents, local anesthetics or anti-inflammatory agents, or may contain additional materials useful in physically formulating various dosage forms of the compositions of the present invention, such as dyes, flavoring agents, preservatives, antioxidants, opacifiers,
30 thickening agents and stabilizers. However, such materials, when added, should not unduly interfere with the biological activities of the components of the compositions of the present invention. The formulations can be sterilized and,

if desired, mixed with auxiliary agents, e.g., lubricants, preservatives, stabilizers, wetting agents, emulsifiers, salts for influencing osmotic pressure, buffers, colorings, flavorings and/or aromatic substances and the like which do
5 not deleteriously interact with the nucleic acid(s) of the formulation.

Aqueous suspensions may contain substances which increase the viscosity of the suspension including, for example, sodium carboxymethylcellulose, sorbitol and/or
10 dextran. The suspension may also contain stabilizers.

Certain embodiments of the invention provide pharmaceutical compositions containing (a) one or more antisense compounds and (b) one or more other chemotherapeutic agents which function by a non-antisense
15 mechanism. Examples of such chemotherapeutic agents include but are not limited to daunorubicin, daunomycin, dactinomycin, doxorubicin, epirubicin, idarubicin, esorubicin, bleomycin, mafosfamide, ifosfamide, cytosine arabinoside, bis-chloroethylnitrosurea, busulfan, mitomycin
20 C, actinomycin D, mithramycin, prednisone, hydroxyprogesterone, testosterone, tamoxifen, dacarbazine, procarbazine, hexamethylmelamine, pentamethylmelamine, mitoxantrone, amsacrine, chlorambucil, nitrogen mustards, methylcyclohexylnitrosurea, melphalan, cyclophosphamide, 6-
25 mercaptopurine, 6-thioguanine, cytarabine, 5-azacytidine, hydroxyurea, 4-hydroxyperoxycyclophosphoramidate, deoxycoformycin, 5-fluorouracil (5-FU), 5-fluorodeoxyuridine (5-FUdR), methotrexate (MTX), colchicine, taxol, vincristine, vinblastine, etoposide (VP-16), trimetrexate, irinotecan,
30 topotecan, gemcitabine, teniposide, cisplatin and diethylstilbestrol (DES). See, generally, *The Merck Manual of Diagnosis and Therapy*, 15th Ed. 1987, pp. 1206-1228, Berkow et al., eds., Rahway, N.J. When used with the

-65-

compounds of the invention, such chemotherapeutic agents may be used individually (e.g., 5-FU and oligonucleotide), sequentially (e.g., 5-FU and oligonucleotide for a period of time followed by MTX and oligonucleotide), or in combination with one or more other such chemotherapeutic agents (e.g., 5-FU, MTX and oligonucleotide, or 5-FU, radiotherapy and oligonucleotide). Anti-inflammatory drugs, including but not limited to nonsteroidal anti-inflammatory drugs and corticosteroids, and antiviral drugs, including but not limited to ribivirin, vidarabine, acyclovir and ganciclovir, may also be combined in compositions of the invention. See, generally, *The Merck Manual of Diagnosis and Therapy*, 15th Ed., Berkow et al., eds., 1987, Rahway, N.J., pages 2499-2506 and 46-49, respectively). Other non-antisense chemotherapeutic agents are also within the scope of this invention. Two or more combined compounds may be used together or sequentially.

In another related embodiment, compositions of the invention may contain one or more antisense compounds, particularly oligonucleotides, targeted to a first nucleic acid and one or more additional antisense compounds targeted to a second nucleic acid target. Numerous examples of antisense compounds are known in the art. Two or more combined compounds may be used together or sequentially.

The formulation of therapeutic compositions and their subsequent administration is believed to be within the skill of those in the art. Dosing is dependent on severity and responsiveness of the disease state to be treated, with the course of treatment lasting from several days to several months, or until a cure is effected or a diminution of the disease state is achieved. Optimal dosing schedules can be calculated from measurements of drug accumulation in the body of the patient. Persons of ordinary skill can easily

-66-

determine optimum dosages, dosing methodologies and repetition rates. Optimum dosages may vary depending on the relative potency of individual oligonucleotides, and can generally be estimated based on EC₅₀s found to be effective in *in vitro* and *in vivo* animal models. In general, dosage is from 0.01 ug to 100 g per kg of body weight, and may be given once or more daily, weekly, monthly or yearly, or even once every 2 to 20 years. Persons of ordinary skill in the art can easily estimate repetition rates for dosing based on measured residence times and concentrations of the drug in bodily fluids or tissues. Following successful treatment, it may be desirable to have the patient undergo maintenance therapy to prevent the recurrence of the disease state, wherein the oligonucleotide is administered in maintenance doses, ranging from 0.01 ug to 100 g per kg of body weight, once or more daily, to once every 20 years.

While the present invention has been described with specificity in accordance with certain of its preferred embodiments, the following examples serve only to illustrate the invention and are not intended to limit the same.

EXAMPLES

Example 1

25 Nucleoside Phosphoramidites for Oligonucleotide Synthesis Deoxy and 2'-alkoxy amidites

2'-Deoxy and 2'-methoxy beta-cyanoethyl diisopropyl phosphoramidites were purchased from commercial sources (e.g. Chemgenes, Needham MA or Glen Research, Inc. Sterling VA). Other 2'-O-alkoxy substituted nucleoside amidites are prepared as described in U.S. Patent 5,506,351, herein incorporated by reference. For oligonucleotides synthesized using 2'-alkoxy amidites, optimized synthesis cycles were

-67-

developed that incorporate multiple steps coupling longer wait times relative to standard synthesis cycles.

The following abbreviations are used in the text: thin layer chromatography (TLC), melting point (MP), high pressure liquid chromatography (HPLC), Nuclear Magnetic Resonance (NMR), argon (Ar), methanol (MeOH), dichloromethane (CH₂Cl₂), triethylamine (TEA), dimethyl formamide (DMF), ethyl acetate (EtOAc), dimethyl sulfoxide (DMSO), tetrahydrofuran (THF).

Oligonucleotides containing 5-methyl-2'-deoxycytidine (5-Me-dC) nucleotides were synthesized according to published methods (Sanghvi, et. al., *Nucleic Acids Research*, 1993, 21, 3197-3203) using commercially available phosphoramidites (Glen Research, Sterling VA or ChemGenes, Needham MA) or prepared as follows:

15

Preparation of 5'-O-Dimethoxytrityl-thymidine intermediate for 5-methyl dC amidite

To a 50 L glass reactor equipped with air stirrer and Ar gas line was added thymidine (1.00 kg, 4.13 mol) in anhydrous pyridine (6 L) at ambient temperature. Dimethoxytrityl (DMT) chloride (1.47 kg, 4.34 mol, 1.05 eq) was added as a solid in four portions over 1 h. After 30 min, TLC indicated approx. 95% product, 2% thymidine, 5% DMT reagent and by-products and 2 % 3',5'-bis DMT product (R_f in EtOAc 0.45, 0.05, 0.98, 0.95 respectively). Saturated sodium bicarbonate (4 L) and CH₂Cl₂ were added with stirring (pH of the aqueous layer 7.5). An additional 18 L of water was added, the mixture was stirred, the phases were separated, and the organic layer was transferred to a second 50 L vessel. The aqueous layer was extracted with additional CH₂Cl₂ (2 x 2 L). The combined organic layer was washed with water (10 L) and then concentrated in a rotary evaporator to approx. 3.6 kg total weight. This was redissolved in CH₂Cl₂ (3.5 L), added to the

-68-

reactor followed by water (6 L) and hexanes (13 L). The mixture was vigorously stirred and seeded to give a fine white suspended solid starting at the interface. After stirring for 1 h, the suspension was removed by suction
5 through a 1/2" diameter teflon tube into a 20 L suction flask, poured onto a 25 cm Coors Buchner funnel, washed with water (2 x 3 L) and a mixture of hexanes- CH₂Cl₂ (4:1, 2x3 L) and allowed to air dry overnight in pans (1" deep). This was further dried in a vacuum oven (75°C, 0.1 mm Hg, 48 h) to a
10 constant weight of 2072 g (93%) of a white solid, (mp 122-124°C). TLC indicated a trace contamination of the bis DMT product. NMR spectroscopy also indicated that 1-2 mole percent pyridine and about 5 mole percent of hexanes was still present.

15

Preparation of 5'-O-Dimethoxytrityl-2'-deoxy-5-methylcytidine intermediate for 5-methyl-dC amidite

To a 50 L Schott glass-lined steel reactor equipped with an electric stirrer, reagent addition pump (connected to an
20 addition funnel), heating/cooling system, internal thermometer and an Ar gas line was added 5'-O-dimethoxytrityl-thymidine (3.00 kg, 5.51 mol), anhydrous acetonitrile (25 L) and TEA (12.3 L, 88.4 mol, 16 eq). The mixture was chilled with stirring to -10°C internal
25 temperature (external -20°C). Trimethylsilylchloride (2.1 L, 16.5 mol, 3.0 eq) was added over 30 minutes while maintaining the internal temperature below -5°C, followed by a wash of anhydrous acetonitrile (1 L). The reaction is mildly
30 exothermic and copious hydrochloric acid fumes form over the course of the addition. The reaction was allowed to warm to 0°C; and the reaction progress was confirmed by TLC (EtOAc-hexanes 4:1; R_f 0.43 to 0.84 of starting material and silyl product, respectively). Upon completion, triazole (3.05 kg,

-69-

44 mol, 8.0 eq) was added the reaction was cooled to -20°C internal temperature (external -30°C). Phosphorous oxychloride (1035 mL, 11.1 mol, 2.01 eq) was added over 60 min so as to maintain the temperature between -20°C and -10°C during the strongly exothermic process, followed by a wash of anhydrous acetonitrile (1 L). The reaction was warmed to 0 °C and stirred for 1 h. TLC indicated a complete conversion to the triazole product (R_f 0.83 to 0.34 with the product spot glowing in long wavelength UV light). The reaction mixture was a peach-colored thick suspension, which turned darker red upon warming without apparent decomposition. The reaction was cooled to -15°C internal temperature and water (5 L) was slowly added at a rate to maintain the temperature below +10°C in order to quench the reaction and to form a homogenous solution. Caution must be used because this reaction is initially very strongly exothermic. Approximately one-half of the reaction volume (22 L) was transferred by air pump to another vessel, diluted with EtOAc (12 L) and extracted with water (2 x 8 L). The combined water layers were back-extracted with EtOAc (6 L). The water layer was discarded and the organic layers were concentrated in a 20 L rotary evaporator to an oily foam. The foam was coevaporated with anhydrous acetonitrile (4 L) to remove EtOAc. Dioxane may be used instead of anhydrous acetonitrile if dried to a hard foam. The second half of the reaction was treated in the same way. Each residue was dissolved in dioxane (3 L) and concentrated ammonium hydroxide (750 mL) was added. A homogenous solution formed in a few minutes and the reaction was allowed to stand overnight (although the reaction is complete within 1 h).

TLC indicated a complete reaction (product R_f 0.35 in EtOAc-MeOH 4:1). The reaction solution was concentrated on a rotary evaporator to a dense foam. Each foam was slowly

-70-

redissolved in warm EtOAc (4 L; 50°C), combined in a 50 L glass reactor vessel, and extracted with water (2 x 4L) to remove the triazole by-product. The water was back-extracted with EtOAc (2 L). The organic layers were combined and
5 concentrated to about 8 kg total weight, cooled to 0°C and seeded with crystalline product. After 24 hours, the first crop was collected on a 25 cm Coors Buchner funnel and washed repeatedly with EtOAc (3 x 3L) until a white powder was left and then washed with ethyl ether (2 x 3L). The solid was put
10 in pans (1" deep) and allowed to air dry overnight. The filtrate was concentrated to an oil, then redissolved in EtOAc (2 L), cooled and seeded as before. The second crop was collected and washed as before (with proportional solvents) and the filtrate was first extracted with water (2
15 x 1L) and then concentrated to an oil. The residue was dissolved in EtOAc (1 L) and yielded a third crop which was treated as above except that more washing was required to remove a yellow oily layer.

After air-drying, the three crops were dried in a vacuum
20 oven (50°C, 0.1 mm Hg, 24 h) to a constant weight (1750, 600 and 200 g, respectively) and combined to afford 2550 g (85%) of a white crystalline product (MP 215-217°C) when TLC and NMR spectroscopy indicated purity. The mother liquor still contained mostly product (as determined by TLC) and a small
25 amount of triazole (as determined by NMR spectroscopy), bis DMT product and unidentified minor impurities. If desired, the mother liquor can be purified by silica gel chromatography using a gradient of MeOH (0-25%) in EtOAc to further increase the yield.

30

Preparation of 5'-O-Dimethoxytrityl-2'-deoxy-N4-benzoyl-5-methylcytidine penultimate intermediate for 5-methyl dC amidite

-71-

Crystalline 5'-O-dimethoxytrityl-5-methyl-2'-
deoxycytidine (2000 g, 3.68 mol) was dissolved in anhydrous
DMF (6.0 kg) at ambient temperature in a 50 L glass reactor
vessel equipped with an air stirrer and argon line. Benzoic
5 anhydride (Chem Impex not Aldrich, 874 g, 3.86 mol, 1.05 eq)
was added and the reaction was stirred at ambient temperature
for 8 h. TLC (CH₂Cl₂-EtOAc; CH₂Cl₂-EtOAc 4:1; R_f 0.25)
indicated approx. 92% complete reaction. An additional
amount of benzoic anhydride (44 g, 0.19 mol) was added.
10 After a total of 18 h, TLC indicated approx. 96% reaction
completion. The solution was diluted with EtOAc (20 L), TEA
(1020 mL, 7.36 mol, ca 2.0 eq) was added with stirring, and
the mixture was extracted with water (15 L, then 2 x 10 L).
The aqueous layer was removed (no back-extraction was needed)
15 and the organic layer was concentrated in 2 x 20 L rotary
evaporator flasks until a foam began to form. The residues
were coevaporated with acetonitrile (1.5 L each) and dried
(0.1 mm Hg, 25°C, 24 h) to 2520 g of a dense foam. High
pressure liquid chromatography (HPLC) revealed a
20 contamination of 6.3% of N4, 3'-O-dibenzoyl product, but very
little other impurities.

The product was purified by Biotage column
chromatography (5 kg Biotage) prepared with 65:35:1 hexanes-
EtOAc-TEA (4L). The crude product (800 g), dissolved in CH₂Cl₂
25 (2 L), was applied to the column. The column was washed with
the 65:35:1 solvent mixture (20 kg), then 20:80:1 solvent
mixture (10 kg), then 99:1 EtOAc:TEA (17kg). The fractions
containing the product were collected; and any fractions
containing the product and impurities were retained to be
30 resubjected to column chromatography. The column was re-
equilibrated with the original 65:35:1 solvent mixture (17
kg). A second batch of crude product (840 g) was applied to
the column as before. The column was washed with the

-72-

following solvent gradients: 65:35:1 (9 kg), 55:45:1 (20 kg), 20:80:1 (10 kg), and 99:1 EtOAc:TEA(15 kg). The column was reequilibrated as above, and a third batch of the crude product (850 g) plus impure fractions recycled from the two
5 previous columns (28 g) was purified following the procedure for the second batch. The fractions containing pure product combined and concentrated on a 20L rotary evaporator, co-evaporated with acetone (3 L) and dried (0.1 mm Hg, 48 h, 25°C) to a constant weight of 2023 g (85%) of white foam
10 and 20 g of slightly contaminated product from the third run. HPLC indicated a purity of 99.8% with the balance as the diBenzoyl product.

**[5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-deoxy-N⁴-
15 benzoyl-5-methylcytidin-3'-O-yl]-2-cyanoethyl-N,N-
diisopropylphosphoramidite (5-methyl dC amidite)**

5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-deoxy-N⁴-benzoyl-5-methylcytidine (998 g, 1.5 mol) was dissolved in anhydrous DMF (2 L). The solution was co-evaporated with toluene (300
20 ml) at 50°C under reduced pressure, then cooled to room temperature and 2-cyanoethyl tetraisopropylphosphorodiamidite (680 g, 2.26 mol) and tetrazole (52.5 g, 0.75 mol) were added. The mixture was shaken until all tetrazole was dissolved, N-methylimidazole (15 ml) was added and the
25 mixture was left at room temperature for 5 hours. TEA (300 ml) was added, the mixture was diluted with DMF (2.5 L) and water (600 ml), and extracted with hexane (3 x 3 L). The mixture was diluted with water (1.2 L) and extracted with a mixture of toluene (7.5 L) and hexane (6 L). The two layers
30 were separated, the upper layer was washed with DMF-water (7:3 v/v, 3 x 2 L) and water (3 x 2 L), and the phases were separated. The organic layer was dried (Na₂SO₄), filtered and rotary evaporated. The residue was co-evaporated with

-73-

acetonitrile (2 x 2 L) under reduced pressure and dried to a constant weight (25 °C, 0.1mm Hg, 40 h) to afford 1250 g an off-white foam solid (96%).

5 **2'-Fluoro amidites**

2'-Fluorodeoxyadenosine amidites

2'-fluoro oligonucleotides were synthesized as described previously [Kawasaki, *et. al.*, *J. Med. Chem.*, **1993**, *36*, 831-841] and United States patent 5,670,633, herein incorporated
10 by reference. The preparation of 2'-fluoropyrimidines containing a 5-methyl substitution is described in US Patent 5,861,493. Briefly, the protected nucleoside N6-benzoyl-2'-deoxy-2'-fluoroadenosine was synthesized utilizing commercially available 9-beta-D-arabinofuranosyladenine as
15 starting material and whereby the 2'-alpha-fluoro atom is introduced by a S_N2-displacement of a 2'-beta-triflate group. Thus N6-benzoyl-9-beta-D-arabinofuranosyladenine was selectively protected in moderate yield as the 3',5'-ditetrahydropyranyl (THP) intermediate. Deprotection of the
20 THP and N6-benzoyl groups was accomplished using standard methodologies to obtain the 5'-dimethoxytrityl-(DMT) and 5'-DMT-3'-phosphoramidite intermediates.

2'-Fluorodeoxyguanosine

25 The synthesis of 2'-deoxy-2'-fluoroguanosine was accomplished using tetraisopropylidisiloxanyl (TPDS) protected 9-beta-D-arabinofuranosylguanine as starting material, and conversion to the intermediate isobutyryl-arabinofuranosylguanosine. Alternatively, isobutyryl-
30 arabinofuranosylguanosine was prepared as described by Ross *et al.*, (*Nucleosides & Nucleosides*, *16*, 1645, **1997**). Deprotection of the TPDS group was followed by protection of the hydroxyl group with THP to give isobutyryl di-THP

-74-

protected arabinofuranosylguanine. Selective O-deacylation and triflation was followed by treatment of the crude product with fluoride, then deprotection of the THP groups. Standard methodologies were used to obtain the 5'-DMT- and 5'-DMT-3'-
5 phosphoramidites.

2'-Fluorouridine

Synthesis of 2'-deoxy-2'-fluorouridine was accomplished by the modification of a literature procedure in which 2,2'-
10 anhydro-1-beta-D-arabinofuranosyluracil was treated with 70% hydrogen fluoride-pyridine. Standard procedures were used to obtain the 5'-DMT and 5'-DMT-3'phosphoramidites.

2'-Fluorodeoxycytidine

15 2'-deoxy-2'-fluorocytidine was synthesized via amination of 2'-deoxy-2'-fluorouridine, followed by selective protection to give N4-benzoyl-2'-deoxy-2'-fluorocytidine. Standard procedures were used to obtain the 5'-DMT and 5'-
DMT-3'phosphoramidites.

20

2'-O-(2-Methoxyethyl) modified amidites

2'-O-Methoxyethyl-substituted nucleoside amidites (otherwise known as MOE amidites) are prepared as follows, or alternatively, as per the methods of Martin, P., (Helvetica
25 Chimica Acta, 1995, 78, 486-504).

Preparation of 2'-O-(2-methoxyethyl)-5-methyluridine intermediate

2,2'-Anhydro-5-methyl-uridine (2000 g, 8.32 mol),
30 tris(2-methoxyethyl)borate (2504 g, 10.60 mol), sodium bicarbonate (60 g, 0.70 mol) and anhydrous 2-methoxyethanol (5 L) were combined in a 12 L three necked flask and heated to 130 °C (internal temp) at atmospheric pressure, under an

-75-

argon atmosphere with stirring for 21 h. TLC indicated a complete reaction. The solvent was removed under reduced pressure until a sticky gum formed (50-85°C bath temp and 100-11 mm Hg) and the residue was redissolved in water (3 L) and heated to boiling for 30 min in order to hydrolyze the borate esters. The water was removed under reduced pressure until a foam began to form and then the process was repeated. HPLC indicated about 77% product, 15% dimer (5' of product attached to 2' of starting material) and unknown derivatives, and the balance was a single unresolved early eluting peak.

The gum was redissolved in brine (3 L), and the flask was rinsed with additional brine (3 L). The combined aqueous solutions were extracted with chloroform (20 L) in a heavier-than continuous extractor for 70 h. The chloroform layer was concentrated by rotary evaporation in a 20 L flask to a sticky foam (2400 g). This was coevaporated with MeOH (400 mL) and EtOAc (8 L) at 75°C and 0.65 atm until the foam dissolved at which point the vacuum was lowered to about 0.5 atm. After 2.5 L of distillate was collected a precipitate began to form and the flask was removed from the rotary evaporator and stirred until the suspension reached ambient temperature. EtOAc (2 L) was added and the slurry was filtered on a 25 cm table top Buchner funnel and the product was washed with EtOAc (3 x 2 L). The bright white solid was air dried in pans for 24 h then further dried in a vacuum oven (50°C, 0.1 mm Hg, 24 h) to afford 1649 g of a white crystalline solid (mp 115.5-116.5°C).

The brine layer in the 20 L continuous extractor was further extracted for 72 h with recycled chloroform. The chloroform was concentrated to 120 g of oil and this was combined with the mother liquor from the above filtration (225 g), dissolved in brine (250 mL) and extracted once with chloroform (250 mL). The brine solution was continuously

-76-

extracted and the product was crystallized as described above to afford an additional 178 g of crystalline product containing about 2% of thymine. The combined yield was 1827 g (69.4%). HPLC indicated about 99.5% purity with the balance
5 being the dimer.

Preparation of 5'-O-DMT-2'-O-(2-methoxyethyl)-5-methyluridine penultimate intermediate

In a 50 L glass-lined steel reactor, 2'-O-(2-
10 methoxyethyl)-5-methyl-uridine (MOE-T, 1500 g, 4.738 mol),
lutidine (1015 g, 9.476 mol) were dissolved in anhydrous
acetonitrile (15 L). The solution was stirred rapidly and
chilled to -10°C (internal temperature).
Dimethoxytriphenylmethyl chloride (1765.7 g, 5.21 mol) was
15 added as a solid in one portion. The reaction was allowed to
warm to -2°C over 1 h. (Note: The reaction was monitored
closely by TLC (EtOAc) to determine when to stop the reaction
so as to not generate the undesired bis-DMT substituted side
product). The reaction was allowed to warm from -2 to 3°C
20 over 25 min. then quenched by adding MeOH (300 mL) followed
after 10 min by toluene (16 L) and water (16 L). The
solution was transferred to a clear 50 L vessel with a bottom
outlet, vigorously stirred for 1 minute, and the layers
separated. The aqueous layer was removed and the organic
25 layer was washed successively with 10% aqueous citric acid (8
L) and water (12 L). The product was then extracted into the
aqueous phase by washing the toluene solution with aqueous
sodium hydroxide (0.5N, 16 L and 8 L). The combined aqueous
layer was overlaid with toluene (12 L) and solid citric acid
30 (8 moles, 1270 g) was added with vigorous stirring to lower
the pH of the aqueous layer to 5.5 and extract the product
into the toluene. The organic layer was washed with water

-77-

(10 L) and TLC of the organic layer indicated a trace of DMT-O-Me, bis DMT and dimer DMT.

The toluene solution was applied to a silica gel column (6 L sintered glass funnel containing approx. 2 kg of silica gel slurried with toluene (2 L) and TEA (25 mL)) and the fractions were eluted with toluene (12 L) and EtOAc (3 x 4 L) using vacuum applied to a filter flask placed below the column. The first EtOAc fraction containing both the desired product and impurities were resubjected to column chromatography as above. The clean fractions were combined, rotary evaporated to a foam, coevaporated with acetonitrile (6 L) and dried in a vacuum oven (0.1 mm Hg, 40 h, 40°C) to afford 2850 g of a white crisp foam. NMR spectroscopy indicated a 0.25 mole % remainder of acetonitrile (calculates to be approx. 47 g) to give a true dry weight of 2803 g (96%). HPLC indicated that the product was 99.41% pure, with the remainder being 0.06 DMT-O-Me, 0.10 unknown, 0.44 bis DMT, and no detectable dimer DMT or 3'-O-DMT.

20 **Preparation of [5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-O-(2-methoxyethyl)-5-methyluridin-3'-O-yl]-2-cyanoethyl-N,N-diisopropylphosphoramidite (MOE T amidite)**

5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-O-(2-methoxyethyl)-5-methyluridine (1237 g, 2.0 mol) was dissolved in anhydrous DMF (2.5 L). The solution was co-evaporated with toluene (200 ml) at 50°C under reduced pressure, then cooled to room temperature and 2-cyanoethyl tetraisopropylphosphorodiamidite (900 g, 3.0 mol) and tetrazole (70 g, 1.0 mol) were added. The mixture was shaken until all tetrazole was dissolved, N-methylimidazole (20 ml) was added and the solution was left at room temperature for 5 hours. TEA (300 ml) was added, the mixture was diluted with DMF (3.5 L) and water (600 ml) and extracted with hexane (3 x

-78-

3L). The mixture was diluted with water (1.6 L) and extracted with the mixture of toluene (12 L) and hexanes (9 L). The upper layer was washed with DMF-water (7:3 v/v, 3x3 L) and water (3x3 L). The organic layer was dried (Na_2SO_4), filtered and evaporated. The residue was co-evaporated with acetonitrile (2 x 2 L) under reduced pressure and dried in a vacuum oven (25°C, 0.1mm Hg, 40 h) to afford 1526 g of an off-white foamy solid (95%).

10 **Preparation of 5'-O-Dimethoxytrityl-2'-O-(2-methoxyethyl)-5-methylcytidine intermediate**

To a 50 L Schott glass-lined steel reactor equipped with an electric stirrer, reagent addition pump (connected to an addition funnel), heating/cooling system, internal thermometer and argon gas line was added 5'-O-dimethoxytrityl-2'-O-(2-methoxyethyl)-5-methyl-uridine (2.616 kg, 4.23 mol, purified by base extraction only and no scrub column), anhydrous acetonitrile (20 L), and TEA (9.5 L, 67.7 mol, 16 eq). The mixture was chilled with stirring to -10°C internal temperature (external -20°C). Trimethylsilylchloride (1.60 L, 12.7 mol, 3.0 eq) was added over 30 min. while maintaining the internal temperature below -5°C, followed by a wash of anhydrous acetonitrile (1 L). The reaction is mildly exothermic and copious hydrochloric acid fumes form over the course of the addition. The reaction was allowed to warm to 0°C; and the reaction progress was confirmed by TLC (EtOAc, R_f 0.68 and 0.87 for starting material and silyl product, respectively). Upon completion, triazole (2.34 kg, 33.8 mol, 8.0 eq) was added the reaction was cooled to -20°C internal temperature (external -30°C). Phosphorous oxychloride (793 mL, 8.51 mol, 2.01 eq) was added slowly over 60 min so as to maintain the temperature between -20°C and -10°C (strongly exothermic), followed by a wash of anhydrous

-79-

acetonitrile (1 L). The reaction was warmed to 0°C and stirred for 1 h, at which point it was an off-white thick suspension. TLC indicated a complete conversion to the triazole product (EtOAc, R_f 0.87 to 0.75 with the product spot glowing in long wavelength UV light). The reaction was cooled to -15°C; and water (5 L) was slowly added at a rate to maintain the temperature below +10°C in order to quench the reaction and to form a homogenous solution. Caution must be used because this reaction is initially very strongly exothermic. Approximately one-half of the reaction volume (22 L) was transferred by air pump to another vessel, diluted with EtOAc (12 L) and extracted with water (2 x 8 L). The second half of the reaction was treated in the same way. The combined aqueous layers were back-extracted with EtOAc (8 L). The organic layers were combined and concentrated in a 20 L rotary evaporator to an oily foam. The foam was coevaporated with anhydrous acetonitrile (4 L) to remove EtOAc. Dioxane may be used instead of anhydrous acetonitrile if dried to a hard foam. The residue was dissolved in dioxane (2 L) and concentrated ammonium hydroxide (750 mL) was added. A homogenous solution formed in a few minutes and the reaction was allowed to stand overnight

TLC indicated a complete reaction (CH_2Cl_2 -acetone-MeOH, 20:5:3, R_f 0.51). The reaction solution was concentrated on a rotary evaporator to a dense foam and slowly redissolved in warm CH_2Cl_2 (4 L, 40°C) and transferred to a 20 L glass extraction vessel equipped with a air-powered stirrer. The organic layer was extracted with water (2 x 6 L) to remove the triazole by-product. (Note: In the first extraction an emulsion formed which took about 2 h to resolve). The water layer was back-extracted with CH_2Cl_2 (2 x 2 L), which in turn was washed with water (3 L). The combined organic layer was concentrated in 2 x 20 L flasks to a gum and then

-80-

recrystallized from EtOAc seeded with crystalline product. After sitting overnight, the first crop was collected on a 25 cm Coors Buchner funnel and washed repeatedly with EtOAc until a white free-flowing powder was left (about 3 x 3 L).
5 The filtrate was concentrated to an oil recrystallized from EtOAc, and collected as above. The solid was air-dried in pans for 48 h, then further dried in a vacuum oven (50°C, 0.1mm Hg, 17 h) to afford 2248 g of a bright white, dense solid (86%). An HPLC analysis indicated both crops to be
10 99.4% pure and NMR spectroscopy indicated only a faint trace of EtOAc remained.

Preparation of 5'-O-dimethoxytrityl-2'-O-(2-methoxyethyl)-N4-benzoyl-5-methyl-cytidine penultimate intermediate:

15 Crystalline 5'-O-dimethoxytrityl-2'-O-(2-methoxyethyl)-5-methyl-cytidine (1000 g, 1.62 mol) was suspended in anhydrous DMF (3 kg) at ambient temperature and stirred under an Ar atmosphere. Benzoic anhydride (439.3 g, 1.94 mol) was added in one portion. The solution clarified after 5 hours
20 and was stirred for 16 h. HPLC indicated 0.45% starting material remained (as well as 0.32% N4, 3'-O-bis Benzoyl). An additional amount of benzoic anhydride (6.0 g, 0.0265 mol) was added and after 17 h, HPLC indicated no starting material was present. TEA (450 mL, 3.24 mol) and toluene (6 L) were
25 added with stirring for 1 minute. The solution was washed with water (4 x 4 L), and brine (2 x 4 L). The organic layer was partially evaporated on a 20 L rotary evaporator to remove 4 L of toluene and traces of water. HPLC indicated that the bis benzoyl side product was present as a 6%
30 impurity. The residue was diluted with toluene (7 L) and anhydrous DMSO (200 mL, 2.82 mol) and sodium hydride (60% in oil, 70 g, 1.75 mol) was added in one portion with stirring at ambient temperature over 1 h. The reaction was quenched

-81-

by slowly adding then washing with aqueous citric acid (10%, 100 mL over 10 min, then 2 x 4 L), followed by aqueous sodium bicarbonate (2%, 2 L), water (2 x 4 L) and brine (4 L). The organic layer was concentrated on a 20 L rotary evaporator to about 2 L total volume. The residue was purified by silica gel column chromatography (6 L Buchner funnel containing 1.5 kg of silica gel wetted with a solution of EtOAc-hexanes-TEA(70:29:1)). The product was eluted with the same solvent (30 L) followed by straight EtOAc (6 L). The fractions containing the product were combined, concentrated on a rotary evaporator to a foam and then dried in a vacuum oven (50°C, 0.2 mm Hg, 8 h) to afford 1155 g of a crisp, white foam (98%). HPLC indicated a purity of >99.7%.

15 **Preparation of [5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-O-(2-methoxyethyl)-N⁴-benzoyl-5-methylcytidin-3'-O-yl]-2-cyanoethyl-N,N-diisopropylphosphoramidite (MOE 5-Me-C amidite)**

20 5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-O-(2-methoxyethyl)-N⁴-benzoyl-5-methylcytidine (1082 g, 1.5 mol) was dissolved in anhydrous DMF (2 L) and co-evaporated with toluene (300 ml) at 50 °C under reduced pressure. The mixture was cooled to room temperature and 2-cyanoethyl tetraisopropylphosphorodiamidite (680 g, 2.26 mol) and 25 tetrazole (52.5 g, 0.75 mol) were added. The mixture was shaken until all tetrazole was dissolved, N-methylimidazole (30 ml) was added, and the mixture was left at room temperature for 5 hours. TEA (300 ml) was added, the mixture was diluted with DMF (1 L) and water (400 ml) and extracted 30 with hexane (3 x 3 L). The mixture was diluted with water (1.2 L) and extracted with a mixture of toluene (9 L) and hexanes (6 L). The two layers were separated and the upper layer was washed with DMF-water (60:40 v/v, 3 x 3 L) and

-82-

water (3 x 2 L). The organic layer was dried (Na_2SO_4), filtered and evaporated. The residue was co-evaporated with acetonitrile (2 x 2 L) under reduced pressure and dried in a vacuum oven (25 °C, 0.1mm Hg, 40 h) to afford 1336 g of an off-white foam (97%).

Preparation of [5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-O-(2-methoxyethyl)-N⁶-benzoyladenoin-3'-O-yl]-2-cyanoethyl-N,N-diisopropylphosphoramidite (MOE A amidite)

5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-O-(2-methoxyethyl)-N⁶-benzoyladenoin (purchased from Reliable Biopharmaceutical, St. Lois, MO), 1098 g, 1.5 mol) was dissolved in anhydrous DMF (3 L) and co-evaporated with toluene (300 ml) at 50 °C. The mixture was cooled to room temperature and 2-cyanoethyl tetraisopropylphosphorodiamidite (680 g, 2.26 mol) and tetrazole (78.8 g, 1.24 mol) were added. The mixture was shaken until all tetrazole was dissolved, N-methylimidazole (30 ml) was added, and mixture was left at room temperature for 5 hours. TEA (300 ml) was added, the mixture was diluted with DMF (1 L) and water (400 ml) and extracted with hexanes (3 x 3 L). The mixture was diluted with water (1.4 L) and extracted with the mixture of toluene (9 L) and hexanes (6 L). The two layers were separated and the upper layer was washed with DMF-water (60:40, v/v, 3 x 3 L) and water (3 x 2 L). The organic layer was dried (Na_2SO_4), filtered and evaporated to a sticky foam. The residue was co-evaporated with acetonitrile (2.5 L) under reduced pressure and dried in a vacuum oven (25 °C, 0.1mm Hg, 40 h) to afford 1350 g of an off-white foam solid (96%).

Preparation of [5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-O-(2-methoxyethyl)-N⁴-isobutyrylguanosin-3'-O-yl]-2-cyanoethyl-N,N-diisopropylphosphoramidite (MOE G amidite)

- 83 -

5'-O-(4,4'-Dimethoxytriphenylmethyl)-2'-O-(2-methoxyethyl)-N⁴-isobutylguanosine (purchased from Reliable Biopharmaceutical, St. Louis, MO, 1426 g, 2.0 mol) was dissolved in anhydrous DMF (2 L). The solution was co-
5 evaporated with toluene (200 ml) at 50 °C, cooled to room temperature and 2-cyanoethyl tetraisopropylphosphorodiamidite (900 g, 3.0 mol) and tetrazole (68 g, 0.97 mol) were added. The mixture was shaken until all tetrazole was dissolved, N-methylimidazole (30 ml) was added, and the mixture was left
10 at room temperature for 5 hours. TEA (300 ml) was added, the mixture was diluted with DMF (2 L) and water (600 ml) and extracted with hexanes (3 x 3 L). The mixture was diluted with water (2 L) and extracted with a mixture of toluene (10 L) and hexanes (5 L). The two layers were separated and the
15 upper layer was washed with DMF-water (60:40, v/v, 3x3 L). EtOAc (4 L) was added and the solution was washed with water (3 x 4 L). The organic layer was dried (Na₂SO₄), filtered and evaporated to approx. 4 kg. Hexane (4 L) was added, the mixture was shaken for 10 min, and the supernatant liquid was
20 decanted. The residue was co-evaporated with acetonitrile (2 x 2 L) under reduced pressure and dried in a vacuum oven (25 °C, 0.1mm Hg, 40 h) to afford 1660 g of an off-white foamy solid (91%).

25 **2'-O-(Aminoxyethyl) nucleoside amidites and 2'-O-(dimethylaminoxyethyl) nucleoside amidites**

2'-(Dimethylaminooxyethoxy) nucleoside amidites

2'-(Dimethylaminooxyethoxy) nucleoside amidites (also known in the art as 2'-O-(dimethylaminoxyethyl) nucleoside
30 amidites) are prepared as described in the following paragraphs. Adenosine, cytidine and guanosine nucleoside amidites are prepared similarly to the thymidine (5-methyluridine) except the exocyclic amines are protected with

-84-

a benzoyl moiety in the case of adenosine and cytidine and with isobutyryl in the case of guanosine.

5'-O-tert-Butyldiphenylsilyl-O²-2'-anhydro-5-methyluridine

5 **5'-O-tert-Butyldiphenylsilyl-O²-2'-anhydro-5-methyluridine**
 O²-2'-anhydro-5-methyluridine (Pro. Bio. Sint., Varese, Italy, 100.0g, 0.416 mmol), dimethylaminopyridine (0.66g, 0.013eq, 0.0054mmol) were dissolved in dry pyridine (500 ml) at ambient temperature under an argon atmosphere and with
10 mechanical stirring. tert-Butyldiphenylchlorosilane (125.8g, 119.0mL, 1.1eq, 0.458mmol) was added in one portion. The reaction was stirred for 16 h at ambient temperature. TLC (R_f 0.22, EtOAc) indicated a complete reaction. The solution was concentrated under reduced pressure to a thick oil. This was
15 partitioned between CH₂Cl₂ (1 L) and saturated sodium bicarbonate (2 x 1 L) and brine (1 L). The organic layer was dried over sodium sulfate, filtered, and concentrated under reduced pressure to a thick oil. The oil was dissolved in a 1:1 mixture of EtOAc and ethyl ether (600mL) and cooling the
20 solution to -10°C afforded a white crystalline solid which was collected by filtration, washed with ethyl ether (3 x 200 mL) and dried (40°C, 1mm Hg, 24 h) to afford 149g of white solid (74.8%). TLC and NMR spectroscopy were consistent with pure product.

25

5'-O-tert-Butyldiphenylsilyl-2'-O-(2-hydroxyethyl)-5-methyluridine

 In the fume hood, ethylene glycol (350 mL, excess) was added cautiously with manual stirring to a 2 L stainless
30 steel pressure reactor containing borane in tetrahydrofuran (1.0 M, 2.0 eq, 622 mL). Caution must be used as this reaction evolves hydrogen gas. 5'-O-tert-Butyldiphenylsilyl-O²-2'-anhydro-5-methyluridine (149 g, 0.311 mol) and sodium

-85-

bicarbonate (0.074 g, 0.003 eq) were added with manual stirring. The reactor was sealed and heated in an oil bath until an internal temperature of 160 °C was reached and then maintained for 16 h (pressure < 100 psig). The reaction vessel was cooled to ambient temperature and opened. TLC (EtOAc, R_f 0.67 for desired product and R_f 0.82 for ara-T side product) indicated about 70% conversion to the product. The solution was concentrated under reduced pressure (10 to 1mm Hg) in a warm water bath (40-100°C) with the more extreme conditions used to remove the ethylene glycol. (Alternatively, once the THF has evaporated the solution can be diluted with water and the product extracted into EtOAc). The residue was purified by column chromatography (2kg silica gel, EtOAc-hexanes gradient 1:1 to 4:1). The appropriate fractions were combined, evaporated and dried to afford 84 g of a white crisp foam (50%), contaminated starting material (17.4g, 12% recovery) and pure reusable starting material (20g, 13% recovery). TLC and NMR spectroscopy were consistent with 99% pure product.

20

2'-O-([2-phthalimidoxy)ethyl]-5'-*t*-butyldiphenylsilyl-5-methyluridine

5'-O-tert-Butyldiphenylsilyl-2'-O-(2-hydroxyethyl)-5-methyluridine (20g, 36.98mmol) was mixed with triphenylphosphine (11.63g, 44.36mmol) and N-hydroxyphthalimide (7.24g, 44.36mmol) and dried over P₂O₅ under high vacuum for two days at 40°C. The reaction mixture was flushed with argon and dissolved in dry THF (369.8mL, Aldrich, sure seal bottle). Diethyl-azodicarboxylate (6.98mL, 44.36mmol) was added dropwise to the reaction mixture with the rate of addition maintained such that the resulting deep red coloration is just discharged before adding the next drop. The reaction mixture was stirred for 4

-86-

hrs., after which time TLC (EtOAc:hexane, 60:40) indicated that the reaction was complete. The solvent was evaporated in vacuo and the residue purified by flash column chromatography (eluted with 60:40 EtOAc:hexane), to yield 2'-O-([2-phthalimidoxy)ethyl]-5'-*t*-butyldiphenylsilyl-5-methyluridine as white foam (21.819 g, 86%) upon rotary evaporation.

10 **5'-O-*tert*-butyldiphenylsilyl-2'-O-[(2-formadoximinoxy)ethyl]-5-methyluridine**

2'-O-([2-phthalimidoxy)ethyl]-5'-*t*-butyldiphenylsilyl-5-methyluridine (3.1g, 4.5mmol) was dissolved in dry CH₂Cl₂ (4.5mL) and methylhydrazine (300mL, 4.64mmol) was added dropwise at -10°C to 0°C. After 1 h the mixture was filtered, 15 the filtrate washed with ice cold CH₂Cl₂, and the combined organic phase was washed with water and brine and dried (anhydrous Na₂SO₄). The solution was filtered and evaporated to afford 2'-O-(aminooxyethyl) thymidine, which was then dissolved in MeOH (67.5mL). Formaldehyde (20% aqueous 20 solution, w/w, 1.1 eq.) was added and the resulting mixture was stirred for 1 h. The solvent was removed under vacuum and the residue was purified by column chromatography to yield 5'-O-*tert*-butyldiphenylsilyl-2'-O-[(2-formadoximinoxy)ethyl]-5-methyluridine as white foam (1.95 g, 78%) upon 25 rotary evaporation.

5'-O-*tert*-Butyldiphenylsilyl-2'-O-[N,N dimethylaminoxyethyl]-5-methyluridine

30 5'-O-*tert*-butyldiphenylsilyl-2'-O-[(2-formadoximinoxy)ethyl]-5-methyluridine (1.77g, 3.12mmol) was dissolved in a solution of 1M pyridinium *p*-toluenesulfonate (PPTS) in dry MeOH (30.6mL) and cooled to 10°C under inert

-87-

atmosphere. Sodium cyanoborohydride (0.39g, 6.13mmol) was added and the reaction mixture was stirred. After 10 minutes the reaction was warmed to room temperature and stirred for 2 h. while the progress of the reaction was monitored by TLC
5 (5% MeOH in CH₂Cl₂). Aqueous NaHCO₃ solution (5%, 10mL) was added and the product was extracted with EtOAc (2 x 20 mL). The organic phase was dried over anhydrous Na₂SO₄, filtered, and evaporated to dryness. This entire procedure was repeated with the resulting residue, with the exception that
10 formaldehyde (20% w/w, 30 mL, 3.37 mol) was added upon dissolution of the residue in the PPTS/MeOH solution. After the extraction and evaporation, the residue was purified by flash column chromatography and (eluted with 5% MeOH in CH₂Cl₂) to afford 5'-O-*tert*-butyldiphenylsilyl-2'-O-[N,N-
15 dimethylaminoxyethyl]-5-methyluridine as a white foam (14.6g, 80%) upon rotary evaporation.

2'-O-(dimethylaminoxyethyl)-5-methyluridine

Triethylamine trihydrofluoride (3.91mL, 24.0mmol) was
20 dissolved in dry THF and TEA (1.67mL, 12mmol, dry, stored over KOH) and added to 5'-O-*tert*-butyldiphenylsilyl-2'-O-[N,N-dimethylaminoxyethyl]-5-methyluridine (1.40g, 2.4mmol). The reaction was stirred at room temperature for 24 hrs and monitored by TLC (5% MeOH in CH₂Cl₂). The solvent was removed
25 under vacuum and the residue purified by flash column chromatography (eluted with 10% MeOH in CH₂Cl₂) to afford 2'-O-(dimethylaminoxyethyl)-5-methyluridine (766mg, 92.5%) upon rotary evaporation of the solvent.

30 5'-O-DMT-2'-O-(dimethylaminoxyethyl)-5-methyluridine

2'-O-(dimethylaminoxyethyl)-5-methyluridine (750 mg, 2.17 mmol) was dried over P₂O₅ under high vacuum overnight at 40°C, co-evaporated with anhydrous pyridine (20 mL), and

- 88 -

dissolved in pyridine (11 mL) under argon atmosphere. 4-dimethylaminopyridine (26.5 mg, 2.60 mmol) and 4,4'-dimethoxytrityl chloride (880 mg, 2.60 mmol) were added to the pyridine solution and the reaction mixture was stirred at room temperature until all of the starting material had reacted. Pyridine was removed under vacuum and the residue was purified by column chromatography (eluted with 10% MeOH in CH₂Cl₂ containing a few drops of pyridine) to yield 5'-O-DMT-2'-O-(dimethylamino-oxyethyl)-5-methyluridine (1.13g, 80%) upon rotary evaporation.

5'-O-DMT-2'-O-(2-N,N-dimethylaminooxyethyl)-5-methyluridine-3'-[(2-cyanoethyl)-N,N-diisopropylphosphoramidite]

5'-O-DMT-2'-O-(dimethylaminooxyethyl)-5-methyluridine (1.08 g, 1.67 mmol) was co-evaporated with toluene (20 mL), N,N-diisopropylamine tetrazonide (0.29 g, 1.67 mmol) was added and the mixture was dried over P₂O₅ under high vacuum overnight at 40°C. This was dissolved in anhydrous acetonitrile (8.4 mL) and 2-cyanoethyl-N,N,N¹,N¹-tetraisopropylphosphoramidite (2.12 mL, 6.08 mmol) was added. The reaction mixture was stirred at ambient temperature for 4 h under inert atmosphere. The progress of the reaction was monitored by TLC (hexane:EtOAc 1:1). The solvent was evaporated, then the residue was dissolved in EtOAc (70mL) and washed with 5% aqueous NaHCO₃ (40mL). The EtOAc layer was dried over anhydrous Na₂SO₄, filtered, and concentrated. The residue obtained was purified by column chromatography (EtOAc as eluent) to afford 5'-O-DMT-2'-O-(2-N,N-dimethylaminooxyethyl)-5-methyluridine-3'-[(2-cyanoethyl)-N,N-diisopropylphosphoramidite] as a foam (1.04g, 74.9%) upon rotary evaporation.

-89-

2'-(Aminooxyethoxy) nucleoside amidites

2'-(Aminooxyethoxy) nucleoside amidites (also known in the art as 2'-O-(aminooxyethyl) nucleoside amidites) are prepared as described in the following paragraphs. Adenosine, 5 cytidine and thymidine nucleoside amidites are prepared similarly.

N2-isobutyryl-6-O-diphenylcarbamoyl-2'-O-(2-ethylacetyl)-5'-O-(4,4'-dimethoxytrityl)guanosine-3'-
10 **[(2-cyanoethyl)-N,N-diisopropylphosphoramidite]**

The 2'-O-aminooxyethyl guanosine analog may be obtained by selective 2'-O-alkylation of diaminopurine riboside. Multigram quantities of diaminopurine riboside may be purchased from Schering AG (Berlin) to provide 2'-O-(2-ethylacetyl) diaminopurine riboside along with a minor amount 15 of the 3'-O-isomer. 2'-O-(2-ethylacetyl) diaminopurine riboside may be resolved and converted to 2'-O-(2-ethylacetyl)guanosine by treatment with adenosine deaminase. (McGee, D. P. C., Cook, P. D., Guinosso, C. J., WO 94/02501 20 A1 940203.) Standard protection procedures should afford 2'-O-(2-ethylacetyl)-5'-O-(4,4'-dimethoxytrityl)guanosine and 2-N-isobutyryl-6-O-diphenylcarbamoyl-2'-O-(2-ethylacetyl)-5'-O-(4,4'-dimethoxytrityl)guanosine which may be reduced to provide 2-N-isobutyryl-6-O-diphenylcarbamoyl-2'-O-(2- 25 hydroxyethyl)-5'-O-(4,4'-dimethoxytrityl)guanosine. As before the hydroxyl group may be displaced by N-hydroxyphthalimide via a Mitsunobu reaction, and the protected nucleoside may be phosphitylated as usual to yield 2-N-isobutyryl-6-O-diphenylcarbamoyl-2'-O-([2-phthalamidoxy]ethyl)-5'-O-(4,4'- 30 dimethoxytrityl)guanosine-3'-[(2-cyanoethyl)-N,N-diisopropylphosphoramidite].

-90-

2'-dimethylaminoethoxyethoxy (2'-DMAEOE) nucleoside amidites

2'-dimethylaminoethoxyethoxy nucleoside amidites (also known in the art as 2'-O-dimethylaminoethoxyethyl, i.e., 2'-O-CH₂-O-CH₂-N(CH₂)₂, or 2'-DMAEOE nucleoside amidites) are prepared as follows. Other nucleoside amidites are prepared similarly.

2'-O-[2(2-N,N-dimethylaminoethoxy)ethyl]-5-methyl uridine

2[2-(Dimethylamino)ethoxy]ethanol (Aldrich, 6.66 g, 50 mmol) was slowly added to a solution of borane in tetrahydrofuran (1 M, 10 mL, 10 mmol) with stirring in a 100 mL bomb. Hydrogen gas evolves as the solid dissolves. O²-,2'-anhydro-5-methyluridine (1.2 g, 5 mmol), and sodium bicarbonate (2.5 mg) were added and the bomb was sealed, placed in an oil bath and heated to 155°C for 26 hours, then cooled to room temperature. The crude solution was concentrated, the residue was diluted with water (200 mL) and extracted with hexanes (200 mL). The product was extracted from the aqueous layer with EtOAc (3 x 200 mL) and the combined organic layers were washed once with water, dried over anhydrous sodium sulfate, filtered and concentrated. The residue was purified by silica gel column chromatography (eluted with 5:100:2 MeOH/CH₂Cl₂/TEA) as the eluent. The appropriate fractions were combined and evaporated to afford the product as a white solid.

5'-O-dimethoxytrityl-2'-O-[2(2-N,N-dimethylaminoethoxy)ethyl]-5-methyl uridine

To 0.5 g (1.3 mmol) of 2'-O-[2(2-N,N-dimethylaminoethoxy)ethyl]-5-methyl uridine in anhydrous pyridine (8 mL), was added TEA (0.36 mL) and dimethoxytrityl chloride (DMT-Cl, 0.87 g, 2 eq.) and the reaction was stirred for 1 h. The reaction mixture was poured into water (200 mL) and extracted

-91-

with CH₂Cl₂ (2 x 200 mL). The combined CH₂Cl₂ layers were washed with saturated NaHCO₃ solution, followed by saturated NaCl solution, dried over anhydrous sodium sulfate, filtered and evaporated. The residue was purified by silica gel
5 column chromatography (eluted with 5:100:1 MeOH/CH₂Cl₂/TEA) to afford the product.

5'-O-Dimethoxytrityl-2'-O-[2(2-N,N-dimethylaminoethoxy)-ethyl]-5-methyl uridine-3'-O-(cyanoethyl-N,N-
10 diisopropyl)phosphoramidite

Diisopropylaminotetrazolide (0.6 g) and 2-cyanoethoxy-N,N-diisopropyl phosphoramidite (1.1 mL, 2 eq.) were added to a solution of 5'-O-dimethoxytrityl-2'-O-[2(2-N,N-dimethylaminoethoxy)ethyl]-5-methyluridine (2.17 g, 3 mmol)
15 dissolved in CH₂Cl₂ (20 mL) under an atmosphere of argon. The reaction mixture was stirred overnight and the solvent evaporated. The resulting residue was purified by silica gel column chromatography with EtOAc as the eluent to afford the title compound.

20

Example 2

Oligonucleotide synthesis

Unsubstituted and substituted phosphodiester (P=O) oligonucleotides are synthesized on an automated DNA
25 synthesizer (Applied Biosystems model 394) using standard phosphoramidite chemistry with oxidation by iodine.

Phosphorothioates (P=S) are synthesized similar to phosphodiester oligonucleotides with the following exceptions: thiation was effected by utilizing a 10% w/v
30 solution of 3H-1,2-benzodithiole-3-one 1,1-dioxide in acetonitrile for the oxidation of the phosphite linkages. The thiation reaction step time was increased to 180 sec and preceded by the normal capping step. After cleavage from the

-92-

CPG column and deblocking in concentrated ammonium hydroxide at 55°C (12-16 hr), the oligonucleotides were recovered by precipitating with >3 volumes of ethanol from a 1 M NH₄OAc solution. Phosphinate oligonucleotides are prepared as
5 described in U.S. Patent 5,508,270, herein incorporated by reference.

Alkyl phosphonate oligonucleotides are prepared as described in U.S. Patent 4,469,863, herein incorporated by reference.

10 3'-Deoxy-3'-methylene phosphonate oligonucleotides are prepared as described in U.S. Patents 5,610,289 or 5,625,050, herein incorporated by reference.

Phosphoramidite oligonucleotides are prepared as described in U.S. Patent, 5,256,775 or U.S. Patent 5,366,878,
15 herein incorporated by reference.

Alkylphosphonothioate oligonucleotides are prepared as described in published PCT applications PCT/US94/00902 and PCT/US93/06976 (published as WO 94/17093 and WO 94/02499, respectively), herein incorporated by reference.

20 3'-Deoxy-3'-amino phosphoramidate oligonucleotides are prepared as described in U.S. Patent 5,476,925, herein incorporated by reference.

Phosphotriester oligonucleotides are prepared as described in U.S. Patent 5,023,243, herein incorporated by
25 reference.

Borano phosphate oligonucleotides are prepared as described in U.S. Patents 5,130,302 and 5,177,198, both herein incorporated by reference.

30 **Example 3**

Oligonucleoside Synthesis

Methylenemethylimino linked oligonucleosides, also identified as MMI linked oligonucleosides, methylenedimethyl-

-93-

hydrazo linked oligonucleosides, also identified as MDH
linked oligonucleosides, and methylenecarbonylamino linked
oligonucleosides, also identified as amide-3 linked
oligonucleosides, and methyleneaminocarbonyl linked oligo-
5 nucleosides, also identified as amide-4 linked oligonucleo-
sides, as well as mixed backbone compounds having, for
instance, alternating MMI and P=O or P=S linkages are
prepared as described in U.S. Patents 5,378,825, 5,386,023,
5,489,677, 5,602,240 and 5,610,289, all of which are herein
10 incorporated by reference.

Formacetal and thioformacetal linked oligonucleosides
are prepared as described in U.S. Patents 5,264,562 and
5,264,564, herein incorporated by reference.

Ethylene oxide linked oligonucleosides are prepared as
15 described in U.S. Patent 5,223,618, herein incorporated by
reference.

Example 4

PNA Synthesis

20 Peptide nucleic acids (PNAs) are prepared in accordance
with any of the various procedures referred to in Peptide
Nucleic Acids (PNA): Synthesis, Properties and Potential
Applications, *Bioorganic & Medicinal Chemistry*, 1996, 4, 5-
23. They may also be prepared in accordance with U.S.
25 Patents 5,539,082, 5,700,922, and 5,719,262, herein
incorporated by reference.

Example 5

Synthesis of Chimeric Oligonucleotides

30 Chimeric oligonucleotides, oligonucleosides or mixed
oligonucleotides/oligonucleosides of the invention can be of
several different types. These include a first type wherein
the "gap" segment of linked nucleosides is positioned between

-94-

5' and 3' "wing" segments of linked nucleosides and a second "open end" type wherein the "gap" segment is located at either the 3' or the 5' terminus of the oligomeric compound. Oligonucleotides of the first type are also known in the art as "gapmers" or gapped oligonucleotides. Oligonucleotides of the second type are also known in the art as "hemimers" or "wingmers".

[2'-O-Me]--[2'-deoxy]--[2'-O-Me] Chimeric

10 **Phosphorothioate Oligonucleotides**

Chimeric oligonucleotides having 2'-O-alkyl phosphorothioate and 2'-deoxy phosphorothioate oligonucleotide segments are synthesized using an Applied Biosystems automated DNA synthesizer Model 394, as above. Oligonucleotides are synthesized using the automated synthesizer and 2'-deoxy-5'-dimethoxytrityl-3'-O-phosphoramidite for the DNA portion and 5'-dimethoxytrityl-2'-O-methyl-3'-O-phosphoramidite for 5' and 3' wings. The standard synthesis cycle is modified by incorporating coupling steps with increased reaction times for the 5'-dimethoxytrityl-2'-O-methyl-3'-O-phosphoramidite. The fully protected oligonucleotide is cleaved from the support and deprotected in concentrated ammonia (NH₄OH) for 12-16 hr at 55°C. The deprotected oligo is then recovered by an appropriate method (precipitation, column chromatography, volume reduced *in vacuo* and analyzed spectrophotometrically for yield and for purity by capillary electrophoresis and by mass spectrometry.

30 **[2'-O-(2-Methoxyethyl)]--[2'-deoxy]--[2'-O-(Methoxyethyl)] Chimeric Phosphorothioate**

Oligonucleotides

[2'-O-(2-methoxyethyl)]--[2'-deoxy]--[2'-O-

-95-

(methoxyethyl)] chimeric phosphorothioate oligonucleotides were prepared as per the procedure above for the 2'-O-methyl chimeric oligonucleotide, with the substitution of 2'-O-(methoxyethyl) amidites for the 2'-O-methyl amidites.

5

[2'-O-(2-Methoxyethyl)Phosphodiester]--[2'-deoxy Phosphorothioate]--[2'-O-(2-Methoxyethyl) Phosphodiester] Chimeric Oligonucleotides

10 [2'-O-(2-methoxyethyl phosphodiester]--[2'-deoxy phosphorothioate]--[2'-O-(methoxyethyl) phosphodiester] chimeric oligonucleotides are prepared as per the above procedure for the 2'-O-methyl chimeric oligonucleotide with the substitution of 2'-O-(methoxyethyl) amidites for the 2'-O-methyl amidites, oxidation with iodine to generate the
15 phosphodiester internucleotide linkages within the wing portions of the chimeric structures and sulfurization utilizing 3, H-1,2 benzodithiole-3-one 1,1 dioxide (Beaucage Reagent) to generate the phosphorothioate internucleotide linkages for the center gap.

20 Other chimeric oligonucleotides, chimeric oligonucleosides and mixed chimeric oligonucleotides/oligonucleosides are synthesized according to United States patent 5,623,065, herein incorporated by reference.

25

Example 6

Oligonucleotide Isolation

After cleavage from the controlled pore glass solid support and deblocking in concentrated ammonium hydroxide at
30 55°C for 12-16 hours, the oligonucleotides or oligonucleosides are recovered by precipitation out of 1 M NH₄OAc with >3 volumes of ethanol. Synthesized oligonucleotides were analyzed by electrospray mass spectroscopy (molecular weight

-96-

determination) and by capillary gel electrophoresis and judged to be at least 70% full length material. The relative amounts of phosphorothioate and phosphodiester linkages obtained in the synthesis were determined by the ratio of correct molecular weight relative to the -16 amu product (+/- 32 +/-48). For some studies oligonucleotides were purified by HPLC, as described by Chiang *et al.*, *J. Biol. Chem.* **1991**, *266*, 18162-18171. Results obtained with HPLC-purified material were similar to those obtained with non-HPLC purified material.

Example 7

Oligonucleotide Synthesis - 96 Well Plate Format

Oligonucleotides were synthesized via solid phase P(III) phosphoramidite chemistry on an automated synthesizer capable of assembling 96 sequences simultaneously in a 96-well format. Phosphodiester internucleotide linkages were afforded by oxidation with aqueous iodine. Phosphorothioate internucleotide linkages were generated by sulfurization utilizing 3,4-dihydro-2H-benzothiole-3-one 1,1-dioxide (Beaucage Reagent) in anhydrous acetonitrile. Standard base-protected beta-cyanoethyl-diisopropyl phosphoramidites were purchased from commercial vendors (e.g. PE-Applied Biosystems, Foster City, CA, or Pharmacia, Piscataway, NJ). Non-standard nucleosides are synthesized as per standard or patented methods. They are utilized as base protected beta-cyanoethyl-diisopropyl phosphoramidites.

Oligonucleotides were cleaved from support and deprotected with concentrated NH_4OH at elevated temperature (55-60°C) for 12-16 hours and the released product then dried *in vacuo*. The dried product was then re-suspended in sterile water to afford a master plate from which all analytical and

-97-

test plate samples are then diluted utilizing robotic pipettors.

Example 8

5 Oligonucleotide Analysis - 96-Well Plate Format

The concentration of oligonucleotide in each well was assessed by dilution of samples and UV absorption spectroscopy. The full-length integrity of the individual products was evaluated by capillary electrophoresis (CE) in
10 either the 96-well format (Beckman P/ACE™ MDQ) or, for individually prepared samples, on a commercial CE apparatus (e.g., Beckman P/ACE™ 5000, ABI 270). Base and backbone composition was confirmed by mass analysis of the compounds utilizing electrospray-mass spectroscopy. All assay test
15 plates were diluted from the master plate using single and multi-channel robotic pipettors. Plates were judged to be acceptable if at least 85% of the compounds on the plate were at least 85% full length.

20 Example 9

Cell culture and oligonucleotide treatment

The effect of antisense compounds on target nucleic acid expression can be tested in any of a variety of cell types provided that the target nucleic acid is present at
25 measurable levels. This can be routinely determined using, for example, PCR or Northern blot analysis. The following cell types are provided for illustrative purposes, but other cell types can be routinely used, provided that the target is expressed in the cell type chosen. This can be readily
30 determined by methods routine in the art, for example Northern blot analysis, ribonuclease protection assays, or RT-PCR.

-98-

T-24 cells:

The human transitional cell bladder carcinoma cell line T-24 was obtained from the American Type Culture Collection (ATCC) (Manassas, VA). T-24 cells were routinely cultured in complete McCoy's 5A basal media (Invitrogen Corporation, Carlsbad, CA) supplemented with 10% fetal calf serum (Invitrogen Corporation, Carlsbad, CA), penicillin 100 units per mL, and streptomycin 100 micrograms per mL (Invitrogen Corporation, Carlsbad, CA). Cells were routinely passaged by trypsinization and dilution when they reached 90% confluence. Cells were seeded into 96-well plates (Falcon-Primaria #3872) at a density of 7000 cells/well for use in RT-PCR analysis.

For Northern blotting or other analysis, cells may be seeded onto 100 mm or other standard tissue culture plates and treated similarly, using appropriate volumes of medium and oligonucleotide.

A549 cells:

The human lung carcinoma cell line A549 was obtained from the American Type Culture Collection (ATCC) (Manassas, VA). A549 cells were routinely cultured in DMEM basal media (Invitrogen Corporation, Carlsbad, CA) supplemented with 10% fetal calf serum (Invitrogen Corporation, Carlsbad, CA), penicillin 100 units per mL, and streptomycin 100 micrograms per mL (Invitrogen Corporation, Carlsbad, CA). Cells were routinely passaged by trypsinization and dilution when they reached 90% confluence.

30 NHDF cells:

Human neonatal dermal fibroblast (NHDF) were obtained from the Clonetics Corporation (Walkersville, MD). NHDFs were routinely maintained in Fibroblast Growth Medium

-99-

(Clonetics Corporation, Walkersville, MD) supplemented as recommended by the supplier. Cells were maintained for up to 10 passages as recommended by the supplier.

5 HEK cells:

Human embryonic keratinocytes (HEK) were obtained from the Clonetics Corporation (Walkersville, MD). HEKs were routinely maintained in Keratinocyte Growth Medium (Clonetics Corporation, Walkersville, MD) formulated as recommended by
10 the supplier. Cells were routinely maintained for up to 10 passages as recommended by the supplier.

Treatment with antisense compounds:

When cells reached 70% confluency, they were treated
15 with oligonucleotide. For cells grown in 96-well plates, wells were washed once with 100 μ L OPTI-MEM™-1 reduced-serum medium (Invitrogen Corporation, Carlsbad, CA) and then treated with 130 μ L of OPTI-MEM™-1 medium containing 3.75 μ g/mL LIPOFECTIN™ reagent (Invitrogen Corporation, Carlsbad,
20 CA) and the desired concentration of oligonucleotide. After 4-7 hours of treatment, the medium was replaced with fresh medium. Cells were harvested 16-24 hours after oligonucleotide treatment.

The concentration of oligonucleotide used varies from
25 cell line to cell line. To determine the optimal oligonucleotide concentration for a particular cell line, the cells are treated with a positive control oligonucleotide at a range of concentrations. For human cells the positive control oligonucleotide is selected from either ISIS 13920
30 (TCCGTCATCGCTCCTCAGGG, SEQ ID NO: 1) which is targeted to human H-ras, or ISIS 18078, (GTGCGCGAGCCCGAAATC, SEQ ID NO: 2) which is targeted to human Jun-N-terminal kinase-2 (JNK2).

-100-

Both controls are 2'-O-methoxyethyl gapmers (2'-O-methoxyethyls shown in bold) with a phosphorothioate backbone. For mouse or rat cells the positive control oligonucleotide is ISIS 15770, **ATGCATTCTGCCCCCAAGGA**, SEQ ID NO: 3, a 2'-O-methoxyethyl gapmer (2'-O-methoxyethyls shown in bold) with a phosphorothioate backbone which is targeted to both mouse and rat c-raf. The concentration of positive control oligonucleotide that results in 80% inhibition of c-Ha-ras (for ISIS 13920) or c-raf (for ISIS 15770) mRNA is then utilized as the screening concentration for new oligonucleotides in subsequent experiments for that cell line. If 80% inhibition is not achieved, the lowest concentration of positive control oligonucleotide that results in 60% inhibition of H-ras or c-raf mRNA is then utilized as the oligonucleotide screening concentration in subsequent experiments for that cell line. If 60% inhibition is not achieved, that particular cell line is deemed as unsuitable for oligonucleotide transfection experiments. The concentrations of antisense oligonucleotides used herein are from 50 nM to 300 nM.

Example 10

Analysis of oligonucleotide inhibition of VEGF-C expression

Antisense modulation of VEGF-C expression can be assayed in a variety of ways known in the art. For example, VEGF-C mRNA levels can be quantitated by, e.g., Northern blot analysis, competitive polymerase chain reaction (PCR), or real-time PCR (RT-PCR). Real-time quantitative PCR is presently preferred. RNA analysis can be performed on total cellular RNA or poly(A)+ mRNA. The preferred method of RNA analysis of the present invention is the use of total cellular RNA as described in other examples herein. Methods of RNA isolation are taught in, for example, Ausubel, F.M. et

-101-

al., *Current Protocols in Molecular Biology*, Volume 1, pp. 4.1.1-4.2.9 and 4.5.1-4.5.3, John Wiley & Sons, Inc., 1993. Northern blot analysis is routine in the art and is taught in, for example, Ausubel, F.M. et al., *Current Protocols in Molecular Biology*, Volume 1, pp. 4.2.1-4.2.9, John Wiley & Sons, Inc., 1996. Real-time quantitative (PCR) can be conveniently accomplished using the commercially available ABI PRISM™ 7700 Sequence Detection System, available from PE-Applied Biosystems, Foster City, CA and used according to manufacturer's instructions.

Protein levels of VEGF-C can be quantitated in a variety of ways well known in the art, such as immunoprecipitation, Western blot analysis (immunoblotting), ELISA or fluorescence-activated cell sorting (FACS). Antibodies directed to VEGF-C can be identified and obtained from a variety of sources, such as the MSRS catalog of antibodies (Aerie Corporation, Birmingham, MI), or can be prepared via conventional antibody generation methods. Methods for preparation of polyclonal antisera are taught in, for example, Ausubel, F.M. et al., (*Current Protocols in Molecular Biology*, Volume 2, pp. 11.12.1-11.12.9, John Wiley & Sons, Inc., 1997). Preparation of monoclonal antibodies is taught in, for example, Ausubel, F.M. et al., (*Current Protocols in Molecular Biology*, Volume 2, pp. 11.4.1-11.11.5, John Wiley & Sons, Inc., 1997).

Immunoprecipitation methods are standard in the art and can be found at, for example, Ausubel, F.M. et al., (*Current Protocols in Molecular Biology*, Volume 2, pp. 10.16.1-10.16.11, John Wiley & Sons, Inc., 1998). Western blot (immunoblot) analysis is standard in the art and can be found at, for example, Ausubel, F.M. et al., (*Current Protocols in Molecular Biology*, Volume 2, pp. 10.8.1-10.8.21, John Wiley & Sons, Inc., 1997). Enzyme-linked immunosorbent assays

-102-

(ELISA) are standard in the art and can be found at, for example, Ausubel, F.M. et al., (*Current Protocols in Molecular Biology*, Volume 2, pp. 11.2.1-11.2.22, John Wiley & Sons, Inc., 1991).

5

Example 11**Poly(A)+ mRNA isolation**

Poly(A)+ mRNA was isolated according to Miura et al., (*Clin. Chem.*, 1996, 42, 1758-1764). Other methods for
10 poly(A)+ mRNA isolation are taught in, for example, Ausubel, F.M. et al., (*Current Protocols in Molecular Biology*, Volume 1, pp. 4.5.1-4.5.3, John Wiley & Sons, Inc., 1993). Briefly, for cells grown on 96-well plates, growth medium was removed from the cells and each well was washed with 200 μ L cold PBS.
15 60 μ L lysis buffer (10 mM Tris-HCl, pH 7.6, 1 mM EDTA, 0.5 M NaCl, 0.5% NP-40, 20 mM vanadyl-ribonucleoside complex) was added to each well, the plate was gently agitated and then incubated at room temperature for five minutes. 55 μ L of lysate was transferred to Oligo d(T) coated 96-well plates
20 (AGCT Inc., Irvine CA). Plates were incubated for 60 minutes at room temperature, washed 3 times with 200 μ L of wash buffer (10 mM Tris-HCl pH 7.6, 1 mM EDTA, 0.3 M NaCl). After the final wash, the plate was blotted on paper towels to remove excess wash buffer and then air-dried for 5 minutes.
25 60 μ L of elution buffer (5 mM Tris-HCl pH 7.6), preheated to 70°C, was added to each well, the plate was incubated on a 90°C hot plate for 5 minutes, and the eluate was then transferred to a fresh 96-well plate.

Cells grown on 100 mm or other standard plates may be
30 treated similarly, using appropriate volumes of all solutions.

-103-

Example 12**Total RNA Isolation**

Total RNA was isolated using an RNEASY 96™ kit and buffers purchased from Qiagen Inc. (Valencia, CA) following the manufacturer's recommended procedures. Briefly, for cells grown on 96-well plates, growth medium was removed from the cells and each well was washed with 200 µL cold PBS. 150 µL Buffer RLT was added to each well and the plate vigorously agitated for 20 seconds. 150 µL of 70% ethanol was then added to each well and the contents mixed by pipetting three times up and down. The samples were then transferred to the RNEASY 96™ well plate attached to a QIAVAC™ manifold fitted with a waste collection tray and attached to a vacuum source. Vacuum was applied for 1 minute. 500 µL of Buffer RW1 was added to each well of the RNEASY 96™ plate and incubated for 15 minutes and the vacuum was again applied for 1 minute. An additional 500 µL of Buffer RW1 was added to each well of the RNEASY 96™ plate and the vacuum was applied for 2 minutes. 1 mL of Buffer RPE was then added to each well of the RNEASY 96™ plate and the vacuum applied for a period of 90 seconds. The Buffer RPE wash was then repeated and the vacuum was applied for an additional 3 minutes. The plate was then removed from the QIAVAC™ manifold and blotted dry on paper towels. The plate was then re-attached to the QIAVAC™ manifold fitted with a collection tube rack containing 1.2 mL collection tubes. RNA was then eluted by pipetting 170 µL water into each well, incubating 1 minute, and then applying the vacuum for 3 minutes.

The repetitive pipetting and elution steps may be automated using a QIAGEN Bio-Robot 9604 apparatus (Qiagen, Inc., Valencia CA). Essentially, after lysing of the cells

-104-

on the culture plate, the plate is transferred to the robot deck where the pipetting, DNase treatment and elution steps are carried out.

5 **Example 13**

Real-time Quantitative PCR Analysis of VEGF-C mRNA Levels

Quantitation of VEGF-C mRNA levels was determined by real-time quantitative PCR using the ABI PRISM™ 7700 Sequence Detection System (PE-Applied Biosystems, Foster City, CA) according to manufacturer's instructions. This is a closed-tube, non-gel-based, fluorescence detection system which allows high-throughput quantitation of polymerase chain reaction (PCR) products in real-time. As opposed to standard PCR in which amplification products are quantitated after the PCR is completed, products in real-time quantitative PCR are quantitated as they accumulate. This is accomplished by including in the PCR reaction an oligonucleotide probe that anneals specifically between the forward and reverse PCR primers, and contains two fluorescent dyes. A reporter dye (e.g., FAM or JOE, obtained from either PE-Applied Biosystems, Foster City, CA, Operon Technologies Inc., Alameda, CA or Integrated DNA Technologies Inc., Coralville, IA) is attached to the 5' end of the probe and a quencher dye (e.g., TAMRA, obtained from either PE-Applied Biosystems, Foster City, CA, Operon Technologies Inc., Alameda, CA or Integrated DNA Technologies Inc., Coralville, IA) is attached to the 3' end of the probe. When the probe and dyes are intact, reporter dye emission is quenched by the proximity of the 3' quencher dye. During amplification, annealing of the probe to the target sequence creates a substrate that can be cleaved by the 5'-exonuclease activity of Taq polymerase. During the extension phase of the PCR amplification cycle, cleavage of the probe by Taq polymerase releases the reporter

-105-

dye from the remainder of the probe (and hence from the quencher moiety) and a sequence-specific fluorescent signal is generated. With each cycle, additional reporter dye molecules are cleaved from their respective probes, and the fluorescence intensity is monitored at regular intervals by laser optics built into the ABI PRISM™ 7700 Sequence Detection System. In each assay, a series of parallel reactions containing serial dilutions of mRNA from untreated control samples generates a standard curve that is used to quantitate the percent inhibition after antisense oligonucleotide treatment of test samples.

Prior to quantitative PCR analysis, primer-probe sets specific to the target gene being measured are evaluated for their ability to be "multiplexed" with a GAPDH amplification reaction. In multiplexing, both the target gene and the internal standard gene GAPDH are amplified concurrently in a single sample. In this analysis, mRNA isolated from untreated cells is serially diluted. Each dilution is amplified in the presence of primer-probe sets specific for GAPDH only, target gene only ("single-plexing"), or both (multiplexing). Following PCR amplification, standard curves of GAPDH and target mRNA signal as a function of dilution are generated from both the single-plexed and multiplexed samples. If both the slope and correlation coefficient of the GAPDH and target signals generated from the multiplexed samples fall within 10% of their corresponding values generated from the single-plexed samples, the primer-probe set specific for that target is deemed multiplexable. Other methods of PCR are also known in the art.

PCR reagents were obtained from Invitrogen Corporation, (Carlsbad, CA). RT-PCR reactions were carried out by adding 20 μ L PCR cocktail (2.5x PCR buffer (-MgCl₂), 6.6 mM MgCl₂,

-106-

375 μ M each of dATP, dCTP, dGTP and dTTP, 375 nM each of forward primer and reverse primer, 125 nM of probe, 4 Units RNase inhibitor, 1.25 Units PLATINUM[®] Taq, 5 Units MuLV reverse transcriptase, and 2.5x ROX dye) to 96-well plates
5 containing 30 μ L total RNA solution. The RT reaction was carried out by incubation for 30 minutes at 48°C. Following a 10 minute incubation at 95°C to activate the PLATINUM[®] Taq, 40 cycles of a two-step PCR protocol were carried out: 95°C for 15 seconds (denaturation) followed by 60°C for 1.5 minutes
10 (annealing/extension).

Gene target quantities obtained by real time RT-PCR are normalized using either the expression level of GAPDH, a gene whose expression is constant, or by quantifying total RNA using RiboGreen[™] quantification reagent (Molecular Probes, Inc. Eugene, OR). GAPDH expression is quantified by real time
15 RT-PCR, by being run simultaneously with the target, multiplexing, or separately. Total RNA is quantified using RiboGreen[™] RNA reagent and methods of RNA quantification by RiboGreen[™] reagent are taught in Jones, L.J., et al,
20 (Analytical Biochemistry, 1998, 265, 368-374).

In this assay, 170 μ L of RiboGreen[™] working reagent (RiboGreen[™] reagent diluted 1:350 in 10mM Tris-HCl, 1 mM EDTA, pH 7.5) is pipetted into a 96-well plate containing 30
25 μ L purified, cellular RNA. The plate is read in a CytoFluor 4000 instrument (PE Applied Biosystems) with excitation at 480nm and emission at 520nm.

Probes and primers to human VEGF-C were designed to hybridize to a human VEGF-C sequence, using published sequence information (GenBank accession number NM_005429.1,
30 incorporated herein as SEQ ID NO:4). For human VEGF-C the PCR primers were:

-107-

forward primer: TCAGGCAGCGAACAAGACCT (SEQ ID NO: 5)
reverse primer: TTCCTGAGCCAGGCATCTG (SEQ ID NO: 6) and the
PCR probe was: FAM-CCCCACCAATTACATGTGGAATAATCACATCT-TAMRA
(SEQ ID NO: 7) where FAM is the fluorescent dye and TAMRA is
5 the quencher dye. For human GAPDH the PCR primers were:
forward primer: GAAGGTGAAGGTCCGAGTC (SEQ ID NO: 8)
reverse primer: GAAGATGGTGATGGGATTTTC (SEQ ID NO: 9) and the
PCR probe was: 5' JOE-CAAGCTTCCCGTTCTCAGCC- TAMRA 3' (SEQ ID
NO: 10) where JOE is the fluorescent reporter dye and TAMRA
10 is the quencher dye.

Example 14**Northern blot analysis of VEGF-C mRNA levels**

Eighteen hours after antisense treatment, cell
15 monolayers were washed twice with cold PBS and lysed in 1 mL
RNAZOL™ reagent (TEL-TEST "B" Inc., Friendswood, TX). Total
RNA was prepared following manufacturer's recommended
protocols. Twenty micrograms of total RNA was fractionated
by electrophoresis through 1.2% agarose gels containing 1.1%
20 formaldehyde using a MOPS buffer system (AMRESCO, Inc. Solon,
OH). RNA was transferred from the gel to HYBOND™-N+ nylon
membranes (Amersham Pharmacia Biotech, Piscataway, NJ) by
overnight capillary transfer using a Northern/Southern
Transfer buffer system (TEL-TEST "B" Inc., Friendswood, TX).
25 RNA transfer was confirmed by UV visualization. Membranes
were fixed by UV cross-linking using a STRATALINKER™ UV
Crosslinker 2400 instrument (Stratagene, Inc, La Jolla, CA)
and then probed using QUICKHYB™ hybridization solution
(Stratagene, La Jolla, CA) using manufacturer's
30 recommendations for stringent conditions.

To detect human VEGF-C, a human VEGF-C specific probe
was prepared by PCR using the forward primer

-108-

TCAGGCAGCGAACAAAGACCT (SEQ ID NO: 5) and the reverse primer
TTCCTGAGCCAGGCATCTG (SEQ ID NO: 6). To normalize for
variations in loading and transfer efficiency membranes were
stripped and probed for human glyceraldehyde-3-phosphate
5 dehydrogenase (GAPDH) RNA (Clontech, Palo Alto, CA).

Hybridized membranes were visualized and quantitated
using a PHOSPHORIMAGER™ instrument and IMAGEQUANT™ Software
V3.3 (Molecular Dynamics, Sunnyvale, CA). Data was
normalized to GAPDH levels in untreated controls.

10

Example 15**Antisense inhibition of human VEGF-C expression by chimeric
phosphorothioate oligonucleotides having 2'-MOE wings and a
deoxy gap**

15 In accordance with the present invention, a series of
oligonucleotides were designed to target different regions of
the human VEGF-C RNA, using published sequences (GenBank
accession number NM_005429.1, incorporated herein as SEQ ID
NO: 4, GenBank accession number AF020393.1, incorporated
20 herein as SEQ ID NO: 11, GenBank accession number AI342741.1,
incorporated herein as SEQ ID NO: 12, and residues 852000-
911000 of GenBank accession number NT_006118.4, incorporated
herein as SEQ ID NO: 13). The oligonucleotides are shown in
Table 1. "Target site" indicates the first (5'-most)
25 nucleotide number on the particular target sequence to which
the oligonucleotide binds. All compounds in Table 1 are
chimeric oligonucleotides ("gapmers") 20 nucleotides in
length, composed of a central "gap" region consisting of ten
2'-deoxynucleotides, which is flanked on both sides (5' and
30 3' directions) by five-nucleotide "wings". The wings are
composed of 2'-methoxyethyl (2'-MOE)nucleotides. The
internucleoside (backbone) linkages are phosphorothioate
(P=S) throughout the oligonucleotide. All cytidine residues

-109-

are 5-methylcytidines. The compounds were analyzed for their effect on human VEGF-C mRNA levels by quantitative real-time PCR as described in other examples herein. Data are averages from two experiments in which T-24 cells were treated with the antisense oligonucleotides of the present invention. The positive control for each datapoint is identified in the table by sequence ID number. If present, "N.D." indicates "no data".

10

Table 1

Inhibition of human VEGF-C mRNA levels by chimeric phosphorothioate oligonucleotides having 2'-MOE wings and a deoxy gap

ISIS #	REGION	TARGET SEQ ID NO	TARGET SITE	SEQUENCE	% INHIB	SEQ ID NO	CONTROL SEQ ID NO
158101	Coding	4	1257	gcaggccgaagccccgctct	51	14	2
158102	Coding	4	1597	tctttccaatatgaagga	52	15	2
158103	5'UTR	4	194	ggctccgcggttccaacttt	60	16	2
158104	3'UTR	4	1787	ctttgcagatgagctccag	80	17	2
158105	Coding	4	513	ttgctgcataagccgtggc	72	18	2
158106	Coding	4	1203	ttggttggtccacagatgtc	70	19	2
158107	Coding	4	1028	taataatggaatgaacttgt	51	20	2
158108	Coding	4	533	gtaactgctcctccagatct	69	21	2
158109	3'UTR	4	1773	ctccagtcatttctgtaaa	42	22	2
158110	5'UTR	4	30	ggtagctttttggagagg	51	23	2
158111	Coding	4	1520	tcgtacatggccgtctgtaa	79	24	2
158112	Coding	4	1602	tgtggtctttccaatatga	57	25	2
158113	Coding	4	510	ctgcataagccgtggcctc	69	26	2
158114	Coding	4	1184	catggaatccatctgttgag	38	27	2
158115	3'UTR	4	1825	gtttggtcattggcagaaaa	65	28	2
158116	Coding	4	1532	ccttctggcggttcgtacat	46	29	2
158117	Coding	4	1319	gtttgttttacagacacac	67	30	2
158118	Coding	4	1071	ttcgctgcctgacactgtgg	52	31	2
158119	Coding	4	1576	acaacgacacacttcttcac	25	32	2
158120	Coding	4	1010	gtctgtaaaccatccagtta	74	33	2
158121	3'UTR	4	1972	acatattttgcatgatataa	50	34	2
158122	3'UTR	4	1938	gtgagttttaccaattgttg	73	35	2
158123	3'UTR	4	1699	caaggtctctctgttcaca	54	36	2
158124	5'UTR	4	127	gtaaaagcctcacaggaaac	61	37	2
158125	Coding	4	1555	atatgaaaatcctggctcac	15	38	2
158126	Coding	4	825	gacacacatggaggtttaa	72	39	2
158127	Coding	4	743	attgagctttctccactca	75	40	2
158128	Coding	4	1136	aatcttctgagccaggcat	35	41	2
158129	3'UTR	4	1780	agatgagctccagtcattt	34	42	2

-110-

158130	3'UTR	4	1831	ttggctgtttggtcattggc	74	43	2
158131	Start Codon	4	362	gcaagtgcattggtggaagga	81	44	2
158132	3'UTR	4	1905	gaatgcagaaacaatatttt	25	45	2
158133	3'UTR	4	1867	tagtcattcttttaagaaa	34	46	2
158134	3'UTR	4	1832	cttggctgtttggtcattgg	75	47	2
158135	3'UTR	4	1840	aggaaaatcttggctgtttg	55	48	2
158136	Coding	4	615	ccttccttagctgacactt	45	49	2
158137	Coding	4	1005	taaacaatccagtttagacat	32	50	2
196823	5'UTR	4	138	gcgggtgtcaggtaaaagcc	70	51	2
196824	5'UTR	4	282	aggccgccccctcctgg	62	52	2
196825	Start Codon	4	372	aagaagcccagcaagtgcatt	78	53	2
196826	Coding	4	548	cactggacacagaccgtaac	88	54	2
196827	Coding	4	663	tctgtccttgagttgaggtt	73	55	2
196828	Coding	4	717	atacttttcaagatctctgt	79	56	2
196829	Coding	4	723	ttatcaatacttttcaagat	38	57	2
196830	Coding	4	856	actattgcagcaacccccac	78	58	2
196831	Coding	4	1049	gtgttgctggcaggaacgt	78	59	2
196832	Coding	4	1163	catctccagcatccgaggaa	11	60	2
196833	Coding	4	1215	tcataccagctccttgtttgg	33	61	2
196834	Coding	4	1268	gtccacagctggcaggccga	62	62	2
196835	Coding	4	1388	ttcttttacatacactgg	66	63	2
196836	Coding	4	1438	acattcacaggcacattttc	74	64	2
196837	Coding	4	1487	ggtggtggaacttctttcct	77	65	2
196838	Stop Codon	4	1620	gtacaatcttagctcatttg	80	66	2
196839	3'UTR	4	1683	cacagacagttctactgtgg	76	67	2
196840	3'UTR	4	1878	aataaattatagtcattc	0	68	2
196841	3'UTR	4	1926	aattgttgctgctataaaaa	2	69	2
196842	5'UTR	11	3	agtgcctgatgatccaaga	27	70	2
196843	5'UTR	11	156	ttatcctcggcactcccg	32	71	2
196844	5'UTR	11	233	tttagaggtgatgacaccac	9	72	2
196845	5'UTR	11	389	ctccctggagctccccgttt	55	73	2
196846	5'UTR	11	408	actctccctcggaagccgtc	0	74	2
196847	5'UTR	11	684	cctccccgaagtggagga	68	75	2
196848	intron	12	2	tgacgaaattgttaaaaggt	11	76	2
196849	intron	12	33	atttcagactgaaatacaat	22	77	2
196850	exon: intron junction	13	2359	gcagacctaccgtggcctcg	82	78	2
196851	exon: intron junction	13	12230	ccatacttacttttcaagat	9	79	2
196852	exon: intron junction	13	13985	caatacccaccgtcttgctg	76	80	2
196853	intron	13	16405	acacattttgtacaggtatc	89	81	2
196854	intron	13	16859	aggaaacacgatgatgcca	70	82	2
196855	intron	13	22515	ggagaactttgaagcagttt	48	83	2
196856	exon: intron junction	13	30058	tcatactcactgtggtagtg	45	84	2
196857	intron	13	41999	attctttcattgtcagagct	66	85	2

As shown in Table 1, SEQ ID NOs 16, 17, 18, 19, 21, 24,

-111-

25, 26, 28, 30, 33, 35, 37, 39, 40, 43, 44, 47, 48, 51, 52, 53, 54, 55, 56, 58, 59, 62, 63, 64, 65, 66, 67, 73, 75, 78, 80, 81, 82 and 85 demonstrated at least 55% inhibition of human VEGF-C expression in this assay and are therefore preferred. The target sites to which these preferred sequences are complementary are herein referred to as "preferred target regions" and are therefore preferred sites for targeting by compounds of the present invention. These preferred target regions are shown in Table 2. The sequences represent the reverse complement of the preferred antisense compounds shown in Table 1. "Target site" indicates the first (5'-most) nucleotide number of the corresponding target nucleic acid. Also shown in Table 2 is the species in which each of the preferred target regions was found.

15

Table 2
Sequence and position of preferred target regions identified in VEGF-C.

SITE ID	TARGET SEQ ID NO	TARGET SITE	SEQUENCE	REV COMP OF SEQ ID	ACTIVE IN	SEQ ID NO
73604	4	194	aaagttgggaacgcgagcc	16	<i>H. sapiens</i>	86
73605	4	1787	ctggagctcatctgcaaaag	17	<i>H. sapiens</i>	87
73606	4	513	gccacggcttatgcaagcaa	18	<i>H. sapiens</i>	88
73607	4	1203	gacatctgtggaccaacaa	19	<i>H. sapiens</i>	89
73609	4	533	agatctggaggagcagttac	21	<i>H. sapiens</i>	90
73612	4	1520	ttacagacggccatgtacga	24	<i>H. sapiens</i>	91
73613	4	1602	tcatattggaaaagaccaca	25	<i>H. sapiens</i>	92
73614	4	510	gaggccacggcttatgcaag	26	<i>H. sapiens</i>	93
73616	4	1825	ttttctgccaatgaccaaac	28	<i>H. sapiens</i>	94
73618	4	1319	gtgtgtctgtaaaaacaaac	30	<i>H. sapiens</i>	95
73621	4	1010	taaactggatgtttacagac	33	<i>H. sapiens</i>	96
73623	4	1938	caacaattggtaaaactcac	35	<i>H. sapiens</i>	97
73625	4	127	gtttcctgtgaggcttttac	37	<i>H. sapiens</i>	98
73627	4	825	tttaaacctccatgtgtgtc	39	<i>H. sapiens</i>	99
73628	4	743	tgagtggagaagactcaat	40	<i>H. sapiens</i>	100
73631	4	1831	gccaatgaccaaactgcaaa	43	<i>H. sapiens</i>	101
73632	4	362	tccttccaccatgcacttgc	44	<i>H. sapiens</i>	102
73635	4	1832	ccaatgaccaaactgcaaa	47	<i>H. sapiens</i>	103
73636	4	1840	caaacagccaagattttcct	48	<i>H. sapiens</i>	104
114955	4	138	ggcttttacctgacacccgc	51	<i>H. sapiens</i>	105
114956	4	282	ccaggaggggcccggcct	52	<i>H. sapiens</i>	106

-112-

114957	4	372	atgcacttgctgggcttctt	53	<i>H. sapiens</i>	107
114958	4	548	gttacgggtctgtgtccagtg	54	<i>H. sapiens</i>	108
114959	4	663	aacctcaactcaaggacaga	55	<i>H. sapiens</i>	109
114960	4	717	acagagatcttgaaaagtat	56	<i>H. sapiens</i>	110
114962	4	856	gtgggggttgctgcaatagt	58	<i>H. sapiens</i>	111
114963	4	1049	acgttccctgccagcaacac	59	<i>H. sapiens</i>	112
114966	4	1268	tcggcctgccagctgtggac	62	<i>H. sapiens</i>	113
114967	4	1388	ccagtgtgtatgtaaaagaa	63	<i>H. sapiens</i>	114
114968	4	1438	gaaaatgtgcctgtgaatgt	64	<i>H. sapiens</i>	115
114969	4	1487	aggaaagaagttccaccacc	65	<i>H. sapiens</i>	116
114970	4	1620	caaatgagctaagattgtac	66	<i>H. sapiens</i>	117
114971	4	1683	ccacagtagaactgtctgtg	67	<i>H. sapiens</i>	118
114977	11	389	aaacggggagctccagggag	73	<i>H. sapiens</i>	119
114979	11	684	tcctctcacttcggggaagg	75	<i>H. sapiens</i>	120
114982	13	2359	cgaggccacggtaggtctgc	78	<i>H. sapiens</i>	121
114984	13	13985	gagcaagacgggtgggtattg	80	<i>H. sapiens</i>	122
114985	13	16405	gatacctgtacaaaatgtgt	81	<i>H. sapiens</i>	123
114986	13	16859	tgggcatcatcgtgtttcct	82	<i>H. sapiens</i>	124
114989	13	41999	agctctgacaatgaaagaat	85	<i>H. sapiens</i>	125

As these "preferred target regions" have been found by experimentation to be open to, and accessible for, hybridization with the antisense compounds of the present invention, one of skill in the art will recognize or be able to ascertain, using no more than routine experimentation, further embodiments of the invention that encompass other compounds that specifically hybridize to these sites and consequently inhibit the expression of VEGF-C.

In one embodiment, the "preferred target region" may be employed in screening candidate antisense compounds. "Candidate antisense compounds" are those that inhibit the expression of a nucleic acid molecule encoding VEGF-C and which comprise at least an 8-nucleobase portion which is complementary to a preferred target region. The method comprises the steps of contacting a preferred target region of a nucleic acid molecule encoding VEGF-C with one or more candidate antisense compounds, and selecting for one or more candidate antisense compounds which inhibit the expression of a nucleic acid molecule encoding VEGF-C. Once it is shown that the candidate antisense compound or compounds are capable of inhibiting the expression of a nucleic acid

-113-

molecule encoding VEGF-C, the candidate antisense compound may be employed as an antisense compound in accordance with the present invention.

According to the present invention, antisense compounds include ribozymes, external guide sequence (EGS) oligonucleotides (oligozymes), and other short catalytic RNAs or catalytic oligonucleotides which hybridize to the target nucleic acid and modulate its expression.

10 **Example 16**

Western blot analysis of VEGF-C protein levels

Western blot analysis (immunoblot analysis) is carried out using standard methods. Cells are harvested 16-20 h after oligonucleotide treatment, washed once with PBS, suspended in Laemmli buffer (100 ul/well), boiled for 5 minutes and loaded on a 16% SDS-PAGE gel. Gels are run for 1.5 hours at 150 V, and transferred to membrane for western blotting. Appropriate primary antibody directed to VEGF-C is used, with a radiolabeled or fluorescently labeled secondary antibody directed against the primary antibody species. Bands are visualized using a PHOSPHORIMAGER™ instrument (Molecular Dynamics, Sunnyvale CA).

-114-

What is claimed is:

1. A compound 8 to 80 nucleobases in length targeted to a nucleic acid molecule encoding VEGF-C, wherein said compound specifically hybridizes with said nucleic acid molecule encoding VEGF-C and inhibits the expression of VEGF-C.
2. The compound of claim 1 which is an antisense oligonucleotide.
3. The compound of claim 2 wherein the antisense oligonucleotide comprises at least one modified internucleoside linkage.
4. The compound of claim 3 wherein the modified internucleoside linkage is a phosphorothioate linkage.
5. The compound of claim 2 wherein the antisense oligonucleotide comprises at least one modified sugar moiety.
6. The compound of claim 5 wherein the modified sugar moiety is a 2'-O-methoxyethyl sugar moiety.
7. The compound of claim 2 wherein the antisense oligonucleotide comprises at least one modified nucleobase.
8. The compound of claim 7 wherein the modified nucleobase is a 5-methylcytosine.
9. The compound of claim 2 wherein the antisense oligonucleotide is a chimeric oligonucleotide.

-115-

10. A compound 8 to 80 nucleobases in length which specifically hybridizes with at least an 8-nucleobase portion of a preferred target region on a nucleic acid molecule encoding VEGF-C.
11. A composition comprising the compound of claim 1 and a pharmaceutically acceptable carrier or diluent.
12. The composition of claim 11 further comprising a colloidal dispersion system.
13. The composition of claim 11 wherein the compound is an antisense oligonucleotide.
14. A method of inhibiting the expression of VEGF-C in cells or tissues comprising contacting said cells or tissues with the compound of claim 1 so that expression of VEGF-C is inhibited.
15. A method of treating an animal having a disease or condition associated with VEGF-C comprising administering to said animal a therapeutically or prophylactically effective amount of the compound of claim 1 so that expression of VEGF-C is inhibited.
16. The method of claim 15 wherein the disease or condition is a cardiovascular disorder.
17. The method of claim 16 wherein the cardiovascular disorder is atherosclerosis or diabetic retinopathy.
18. The method of claim 15 wherein the disease or condition is an inflammatory disorder.

-116-

19. The method of claim 15 wherein the disease or condition is an autoimmune disorder.
20. A method of screening for an antisense compound, the method comprising the steps of:
 - a. contacting a preferred target region of a nucleic acid molecule encoding VEGF-C with one or more candidate antisense compounds, said candidate antisense compounds comprising at least an 8-nucleobase portion which is complementary to said preferred target region, and
 - b. selecting for one or more candidate antisense compounds which inhibit the expression of a nucleic acid molecule encoding VEGF-C.
21. Use of a therapeutically or prophylactically effective amount of the compound of claim 1 in a medicament for treating an animal having a disease or condition associated with VEGF-C so that expression of VEGF-C is inhibited.
22. Use according to claim 21 wherein the disease or condition is a cardiovascular disorder.
23. Use according to claim 22 wherein the cardiovascular disorder is atherosclerosis or diabetic retinopathy.
24. 22. Use according to claim 21, wherein the disease or condition is an inflammatory disorder.
25. Use according to claim 21, wherein the disease or condition is an autoimmune disorder.

SEQUENCE LISTING

<110> Hong Zhang
Kenneth W. Dobie
Isis Pharmaceuticals, Inc.

<120> ANTISENSE MODULATION OF VEGF-C EXPRESSION

<130> PTS-0036WO

<150> 10/173,718

<151> 2002-06-17

<160> 125

<210> 1

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 1

tccgcatcgc ctcctcaggg

20

<210> 2

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 2
 gtgcgcgcgga gcccgaaatc 20

<210> 3
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 3
 atgcattctg cccccaagga 20

<210> 4
 <211> 2015
 <212> DNA
 <213> H. sapiens

<220>

<220>
 <221> CDS
 <222> (372)...(1631)

<400> 4
 cgcggggtgt tctggtgtcc cccgccccgc ctctccaaaa agctacaccg acgcggaccg 60
 cggcggcgct ctccctcgcc ctcgcttcac ctcgcgggct ccgaatgcgg ggagctcgga 120
 tgtccggttt cctgtgaggc ttttacctga cacccgccgc ctttccccgg cactggctgg 180
 gagggcgccc tgcaaagtgt ggaacgcgga gccccggacc cgctccccgcc gcctccggct 240
 cgcccagggg gggtcgcccg gaggagcccg ggggagaggg accaggaggg gcccgcggcc 300
 tcgcaggggc gcccgcgccc ccaccctgc ccccgccagc ggaccggtcc cccacccccg 360
 gtccttcac c atg cac ttg ctg ggc ttc ttc tct gtg gcg tgt tct ctg 410
 Met His Leu Leu Gly Phe Phe Ser Val Ala Cys Ser Leu
 1 5 10

ctc gcc gct gcg ctg ctc ccg ggt cct cgc gag gcg ccc gcc gcc gcc 458
 Leu Ala Ala Ala Leu Leu Pro Gly Pro Arg Glu Ala Pro Ala Ala Ala
 15 20 25

gcc gcc ttc gag tcc gga ctc gac ctc tcg gac gcg gag ccc gac gcg 506
 Ala Ala Phe Glu Ser Gly Leu Asp Leu Ser Asp Ala Glu Pro Asp Ala
 30 35 40 45

ggc gag gcc acg gct tat gca agc aaa gat ctg gag gag cag tta cgg 554
 Gly Glu Ala Thr Ala Tyr Ala Ser Lys Asp Leu Glu Glu Gln Leu Arg
 50 55 60

tct gtg tcc agt gta gat gaa ctc atg act gta ctc tac cca gaa tat 602
 Ser Val Ser Ser Val Asp Glu Leu Met Thr Val Leu Tyr Pro Glu Tyr
 65 70 75

tgg aaa atg tac aag tgt cag cta agg aaa gga gcc tgg caa cat aac 650
 Trp Lys Met Tyr Lys Cys Gln Leu Arg Lys Gly Gly Trp Gln His Asn
 80 85 90

aga gaa cag gcc aac ctc aac tca agg aca gaa gag act ata aaa ttt 698
 Arg Glu Gln Ala Asn Leu Asn Ser Arg Thr Glu Glu Thr Ile Lys Phe
 95 100 105

gct gca gca cat tat aat aca gag atc ttg aaa agt att gat aat gag 746
 Ala Ala Ala His Tyr Asn Thr Glu Ile Leu Lys Ser Ile Asp Asn Glu
 110 115 120 125

tgg aga aag act caa tgc atg cca cgg gag gtg tgt ata gat gtg ggg 794
 Trp Arg Lys Thr Gln Cys Met Pro Arg Glu Val Cys Ile Asp Val Gly
 130 135 140

aag gag ttt gga gtc gcg aca aac acc ttc ttt aaa cct cca tgt gtg 842
 Lys Glu Phe Gly Val Ala Thr Asn Thr Phe Phe Lys Pro Pro Cys Val
 145 150 155

tcc gtc tac aga tgt ggg ggt tgc tgc aat agt gag ggg ctg cag tgc 890

Ser Val Tyr Arg Cys Gly Gly Cys Cys Asn Ser Glu Gly Leu Gln Cys	
160	165 170
atg aac acc agc acg agc tac ctc agc aag acg tta ttt gaa att aca	938
Met Asn Thr Ser Thr Ser Tyr Leu Ser Lys Thr Leu Phe Glu Ile Thr	
175	180 185
gtg cct ctc tct caa ggc ccc aaa cca gta aca atc agt ttt gcc aat	986
Val Pro Leu Ser Gln Gly Pro Lys Pro Val Thr Ile Ser Phe Ala Asn	
190	195 200 205
cac act tcc tgc cga tgc atg tct aaa ctg gat gtt tac aga caa gtt	1034
His Thr Ser Cys Arg Cys Met Ser Lys Leu Asp Val Tyr Arg Gln Val	
	210 215 220
cat tcc att att aga cgt tcc ctg cca gca aca cta cca cag tgt cag	1082
His Ser Ile Ile Arg Arg Ser Leu Pro Ala Thr Leu Pro Gln Cys Gln	
	225 230 235
gca gcg aac aag acc tgc ccc acc aat tac atg tgg aat aat cac atc	1130
Ala Ala Asn Lys Thr Cys Pro Thr Asn Tyr Met Trp Asn Asn His Ile	
	240 245 250
tgc aga tgc ctg gct cag gaa gat ttt atg ttt tcc tcg gat gct gga	1178
Cys Arg Cys Leu Ala Gln Glu Asp Phe Met Phe Ser Ser Asp Ala Gly	
	255 260 265
gat gac tca aca gat gga ttc cat gac atc tgt gga cca aac aag gag	1226
Asp Asp Ser Thr Asp Gly Phe His Asp Ile Cys Gly Pro Asn Lys Glu	
270	275 280 285
ctg gat gaa gag acc tgt cag tgt gtc tgc aga gcg ggg ctt cgg cct	1274
Leu Asp Glu Glu Thr Cys Gln Cys Val Cys Arg Ala Gly Leu Arg Pro	
	290 295 300
gcc agc tgt gga ccc cac aaa gaa cta gac aga aac tca tgc cag tgt	1322
Ala Ser Cys Gly Pro His Lys Glu Leu Asp Arg Asn Ser Cys Gln Cys	
	305 310 315

gtc tgt aaa aac aaa ctc ttc ccc agc caa tgt ggg gcc aac cga gaa 1370
 Val Cys Lys Asn Lys Leu Phe Pro Ser Gln Cys Gly Ala Asn Arg Glu
 320 325 330

ttt gat gaa aac aca tgc cag tgt gta tgt aaa aga acc tgc ccc aga 1418
 Phe Asp Glu Asn Thr Cys Gln Cys Val Cys Lys Arg Thr Cys Pro Arg
 335 340 345

aat caa ccc cta aat cct gga aaa tgt gcc tgt gaa tgt aca gaa agt 1466
 Asn Gln Pro Leu Asn Pro Gly Lys Cys Ala Cys Glu Cys Thr Glu Ser
 350 355 360 365

cca cag aaa tgc ttg tta aaa gga aag aag ttc cac cac caa aca tgc 1514
 Pro Gln Lys Cys Leu Leu Lys Gly Lys Lys Phe His His Gln Thr Cys
 370 375 380

agc tgt tac aga cgg cca tgt acg aac cgc cag aag gct tgt gag cca 1562
 Ser Cys Tyr Arg Arg Pro Cys Thr Asn Arg Gln Lys Ala Cys Glu Pro
 385 390 395

gga ttt tca tat agt gaa gaa gtg tgt cgt tgt gtc cct tca tat tgg 1610
 Gly Phe Ser Tyr Ser Glu Glu Val Cys Arg Cys Val Pro Ser Tyr Trp
 400 405 410

aaa aga cca caa atg agc taa gattgtactg ttttccagtt catogatttt 1661
 Lys Arg Pro Gln Met Ser
 415

ctattatgga aaactgtggt gccacagtag aactgtctgt gaacagagag acccttgtgg 1721
 gtccatgcta acaaagacaa aagtctgtct ttctgaacc atgtggataa ctttacagaa 1781
 atggactgga gctcatctgc aaaaggcctc ttgtaaagac tggttttctg ccaatgacca 1841
 aacagccaag attttctct tgtgatttct ttaaaagaat gactatataa tttatttcca 1901
 ctaaaaatat tgtttctgca ttcattttta tagcaacaac aattggtaaa actcactgtg 1961
 atcaatattt ttatatcatg caaaatatgt ttaaaataaa atgaaaattg tatt 2015

<210> 5

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 5

tcaggcagcg aacaagacct

20

<210> 6

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 6

ttcctgagcc aggcattctg

19

<210> 7

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Probe

<400> 7

ccccaccaat tacatgtgga ataatcacat ct

32

<210> 8

<211> 19

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 8

gaaggtgaag gtcggagtc

19

<210> 9

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 9

gaagatggtg atgggatttc

20

<210> 10

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Probe

<400> 10

caagcttccc gttctcagcc

20

<210> 11
 <211> 1127
 <212> DNA
 <213> H. sapiens

<220>

<220>
 <221> CDS
 <222> (1125) ... (1127)

<400> 11
 gttcttggat catcaggcaa ctttcaacta cacagaccaa gggagagagg ggacccctcc 60
 gaggtcccat agggttctct gacatagtga tgaccttttt ccaaactttg agcagggcgc 120
 tgggggcccag gcggtgcggga gggaggacaa gaactcggga gtggccgagg ataaagcggg 180
 ggctccctcc accccacggg gccagtttc tccccgctgc acgtggtcca gggtggtcgc 240
 atcacctcta aagccgggtcc cgccaaccgc cagccccggg actgaacttg cccctccggc 300
 cgcccgctcc ccgcagggga caggggcggg gagggagaga tccagagggg ggctggggga 360
 ggtggggccg ccggggagga ggcgagggaa acggggagct ccagggagac ggcttccgag 420
 ggagagtgag aggggagggc agcccgggct cggcacgctc cctccctcgg ccgctttctc 480
 tcacataagc gcaggcagag ggcgctcag tcatgccctg cccctgcgcc cgccgcccgc 540
 gccgcccgcg ctccagcccg cgcgctctgg aggatcctgc gccgcccgcg tcccgggccc 600
 cgccgcccgc agccgccccg gcggccctcc tcccgcccc ggcaaccgcg ccagcgcccc 660
 cgccgagcgc cccgcccgcg ggctcctctc acttcgggga aggggagggg ggagggggac 720
 gagggctctg gcggggttg aggggctgaa catcgcgggg tgttctggtg tccccgccc 780
 cgcctctcca aaaagctaca ccgacgcgga ccgcccggc gtccctccctc gccctcgctt 840
 cacctcgcgg gctccgaatg cggggagctc ggatgtccgg tttcctgtga ggcttttacc 900
 tgacaccgcg cgctttccc cggcaactggc tgggagggcg ccttgcaaag ttgggaacgc 960
 ggagcccccg acccgctccc gccgcctccg gctcgcccag ggggggtcgc cgggaggagc 1020
 ccgggggaga gggaccagga ggggcccgcg gcctcgcagg ggcgcccgcg cccccacccc 1080
 tgccccgcg agcggaccgg tccccaccc ccggtccttc caccatg 1127

<210> 12
 <211> 409
 <212> DNA
 <213> H. sapiens

<220>

<400> 12

```

aaccttttaa caatttcgtc ataaaatgag taattgtatt tcagtctgaa atttaaaaac 60
acagaaatac cttggtagca tgatgaatcc attgccttga tctttaacgt aagtgtgttc 120
ttgtttgttc tcctagctgt tacagacggc catgtacgaa ccgccagaag gcttgtgagc 180
caggattttc ataatgtgaa gaagtgtgtc gttgtgtccc ttcataattgg aaaagaccac 240
aaatgagcta agattgtact gttttccagt tcatcgattt tctattatgg aaaactgtgt 300
tgccacagta gaactgtctg tgaacagaga gacccttgtg ggtccatgct aacaaagaca 360
aaagtctgtc tttcctgaac catgtggata actttacaga aatggactg 409
    
```

<210> 13

<211> 59001

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> 10057-10156

<223> n = A,T,C or G

<221> misc_feature

<222> 53857-53956

<223> n = A,T,C or G

<400> 13

```

gagtctgatg ggatggaatt tcataaagat acataaaaaa gcatcttggg tacagtaaac 60
ttaactccac aaatacaggg gaatttagac gtgactaagt agcagtacat atgaaaaatt 120
attgaggaat tttgttgact ttaagggtag tgtgagtcaa cactgtgatt tggctgccag 180
aaaataaact caatccaagg ctgtatcaac aaaggcatac tgtccattct gcatgctcat 240
tacagcacta agtaccgagc catgtttctca accgcatact tcatgaacat ggaaagctaa 300
cagtatggtt aaggggggaa actggaactg tcatcttggg gaataaaagg gatatttagc 360
caggagtaaa gtttagcttag ggagaccatg ataaatattt tcaaaatatt tgaaggactc 420
agttgtggaa gtgagattag atttattgtg taaaactcca ggagtcaaaa gcaatagaga 480
    
```

gatagaagga aatgcttttc agcagtgttg ctcatcaata aagggagtga acagccacac 540
agaatggaag gttccctgtc ctttgagata ttttaagcctt caagtaaatt atgggtgagg 600
agtttcaaat ctagagttga accagataag aaagtctctt cttccggtaa gatattatgg 660
acctataaca tctgtgtact taaaagtaga ttgggagtga aaggcagact tttgatgttc 720
tgtacactgt tgaaaccctc tagcgtggtc ctctgtaacc tgctcaccct gccccaggga 780
ggcagctagc caatgccacc agcccaacgg aaaccccagt gcttttcaa tggggaaatg 840
cagtcacttt tctttggatg ctacacatcc tttctggaat atgtctcaca cacatctctc 900
tttatcacc cctttttcaa gtaaaccaac ttcttgaga agctgacaat gtgtctcttt 960
actctccacg aagattctgg cccttctctt cacctgtcag aagtttagga ttccaaaggg 1020
atcattagca tccatcccaa cagcctgcac tgcctcctga gaactgcggg tcttgatca 1080
tcaggcaact ttcaactaca cagaccaagg gagagagggg acccctccga ggtcccatag 1140
ggttctctga catagtgtg accttttctt ggaactttta caaccccag gacatttcca 1200
aactttgagc agggctctgg gggccaggcg tgcgggaggg aggacaagaa ctcgggagtg 1260
gccgaggata aagcgggggc tccctccacc ccacggtgcc cagtttctcc ccgctgcacg 1320
tggtccaggg tggctgcac acctctaaag ccggtcccgc caaccgccag ccccgggact 1380
gaacttgccc ctccggccgc ccgctccccg caggggacag gggcggggag ggagagatcc 1440
agagggggggc cgggggaggt ggggcccggc gggaggaggc gagggaaacg gggagctcca 1500
gggagacggc ttccgagggg gagtgagagg ggagggcagc ccgggctcgg cacgctccct 1560
ccctcggccg ctttctctca cataagcgca ggcagagggc gcgtcagtca tgccctgcc 1620
ctgcgcccgc cgcgcgccgc gccgcgcctc agcccggcgc gctctggagg atcctgcgcc 1680
gcggcgctcc cgggccccgc cgcgcgcagc cgcgcccgcc cctcctccc gccccggca 1740
ccgcgcagc cgcgcccgcc gcagcgcggc cggcccggct cctctcactt cggggaaggg 1800
gagggaggag ggggacgagg gctctggcgg gtttgagggg gctgaacatc gcgggtgtt 1860
ctggtgtccc ccgccccgcc tctccaaaa gctacaccga cgcggaccgc ggcgcgctcc 1920
tccctcgccc tcgcttcacc tcgcgggctc cgaatgcggg gagctcggat gtccggtttc 1980
ctgtgaggct tttacctgac acccgccgcc tttcccggc actggctggg agggcgccct 2040
gcaaagtgg gaacgcggag ccccgacc cgtcccgc cctccggctc gccaggggg 2100
gggtcgccgg gaggagccc ggggagaggg accaggagg gcccgcggcc tcgcaggggg 2160
gcccgcgccc ccaccctgc ccccgccagc ggaccggtcc cccacccccg gtccttccac 2220
catgcacttg ctgggcttct tctctgtggc gtgttctctg ctgcgcctg cgtgtctccc 2280
gggtcctcgc gaggcggccc ccgcccgcgc cgccttcgag tccggactcg acctctcgga 2340
cgcgagagccc gacgcggggc aggccacggt aggtctgct tagggtttgc ggagaacccg 2400
agagtttgcg ttagggtctg cggcgacc gagaacctgc gcgggggaaa gtgtgtgtgc 2460
tttaagcttg tgtacgtggg atccaaagt actgagctca gtgcacgctg ctttgagaaa 2520
aaatcttctt ctttttaaat agaaagtgt tactagagag gcaagcaagt tacacgagtg 2580
aagggcccgg agaggtgccc agtgagagat cccgaaatct atttcagact ggtttcctct 2640
ggggcaacca aggggtcttg aaccctgcc agtcagcggg gctctggaga gtatgagttc 2700
attttggtcg ggaaatgctc gtttcttcc ccagctgatt catgggactc caaacagatt 2760

ctgggacact ggtgatcagt caaccagcc ttaocttctt ggagtgttca tagtctgcag 2820
 agcaccaggc gctgtgagcg actttagaaa aaaagtggtca gggacttttag taaccaggct 2880
 ccagagcttt cagagttcac ttgaagttgg tcagacttga gatgtaacgg gagattagag 2940
 tcagtttagat cacaatcaag caatgcaagg cttgtttcaa ctttataagc tcttcattcc 3000
 taaaatctgg tcatgagata ctgcaaaacta agtttttttt tttcttgttt gggttttttg 3060
 tttgttttgt tttcttttaa ataagaggtg tttaatcctt tgcctgaaaa tgttgccaaa 3120
 atacttttga ggcacctga ttttgaaaaa ggatttgtgt gtgttcctac ttcttactgg 3180
 ctcttcctaaa agagtatatc tttatctaaa aagttcatac ctgccataac catatagagt 3240
 atgcttaaga gagttcctaa agaagttata ttcgacattt cagttcaaac acttgacagta 3300
 ccccttgctg gaactatact gggtagtgg atttcaactg gtggcaagtg gagaggtagc 3360
 tgtggtgtgt tacagtcaac tttaatgtga ctgttgatta acacatacac aatgtgtaga 3420
 aaatagcctt atattttgaa atattttatt tatatttata ttttgaaata cacttaaaga 3480
 ttgagaagac taatttttgg aaatcaaact acctttgcat ttaaattttg ggaaaacatt 3540
 aaaatgttgg aatgtgtaat aatttaatat aggggtttaa ggaatgcctc ttgagtaaaa 3600
 acaatataca tgaaatagaa cagactgcat tctgtgaatc acaaaaataa ttttcagtgc 3660
 ctatacttac attgccgtaa ctatagtgat gaaaatatct ttctgttctt aaataccgca 3720
 gaaatgtaat aatcggctca ataacgcttc tgataatttg agtccttgat tttgcagtac 3780
 ctatttgcta tttctgcaaa gtcaaaacta gtaaaagtat gatttagtaa ggcaacatct 3840
 ctgtattggt tacggctcaact tactttgggt tacagtaaga aatgtgcag tagagagtaa 3900
 aatgagaaac attcctgat aagccaatag cctgtatgag agtcactcac tgctgggtga 3960
 gtctagaatt aatgtacctt gagtgtgtgt gtgtgtgtgt ttacgtgtgt tatgtatagg 4020
 tatgggattg gaggattaaa aaagttaact tatttttaac tttcaaaagt gtcttctgtg 4080
 gtgatgtaaa atttgccatt ggtgatatgg caaattttac cattgctgtg aaaatcacta 4140
 tgcatttttg tttgcaaatg tcgcttttca tatttaacca tttttatatt ctgtatgggtg 4200
 aatatgggtt agtattctgc atgtgatgtg ggttcctttc ctatgtgatg aaacatcatt 4260
 tcccatatga aacatcatct ctgagttcca tagaccacaa agcagtattt cattgctcca 4320
 ttaaagtgtc gttccccagt ggggtgaatag agccttgaca gggcatgtgt cattaatgca 4380
 ttcagtgaaa ctgctgtgtc cttgctatga aacatgtcac acaactgcca agaaactgat 4440
 atgaagactc tatgtctctc cctagtggcc tgtacaagca ccacatttag atttaagcac 4500
 tagagaagaa ctggggagaa atcaattggt aattttctat aactacatga accccaagaa 4560
 aatatttcag aaggtaaaca gaagatgcta aatatctata atccaatag atgcaatggt 4620
 ttcataactt tcacctagga acttttttaa tatttaaaaa atcaaaaata gcacacaaac 4680
 tatgaaccaa ggtagccat ccagcccttc agcccaccg taaagccatg caaaagattg 4740
 taggtaaaaa tcacaaatag aatgaagatc tccttattcc tcgacttata tttgcattta 4800
 ctggattcct atttctagtt atttgtgggt gcatgactta ctgaaaagtc ttaaaatggt 4860
 tctctgaagg atgttaacac aaacaagaca gtgtcaaaaa agtaaggcag gcaaaagtaa 4920
 aattaactgg acatgatggg aacaaaagag aatagcaagt ctgaagctgt tctatttaca 4980
 atttttctta tagtatatga ttatgtttgt caacatctac tcttataaga gatttctgac 5040

attcctctga gtgatgctac ttatatgagg ttttagatat agaaagttca attcatcttt 5100
cattttgatt tccactccat tatagtcttc tgaactatga atatctgatt gatagctatt 5160
tattcacaag acagttatca aaattgcttc atagaaagca gttaattact gatcttagtg 5220
gcatatgtct aagaccgta gatctcctaat cctgtccttt tcagttgagt tgagaaagac 5280
ggagggtagt ggatttagat ttcattttgg acttccagtt ctggttccgg ttcctcttga 5340
gaggaaggtc ctggcctcct agttatctcc atgtgtccaa cccagaatgc tgtttctcag 5400
ctcttgctgg ggttttcacc aacagaggaa ctggcgtgag actaatgaag gtttcccgaa 5460
tcctgggac tgtgcttgcc tgcttgccct tggcagtcct ggggttgta ttgagagtag 5520
gagtttataa tcattatctg tgatgttcat aacgtgtggt aagcacttgc caaagagctt 5580
tgcaagttag ttttgttttt cactacacct ccataatatg gacattattt tatttcat 5640
aggatgatga acccaagatc tgaatgactt aaccaggaaa aagaatagat gtttttttcc 5700
tttttttctt ttaaaaaaag ttttatggca tataccaaca gtccccgact tagaatgggt 5760
ctattttagga cttttggact ttacaatggt gcaaagtcca tagacattca gtagaaacca 5820
tacttcaagg tttgaatttt gatcttttcc tgggctaata atatgcagta aaatactttc 5880
tctggatgct ggacaatgga agtgagctgt gaccagtcag ccatgtgatc gtgagagtaa 5940
acaacaggta ctctactgta tattaaatat attttcaact caggaggaat ttgtcaggat 6000
gtagcccat cgtaaateca gaagcatgta tgttttagag aaatgtggta atgggtttgg 6060
gcctttgggc atgtgattca gactcctttc agcccctctg ctggtctcct ttggcacact 6120
atagcagcgt gcaggaaggc agagaaaagc actgcgtgtg actttgcagt taggggcacc 6180
ttgttgcat ttttaggtc ctaattcccc acaacgggac aaacggaaat aaaccagac 6240
tttcttcag gctttaccac atgggtgttc tttgactaaa gataaattat aaatgacct 6300
atccatggta atattataag aacgactgct taactggagg cttacaaga tttctgaca 6360
aacggttttt caaagttgtg ctcaagtaac ttcagtgaag gattgtgatt ttacctctgc 6420
caaatgacct ttcacaaata agtggatttg gaaagaaca atttaagttg gcagaagcgt 6480
tatacttctc ctgacttctg agcatatatt gtttcacctt tcaccagtta cattattcct 6540
tgtagatgtg cattttataa atactatatg ggtctacctt ttgttacaaa cacatatttg 6600
tatgtacaga tttgtataaa ttgtggatta agttatcatt gatttttctc attagaaaaa 6660
acagatatag cagagaactt taaacaagga gagatgctga tgagtttgag ttatagtgt 6720
attataaaat ttactataaa tatttttaggt ttgtgtaatt gtttttctga aacttactag 6780
tgtaaatttt tgggggtggct tctttcctcg gttgattaat ttttttttcc agagtttatg 6840
tttaaccatg agaaagcaga gactgttttt atttaatttg ctttattggg catccatcat 6900
ggaattaatc aaattgtaca ttaaagctt tgctcagtga ttagtactgt ttccttatac 6960
tgctcgcgcc ctagtggctt tcagggtgta ggagttatcc acgatagttt tttttccaat 7020
tttgaaatcc aaaaactcat tttctttttc tttctttcct tttttttttt ttttttgaga 7080
cagggtcatg ctctattgcc caggctggag tgcagtggca tgatgttggc tctactgcaac 7140
ctctgcctcg tgggttcaag cgattctcgt gcctcagcct cccaagtagc tggaattaca 7200
ggcatgtgcc actagagcag gctaattttt gtatttttca tagagagagg gttttgtcct 7260
gtaggcgaag ctggtttcaa actcctgact tcagggtgatc cacttgatct cctaaagtgc 7320

tgggattaca ggcattgagcc actgcgcccc gccccaaaag ctaattttct taaattattt 7380
gggtgctataa tcctgatctg aactgacata aggctattta tagtcattat tatctaattt 7440
aaatattttct aatcttcact gggaaaatag taatgtatth atttaccatga tgctgcatc 7500
agtcctgctt gtttttgggg tgtgagtatt aatatgtcta tgcacttggg ttccagaaaa 7560
attccagggt cagagactct tgcagtctcc aggcattttg gataagcaga tggattgtag 7620
ccctgttcca caggccttcc tccattgctt tattgccttg gggctgctg gaagaggcag 7680
ggaaaggtaa tccaggaatt caagaaagaa gccatgtgga tgctcctaatt tcctttttac 7740
ccattagtac caggtttatt tgctctttaa tctacttcaa ctcatgttaa tcagttctta 7800
ctagtggcag cttcagttct aagcaggaaa ttacagttca gtgtgtgatg caaaaaatgt 7860
agagtgagtg atcttgtcta taacaagctt ataatcattt tggagaaaat aaatttactc 7920
attaaaaata ctacaagtga attaaactat aaatccagac catgtaagtg gttttgtaag 7980
agagtcttgt gagttaaccac cagttgttgg ctagccttct gggagtgtga tgggcaggcc 8040
tgggaaagggt ggaatgatata gagacccaat gttaaaaatg agtagatgct aattacagaa 8100
atgggggaga gagggatggt ttgaatattg aagaatgaca gcaaagtaca gaggaagtaa 8160
aggtcaggga tatttgggga actagaagca aacccttgg tgcaaaagtc aagagcttta 8220
agggaaactg aagggagcat ggtgggaagt gtgaaaacca aagagaaagt ctgagtttga 8280
gttgaaagca gcaggagcca atgaagattc ttttagtggt tgctgaaagt gaggttttta 8340
aggagcttaa cctggcattg ttgcacagaa gtgtctaaga aagctgacca gaagcttctc 8400
agaagcttct cagaagccat tttcaactgc agggtttctt aaagtgtctg gcagtgactc 8460
tgcttttctc attacagtgg ccaccatata atgggtattcg aactcctcta tatgctcctg 8520
attgctggcg atgatagtgg tgattacttt tgacttactg taggcacca tcattgtgct 8580
atctggatac atttacaata gtatagaaat agacatctct gccattatgg tttgcatc 8640
aaacgaagca agcagtgtaa acagtgttca tgcttagata cacatacaca gaacccccca 8700
cacacatatt catgcataca ccgcttaata gttcttcaga ctgccaaagt tgacttacgt 8760
gttctttctt gcatggtaaa gaattttag acagttcatt gaaatagaat atttagacaa 8820
tcagaagtgt acaagattag aatctctata tttgcacaca catttatttg taggtcagt 8880
ggcaattgaa aaaaaagaga aggagagaac cctaaagtgt cctggaaatt cctgcttttt 8940
aaaatctgct aaaaatgaga ccagaaagggt gggagtgggg atgtgaggag gtgggtaaac 9000
taccaaaata acaatataaa tactttggct gttcttataa aaggtttttt taaatatggt 9060
ggaataattt actcaaactc aagagatgcc acctactaga gaaaggacat actgaaagag 9120
gaaatattca aatgcacaac tttgtgtaaa aggtaatact tacaagtta aaagactcac 9180
ttctaaagaa gtttggcttc aacctgtcct catttggaca ggtggaaaga tgtatatttt 9240
gggtagacc agaaagatgt ggtttgactt ctatagttga aaggtttcta tttagacatt 9300
tattttgtaa ttttaatttac ctaaaacttc atccctaat taccattttc tcttactttt 9360
atgcaatagt aaagtatgca gtcataagt ataaaggctg ggatcaaagc tagctatth 9420
gctaactgca gaacctatgc actctgttta cactgcctgt gatagcgcct ggaagaactc 9480
actgcaata tttctgcttg tttcagtcac ataactgttg cttgtgtcag tgacgtaact 9540
tgtttagaga tgcagtgcag tggataagag cacagggtat ggacatgcct ttctgtattc 9600

atgtcctgag gatgcctctt agtacctgtg tgaccttggga caagttactt agtttctctc 9660
tgcttctttg tagcacaatg agtagatgct aatttttaat ttttaatttt ttctttttct 9720
ttgtgtctgt tttgtgtatt tagttacttt atcttttttt tttttttttt ttttttgaga 9780
cagactctgg ctctgtagcc caggctggat ggaatgcagt ggtgctatct cagctcactg 9840
caagctccgc ctcccagggt cacgccatcc tcctgcctca gcctcctgag tagctggaac 9900
tacaggtgcc caccaccaca cccagctaat ttttttttta ttttttttta gtagagacgg 9960
ggtttgacca tgtagccag aatggctca atctcctgac ctctgatct gccgcctcg 10020
gcctcccaaa gtgctgggat tacaggtgtg aaccacnnnn nnnnnnnnnn nnnnnnnnnn 10080
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10140
nnnnnnnnnn nnnnnncaat cacacacagg tcctattgc tccttttctt ctagggtat 10200
tcaacaagca gagacaaaca aaaactgac tgtcacattt tactgcaggt aacagtgacc 10260
acaatacac ctgtacactg ttttttttcc aacattattg aggtgtgact gacaaataaa 10320
aatggtatat atttaaggta tgtgatgca tgttttgctg tacatgtaca ttgtgaaatg 10380
attatcataa tcatgttaac atatccatca cctcacatat tgttattttt ataactactt 10440
ttataaattt tcccttcagt agtaacctt tgagattctg aaacactttt gagttttgaa 10500
tgttgtgtca taagtgagtt tttttggcaa atattttata aacatctaata tgggaaactg 10560
tgagaggatg ttttgaact ttcatttaa ctaagatctc ctctccacct ccacccttcc 10620
catattcagt agattttcta tggctgtatg gtgtatcagg aatcatccca tatacttcag 10680
aaagtaatgg aataatataa atatttttct aacatagacc tacagagatt aagtgatata 10740
ctcaaatcat accataaatg gtagtgaaag ggtaaactca acattcacag ccaacttcat 10800
tatctaatag ggaacaacac atttcatatt aataattgac ttttattttt aggaggctct 10860
gttataatga attaaaaaat gagattagat tatgtttctc aatatatggg acaaaaaatca 10920
cttacaataga ataacttoga ttattaaaaa tagaagtgaa aaatactttt cagattaaat 10980
aagatttcaa ttttaaggca aatattgttca ttttgtgctc aaaaattaga caggaaaact 11040
ataatctggt aaaatgtagg ttcttgatcc acccaagaaa ccatggagag gtgggggaca 11100
gttacctgca tttctgacag ttagttcatt gactcttatt cgactaaact ttgagaacca 11160
gtgtccaaag tatttactta catatttata tttaaaagct gggataatca tgacaatata 11220
aaataaacia ttgatgttgt attttacttg gcagatagat aaagctatac aagtttcctt 11280
tttcttaatt ctaaagcata attagaaaaa aattgtattt atttcttaga tatggtaact 11340
actaaaatat aatttggat gttctactta aatgagtata cttatgtatg cagttagccg 11400
tcaataagat ttagtgttct ctgaactatt tcatttctat gaaggtggta ttaatcttga 11460
tgataaaatt ttctttttaa tattgagaga ttctgcaaaa atataaatct tactgtgaca 11520
atgttgttga aattcctttc tcatatgat ttttgttct aaaatagctc ccatgaaagt 11580
aaaattccta ctctcaagca tcctgcttaa gcatattgta ttgttgctgg taactaaaaa 11640
taaagcaaag cattaggggtg gatttcagt taattttatc tgtttagtag agtgatacaa 11700
atgaataaat tttcaaacat acaaatgac caattatttg aaagagctat gaataaaata 11760
aataagctac cttaaaatta ttttaaaatt atatcaciaa atttttttcc taacaacaaa 11820
atatgtattt ttaattgttt tggttcatac aaattaaaat acaaatataa ttttaccag 11880

gaaagtgaag taatthagga gtaattaata atatttgtgt tagggaacgg agcatagata 11940
ctattaaacc taaccgtttt atagtgcatt tgtttctaaa tcccttaagc tggtaagtca 12000
ttgatcttgc atttttcttt gacaggctta tgcaagcaaa gatctggagg agcagttacg 12060
gtctgtgtcc agtgtagatg aactcatgac tgtactctac ccagaatatt ggaaaatgta 12120
caagtgtcag ctaaggaaag gaggtctggca acataacaga gaacaggcca acctcaactc 12180
aaggacagaa gagactataa aatttgctgc agcacattat aatacagaga tcttgaaaag 12240
taagtatggg aaataaaatg tatagtaagc cttatatatc aaaccagtag aagttgccag 12300
cctggttctt taatctcagt gtttctatca ccgaatacac ctttagggcg gaaagcaaat 12360
agaaaataaa gaattacttc aaaatacaac attttatgat ttcactttga gtattaagat 12420
gacttcataa aatacaggat agataggctc atgcataaac aacacacatt tattttttca 12480
ggtgatagcc attggaaaat ttttaaaaca gatggctata ttttacatac tataatctgtt 12540
taataactta aacattcttt atcttatcaa ataattatac ctttaattcc tatctgattt 12600
ggattgagaa atgcttcatt tgaaatcacc ttttcagatg tcacttggtg ctttgctgct 12660
tcacagataa ctggctcaga gcgggactta agctgatgag gtatctagca taaagcttca 12720
caaaaaatga tcatgttgca taaaatgtca tagcatcaat tgggtgggta tgactgggca 12780
cagaggttca tgtctgtaat ccgagaccat tgggaggcca aagtgggagg attgcttgaa 12840
gccagagtca gagaccagcc tgaacaacac agcaagacct catctctatt aaaaaaatt 12900
aaaagtggg ggagtaattt aaaaacatat ttgccttcag ttggtagctc ttaatcgcta 12960
taaaagtttg atattgaaa gtattttgag aacagaggca acaaaaatgg cttcatgggt 13020
gtataaccaa tagtttatgt cttttccaa aaattctatt ttaggggaaa tatgggataa 13080
gtaaactagt atacttagga agtacattat ttgtaatgca gaatttaaaa agtcaacagt 13140
attataggta agctactaac tcatcagaaa caaaattcca gaggagagt aaaagggata 13200
ggaaatattg actttgctga agaaagttca gaatgagtcc cagttggtga aaaatagcct 13260
tgtattttac gtagctgagg ccacgacatt actcaagggt gccaaatagt ccgcctgaaa 13320
aggacgtttg acttgaatct aaaaccatag ctgtcttctc atatgcaact cagttttaaa 13380
gaaaaataaa attatattta tatttgcatt gagagcacia aatggccca tactatagag 13440
aatgaacatg actctctctg cacatcttat aacattcat atattaaatt accattttcc 13500
taagaaagaa aaaccaaaca ttagtgttgt gggatattgg tgtactttta ataaagcata 13560
atgctaattt ttcttattgc ctagctatag aatgagttt aagaaatc cttagcttgt 13620
cttttcaatg ttatatcatt ctcatcaca agtcaattta ggatcccatc agctgtctct 13680
ttagttatat attgaaacta aatgttaaaa tctgaaatga atgtgagtga tttgcagctg 13740
ctgtgggctt tttttatctt ggtggtttcc gtaaaactta ctgatttccc tatgacttga 13800
caggtattga taatgagtgg agaaagactc aatgcatgcc acgggagggtg tgtatagatg 13860
tggggaagga gtttgagtc gcgacaaaca cttctttta acctccatgt gtgtccgtct 13920
acagatgtgg gggttgctgc aatagtgagg ggctgcagtg catgaacacc agcacgagct 13980
acctcagcaa gacgggtggg attgcccatt cctgctacct gcgagagccc ccttaatcag 14040
aagctttaac tttggggatg atgtgtttaa acttctgggt gctttaaact aacttggttc 14100
aactcaggg ctcttatatt ttcttaggg gaatgaagtc agtttttcag tgttttgtgt 14160

gtttgacaac acagtgtcac aaagaaagta attgatgctt tattactttg caaactcttc 14220
atatagtggg gtcaaagcct catgcagaaa acaataaata tacatatgta tgtgtgtaca 14280
tgtacgtgtg tatgtgtgtg tgttgatctt caagaataaa catgttcatt tggcaataaa 14340
atacttttaa atgggtgagt tttcagatgg tggagataac agtgcaaca gtattcaaag 14400
tgtaaaatta aatgcgtgtg ttctgtgaaa cagtgggata gtattgaaat aagccatctc 14460
actgaagaaa tttattgtct tcctgtcgaa agcaaaggaa ataactctga cactaaatct 14520
cagatgtttt gaataaactg ggaagtttat ctcatataat ggtaagcttc tgataaccag 14580
gaaatgtctc ttaaggttgg ctttgatca gggattttca gaccagacac ataaacccgc 14640
tttgtgggccc ttacgtggta gaggaggctg gctgtggaac acgcgatgt gctggcgata 14700
accatgaaag tgttgtttga aattgtgtgc acctctggcc tctgaacaga aatgtttgga 14760
gaaatccagt cccgtgggct ggactgaagc aaacttctcc cgaaagaca aactgtgggt 14820
gtcctaataa ctttctaaag ctatttttgg agtttagactc acacacgtaa atctggctaa 14880
attcagaact gttctttgac ctggagaatg aaaaatgaaa ttgaaagatg tgggtgttaga 14940
tttgatttgt atgatgtgg atgacagaaa attcacatct taattttatg caaactttag 15000
ctgatctaca tggttttcag taataaaacc tactcttatt tcaatatagt attaataaga 15060
atcagagtta agattactat ctcaattcca tgaactacat tttttcagct agctgaatat 15120
ttccaaaatt tttactgtta cgatcatttg gcctctgagt ttggagacat tgtcttatta 15180
tacaacccaa attaattgct ttcttttcaa aaattgagtt tctttatatt gctcaccac 15240
ttgataaaga aacaagccat tttagattaa gacttgcaag aagaaacaaa aaataaatta 15300
taaagctaaa ttttctattc agcttgctc tgcctcaact ctaaagggta aaaaaaatt 15360
aaaagggtgat gctttggggc ctgctatcat ttcttttagt atcatagtta gtttaagctta 15420
gccagatgca attcattctg attatttctt aaaagcacag caaaataatg taggggtcgt 15480
tccaacatgt gttctcaagg aatacaatgt gactctaac atagtcctgc atttagaac 15540
tttaccagaa aatgtacag ttgttagaac atgacaatat ttcagtgggt gccaaaatta 15600
taatattttt gatagactaa catataatta tactattgga catagcaaaa gttatctta 15660
gttcatttat cctccctgag tctttatcaa acaacaataa acaataagac gttatagttt 15720
tgtacttact attctaaata aattaatact acttgaacag gttttattgg tattatgtat 15780
tttatcataa ccctgaaata aacaggattc ttttaagatcc tctaaattct agaactgcat 15840
actgtaaaat aaattttgga atgccgttaa tgatttatat tttaaaataa ttcagttaag 15900
caatatgatg tattctagat tatattccta atatgggtatt tgatttaaat caaatagtaa 15960
accaaggtac tttggatact gtttccaact aataagataa ccttctctct gctcttctc 16020
cttccatttc ctttcttcat aatcttttct ataaattact ttcaagcata ttttttaaaa 16080
acatagatcc aatctatgg ctaggcatgg tggctcacgc ctgtggctc agctgctgag 16140
gaggcagagg caggaggatc gctggagccc aggagttcag ggctgcagtg tgctaggatt 16200
gcaccactgc actcccagcc taggtgacag agcatgacct tgtttcaaaa acaaaaaaac 16260
cccacagagc tgatgtcagt ctccagcata aaggcctttc agtgatttcc cattgcttgt 16320
aaaataagac aaaaatcttt tttgaaggtt atatttattg acatatcata gctgtacata 16380
ttttggggtc atgtgttata tgttgatacc tgtacaaaat gtgtaatgat caaatcaggg 16440

taattgggat atccatcatc tcaaacatth atcttttctt tctattggga acaatacagt 16500
acttctcttg cagcttttta aaaaatatac aatataatcgt tattcactcc aatttcctg 16560
atgtactatt gaacactaaa acttactcct tctatttaac tgtgcttttg taccctcaa 16620
ccatcctctc tccatcctcc ctacccttgc cctcccagc ctctgataat cagtattctt 16680
ctctctactt ccatgagatc cactcgttta gctcccacat aatgagaac atgcaggatt 16740
ttcaagacca catctttatc acaacctgca atctctgccc catagcccct gttcagcatc 16800
atctaagtga atgattcttg atcttgtgct tccagtgtc tagactttct ccaattcctg 16860
ggcatcatcg tgtttcctct gctcctgggc ccttgcacac acaccttag gttaacctgg 16920
ccttctgtc tctctttcc tgagggtgcta gctcttcaga tcttaatttg ggtgtcaagt 16980
ggctttggca aggcctcagg caagactttc tgacctattt ggattagacc caatatttgt 17040
agtaaagtct ttatagtttc ccatattttt aaaatagtat ttattacagc ctataatcac 17100
atagtttata gttatgtagt cattgttact tgattaatgt ctccctcctc cacaaaggca 17160
aaaacagatc cgttcactac ttatcccagt gcctaccaca gaatcagtgc acaaatctct 17220
tttagtatca cataaaaact gtggaccctt tcatgagaac aatgaaaat atagacacat 17280
ccacagcatt tccccccaca gttttagga gtttgtccc tgaccctag cttttccct 17340
ttaaaccagg ctttaaaact gctgccttga agattattaa taataaaata gctgattaac 17400
acacttggg aattttcaa taattggtct ttttgcata caaacgttta tgctatcaat 17460
ttagcttgat cttttaaaa atatctgttt ctggagagga acagaaaagg taactattgg 17520
gtactgggct taattcctgg gtgatgaaat aatctgtgca acaaccccca tgacatgagt 17580
ttacctatgt cacaaacctg cacatgcacc cccaaaccta aaataaaagt ttacctattt 17640
aacaacctg cacatggacc cccaaaccta aaataaaagt ttacctatgt aacaacctg 17700
cacatggacc cccaaaccta aaataaaacc ccaaacctaa aataaaagtt tatgtcaca 17760
acctgcacat gcaccccaa acctaaaata aaaccccaaa cctaaaataa aagtttacct 17820
atgtcacaaa cctgcacatg ccccccaaa cctaaaataa aagtttacct gtgtcgcaaa 17880
tctgcatatg caccgcgaaa cctaaaataa aagtttacct atgtcacaaa tctgcacatg 17940
gaccccaaa cctaaaataa aagtttacct atgtcacaaa cctgcacatg gaccccaaa 18000
cctaaaataa aaccccaaac ttaaaataaa agtttaccta tgtcacaac ctgcacatgg 18060
accacaaac ctgcacatgg acccccaaac ctaaaataaa acccaaac taaaataaaa 18120
gtttacctat gtcacaaacc tgcacatgca ccccaaac taaaataaaa gtttacgtat 18180
gtaacaaacc tgcacatgga ccccaaac taaatataaa ctttatatat gtatgtatat 18240
atatacatat gtatatacag atatacatat atgtatctgt aaaagaaaca tatctatcta 18300
tatctatc tagatatcca gatgactaga tgagatatag atatctgttt ctcttacaga 18360
taaggaagaa aattcttctt tattctctca ttcattagtt aatagagca aaagccaaca 18420
ttatatcttt ctactattg gtcaaccaa taaatgcata ctttacatat gggccaaata 18480
gatacctaag cataccttag gtaaaatcta aacttggaac acagaaaatc agtacattgt 18540
tagtaagcaa agtaaaataaa tttagtcttc tattaacct ttgagttttt tctgtgacca 18600
ttgcatattc atttctgca ttggtotaat taggtgtcac attcaccag tgggtattgg 18660
agtgaagta tttattcaag ggtagggtgtg tatgtcagc agcataaaaa ttgctctatc 18720

aagagggagg tgttgagag gtttaggctt tagaaattgc aaggcttcgc tgggactgt 18780
ggcttacacc tgtaatccca gcactttggg atgccaagac aggcagatct cttgagctta 18840
ggaattggag accagcctgg gcaacatagt gaaaccctgt ctccccaaaa aagtataaaa 18900
attagccagg catggtgttg tgcacctgtg gtcccagcta cgcgggagggt agctgagggtg 18960
ggaggattgc tggatcctgg aaggcagagg ctgcagttag taaagatcgc agcactgcac 19020
tccagcctgg gtgacagagc cagaccctgc ctcaaaaaaa aaaaaaaca aaaaaaaca 19080
aaagaaaaaa attacaaggg gccaggggta agatgagcat ttagtgctaa gtgtaagctg 19140
catatagggtg aaataatddd gccaggtcta cttagtactt aaacatcaga tatttgctat 19200
aagaaaattg tctacttttag atcacatgga atcttttctt gctgtatttc atctctttca 19260
cctcccagct ttgcagaatt taggcttact tttataatag atttacaatg ttgatatgag 19320
ttttgtcaac ttcttcccaa tcatttgctt gtgtgaaaaa tcaggacttg cccatgaaat 19380
ttctgaagtt tgtctgcttt tttgttattg gttgattttt ttttcagttt tttttctatc 19440
ctattggtgg gctatatatt tacaataatc acaatattta tacattctga tctgtaatgc 19500
agtcatattd taaaattdat atagcattag tttttatddd gtcatttdta ttttdtaag 19560
attdgaagat ttdgagaatt aaaaattdgt agtdcatcaa aatatggta ttdctctga 19620
gaattctgac ctaagacacc ttdctctgag actgtaactc tccttgattt ttdctcatc 19680
cttdtdtdtda ctdtdtdccaa accttdtdct ccaatdcagg ctdtdtdctt gatttdtda 19740
ctgtccactc aaacttdtga ctdggtctgc ctdcagacag gtcagtctgt ctdacttaac 19800
agtdtagaac tattccctac caccttdccc aaataagcaa gtaagtgaat ctdattcttd 19860
tdccagtdgt ctdgtcttda gtaaatgaca tcatcattca gccagtdcgt caagccaagt 19920
tatttdtdct ttdaccaaaa ctaattdgta ttdaatctac ctdgcaaaaca ttdctaggac 19980
ctatccagtd ttdccacgct gctaattccc ctdctgatc ggtcacattt acccttdacc 20040
tagaccattc ctdctdggaca agcctaataa ctdggtctcct acctccattc ttdcttdtdct 20100
ccaatctatt tgdccacttdt tagcagtdgat ttdtdtdtdt ttdcttdaatg ctaagtdcaga 20160
tactgctaaa aactccttdga taacattgdc ttdattdggat ggagccagca ttdcaactagt 20220
gtggtgataa gcttatcttd tctaattggg ttdcttaggc cagcatctgg tdcagattgat 20280
tggtdcaaat gtdtdtaagtd gtdagtdgtd tgaatdgaac ttdtdtdtdca gacaaaatcc 20340
cttdaacata attdcttdgg ccctcaatga ttdggtatgct acctaagtdt ccagcctcaa 20400
atgtdcaatgc ttdtdccccc aaacactact taccactaag acccaaatc ttdtdacttd 20460
tdggttdtat tcaatatctg tdcactccag ttdtdcaatg ctdgtactctg tgdtdtdaac 20520
attdccctac acatgatccc gtdatctgat gtdatctagtd aatgtdtact catccttdag 20580
cctaaacata aatgtdactt ccacgacatt ctdgtdatgtd ctdccacagc atcctatgtd 20640
gctgtdatct agcatcatta ctdtdatctgtd cagattgtdt tatcttdtdt taattdccta 20700
ggttdtdtdtd ttdtdtdtdtd ttdtaaaagg agtdtaactt ttdgtdtdcca ggggggggag 20760
tdcaggggca ggaactaagc taactdggac ctdtdgctcc tgggttdaaag ggattattdtd 20820
gcctcaccct cccaagtdacc tgggattgca gtdgtdgcacc accacacctg gtdaatttdtd 20880
tdtatttdga gaaaaaagg ggttdcaca tgdtdaccag gtdggttdta aactcctga 20940
cctcagggga tcttdcccac ttdgcttdcc aaaggtctgd gattacaggg gtdgaccact 21000

gccccagtt ttattccata ggttctaacc tccataaaaa caaaaattgc gtttgtttgt 21060
tacctgaac aataccacag tagttggccc ataccaaaca cttaaagtga aataaaatga 21120
atTTTTgaat aaaaaaaaaac aaatttgggg tcaaaggtca aattttgaat ttgctgattt 21180
aaaaaagtct gaatggtggt tagggggaaa taccaaaagg tttgatgtac aattctaaaa 21240
ttttaaaca ataactgggt ctttaaatat agatttggga ggaagcatca ttctgataaa 21300
tgtcatttga agcctggtgt gcttattaat tcttagggta ttgatgtatc acaagagaag 21360
agcaaacatc agtaccaagg gaaagccaac attcatcatt agctgtgccc tgcatttccc 21420
tctttgattt tgtattatca tgccttttc ctatccaaag gattcgtta tatttttgtt 21480
tatagtataa ggaatttttt ttttttgaa acggagtctc actcttgtca cccaggctgg 21540
agtgcaatgg tgtgatcttg gctcactgca acctccacct cctgggttca agcgattctc 21600
ctgcctcagc ctcccagta gctgggatta cagggtgtcca ccaccagcc cagctgattt 21660
tgtattttaa tagagatagg tttcaccac gttggccagg ctggtctcga actcctgact 21720
tcaagtgatc cactgcctc ggccacccaa tgtgctggga ttacaggcat gagccaccgc 21780
accggcaga tagttttaat ttaaattcag ttagacaagt gcatgtggta caccattat 21840
gtgacgctc atattatgta ctaaaagcat gctatgcaca aagtgtagg gtacagtgat 21900
ggatagatcc tgcctatgg aactaaciaa ttagcaagcg aatcaacat ttatcaagta 21960
attaaaagta gtttaagtgt tgtgaaggaa aagtacaaag gcttatggaa gtgcataatc 22020
aggatgttta taaaccaagg gaggtcgagg cataaaattt ctgaaagcaa gaatgcagtg 22080
ttcctacatt tttactaata ttccaagttc tgaaaacgtg tacttaagta gcagcttttc 22140
ctgagtttat tcaatttatt gcattatcag tgtagtatga atgggagtgt gtgatgattt 22200
ccttcaaata taaaaggaaa tttcttaatg ctgtgttttc ttaaggtaaa aaaaacaaaa 22260
aaaagacttt tggttttgtt aaaagagttc caggcacatt ttgtagcaa gccaccaatt 22320
ggcatgatgt aaattgtaag gacaatgtac cttaggtaat gttaaactca agttttttat 22380
agatacttca tagaaatgct tttttaaata gggactcaat gaaggggagg ttatcagggc 22440
gtgtgaggaa gccagctggt cattgaaaag taaaagcata tgcactgtgc attccaaagt 22500
ataactaatg ataaaaactg cttcaaagtt ctccacacia gtatctaagt ggctttgaaa 22560
agttaagtag attggatgtc ttttcaacct ctttcataaa catacacttt gagtatattt 22620
tcttatgaga agttgtatgt ttaagagata taaccaaaaa gtagaaaatg aattttgcag 22680
ggggtaaatt atttggtgat gaaagcagaa gcagaagaaa gtcatctgct tttgaggcac 22740
acatttggat taaaagttac ttgaatgtta gtatttacct tatatatttt actaaatcta 22800
ggaaagaata aagaaaaaca gctcaatctc tagagtcttc ctgaaaatgg tggggttaa 22860
tcaagctcaa ggaactagaa gttctttgct gtgttcatta gtgccagact aaagaagcca 22920
gtttagttaa gaagctagct tgctatgagc taattctgat acatagaagc agcattattt 22980
taaattctac attgccatct taaaaacaag cagctacttc ttttccatta ttttataaag 23040
caataatag agcaatcaat aatatgagca tgtatattaa caaagacatg ctcgattata 23100
tttaagttgg atgtttcata aaattttaca acaaaaaata tttaaaccta agtaaaatat 23160
gcaactatat aatgaaatta gaggtattat tttgactgc agtcagagat atttcttcca 23220
ttaattaaac agaactctgc agcagctctgg aagattaatt tattacattc ctttgtattc 23280

agcaagtgga acgtgatttg tatcttttagc aagttaatat ttggcaagca cagtttttgg 23340
aatcagggtt ttatcctgat gtggaaaaag aactctatgc cagtaccagt tatcagaagt 23400
ttctttaaga ataacatttg cottaagaac ataattaatg catctgtgaa aaataccact 23460
ttttccaac ccaaggtaca atgagattca cgtattttat tcagtatttc taaaactcct 23520
caattgaggt ttttatttta attgtggtac tgtgtatact ttttatgcat aacaatatgt 23580
taaaattaat tcagtaccag atttaaatct gttaggctat ctgttattcc attttctttc 23640
ccaaacccaa ttaagttcc acaggccttt tatagctggc catgtggttg caatgctaag 23700
tcaaaatgct aagacotttt ccatttttg ggcatattt ctttaaattc gttgtagttt 23760
taaaatctga ttttatagtt attaatcata ttcttctggg ggtacatata tctgtactta 23820
atggactaaa gtttgaata tttaaattta aattaaatat ggtagtgaga tttgttctga 23880
agagaaactg gttatattta ggatataatg gtgttttaaa ggaaggcatt ttttgtgtca 23940
tatttgtatc cacacaaaca ttacttataa tttagtacta aattataact caagatcctc 24000
tcagtaagga agttctacac ctttttcttg aatagccagc aagataagta ctattaggaa 24060
ttgttgcttt atagccgttg ctatgctgct gtaacaacc tattgtttca aaatgtggct 24120
ttatttgaag tccattcttg ttggtgagaa ctgactatta aatattccta ttaaagaatg 24180
ccagtagatt acaactggga acctctcagt tggaaacact tgccatttcc tcttgtttat 24240
tatagtttta cagaatatac agtcagattt ttttttcagt ttccaacctt gcttaaatgc 24300
ctttttttat atgtcctttt gatcttctcc aaaaccaaca actggtatta tatgcctaaa 24360
gtatatgtca ggcttcataa tgaattgat tattacaact ttttttgtgt cctacattgg 24420
acacaagtta atattactta ttggttcaga acaattcact ttattgtata aaactgtttt 24480
agaaaaatca taattttcaa aatatcaaat tataatagat acattatcag ttaaatacat 24540
catcaatcag tatttactaa ataaacatac ttggagctat tagctagtgc tgaataatac 24600
ataaatttta tgtgctatga acttaaatgt agaacagaaa atactattat atacacaata 24660
aagcacaaaa tagtctgaca atagtatcat gattcatatt ttcttaaaat aatctactta 24720
ctgtttcaaa cattgatatg tttggataga tctaaactct tttgagagga aattctagtt 24780
actatgtagg ttcagatcat tattaatggc aagaacaaat tgcataaggg aattaaactc 24840
cagtgtttac ggtaagctat ttggtatctc agaaattcta ccatgggtatc tctctctgtg 24900
tatatgtatg ttctcattca ccctcacctg gaggtctctt tttttttctt tttctttttt 24960
tttttttttt tttttttcgg agagggagtc cacctctggt ggccaggctg gagtgcaggt 25020
gcagtgggtc aatctgggct cactgcaacc tctgcctccc gggttcaaga aattctcctg 25080
cctcagcctc ccaagtagct gggattacag gcatccgcca cgccaccag ctaatttttag 25140
tatttttagt agaaacggga tttcaccatg ttgggcccggc tggctctgaa ctctgacct 25200
caagtgatcc tcccacctca gcctccaaa gtgctggggg tacaggcgtg agtcaccatg 25260
ccccgcctc attgttaggt ttctgatcag catagaggca aggtcattac tgatgactgt 25320
tatcctgaga ttctagcaag gtgcttggtt tcatgtaggt actgaataaa tagttactga 25380
aaactaaaca aggaaagatt ttaaataaaa attttaaaat ttagaatttc ttaacacaa 25440
atttaactaa atcttggctg atgtgaatga tgataaagaa caagagaaag aagaggagga 25500
agagaaatcc aaacattaca ctgaaatcac cattattatt atccctattt tataggtaat 25560

gaaatggagg cttggggaag ttaggtaatt tttccaagtc tcagaagtag taaattgaag 25620
atccaggact ctaattactg tctgtttctc aagactaagt aaactcttgt ttattcttac 25680
atacttcgtg gagtagggaa gttgccagtg atttgaaaat gttatgtatt tttttggga 25740
ggcgccatg caagtttata atgcacttct aattattatt ttttcagttc aaaaaatgta 25800
tattttctaa gcacatacta tctaggaaaa aagcttacgg aaattcttag taacgtaggg 25860
attaaatata gctatgtaaa aattcacaaa gtggctctgta aacaagaaaa tagaatttac 25920
tatatttatt gaaaaggcag tagaaactgt aatcatgtct aactgataag agtaagtcag 25980
tcaatataca cacatgagac cagcacagaa agaattagaa aaattaaaga ggctcttta 26040
gtatgtgaca agagctggta taggtacaat tatacttaa tacataccca gatggcatat 26100
tcccagtaga tattactaat tttcctgaag ttgatagata aataaaactc aattatcaag 26160
actctttcta cttttaatat tatcaaaggc atctttctct ggactcttaa aatTTTTTca 26220
cactggatc atcaagtagc aggtcaaaaat ttagggatag aggatgtgtt tcaagagaag 26280
aacaaaattt tttacttta atttttcaat ctctgttctt tttgacaaa ttaactgaat 26340
tcagcctgcc aattgctact tagacacaaa aaagagaaat gagacattag ttattaagat 26400
tttgattttg attcctttat gttgtgattt cctttatttc ctttaagtag ttaaactgca 26460
acaaagattt tttctctcc aactgtaaaa aagtcttcca ttttaccact tataaaagtc 26520
ttcaacttag aggaaaagac agttctctaa acagtgattt ggTTTTtagtg ctgTTTTacc 26580
tgtctctta ttcttttagt cagtgtattc aaaaacacat ttcttgagc ctttgaaagc 26640
atagagccac acttagtttt ccaggaagat gctttcttaa tgtgaattag ctattcttca 26700
tctcaccaga acattottca aagctgctgt tctgaattc atgcttgata cccatgggtg 26760
gatcctgatt tgtgaacatt atctacctag gcatggatga ggtctctcca tgatcactga 26820
tactgctcac ggcaatctgt ctaccatttg gaaaacaggc tggggcagaa tctggaagg 26880
cagccatcag tgggagcagg aatgatgctg ccaagggtta ggggtgggaa aacataaact 26940
cttctacaat agctgttagt tatctcaaat agtgccttac cttcactatc aggctctgtc 27000
ctgctgggtt tcaactggaga caatgcaaaa tctgctgtg cttacttatt actttgtttt 27060
tgcattgcta taaaggaata cctaagactg ggtagtttat aaagaaaata ggcttaattg 27120
gcttatgggt ctgcaggctg tacaggaagc atgatgctgg catttgctca ccttctgggtg 27180
atgcctcagg aagcttgcaa tcttggtgga agggggcgct ggtgtgtcat atggccagag 27240
tgtgagcaag agagggagg gagaggtctc agactcctt aaacaaccag ctctagcatg 27300
aactaacagg aaaatctcac ttcttaccga aggatgggtc taagccattc atgagggatc 27360
tggctccatg acccattcat ctcccgcag ctcccacctc taacattggg aatcacattt 27420
caacatgaga tttggaggag acacagatcc aaacatatac agattacaaa cactgtgcgt 27480
gtaagcaaaa ttgtttctgc tggctctgagc cacttgagcc ctttcccact agtcacatgg 27540
agcacctctt cttgtcccc gtctccactg ctgcttctt cctacagccc ctgccctgcc 27600
acagcccttg gaaccacatc gggctgctgt tatctctgt caattcttgt tacatttgac 27660
ttcccgaaga catgtggccc tgtgaccaca cacacctgtt tgaagctctc cctctttct 27720
ctggcaccg ttggcttgac agtcagtctc cctcctgat gctccatcat tgttatgctt 27780
cttcccagac atattcacta gttcctttgc ctcttcccac ctcatagatg ttctgtgcc 27840

ccaggtttct attccttagac attttctctt ttcattttaga gtgctcatca agtaatctca 27900
tcaatgtcca agactccaat tacagcttaa atgttcatga cttttcaatg gctctcctaa 27960
actctagaca cacatatgtg actgcctact agacctaaat tcttaggtac accttcctac 28020
catctacatt attgcagagc ttaaataatc aaattataaa ttatacaaat ataaaaaata 28080
agtaatttta taaatttatt aaaaattgct ttttaaaacc actttgtatg tcttctgtag 28140
tgctagttct taatatagat tccatagatg tcttaaaaga tagaaaaatt gtagcttatg 28200
aagtcactgc agaattcctc tcttaaaata ggtatatata gtgcaaagga ttactgccat 28260
tgcaatattt cacctaatca aagattttat ccaacgtgta taaacgggta tctttttaga 28320
taagaaactg gtcaaaatca taatagtggc atatgttgag gtcacatagt gaagttatct 28380
atgcagatca gacagtattg tggagaatac ctagaatctt gagaatttac acaaatagta 28440
ataacttaga tctgtatagc aaataatata catatcagat tcatttagtt tatgtaatct 28500
ttataatctg agaataatga gtcaagtgtt atccttagtt tactgattat caaagaaaaa 28560
tacagcttgg agaagtcaaa tgaattttcc aaatttatgc tacttataaa tgggagagct 28620
aataaaagcc acctgtctgg gtataaccct gtacctatct cttctaaagc atattgtttt 28680
ctaagaaaag tagttacatt ttaaataaat atataaggaa tacatcttga atgaagacaa 28740
tatgcatttt ataatttaca tatttctgta tttcattatt cagaattcat ttaaaaagtc 28800
tccttgttgg catggacaaa attggcagct aatcctatcc cataaatatt ttcaatggtc 28860
ttcagctgat ggcgtgaaag ttcagcaaag gccagtcac aatgctttga agtttgtcat 28920
cttttttatc actgctcaga ttgtcatgaa gacaccaagt aatacatgga tcaattatag 28980
ttaaatagact gcaagttgat cttatcaagc agttagtggg caagatttca cttttaccat 29040
tttttctggg aatgatttag agtgtaagaa cactaggaat ttaaacttac ttagttttac 29100
ttaatcaatg atgtatttgt ttatttaatt taaaacattg tttttggaga aggaagagaa 29160
cacaagaagt tcttgatcca taaataaggc gttcagaagg agaccagtta aactaaattt 29220
tgactgtaat tcagaattgt gtaaaatgcc ctactttatt ttaagctcat tattttcata 29280
attcctttac caagatagca gaaacatggc tttatctatt aaattataaa tgtagttttt 29340
tttgttttaa acataatcaa tttataagag atctaagaca tataaaacta cattgaggta 29400
gaaatttaat aaaaattaac tttacatcta tgtaacttac gtctatgtat tcagtttatt 29460
taaagtaggt cataaccaag tatttctggg ttgtgtggcc aatagctttg ctctaaatc 29520
acaagatgt ttctgtgtgt gacacttttt aaatccctgg aatattttaa atgttcattt 29580
caaataatct ataactttga aacagttctt ttttcttgaa cacattgttc tgtcacgtct 29640
aataggataa ccaattagtt ggttatccta attctaatat gcttcataac tagtgacata 29700
agataaacat atggagtatt tattactctg ataaaagata ctgggttagga gcaagacagt 29760
tttactttg taatcataaa aacatagcgt cctgcgtaca ctaggcagct acttatatag 29820
agctaatoct actctttttc ggactccaat atttaccata aaagtgcttt acagcatcta 29880
tataattttt tcttcttttt gtcccttctt tgtagttatt tgaaattaca gtgcctctct 29940
ctcaaggccc caaaccagta acaatcagtt ttgccaatca cacttcctgc cgatgcatgt 30000
ctaaactgga tgtttacaga caagttcatt ccattattag acgttcctctg ccagcaacac 30060
taccacagtg agtatgaatt aaatctatct tttctgcatc ctctactaat attaaattat 30120

tttcaagtaa acataagtta aacaaataac ctctgtgaaa tttactgtaa tatactctgc 30180
ttcaaataca aaacaaagac aaaatthttgt atagtataca atgaaaataa actgaaaatt 30240
aaaattcagtg ggggaagcgtg ttttagaggc aaagtaattc acaatataaa agaagtatgt 30300
ctgtgaaata aaacacattc caccagtgtt cataaaatat cagtaaggta gcttaattga 30360
tttaacattc ctgaaatcag gtatgctaac tcataaatca tattaagtt caaacattt 30420
cagccttgat aagatthttag gccaaaacaa aaagtaaata taaataattt tatatctatt 30480
agcattcagtg tgctthttc aatctcaaac ataaaaatat atatgtgtgt gtgtgtatat 30540
gtgtatataat gtatctccat gthttcttatt tcagcaatat ttgaatagtg aaatagtgtc 30600
taattacttg aaaaaattt ctgttaatat cthttcaaat tgacattcaa accattattt 30660
gttattcagg tthctthttc atctatthtt tthggagggg caaatggcag catatactc 30720
atgtgatcat ataggagaac ttatcacgtt ctctgtatg atcatgttct gacctgtga 30780
atagctgata atgaattaat atgcattact gcctaagatc tgagagthtt ctgccccaaa 30840
gtagggatgc aacctthttg tcttaggat atccggtgca tatggatgaa acttacataa 30900
ctctagcatt tgtaacaaat gtataaataa attgatcaaa atacatataat attacaaaaa 30960
agaaaacctt acathtttaa aagactcaac aataaatgta atatgcaaac catggthttc 31020
cccactgaat cacgttgtag tactctgctt agtgtggcag atagggggac ccttgthttg 31080
tccagggaca gthttagaaa acctggaaga tagcctaatt tgagtgcag agatctgtcc 31140
cctthttatgt gagaaagcaa agatcttaca gacathttgt aataaatgta tgctaataca 31200
agcaagagac cactaccatc tgaactcaag aggatgccgt atgtgcaatt tcatccccgc 31260
aggctagaga gthctgttca cthctctcag cctgagcag aagccctcag tgactctcc 31320
attctgagag cthtagccat tatctccata ctcatggctt ccaaatgac atthgaagat 31380
ctaaccatata thttgactta taatathtt tctaaatcta tcttctattg cthtgctgtg 31440
aattatatta gctgtaagta caaagagaaa gacttgthta atctthttatt tctacattat 31500
tcatggcagtg gthttctcat thttcttgag gcctgaaat cacagagthg tcatthctc 31560
tctthcttct tctatcccat ctaaacacaa thgattath ttggaggagg ggagatatgt 31620
gttccaagct aaacttatata thcathttca tctactgcat tatcatccac thcattatgc 31680
atgtctctga tctaagcag tgatathaat actagthttc aagcatcaag tctthctctc 31740
gttgatataat gctactgtht tatgaagta gctthaaaag gcatcgtht atctatctca 31800
ccagaaacag tgactthcag thgtcaag aactgagcac aaththctca ccaacctatt 31860
ctgggtgcag cthctctthc thgtthttgt thcattatt caccagtht gtgthttacc 31920
aagthcgatt cccagaagc caagctcatt tctgctctg tatctgtht ctathctggc 31980
cctggctthg gthctgctt thcaaagccc thctcaththc tcagagthca ggtaaaggtc 32040
thtctctgc atthgtctg thcgaatacc acggccaact thgtctctg aataccata 32100
gcactcactg gathactthc atctcacat ggctthtca ggtctgcaa atgthgcatg 32160
tggaactgga thtatgagtg gththcagta agthtatatt gtctgcaaca thcagctaac 32220
aacactcaaa aatactgag cthctctgt gctctcagcc cththacctg cathgtctga 32280
thtaacctt acaaaatca thtagaagtg cagcagcagc atthgctth ththacagat 32340
aaaggtggaa atcacactcc aaagctgthg gctaaataat gtacataaat gctgththaa 32400

taaaatgact catttcctgg atgtatttta tagcatttgc catgttttca tcattactta 32460
ttcatataac tactatccta tacactgttt gaggtcaaga agagacagga agctacagca 32520
aaaagagcct gagtcctaga gtaagccata cctgattctg aattcagatt tgccactaat 32580
tcattctctg tgatcaggta ggtgttttat gtctctgaaa ctcagtttac ttatatataa 32640
actgtaaaga tgtaaagagg attgaatgaa gtactgaatg cttatgcaat gcttagcaga 32700
gtgtctggta tccagaagg agtcattaaa gggcaaattt attattacca ttaatatcta 32760
agtctgattc atttctcac ccatgtgttt ttcagttttt ttaaaaccac tatctacagt 32820
aagaaaaaca ttttatatc caaaccagga tgtgtggtat acacacacat agtgtgtctc 32880
tgtgtatata tatatgtcga tattagagaa ttttaaatta gctagataac ttcacatgta 32940
attaattatg tataagcata aaactaaaac aaattaccat ttaaaaatat tgccttaat 33000
tatggtgctg ctcaagcttt ctcatatca caacccaag gatcctttta taaacatttt 33060
ttccctaate actttcccca taaaatttta atgtaacaaa tatactctgt ctttatacta 33120
tatgtatctc tgtgctttat acataaaaag gaagtgtaaa gtttactttt ttgttcaatg 33180
tagtgttaaa taacaaataa aatgtttaag ttaacattta aattaaaaat ttttaatctc 33240
ttaacagcta acataaattt ataattgtat ttgccaggaa acttttaagg taatctgttt 33300
acttatataa attgtaatat atcaaattac atatgcaaag taacaattga tatgaatata 33360
taatacaaat ataataattt ataaaaacat tcaaaatcac tatttttcca gcaaaagtaa 33420
aataattttt aaaataactt ctttaaaaat atttggagag cttaactatt aactttgctc 33480
agtatgagaa aacttttaac ttatatagat gataaatatt aatttagata ttgtcaactt 33540
tttcttttca tcatttgaca tgtaatggta tgactaacat ttgtagaatt aataataaaa 33600
tataaattat tattatttaa ttctcaaagc tcaagtcaca catgaaaaca ccaaggacca 33660
agcataagca atatccacag ggcaaccaat ggcagccaat aggccacat aggtaccact 33720
cagtatatgt ttgaaatcac tcagtcaatt aagagaaaat cctccagtca tagccatcta 33780
gtcaccataa tttcatagct catatattgc aaacacttgg ggtgtgttcc gttgccaatg 33840
ctgttcatgt aatggccaat aaagaccaa tagaattta gtattaagga atattatttt 33900
cttgggtttt ggagagccaa aatcattga gatattaaga ttttagcccc atccaggcct 33960
cataccaat ttttaccccc ttagagatga ttttactccc tttatgaata tttgccttat 34020
tattacgtgt gatataccca aataattttt ctattctggt tttctttgtg aaaatattgg 34080
tctcaaccct ttatgaatta tggcccacaa ttcaataaaa gctacactag tgataatata 34140
tatatgatgg acatgtgact atgttataaa attctaacat ttgttaagat ttaagaggaa 34200
caatcaccac cacggtggct aaaattaaaa acactgacaa aaacaagtgt tagaatgtgg 34260
tgcaaccaa actcctatgt caccaatggg cattaaaatt gttcaatcat ttttcaaac 34320
agttgacagt ttcttatgaa tttaacatgt acatatttcc acacaataac ttgtgtacaa 34380
atgttcttta ttagtcaata actggaacaa actcaaatgc ccatgaaaga tgagtggatt 34440
aacacatctt agtatatgca tacagtggaa cattgctgag caatgtaaaa agaatgaact 34500
actgatacac acaagaacat aaaaggatct cagagacatt aagttgagtg aaataagctt 34560
ggctaaaaag agtacattct ctatgatattg atttacatga ggcccttaa agacgcaaat 34620
gtactctaga gtgagagaac gcaattcagt gtttgtttga ggatttggat caggattgac 34680

tgacaatggg caagagagac attgtgaggt gttagaaaca ttctttatntt tgattgtgac 34740
agtagttaat gtggatacat aaacttacga aaagatgata aattcaaaat aggagcatgt 34800
tattgtatgc aaattattcc ccaataaaaat tgatttttga aaagcctaga aaagaattcc 34860
tttattaagt acaatttttt aagtagctta ggtttcaaat agtagatttc tctttttttt 34920
tttttttttt tttttttttt tgagacagat tctcactcta tcaccagggc tggagtacaa 34980
tgacgtgatc tcagctcact gcaacctcct cctcctgggt tcaagcaatt ctctgcctca 35040
tctcccagag tagctgtgat tacagttgcc cgccaccatg cctgggtaat ttttgtatntt 35100
ttagtagaga tggggtttca ccattctggc caggctggtc ttgaactcct gaccttgtga 35160
tctaccgcc tcggcctccc aaagtgtgg gattgcaggc atgagccacc acgccagcc 35220
aaaatagtag atttttttaa acctttaca taacatttta agtcaaatct ttgattaggg 35280
aggagatntt cccacagtca tccaaaaact atacgttgag tctaagat aaagaatgct 35340
tctagatcct attagagaaa acaagatgc gcttggctgg ctctgaactt agagcactta 35400
tagtcaagta agaacaacac atattatagt tacattatta ttgttttatt ttactaaata 35460
ttacaaagca ccaagcatct gccacacatt gtctcatgnt catacaaata ctgccgggtg 35520
ctccattcgt aatacacatc acatagtaca aggctgaaga gtgtggatgc tgtagggagc 35580
ttccagttntt gagaatttag tctctgattt ctctttgagt ttgagtaggt tctgtaaact 35640
ttcttgtgac acattttttt taatctaca aagagcacat aagtaagaat gttacatata 35700
acatcataca gaatatttat aagctaact tcaactgaaat caatctgttc aatagcatta 35760
taccatattt gacataccat agccatgnta atctgatatt gtagaatagc atagtataat 35820
aataataact cctaactcaa ggatgttgtg atctttataa ccagcaatcc atgttaaata 35880
ttagcacagt gcctaaaaca tattaagcat tcaataaatg atcgtacta tttttactaa 35940
catcctacag atttggaaat tgagtcttag aaatgttaat gtgtaaaatg ctaaagagcc 36000
aaaaaaaaactg ccaggaaaat ataaaaatta aatcattntt atttctgaaa cccatgtgnt 36060
cccccaata tcttctaaca tttctagnt ttacagaaaa actttcaagt ctcaatatca 36120
gaaagtttca taaaagccag aggaagtagc aattctcttt agcagacaga gntagatacc 36180
aatnttcata ttgggtgntct cacagattat tttttccaat tnttctctct catatnttct 36240
tcaaatntta ctccatgnt ttccatgggc cctnttagnt agactntctt tntcatggag 36300
gagtcattntg tgattaagnt gccatcccct agatnttct ccttggtagt agtnttntgnt 36360
tnttcattca tgagacatt atgtcaggnt ttctatctct ataactgt attaggaat 36420
gtgggntat tntaaatgaa taatataat ctttagntaa aataactcat gtactagagt 36480
acaaacaaac aatattntcta gntagntaag tntctaaaac actatgntgt actntcattca 36540
ctcattcaac aaatagcatt tntgntntctg ctgataacta ataagntgnt tagaccacag 36600
agatagaaag atacatagaa ttgntactcc agaagaaat ccatntctga accagacatg 36660
taaacatag actgaaatca gntgaaatcag tntggntgag aagcattntg gagacatggc 36720
tacagntag aggggataca tgagtctgga gaggntcaaca gaatcacata tntcccagnt 36780
aggatggnt gnttctntaa gntctntgag aactactaaa agntatntatg cagagatntg 36840
gntatctcag tnttntnta tagaaagaa agnttntgnt caagggagnt tntgntntag 36900
aggaagcaaa acgggagntg gaaaaatntg gntagntaacta tntgaaataa tntgagntgaa 36960

aaataataaaa gatggtggcc acagcagtgg atatggggat agaaagaaaa gaggtggaaa 37020
agacgtttga aatacaagaa agcatatgac ttgggaaccg catatggggg atggtaaagt 37080
gggagaagtt gaggatgaaa ctgtatggat gacattcaaa agacaaggag cacctgttca 37140
tagtaggaac ttacctgaaa aggagaactg agtttttaac acatgttagc ttgagtgcct 37200
gtaatatttg taagacatca gagtggtaga gtaattggaa gagatgacct caaccaaag 37260
aaatatagca agtaagaaga caagatgagc cacagatgaa actgtgggga aatcaatatt 37320
taagaaacag ttatctgtag agcagacaca ataactctgga gaaacagtaa tctgtacggg 37380
agactgagaa gaagcaatca gaacaatcgg aataacaaca gagagagaga ccacaaaagc 37440
cagtggagac atgagttcaa ggaaggagga agtggttaaga ggtgccaaat attagaaagg 37500
ttcaagaata actgggactt actctgttag ctgaacagca tgaaatagag atgtttattt 37560
actaataact taccatactc tgggcacaga gagggagaat gtgttttgag aaagggggag 37620
agaaagcaag acagtgaaag aaattcacat ttacacatga gaaactgaat cagaaagggg 37680
aagtaaattt ccccaatgtc acacaactag aaaagtagca aagcagagat tcaaaactaaa 37740
tttgtttgga tataatccat ttcttttgag gtgtggctat gactgcattt ttattgtttt 37800
ttagttgata cataattgta catatttatg gggatatatg gatattttga tacatacata 37860
taatgcgtaa tggctagctc aggatatttt ggatatccat cacctcaaac atttatcatt 37920
tctttgtgtt gtaaacattt taaatcttct agctattttg aaatatataa taaaatatta 37980
ttaactattg ttacctact gtgctgtgga aactagaac ttgtttcttc tatataactg 38040
tactttggaa tgcattacca acctctcttc attccccca caactgacct acacttccca 38100
gactctagta accatcattc tactctctac ttccgtgaga ttcacttttt ggctctcata 38160
aatgagttag aacatgagat atgtgtcttt ctgtgcccag cttatttcac ttaacataat 38220
agcctccagt tcctccatg ttgatgcaaa tgacaggatt gcattctttt tatggctaaa 38280
tagtactccg ttgtgtatat ataacagatt ttctttattt ttaacttcta taaattaa 38340
aatatcaaac atcaaggaca aaacaaaaac aaaacaaaaa agttcccttc taccccaatt 38400
tgaaatgtaa caactatttt tttgtattta ctaaagatac aacccaaatc cccagtcac 38460
attcttcact cttcattccc agaagtaacc accactcaaa aattgtatta ttctcactct 38520
tctgtatttt tatggcatct ggatttatta acataccaga agatattata cagtgcagct 38580
ttgagtgttt tgtattttta gttaaaatgt tggcattact ttttttctct tgacattatg 38640
ttttacagat atagtcatat tgctatctgt acacctgggtg tattcattta ataactatat 38700
agcattccat taaataaaca caagttttta tgtttttttt attctgtgac gatgtactat 38760
agggaggtca atatacacta cagattagca agagagtaca atcaaagagg agagagtgat 38820
gtcaaataga caaggacctt gaagttaata gtgctaaatt tattagttat ttgggttagca 38880
tggacaatgt ccctaaagat gtaataaaat aacacgttct gtgcacagtt aataaaccag 38940
gttgatttca gtcttaagct gtgaaagcaa ttggatatta aatgaaaaca caatactgtg 39000
attctaaagc tttgagcagc agcctagggg atttcaccta ttctgtagag gtattagggg 39060
agcagaggtt gtttctgggt agaggaatga ttgacaagac aatgttcagc atagccaaga 39120
aatcttgaat tatttttaat ggataataaa tatagaagat aggtattatt tcaatttggg 39180
agacagtcgc ctctgtgtga aaaagagaaa agggatgaaa aaagatgagg aaaaaaatg 39240

gataaatcta gcctgaaggc taatttgc atttatttat ttgctgtcctt ttatttcctt 39300
tattaaatg ttgacattgt ttactcagta aaacttattt gcttgcaaga agttacactt 39360
actagggttt atttaagaag attttattct tatgagaaca tgtcttattt ccaaggaata 39420
gagtaagtag atttatagcc aaggaaacat ttgatttggga acttgaggtc tccgtgtttt 39480
tggcttatta atatttaaca gatcctcagt ccaaatgcct agatgcttaa atatccataa 39540
gttgacttt tagccataat gcttcactct ttccttttta atttgaaggt tatacattat 39600
tgctttacca aaagacaatg aagcaaaata gcctaataac ataaatttaa taaactgtat 39660
cagagcagaa atggagagat agcttagctt ccagttaatt ggaccacagt ctcgtagtag 39720
atcatcagaa actactgttt ttcagcattt taaaacatgc atcttttagt tcccctgaaa 39780
ttatcctctt tgccagttca aatctttat gtccatatac ggttattttt gaaatcgtaa 39840
gtcatatgtc tgttcatccc ctttcctata aagctcatta tctaaacatt gtctcacact 39900
tcaaaggatc acctgttgat gtgaatggat ttaccagtgg aactcaggca agagattcca 39960
ggctgtgagc gttactgttt ttgtaaaggg aaggttaactt ttatatatag tttgcctgcc 40020
tttggacagt cttgtttcta ctttctaaaa tacgataatt gaaagcacag cgttcctatg 40080
gcttgctgat agattgtgag aaataacagc aaaatatata cagctgagga gattgcatca 40140
atacatctct gcaccgaaat tatggcagtg atttgcctag aattgggggc tattagatca 40200
aagattgtcc taaagtttaa agttacattc aaccaaacac atggacaatt tttgttaata 40260
tcatattaga ccagatttac cagttgtttt ctaaaattta atatacattt ttttaaccatt 40320
ttaggttaca gaaaaagcag aaagtataga gagttttcat ataacatcct ctttcacgag 40380
tttctctag ttttaacatc ttgcattagc atggtagatt tgtaacagtt gaggagccga 40440
tactgacatg tgatcgtaa ctaaaatcca tagcttacat tagggttccc tcttttttgg 40500
gtgtaagttc tgtggatttt gacaaaggta taaatgatgt gtcaatcatt acagtatcat 40560
gcagaatact tttactgcc taaaagtctc cgtgttccgc ctgttcattt ctctcccct 40620
gtattcctgg catcaaccag ccactttgat gtcatatcaa ctcttatacc ctcataacgt 40680
gttttttatt gtcatcattg ttttgtttt gtttttctgt taatgacgta tggctgtgac 40740
aaggtaatga gaccagcggc tgttgttggga ggacctttt tcttaattcc tggttactta 40800
tccttgaatt atgtgacaag ctgataaatc caataacat ggattttcca ttgtttcaac 40860
ctgccaatac ttctctcaag cttgtatata caagaaaata taaaaaatag tagagcagaa 40920
aattacaagc ttccatactt atcctactct ccctgtgtcc ccatccccac ttctgggtggc 40980
tttccaactc agttaagtcc tgtacttagc cagttggatga ccacacaatg tgacatgttt 41040
tatatccaag tcagtctcag actaggtgag tttttatgt ggagattac tgcttagttt 41100
atactcagac aaaaaggaaa aaattaaaca tataaacacc cttttttttt tttcgagagt 41160
ctcatgtgtg tgcccaggct agagtgcagt ggcgtgatct cagctcactg catcctctgc 41220
ctcccgggtt ccagcgattc tcctgcctca gcctcctgag taactgggat tacaagcacg 41280
tgccaccaca cctggctaat ttgtatattt ttagtagaga tggggtttcg ccatattggc 41340
caggcttgtc tcaaaactct gacctcaggt gatctgccca tctcgacctc ccaaaatgct 41400
gggattacag gcgtaagcca ccgtgccag cctaaacca tgttttaaat taatgcaatg 41460
attgtattag cttttcaaac attaagactg gcaaaattgt tacaactatg gagttttcat 41520

tgattcatcc tactcaccat cttttccttt aatatctgaa caaaccccaa ctctgttcac 41580
tgttctccct gtgagtagac taatggaaca agagcaaca gtaacacaaa cagacgctaa 41640
cacaagttca tgcagaatca ataaaaatca gtacaacaga agataagtct atgtgttttt 41700
gatacatacg attaagcatg tgccttttta acaatttata taaaacctaa aatatatgta 41760
tttctgattt ttacctgtag tactgaagaa gttacttaat aacgttgaat ataaaggcca 41820
cttttaotta accaccttcc attcactatt aactgctggt tccaaagtgt aagcaaatca 41880
gtctctgtgt acatagtcaa gtgtatacaa gcgtcaggcg taacaaactc aagatgaaca 41940
tgacagttca aagatatttg ggacaaaatt gttgaagcat ttttaccag ggctctgtag 42000
ctctgacaat gaaagaatat agttgctctt ccagctgcta ttcagacaga aagcttgggc 42060
aagaagggtc tgtatctatg ttcttcataa tacaattaca agtttgaact tcagataaca 42120
tcagcagttg gcatgtggaa aaccaaacc ctattttgggt atttatcaag attgttaatg 42180
gagtcagggt tcccttattt gtttcttta tggggtacag aacatcctgt tggataaacc 42240
gctgagtgac atgacgatgc tctgaaggaa tgcagtagag cttgtggtac ctgccttgaa 42300
acatggggtt cactaatgct ggtggctcac acttcccatt gaacaagact agagatagga 42360
aggctatttg agggacacag ctatggaacc ataggtgcca catggtaagc caaattttta 42420
tttgttgtgt tgttgtgaaa aaccttatta aaagagcttc caatgagagt acttgattaa 42480
taacacagtt cgtatctata gaaataattt gcttttcaag aaaatcatca tgtgctacag 42540
ttgaattgac attaatgta tcattcattt tgaatgatct gtgaagtatt ttaagagaga 42600
tctggggaag taaatcaata agtagttta cataaatcaa ggtgcagtta tttttttcaa 42660
ttagaaatat attacaaaga ttctgtcatt tccaacaacg tgaataaacc tggaggaaat 42720
tgtgctgagt gaaataagcc agacacagaa agatgcatat tactgatctc ccttatatat 42780
ggaatctaga aaagtagaat tcatagatac agaatagaag agtggtcacc acggctcttg 42840
gggtggggga catggggaaa tgttagagtg tgacaacttg cagttacaag gtgaatatgc 42900
tctggagacc taacgtatag tacagcatag ttactctact caatagtatt ctactgcac 42960
cccccaactc cccacacac acagtaactc tttgaggtgt ggatatgta attagcttga 43020
ttttggtaat catttcacag tgtatacata taccaaaaca tcactttgta taccttatat 43080
atgtacaata tttatcaatc atacttcaat aaagctggaa aatgcatga atattatata 43140
tatgtatatg tatatacaaa tgtataagag attatagcag attatagctc tttgaaaaag 43200
aataacattt cagcccagtt ctgactaaga ccaaacaaaa ggtgatagca tgttttagtt 43260
ccttaaagt ggatttgagg agtcaagaaa tctccaagt taggaaaacc tccgtggcaa 43320
agagtgtaga atatgagaat cacaaaaata gcatgcacaa aatagcctgc tgtgatgttg 43380
aaagtataag agctagtaat tattattatt attattattt ttatttttat tattatactt 43440
taagttttag ggtacatgtg cacattgtgc aggttagtta catatgtata catgtgccat 43500
getgggtgtg tgcaccact aacttgtcat ctaccattag gtatatctcc caatgctatc 43560
cctccccctc cctccacccc acaacagtcc ccagagtgtg atgttcccct tcctgtgtcc 43620
atgtgatctc attgttcaat tcccacctat gagtgagaat atgcggtgtt tggttttttg 43680
ttcttgtgat agtttactga gaatgatgac ttccaatttc atccatgtcc ctacaaagga 43740
catgaactca tcatttttta tggctgcata gtattccatg gtgtatatgt gccacatttt 43800

cttaatccag tctatcattg ttggacattt gggttggttc caagtctttg ctatcgtgaa 43860
taatgccgca ataaacatac gtgtgcatgt gtctttatag cagcatgatt tatagttctt 43920
tgggtatata cccagtaatg ggatggctgg gtcaaatggg atttccagtt ctagatccct 43980
gaggaatcgc cacactgact tccacaatgg ttgaactagt ttacagtccc accaacagtg 44040
taaaagtgtt cctatctctc cacatcctct ccagcacctg ttgtttcctg acttttgaat 44100
gattgccatt ctacgtggtg tgagatggta tctcattgta gttttgattt gcatttctct 44160
gatgaccagt gatggtgagc attttttcat gtgttttttg gctgcataaa tgtcttcttt 44220
tgagaagtgt ctgttcattg ccttcgcccc ctttttgatg gggttgtttg ttttttctt 44280
gtaaatttgt ttgagttcat tgtagattct ggatattagc cttttgtcag atgagtaggt 44340
tgtgaaaatt ttctcccatt ttgtagggtg cctgttcact ctgatggtag tttcttttgc 44400
tgtgcagaag ctcttttagtt taattagatc ccatttatca attttgcctt ttgttgccat 44460
tgcttttggg gtttttagacg tgaagtcctt gcccatgcct atgtcctgaa tggtaatgcc 44520
taggttttct tctaggggtt ttatggtttt aggtctaacg ttaagtctt taatccatct 44580
tgaactgatt tttgtataag gtgtaaggaa gggatccagt ttcagcttcc tacatatggc 44640
tagccagttt tcccagcacc atttattaaa tagggcatcc tttccccatt gcttgttttt 44700
ctcaggtttg tcaaagatca gattgttgta gatatgcggc gttatctctg agggctctgt 44760
tctgttccat tgatctatat ctctgttttg aaccagtacc atgtctgttt ggttactgta 44820
gccttgtggg atagtttgaa gtcaggtagg gtgatgcctc cagctttgtt cttttggctt 44880
agggttgact tggatgatgca ggctcttttt tggttccata tgaactttaa agtagttttt 44940
tccagttctg tgaagaaagt cattggtagc ttgatgggga tggcattgaa tctgtaaatt 45000
accttgggca gtatggccat tttcacgata ttgattcttc ctacccatga gcatggaatg 45060
ttcttccatt tgtttgatc ctcttttatt tccttgagca gtggtttgta gttctccttg 45120
aagaggtcct tcacatccct tgtaagttgg attcctaggt attttattct ctttgaagca 45180
attgtgaatg ggagttcact catgatttgg ctctctgttt gtctgttggt ggtgtataag 45240
agtgttgtg atttttgtac attgattttg tatctggaga ctttgctgaa gttgottatc 45300
agcttaagga gatttttggc tgagacaatg gggttttcta gatatacaat catgttgtct 45360
gcaaacaggg acaatttgac ttctctttt cctaattgaa taocctttat ttcttctcc 45420
tgccaatg ccttggccag aactccaac actatgttga atagtagtgg tgagagaggg 45480
catccctgtc ttgtgccagt tttcaaaggg aatgcttcca gtttttgccc attcagtatg 45540
atattggctg tgggtttgtc atagatagct cttattattt tgaatatatg cccatcaata 45600
cctaatttat tgagagtttt tagcatgaag ggttgttgaa ttttgcataa ggctttttct 45660
gcatctattg agataatcat gtggtttttg tctttggctc tgtttatatg ctggattaca 45720
tttatttatt gatttgcata tattgaacca gtcttgcato ccagggatga agcccgttg 45780
atcatggtgg ataagctttt tgatgtgctg ctggattcgt tttgccagta ttttattgag 45840
gatttttgca tcaatgttca tcaaggatat tgggtctaaaa ttctcttttt ttgttgtgtc 45900
tctgcctggc tttcatatca gaatgatgct ggcctcataa aatgagttag ggaggattcc 45960
ctctttttct attgattgga atagtttcag aaggaatggg accagttcct ccttgtacct 46020
ctggtagaat tgggctgtga atccgtctgg tcttggactc tttttggttg gtaagctatt 46080

gattattgcc acaatttcag atcctgttat tggctatc agagattcaa cttcttcctg 46140
gtttagtctt gggagagtgt atgtgtcgag gaatttattc atttcttcta gattttctag 46200
tttatttgca tagagggtt ttagtattc tctgatggta gtttgattt ctgtgggatc 46260
ggtggtgata tcccctttat cttttttat tgcgtctatt tgattctct ctctttttt 46320
ctttattagt cttgctagca gtctatcaat tttgttgatc cttcaaaaa accagctcct 46380
ggattcatta atttttgaag ggtttttgt gtctctattt cttcagttc tgctctgatt 46440
ttagttattt cttgccttct gctagctttt gaatgtggtt gctcttgctt ttctagttct 46500
tttaattgtg atgttagggg gtcaattttg gatctttcct gctttctctt gtgggcattt 46560
agtgtataa atttccctct acacactgct ttgaatgctt cccagagatt ctggtagtt 46620
gtgtctttgt tctcgttggg ttcaaagaac atctttattt ctgccttcat ttctgttacgt 46680
accagtagt cattcaggag caggttgttc agtttccatg tagttgagca gttttgagtg 46740
agattcttaa tcctgagttc tagtttgatt gcactgtggg ctgagagata gtttgattata 46800
atctctgttc ttttacattt gctgaggaga gctttacttc caagtatgtg gtcaattttg 46860
gaatagggtg ggtgtgggtg tgaaaaaat atatattcca ttgatttggg gtggagagtt 46920
ctgtagatgt ctattaggtc cgcttgggtg agaactgagt tcaattcctg cgtatccttg 46980
ttgactttct gtctcgttga tctgtctaat gttgacagtg ggtgttaaag tctccatta 47040
ttaatgtgtg ggagtctaag tctctttgta ggtcactcag gacttgctt atgaatcctg 47100
gtgctcctgt attgggtgca tatatattta ggatagttag ctcttcttgt tgaattgatc 47160
cctttacgat tatgtaatgg cctttttgt ctcttttgat ctttgttggt ttaaagtctg 47220
ttttatcaga gactaggatt acaaccctg ctttttttg tttccattt gcttggtaga 47280
tcttctoca tccttttatt ttgagcctat gtgtgtctct gcacgtgaga tgggtttcct 47340
gaatacagca cactgatggg tcttgactct ttatccaatt tgccagtctg tgtcttttaa 47400
ttggagcatt tagtccattg acatttaaag ttaatattgt tatgtgtgaa tttgatcctg 47460
tcattatgat gtttagctggg tattttgctc attagttgat gcagttctt cctagtctcg 47520
atggctctta ctttttgca tgattttgca gcggtggta ccagttgttc ctttccatat 47580
ttagcgttc cttcaggagc tcttttaggg caggcctggg ggtgacaaaa tctctcagca 47640
tttgcctgtc tgtaaagtat tttatttctc cttcacttat atgaagctta ttttgctgg 47700
atatgaaatt ctgggttgaa aattcttttc ttaagaatg ttgaatattg gccccactc 47760
tcttctggct ttaggggtt ctgccgagag atcctctgtt agtctgatgg gcttccctt 47820
gagggtaacc agacctttct ctctggctgc cctaacatt tttccttca tttcaacttt 47880
ggtgaatctg acagttatgt gtcttgaggt tgctctctc gaggagtatc tttgtggcat 47940
tctctgtatt tcctgaatct gaacgttggc ctgccttgct agattgggga agttctcctg 48000
gataatatcc tgcagagtgt tttccaactt gggtccattc tccccatcac tttcaggtag 48060
accaatcaga catagatttg gtctttcac atagtcccat atttcttggg ggtttgctc 48120
atctctttt attcttttt ctctaaactt ccttctctc ttcatttcat tcaatttcat 48180
ttccattgct gatactctt cttccagttg attgcatcgg ctctgagggc ttctgcattc 48240
ttcacgtagt tctcagcct tggttttcag ctccatcagc tcctttaagc acttctctgt 48300
attggttatt ctagttatac attcttctaa attttttca aagttttcaa cttctttgcc 48360

tttggtttga atgtcctcct gtagctcaga gtaatttgat cgtctgaagc cttcttctcg 48420
cagctcgtca aagtatttct ccatccagct ttgttccatt gctggttaagg aactgtgttt 48480
ctttggagga ggagagggcac tctgcttttt agactttcca gtttttctgt tctgtttttt 48540
ccccatcttt gtggttttat cgacttttgg tctttgatga tggatgatga cagatgggtt 48600
ttcgggtggt atgtcctttc tgtttgtag ttttcttct aacagacagg accctcagct 48660
gcaggtctgt tggaaataccc tgctgtgtga ggtgtcagtg tgcccctgct ggggggtgcc 48720
tcccagttag gctgctcggg ggtcaggggt cagggacca cttgaggaag cagtctgcc 48780
gttctcagat ctccagctgc gtgctgggag aaccactgct ctcttcaaag ctgtcagaca 48840
gggacattta agtctgcaga agttaacttct gtctttttgt ttgtctgtgc cctgccccca 48900
gaggtggagc ctacagaggc aggcaggcct ccttgagctg tgggtgggctc caccagttc 48960
gagcttcccg gctgctttgt ttacctaatc aagcctgggc aatggcgggc gtccctcccc 49020
cagcctcatt gccgccttgc agtttgatct cagactgctg tgctagcaat cagcgagact 49080
ccgtggggtg aggaccctcg agccagggtgc ggggtataatc tcgtggtgtg ccgttttttg 49140
tttttttttt gtttgtttgt ttgtttgtt ttgagacgga gtctcgctct gtcgcccagg 49200
ctggagtgca gtggcgggat ctcggtctac tgcaagctcc gcctgccggg ttcacgccat 49260
tctcctgcct cagcctccca agtagctggg actacaggcg cccgccacta cgcccggcta 49320
attttttgta ttttttagtag agacgggggtt tcaccgtttt agccgggatg gtctcgatct 49380
cctgacctcg tgatccgccc gcctcggcct cccaaagtgc tgggattaca ggcgtgagcc 49440
accgcgcgg gccgggtgccc cgttttttaa gcccgtcggg aaagcgcagt attcgggtgg 49500
gagtgacctg attttccagg tgctgcccct caccctttc tttgactagg aaagggaact 49560
ccctgaccac ttgtgcttcc caagtgaggc aatgcctcgc cctgcttcgg ctcgcgcacg 49620
gtgcgtgac cactgacct gcgcccactg tctggcactc ctagtgaga tgaacccggg 49680
acctcagatg gaaatgcaga aatcacccgt cttctgcgtt gctcaggctg ggagctgtag 49740
accggagctg ttcctatttg gccatcttgg ctctcccca agagctggta attattacc 49800
aacgtttaga atataatgta acagaacgta tttggaatac atttatcttt taatatattca 49860
aaggagagat aatgcaaac taagatttta tttatcatta aattcattgt tcttttagttt 49920
gaaaaacaca tatactcata tttttgtcac gtgattttta ccagcataaa gcaatagctg 49980
ccacaccaa gtcaagtga gctgtgaaaa taattaatga aacaattctg ataaaagcat 50040
ctttatcctt catacaggtt ataggcagga ttgtggtgt gcctgtgtgt gtgtgtggtg 50100
gggtcgggtg ggggatagga aagttagttt attcacatga ttctcaacat agacatacct 50160
tgcaaataat tgtgcaattc cctgctttgc tttgcagaca gaatgtcttt gaatcaatgt 50220
cttttataaa tacactatac aaaaaacata attataaacc tttaaagtat ttatgcatac 50280
tttaatatc agaactagaa gagaaatttc tgtaaggaga aatgctttgg agattaaaac 50340
gtaaaggaga taaccaaatt ttactaaacc tctatccatc atgaccacat ttcctttcat 50400
tctaaaataa aatattttg gaaatgtga cccctgccc accatttttg cctgagaagt 50460
tttaaagtgc aagaacttgc tcctcacttt tgaagcacgg tgctttcttt aggaaccaat 50520
ctgtatcaca tgcattctat tggtttgtat tttattacat atcattctat tacattttaa 50580
aatgcatttc tgagcccatg tcattgattt catgattcat taatgggttg caagcagcag 50640

ctggaataac actggccttag gtaatggcag tataaatcac atgttggttaa aatgaaagat 50700
gcctttctac caacacgtta taatggccac tcgatgccaa catggcagtg tgttacttaa 50760
accatatatc aatcaacaca gaagcgacat aaccaaagt tttttacat catttataga 50820
agattatgag tttcatgcag ccaggattaa ttgcattgct attatagaac ttaataaag 50880
ttgtatcata tttgatggaa ggatgctaac tatcaagagg atataaattt taaaaacaat 50940
gatactgttt tggataagtc ctgactgtct gcattttatc tagaacaata tatgttcatg 51000
attaacctag aaaataatga caaaaatgcg caatgagtag agacttgccc ttggaattcc 51060
taaaattatt aggaaggtag aatgtttctc aagaaatatt ggtaagttaa aaaaacaagg 51120
tttacaataa tgtgtaacgg atgctatctg tgtaacaaaa tgtgaaaaaa catatatgca 51180
atTTTTTggc aagaagatac cctgaaaaca taaacaaaa agccaaaaaa aaaaaaagaa 51240
acttgctgat gatacagaac ttctctgact gtacattttt ctcatagtct taactctgat 51300
aattacatac atgtttgaca ttctcaaaag taagaacata gaagtacata atatatgcag 51360
gataaagcaa ataacagttt atgtgaagat tgtagaagc caaggttttc atggaagtta 51420
ttagtatgaa gtcaaagtag tacattaana tccagtaatg ttaattggat ttggaagtaa 51480
cagtatgaaa ttttttctct aaaaatgtgt gcaggaatct cttgcttaag atcagtaatc 51540
tattcccaaa aatagatttt gagagcacat ttctgttcta aatgggaaca attatagtga 51600
atggaaggta aacaaaaaca tagctatttc ctacaaaata attgatacct aataatgcct 51660
gtagaaaagc atagaaaagt gacaaactct cttgataaaa tctaaaatgt atattgtttc 51720
ctaactcagca attattgatg ttgctaacat acagttgctt tgtatctagt tacatggacc 51780
gtgtatatgc tgtatgagaa attgtcactt tgtggtatag catagcaaat aacacagcaa 51840
tcaaaaagtt aaaagtcaaa attaaaataa ataaaagacc cataatagac tgtttttcac 51900
cttgctaaag gagaaatgac tcttgaagaa actcaagaag ggacaaaagaa aggtcaaatc 51960
agtcattgct atttgcagaa ggtgatgggt ggatggagga caagggggca atcttcgaca 52020
aactgtcata tgaagaggaa tgaagagacc ttccccatgt ctgtgtctag tagatgctgt 52080
agagacttgc tggacaggaa agctttggta aatcttcaga aatgatgttc aagctgaagg 52140
aagatgttca agctgcattt tggttgcaaa atgcagattt gcttgttttt ctggtgaaaa 52200
attcctctat gtgatgctgt aagagttacc aggctattat taaaagaaag atgtattgtc 52260
aactgagta tgtctgcaaa actttcttag taactggcct gtttgtttat actcaatgca 52320
catgttaaga tggctacata ggattcttac tatgttctat caatttggaa tgtaactcaa 52380
cattcttcca gcacataaag ttgtttacaa tatataatga gtgtgaaaga tataaataaa 52440
tacttctatt tgatgtctac accaacatta gatgaaaaaa tgaggtagtg agaagttgat 52500
atatatcctg taaattagca tatttaaaac caaaaatcct tctactgggt aaatattaac 52560
taggaaaaca ttctacagga tataatgcat ttttctagct ctgccccctg aaaaggctta 52620
gtagtaatgt gatctgatgg cactaaacac ttctgatctc aatgatagc ctctaaaatg 52680
ctgttttcca acaaagagaa ttagggctcc tgaagaaatg gctgatgcca ggtgtggagg 52740
agaaaatgag caagctgagc ttgtgatata tgttcacccc agagagggt catgtcacgt 52800
gggcaccaac agagaggggt ctcaggtagc aacagtgggg ccattggacc atcaaaggag 52860
caattagggt acttgactga agcctgtagg ttcataattt ttactctaaa acaaggaagg 52920

aaggtgggaa gaaaggggaa aggaggggaa ggggaaggaag aaaggagggga gggaggggaaag 52980
gaaggaagta aagagctaga acaacaatga aacaaaatat ctttgataac tttttggggg 53040
agtacagagt ggaaatthttg cttttctgga ggctttacaa gatggggatc tgtgcttcag 53100
aaataagtca gtaaagatac atttacatag atgtttcctc cagtattgct tggaaaagcc 53160
aaaatcaaag tgatgtagat agttacctaa gagatgaata aaataaataa tggtagatct 53220
tttgtagtga atattacata gccactaaaa ataatggagc agaactctac tcaatgacat 53280
ataaaaatgt ccatgatata ttgtataatg gaaaaaaatt ttaaagcaaa aggttttacc 53340
ttatthtttt cacaagctaa aaccctatat gattatacaa accaatgtgt ttgccttaaa 53400
aagataaaag tctgggtgtcc tctaagacca taccacaaac agtgatcacc tctgggattg 53460
gaatthttata gcagttgtgg gtagagagag tttgggagaa aacattcaca ttttatctca 53520
taattcttat atthtttaaag cttttctatg agcattatth tgatacatat gtgagcagag 53580
atggctcctt ctaggtacca gttacaccat tgctaaacct cgtctgtttc ttccaaaggt 53640
gtcaggcagc gaacaagacc tgccccacca attacatgtg gaataatcac atctgcagat 53700
gctggctca ggaagatthtt atgtthttcct cggatgctgg agatggtaag cagaatgatg 53760
ctthaaagac taaacgccag gggctatat tgtgattaac atactctaat attaaaacca 53820
cacatataat atthcatcatt cttctaaata agtgacnnnn nnnnnnnnnn nnnnnnnnnn 53880
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 53940
nnnnnnnnnn nnnnnngggg cttcggctg ccagctgtgg accccacaaa gaactagaca 54000
gaaactcatg ccagtgtgtc tgtaaaaaca aactcttccc cagccaatgt ggggccaacc 54060
gagaatthga tgaaaacaca tgccagtgtg tatgtaaaag aacctgcccc agaatcaac 54120
ccataaatcc tggaaaatgt gcctgtgaat gtacagaaaag tccacagaaa tgcttgtaa 54180
aaggaaaagaa gttccaccac caaacatgca ggtaagagat cctcatgaag aatattaatt 54240
tatcttacta gaaatatgct ctaaaccaaa agcagthttc ttaaccaaag taacaattgc 54300
ttgttagaat gaggccaata cattctgaaa gacaaaaatc ttatathttt ttaaccaatg 54360
ctaaacttgc atgagactgg atthgttcc aaaaaagcaa gacaggagcc ttgaggatca 54420
aaattagatc caagctcaga atatactota aattagthgt tctgtggga tgtthttcaa 54480
atacctaaag gcaatccact tcacccccac caaaaagtag gatgccttca gacataatth 54540
aggacaactg ttctaaatgc agacaattht tatgcaaaac ttacaaaca aggacaggag 54600
agtaccactg gtatagaatt atatctggaa atgaaaatac atagctaatg atthgataga 54660
attagagaaa taaatgaata tgttgtaaa ggtacccaag actcatgtct taaactthgt 54720
thtgactctt ctgcatcaag aaaaaaaaaa tgaagthatt cgactcaagc acatthcagth 54780
ttaatcccga aacctcttcc tctthaatct cagagaataa atthgatccta ctatthtaagc 54840
cagacctgtc cgthtgacat thttctccaca thcttcaac atthgttcat gccatthctt 54900
ctthctctg taactthaac atcttcaact ctactthttt tccccttagt ttataaaagt 54960
tcccaaatht cttctathtt atthagctct tactgtgaag thgagcaaac thtagcaacc 55020
caggtctgth ctctthcaag ctgctcaaaa acgaggcatt tgtthttcacc tctthcacct 55080
accatgtht caaccaacta aaacctgtht tatggcttac agthgtatgt taaacctctg 55140
thtaacgtht catthgcctc atagthgcca aatccaacag agaaththgt cctthagthaa 55200

cctgtccttt ccatggcctt tttctttaag ctcacctctt ctatgacatg acattgtact 55260
ttcctagttg tctacttaat tttcagacca tttgttttta atcaactttca ctggattctc 55320
tttttacatc cattctttac atatgggtag ataatgttct ttttaacaga aaaaagcttg 55380
tatttcattt atttactatc ttcactaatg ttcattatta tttttccaca gtatataacg 55440
agttcatctt tttattctca agagttttgt gtccttattt taatgtatat ttttgggtgg 55500
tttcctaaac taatatttta gctactggat aaaaaggctt tagtggatct ggctatctca 55560
ttgattatat ttttcttccc ttatactgca aagtacatca taggtataat tttattttga 55620
gaaagaacta tattccaatt ttacctatat attcctactg taataaccaa gaggacattt 55680
caaaaatgta tatgctacta ttttcagaag gatatagata ctaaactctt ttaagcctta 55740
ttaaactatg ttttatattg aattcttatt gagtgaattt agtatacatt gaaaaatctg 55800
aaattcttat gtctgcatac ctacttttct acagaattat atgaacacag acaatatttc 55860
tatcattgca tttctcctaa tactatacat tctttcaaaa agactctgtt tactgcaact 55920
catcctagat tcatttttct aagtgtataa ttccaagctt tgtgatataa attggagcaa 55980
aagagaaaat atcatatfff gaaaatatat attctcctca ttctctctta atatctatta 56040
caaagtgcag taggcagagt atatattctt tggttcattg ccttaattaa cagcttttat 56100
ggatcattaa gtatcatatc cttttttagt agaggaatac cagtagtcat aactatgaat 56160
ttgcatttgc aaagtftaag aaaaattaac tagacatgct ttcttttcta gagccaaata 56220
aatgaaatgt caaaaccaa ttgtgtaggg ttttttatag cacacttttg atttccacct 56280
caggtcatct tgcacttttc cttttacacc tattcccaca cattaggcat gctcactctc 56340
caaaatgttt ttaaaagata gctccaattc tcaccaggcg tgggtggctca cacctgtaat 56400
cctagcactt tgggaaggcca aggcgggcag atcacttgag gccaggagt caagaccagc 56460
ccagccaaca tggcaaaacc ctggctctat gaaaaatata gaaacttagc caggcgtggc 56520
cacacatacc tgtaatcca gctactcggg aggctgagac acaagaatgg cttgagccca 56580
ggaggcggag gttgcagtaa gctgagattg cccactgcg ctccagcctg ggcgacagaa 56640
caagactctg tctcaaaaaa tgaagcctca attcactgga gcaacagggc aggagacaca 56700
ctgaccaatt gataatttgt cctctcaccg ctactggaaa ctggaagtct aggatgggag 56760
gaaggggacc ctgctgctg agtattgtcc ttctcacctc cctgccccag aatgcctctg 56820
tttcacctgg agccaaggtg gaaaggaaga gttagtcacc gtcatgcagc cagctgcca 56880
gtagagacac aagagtgagg gtggggctgc tcaactgagga ccacagttct gccaacatgg 56940
ctgggtttta ttgcctggaa gtctgggtcc catacaagtc ccagactcat gtgaattctg 57000
ccatgcttat gtgatcacag taagtcagca gaccttcagt gttcactttc agaatttctg 57060
ttttatgttt ctttcacca tgaagcaaag agcacttcaa aggaactaat gcttacaata 57120
cctcatcctt ccatttgtta tatcagcaag caagactttt attttaaggc ttttccattc 57180
ctcttaaagt ttgtaattct aagcagaaac aaactttttt aaacctatga acaatttctg 57240
cataaaatta gtaattttat tccagtctga aatttaaaaa cacagaaata cttgggtagc 57300
atgatgaatc cattgccttg atctttaacg taagtgtgtt cttgtttgtt ctcttagctg 57360
ttacagacgg ccatgtacga accgccagaa ggcttgtgag ccaggatttt catatagtga 57420
agaagtgtgt cgttgtgtcc cttcatattg gaaaagacca caaatgatga gctaagattg 57480

tactgttttc cagttcatcg attttctatt atggaaaact gtggtgccac agtagaactg 57540
tctgtgaaca gagagaccct tgtgggtcca tgctaacaaa gacaaaagtc tgtctttcct 57600
gaacctggtg gataacttta cagaaatgga ctggagctca tctgcaaaag gcctcttgta 57660
aagactgggtt ttctgccaat gaccaaacag ccaagatfff cctcttggtga tttctttaaa 57720
agaatgacta tataatftat ttccactaaa aatattgfff ctgcattcat ttttatagca 57780
acaacaattg gtaaaactca ctgtgatcaa tattttttata tcatgcaaaa tatgftttaaa 57840
ataaaatgaa aattgtatta taagctgcta agttcagfcc attatcatct tacatgatga 57900
acgaaaaacta ctatcatgaa gacactgatc tttctctgcc cttttttgft ctctaaccag 57960
atgtcacata tgtattacta tgataaaaag tatgatcctg tgaaagagag tgtcagagga 58020
caacagaatg ctattgcttc atctcttata tgtttaatga ttataaacat tttagttacat 58080
gatactttttg aatttatgac caagtgaatc aatatgaaac atcttgtaag atagactact 58140
tagcattgtg attaaaagtc attcagtgct ctgagaacat tcagaatctt acgttggtag 58200
aaaatcctgc agtatatatt aaaatggctt taaatatttt ctcaaaaata atcttttcca 58260
aatatttgac tttttctggc cagctaaaat actttttgtg agtggaaagt ctcctatcca 58320
atacatttta aaaaaaact aaaaataata tttaaact catgcatgft taaaaggaaa 58380
catttcagcc acacaattga agtgcttftt tcattttcaa aaaggcatta cactaaaata 58440
cctttatctt tttttttttt tttcttcaga tagtfttft ctgfttgccca ggctggagtg 58500
cagtggcgtg atctcagctc actgccccta acctctgctt cccggattca agtgattctc 58560
ctgcctcagc ctctgagta gctgggatta taggcacgtg ccggcacacc tggctaattt 58620
ttgtatfttt agtgcagatg gggtttcacc atgfttgcca gtctgatcac gaactcctga 58680
cctcaagtga tccacagacg tcggcctcaa agtgctggga ttacaggcat gaaccaccgt 58740
gcccagccta aaataccttt tactaaaaca aagattttgc ctgcataaat aaaaccaaat 58800
aagtggftttg ttcgfttgaca gagaattata tactggtaca tcataactac tgtataaaaa 58860
taattagfcc tgaaaagaga aatattcct cataagcatg taaaggtagc acattaattt 58920
aactaaaaca tttttgfttt tttgaaacgg aatctcgtc tfttgccag gctggagtgc 58980
agtggtgoga tctcggctca c 59001

<210> 14

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 14

gcaggccgaa gccccgctct 20

<210> 15

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 15

tcttttccaa tatgaaggga 20

<210> 16

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 16

ggctccgcgt tcccaacttt 20

<210> 17

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 17
cttttgcaga tgagctccag 20

<210> 18
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 18
ttgcttgcatt aagccgtggc 20

<210> 19
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 19
ttgtttggtc cacagatgtc 20

<210> 20
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 20
taataatgga atgaacttgt 20

<210> 21
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 21
gtaactgctc ctccagatct 20

<210> 22
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 22
ctccagtcca tttctgtaaa 20

<210> 23
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 23

ggtgtagctt tttggagagg

20

<210> 24

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 24

tcgtacatgg ccgtctgtaa

20

<210> 25

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 25

tgtggctttt tccaatatga

20

<210> 26

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

Ro

<223> Antisense Oligonucleotide

<400> 26

cttgcataag ccgtggcctc

20

<210> 27

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 27

catggaatcc atctggtgag

20

<210> 28

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 28

gtttggtcat tggcagaaaa

20

<210> 29

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 29

ccttctggcg gttcgtacat

20

<210> 30

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 30

gtttgttttt acagacacac

20

<210> 31

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 31

ttcgctgcct gacactgtgg

20

<210> 32

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 32

acaacgacac acttcttcac

20

<210> 33

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 33

gtctgtaaac atccagttta

20

<210> 34

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 34

acatattttg catgatataa

20

<210> 35

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 35

gtgagtttta ccaattggtg

20

<210> 36

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 36

caagggtctc tctgttcaca

20

<210> 37

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 37

gtaaaagcct cacaggaac

20

<210> 38

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 38

atatgaaaat cctggctcac

20

<210> 39

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 39

gacacacatg gaggtttaa

20

<210> 40

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 40

attgagtctt tctccactca

20

<210> 41

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 41

aatcttcctg agccagggcat

20

<210> 42

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 42

agatgagctc cagtccattt

20

<210> 43

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 43

ttggctgttt ggtcattggc

20

<210> 44

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 44

gcaagtgcac ggtggaagga

20

<210> 45

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 45

gaatgcagaa acaatatttt

20

<210> 46

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 46

tagtcattct tttaaagaaa

20

<210> 47

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 47

cttggctggt tggtcattgg

20

<210> 48

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 48

aggaaaatct tggctgtttg

20

<210> 49

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 49

cctttcctta gctgacactt

20

<210> 50

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 50

taaacatcca gtttagacat

20

<210> 51

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 51

gcgggtgtca ggtaaaagcc

20

<210> 52

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 52

aggccgcggg cccctcctgg 20

<210> 53

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 53

aagaagccca gcaagtgcac 20

<210> 54

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 54

cactggacac agaccgtaac 20

<210> 55

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 55
 tctgtccttg agttgaggtt 20

<210> 56
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 56
 atactttttca agatctctgt 20

<210> 57
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 57
 ttatcaatac ttttcaagat 20

<210> 58
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 58

actattgcag caacccccac

20

<210> 59

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 59

gtggtgctgg cagggaacgt

20

<210> 60

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 60

catctccagc atccgaggaa

20

<210> 61

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 61

tcacccagct ccttgtttgg

20

<210> 62

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 62

gtccacagct ggcaggccga

20

<210> 63

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 63

ttcttttaca tacacactgg

20

<210> 64

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 64

acattcacag gcacattttc

20

<210> 65

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 65

ggtggtggaa cttctttcct

20

<210> 66

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 66

gtacaatctt agctcatttg

20

<210> 67

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 67

cacagacagt tctactgtgg

20

<210> 68

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 68

aataaattat atagtcattc

20

<210> 69

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 69

aattggtggt gctataaaaa

20

<210> 70

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 70

agttgcctga tgatccaaga

20

<210> 71

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 71

tttatcctcg gccactcccg

20

<210> 72

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 72

tttagagggtg atgcgaccac

20

<210> 73

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 73

ctccctggag ctccccgtt

20

<210> 74

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 74

actctccctc ggaagcctc

20

<210> 75

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 75

ccttccccga agtgagagga

20

<210> 76

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 76

tgacgaaatt gttaaaaggt

20

<210> 77

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 77

atttcagact gaaatacaat

20

<210> 78

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 78

gcagacctac cgtggcctcg

20

<210> 79

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 79

ccatacttac ttttcaagat

20

<210> 80

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 80

caatacccac cgtcttgctg

20

<210> 81

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 81

acacattttg tacaggtatc

20

<210> 82

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 82

aggaaacacg atgatgccca

20

<210> 83

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 83

ggagaacttt gaagcagttt

20

<210> 84

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 84

tcatactcac tgtggtagtg

20

<210> 85

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Antisense Oligonucleotide

<400> 85

attctttcat tgtcagagct

20

<210> 86

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 86

aaagttggga acgcggagcc

20

<210> 87

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 87

ctggagctca tctgcaaaag

20

<210> 88

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 88

gccacggctt atgcaagcaa

20

<210> 89

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 89

gacatctgtg gaccaaaca

20

<210> 90

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 90

agatctggag gagcagttac

20

<210> 91

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 91
ttacagacgg ccatgtacga 20

<210> 92
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 92
tcatattgga aaagaccaca 20

<210> 93
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 93
gaggccacgg cttatgcaag 20

<210> 94
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 94
ttttctgcca atgaccaaac 20

<210> 95

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 95

gtgtgtctgt aaaaacaaac

20

<210> 96

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 96

taaactggat gtttacagac

20

<210> 97

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 97

caacaattgg taaaactcac

20

<210> 98

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 98
gtttcctgtg aggcttttac 20

<210> 99
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 99
tttaaacctc catgtgtgtc 20

<210> 100
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 100
tgagtggaga aagactcaat 20

<210> 101
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 101
gccaatgacc aaacagccaa 20

<210> 102

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 102

tccttcacc atgcacttgc

20

<210> 103

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 103

ccaatgacca aacagccaag

20

<210> 104

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 104

caaacagcca agattttcct

20

<210> 105

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 105

ggcttttacc tgacacccgc

20

<210> 106

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 106

ccaggagggg cccgcggcct

20

<210> 107

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 107

atgcacttgc tgggcttctt

20

<210> 108

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 108

gttacggtct gtgtccagtg

20

<210> 109

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 109

aacctcaact caaggacaga

20

<210> 110

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 110

acagagatct tgaaaagtat

20

<210> 111

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 111

gtggggggtg ctgcaatagt

20

<210> 112

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 112

acgttccctg ccagcaacac

20

<210> 113

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 113

tgggcctgcc agctgtggac

20

<210> 114

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 114

ccagtgtgta tgtaaaagaa

20

<210> 115

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 115

gaaaatgtgc ctgtgaatgt

20

<210> 116
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 116
aggaaagaag ttccaccacc 20

<210> 117
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 117
caaatgagct aagattgtac 20

<210> 118
<211> 20
<212> DNA
<213> H. sapiens

<220>

<400> 118
ccacagtaga actgtctgtg 20

<210> 119
<211> 20
<212> DNA

<213> H. sapiens

<220>

<400> 119

aaacggggag ctccaggag

20

<210> 120

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 120

tcctctcact tcggggaagg

20

<210> 121

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 121

cgaggccacg gtaggtctgc

20

<210> 122

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 122

cagcaagacg gtgggtattg

20

<210> 123

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 123

gatacctgta caaaatgtgt

20

<210> 124

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 124

tgggcatcat cgtgtttcct

20

<210> 125

<211> 20

<212> DNA

<213> H. sapiens

<220>

<400> 125

agctctgaca atgaaagaat

20