(19) World Intellectual Property Organization International Bureau





(43) International Publication Date 27 February 2003 (27.02.2003)

PCT

(10) International Publication Number WO 03/015777 A1

(51) International Patent Classification7: A61K 31/426, C07D 277/40, 277/42, 277/44, 277/46

(21) International Application Number: PCT/US02/25438

(22) International Filing Date: 13 August 2002 (13.08.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

01119473.5 13 August 2001 (13.08.2001) EP 10/185,731 1 July 2002 (01.07.2002)

(71) Applicant (for all designated States except US): LION BIOSCIENCE AG [DE/DE]; Waldhofer Strasse 98, D-69123 Heidelberg (DE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): BAUER, Ulrike [—/DE]; Wingerstr. 24, D-69207 Sandhausen (DE). CHERUVALLATH, Zach [-/US]; 7813 Roan Road, San Diego, CA 92129 (US). DEUSCHLE, Ulrich [—/DE]; Lilienstr. 19, D-69245 Bammental (DE). **DNE-**PROVSKAIA, Elena [—/US]; 4080 Huerfano Avenue, #351, San Diego, CA 92117 (US). GAHMAN, Tim [—/US]; 262 Chapalita, Encinitas, CA 92024 (US). **GIEGRICH, Kristina** [—/DE]; Firnheimerstr. D-68623 Lampertheim (DE). HANECAK, Ronnie [-/US]; 904 Calle Venezia, San Clemente, CA 92672 (US). HEBERT, Normand [—/US]; 1861 Montgomery Avenue, Cardiff, CA 92007 (US). **KIELY, John** [—/US]; 4230 Corte Facil, San Diego, CA 92130 (US). KOBER, Ingo [—/US]; Am Grossen Wald 30, D-69251 Gaiberg (DE). KOGL, Manfred [-/DE]; Hauptstr. 131/4. D-69214 Eppelheim (DE). KRANZ, Harald [—/DE]; Barenpfad 9, D-69181 Leimen (DE). KREMOSER,

Claus [—/DE]; Mozartstr. 29, D-69121 Heidelberg (DE). LEE, Matthew [—/US]; 1544 Santa Elena Court, Solana Beach, CA 92075 (US). OTTE, Kerstin [—/DE]; Biethstr. 20, D-69121 Heidelberg (DE). SAGE, Carlton [—/US]; 1232 Greenlake Drive, Cardiff, CA 92007 (US). SUD, Manish [—/US]; 4411 Cather Avenue, San Diego, CA 92122 (US).

(74) Agents: TOFENETTI, Judith, L. et al.; McDermott, Will & Emery, 600 13th Street, N.W., Washington, DC 20005 (US).

(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, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(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, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

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.

(54) Title: NR1H4 NUCLEAR RECEPTOR BINDING COMPOUNDS

(57) Abstract: The present invention relates to compounds according to the general formula (I) which bind to the NR1H4 receptor and act as agonists of the NR1H4 receptor. The invention further relates to the treatment of diseases and/or conditions through binding of the nuclear receptor by the compounds and the production of medicaments using the compounds.

NR1H4 NUCLEAR RECEPTOR BINDING COMPOUNDS

Field of the Invention

[0001] The present invention relates to compounds that bind to the NR1H4 nuclear receptor and methods of treating diseases or pathological conditions influenced by NR1H4.

Background of the Invention

[0002] Multicellular organisms are dependent on advanced mechanisms of information transfer between cells and body compartments. The information that is transmitted can be highly complex and can result in the alteration of genetic programs involved in cellular differentiation, proliferation, or reproduction. The signals, or hormones, are often simple molecules, such as peptides, fatty acid, or cholesterol derivatives.

[0003] Many of these signals produce their effects by ultimately changing the transcription of specific genes. One well-studied group of proteins that mediate a cell's response to a variety of signals is the family of transcription factors known as nuclear receptors, hereinafter referred to often as "NR". Members of this group include receptors for steroid hormones, vitamin D, ecdysone, cis and trans retinoic acid, thyroid hormone, bile acids, cholesterol-derivatives, fatty acids (and other peroxisomal proliferators), as well as so-called orphan receptors, proteins that are structurally similar to other members of this group, but for which no ligands are known (Escriva, H. et al., Ligand binding was acquired during evolution of nuclear receptors, PNAS, 94, 6803-6808, 1997). Orphan receptors may be indicative of unknown signaling pathways in the cell or may be nuclear receptors that function without ligand activation. The activation of transcription by some of these orphan receptors may occur in the absence of an exogenous ligand and/or through signal transduction pathways originating from the cell surface (Mangelsdorf, D. J. et al., The nuclear receptor superfamily: the second decade, Cell 83, 835-839, 1995).

In general, three functional domains have been defined in NRs. An [0004] amino terminal domain is believed to have some regulatory function. A DNA-binding domain hereinafter referred to as "DBD" usually comprises two zinc finger elements and recognizes a specific Hormone Responsive Element (hereinafter referred to as "HRE") within the promoters of responsive genes. Specific amino acid residues in the "DBD" have been shown to confer DNA sequence binding specificity (Schena, M. & Yamamoto, K.R., Mammalian Glucocorticoid Receptor Derivatives Enhance Transcription in Yeast, Science, 241:965-967, 1988). A ligand-binding-domain (hereinafter referred to as "LBD") is at the carboxy-terminal region of known NRs. In the absence of hormone, the LBD appears to interfere with the interaction of the DBD with its HRE. Hormone binding seems to result in a conformational change in the NR, and thus, opens this interference (Brzozowski et al., Molecular basis of agonism and antagonism in the oestogen receptor, Nature, 389, 753-758, 1997; Wagner et al., A structural role for hormone in the thyroid hormone receptor, Nature, 378, 690 -697. 1995). A NR without the LBD constitutively activates transcription, but at a low level.

[0005] Coactivators or transcriptional activators are proposed to bridge between sequence specific transcription factors, the basal transcription machinery, and in addition, to influence the chromatin structure of a target cell. Several proteins like SRC-1, ACTR, and Grip1 interact with NRs in a ligand enhanced manner (Heery et al., A signature motif in transcriptional coactivators mediates binding to nuclear receptors, Nature, 387, 733-736; Heinzel et al., A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression, Nature 387, 43-47, 1997). Furthermore, the physical interaction with negative receptor-interacting proteins or corepressors has been demonstrated (Xu et al., Coactivator and Corepressor complexes in nuclear receptor function, Curr Opin Genet Dev, 9 (2), 140-147, 1999).

[0006] Nuclear receptor modulators like steroid hormones affect the growth and function of specific cells by binding to intracellular receptors and forming nuclear receptor-ligand complexes. Nuclear receptor-hormone complexes then interact with a HRE in the control region of specific genes and alter specific gene expression.

[0007] The Farnesoid X Receptor alpha (hereinafter to as "FXR" and also often referred to as "NR1H4" when referring to the human receptor) is a prototypical type 2 nuclear receptor which activates genes upon binding to promoter region of target genes in a heterodimeric fashion with Retinoid X Receptor (hereinafter referred to as "RXR") (Forman et al., Cell, 81, 687-93, 1995). The relevant physiological ligands of NR1H4 seem to be bile acids (Makishima et al., Science, 284, 1362-65, 1999; Parks et al., Science, 284, 1365-68, 1999). The most potent is chenodeoxycholic acid, which regulates the expression of several genes that participate in bile acid homeostasis. Farnesoid, originally described to activate the rat ortholog at high concentration does not activate the human or mouse receptor. FXR is expressed in the liver, small intestine, colon, ovary, adrenal gland and kidney. Like FXRα, NR1H4 is involved in intracrine signaling.

[0008] FXR is proposed to be a nuclear bile acid sensor. As a result, it modulates both, the synthetic output of bile acids in the liver and their recycling in the intestine (by regulating bile acid binding proteins). Upon activation (e.g. binding of chenodeoxycholic acid) it influences the conversion of dietary cholesterol into bile acids by inhibiting the transcription of key genes which are involved in bile acid synthesis such as CYP7A1. This seems to be a major mechanism of feedback regulation onto bile acid synthesis.

[0009] The synthetic compounds, 1,1-bisphosphonate esters, appear to display a number of similar activities to the two identified prototypes of natural FXR agonists, farnesol, and chenodeoxycholc acid. Like farnesol, the 1,1-bisphosphonate esters increase the rate of 3-Hydroxy-3-methylglutaryl-CoA (HMG-CoA) reductase degradation and like bile acids they induce the expression of the intestinal bile acid binding protein (hereinafter referred to as "I-BABP") and repress the cholesterol 7 α -hydroxylase gene. Certain 1,1-bisphosphonate esters also bind to FXR (Niesor et al., Curr Pharm Des,7(4):231-59, 2001). That means that activation of FXR could lead to opposing effects (lowering the rate of cholesterol synthesis by increasing degradation of HMG-CoA reductase and increasing the cholesterol pool by inhibition of cholesterol degradation into bile acids). The FXR

agonist, chenodeoxycholic acid, does not change cholesterol and lipoprotein levels significantly in patients, although a repression of bile acid synthesis as well as a decreased HMG-CoA reductase activity was observed (Einarsson et al., Hepatology, 33(5), 1189-93, 2001) confirming that cellular cholesterol synthesis and degradation are controlled by numerous regulatory loops including the coordinate regulation of HMGCoA reductase and cholesterol 7α -hydroxylase and that compounds modulating FXR acitvity might have different effects on blood lipid parameters.

[0010] In the course of functional analysis of certain 1,1-bisphosphonate esters, it was shown that these compounds which are known to bind to FXR also induce apoptosis in a variety of cell types, similar to the isporenoids farnesol and geranylgeraniol which are also known as weak FXR binders (Flach et al., Biochem Biophys Res Com, 270, 240-46, 2000).

[0011] To date only very few compounds have been described which bind the NR1H4 receptor, and thus, show utility for treating diseases or conditions which are due to or influenced by this nuclear receptor (Maloney at al., J Med Chem, 10; 43(16): 2971-2974, 2000).

[0012] It is an object of the present invention to provide for a novel NR1H4 binding compound. It was also an object of the present invention to provide for compounds which, by means of binding the NR1H4 receptor, act as an agonist or antagonist of said receptor, and thus, show utility for treating diseases or conditions which are due to or influenced by said nuclear receptor.

[0013] It is a further object of the invention to provide for compounds which may be used for the manufacture of a medicament for the treatment of cholesterol associated conditions or diseases. In a preferred embodiment of the invention it was an object of the invention to provide for cholesterol lowering or cholestatic compounds. It was also an object of the invention to provide for compounds may be used for the manufacture of antitumor medicaments.

Summary of the Invention

[0014] The present invention provides, *inter alia*, novel NR1H4 nuclear receptor protein binding compounds according to the general formula (I) shown below. These compounds are also binders of mammalian homologues of the receptor. Further the object of the invention is solved by providing for, amongst the NR1H4 nuclear receptor protein binding, compounds according to the general formula (I) which act as agonists and compounds which act as antagonists of the human FXR receptor or a mammalian homologue thereof.

[0015] The invention provides for FXR agonists which may be used for the treatment of cholesterol associated conditions or diseases. In a preferred embodiment of the invention cholesterol lowering or cholestatic compounds are disclosed. The compounds according to the invention may be used for manufacture of antitumor medicaments and/or for the treatment of diseases such as cancer.

[0016] The foregoing merely summarizes certain aspects of the present invention and is not intended, nor should it be construed, to limit the invention in any manner. All patents and other publications recited herein are hereby incorporated by reference in their entirety.

Brief Description of the Drawings

[0017] The various novel features of this invention, along with the foregoing and other objects as well as the invention itself may be more fully understood from the following description when read in conjunction with the accompanying drawings.

[0018] Figs. 1A and 1B show the synthesis of the compounds of the invention described in Example 2.

[0019] Fig. 2 shows the synthesis of the compounds of the invention described in Example 3.

[0020] Fig. 3 shows the synthesis of the compounds of the invention described in Example 4.

[0021] Fig. 4A shows SEQ ID No. 1 which is a protein sequence of the FRX protein, a portion of which can be used for cloning.

[0022] Fig. 4B shows SEQ ID NO. 2 which is the mRNA sequence encoding the FRX protein.

[0023] Fig. 4C shows SEQ ID NO. 3 which shows the protein sequence of TIF2 (Acc. No.: XM 011633 RefSeq DB).

[0024] Fig. 4D shows SEQ ID NO. 4 which is respective mRNA sequence corresponding to the TIF2 protein.

[0025] Fig 5 A shows a dose response with LN6348 in the HEK293-FXR reporter assay

[0026] Fig 5B shows a dose response with LN6316 in the HEK293-FXR reporter assay

[0027] Fig 5C shows a dose response with LN6365 in the HEK293-FXR reporter assay

[0028] Fig 5D shows a dose response with LN6322 in the HEK293-FXR reporter assay

Detailed Description of the Invention

[0025] The invention provides for a compound including resolved diastereoisomers and enantiomers, and tautomers, pharmaceutical acceptable salts or solvates thereof (hereinafter also referred to as the "compounds according to the invention"), having the following formula (I):

$$R_3 N$$
 R_4
 R_5
 R_1
 R_2
(I)

wherein

[0026] R₁ and R₂ are independently selected from the group consisting of C₁ to C₈ alkyl, C₁ to C₈ substituted alkyl, phenyl, substituted phenyl, C₇ to C₁₂ phenylalkyl, C₇ to C₁₂ substituted phenylalkyl, heterocyclic ring, substituted heterocyclic ring, heteroaryl, and substituted heteroaryl; and

[0027] R_3 and R_4 are independently selected from the group consisting of hydrogen, C_1 to C_8 alkyl, C_1 to C_8 substituted alkyl, phenyl, substituted phenyl, C_7 to C_{12} phenylalkyl, C_7 to C_{12} substituted phenylalkyl, naphthyl, substituted naphthyl, C_1 to C_8 alkanesulfonyl, C_1 to C_8 substituted alkanesulfonyl, benzenesulfonyl, substituted benzenesulfonyl, C_1 to C_8 acyl, and C_1 to C_8 substituted acyl; where R_3 and R_4 may be taken together with nitrogen to form a heterocycle or substituted heterocycle or a heteroaryl or substituted heteroaryl ring.

In one embodiment of the present invention, R_1 and R_2 in formula (I) are independently selected from the group consisting of phenyl, substituted phenyl, C_7 to C_{12} phenylalkyl, C_7 to C_{12} substituted phenylalkyl, heteroaryl, and substituted heteroaryl; R_3 and R_4 are independently selected from the group consisting of hydrogen, C_1 to C_8 alkyl, C_1 to C_8 substituted alkyl, substituted phenyl, C_7 to C_{12} phenylalkyl, C_7 to C_{12} substituted phenylalkyl, substituted naphthyl, C_1 to C_8 substituted alkanesulfonyl, and substituted benzenesulfonyl; where R_3 and R_4 may be taken together with nitrogen to form a heterocycle or a substituted heterocycle or a heteroaryl or a substituted heteroaryl ring.

[0029] In a more preferred embodiment of the present invention, R_1 and R_2 are independently selected from the group consisting of phenyl, substituted phenyl,

 C_7 to C_{12} phenylalkyl, C_7 to C_{12} substituted phenylalkyl, heteroaryl, and substituted heteroaryl; R_3 and R_4 are independently selected from the group consisting of hydrogen, C_1 to C_8 substituted alkyl, substituted phenyl, C_7 to C_{12} substituted phenylalkyl, substituted naphthyl, C_1 to C_8 substituted alkanesulfonyl, and substituted benzenesulfonyl, wherein at least one of the mentioned above groups is substituted with carboxylic acid functionality as shown in formula (II) below:

COOH
$$R_3 \qquad R_4$$

$$R_1 \qquad R_2 \qquad (II)$$

where R₃ and R₄ may be taken together with nitrogen to form a heterocycle or a substituted heterocycle or a heteroaryl or substituted heteroaryl ring, also substituted with carboxylic acid functionality.

[0030] In a more preferred embodiment of the invention compounds are claimed, or pharmaceutical acceptable salts or solvates thereof, wherein R_1 and R_2 are independently selected from the group consisting of phenyl, substituted phenyl, C_7 to C_{12} phenylalkyl, C_7 to C_{12} substituted phenylalkyl, heteroaryl, and substituted heteroaryl; R_3 is hydrogen, C_1 to C_8 alkyl, C_1 to C_8 substituted alkyl, C_1 to C_8 acyl, C_1 to C_8 substituted alkanesulfonyl, benzenesulfonyl, and substituted benzenesulfonyl; and R_4 is one of the following structural formulas:

wherein n is an integer from 0 to 3. The symbol:

indicates the point of attachment of R₄ where R₄ is covalently bonded to Formula (I).

[0031] In a more preferred embodiment of the invention, there are provided compounds including resolved diastereoisomers and enantiomers, and tautomers, pharmaceutical acceptable salts or solvates thereof, wherein R_1 and R_2 are independently selected from the group consisting of substituted phenyl, C_7 to C_{12} substituted phenylalkyl, and substituted heteroaryl, where preferred substituents are taken from hydrogen, halogen, hydroxy or alkoxy groups; R_3 is hydrogen, C_1 to C_8 alkyl, C_1 to C_8 substituted alkyl, C_1 to C_8 acyl, and C_1 to C_8 substituted acyl; and C_9 is one of the structures set forth above.

[0032] A particularly preferred compound which may act as agonist of NR1H4 is shown in formula (III) below. The inventors have been able to demonstrate that the compound according to formula (III) has a low effective binding concentration at FXR with an EC $_{50}$ of 0.2 μ M wherein the EC $_{50}$ reflects the half-maximal effective

concentration, and which is higher than the EC $_{50}$ of 0.015 μ M for the published FXR agonist GW4064 (B.Goodwin et al., Molecular Cell 6, 517-526, 2000).

[0033] The inventors have also found the compounds according to formulas (IV), (V) and (VI) shown below to be active as agonist of the NR1H4 human nuclear receptor (see figures for details).

[0034] The inventors have identified the compounds as well as the general structure capable of effectively binding FXR.

[0035] The compounds of the invention can also exist as solvates and hydrates. Thus, these compounds may crystallize with, for example, waters of hydration, or one, a number of, or any fraction thereof of molecules of the mother liquor solvent. The solvates and hydrates of such compounds are included within the scope of this invention.

[0036] A solid carrier can be one or more substances which can also act as diluents, flavoring agents, solubilizers, lubricants, suspending agents, binders, or tablet disintegrating agents; it can also be an encapsulating material.

[0037] In powders, the carrier is generally a finely divided solid, which is in a mixture with the finely divided active component. In tablets, the active compound or combination of active compounds is mixed with the carrier having the necessary binding properties in suitable proportions and compacted in the shape and size desired.

[0038] For preparing a pharmaceutical composition in the form of suppositories, a low-melting wax such as a mixture of fatty acid glycerides and cocoa butter is first melted and the active ingredient or combination of active ingredients is dispersed therein by, for example, stirring. The molten homogeneous mixture is then poured into convenient-sized molds and allowed to cool and solidify.

[0039] Powders and tablets preferably contain between about 5% to about 70% by weight of the active ingredient. Suitable carriers include, for example, magnesium carbonate, magnesium stearate, talc, lactose, sugar, pectin, dextrin, starch, tragacanth, methyl cellulose, sodium carboxymethyl cellulose, a low-melting wax, cocoa butter and the like.

[0040] The pharmaceutical compositions can include the formulation of the active compound(s) with encapsulating material as a carrier providing a capsule in which the active component (with or without other carriers) is surrounded by a carrier, which is thus in association with it. In a similar manner, cachets are also included. Tablets, powders, cachets, and capsules can be used as solid dosage forms suitable for oral administration.

[0041] Liquid pharmaceutical compositions include, for example, solutions suitable for oral or parenteral administration, or suspensions, and emulsions suitable for oral administration. Sterile water solutions of the active component or sterile

solutions of the active component in solvents comprising water, ethanol, or propylene glycol are examples of liquid compositions suitable for parenteral administration.

[0042] Sterile solutions can be prepared by dissolving the active component(s) in the desired solvent system, and then passing the resulting solution through a membrane filter to sterilize it or, alternatively, by dissolving the sterile compound in a previously sterilized solvent under sterile conditions.

[0043] The term "halogen" refers to the fluoro, chloro, bromo or iodo atoms. There can be one or more halogen, which are the same or different. Preferred halogens are chloro and fluoro.

[0044] The term " C_1 to C_8 alkyl" denotes such radicals as methyl, ethyl, n-propyl, isopropyl, n-butyl, iso-butyl, sec-butyl, tert-butyl, amyl, tert-amyl, hexyl, n-heptyl, 2-heptyl, 3-heptyl, 4-heptyl, 2-methyl-1-hexyl, 2-methyl-2-hexyl, 2-methyl-3-hexyl, n-octyl, and the like.

[0045] The term " C_1 to C_8 substituted alkyl" denotes that the above C_1 to C_8 alkyl groups are substituted by one or more, and preferably one or two, halogen, hydroxy, protected hydroxy, oxo, protected oxo, C_3 to C_7 cycloalkyl, phenyl, naphthyl, amino, protected amino, monosubstituted amino, protected monosubstituted amino, disubstituted amino, guanidino, protected guanidino, heterocyclic ring, substituted heterocyclic ring, C_1 to C_8 alkoxy, C_1 to C_8 acyl, C_1 to C_8 acyloxy, nitro, carboxy, protected carboxy, carbamoyl, carboxamide, protected carboxamide, N-(C_1 to C_6 alkyl)carboxamide, N,N-di(C_1 to C_6 alkyl)carboxamide, cyano, methylsulfonylamino, thiol, C_1 to C_4 alkylthio or C_1 to C_4 alkylsulfonyl groups. The substituted alkyl groups may be substituted once or more, and preferably once or twice, with the same or with different substituents.

[0046] Examples of the above substituted alkyl groups include the 2-oxo-prop-1-yl, 3-oxo-but-1-yl, cyanomethyl, nitromethyl, chloromethyl, hydroxymethyl, tetrahydropyranyloxymethyl, trityloxymethyl, propionyloxymethyl, amino, methylamino, aminomethyl, dimethylamino, carboxymethyl, allyloxycarbonylmethyl,

allyloxycarbonylaminomethyl, methoxymethyl, ethoxymethyl, t-butoxymethyl, acetoxymethyl, 4-carboxybutyl, 5-carboxypentyl, 6-carboxyhexyl, chloromethyl, bromomethyl, iodomethyl, trifluoromethyl, 6-hydroxyhexyl, 2,4-dichloro(n-butyl), 2-aminopropyl, 1-chloroethyl, 2-chloroethyl, 1- bromoethyl, 2-chloroethyl, 1-fluoroethyl, 2-fluoroethyl, 1- iodoethyl, 2-iodoethyl, 1-chloropropyl, 2-chloropropyl, 3-chloropropyl, 1-bromopropyl, 2-bromopropyl, 3-bromopropyl, 1-fluoropropyl, 2-fluoropropyl, 3-fluoropropyl, 1- iodopropyl, 2-iodopropyl, 3-iodopropyl, 2-aminoethyl, 1- aminoethyl, N-benzoyl-2-aminoethyl, N-acetyl-2-aminoethyl, N-benzoyl-1-aminoethyl, N-acetyl-1-aminoethyl, and the like.

[0047] The term "substituted phenyl" specifies a phenyl group substituted with one or more, and preferably one or two, moieties chosen from the groups consisting of halogen, hydroxy, protected hydroxy, cyano, nitro, C_1 to C_8 alkyl, C_1 to C_8 substituted alkyl, C_1 to C_8 alkoxy, C_1 to C_8 substituted alkoxy, C_1 to C_8 acyloxy, carboxy, protected carboxy, carboxymethyl, protected carboxymethyl, hydroxymethyl, protected hydroxymethyl, amino, protected amino, monosubstituted amino, protected monosubstituted amino, disubstituted amino, carboxamide, protected carboxamide, N-(C_1 to C_6 alkyl)carboxamide, protected N-(C_1 to C_6 alkyl)carboxamide, trifluoromethyl, N-((C_1 to C_6 alkyl)sulfonyl)amino, N- (phenylsulfonyl)amino or phenyl, wherein the phenyl is substituted or unsubstituted, such that, for example, a biphenyl results.

[0048] Examples of the term "substituted phenyl" includes a mono- or di(halo)phenyl group such as 2, 3 or 4-chlorophenyl, 2,6-difluorophenyl, 2,3-difluorophenyl, 2,6-dichlorophenyl, 2,5-dichlorophenyl, 3,4-dichlorophenyl, 2, 3 or 4-bromophenyl, 3,4-dibromophenyl, 3-chloro-4-fluorophenyl, 2, 3 or 4-fluorophenyl and the like; a mono or di(hydroxy)phenyl group such as 2, 3 or 4-hydroxyphenyl, 2,4-dihydroxyphenyl, the protected-hydroxy derivatives thereof and the like; a nitrophenyl group such as 2, 3 or 4-nitrophenyl; a cyanophenyl group, for example, 2, 3 or 4-cyanophenyl; a mono- or di(alkyl)phenyl group such as 2, 3 or 4-methylphenyl, 2,4-dimethylphenyl, 2, 3 or 4-(iso-propyl)phenyl, 2, 3 or 4-ethylphenyl, 2, 3 or 4-(n-propyl)phenyl and the like; a mono or di(alkoxyl)phenyl group, for example,

2,6-dimethoxyphenyl, 2, 3 or 4-methoxyphenyl, 2, 3 or 4-ethoxyphenyl, 2, 3 or 4-(isopropoxy)phenyl, 2, 3 or 4-(t-butoxy)phenyl, 3-ethoxy-4-methoxyphenyl and the like; 2, 3 or 4-trifluoromethylphenyl; a mono- or dicarboxyphenyl or (protected carboxy)phenyl group such as 2, 3 or 4-carboxyphenyl or 2,4-di(protected carboxy)phenyl; a mono- or di(hydroxymethyl)phenyl or (protected hydroxymethyl)phenyl such as 2, 3, or 4-(protected hydroxymethyl)phenyl or 3,4-di(hydroxymethyl)phenyl; a mono- or di(aminomethyl)phenyl or (protected aminomethyl)phenyl such as 2, 3 or 4-(aminomethyl)phenyl or 2,4-(protected aminomethyl)phenyl; or a mono- or di(N-(methylsulfonylamino))phenyl such as 2, 3 or 4-(N-(methylsulfonylamino))phenyl. Also, the term "substituted phenyl" represents disubstituted phenyl groups wherein the substituents are different, for example, 3-methyl-4-hydroxyphenyl, 3-chloro-4-hydroxyphenyl, 2-methoxy-4-bromophenyl, 4-ethyl-2-hydroxyphenyl, 3-hydroxy-4-nitrophenyl, 2-hydroxy 4-chlorophenyl, and the like.

[0049] The term "C₇ to C₁₂ phenylalkyl" denotes a C₁ to C₆ alkyl group substituted at any position by a phenyl, substituted phenyl, heteroaryl or substituted heteroaryl. Examples of such a group include benzyl, 2-phenylethyl, 3-phenyl(n-propyl), 4-phenylhexyl, 3-phenyl(n-amyl), 3-phenyl(sec-butyl), and the like. Preferred C₇ to C₁₂ phenylalkyl groups are the benzyl and the phenylethyl groups.

[0050] The term " C_7 to C_{12} substituted phenylalkyl" denotes a C_7 to C_{12} phenylalkyl group substituted on the C_1 to C_6 alkyl portion with one or more, and preferably one or two, groups chosen from halogen, hydroxy, protected hydroxy, oxo, protected oxo, amino, protected amino, monosubstituted amino, protected monosubstituted amino, disubstituted amino, guanidino, protected guanidino, heterocyclic ring, substituted heterocyclic ring, C_1 to C_8 alkyl, C_1 to C_8 substituted alkyl, C_1 to C_8 alkoxy, C_1 to C_8 substituted alkoxy, C_1 to C_8 acyloxy, nitro, carboxy, protected carboxy, carbamoyl, carboxamide, protected carboxamide, N-(C_1 to C_6 alkyl)carboxamide, protected N-(C_1 to C_6 alkyl)carboxamide, cyano, N-(C_1 to C_6 alkylsulfonyl)amino, thiol, C_1 to C_4 alkylthio, C_1 to C_4 alkylsulfonyl groups; and/or the phenyl group may be substituted with one or more, and preferably one or two,

substituents chosen from halogen, hydroxy, protected hydroxy, cyano, nitro, C_1 to C_8 alkyl, C_1 to C_8 substituted alkyl, C_1 to C_8 alkoxy, C_1 to C_8 substituted alkoxy, C_1 to C_8 acyloxy, carboxy, protected carboxy, carboxymethyl, protected carboxymethyl, hydroxymethyl, protected hydroxymethyl, amino, protected amino, (monosubstituted)amino, protected (monosubstituted)amino, (disubstituted)amino, carboxamide, protected carboxamide, N-(C_1 to C_6 alkyl) carboxamide, protected N-(C_1 to C_6 alkyl) carboxamide, N, N-di(C_1 to C_6 alkyl)carboxamide, trifluoromethyl, N-((C_1 to C_6 alkyl)sulfonyl)amino, N-(phenylsulfonyl)amino, cyclic C_2 to C_8 alkylene or a phenyl group, substituted or unsubstituted, for a resulting biphenyl group. The substituted alkyl or phenyl groups may be substituted with one or more, and preferably one or two, substituents which can be the same or different.

[0051] Examples of the term "C₇ to C₁₂ substituted phenylalkyl" include groups such as 2-hydroxyphenylmethyl, 3-hydroxyphenylmethyl, 2-methoxyphenylmethyl, 3-methoxyphenylmethyl, 2,6-difluorophenylmethyl, 2,3-difluorophenylmethyl, 2,6-dichlorophenylmethyl, 2,3-dichlorophenylmethyl, 3,5-dichlorophenylmethyl, 2-hydroxyphenylethyl, 3-hydroxyphenylethyl, 2-methoxyphenylethyl, 3-methoxyphenylethyl, 2,6-difluorophenylethyl, 2,3-dichlorophenylethyl, 2,3-dichlorophenylethyl, 3,5-dichlorophenylmethyl, 2,6-dichlorophenylethyl, 2,3-dichlorophenylethyl, 3,5-dichlorophenylmethyl 2-phenyl-1-chloroethyl, 2-(4-methoxyphenyl)ethyl, 4-(2,6-dihydroxy phenyl)n-hexyl, 2-(5-cyano-3-methoxyphenyl)n-pentyl, 3-(2,6-dimethylphenyl)n-propyl, 4-chloro-3-aminobenzyl, 6-(4-methoxyphenyl)-3-carboxy(n-hexyl), 5-(4-aminomethylphenyl)- 3-(aminomethyl)n-pentyl, 5-phenyl-3-oxo-n-pent-1-yl, and the like.

[0052] The term "heterocycle" or "heterocyclic ring" denotes optionally substituted five-membered to eight-membered rings that have 1 to 4 heteroatoms, such as oxygen, sulfur and/or nitrogen, in particular nitrogen, either alone or in conjunction with sulfur or oxygen ring atoms. These five-membered to eight-membered rings may be saturated, fully unsaturated or partially unsaturated, with fully saturated rings being preferred. Preferred heterocyclic rings include morpholino,

piperidinyl, piperazinyl, 2-amino-imidazoyl, tetrahydrofurano, pyrrolo, tetrahydrothiophen-yl, hexamethyleneimino and heptamethyleneimino.

[0053] The term "substituted heterocycle" or "substituted heterocyclic ring" means the above-described heterocyclic ring is substituted with, for example, one or more, and preferably one or two, substituents which are the same or different which substituents can be halogen, hydroxy, protected hydroxy, cyano, nitro, C₁ to C₈ alkyl, C₁ to C₈ alkoxy, C₁ to C₈ substituted alkoxy, C₁ to C₈ acyl, C₁ to C₈ acyloxy, carboxy, protected carboxy, carboxymethyl, protected carboxymethyl, hydroxymethyl, protected hydroxymethyl, amino, protected amino, (monosubstituted)amino, protected (monosubstituted)amino, (disubstituted)amino carboxamide, protected carboxamide, N-(C₁ to C₁₂ alkyl)carboxamide, protected N-(C₁ to C₆ alkyl)carboxamide, trifluoromethyl, N-((C₁ to C₆ alkyl)sulfonyl)amino, N-(phenylsulfonyl)amino, heterocycle or substituted heterocycle groups.

[0054] The term "heteroaryl" means a heterocyclic aromatic derivative which is a five-membered or six-membered ring system having from 1 to 4 heteroatoms, such as oxygen, sulfur and/or nitrogen, in particular nitrogen, either alone or in conjunction with sulfur or oxygen ring atoms. Examples of heteroaryls include pyridinyl, pyrimidinyl, and pyrazinyl, pyridazinyl, pyrrolo, furano, thiopheno, oxazolo, isoxazolo, phthalimido, thiazolo, and the like.

[0055] The term "substituted heteroaryl" means the above-described heteroaryl is substituted with, for example, one or more, and preferably one or two, substituents which are the same or different which substituents can be halogen, hydroxy, protected hydroxy, cyano, nitro, C₁ to C₈ alkyl, C₁ to C₈ alkoxy, C₁ to C₈ substituted alkoxy, C₁ to C₈ acyl, C₁ to C₈ substituted acyl, C₁ to C₈ acyloxy, carboxy, protected carboxy, carboxymethyl, protected carboxymethyl, hydroxymethyl, protected hydroxymethyl, amino, protected amino, (monosubstituted)amino, protected (monosubstituted)amino, (disubstituted)amino, carboxamide, protected carboxamide, N-(C₁ to C₆ alkyl)carboxamide, protected N-(C₁ to C₆

alkyl)carboxamide, N, N-di(C_1 to C_6 alkyl)carboxamide, trifluoromethyl, N-((C_1 to C_6 alkyl)sulfonyl)amino or N-(phenylsulfonyl)amino groups.

[0056] The term "substituted naphthyl" specifies a naphthyl group substituted with one or more, and preferably one or two, moieties either on the same ring or on different rings chosen from the groups consisting of halogen, hydroxy, protected hydroxy, cyano, nitro, C₁ to C₈ alkyl, C₁ to C₈ alkoxy, C₁ to C₈ acyl, C₁ to C₈ acyloxy, carboxy, protected carboxy, carboxymethyl, protected carboxymethyl, hydroxymethyl, protected hydroxymethyl, amino, protected amino, monosubstituted amino, protected monosubstituted amino, disubstituted amino, carboxamide, protected carboxamide, N-(C₁ to C₆ alkyl)carboxamide, N, N-di(C₁ to C₆ alkyl)carboxamide, trifluoromethyl, N-((C₁ to C₆ alkyl)sulfonyl)amino or N-(phenylsulfonyl)amino.

[0057] Examples of the term "substituted naphthyl" includes a mono or di(halo)naphthyl group such as 1, 2, 3, 4, 5, 6, 7 or 8-chloronaphthyl, 2, 6dichloronaphthyl, 2, 5-dichloronaphthyl, 3, 4-dichloronaphthyl, 1, 2, 3, 4, 5, 6, 7 or 8bromonaphthyl, 3, 4-dibromonaphthyl, 3-chloro-4-fluoronaphthyl, 1, 2, 3, 4, 5, 6, 7 or 8-fluoronaphthyl and the like; a mono or di(hydroxy)naphthyl group such as 1, 2, 3, 4, 5, 6, 7 or 8-hydroxynaphthyl, 2, 4-dihydroxynaphthyl, the protected-hydroxy derivatives thereof and the like; a nitronaphthyl group such as 3- or 4-nitronaphthyl; a cyanonaphthyl group, for example, 1, 2, 3, 4, 5, 6, 7 or 8-cyanonaphthyl; a mono- or di(alkyl)naphthyl group such as 2, 3, 4, 5, 6, 7 or 8-methylnaphthyl, 1, 2, 4-dimethylnaphthyl, I, 2, 3, 4, 5, 6, 7 or 8-(isopropyl)naphthyl, I, 2, 3, 4, 5, 6, 7 or 8-ethylnaphthyl, I, 2, 3, 4, 5, 6, 7 or 8-(n-propyl)naphthyl and the like; a mono or di(alkoxy)naphthyl group, for example, 2, 6-dimethoxynaphthyl, 1, 2, 3, 4, 5, 6, 7 or 8-methoxynaphthyl, 1, 2, 3, 4, 5, 6, 7 or 8-ethoxynaphthyl, I, 2, 3, 4, 5, 6, 7 or 8-(isopropoxy)naphthyl, 1, 2, 3, 4, 5, 6, 7 or 8-(t-butoxy)naphthyl, 3-ethoxy-4methoxynaphthyl and the like; 1, 2, 3, 4, 5, 6, 7 or 8-trifluoromethylnaphthyl; a monoor dicarboxynaphthyl or (protected carboxy)naphthyl group such as 1, 2, 3, 4, 5, 6, 7 or 8-carboxynaphthyl or 2, 4-di(-protected carboxy)naphthyl; a mono-or di(hydroxymethyl)naphthyl or (protected hydroxymethyl)naphthyl such as 1, 2, 3, 4, 5, 6, 7 or 8-(protected hydroxymethyl)naphthyl or 3, 4-di(hydroxymethyl)naphthyl; a

mono- or di(amino)naphthyl or (protected amino)naphthyl such as 1, 2, 3, 4, 5, 6, 7 or 8-(amino)naphthyl or 2, 4-(protected amino)-naphthyl, a mono- or di(aminomethyl)naphthyl or (protected aminomethyl)naphthyl such as 2, 3, or 4-(aminomethyl)naphthyl or 2, 4-(protected aminomethyl)-naphthyl; or a mono- or di- (N-methylsulfonylamino) naphthyl such as 1, 2, 3, 4, 5, 6, 7 or 8-(N-methylsulfonylamino)naphthyl. Also, the term "substituted naphthyl" represents disubstituted naphthyl groups wherein the substituents are different, for example, 3-methyl-4-hydroxynaphth-1-yl, 3-chloro-4-hydroxynaphth-2-yl, 2-methoxy-4-bromonaphth-1-yl, 4-ethyl-2-hydroxynaphth-1-yl, 3-hydroxy-4-nitronaphth-2-yl, 2-hydroxy-4-chloronaphth-1-yl, 2-methoxy-7-bromonaphth-1-yl, 4-ethyl-5-hydroxynaphth-2-yl, 3-hydroxy-8-nitronaphth-2-yl, 2-hydroxy-5-chloronaphth-1-yl and the like.

[0058] As outlined above R_3 and R_4 may be taken together with nitrogen to form a heterocycle or substituted heterocycle of the following kind aziridine, azetidine, pyrrolidine, 3-methylpyrrolidine, 3-aminopyrrolidine, 3-hydroxypyrrolidine, pyrazolidine, imidazolidine, piperidine, 2-methylpiperidine, 4-carboxypiperidine, 4-(carboxymethyl)piperidine, piperazine, morpholine, azepine, and tetrahydroisoquinoline.

[0059] The term "C₁ to C₈ acyl" encompasses groups such as formyl, acetyl, propionyl, butyryl, pentanoyl, pivaloyl, hexanoyl, heptanoyl, benzoyl and the like. Preferred acyl groups are acetyl and benzoyl.

[0060] The term " C_1 to C_8 substituted acyl" denotes the acyl group substituted by one or more, and preferably one or two, halogen, hydroxy, protected hydroxy, oxo, protected oxo, cyclohexyl, naphthyl, amino, protected amino, monosubstituted amino, protected monosubstituted amino, disubstituted amino, guanidino, heterocyclic ring, substituted heterocyclic ring, imidazolyl, indolyl, pyrrolidinyl, C_1 to C_8 alkoxy, C_1 to C_8 acyl, C_1 to C_8 acyloxy, nitro, C_1 to C_8 alkyl ester, carboxy, protected carboxy, carbamoyl, carboxamide, protected carboxamide, N-(C_1 to C_6 alkyl)carboxamide, protected N-(C_1 to C_6 alkyl)carboxamide, cyano, methylsulfonylamino, thiol, C_1 to C_4 alkylthio or C_1 to C_4 alkylsulfonyl groups. The

substituted acyl groups may be substituted once or more, and preferably once or twice, with the same or with different substituents.

[0061] Examples of C₁ to C₈ substituted acyl groups include 4-phenylbutyroyl, 3-phenylbutyroyl, 3-phenylpropanoyl, 2- cyclohexanylacetyl, cyclohexanecarbonyl, 2-furanoyl and 3-dimethylaminobenzoyl and the like.

[0062] The term " C_1 to C_8 alkoxy" as used herein denotes groups such as methoxy, ethoxy, n-propoxy, isopropoxy, n-butoxy, t-butoxy and like groups. A preferred alkoxy is methoxy. The term " C_1 to C_8 substituted alkoxy" means the alkyl portion of the alkoxy can be substituted in the same manner as in relation to C_1 to C_8 substituted alkyl.

[0063] The term " C_1 to C_8 substituted aminoacyl" denotes the acyl group substituted by one or more, and preferably one or two, halogen, hydroxy, protected hydroxy, oxo, protected oxo, cyclohexyl, naphthyl, amino, protected amino, monosubstituted amino, protected monosubstituted amino, disubstituted amino, guanidino, heterocyclic ring, substituted heterocyclic ring, imidazolyl, indolyl, pyrrolidinyl, C_1 to C_8 alkoxy, C_1 to C_8 acyl, C_1 to C_8 acyloxy, nitro, C_1 to C_8 alkyl ester, carboxy, protected carboxy, carbamoyl, carboxamide, protected carboxamide, N-(C_1 to C_6 alkyl)carboxamide, protected N-(C_1 to C_6 alkyl)carboxamide, cyano, methylsulfonylamino, thiol, C_1 to C_6 alkylthio or C_1 to C_6 alkylsulfonyl groups. The substituted acyl groups may be substituted once or more, and preferably once or twice, with the same or with different substituents.

[0064] This invention provides a pharmaceutical composition comprising an effective amount of a compound according to the invention. Such compositions can be administered by various routes, for example oral, rectal, subcutaneous, intramuscular, intravenous or intracerebral. The preferred route of administration would be oral at daily doses of the compound for adult human treatment of about 0.01-5000 mg, preferably 1-1500 mg per day. The appropriate dose may be administered in a single dose or as divided doses presented at appropriate intervals for example as two, three four or more subdoses per day.

[0065] For preparing pharmaceutical compositions containing compounds of the invention, inert, pharmaceutically acceptable carriers are used. The pharmaceutical carrier can be either solid or liquid. Solid form preparations include, for example, powders, tablets, dispersible granules, capsules, cachets, and suppositories.

[0066] In particular, the invention relates to compounds as described above wherein the compounds are capable of binding the NR1H4 receptor protein or a portion thereof according to SEQ ID NOS 1-4 shown in Figs. 4A-4D, respectively, or a mammalian homologue thereof. The claimed compound can bind to the NR1H4 receptor protein or a portion thereof in a mixture comprising 10-200 ng of NR1H4 receptor protein or a portion thereof, preferably the ligand binding domain, 20 mM Tris /HCl at pH 7.9; 60 mM KCl; 5 mM MgCl₂; 160ng/μl BSA in a total volume of preferably about 25 μl.

[0067] A mammalian receptor protein homologue of the protein according to SEQ ID NO. 1 shown in Fig. 4A as used herein is a protein that performs substantially the same function as NR1H4 does in humans and shares at least 40% sequence identity at the amino acid level, preferably 50% sequence identity at the amino acid level more preferably 65% sequence identity at the amino acid level and most preferably over 85% sequence identity at the amino acid level.

[0068] The invention in particular concerns a method for prevention or treatment of a NR1H4 receptor protein- or NR1H4 receptor protein homologue-mediated disease or condition in a mammal comprising administration of a therapeutically effective amount of a compound according to the invention wherein the prevention or treatment is directly or indirectly accomplished through the binding of a compound according to the invention to the NR1H4 receptor protein or to the NR1H4 receptor protein homologue.

[0069] The term mediated herein means that the physiological pathway in which the NR1H4 receptor protein acts is either directly or indirectly involved in the disease or condition to be treated or prevented. In the case where it is indirectly involved it could be that, e.g. modulating the activity of NR1H4 by a compound according to the invention influences a parameter which has a beneficial effect on a disease or a condition. One such example is that modulation of NR1H4 activity leads to decreased levels of serum cholesterol or certain lipoproteins, which, in turn, have a beneficial effect on the prevention and treatment of artherosclerosis. Herein a condition is a physiological or phenotypic state which is desirably altered. One such example would be obesity which is not necessarily medically harmful but nonetheless a non desirable phenotypic condition. In a preferred embodiment of the invention the method for prevention or treatment of a NR1H4 receptor protein mediated disease or condition is applied to a human. This may be male or female.

[0070] Pharmaceutical compositions generally are administered in an amount effective for treatment or prophylaxis of a specific condition or conditions. Initial dosing in a human is accompanied by clinical monitoring of symptoms, such symptoms being determined for the selected condition. In general, the compositions are administered in an amount of active agent of at least about 100 µg/kg body weight. In most cases they will be administered in one or more doses in an amount not in excess of about 20 mg/kg body weight per day. Preferably, in most cases, the dose is from about 100 µg/kg to about 5 mg/kg body weight, daily.

[0071] For administration particularly to mammals, and particularly humans, it is expected that the daily dosage level of active agent will be 0.1 mg/kg to 10 mg/kg and typically around 1 mg/kg.

[0072] By "therapeutically effective amount" is meant a symptom- alleviating or symptom -reducing amount, a cholesterol-reducing amount, a fatty acid absorption blocking amount, a protein and/or carbohydrate digestion-blocking amount and/or a de novo cholesterol biosynthesis-blocking amount of a compound according to the invention. The term "blocking" as used herein means either total blockage or partial blockage.

[0073] FXR is proposed to be a bile acid sensor. As a result, it modulates both the synthetic output of bile acids in the liver and their recycling in the intestine, by regulating bile acid binding proteins. In one embodiment of the invention the invention concerns a method for regulating bile transport in a mammal, in a preferred embodiment a human, which comprises activating the NR1H4 receptor with a therapeutically effective amount of a compound according to the invention.

[0074] Likewise the invention concerns a method of treating in mammal a disease which is affected by cholesterol, triglyceride, or bile acid levels comprising administering to a mammal in need of such treatment a therapeutically effective amount of a compound according to the invention.

[0075] Accordingly, the compounds according to the invention may also be used as a method of prevention or treatment of mammalian atherosclerosis, gallstone disease, lipid disorders, obesity or cardiovascular disorders such as coronary heart disease or stroke.

[0076] The invention further concerns a method of blocking fatty acid absorption in the intestine of a mammal comprising administering to the mammal a therapeutically effective amount of a compound according to the invention. The invention may also be used to treat obesity in a mammal, particularly in humans.

[0077] FXR alpha is a prototypical type 2 nuclear receptor which activates genes upon binding to the promoter region of target genes in a heterodimeric fashion with RXR. The relevant physiological ligands of NR1H4 are bile acids. The present compounds according to the invention have been demonstrated to have high binding efficacy binding coefficients measured as IC50 in the range 400 nM to 1000 nM as well as agonistic and/or antagonistic properties. Consequently they may be applied to regulate genes that participate in bile acid homeostasis as well as other downstream regulated genes. Examples of such genes are, but are not limited to, genes encoding proteins or factors involved directly or indirectly in lipid absorption, cholesterol biosynthesis, cholesterol transport or binding, bile acid transport or

binding, proteolysis, amino acid metabolism, glucose biosynthesis, protein translation, electron transport, and hepatic fatty acid metabolism. FXR often functions in vivo as a heterodimer with the RXR. Published FXR agonists such as the Glaxo SmithKline compound "GW 4064" are known to influence the regulation of various liver genes. Examples of known agonists are showin in Table 1 below.

TABLE 1

[0078] Genes found to be regulated by GW 4064 are genes down-regulated in the liver, genes up-regulated in the liver and genes having altered expression in the intestine.

[0079] Genes down-regulated in the liver include apolipoprotein B; plasma proteinase inhibitor alpha-1-inhibitor III group 3(m22360); L-glucono-gamma-lactone oxidase (d12754); peroxisomal enoyl-CoA:hydrotase-3-hydroxyacyl-CoA bifunctional enzyme (k03249); liver fatty acid binding protein (L-FABP, m13501), CYP4A2 (m57719), CYP3A23 (x96721) and CYP3A1 (x64401); Cholesterol-7-alpha-hydroxylase, CYP7A1 (RefSeq NM000780, XM 005022, XM 044651, and XM 044652); and sodium-taurocholate cotransport protein, ntcp (RefSeq NM003049, XM007466).

[0080] Genes that up-regulate in the liver include small heterodimer partner homolog (d86580); bile salt export pump, bsep (RefSeq NM 003742, XM 003644,

and XM 033122); phospholipid transfer protein, PLTP (RefSeq NM 006227, XM 009490, XM 029929, and XM 029930); arnithine palmitoyltransferase II, CPTII (RefSeq NM 000098, XM 001758, XM 038866, and XM 038867); phenylethanolamine-N-methyltransferase, PNMT (RefSeq NM 002686, XM 008597, andXM 049837); insulin-induced growth-response protein CL-6 (I13619); elongation factor 2, EF-2 (y07504); mouse cornichon; protein kinase C receptor (u03390); mitochondrial cytochrome c oxidase (m27315); cystathione gamma-lyase (x53460, d17370); cytosolic phosphoenolypyruvate carboxykinase (k03243); histidase (m58308); S-adenosylmethionine synthetase (x60822); lanosterol 14-alphademethylase (u17697); G protein-coupled purinoceptor P2U (146865) and hepatic squalene synthetase (m95591).

[0081] Genes having altered expression in the intestine include lipase (x61925); pancreatic lipase (d88534); colipase (m58370); pancreatic phospholipase A-2 (d00036); pancreatic amylase (m24962); carboxypeptidase A1 (m23986); carboxypeptidase A2 (m23721): carboxypeptidase B (m23959): pancreatic trypsin I (j00778): pancreatic cationic trypsinogen (m16624); pancreatic trypsinogen II (v01274); elastase I (v01234, I00112); elastase II (I00118, I00124); I-BABP (I22788); intestinal fatty acid binding protein (FABP, k01180); hepatic squalenesynthetase (m95591); protein kinase C receptor (u003390): longation factor 2, EF-2 (y07504) and small heterodimer partner homolog (d86580).

[0082] Thus, the invention also concerns a method of modulating a gene whose expression is regulated by the NR1H4 receptor in a mammal comprising administration of a therapeutically effective amount of a compound according to the invention to the mammal.

[0083] It is known that the orphan receptor FXR can bind the response element of the SHP gene as a heterodimer with RXR (9-cis retinoic acid receptor) and the SHP-protein, in turn, prevents efficient transcription from the cyp7a1 promoter (Lu et al., Mol Cell, 6(3):505-17; Goodwin et al. Mol Cell, 6(3), 717-26, 2000). Another gene that is repressed via SHP upon FXR activation is the sodium/bile acid cotransporter gene, NTCP, a membrane transport protein which is

required for the import of conjugated bile acids in the hepatocyte (Denson et al., Gastroenterology;121(1):218-20, 2001). The gene for the bile salt export pump, a membrane transporter responsible for the secretion of bile acids into the gall is directly activated by FXR (Ananthanarayanan et al., J Biol Chem, 3;276(31):28857-28865, 2001). Consequently, the invention likewise concerns a method for lowering the expression of cholesterol 7-alpha-hydroxylase and NTCP and increasing expression of BSEP in parallel by use of the compounds according to the invention. In one embodiment the invention concerns a method for enhancing the expression of the Intestinal Bile Acid Binding Protein (I-BABP) (Grober et al., J Biol Chem, 15;274(42):29749-54, 1999 and/or the activity of the canicular bile salt excretion pump.

[0084] The compounds according to the invention may be used as medicaments, in particular for the prevention or treatment of a NR1H4 receptor protein- or NR1H4 receptor protein homologue- mediated disease or condition in a mammal wherein the prevention or treatment is directly or indirectly accomplished through the binding of the compound according to the invention to the NR1H4 receptor protein or NR1H4 receptor protein homologue. These pharmaceutical compositions contain 0.1% to 99.5% of the compound according to the invention, more particularly 0.5% to 90% of the compound according to the invention in combination with a pharmaceutically acceptable carrier.

[0085] The invention also concerns the use of a compound or combination of compounds according to the invention for the prevention or treatment of a NR1H4 receptor protein-mediated disease or condition wherein the mammal described above is a human. The medicament may be used for regulating the bile transport system in a mammal, preferentially a human, by activating the NR1H4 receptor, for regulating levels of cholesterol, triglyceride, and/or bile acid. For example, the medicament may be used for the treatment of atherosclerosis, gallstone disease, lipid disorders, obesity or a cardiovascular disorder.

[0086] The invention further concerns the use of a compound or combination of compounds according to the invention for blocking in a mammal, preferentially a

human, fatty acid absorption in the intestine. Further the compounds of the invention may be used for treating obesity in humans and for modulating a gene whose expression is regulated by the NR1H4 receptor (see details above and figures). The invention also further concerns the use of a compound or combination of compounds as antitumor medicaments. The antitumor effects of such medicaments could be exerted by selective inhibition of cell proliferation and induction of apotptosis of tumor cells in a way similar to described activities for certain bisphosphonates (Alberts DS, et al., Clin Cancer Res 2001 May;7(5):1246-50).

[0087] The following examples illustrate a specific embodiment of the invention, but they are not to be considered as limiting the invention in any manner.

EXAMPLE 1

[8800] For in vitro screening for compounds which influence FXR binding to coactivators, a fragment of the open reading frame of human FXR alpha (NR1H4 -(Acc. No:AF384555)) encoding aminoacids 187-472 was amplified by standard RT PCR procedures (see SEQ ID NOS. 1 and 2 in Figs. 4A and 4B, respectively). The starting material was total RNA derived from human liver. The resulting cDNA obtained after reverse transcription was subsequently cloned using the Gateway[™] recombination technology (Invitrogen, USA) into the expression plasmid pDest15 (Invitrogen, USA). This construct was used to express a recombinant GST-FXR fusion protein in E.coli (BL21 strain). A pDEST 17 derivative clone harboring an additional sequence encoding amino acids 548-878 of human TIF2 (Acc. No: XM 011633 RefSeq) was constructed using Gateway[™] recombination technology (Invitrogen, USA) in order to obtain a construct which was used to express recombinant His-tagged TIF2 fragment in E. coli. For E. coli expression of both constructs, plasmid DNA was transformed into chemically competent E. coli BL21 (Invitrogen, USA) and cells were grown to an OD600 of 0.4-0.7 before expression was induced by addition of 0.5 mM IPTG according instructions of the manufacturer (Invitrogen). After induction for 8 hours at 30°C, the cells were harvested by centrifugation for 10 minutes at 5000 x g. Fusion proteins were affinity purified using Glutathion sepharose (Pharmacia) or Ni-NTA Agarose (QIAGEN) according to the

instructions of the respective manufacturer. Recombinant proteins were dialyzed against 20 mM Tris/HCL pH 7.9; 60 mM KCl; 5 mM MgCl₂; 1 mM DTT, 0.2 mM PMSF; 10% glycerol. The TIF2 fragment was subsequently biotinylated by addition of 40-120 µl of a biotinamidocaproate N-hydroxysuccinimide-ester (Sigma) solution (20 mg/ml in DMSO). Overhead rotating samples were incubated for 2 hours at room temperature. Unincorporated label was then separated using G25 Gel filtration chromatography (Pharmacia Biotech, Sweden). Protein containing fractions from the column were pooled and tested for activity in the assay as described below.

[0089] For screening of the compound libraries as provided for by the methods described in Examples 2, 3 and 4 below for substances which influence the FXR/Tif 2 interaction, Perkin Elmer LANCE technology was used. This technoligy relies on the binding dependent energy transfer from a donor to an acceptor fluorophore attached to the binding partners of interest. For ease of handling and reduction of background from compound fluorescence, LANCE technology makes use of generic fluorophore labels and time resoved detection. For a detailed description of this technoligy, see Hemmilä I, Blomberg K and Hurskainen P, Timeresolved resonance energy transfer (TR-FRET) principle in LANCE, Abstract of Papers Presented at the 3 rd Annual Conference of the Society for Biomolecular Screening, Sep., California (1997).

[0090] For screening, 20-200 ng of biotinylated Tif 2 fragment and 10-200 ng of GST-FXR fragment were combined with 0.5-2 nM LANCE Eu-(W1024) labelled anti-GST antibody (Perkin Elmer) and 0.5-2 μg of highly fluorescent APC-labeled streptavidin (Perkin Elmer) in the presence of 50 μM of individual compounds to be screened in a total volume of 25 μl of 20 mM Tris /HCl pH 7.9; 60 mM KCl; 5 mM MgCl2; and 160 ng/μl BSA. DMSO content of the samples was kept below 4%. The samples were incubated for a minimum of 60 minutes in the dark at room temperature in FIA-Plates black 384well med. binding (Greiner).

[0091] The LANCE signal was detected by a Perkin Elmer VICTOR2V™ Multilabel Counter applying the detection parameters listed in Table 2 below.

TABLE 2

I ABLE 2
Number of repeats 1
plate: GREINER FIA-Plate black 384 well med. binding
Measurement height 3.50 mm
Label technology TR-F Lance
Emission filter name D615
Emission filter slot A1
Emission aperture Normal
Excitation filter D340
Delay 50 μs
Window time 400 μs
Cycle 1000 µs
Light integrator capacitors 1
Light integrator ref. level 95
Flash energy area High
Flash energy level 223
Flash absorbance measurement No
Beam Normal
Label technology TR-F Lance
Emission filter name D665
Emission filter slot A8
Emission aperture Normal
Excitation filter D340
Delay 50 μs
Window time 400 μs
Cycle 1000 μs
Light integrator capacitors 1
Light integrator ref. level 95
Flash energy area High
Flash energy level 223
Flash absorbance measurement No
Beam Normal

[0092] The results were visualized by plotting the ratio between the emitted light at 665 nm and at 615 nm. For every batch of recombinant proteins amount of proteins and labeling reagents giving the most sensitive detection of hits were determined individually by analysis of dose response curves for chenodeoxycholic acid.

[0093] The methods of preparing 2-aminothiazole derivative compounds and combinatorial libraries are set forth in Examples 2, 3 and 4 below.

EXAMPLE 2

[0094] The following steps describe the experimental procedure for the preparation of the compounds according to the invention utilizing ArgoGel-MB-CHO resin. The synthesis scheme is shown in Figs. 1A and 1B.

Step 1. Synthesis of Chalcone (compound 1)

[0095] To a solution of acetophenone (0.05 mol) and aldehyde (0.05 mol) in 150 ml of MeOH was added 2.5 eq of NaOH as pellets. The reaction bottle was placed on a shaker (slightly exothermic reaction) for 24-48 hrs. Then the reaction container was placed in an acetone-ice bath (-5 to 0 °C) and the reaction mixture was quenched with 10.3 ml of 37 % aqueous HCl (2.5 eq). Solvent was removed *in vacuo* and the residue was partitioned between EtOAc and water. The aqueous layer was discarded and the organic layer was washed with water, brine and dried over Na₂SO₄. Solvent was removed *in vacuo* and the crude chalcone (compound 1) was recrystallized from methanol (MeOH) or a mixture of hexane and ethyl acetate (Hex:EtOAc) 3:1 or purified by silica gel chromatography using a mixture of Hex:EtOAc = 5:1 as an eluent.

Step 2. Preparation of 1,3-Diphenylpropan-1-one (compound 2)

[0096] Chalcone (compound 1) (37.6 mmol) was dissolved in 250 ml of toluene (in some cases 25 ml of MeOH was added to insure a complete dissolution) and 400 mg of 10% Pd/C (0.01 eq) was added to the mixture. The compound was

hydrogenated at 45 psi in a Parr apparatus for 40 min to 3 hrs to form 1,3-diphenylpropan-1-one (compound 2), during which time the reaction mixture was periodically checked for the starting material by TLC (Hex:EtOAc=1:1) to prevent over-reduction. Upon completion of reduction the reaction mixture was filtered through a short pad (2-3 inch) of silica gel and solvent was removed *in vacuo* to afford a crude product, which was either sufficiently pure to use in the next step or was purified by recrystallization from MeOH or Hex:EtOAc=3:1 or, in the case where over-reduction has occurred, by silica gel column chromatography using Hex:EtOAc=5:1 as an eluent.

Step 3. Preparation of 2-Bromo-1,3-Diphenylpropan-1-one (compound 3)

[0097] To a solution of 1,3-diphenylpropan-1-one (compound 2) (7.0 mmol) in 25 ml of anhydrous dioxane, a solution of bromine (7.0 mmol) in 10 ml of dioxane was slowly added dropwise over 30 min. The reaction mixture was left to stir at room temperature for 24 hrs to form 2-bromo-1,3-diphenylpropan-1-one (compound 3). Solvent was removed *in vacuo*. The 2-bromo-1,3-diphenylpropan-1-one product formed was determined by ¹H NMR to be sufficiently pure to be used in the next step.

Step 4. Reductive Amination of Argogel-MB-CHO Resin (4) with Primary Amines to form Aminated Argogel-MB-HCO resin (5)

[0098] Argogel-MB-CHO resin (Argonaut Technologies Inc.) (100 mg each tea-bag [Houghten, U.S. Patent No. 4,631,211], 0.41 mmol/g substitution) (compound 4) was swollen in 1% acetic acid (AcOH) in DMF (by volume). The amine (10 eq.) was added and the bottle(s) placed on a shaker for 30 min. Solid NaBH₃CN (20 eq) was added and the reaction bottle(s) placed on a shaker at room temperature for 18 hrs to form resin aminated Argogel-MB-HCO resin (5). The resin was washed as follows: DMF (4x), MeOH (4x), CH₂Cl₂ (2x) and then allowed to air dry. For the amines that were hydrochloride salts, 1 eq of Et₃N was added.

Step 5. Preparation of a resin-bound thiourea (compound 7)

[0099] A 0.2 M solution of Fmoc-NCS (5 eq) in anhydrous CH_2Cl_2 was added to a bottle containing the aminated Argogel-MB-HCO resin (resin compound 5). The bottle was placed on a shaker for one hour to form resin compound 6. The resin was washed with CH_2Cl_2 (3x) and DMF (3x) and subsequently reacted with 20 % piperidine in DMF (5 eq) for one hour to produce a resin-bound thiourea (resin compound 7). The resin was then washed with DMF (3x) and MeOH (3x) and used directly in the next step.

Step 6. Preparation of the resin-bound 2-aminothiazole (compound 8)

[00100] The resin was placed a reaction bottle and the resin was swollen in MeOH and NaHCO $_3$ (10 eq) was added to the solution. The reaction bottle was placed on a shaker for 10 min, then a 0.2 M solution of 2-bromo-1,3-diphenylpropan-1-one (15 eq) (compound 3) in MeOH was added to the mixture and the bottle was placed on a shaker for 24 hrs to form resin-bound 2-aminothiazole. Then the solution mixture was decanted and the resin was washed with MeOH (4x), $CH_3CN/H_2O=1/1$ (3x), CH_3CN (2x) and was taken directly to the next step.

Step 7. Hydrolysis of the ester 8 to the acid (compound 9)

[0101] Tea-bags were placed in a bottle containing 0.5 M solution of LiOH*H $_2$ O (100 eq) in a mixture CH $_3$ CN/H $_2$ O=3/1. For the purpose of this example, R $_3$ in compound 8 in Fig. 1B is -(CH $_2$ -CH $_2$) $_n$ -CH $_2$ -CO $_2$ CH $_3$ where n is an integer from 0 to 8, preferably 1 to 6 and more preferably 1 to 4. The bottle was placed on a shaker for 24 hrs. Then bags were washed with CH $_3$ CN/H $_2$ O=3/1 (3x), H $_2$ O (2x), then briefly treated with 1N HCl (1 min, 1x), then washed with water (3x), CH $_3$ CN/H $_2$ O=3/1 (3x), CH $_3$ CN (2x), CH $_2$ Cl $_2$ (2x) and air-dried to form acid compound 9.

Step 8. Cleavage of 2-aminothiazole (compound 10) from resin

[0102] Tea-bags (100 mg of resin each) were placed in vials and 3 ml of 95 % aqueous TFA were added to each vial. The vials were tightly capped and placed in an oven-shaker at 50 $^{\circ}$ C for 3 hrs. The eluate was collected, combined with one

subsequent TFA wash and the solvent was removed with a Genevac. The crude material was analyzed by LCMS (Thermo Finnigan LCQ-classic) and if necessary purified by HPLC. The product from was a 2-aminothiazole (compound 10).

EXAMPLE 3

[0103] Experimental procedure for the preparation of the compounds according to the invention utilizing Wang resin. The synthesis scheme is shown in Fig. 2.

Step 1. Immobilization of a carboxylic acid (compound 11) on Bromo-Wang resin

[0104] Bromo-Wang resin (100 mg each tea-bag, 1.00 mmol/g substition) was swollen in anhydrous DMF. A carboxylic acid (compound 11) (5 eq.) was added, followed by potassium iodide (5 eq) and cesium carbonate (5 eq), and the bottle(s) placed in an oven-shaker at 80 °C for 3 days. Then the reaction mixture was decanted and the resin was washed as follows: DMF (4x), MeOH (4x), CH_2CI_2 (2x) and then allowed to air dry to form resin compound 12.

Step 2. Preparation of the resin-bound thioureas (compound 15)

[0105] Resin compound 12 was placed in a bottle and the resin compound was swollen in 20 % piperidine in DMF (5 eq). Tea-bags were placed in the bottle and the bottle was placed on a shaker at room temperature for one hour. Then the resin was washed with DMF (3x), MeOH (3x), CH_2Cl_2 (3x) to form resin compound 13. Then a 0.1 M solution of Fmoc-NCS (5 eq) in anhydrous CH_2Cl_2 was applied to resin compound 13 in the bottle and the bottle was placed on a shaker for one hour. The resin was then washed with CH_2Cl_2 (3x) and DMF (3x) and subsequently reacted again with 20 % piperidine in DMF (5 eq) for one hour to produce the resin-bound thiourea (compound 15). The resin was then washed with DMF (3x) and MeOH (3x) and used directly in the next step.

Step 3. Preparation of a Resin-Bound 2-Aminothiazole (compound 16)

[0106] The resin-bound thiourea was placed in a reaction bottle and the resin was swollen in MeOH and NaHCO₃ (5 eq) was added to the solution. The reaction

bottle was placed on a shaker for 10 min, then a 0.1 M solution of 2-bromo-1,3-diphenylpropan-1-one (compound 3) (5 eq) in MeOH was added to the mixture and the bottle was placed on a shaker for 2 days. Then the solution mixture was decanted and the resin was washed with MeOH (4x), CH₃CN/H₂O=1/1 (3x), CH₃CN (2x), CH₂Cl₂ (2x) and air-dried to form resin-bound 2-aminothiazole (compound 16).

Step 4. Cleavage of 2-aminothiazole to form products 10 from resin.

[0107] Tea-bags (100 mg of resin each) were placed in vials and 3 ml of 50 % TFA in CH_2Cl_2 were added to each vial. The vials were tightly capped and placed on a shaker at room temperature for 3 hrs. The eluate was collected and combined with one subsequent TFA wash. Then the solvent was removed with a Genevac. The crude material was analyzed by LCMS (Thermo Finnigan LCQ-classic), and if necessary purified by HPLC-MS, to form product 10.

EXAMPLE 4

[0108] Experimental procedure for the preparation of the compounds according to the invention utilizing solution-phase chemistry is shown in Fig. 3.

Step 1. Preparation of a Fmoc-Protected Thiourea (compound 18)

[0109] To a solution of an HCl salt of amine compound 17 (25 mmol) in 200 ml of anhydrous CH_2Cl_2 , Et_3N (25 mmol) was added which was followed by FmocNCS (25 mmol). The reaction mixture was stirred at room temperature for 2 hrs. Then solvent was removed *in vacuo* to give a crude bright-yellow solid Fmocprotected thiorurea product (compound 18) in a quatitative yield.

Step 2. Fmoc-deprotection of Thiourea Compound 18

[0110] The crude product (compound 18) from the previous step (25 mmol) was dissolved in 20 ml of DMF and 1 ml of piperidine was added. Evolution of gas (CO₂) was observed within 5 min and the reaction was slightly exothermic. The

reaction mixture was left to stir at room temperature overnight. Then 300 ml of H_2O were added and the mixture was extracted extensively with EtOAC (150 ml, 5x). The combined extracts were dried over Na_2SO_4 , solvent was removed *in vacuo* to give a yellow solid Fmoc-deprotected thiourea compound 19, to which 50 ml of CH_2Cl_2 were added and the suspension was filtered. The precipitate, compound 19, was collected on a filter, washed with a small amount of CH_2Cl_2 and dried *in vacuo*.

Step 3. Formation of a 2-aminothiazole (compound 20)

[0111] To a solution comprising 2-bromo-1,3-diphenylpropan-1-one (compound 3) (1.0 mmol) and Fmoc-deprotected thiourea (compound19) (1.0 mmol) in 15 ml of MeOH, NaHCO $_3$ (1.0 mmol) was added and the reaction mixture was heated at 50°C for 24 hours. Then solvent was removed *in vacuo*, and the residue was partitioned between EtOAc and water to form 2-aminothiazole (compound 20) which was washed with brine, dried over Na $_2$ SO $_4$. Solvent was removed *in vacuo* and the residue was purified by silica gel chromatography using Hex:EtOAc = 3:1 as eluent.

Step 4. Hydrolysis of Ester Compound 20 to Acid Product 10

[0112] To a solution of the 2-aminothiazole compound 20 (2.0 mmol) in 9 ml of CH₃CN, a solution of LiOH*H₂O in 3 ml of H₂O was added. The mixture was stirred at room temperature for 24 hours, then it was acidified to pH 3 by 1 N HCl and extracted with EtOAc (3x). The extracted product was washed with brine, dried over Na_2SO_4 and solvent was removed *in vacuo* to give the product 10 as a light-yellow solid, which was purified by HPLC.

[0113] Tables 3, 4, 5 and 6 illustrate the preferred compounds according to the invention that can mediate transactivation of FXR mediated transcription in a HEK293 reporter cell line. The data summarized in the Tables below which shows the internal molecular name used by Applicant (MOLNAME) as well as the corresponding structures of preferred compounds according to the invention. The

Table provides the EC_{50} values (EC50 AGV) as established as well as their respective average efficacy (% activity relative to CDCA control agonist).

TABLE 3

MOLNAME	MOLECULAR STRUCTURE	EC50 AVG	EFFIC AVG	EXPECTE D MASS	FOUND MASS
LN0000006316	H H H H H H H H H H H H H H H H H H H	0.20	110	458.53	459.28
LN0000006322	HO N S F F OH	1.45	115	459.53	459.29
LN0000006323	HO N S F F OH	0.44	86	452.48	453.22
LN0000006365	HN S CI OH	0.36	92	491.44	492.21

TABLE 4

						-
MOLNAME	MOLECULE STRUCTURE	EC50 AVG	EFFIC AVG	EXPECTED MASS	FOUND MASS	
LN0000006317	HO OH	1.45	115	452.48	453.22	
LN0000006328	O HO DO	4.50	116	452.58	459.26	
LN0000006329	HN S OH	0.73	105	446.53	447.23	
LN0000006339	HN S OH	2.32	86	458.43	459.28	

TABLE 5

				T	
MOLNAME	MOLECULAR STRUCTURE	EC50 AVG	EFFIC AVG	EXPECTE D MASS	FOUND MASS
LN0000006346	HN S F F	5.12	63	432.49	433.25
LN0000006347	HO N S O	3.10	66	426.54	427.26
LN0000006348	HN OH OH	2.39	109	458.53	459.29
LN0000006349	HN O OH	4.05	66	452.48	453.22

TABLE 6

MOLNAME	MOLECULAR STRUCTURE	ED50 AVG	EFFIC AVG
LN0000006316	HO S F OH	12	309
LN0000006317	HN S L	15	223
LN0000006339	HN SOH	20	323
LN0000006365	HN S CI OH	15	340

[0114] Stable HEK293FXR reporter cell lines were generated by stably transfecting with the pTRexDest30 (Invitrogen) derivatives pTRexDest30-hFXR, pTRexDest30-hRXR and the pGL2promoter (Promega) derivative pGL2promoter-FXRRE. The full length human FXR (accession U68233) and the full length human RXR α (accession P19793) were cloned into the pTRexDest30 applying the manufacturer protocols for the GatewayTM system (Invitrogen).

(undrelined). 5' - cccaGGTGAaTAACCTcggggctctgtccctccaatcccaGGGTGAaTAACCTcggg 3' (SEQ ID NO. 5) was created from the human IBAB-P promoter (Grober et al 1999, JBC 274, pp. 29749-29754). A stable clone was selected and seeded at a density of 1x10⁴ cells per well in 96 well plates. Luciferase reporter activity was measured in triplicates from extracts of cells after incubating cells in culture medium (DMEM [Gibco-BRL] + 10% FCS [PAA laboratories]) for 16 hours (5% CO₂, 37°C) containing 0.5% DMSO (control) or 0.5% DMSO with increasing concentrations of compounds. Examples of such dose response assays in the HEK293-FXR cell line are shown in Fig. 5A for LN6348, in Fig 5B for LN6316, in Fig.5C for LN6365 and in Fig. 5D for LN6322. One can derive EC50 values for the potency in the cellular reporter assay and as an example LN6348 can be determined in such an experiment with an EC50 of 1.3 μM and an relative efficacy compared to GW4064 of 109%.

[0116] While the salient features have been illustrated and described with respect to particular embodiments, it should be readily apparent that modifications can be made within the spirit and scope of the invention, and it is, therefore, not desired to limit the invention to the exact details shown and described.

What is claimed is:

1. A compound including resolved diastereoisomers and enantiomers, and tautomers, pharmaceutical acceptable salts or solvates thereof, having the following formula (I):

$$R_3$$
 R_4 R_5 R_1 R_2 (I)

wherein:

 R_1 is C_1 to C_7 alkyl, C_1 to C_7 substituted alkyl, phenyl, substituted phenyl, alkylphenyl, substituted alkylphenyl, C_2 to C_7 heterocyclic ring, C_2 to C_7 substituted heterocyclic ring, C_5 to C_6 heteroaryl, and C_5 to C_6 substituted heteroaryl;

 R_2 is C_1 to C_7 alkyl, C_1 to C_7 substituted alkyl, phenyl, substituted phenyl, alkylphenyl, substituted alkylphenyl, C_2 to C_7 heterocyclic ring, C_2 to C_7 substituted heterocyclic ring, C_5 to C_6 heteroaryl, and C_5 to C_6 substituted heteroaryl.

 R_3 is hydrogen, C_1 to C_8 alkyl, C_1 to C_8 substituted alkyl, C_7 to C_{12} alkylphenyl, C_7 to C_{12} substituted alkylphenyl, naphthyl, substituted naphthyl, C_1 to C_8 alkanesulfonyl, benzenesulfonyl, substituted benzenesulfonyl, and C_1 to C_8 substituted acyl;

 R_4 is hydrogen, C_1 to C_8 alkyl, C_1 to C_8 substituted alkyl, C_7 to C_{12} alkylphenyl, C_7 to C_{12} substituted alkylphenyl, naphthyl, substituted naphthyl, C_1 to C_8 alkanesulfonyl, benzenesulfonyl, substituted benzenesulfonyl, C_1 to C_8 substituted acyl;

where R₃ and R₄ may be taken together with nitrogen to form a heterocycle or substituted heterocycle.

2. The compound of claim 1, wherein:

 R_1 is C_1 to C_7 alkyl, C_1 to C_7 substituted alkyl, phenyl, substituted phenyl, alkylphenyl, substituted alkylphenyl, C_2 to C_7 heterocyclic ring, C_2 to C_7 substituted heterocyclic ring, C_5 to C_6 heteroaryl, or C_5 to C_6 substituted heteroaryl;

 R_2 is C_1 to C_7 alkyl, C_1 to C_7 substituted alkyl, phenyl, substituted phenyl, alkylphenyl, substituted alkylphenyl, C_2 to C_7 heterocyclic ring, C_2 to C_7 substituted heterocyclic ring, C_5 to C_6 heteroaryl, or C_5 to C_6 substituted heteroaryl;

 R_3 is hydrogen, C_1 to C_8 alkyl, C_1 to C_8 substituted alkyl, C_7 to C_{12} alkylphenyl, C_7 to C_{12} substituted alkylphenyl, naphthyl, substituted naphthyl, C_1 to C_8 alkanesulfonyl, benzenesulfonyl, substituted benzenesulfonyl, or C_1 to C_8 substituted acyl; and

R₄ is:

3. The compound of claim 2, wherein:

 R_1 is C_1 to C_7 alkyl, C_1 to C_7 substituted alkyl, phenyl, substituted phenyl, alkylphenyl, substituted alkylphenyl, C_2 to C_7 heterocyclic ring, C_2 to C_7 substituted heterocyclic ring, C_5 to C_6 heteroaryl, or C_5 to C_6 substituted heteroaryl;

 R_2 is C_1 to C_7 alkyl, C_1 to C_7 substituted alkyl, phenyl, substituted phenyl, alkylphenyl, substituted alkylphenyl, C_2 to C_7 heterocyclic ring, C_2 to C_7 substituted heterocyclic ring, C_5 to C_6 heteroaryl, or C_5 to C_6 substituted heteroaryl;

 R_3 is hydrogen, C_1 to C_8 alkyl, C_1 to C_8 substituted alkyl, C_7 to C_{12} alkylphenyl, C_7 to C_{12} substituted alkylphenyl, naphthyl, substituted naphthyl, C_1 to C_8

alkanesulfonyl, benzenesulfonyl, substituted benzenesulfonyl, or C_1 to C_8 substituted acyl; and

R₄ is:

4. The compound according to claim 1, wherein the compound is:

5. The compound according to claim 1, wherein the compound is:

6. The compound according to claim 1, wherein the compound is:

7. The compound according to claim 1, wherein the compound is:

8. The compound according to claim 1, wherein the compound is:

9. The compound according to claim 1, wherein the compound has the following formula:

10. The compound according to claim 1, wherein the compound has the following formula:

12. The compound according to claim 1, wherein the compound has the following formula:

14. The compound according to claim 1, wherein the compound has the following formula:

16. The compound according to claim 1, wherein the compound has the following formula:

18. A compound of any one of claims 1 to 17 wherein said compound is capable of binding the NR1H4 receptor protein or a portion thereof according to SEQ ID NO. 1 or a mammalian homologue thereof.

- 19. A therapeutic composition comprising at least one compound according to any one of claims 1 to 17 in admixture with a pharmaceutically acceptable carrier, adjuvant or vehicle.
- 20. A method for prevention or treatment of a NR1H4 receptor protein- or NR1H4 receptor protein homologue-mediated disease or condition in a mammal comprising administration of a therapeutically effective amount of a compound or combination of compounds according to any one of claims 1 to 17 wherein the prevention or treatment is directly or indirectly accomplished through the binding of the compound to the NR1H4 receptor protein or to the NR1H4 receptor protein homologue.
- 21. The method for prevention or treatment of a NR1H4 receptor protein mediated disease or condition according to claim 20 wherein the mammal is a human.
- 22. A method for regulating bile transport in a mammal which comprises activating the NR1H4 receptor with a therapeutically effective amount of a compound according to any one claims 1 to 17.
- 23. A method of treating a disease which is affected by cholesterol, triglyceride, or bile acid levels comprising administering to a mammal in need of such treatment a therapeutically effective amount of a compound or combination of compounds according to any one of claims 1 to 17.
- 24. A method of treating atherosclerosis, gallstone disease, lipid disorders, obesity or a cardiovascular disorder comprising administering to a mammal in need of such treatment a therapeutically effective amount of a compound or combination of compounds according to any one of claims 1 to 17.

25. A method of blocking fatty acid absorption in the intestine of a mammal in need of such blocking comprising administering to the mammal a therapeutically effective amount of a compound or combination of compounds according to any one claims 1 to 17.

- 26. The method of claim 25, wherein the method is for treating obesity in humans.
- 27. A method of modulating a gene whose expression is regulated by the NR1H4 receptor in a mammal comprising administering to the mammal a therapeutically effective amount of a compound or combination of compounds according any one of claims 1 to 17.
- 28. The method according to claim 27 wherein the expression of the gene encoding cholesterol 7-alpha-hydroxylase is lowered.
- 29. The method according to claim 27 wherein the expression of the gene encoding organic anion protein-1 is lowered.
- 30. A method according to claim 27 wherein the expression of the gene encoding cholesterol 7-alpha-hydroxylase and organic anion protein-1 are lowered.
- 31. A method according to claim 27 wherein the expression of the gene encoding intestinal bile acid binding protein is enhanced.
- 32. A method according to claim 27 wherein the expression the activity of the canicular bile salt excretion pump is enhanced.
- 33. A method according to claim 27 wherein the expression of the gene encoding intestinal bile acid binding protein and the activity of the canicular bile salt excretion pump are enhanced.

FIG. 1A

ArgGel-MB-CHO resin

8

CO₂Me
$$\begin{array}{c} CO_2 \text{Me} \\ \\ N \\ S \\ \end{array}$$

$$\begin{array}{c} CO_2 \text{H} \\ \\ CH_3 \text{CN/H}_2 \text{O} = 3/1 \\ \end{array}$$

$$\begin{array}{c} CO_2 \text{H} \\ \\ N \\ \end{array}$$

$$\begin{array}{c} CO_2 \text{H} \\ \\ N \\ \end{array}$$

8

FIG. 2

Bromo-Wang resin

12

FIG. 3

FIG. 4A

SEQ. ID NO. 1

MGSKMNLIEH	SHLPTTDEFS	FSENLFGVLT	EQVAGPLGQN	LEVEPYSQYS	NVQFPQVQPQ	60
ISSSSYYSNL	GFYPQQPEEW	YSPGIYELRR	MPAETLYQGE	TEVAEMPVTK	KPRMGASAGR	120
IKGDELCVVC	GDRASGYHYN	ALTCEGCKGF	FRRSITKNAV	YKCKNGGNCV	MDMYMRRKCQ	180
ECRLRKCKEM	GMLAECMYTG	LLTEIQCKSK	RLRKNVKQHA	DQTVNEDSEG	RDLRQVTSTT	240
KSCREKTELT	PDQQTLLHFI	MDSYNKQRMP	QEITNKILKE	EFSAEENFLI	LTEMATNHVQ	300
VLVEFTKKLP	GFQTLDHEDQ	IALLKGSAVE	AMFLRSAEIF	NKKLPSGHSD	LLEERIRNSG	360
ISDEYITPMF	SFYKSIGELK	MTQEEYALLT	AIVILSPDRQ	YIKDREAVEK	LQEPLLDVLQ	420
KLCKIHOPEN	POHFACLLGR	LTELRTFNHH	HAEMLMSWRV	NDHKFTPLLC	EIWDVQ	476

WO 03/015777 PCT/US02/25438 FIG. 4B

SEQ. ID NO. 2

1431

atgggatcaa aaatgaatct cattgaacat tcccatttac ctaccacaga tgaattttct 60 ttttctgaaa atttatttgg tgttttaaca gaacaagtgg caggtcctct gggacagaac 120 180 ctggaagtgg aaccatactc gcaatacagc aatgttcagt ttccccaagt tcaaccacag atttcctcgt catcctatta ttccaacctg ggtttctacc cccagcagcc tgaagagtgg 240 tactctcctg gaatatatga actcaggcgt atgccagctg agactctcta ccagggagaa 300 actgaggtag cagagatgcc tgtaacaaag aagccccgca tgggcgcgtc agcagggagg 360 atcaaagggg atgagctgtg tgttgtttgt ggagacagag cctctggata ccactataat 420 gcactgacct gtgaggggtg taaaggtttc ttcaggagaa gcattaccaa aaacgctgtg 480 540 tacaaqtqta aaaacqqqqq caactqtqtq atqqatatqt acatqcqaaq aaaqtqtcaa gagtgtcgac taaggaaatg caaagagatg ggaatgttgg ctgaatgtat gtatacaggc 600 ttgttaactg aaattcagtg taaatctaag cgactgagaa aaaatgtgaa gcagcatgca 660 720 gatcagaccg tgaatgaaga cagtgaaggt cgtgacttgc gacaagtgac ctcgacaaca aaqtcatqca qqqaqaaaac tqaactcacc ccaqatcaac aqactcttct acattttatt 780 atggattcat ataacaaaca gaggatgcct caggaaataa caaataaaat tttaaaagaa 840 900 gaattcagtg cagaagaaaa ttttctcatt ttgacggaaa tggcaaccaa tcatgtacag gttcttgtag aattcacaaa aaagctacca ggatttcaga ctttggacca tgaagaccag 960 attgctttgc tgaaagggtc tgcggttgaa gctatgttcc ttcgttcagc tgagattttc 1020 aataagaaac ttccgtctgg gcattctgac ctattggaag aaagaattcg aaatagtggt 1080 atctctgatg aatatataac acctatgttt agtttttata aaagtattgg ggaactgaaa 1140 atgactcaag aggagtatgc tctgcttaca gcaattgtta tcctgtctcc agatagacaa 1200 tacataaagg atagagaggc agtagagaag cttcaggagc cacttcttga tgtgctacaa 1260 aagttgtgta agattcacca gcctgaaaat cctcaacact ttgcctgtct cctgggtcgc 1320 ctgactgaat tacggacatt caatcatcac cacgctgaga tgctgatgtc atggagagta 1380

aacgaccaca agtttacccc acttctctgt gaaatctggg acgtgcagtg a

FIG. 4C

SEQ. ID NO. 3

MLVKPLPDSE	EEGHDNQEAH	QKYETMQCFA	VSQPKSIKEE	GEDLQSCLIC	VARRVPMKER	60
PVLPSSESFT	TRQDLQGKIT	SLDTSTMRAA	MKPGWEDLVR	RCIQKFHAQH	EGESVSYAKR	120
HHHEVLRQGL	AFSQIYRFSL	SDGTLVAAQT	KSKLIRSQTT	NEPQLVISLH	MLHREQNVCV	180
MNPDLTGQTM	GKPLNPISSN	SPAHQALCSG	NPGQDMTLSS	NINFPINGPK	EQMGMPMGRF	240
GGSGGMNHVS	GMQATTPQGS	NYALKMNSPS	QSSPGMNPGQ	PTSMLSPRHR	MSPGVAGSPR	300
IPPSQFSPAG	SLHSPVGVCS	STGNSHSYTN	SSLNALQALS	EGHGVSLGSS	LASPDLKMGN	360
LQNSPVNMNP	PPLSKMGSLD	SKDCFGLYGE	PSEGTTGQAE	SSCHPGEQKE	TNDPNLPPAV	420
SSERADGQSR	LHDSKGQTKL	LQLLTTKSDQ	MEPSPLASSL	SDTNKDSTGS	LPGSGSTHGT	480
SLKEKHKILH	RLLQDSSSPV	DLAKLTAEAT	GKDLSQESSS	TAPGSEVTIK	QEPVSPKKKE	540
NALLRYLLDK	DDTKDIGLPE	ITPKLERLDS	KTDPASNTKL	IAMKTEKEEM	SFEPGDQPGS	600
ELDNLEEILD	DLQNSQLPQL	FPDTRPGAPA	GSVDKQAIIN	DLMQLTAENS	PVTPVGAQKT	660
ALRISQSTFN	NPRPGQLGRL	LPNQNLPLDI	TLQSPTGAGP	FPPIRNSSPY	SVIPQPGMMG	720
NQGMIGNQGN	LGNSSTGMIG	NSASRPTMPS	GEWAPQSSAV	RVTCAATTSA	MNRPVQGGMI	780
RNPAASIPMR	PSSQPGQRQT	LQSQVMNIGP	SELEMNMGGP	QYSQQQAPPN	QTAPWPESIL	840
PIDQASFASQ	NRQPFGSSPD	DLLCPHPAAE	SPSDEGALLD	QLYLALRNFD	GLEEIDRALG	900
IPELVSQSQA	VDPEQFSSQD	SNIMLEQKAP	VFPQQYASQA	QMAQGSYSPM	QDPNFHTMGQ	960
RPSYATLRMQ	PRPGLRPTGL	VQNQPNQLRL	QLQHRLQAQQ	NRQPLMNQIS	NVSNVNLTLR	1020
PGVPTQAPIN	AQMLAQRQRE	ILNQHLRQRQ	MHQQQQVQQR	TLMMRGQGLN	MTPSMVAPSG	1080
MPATMSNPRI	PQANAQQFPF	PPNYGISQQP	DPGFTGATTP	QSPLMSPRMA	HTQSPMMQQS	1140
QANPAYQAPS	DINGWAQGNM	GGNSMFSQQS	PPHFGQQANT	SMYSNNMNIN	VSMATNTGGM	1200
SSMNQMTGQI	SMTSVTSVPT	SGLSSMGPEQ	VNDPALRGGN	LFPNQLPGMD	MIKQEGDTTR	1260
KYC						1263

WO 03/015777 PCT/US02/25438 FIG. 4D

SEQ. ID NO. 4

1 geogeocea gectegota cagettego gegagagte agegagag geagaegeg geagaegeg totagegae 121 tetetgaact tetetagatt tegetacaa etatagatt 121 tetetgaact tetacagae cagttgetga tatgtgetta agatgagtg gatgggaga 181 aatacetetg acceetecag geagagaea agaaagega aggaatgtee tgaccaacti 131 tetetgaactg acceetecag geagagaea agaaagega aggaatgtee tgaccaacti 132 ggaccaacge ceaaaggaa cactgaaaa etatacgtg aacaggaaa taaatatatatatatatagaacaactggaaagagaagaagagaagaagaagagaagaagagaaga	61 cctgacggcg tgaccgaccc gagccgattt ctcttggatt tggctacaca cc 121 ttctgcactg tttacaggca cagttgctga tatgtgttca agatgagtgg g 181 aatacctctg acccctccag ggcagagaca agaaagcgca aggaatgtcc t 241 ggacccagcc ccaaaaggaa cactgaaaaa cgtaatcgtg aacaggaaaa t 301 gaagaacttg cagagttgat ttttgcaaat tttaatgata tagacaactt t 361 cctgacaaat gtgcaatctt aaaagaaact gtgaagcaaa ttcgtcagat c 421 gagaaagcag cagctgccaa catagatgaa gtgcagaagt cagatgtate c 481 cagggtgtca tcgacaagga tgcgctgggg cctatgatgc ttgaggccct t 541 ttctttgtag tgaacctgga aggcaacgtt gtgtttgtg cagagaatgt g 601 ctaaggtata accaagaaga gctgatgaac aaaagtgtat atagcatctt g 661 gaccacacgg aatttgtcaa aaacctgctg ccaaagtcta taggtaaatg g 721 gtctggcgaa cctccgaggc ggaacagcca taccttcaat tgtcggatge t 781 tttacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga a 841 tatgcagtgc ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag a 901 gtcctgcttg atttgcgtgg caagaaggt tcccatgaag gaaagaccag t 961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcacgtctc t	cagccggca
181 aatacetetg acceetecag ggeagagaca agaaagegea aggaatgtee tgaccaacte 241 ggaccacage ceaaaaggaa cactgaaaaa cgtaategg acaggaate taacteaat 301 gaagaacttg cagagttgat ttttgcaaat tttaatgata tagacaactt taacteaat 361 ectgacaaat gtgeaatett aaaagaaact gggaageaa teegteagag caaagagaa catggaagaa teegteagag catagatgaa ggaaagaga cagagagagaa teegteagaga teegtegggg ceatagatge tegaggeect tegategggt tttttettetgag tgaacetega aggeaatgaa gtgeagaaga cagagagate cetetacaggg ttttttettetgag tgaacetega aggeaatgaa gtgeagagat taggaagatg gagaatgt geacaaggat teetacaggga getgatgaa aaaagtgata taagacatet geatgttgg 661 gaccacacgg aatttgteaa aaacetgeetg ceaaaggaa getgataaatg ggggatette 721 gtetggega cetcegagge ggaacagee tacetteaat tgetggatge tggtaaaag 781 tttacetgat teagaagagg agggteatga taaceaggaa geteateaga gagagaagga gggteatga tataceaggaa geteateaga aatatgaaaa 841 tatgcagtge ttegetgte etcaaceaaa gecaatgaaa gaaagaaggag aggatetge 901 gteetgetgt atttegetgg caagaagagt teccatgaag gaaagaacgag tteteceedga gteeteega tttaceace gecaaggaag teceatgaag gaaagacgag tteteceedga gteeteega attacggaga gaagaactgt gteetatgat aagagaaga gtateega 1141 actgagacaa gaactgagaa gacaactgat gteetatgat aagagaaga gtateegaga gecaatgaa geaaactcat cegtteteag actactaaatg aaceteaace 1201 tgttgetgea caaaaggaag agcaactga teetatget aggacactga accetagaag gacaactga teetagaaga accetagaagaga aggacatga geetatgaagaa accetaace 12201 tgttgetgea caaaggaaga agcacatgaa teeaatgage tgtgtgatga atceggate 1321 gactggacaa acgatggaga agcactga teeaaggaa aggacatga geetatgagaa accetaace 1331 gactggacaa acgatgggaaga aggacactga teeaagaaga ggaaatttte 12201 tgttgetgea agtgggaace caggacacagaa cactaceaca ggagaatattte 12201 tgttteeggga agtggaagaa gaagacagaa cactacacaga ggaagaatttte teetaacaga gcaagaagaa gaagacagaa cactacacaga ggaagaacaa acgacagaa gaagaacaa acaagaagaa gaagaacaa acaagaagaa gaagaacaa acaagaagaa gaagaacaa acaacagaaga gaagaacaa acaagaagaa gaagaacaa acaagaagaa gaagaacaa acaacaacaa gaagaacaa gaagaacaa acaacagaagaa gaagaacaa acaacagaagaa gaagaacaacaa acaacagaagaa gaagaacaa acaacagaagaa gaagaacaa acaaga	181 aatacctctg acccctccag ggcagagaca agaaagcgca aggaatgtcc to 241 ggacccagcc ccaaaaggaa cactgaaaaa cgtaatcgtg aacaggaaaa to 301 gaagaacttg cagagttgat tittgcaaat titaatgata tagacaacti to 361 cctgacaaat gtgcaatcti aaaagaaact gtgaagcaaa ticgtcagat codaggaggaaagcag cagctgccaa catagatgaa gtgcagaagt cagatgtate codaggagtgtca tcgacaagga tgcgctgggg cctatgatgc titgaggccct to 541 tictitgtag tgaacctgga aggcaacgit gtgtitgtgi cagagaatgi g601 ctaaggtata accaagaaga gctgatgaac aaaagtgtat atagcatcti g661 gaccacacgg aatitgtcaa aaacctgctg ccaaagtcta taggtaaatgi g721 gtctggcgaa cctccgaggc ggaacagcca taccticaat tgtcggatge t781 titacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga aaac 841 tatgcagtgc titcgctgtc ctcaaccaaa gtccatcaaa gaagaaggag aggi gtccatcaaa gaagaaggag aggi gtccatcaaa gaagaaggag aaggi atcacgaaagt titactactc gccaggatct ccaaggcaag atcacgtctc titaccactagaa gcaccatgaa gcacacagga gaagaccag titaccactaga aaccaggctg ggaggacctg gtaagaaggt gcaagaaggt gaagaaggt gaagaaggt gaagaaggt gaagaaggt gaagaaggt ggaggacctg gtaagaaggt gaagaaggt gaagaaggt ggaagaacgg gaagaaggt gaagaaggt ggaggacctg gtaagaaggt gaagaaggt gaagaaggt ggaagaacgt ggaagaacgg gaagaaggt gaagaaggt gaagaaggt ggaagaacgt ggaagaacgt ggaagaacgt ggaagaaggt gaagaaggt ggaagaacgt ggaagaaggt ggaagaacgt ggaagaaggt ggaagaacgt ggaagaacgt ggaagaaggt ggaagaaggt ggaagaaggt ggaagaaggt ggaagaaggt ggaagaacgt ggaagaaggt ggaagaagaggt ggaagaagaaggt ggaagaaggt ggaagaaggt gaagaaggt ggaagaaggt ggaagaagaggt ggaagaagaggt ggaagaagaaggt ggaagaagaaggt ggaagaagaggagaagaagaggagaagaagagagag	
181 aatacetetg acceetecag ggeagagaca agaaagegea aggaatgtee tgaccaacte 241 ggaccacage ceaaaaggaa cactgaaaaa cgtaategg acaggaate taacteaat 301 gaagaacttg cagagttgat ttttgcaaat tttaatgata tagacaactt taacteaat 361 ectgacaaat gtgeaatett aaaagaaact gggaageaa teegteagag caaagagaa catggaagaa teegteagag catagatgaa ggaaagaga cagagagagaa teegteagaga teegtegggg ceatagatge tegaggeect tegategggt tttttettetgag tgaacetega aggeaatgaa gtgeagaaga cagagagate cetetacaggg ttttttettetgag tgaacetega aggeaatgaa gtgeagagat taggaagatg gagaatgt geacaaggat teetacaggga getgatgaa aaaagtgata taagacatet geatgttgg 661 gaccacacgg aatttgteaa aaacetgeetg ceaaaggaa getgataaatg ggggatette 721 gtetggega cetcegagge ggaacagee tacetteaat tgetggatge tggtaaaag 781 tttacetgat teagaagagg agggteatga taaceaggaa geteateaga gagagaagga gggteatga tataceaggaa geteateaga aatatgaaaa 841 tatgcagtge ttegetgte etcaaceaaa gecaatgaaa gaaagaaggag aggatetge 901 gteetgetgt atttegetgg caagaagagt teccatgaag gaaagaacgag tteteceedga gteeteega tttaceace gecaaggaag teceatgaag gaaagacgag tteteceedga gteeteega attacggaga gaagaactgt gteetatgat aagagaaga gtateega 1141 actgagacaa gaactgagaa gacaactgat gteetatgat aagagaaga gtateegaga gecaatgaa geaaactcat cegtteteag actactaaatg aaceteaace 1201 tgttgetgea caaaaggaag agcaactga teetatget aggacactga accetagaag gacaactga teetagaaga accetagaagaga aggacatga geetatgaagaa accetaace 12201 tgttgetgea caaaggaaga agcacatgaa teeaatgage tgtgtgatga atceggate 1321 gactggacaa acgatggaga agcactga teeaaggaa aggacatga geetatgagaa accetaace 1331 gactggacaa acgatgggaaga aggacactga teeaagaaga ggaaatttte 12201 tgttgetgea agtgggaace caggacacagaa cactaceaca ggagaatattte 12201 tgttteeggga agtggaagaa gaagacagaa cactacacaga ggaagaatttte teetaacaga gcaagaagaa gaagacagaa cactacacaga ggaagaacaa acgacagaa gaagaacaa acaagaagaa gaagaacaa acaagaagaa gaagaacaa acaagaagaa gaagaacaa acaacagaaga gaagaacaa acaagaagaa gaagaacaa acaagaagaa gaagaacaa acaacaacaa gaagaacaa gaagaacaa acaacagaagaa gaagaacaa acaacagaagaa gaagaacaacaa acaacagaagaa gaagaacaa acaacagaagaa gaagaacaa acaaga	181 aatacctctg acccctccag ggcagagaca agaaagcgca aggaatgtcc to 241 ggacccagcc ccaaaaggaa cactgaaaaa cgtaatcgtg aacaggaaaa to 301 gaagaacttg cagagttgat tittgcaaat titaatgata tagacaacti to 361 cctgacaaat gtgcaatcti aaaagaaact gtgaagcaaa ticgtcagat codaggaggaaagcag cagctgccaa catagatgaa gtgcagaagt cagatgtate codaggagtgtca tcgacaagga tgcgctgggg cctatgatgc titgaggccct to 541 tictitgtag tgaacctgga aggcaacgit gtgtitgtgi cagagaatgi g601 ctaaggtata accaagaaga gctgatgaac aaaagtgtat atagcatcti g661 gaccacacgg aatitgtcaa aaacctgctg ccaaagtcta taggtaaatgi g721 gtctggcgaa cctccgaggc ggaacagcca taccticaat tgtcggatge t781 titacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga aaac 841 tatgcagtgc titcgctgtc ctcaaccaaa gtccatcaaa gaagaaggag aggi gtccatcaaa gaagaaggag aggi gtccatcaaa gaagaaggag aaggi atcacgaaagt titactactc gccaggatct ccaaggcaag atcacgtctc titaccactagaa gcaccatgaa gcacacagga gaagaccag titaccactaga aaccaggctg ggaggacctg gtaagaaggt gcaagaaggt gaagaaggt gaagaaggt gaagaaggt gaagaaggt gaagaaggt ggaggacctg gtaagaaggt gaagaaggt gaagaaggt ggaagaacgg gaagaaggt gaagaaggt ggaggacctg gtaagaaggt gaagaaggt gaagaaggt ggaagaacgt ggaagaacgg gaagaaggt gaagaaggt gaagaaggt ggaagaacgt ggaagaacgt ggaagaacgt ggaagaaggt gaagaaggt ggaagaacgt ggaagaaggt ggaagaacgt ggaagaaggt ggaagaacgt ggaagaacgt ggaagaaggt ggaagaaggt ggaagaaggt ggaagaaggt ggaagaaggt ggaagaacgt ggaagaaggt ggaagaagaggt ggaagaagaaggt ggaagaaggt ggaagaaggt gaagaaggt ggaagaaggt ggaagaagaggt ggaagaagaggt ggaagaagaaggt ggaagaagaaggt ggaagaagaggagaagaagaggagaagaagagagag	gatgggagaa
gaacccagcc ccaaaaggaa cactgaaaaa cgtaatcgtg aacaggaaaa taaatatata 301 gaagaacttg cagagttgat titttgcaaat tittaatgata tagacaactt taacttcaa. 301 cctgacaaat gtgcaatctt aaaagaaact gtgaagcaaa ticgtcagat cagagaaca. 421 gagaaagcag cagctgcaa catagatgaa gtgcagaagt cagatgtatc cctcacagg. 421 gagaaagcag cagctgcaa catagatgaa gtgcagaagt cagatgtatc cctcacagg. 421 gagaaagcag cagctgcaa acatagatgaa gtgcagaagt cagatgtatc ctctacagg. 421 dagggtgtca tegacaagga gegacaggt gtgtttgtg cagagaatgt gacacagta. 421 gagaaagaca acacaggaa gctgatgaac aaaagtgtat taggcactt gatgggtg. 541 tictttgtag tgaacctgg aggcaacgca taccttcaat tgtcggatgg ggaacagta. 661 gaccacacgg aatttgtcaa aaacctgctg ccaaagtcat taggtaaatg ggggatctt. 721 gtctggcgaa cctccgaggc ggaacagcca taccttcaat tgtcggatgg tggaaaac. 731 tittacctgat tragagagg agggtcatga taaccaggaa gctcatcaga aatatgaaa. 841 tatgcagtgc titcgctgtc ctcaaccaaa gtccatcaaa gaagaaggag aagatttgc. 961 atcagaaagt tittactactc gccaggatct ccaaggcaag acacaggcag titctccct. 961 atcagaaagt tittactactc gccaggatct gtgcatatga gaagaaggg gttcatcaga. 1021 caccatggag gagaccatga aacaaggcag ggagaatctg gtccataga gaagaaggag tittcaga. 1021 caccatggag gagaccatga agaaaccag gtcatcagaa gaagaaggag tittcagaa. 1021 tgttgctgca caaacgaaga gagaatctgt gtccataga gaagaaggag tactagaa. 1141 actagaacaa ggattggga tcagtcagaa ccatacagac tcaacagac 1261 tgtaatatct ttacatatgc ttcacagaga gcagaatgg tgtgtgatga atccgaacac 1261 tgtaatactc ttacatatgc ttcacagaga gcagaatgg tgtgtgatga atccgaacac 1261 tgtaatactc ttacatatgc ttcacagaga gccaatgaccc agtagcacac accacacacacacacacacacacacacacaca	241 ggacccagcc ccaaaaggaa cactgaaaaa cgtaatcgtg aacaggaaaa t 301 gaagaacttg cagagttgat ttttgcaaat tttaatgata tagacaactt t 361 cctgacaaat gtgcaatctt aaaagaaact gtgaagcaaa ttcgtcagat c 421 gagaaagcag cagctgccaa catagatgaa gtgcagaagt cagatgtatc c 481 cagggtgtca tcgacaagga tgcgctgggg cctatgatgc ttgaggccct t 541 ttctttgtag tgaacctgga aggcaacgtt gtgtttgtgt cagagaatgt g 601 ctaaggtata accaagaaga gctgatgaac aaaagtgtat atagcatctt g 661 gaccacacgg aatttgtcaa aaacctgctg ccaaagtcta taggtaaatg g 721 gtctggcgaa cctccgaagg ggaacagca taccttcaat tgtcggatge t 781 tttacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga a 841 tatgcagtgc ttcgctgtc ctcaaccaaa gtccatcaaa gaagaaggag a 901 gtcctgcttg atttgcgtgg caagaaggt tcccatgaag gaaagaccag t accatgaaag gaaagaccag ttaccagaaag tttactactc gccaggatct ccaaggcaag atcacgtctc t 1021 caccatgaag gcagccatga aaccaggctg ggaggacctg gtaagaaggt g	
361 gaagaacttg cagagttgat tittgcaaat titaatgata tagacaactt taacticaat 361 cctgacaaat gtgcaatctt aaaagaaact gtgaagcaaa ticgtcagat caaagaaca 421 gagaaagcag cagctgccaa catagatgaa gtgcagaagt cagatgtat ctctacagg 481 cagggtgtca tcgacaagga tgcgctgggg cctatgatgc ttgaggccct tgatggggtct 661 ctaaggtata accaagaaga gtgatgaac aaaagtgtat atagcatctt gcatgtggg 661 gaccacacgg aatttgtcaa aaacctgctg ccaaagtcta taggtaatg ggaggatctt 721 gtctgggcga cctccgagge ggaacagcca taccttcaat tgtcggatgc tggtatgaac 781 tttacctgat tcagaagagg agggtatga taaccaggaa gctcatcaga aatatgaaaa 481 tatgcagtgc tctcgctgct ctcaacacaa gtccatcaaa gaagaagagg aggatttg 901 gtcctgcttg atttgcgtgg caagaaggt tccaaggaa gacaacagga agaatttg 901 gtcctgcttg atttgcgtgg caagaaggt tccaatgaa gaagaaggag agaatttg 901 gtcctgcttg atttgcgtgg caagaagat tccaatgaa gaagaaggag agaatttg 911 accaatgaga gaagcaatga aaccaggcag gagaacctg gtaagaaggt gttcaccag 1021 caccatgaga gaagcaatga aaccaggcag gagaactctg tgcatagaa atcaggcat t 1021 tgttgctgca caaacgaaga gagaatctgt gtcctatgct aagaggagat atcacagaa 1141 actgagacaa ggattggaat tcagtcaaaa ctatcgtttt tccttgttg atggagactc 11201 tgttgctgca caaacgaaga gcaaactcat ccgttctcag actacataatg accctaaca 11261 tgtaatactt ttacatatgc ttcacagaga gcagaatgt gtgtgtgatg atccaacag 11261 tgtaatact ttacatatgc ttcacagaga catgacccta agaagcacat atcaatgc 11261 tgtaatacct gtgtcaggaa aagcactga gccatggga aggattgtg gttctgggg 11201 aatgaacaac gtgtgaggaac caggtcagga catgacccta agaagcacat ataaatttc 11261 tgtaatact ttacatatgc tgcaagaaca cactcctcag ggtagtaact atgacacaa 11261 tgtaatact tgcacaagaa gaccactgaa tccaataag gagattggt gttctgggg 11201 aatgaacaac gtgtcaggaa tgcaactga gccatggga ggtggaacc agaagaacaa acgacagaa gacaccaa gacagaca	301 gaagaacttg cagagttgat ttttgcaaat tttaatgata tagacaactt t 361 cctgacaaat gtgcaatctt aaaagaaact gtgaagcaaa ttcgtcagat c 421 gagaaagcag cagctgccaa catagatgaa gtgcagaagt cagatgtate c 481 cagggtgtca tcgacaagga tgcgctgggg cctatgatgc ttgaggccct t 541 ttctttgtag tgaacctgga aggcaacgtt gtgtttgtgt cagagaatgt g 601 ctaaggtata accaagaaga gctgatgaac aaaagtgtat atagcatctt g 661 gaccacacgg aatttgtcaa aaacctgctg ccaaagtcta taggtaaatg g 721 gtctggcgaa cctccgaggc ggaacagcca taccttcaat tgtcggatgc t 781 tttacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga a 841 tatgcagtgc ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag a 901 gtcctgcttg atttgcgtgg caagaagagt tcccatgaag gaaagaccag t 961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcacgtctc t 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt g	
421 gagaaagcag cagctgcaa catagatgaa gtgcagaagt cagatgtate ctctacaggg 481 cagggtgtca tcgacaagga tgcgctgggg cctatgatge ttgaggccet tgatgggtte 541 ttctttgtag tgaacctgga aggcaacgtt gtgtttgtg cagagaatgt gacacagta1 601 ctaaggtata accaagaaga gctgatgaac aaaagtat atagcatctt gcatgttgg 661 gaccacacgg aatttgtcaa aaacctgctg ccaaaggtat atagcatctt gcatgttgg 661 gaccacacgg aatttgtcaa aaacctgctg ccaaaggtat ataggtaaatg gggggatct1 721 gtctggcgaa cctccgaggc ggaacagcac taccttcaat tgtcggatge tggtaaaac 781 tttacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga aatatgaaa 841 tatgcagtge ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag aagatttgc 901 gtcctgcttg atttgcgtgg caagaaggat tcccatgaag gaagaacgag ttcttccct. 961 atcagaaagt tttactactc gccaggatct ccaaggaag atcacgtct tggatacca; 1021 caccatgaga gcagcatga aaccaggctg ggaggacctg gtaatgaagg gtattcaga; 1141 actgagacaa ggattggcat tcagtcaaa ctatcgtttt tccttgtctg atggacacag agacactga tccattagt ttcttcctt. 1201 tgttgctgca caaacgaaga gcaaactcat ccgttctcag actactaatg aacctcaac 1221 tgttaatactt ttacatatgc ttcacaggag gcagaatgtg tgtgtgatga atccggatc 1231 gactggacaa acgatgggga agccactgaa tccaattagc tctatacgc ctgccatca 1231 gactggacaa acgatggggaa agccactgaa tccaattagc tctacacagc ctgccatca 1241 cataaatggc ccaaaaggaac aaggaccag catgacctc agtagcacta ataaattttc 1321 gactggacaa gtgggaacc caggtcagga catgacctc agtagcacta taaattttc 1441 cataaatggc ccaaaaggaac agagccctgg catgaatca aggacacca 1561 aatgaaccat gtgtcagga tgcaagcaac cactcctcag ggtagtaact atgacctca 1561 aatgaaccat gtgtcaggat tgcaagcaac cactcctcag ggtagtaact atgacctca 1561 gttttcccct gcaggaagct tgcatcccc tgtgggagt ttgcagcagc cacggaata 1741 ccatagttat accaacagct cctcaatga gcctggagc ggtggcagc cctcgaatcc caccaggag 1741 ccatagttat accaacagct cctcaatga gattgagaac gcaggagggggggggg	421 gagaaagcag cagctgccaa catagatgaa gtgcagaagt cagatgtate c 481 cagggtgtca tcgacaagga tgcgctgggg cctatgatge ttgaggccct t 541 ttctttgtag tgaacctgga aggcaacgtt gtgtttgtgt cagagaatgt g 601 ctaaggtata accaagaaga gctgatgaac aaaagtgtat atagcatctt g 661 gaccacacgg aatttgtcaa aaacctgctg ccaaagtcta taggtaaatg g 721 gtctggcgaa cctccgaggc ggaacagcca taccttcaat tgtcggatge t 781 tttacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga a 841 tatgcagtgc ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag a 901 gtcctgcttg atttgcgtgg caagaagagt tcccatgaag gaaagaccag t 961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcacgtctc t 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt g	
421 gagaaagcag cagctgcaa catagatgaa gtgcagaagt cagatgtate ctctacaggg 481 cagggtgtca tcgacaagga tgcgctgggg cctatgatge ttgaggccet tgatgggtte 541 ttctttgtag tgaacctgga aggcaacgtt gtgtttgtg cagagaatgt gacacagta1 601 ctaaggtata accaagaaga gctgatgaac aaaagtat atagcatctt gcatgttgg 661 gaccacacgg aatttgtcaa aaacctgctg ccaaaggtat atagcatctt gcatgttgg 661 gaccacacgg aatttgtcaa aaacctgctg ccaaaggtat ataggtaaatg gggggatct1 721 gtctggcgaa cctccgaggc ggaacagcac taccttcaat tgtcggatge tggtaaaac 781 tttacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga aatatgaaa 841 tatgcagtge ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag aagatttgc 901 gtcctgcttg atttgcgtgg caagaaggat tcccatgaag gaagaacgag ttcttccct. 961 atcagaaagt tttactactc gccaggatct ccaaggaag atcacgtct tggatacca; 1021 caccatgaga gcagcatga aaccaggctg ggaggacctg gtaatgaagg gtattcaga; 1141 actgagacaa ggattggcat tcagtcaaa ctatcgtttt tccttgtctg atggacacag agacactga tccattagt ttcttcctt. 1201 tgttgctgca caaacgaaga gcaaactcat ccgttctcag actactaatg aacctcaac 1221 tgttaatactt ttacatatgc ttcacaggag gcagaatgtg tgtgtgatga atccggatc 1231 gactggacaa acgatgggga agccactgaa tccaattagc tctatacgc ctgccatca 1231 gactggacaa acgatggggaa agccactgaa tccaattagc tctacacagc ctgccatca 1241 cataaatggc ccaaaaggaac aaggaccag catgacctc agtagcacta ataaattttc 1321 gactggacaa gtgggaacc caggtcagga catgacctc agtagcacta taaattttc 1441 cataaatggc ccaaaaggaac agagccctgg catgaatca aggacacca 1561 aatgaaccat gtgtcagga tgcaagcaac cactcctcag ggtagtaact atgacctca 1561 aatgaaccat gtgtcaggat tgcaagcaac cactcctcag ggtagtaact atgacctca 1561 gttttcccct gcaggaagct tgcatcccc tgtgggagt ttgcagcagc cacggaata 1741 ccatagttat accaacagct cctcaatga gcctggagc ggtggcagc cctcgaatcc caccaggag 1741 ccatagttat accaacagct cctcaatga gattgagaac gcaggagggggggggg	421 gagaaagcag cagctgccaa catagatgaa gtgcagaagt cagatgtate c 481 cagggtgtca tcgacaagga tgcgctgggg cctatgatge ttgaggccct t 541 ttctttgtag tgaacctgga aggcaacgtt gtgtttgtgt cagagaatgt g 601 ctaaggtata accaagaaga gctgatgaac aaaagtgtat atagcatctt g 661 gaccacacgg aatttgtcaa aaacctgctg ccaaagtcta taggtaaatg g 721 gtctggcgaa cctccgaggc ggaacagcca taccttcaat tgtcggatge t 781 tttacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga a 841 tatgcagtgc ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag a 901 gtcctgcttg atttgcgtgg caagaagagt tcccatgaag gaaagaccag t 961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcacgtctc t 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt g	caaagaacaa
tecttytag tgaacetgaa aggeaacgtt gtgtttgtt cagagaatgt gacacagtat 601 ctaaggtata accaagaaga getgatgaac aaaagtgtat atageatett geatgttgg 661 gaccacacgg aatttgtcaa aaacetgtgt cagaagtat taggtaaatg gaggatett. 721 getggggaa cetecagagg ggaacagca tacetteaat tggeggatg tggaaaaac 781 tttacetgat teagaagaag agggeatgta taaceaggaa getcateaga aatatgaaaa 841 tatgeagtge ttegetgtet etcaaccaaa gecaataaa gaagaaggag aaagatttge 901 gteetgetgtg tetgetggtg eagagaaggt teccatgaag gaagaaggag aagatttge 901 gteetgetgtg atttgegtgg caagaagagg teccatagaa gaagaaggag atteeteege 901 gteetgetgt atttgegtgg caagaagagg ggagacetg gtaagaagg gtatteege 901 gteetgetgt attegetgg caagaagagg ggagacetg gtaagaaggg gtatteege 901 accaagaag geagecatga aaccaggetg ggaggacetg gtaagaaggg gtatteege 1021 caccatgaga geagecatga aacaggetg ggaggacetg gtaagaaggg gtatteegg 1141 actgagacaa ggaattggeat teagteaat etcatgttt teettgteg atggeacet 1201 tgttgetgea caaacgaaga geaaacteat cegtteteag actactaatg aaceteaact 1201 tgttgetgea caaacgaaga geaaacteat cegtteteag actacataag aaceteaact 1201 tgttgatgaca acgatgggga agecactgaa tecaattage tetaacagee etgeceate 1321 gactggacaa acgatgggga agecactgaa tecaattage tetaacagge etgeceate 1321 gactggacaa acgatgggga agecactgaa tecaattage tetaacagee etgeceate 1321 gactggacaa acgatgggga agecactgaa tecaattage tetaacagee etgeceate 1321 gactggacaa gtgtgaggaace caggcaaga catgaceete agtaggaataa taaatttee 1414 cataaatgge cetacacaa geagecetgg catgaacaa ggatgtggtg gttetggggg 1501 aatgaacag eecteacaaa geagecetgg ggetggaage ggaagage aggaagacte tegaagaaga gacgaceetg ettaacaagge etgaagaagatg ggetggaage etgaagaaga gtteecaagaga getgeceagag geetggggg ggetgggggg etteaacaagga geagagage tegaagaagatg ggaagagatg tegaagaaga ggaagagatg tegaagaaga ggaagagaga ggaagagaga ggaagagaga ggaagagaga gaagaa	541 ttctttgtag tgaacctgga aggcaacgtt gtgtttgtgt cagagaatgt g 601 ctaaggtata accaagaaga gctgatgaac aaaagtgtat atagcatctt g 661 gaccacacgg aatttgtcaa aaacctgctg ccaaagtcta taggtaaatg g 721 gtctggcgaa cctccgaggc ggaacagcca taccttcaat tgtcggatgc t 781 tttacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga a 841 tatgcagtgc ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag a 901 gtcctgcttg atttgcgtgg caagaagagt tcccatgaag gaaagaccag t 961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcacgtctc t 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt g	ctctacaggg
601 ctaaggtata accaagaaga gctgatgaac aaaagtgtat atagcatctt gcatgtgg 661 gaccacacgg aatttgtcaa aaacctgctg ccaaagtcta taggtaaatg ggggatcttg 721 gtctggcgaa cctccgaggc ggaacagcca taccttcaat tgtcggatge tggtaaaacc 781 tttacctgat tcagaagaag agggtcatga taaccaaggaa gctcatcaga aatatgaaac 841 tatgcagtgc ttcgctgtct ctcaaccaaa gaccatcaaa gaagaagaga aagattgc 901 gtcctgcttg atttgcgtgg caagaagagt tcccatgaaa gaagaagaga aagattgc 961 atcagaagag tttactactc gccaggatct ccaaggcaag atcacgtctc tgggtaacac 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt gtattcaga 1021 caccatgaga gcagcatga aaccaggctg ggaggacctg gtaagaaggt gtattcaga 1021 caccatgaga gcagcatga gagaatctgt gtcctatgct aagaggagat atcatgag 1141 actgagacaa ggattggcat tcagtcaaat ctatcgtttt tccttgtctg atggcactc' 1201 tgttgctgca caaacgaaga gcaaactcat ccgttctcag actactaatg aacctcaac' 1261 tgtaatatct ttacatatgc ttcacagaga gcagaatgtg tgtgtgatga atccggatc' 1321 gactggacaa acgatgggga agccactgaa tccaattagc tctaacagcc ctgcccatc' 1381 ggccctgtge agtgggaacc caggtcagga catgacctc agtagcaata taaatttc' 1441 cataaatggc ccaaaggaac aaatgggcat gcccatgggc aggtttggtg gttctggggg 1501 aatgaacagc ccctcacaaa gcagcctgg catgaacca ggatagata ttgcagcatca 1561 aatgaacagc ccctcacaaa gcagcctgg ggtgggacg cctccagaacc cctccatgc' 1681 gttttcccct gcaggaagct tgcattccc tgtgggagc cctcagaccc cctccatgc' 1681 gttttcccct gcaggaagct tgcattccc tgtgggagc cctcagaacc cacccagca 1741 ccatagttat accaacagct ccctcaatgc acttcaagacc ggaagaaga ggaagaagagagagagagagag	601 ctaaggtata accaagaaga gctgatgaac aaaagtgtat atagcatctt g 661 gaccacacgg aatttgtcaa aaacctgctg ccaaagtcta taggtaaatg g 721 gtctggcgaa cctccgaggc ggaacagcca taccttcaat tgtcggatgc t 781 tttacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga a 841 tatgcagtgc ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag a 901 gtcctgcttg atttgcgtgg caagaagagt tcccatgaag gaaagaccag t 961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcacgtctc t 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt g	gatgggttc
661 gaccacacgg aatttgtcaa aaacctgctg ccaaagtcta taggtaaatg ggggatcttg 721 gtctggcgaa cctccgaggc ggaacagcca taccttcaat tgtcggatgc tggtaaaacc 781 tttacctgat tcagaaggag agggtcatga taacctacaaa gaagaaggag aagatttgc 781 tttacctgat tcagaaggag agggtcatga taaccacgaaa gctcatcaga aatatgaaac 841 tatgcagtgc ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag aagatttgc 901 gtcctgcttg atttgcgtgg caagaaggt tcccatagaag gaaagaaccag ttcttcccct 961 atcagaaagt tttactactc gccaggatct ccaaggaag atcacgtctc tggataccac 961 atcagaaagt tttactactc gccaggatct ccaaggaag atcacgtctc tggataccac 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt gtattcaga. 1081 gttccatgcg cagcatgaag gaaactcat ccgttctcg aagaaggagt atcatgaag 1141 actgagacaa ggattggcat tcagtcaaat ctatcgtttt tccttgtctg atggcactc 1201 tgttgctgca caaacgaaga gcaaactcat ccgttctcag actactaatg aacctcaacc 1261 tgtaaatatct ttacatatgc ttcacagaga gcagaatgtg tgtgtgatga atccggatci 1321 gactggacaa acgatgggga agccatgaa tccaattagc tctaacagcc ctgccatca 1381 ggccctgtgc agtgggaacc caggtcagga catgacctc agtagcaata taaattttc 1441 cataaatggc ccaaaggaac aaatgggcat gccatggga gggttggtgg gttttgggg 1501 aatgaaccat gtgtcaggac tgcaagcaac acatcctcaag ggaagagcca cctccaaga 1621 ttcaccaagg catcgcatga gccctggagt ggctggagg cctcgaatcc caccaagca 1621 ttcaccaagg catcgcatga gccctggagt ggctggagg cctcgaatcc caccaagca 1621 ttcaccaagg catcgcatga gccctggagt ggctggagg cctcgaatcc caccaagca 1621 ttcaccaagg catcgcatga gccctgaatg ggcaggagc cctcgaatcc caccaagca 1621 ttgacatatat ggggaacct ccccaagaa gatgggagac ggcagagggg 1621 tggactatat ggggaacct ccccaagaa gatgggagac ggcagagac caccacagaga ggcagagagac 1621 tggactatat ggggaacct ccccaagaa gatgggagac ggcagagagac gggaagagac 1621 tggactatat ggggaacct ccccaagaa gatgggaac ggcagagagac gggaagagac 1621 tggactatat ggggaacct ccccaagaa gatgggaacc ggcagagagac 1621 tggagagaaa aagaacagaa atcaccccc acctagaaa gaagagaaca aagacgaga 1621 tggagagaaa aagaacagaa gccacagaa aagagaaca gcagaagaaca 1622 tggagagaac gagaactca ccccaaaaca gcagaagaaca gcagaacaca gcagaagaa gcagaacaca caccaagaa gcacacaagaa gccacaagaa ga	661 gaccacacgg aatttgtcaa aaacctgctg ccaaagtcta taggtaaatg g 721 gtctggcgaa cctccgaggc ggaacagcca taccttcaat tgtcggatgc t 781 tttacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga a 841 tatgcagtgc ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag a 901 gtcctgcttg atttgcgtgg caagaagagt tcccatgaag gaaagaccag t 961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcacgtctc t 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt g	gacacagtat
721 gtctggcgaa cctccgaggc ggaacagcca taccttcaat tgtcggatgc tggtaaaacca 781 tttacctgat tcagaaggg agggtcatga taaccaggaa gctcatcaga aatatgaaaca 841 tatgcagtgc ttcgctgtc ctcaaccaaa gtccatcaaa gaagaaggga aagatttgc 901 gtcctgcttg atttgcgtgg caagaaggat tcccatgaag gaaagaccag ttcttccctc 961 atcagaaagt tttactactc gccaggatct ccaaggaaag atcacgtctc tggataccag 1021 caccatgaga gcagccatgaa aaccaggctg ggaggacctg gtaagaaggt gtattcagat 1141 actgagaacaa ggattggag gagaatctgt gtcctatgct aagaggaat atcatgaag 1141 actgagaacaa ggattgggaa tcagtcaat ctatcgtttt tccttytctg atggcactci 1201 tgttgctgca caaacgaaga gcaactcaat ccgttctcag actactaatg aacctcaacci 1321 gactggacaa acgatgggga agcaactgaa tccaattagc tctaacagcc ctgcccatc. 1331 ggccctgtgc agtgggaacc caggtcagga catgacccc agtagcaata taaattttc 1441 cataaatggc ccaaaggaac aaatgggcaa gccactgaa tccaattagc tctaacagcc ctgcccatc. 1341 cataaatggc ccaaaggaac aaatgggcat gccaatgggc aggttggtg gttttggg gttttggg 1501 aatgaaccat gtgtcaggaa tgcaactgaa ccatcacag ggtaggaacc cctcacagag gccatggac ggcaatggg ggtagtaact atgacacta. 1561 aatgaaccag catcgcatga gccctggagt ggctggagg cctcggatc caccacaggc 1621 ttcaccaagg catcgcatga gccctggagt ggctggagc cctcgaatcc cacccatgc. 1681 gttttcccct gcaggaagct tgcatcccc tgtgggggt tgcagcagca caccacaggg. 1801 ctcattaggg tcatcgttgg cttcaccaca gatgggaagc cctcgaatcc caccaaggc. 1801 ctcattaggg tcatcgttgg cttcaccaca gatgggaagc ttggacacca aaagacacca caccaaggc. 1801 tgaggagcaa aaggaaacaa atcaccaca cctcaagaa gatgggaagc cttggacaa aagaacacca caccaagaa gatggaagaa aagaagaa aagaacacca caccaagaa gcagaagac caccaagaa gaggaacca caccaagaa gaaggaaca caccaagaa gaagaagaa caccaagaa gaagaacac caccaagaa gaagaacac caccaagaa gcagaagaa caccaagaa gcagaacac caccaagaa gaagaagaa caccaagaa gaagaacac caccaagaa gaagaagaa cacacaagaa gcacaagaa aagaaaaa atcaccaa cctgacaa gcagaagaa gcaaacacca caccaagaa gcagaagaa aagaacacca caccaagaa gcagaagaa aagaaaaa atgaccacaa cctgacaa gaagagaac gaagaagaa gcacaagaa gaagaacac caccaagaa gaagaacac caccaagaa gaagaacac caccaagaa gaagaacac cacaagaa gaagaacac cacaagaa gaagaacac cacaagaa gaaga	721 gtctggcgaa cctccgaggc ggaacagcca taccttcaat tgtcggatgc t 781 tttacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga a 841 tatgcagtgc ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag a 901 gtcctgcttg atttgcgtgg caagaagagt tcccatgaag gaaagaccag t 961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcacgtctc t 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt g	jcatgttggg
841 tatgcagtgc ttcgctgct ctcaacaaa gtccatcaaa gagaagagga aagatttgcg 901 gtcctgcttg atttgcgtgg caagaaggt tcccatgaag gaaagagga aagatttgcg 901 gtcctgcttg atttgcgtgg caagaagggt tcccatgaag gaaagacag ttcttccctg 961 atcagaaagt tttactactc gccaggatct ccaaggaag atcacgtctc tggataccag 1021 caccatggaag gcagccatga aaccaggctg ggaggacctg gtaagaaggt gtattcagag 1081 gttccatgcg cagcatgaag gagaatctgt gtcctatgct aagaaggct gtattcagag 1141 actgagacaa ggattggcat tcagtcaaat ctatcgtttt tccttgtctg atggcactcd 1201 tgttgctgca caaacgaagg gcaaaccat ccgttctcag actactaatg aacctcaaccd 1201 tgttgactgca caaacgaaga gcaaaccat ccgttctcag actactaatg aacctcaaccd 1201 tgttgactgca caaacgaaga gcaaaccat ccgttctcag actactaatg aacctcaaccd 1201 tgtaatatct ttacatatgc ttcacagaga gcagaatgtg tgtgtgatga atccggatcd 1321 gactggacaa acgatgggga agccactgaa tccaattagc tctaacagc ctgcccatcd 1381 ggccctgtgc agtgggaacc caggtcagga catgaccctc agtagcaata taaattttcd 1441 cataaatggc ccaaaggaac aactgcagga gccatggga gggttggtg gttctggggg 1501 aatgaaccat gtgtcaggca tgcaagcaac cactcctcag ggtagtaact atgcactca 1561 aatgaacaag ccctcacaaa gcagcctgg catgaatca ggacaggcca cctccatgcd 1621 ttcaccaagg catcgcatga gccctggagt ggctggaagc cctcgaatcc cacccagtc 1681 gttttcccct gcaggaagct tgcatcccc tgtgggagt tgcagcagca caccacagac 1741 ccataagttat accaacagct ccctcaatgc acttcaggcc ctcagaagg ggcaggagg 1801 ctcattaggg tcatcgttgg cttcaccaga actcagaca gagaagagg ggaactggd 1801 ctcattaggg tcatcgttgg cttcaccaga actcagaca gagaagag ggcacgggg 1801 dcactaagtat gaccacagac accacaagac actcagaca gagaagac aagaactgg 1801 ctcattaggg tcatcgttgg cttcaccaa gatggaaca gcaggaagac ggagaactg 1921 tggactatat ggggagacc cttgaagaa acgagaagac accacaagac 181 agttaatata gatcagatgg acctcgcc cttagacaa gaagaggac gtggaagac 1921 tggacaatat gacagaaga acgacacaa accacaagac 2011 caccaaatct gatcagatgg accacaagacaa accacaacaa accacaagac 2101 caccaaatct gatcagaaga gacaccaagacaa accacaacaa accacaagac 211 agaagagaa acaagactga accacaagaa accacaaa aaacacaacaa ccccaaaacaacaa accacaagaa acaacaacaa accacaagaa acaacaacaa accacaagaa acaacaacaa accacaagaa	781 tttacctgat tcagaagagg agggtcatga taaccaggaa gctcatcaga a 841 tatgcagtgc ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag a 901 gtcctgcttg atttgcgtgg caagaagagt tcccatgaag gaaagaccag t 961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcacgtctc t 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt g	gggatcttg
841 tatgcagtgc ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag aagatttgc 901 gtcctgcttg atttgcgtgg caagaaggt tcccatgaag gaaagaccag ttcttccctc 961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcagtctc tggataccag 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt gtattcaga 1081 gttccatgcg cagcatgaag gagaatctgt gtcctatgct aagaggcatc atcatgaag 1141 actgagacaa ggattggcat tcagtcaaat ctatcgtttt tccttgtctg atggcatct 1201 tgttgctgca caaacgaaga gcaaactcat ccgttctcag actactaatg aacctcaact 1261 tgtaatatct ttacatatgc ttcacaggag gcagaatgtg tgtgtgatga atccggatc1 1321 gactggacaa acgatgggga agcactgaa tccaattagc tctaacagcc ctgcccatc 1381 ggccctgtgc agtgggaacc caggtcagga catgaccctc agtagcaata taaattttc 1441 cataaatggc ccaaaggaac aaatggacga catgaccctc agtagcaata taaattttc 1441 cataaatggc ccaaaggaac acaggacagac cactgcaa ggctgggg 1501 aatgaaccat gtgtcaggca tgcaaggaa catgaaccac ggacagcca cctccatgc 1621 ttcaccaagg catcgcatga gccctggagt ggctggcagc cctcgaatcc cacccagtc 1621 ttcaccaagg catcgcatga gccctggagt ggctggcagc cctcgaatcc cacccagtc 1631 gttttcccct gcaggaagct tgcattccc tgtgggagtt tgcagcagca caggaaatat 1741 ccatagttat accaacagct ccctcaatgc acttcaggc ctcaggagg ggcacgggg 1801 ctcattaggg tcatcgttgg cttcaccag acttcaggc ctcaggagg ggcacgggg 1801 ctcattaggg tcatcgttgg cttcaccag acttgagaac tgggaagcc ctcagagag agaacccc 1861 agttaatatg aatcctcccc cactcaagaa actgggaagc ttggactcaa aagaccgt 1821 tggagctatat ggggagcct ctgaaggaa cctagagaac caggagagca ggcacgggg 1801 ctcattaggg catcggtgg ctcacagaa gatgggaagc tttggactcaa aagactgtt 1921 tggactatat ggggagcct ctgaaggaa cctagagaac cacggagagca gctgcaacc 1861 agtaatatg aatcctcccc cactcaagaa gatgggaagc tttggactcaa aagactgt 1821 tggagagcaa aggaaaccaa atgacccaa cctggccccg gccgtgagc gtgagagag 1801 ctcattaggg agcacgagaa accagaa aggagaacca aactccgc ctcaggaga gcagagaga 1801 ctcattagga agcacagaa accagaa aggagaacca aactccca agaccagaa gcacagaacca agacagaaca agacagaaca accagaacaa aggagaacca aaccagaa agacagaacaa agacagaacaa agacagaacaa caccagaa agacagaacaa agacagaacaa agacagaacaa agacagaacaa agaacaacaa agaacaacaa	841 tatgcagtgc ttcgctgtct ctcaaccaaa gtccatcaaa gaagaaggag a 901 gtcctgcttg atttgcgtgg caagaagagt tcccatgaag gaaagaccag t 961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcacgtctc t 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt g	:ggtaaaacc
901 gtcctgcttg atttgcgtgg caagaaggt tcccatgaag gaaagaccag ttcttccct. 961 atcagaaagt tttactactc gccaggatct caaaggcaag atcacgtct tggataccag. 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt gtattcaga. 1081 gttccatgcg cagcatgaag gagaatctgt gtcctatgct aagaggcatc atcatgaagg. 1141 actgagacaa ggattggcat tcagtcaaat ctatcgtttt tccttgtctg atggcactc. 1201 tgttgctgca caaacgaaga gcaaactcat ccgttctcag actactaatg aacctcaact. 1201 tgttaatatct ttaccatatgc ttcaccagaga gcagaatgtg tgtgtgatga atccggatc. 1321 gactggacaa acgatgggaa agccactgaa tccaattagc tctaacagcc ctgcccatc. 1331 ggccctgtgc agtgggaacc caggtcagga ctgaccctc agtagcaata taaattttc. 1441 cataaatggc ccaaaggaac aaatggacat gccaatggc aggtttggtg gtttcggggg. 1501 aatgaaccat gtgtcaggac tgcaagaac cactcctcag ggtagtaact atgcactca. 1561 aatgaacagc cctcacaaag gcagcactga cactcaca ggacagcca cctccatgc. 1621 ttcaccaagg catcgcatga gccctggagt ggctggcagc cctcgaatcc cacccagtc. 1681 gttttcccct gcaggaagct tgcattcccc tgtgggagtt tgcagcagca caggaaatat. 1741 ccatagttat accaacagct ccctcaatgc actcaaggca ggcaggggg. 1801 ctcattaggg tcatcgttgg cttcaccaga acttcaaggc ctcaggagg ggcacgggg. 1801 ctcattaggg tcatcgttgg cttcaccaga acttcaggcc ctcagagag ggcacgggg. 1801 ctcattaggg tcatcgttgg cttcaccaga acttcaggcc ctcagagag ggcacgggg. 1801 ctcattaggg tcatcgttgg cttcaccaga acttcaggcc ctcagcagag ggcacgggg. 1801 ctcattaggg tcatcgttgg cttcaccaga acttgagaac gcagaagaca aagacatgtt. 1921 tggactatat ggggaacct ctgaaggac actgagaac actggaagac aagagagac gctgcatc. 1981 tggagagcaa aagagaacaa atgacccca cctgccccg gccgtgaga gtgagagag. 2041 tgacgggcag agcagatgg agccctggc cttgagaga acagagagac actcttgcg atcacaacaga accagaacac accacaactc gacacaacac accacaactc gcacaacaacac accacaactc cctgagaacac accacaaccac cctggcacc cttgtggac tcaaaaacac 2161 agaactccaca ggtagcttgc ctggttctgg gtctacacac ggaactccc cctgggac tcaaagaga. 2221 gcataaaatt ttgcacaaga cagacctga caggaagtcc cctgtggac tggcacagg. 2281 aacagcagaa gccacagga aagacctga caggaagacc aaactcacc accacacaccacac	901 gtcctgcttg atttgcgtgg caagaagagt tcccatgaag gaaagaccag t 961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcacgtctc t 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt g	atatgaaac
961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcacgtctc tggatacca; 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt gtattcaga; 1081 gttccatgcg cagcatgaag gagaatctgt gtcctatgct aagaggcatc atcatgaag; 1141 actgagacaa ggattggcat tcagtcaaat ctatcgtttt tccttgtctg atggcatct; 1201 tgttgctgca caaacgaaga gcaaactcat ccgttctcag actactaatg aacctcaact; 1221 gactggacaa acgatgggga agccactgaa tccaattagc tctaacagc ctgcccatc; 1381 ggccctgtgc agtgggaacc caggtcagga catgaccctc agtagcaata taaattttc; 1441 cataaatggc ccaaaggaac aaatgggcat gcccatgggc aggtttggtg gttctgggg; 1501 aatgaaccat gtgtcaggca tgcaagcaac cactcctcag ggtagtaact atgcactca; 1561 aatgaacagc ccctcacaaa gcagccctgg catgaatcca ggacagcca cctccatgc; 1621 ttcaccaagg catcgcatga gccctggagt ggctggcagc cctcgaatcc caccaggc; 1681 gttttcccct gcaggaagct tgcattcccc tgtgggggtt tgcagcagca caggaaata; 1741 ccatatgtat accaacagct ccctcaatgc acttcaagca cttcaggag ggcaggagg; 1801 tcattatggg tcatcgtgg cttcaccaga acttgaagcac ttggaagcag aaaacctca; 1681 agttaatatg aatcctccc cactcaagcaa gctgagagg ggcaggagg ggcagggg; 1801 tccattatggg tcatcgttgg cttcaccaga acttcaagaa gcaggaagag ggcacggg; 1801 ttggactatat ggggagcct ctgaaggaac ctaaaaaatg ggcaatttgc aaaacctcc; 1681 gttttccctt gcaggaagct tcatcacaga acttcaagaa gcaggaagag ggcacggg; 1801 tcattatggg tcatcgttgg cttcaccaga acttcaagaa gcaggaagag ggcacggg; 1801 ttggactatat ggggagcct ctgaaggaa atgaggaagc ttggactcaa aagacggag; 1801 tcaatagga aaggaaaca atgacccaa actggacaa gcaggagag ggcacctgg; 1801 tgagagagaa aaggaaaca atgacccaa actggacaa gcaggagaga gcagcaggaga; 1801 tcaatagga aaggaagaca cacccaa cctgacccag gcaggagag gtgcactca; 1801 tgagagagaa aaggaaaca atgacccaa actggacaa gcaggagaa gcagcagga; 1801 taacaaatt tggaagaaca atgacccaa actggacaa gcaggagaac ttgagagaga; 1801 taacaaatt ttgcacagaa atgacccaa cctgaccaa gcaggaaca caccaaccga; 1801 taacaaaatt ttgcacagaa agacccaa accagaaa gaagagaac tcaagaagaa; 1801 taacaaaatt ttgcacagaa agaccccaa ccagaagaa aaaacaacaa ccccaaacaa; 1801 agaagagaa gcacaagaa aagacctga ccagaagaa aagaagaa caccaaaca; 1801 taacaaaca ggaagaaca aag	961 atcagaaagt tttactactc gccaggatct ccaaggcaag atcacgtctc t 1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt g	agatttgca
1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt gtattcaga. 1081 gttccatgcg cagcatgaag gagaatctgt gtcctatgct aagaggactc atcatgaag 1141 actgagacaa ggattggcat tcagtcaaat ctatcgtttt tccttgtctg atggcactc 1201 tgttgctgca caaacgaaga gcaaactcat ccgttctcag actactaatg aacctcaacg 1261 tgtaatatct ttacatatgc ttcacagaga gcagaatgtg tgtgtgatga atccggatc 1321 gactggacaa acgatgggga agccactgaa tccaattagc tctaacagcc ctgcccatc 1381 ggccctgtgc agtgggaacc caggtcagga catgacctc agtagcaata taaattttc 1441 cataaatggc ccaaaaggaac aaatgggcat gcccatgggc aggtttggtg gttctgggg 1501 aatgaaccat gtgtcaggca tgcaagcaca cactcctcag ggtagtaact atgcactca. 1561 aatgaacagc ccctcacaaa gcagccctgg catgaatcca ggacagcca cctccatgc 1621 ttcaccaagg catcgcatga gccctggagt ggctggcagc cctcgaatcc cacccagtc 1681 gttttcccct gcaggaagct tgcatcccc tgtgggagtt tgcagcagca caggaaatag 1741 ccatagttat accaacagct ccctcaatgc actcagcag cctcagcagag ggcaggggg 1801 tcattatggg tcatcgttgg cttcaccaga cctaaaaatg ggcaatttg aaaaccccc 1861 agttaatatg aatcctcccc cactcagcaa gatgggaagc ttggagagggggggggg	1021 caccatgaga gcagccatga aaccaggctg ggaggacctg gtaagaaggt g	tettecete
1081 gttccatgcg cagcatgaag gagaatctgt gtcctatgct aagaggcatc atcatgaag 1141 actgagacaa ggattggcat tcagtcaaat ctatcgtttt tccttgtctg atggcactc 1201 tgttgctgca caaacgaaga gcaaactcat ccgttctcag actactaatg aacctcaacg 1221 tgtaatatct ttacatatgc ttcacagaga gcagaatgtg tgtgtgatga atccggatct 1321 gactggacaa acgatgggga agccactgaa tccaattagc tctaacagcc ctgcccatc 1381 ggccctgtgc agtgggaacc caggtcagga catgacctc agtagcaata taaattttc 1441 cataaatggc ccaaaggaac aaatgggcat gcccatggg aggtttggtg gttctggggg 1501 aatgaaccat gtgtcaggca tgcaagcaac cactcctcag ggtagtaact atgcactca 1561 aatgaacagc ccctcacaaa gcagcctgg catgaatcca ggacagcca cctccatgc 1621 ttcaccaagg catcgcatga gccctgggt ggctggcagc cctcgaatcc cacccagtc 1681 gttttcccct gcaggaagct tgcattcccc tgtgggagtt tgcagcagca caggaaatag 1741 ccatagttat accaacagct ccctcaatgc acttcaggcc ctcagcagg ggcaggggg 1801 ctcattaggg tcatcgttgg cttcaccaga cctaaaaatg ggcaatttgc aaaacccca 1861 agttaatatg aatcctcccc cactcagcaa gatgggaagc ttggactcaa aagactgtt 1921 tggactaatat ggggagccct ctgaaggtac aactggacaa gcagagagc 1981 tggagagcaa aaggaaacaa atgaccccaa cctgccccg 1981 tggaggagaa aaggaacaca 1981 tggagggaga agcagctgc atgacagcaa agggcagacc 1981 tggagggaa agcagctgc 1982 tggactaata ttgacagaa 1983 tggagagcaa agcagactgc 1984 tgacgggaag 1985 agcaggaga 1986 agactcaca 1986 agactcaca 1987 tggagagcaa 1988 tggagagcaa 1989 tggagagcaa 1989 tggagagcaa 1980 tggagagagaa 1980 tggagagaa 1980 tggagagaa 1980 tggagagaaa 1980 tggagaaaaaaaa 1980 tggagaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa		ggataccag
1141 actgagacaa ggattggcat toagtoaaat otategtttt toottgtotg atggcatted 1201 tgttgotgca caaacgaaga gcaaactcat cogttotcag actactaatg aacotcaacd 1201 tgttgotgca caaacgaaga gcaaactcat cogttotcag actactaatg aacotcaacd 1201 tgtaatatot ttacatatgc ttcacagaga gcagaatgtg tgtgtgatga atcoggatod 1221 gactggacaa acgatgggga agccactgaa tocaattago totaacago otgoccatca 1381 ggcoctgtgo agtgggaaco caggtcagga catgacocto agtagcaata taaatttood 1441 cataaatggo ccaaaggaac aaatgggcat gcccatgggo aggtttggtg gttotggggg 1501 aatgaaccat gtgtcaggca tgcaagcaac cactoctcag ggtagtaact atgcactcaa 1561 aatgaacago coctcacaaa gcagcoctgg catgaatcca ggacagcca cotcaatgo 1621 ttcaccaagg catcgcatga gccctggagt ggctggcago cotcgaatco cacccagga 1681 gttttococt gcaggaagot tgcattococ tgtgggagtt tgcagcago caggaaatag 1741 ccataagtat accaacagot coctcaatgo actcaaaaatg ggcaatttgo aaaactcoc 1861 agttaatatg aatoctococ cactcagaa gatgggaago ttggacago ctcaggagg ggcaagggg 1801 ctcattaggg tcatcgttgg cttcaccaga actggacaa ggcagagag ggcaactgg 1921 tggactatat ggggagcoct ctgaaggtac aactggacaa gcagagaga gctgcaatca 1981 tggagagcaa aaggaacaaa atgaccocaa cotgococog gccgtgagca gtggagagg 2041 tgacgggcag agcagactgc atgacagcaa agggcagacc aaactcctgc agctggaga 2101 caccaaatct gatcagatgg agccctcgoc cttagccago tctttgtgg gtctacacaa ggaactcgc atgacagaa aggacctcgo cttagcaga accccaa cotgcococ gccgtgagca gtgagagaga 2221 gcataaaatt ttgcacagac tcttgagaga caggagtcc cctgtggact tggacaacac 2261 agaccacaa gccacaggca agacctgga cccaagaaga accccaagaa accccaagaa accccaagaa accccaagaa accccaagaa accccaagaa gcaacaagaa gccacaggaa acccaagaa accccaagaa acccaagaa acccaagaa acccaagaa acccaagaa acccaagaa accaagaa accaagaa acccaagaa acccaagaa accaagaaa accaagaagaa gcaacaagaa accaagaagaa accaagaagaagaa accaagaagaagaagaagaagaagaagaagaagaagaaga		gtattcagaa
1261 tgtaatatct ttacatatgc ttcacagaga gcagaatgtg tgtgtgatga atccggatci 1321 gactggacaa acgatgggga agccactgaa tccaattagc tctaacagcc ctgcccatci 1331 ggccctgtgc agtgggaacc caggtcagga catgacctc agtagcaata taaattttci 1441 cataaatggc ccaaaggaac aaatgggcat gcccatgggc aggtttggtg gttctggggg 1501 aatgaaccat gtgtcaggca tgcaagcaac cactcctcag ggtagtaact atgacccca 1561 aatgaacagc ccctcacaaa gcagccctgg catgaatcca ggacagcca cctccatggc 1621 ttcaccaagg catcgcatga gccctggagt ggctggcagc cctcgaatcc caccagtci 1681 gttttcccct gcaggaagct tgcattcccc tgtgggagtt tgcagcagca caggaaatag 1741 ccatagtta accaacagct ccctcaatgc acttcaggcc ctcaggagg ggcagggg 1801 ctcattaggg tcatcgttgg cttcaccaga acttcaggcc ctcaggagg ggcaggggg 1801 tggatatat agaaccccc cactcagca acttcaggcc ctcaggagg ggcaggggg 1801 tggatatata gaggagccct ctgaaggta actaaaaatg ggcaatttgc aaaaccccc 1861 agttaatatg aatcctcccc cactcagcaa gatgggaagc ttggaaccaa aagactgtt 1921 tggactatat ggggagccct ctgaaggtac actaaaaatg ggcaattga aaaaccccc 1861 agttaatatg aatcctcccc cactcagcaa gatgggaagc ttggaaccaa aggacggg 2041 tgacgggcag agcagaccg atgacagaa actggacaa gccggtgagag 2041 tgacgggcag agcagactg atgacagcaa aggcagagc aaactcctg agctgctgac 2101 caccaaatct gatcagatgg agccctcgcc cttagccag tcttttgcgg atacaaacac 2161 agactccaa ggtagcttg ctggttctgg gtctacacat ggaacctcgc tcaaggaga 2221 gcataaaatt ttgcacagac tcttgagga caggagtcc agcagcacg ctcctggat 2281 aacagcagaa gccacaggca aagacctgaa caggagtcc cctgtggact tggacagc 2281 aacagcagaa gccacaggca aagacctgaa caccaagaa aagacctga caccaagac 2341 agaagtgact attaaacaa agccggtgag cccaagaag aaagagaatg cactacttc 2401 ctatttgcta gataaagaa atactaaaga tattggttta ccagaaataa cccccaaac 2461 tgagaaggag gagatgagct ttgagcctg tgaccagcc ggcagtagac gacgacagc 2581 ggagagagat ttggatgat ttgaagata cagtctgc cagtaacaca aaattaatag caatagaaa 2581 ggagagaatt ttggatgatt tgcagaatag cagttgaca gcaagccac accccagaac 2641 gccaggcgc cctgctggat cagttgaca gcaagccag aaacccaca cagcttttc cagacaccacacacacacacacacacacacacacacac	1081 gttccatgcg cagcatgaag gagaatctgt gtcctatgct aagaggcatc a	atcatgaagt
1261 tgtaatatct ttacatatgc ttcacagaga gcagaatgtg tgtgtgatga atccggatci 1321 gactggacaa acgatgggga agccactgaa tccaattagc tctaacagcc ctgcccatci 1381 ggccctgtgc agtgggaacc caggtcagga catgacctc agtagcaata taaattttci 1441 cataaatggc ccaaaggaac aaatgggcat gcccatgggc aggtttggtg gttctggggg 1501 aatgaaccat gtgtcaggca tgcaagcaac cactcctcag ggtagtaact atgcactcai 1561 aatgaacagc ccctcacaaa gcagccctgg catgaatcca ggacagcca cctccatggc 1621 ttcaccaagg catcgcatga gccctggagt ggctggcagc cctcgaatcc caccagtci 1681 gttttcccct gcaggaagct tgcattcccc tgtgggagtt tgcagcagca caggaaatai 1741 ccatagtat accaacagct ccctcaatgc acttcaggcc ctcaggagg ggcaggggg 1801 ctcattaggg tcatcgttgg cttcaccaga cctaaaaatg ggcaatttgc aaaactccci 1661 agttaatatg aatcctcccc cactcagcaa gatgggaagc ttggactcaa aagactgti 1921 tggactatat ggggagccct ctgaaggtac actaaaaatg ggcaatttgc aaaactccgi 1981 tggaggaga aaggaacaa atgaccccaa cctgcccccg gccgtgagca gtggagagg 2041 tgacgggcag agcagactgc atgacagaa agggcagacc aaactcctgc agctgctgac 2101 caccaaatct gatcagatgg agccctcgcc cttagccagc tctttgtcgg atacaaacac 2161 agactccaa ggtagcttgc ctggttctgg gtctacacat ggaactcgc tcaaggaga 2221 gcataaaatt ttgcacagac tcttgacgag cccaagagtcc agcagcacag ctccaggaga 2221 gcataaaatt ttgcacagac tcttgagaga caggagtcc cctgggacc 2231 aacagcagaa gccacaggca aagacctga ccaagagatc agcagcacag ctcctggatc 2341 agaagtgact attaaacaag agccggtgag cccaagaag aaagaaatg cactacttc 2401 ctatttgcta gataaagat atactaaaga tattggttta ccagaaataa cccccaaac 2461 tgagaaggag gagatgagct ttgagcctgg tgaccagcc ggcagtgagc tggacaact 2521 tgagaaggag gagatgagct ttgagcctgg tgaccagcc ggcagtgagc tggacaact 2581 ggagagaatt ttggatgatt tgcagaatag caagtccacc agcagccac ccaggagcc 2641 gccaggcgc cctgctggat cagttgaca cagttgacacc agcagccac acaacccaccaccaccaccaccacaccaccaccac	1141 actgagacaa ggattggcat tcagtcaaat ctatcgtttt tccttgtctg a	atggcactct
1321 gactggacaa acgatgggga agccactgaa tccaattagc tctaacagcc ctgccatcd. 1381 ggccctgtgc agtgggaacc caggtcagga catgaccctc agtagcaata taaattttc. 1441 cataaatggc ccaaaggaac aaatgggcat gccatggc aggtttggtg gttctggggg. 1501 aatgaaccat gtgtcaggca tgcaagcac cactcctcag ggtagtaact atgcactca. 1561 aatgaacagc ccctcacaaa gcagccctgg catgaatcca ggacagccca cctccatggc. 1621 ttcaccaagg catcgcatga gccctggagt ggctggcagc cctcgaatcc cacccagtc. 1681 gttttcccct gcaggaagct tgcattcccc tgtgggagtt tgcagcagca caggaaatag. 1741 ccatagttat accaacagct ccctcaatgc acttcaggcc ctcagcagg ggcaggggg. 1801 ctcattaggg tcatcgttgg cttcaccaga cctaaaaatg ggcaatttgc aaaactccc. 1861 agttaatatg aatcctcccc cactcagcaa gatgggaagc ttggactcaa aagactgtt. 1921 tggactatat ggggagccct ctgaaggtac aactggacaa gcagagaga gctgccatc. 1981 tggagagcaa aaggaacaa atgacccaa cctgccccg gccgtgagca gtggagagg. 2041 tgacgggcag agcagactgc atgacagcaa agggcagacc aaactcctgc agctgcatc. 2101 caccaaatct gatcagatgg agccctcgcc cttagccagc tctttgtcgg atacaaaca. 2161 agactccaca ggtagcttgc ctggttctgg gtctacacat ggaacctcgc tcaaggaga. 2221 gcataaaatt ttgcacagac tcttgcagga cagcagttcc cctgtggact tggccaagt. 22341 agaagtgact attaaacaag agccggtgag cccaagaag aaagagaatg cactacttc. 2401 ctatttgcta gataaagatg atactaaaga tattggtta ccagaaataa cccccaaac. 2461 tgagagactg gacagtaaga cagatcctgc cagtaacaca aaattaatag caatacttc. 2581 ggagaggatt ttggatgatt tgcagaatag tcaatacca cagctttcc cagacacca. 2581 ggagagagtt ttggatgatt tgcagaatag tcaatacca cagcttttcc cagacacca. 2581 ggagagagtt ttggatgatt tgcagaatag cagtcacca cagcaccta atcaatgacc tcatgcaacc. 2581 ggagagagtt ttggatgatt tgcagaatag cagaccatc atcaatgacc tcatgcaacc. 2581 gagagagatt ttggatgatt tgcagaatag cagaccatc atcaatgacc tcatgcaacc. 2581 gagagagatt ttggatgatt tgcagaatag cagtagccat atcaatgacc tcatgcaacc. 2581 gagagagata aacagccctg tcacacctg tggagccatc atcaatgacc tcatgcaacc. 2581 gagagagata aacagccctg tcacacctg tggagccatc atcaatgacc tcatgcaacc. 2641 gccaggcgcc cctgctggat cagttgaca gcaagccatc atcaatgacc tcatgcaacc.	1201 tgttgctgca caaacgaaga gcaaactcat ccgttctcag actactaatg a	acctcaact
1321 gactggacaa acgatgggga agccactgaa tccaattagc tctaacagcc ctgccatcd. 1381 ggccctgtgc agtgggaacc caggtcagga catgaccctc agtagcaata taaattttc. 1441 cataaatggc ccaaaggaac aaatgggcat gccatggc aggtttggtg gttctggggg. 1501 aatgaaccat gtgtcaggca tgcaagcac cactcctcag ggtagtaact atgcactca. 1561 aatgaacagc ccctcacaaa gcagccctgg catgaatcca ggacagccca cctccatggc. 1621 ttcaccaagg catcgcatga gccctggagt ggctggcagc cctcgaatcc cacccagtc. 1681 gttttcccct gcaggaagct tgcattcccc tgtgggagtt tgcagcagca caggaaatag. 1741 ccatagttat accaacagct ccctcaatgc acttcaggcc ctcagcagg ggcaggggg. 1801 ctcattaggg tcatcgttgg cttcaccaga cctaaaaatg ggcaatttgc aaaactccc. 1861 agttaatatg aatcctcccc cactcagcaa gatgggaagc ttggactcaa aagactgtt. 1921 tggactatat ggggagccct ctgaaggtac aactggacaa gcagagaga gctgccatc. 1981 tggagagcaa aaggaacaa atgacccaa cctgccccg gccgtgagca gtggagagg. 2041 tgacgggcag agcagactgc atgacagcaa agggcagacc aaactcctgc agctgcatc. 2101 caccaaatct gatcagatgg agccctcgcc cttagccagc tctttgtcgg atacaaaca. 2161 agactccaca ggtagcttgc ctggttctgg gtctacacat ggaacctcgc tcaaggaga. 2221 gcataaaatt ttgcacagac tcttgcagga cagcagttcc cctgtggact tggccaagt. 22341 agaagtgact attaaacaag agccggtgag cccaagaag aaagagaatg cactacttc. 2401 ctatttgcta gataaagatg atactaaaga tattggtta ccagaaataa cccccaaac. 2461 tgagagactg gacagtaaga cagatcctgc cagtaacaca aaattaatag caatacttc. 2581 ggagaggatt ttggatgatt tgcagaatag tcaatacca cagctttcc cagacacca. 2581 ggagagagtt ttggatgatt tgcagaatag tcaatacca cagcttttcc cagacacca. 2581 ggagagagtt ttggatgatt tgcagaatag cagtcacca cagcaccta atcaatgacc tcatgcaacc. 2581 ggagagagtt ttggatgatt tgcagaatag cagaccatc atcaatgacc tcatgcaacc. 2581 gagagagatt ttggatgatt tgcagaatag cagaccatc atcaatgacc tcatgcaacc. 2581 gagagagatt ttggatgatt tgcagaatag cagtagccat atcaatgacc tcatgcaacc. 2581 gagagagata aacagccctg tcacacctg tggagccatc atcaatgacc tcatgcaacc. 2581 gagagagata aacagccctg tcacacctg tggagccatc atcaatgacc tcatgcaacc. 2641 gccaggcgcc cctgctggat cagttgaca gcaagccatc atcaatgacc tcatgcaacc.		
1381 ggccctgtgc agtgggaacc caggtcagga catgacctc agtagcaata taaatttto 1441 cataaatggc ccaaaggaac aaatgggcat gcccatgggc aggtttggtg gttctggggg 1501 aatgaaccat gtgtcaggca tgcaagcaac cactcctcag ggtagtaact atgcactca 1561 aatgaacagc ccctcacaaa gcagccctgg catgaatcca ggacagccca cctccatgc 1621 ttcaccaagg catcgcatga gccctggagt ggctggcagc cctcgaatcc cacccagtc 1681 gttttcccct gcaggaagct tgcattcccc tgtgggagtt tgcagcagca caggaaataa 1741 ccataagttat accaacagct ccctcaatgc acttcaggcc ctcagcgagg ggcacggggg 1801 ctcattaggg tcatcgttgg cttcaccaga cctaaaaatg ggcaatttgc aaaactccc 1861 agttaatatg aatcctccc cactcagcaa gatgggaagc ttggactcaa aagactgtt 1921 tggactatat ggggagccct ctgaaggtac aactggacaa gcagagaga gctgccatc 1981 tggaggagaa aaggaaacaa atgaccccaa cctgccccg gccgtgagca gtggagagg 2041 tgacgggga agcagactgc atgacagaa agggcagaac aaactcctgc agctgctaca 1981 tggaggagaa agcagactgc atgacagaa agggcagacc atcatgacaa gcagagaga gctgccatc 1981 tggactcaat ggatcagatgg agccctcgcc cttagccagc tctttgtcgg atacaaaca 2161 agactccaaa ggtagcttgc ctggttctgg gtctacaacat ggaacctcgc tcaaggaga 2221 gcataaaatt ttgcacagac tcttgcagga cagcagttcc cctgtggact tggccaagt 2281 aacagcagaa gccacaggca aagacctgag cccaaggaac cacaggaac 2281 agaaggact attaaacaag agccggtgag ccccaagaag aaagagaatg cactacttcg 2401 ctatttgcta gataaagaa aagacctgag cccaagaaga aaagagaatg cactacttcg 2401 ctatttgcta gataaagaa aagacctga cccaagaaga aaagagaatg cactacttcg 2401 ctatttgcta gataaagaa cagatcctgc cagtaacaca aaattaatag caatgaaaa 2521 tgagaaggag gagatgagct ttgagcctgg tgaccagcct ggcagtagc tggacaacc 2581 gagagagagt ttggatatt tgcagaatag tcaatacca cagcttttcc cagacacca 2581 gagagagatt ttggatgatt tgcagaatag ccatacctgc cagtaacaca aaattaatag caatgaaaaccaca 2461 gccaggcgc cctgctggat cagttgacaa gcaagccaca aacacacacacacacacacacac	1261 tgtaatatct ttacatatgc ttcacagaga gcagaatgtg tgtgtgatga a	atccggatct
1441 cataaatggc ccaaaggaac aaatgggcat gcccatgggc aggtttggtg gttctggggg 1501 aatgaaccat gtgtcaggca tgcaagcaac cactcctcag ggtagtaact atgcactca 1561 aatgaacagc ccctcacaaa gcagccctgg catgaatcca ggacagcca cctccatgcc 1621 ttcaccaagg catcgcatga gccctggagt ggctggcagc cctcgaatcc cacccagtcc 1681 gttttcccct gcaggaagct tgcattcccc tgtgggagtt tgcagcagca caggaaatac 1741 ccatagttat accaacagct ccctcaatgc acttcaggcc ctcagcagg ggcacggggc 1801 ctcattaggg tcatcgttgg cttcaccaga cctaaaaatg ggcaatttgc aaaactcccc 1861 agttaatatg aatcctcccc cactcagcaa gatgggaagc ttggactcaa aagactgtt 1921 tggactatat ggggagccct ctgaaggtac aactggacaa gcagagagca gctgccatcc 1981 tggaggagca aaggaaacaa atgacccaa cctgccccg gccgtgagca gtggagagg 2041 tgacgggcag agcagactgc atgacagcaa agggcagacc aaactcctgc agctgctgac 2101 caccaaatct gatcagatgg agccctcgcc cttagccagc tcttttgtcgg atacaaacac 2161 agactccaca ggtagcttgc ctggttctgg gtctacacat ggaacctcgc tcaaggaga 2221 gcataaaatt ttgcacagac tcttgcagga cagcagttcc cctgtggact tggccaagt 2281 aacagcagaa gccacaggca aagacctgg ccaggagtcc agcagcacag ctcctgggat 2281 aacagcagaa gccacaggca aagacctgag cccaagaag aaagagaatg cactctgc 2401 ctatttgcta gataaagatg atactaaaga tattggttta ccagaaataa cccccaaac 2461 tgagagactg gacagtaaga cagatcctgc cagtaacaca aaattaatag caatgaaaa 2521 tgagaaaggag gagatgagct ttgagcctgg tgaccagcct ggcagtgagc tggacaacc 2581 ggagaggatt ttggatgatt tgcagaatag tcaattacca cagcttttcc cagacacga 2641 gccaggcgc cctgctggat cagttgacaa gcaagccatc atcaatgac tcatgcaacc 2701 cacagctgaa aacagcctg tcacacctgt tggagccag aaaacagcac tgcgaattt	1321 gactggacaa acgatgggga agccactgaa tccaattagc tctaacagcc c	tgcccatca
1501 aatgaacat gtgtcaggca tgcaagcaac cactectcag ggtagtaact atgcactca. 1561 aatgaacage cectcacaaa geageeetgg catgaateca ggacageeca cetccatge. 1621 ttcaccaagg categcatga geeetgggt ggctggcage cetcgaatee cacecagte. 1681 gttttceeet geaggaaget tgcatteeee tgtgggagt tgcageagea caggaaataa. 1741 ceatagttat accaacaget cectcaatgea acttcaggee etcagegagg ggcacggggg. 1801 etcattaggg teategttgg etteaceagaa cetaaaaatg ggcaatttgeaaaacaece. 1861 agttaatatg aatecteeee cactcageaa gatgggaage ttggactcaa aagactgtt. 1921 tggactatat ggggageeet etgaaggtae aactggacaa geagagagea getgecate. 1981 tggagagcaa aaggaaacaa atgaceecaa eetgeeeeg geegtgagea gtggagagg. 2041 tgacgggeag agcagactgeatgacaaa aggacagaac aaacteetgeagagag. 2101 caccaaatet gateagatgg agceetegee ettageeage tetttgtegg atacaaacaa. 2161 agacteeaca ggtagettge etggttetgg gtetacacat ggaacetege teaagggag. 2221 geataaaatt ttgeacagae tettgeagga eageagtee eetettgegeet tggeeaagt. 2281 aacageagaa geeacaggea aagacetgag eeaggagtee agcageacag eteetggat. 2341 agaagtgact attaaacaag ageeggtgag eecaagagatee ageageacag eteetggat. 2401 etatttgeta gataaagatg atactaaaga tattggttta eeagaaataa eeceeaac. 2461 tgagaagaetg gacagtaaga eagateetge eagtaacaca aaattaatag eaatgaaaa. 2521 tgagaaggag gagatgaget ttgageetgg tgaccageet ggcagtgage tggacaact. 2581 ggaggagatt ttggatgatt tgeagaatag teaattacea eagettttee eagacaga. 2641 geeaggegee eetgetggat eagttgacaa geaageeate ateatgace teatgeace. 2641 geeaggegaa aacageetga eagttgacaa geaageeate ateatgace teatgeace. 2641 geeaggagaa aacageetga eagttgacaa geaageeate ateatgace teatgeace. 2641 geeaggagaa aacageetga eagttgacaa geaageeate aacageace teatgeace. 2701 cacagetgaa aacageetga eagttgacaa gaagee	1381 ggccctgtgc agtgggaacc caggtcagga catgaccctc agtagcaata t	aaattttcc
1561 aatgaacage cecteacaaa geagecetgg catgaateca ggacageeca cetecatge 1621 tteaceaagg categeatga geeetggagt ggetggeage cetegaatec cacceagter 1681 gtttteecet geaggaaget tgeatteece tgtgggagtt tgeageagea caggaaatag 1741 ceatagttat aceaacaget cecteaatge actteaggee etcagegagg ggeacggggg 1801 etcattaggg teategttgg etteaceaga cetaaaaatg ggeaatttge aaaacteee 1861 agttaatatg aateeteece cacteageaa gatgggaage ttggacteaa aagactgtt 1921 tggactatat ggggageecet etgaaggtae aactggacaa geagagagea getgeeatee 1981 tggagggaa aaggaaacaa atgaceecaa eetgeeeeg geegtgagea gtgagagag 2041 tgacggggaa ageagactge atgacageaa agggcagaee aaacteetge agetgetgaa 2101 caceaaatet gateagatgg ageeetegee ettageeage tetttgtegg atacaaacaa 2161 agacteeaca ggtagettge etggttetgg gtetacacat ggaacetege teaaggagaa 2221 geataaaatt ttgeacagae tetttgeagga cagcagttee eetggteega 2221 geataaaatt ttgeacagae tetttgeagga cagcagttee eetggtage 2231 aacageagaa geeacaaggea aagacetgag eeceaaggaa geagageete tggecaagte 22401 etatttgeta gataaacaa ageeggtgag eeceaagaag aaagagaatg cactaettee 2401 etatttgeta gataaaagat gatactaaaga tattggttta eeagaaataa eeceeaace 2461 tgagaaggag gacagtaaga eagateetge eagtaacaca aaattaatag eaatgaaaa 2521 tgagaaaggag gagatgaget ttgagacatg tgacaageet ggacagacet eegsaacet ggaaggagat ttggatgatt tgeagaatag teaattaeca cagettttee eagacagaa 2641 geeaggagee eetgetggat tgacaacet tgagaacace 2641 geeaggagee eetgetggat tgacaacet tgagaaacet tgagaagaa aacageeggee eetgetggat tggacaacet tgagaagaagaa aacageeggaa eagaagaagaa eagaagaacet tgagacaacet eagaacaca aaattaatag eagacaacet 2641 geeagagaga gagatgaget ttgagaatag teaattaeca cagettttee eagacagaa 2641 geeaggagee eetgetggat tegacacettg tgagacaacet teatgeacet t	1441 cataaatggc ccaaaggaac aaatgggcat gcccatgggc aggtttggtg g	gttctggggg
1621 ttcaccaagg categcatga gecetggagt ggetggcage cetegaatee cacccagted 1681 gtttteceet geaggaaget tgeattecee tgtgggagtt tgeageagea caggaaatag 1741 ceatagttat accaacaget ceeteaatge actteaggee etcagegagg ggeaeggggg 1801 etcattaggg teategttgg etteaceaga cetaaaaatg ggeaatttge aaaacteeed 1861 agttaatatg aatecteeee caeteageaa gatgggaage ttggacteaa aagactgtt 1921 tggactatat ggggageeet etgaaggtae aactggacaa geaggaggae getgeeatee 1981 tggaggagaa aaggaaacaa atgaceeeaa eetgeeeeg geegtgagea gtgaggagge 2041 tgaeegggaga ageagaetge atgacageaa agggeagaee aaacteetge agetgetgae 2101 caccaaaatet gateagatgg ageeeteegee ettageeage tetttgtegg atacaaacaa 2161 agaeeteeaa ggtagettge etggteetgg gtetacacat ggaacetege teaaggagaa 2221 geataaaatt ttgeacagae tettgeagga eageagtee ageageacag eteetggate 2341 agaagtgaet attaaacaag ageeggtgag eeceaagaag aaagagaatg eactaettee 2401 etatttgeta gataaagatg atactaaaga tattggttta eeagaaataa eeceeaaace 2461 tgagaagaetg gacagtaaga eagateetge eagtaacaca aaattaatag eaatgaaaac 2521 tgagaaggag gagatgaget ttgageetgg tgaceageet ggeagtgage tggacaacte 2581 ggagaagat ttggatgat ttgageetgg tgaceageet ggeagtgage tggacaace 2581 ggaggagatt ttggatgat ttgageetgg tgaceageet ggeagtgage tegaacace 2641 geeaggeee eetgetggat eagttgacaa geageeete ateaatgaee teatgaaaca 2641 geeaggegee eetgetggat tagacaactgg tagacaace 2641 geeaggegee eetgetggat eagttgacaa geageeete ateaatgaee teatgaacac 2641 geeaggegee eetgetggat tagacaactgg tagaceace 2641 geeaggegee eetgetggat tagacaactg teaattacea cagettttee eagacacgae 2641 geeaggegee eetgetggat tagacaactg teaattacea cagettttee eagacacgae 2641 geeaggegee eetgetggat tagacaactg teaattacea cagettttee eagacacgae 2641 geeaggegee eetgetggat tagacaactg teaattacea eagacacaca aaaacacacacacacacacacacacacaca	1501 aatgaaccat gtgtcaggca tgcaagcaac cactcctcag ggtagtaact a	atgcactcaa
1681 gttttccct gcaggaagct tgcattccc tgtgggagtt tgcagcagca caggaaatag 1741 ccatagttat accaacagct ccctcaatgc acttcaggcc ctcagcgagg ggcacggggg 1801 ctcattaggg tcatcgttgg cttcaccaga cctaaaaatg ggcaatttgc aaaactccc 1861 agttaatatg aatcctccc cactcagcaa gatgggaagc ttggactcaa aagactgtt 1921 tggactatat ggggagccct ctgaaggtac aactggacaa gcagagagca gctgccatc 1981 tggagagcaa aaggaaacaa atgaccccaa cctgcccccg gccgtgagca gtgagagag 2041 tgacgggcag agcagactgc atgacagcaa agggcagacc aaactcctgc agctgctga 2101 caccaaatct gatcagatgg agccctcgcc cttagccagc tctttgtcgg atacaaaca 2161 agactccaca ggtagcttgc ctggttctgg gtctacacat ggaacctcgc tcaaggaga 2221 gcataaaatt ttgcacagac tcttgcagga cagcagttcc cctgtggact tggccaagt 2281 aacagcagaa gccacaggca aagacctgag ccagagttcc cctgtggact tggccaagt 2341 agaagtgact attaaacaag agccggtgag ccccaagaag aaagagaatg cactacttc 2401 ctatttgcta gataaagatg atactaaaga tattggttta ccagaaataa cccccaaacc 2461 tgagaagactg gacagtaaga cagatcctgc cagtaacaca aaattaatag caatgaaaac 2521 tgagaaggag gagatgagct ttgagcctgg tgaccagcct ggcagtgagc tggaccaact 2581 ggaggagatt ttggatgatt tgcagaatag tcaattacca cagcttttcc cagacacgac 2641 gccaggcgcc cctgctggat cagttgacaa gcaagccatc atcaatgacc tcatgcaacc 2701 cacagctgaa aacagccctg tcacacctgt tggagcccag aaaacagcac tgcgaattc	1561 aatgaacago coetcacaaa goagocotgg catgaatoca ggacagocoa o	ctccatgct
1741 ccatagttat accaacaget cceteaatge actteaggee etcagegagg ggeaegggggggggggggggggggggggggg	1621 ttcaccaagg catcgcatga gccctggagt ggctggcagc cctcgaatcc c	cacccagtca
1801 ctcattaggg tcatcgttgg cttcaccaga cctaaaaatg ggcaatttgc aaaactcccquaggga agttaatatg aatcctcccc cactcagcaa gatgggaagc ttggactcaa aagactgtt 1921 tggactatat ggggagccct ctgaaggtac aactggacaa gcagagagca gctgcatca 1981 tggaggacaa aaggaaacaa atgaccccaa cctgcccccg gccgtgagca gtgaggaggc 2041 tgacgggcag agcagatgc atgacagcaa agggcagacc aaactcctgc agctgctgac 2101 caccaaatct gatcagatgg agccctcgcc cttagccagc tctttgtcgg atacaaacaa 2161 agactccaca ggtagcttgc ctggttctgg gtctacacat ggaacctcgc tcaaggagaa 2221 gcataaaatt ttgcacagac tcttgcagga cagcagttcc cctgtggact tggccaagtc 2281 aacagcagaa gccacaggca aagacctgag ccagagatcc agcagcacag ctcctggatc 2341 agaagtgact attaaacaag agccggtgag ccccaagaag aaagagaatg cactacttc 2401 ctatttgcta gataaagat atactaaaga tattggttta ccagaaataa cccccaaacc 2461 tgaggaggag gacagtaaga cagatcctgc cagtaacaca aaattaatag caatgaaaa 2521 tgagaaggag gagatgagct ttgagcctgg tgaccagcct ggcagtgagc tggacaactc 2581 ggaggagatt ttggatgatt tgcagaatag tcaattacca cagcttttcc cagacacga 2641 gccaggcgcc cctgctggat cagttgacaa gcaagccatc atcaatgacc tcatgcaacc 2701 cacagctgaa aacagccctg tcacacctgt tggagcccag aaaacagcac tgcgaatta	1681 gttttcccct gcaggaagct tgcattcccc tgtgggagtt tgcagcagca c	caggaaatag
1861 agttaatatg aatcetecce cacteageaa gatgggaage ttggacteaa aagactgtte 1921 tggactatat ggggageet etgaaggtae aactggacaa geagagagea getgecatee 1981 tggaggaea aaggaaacaa atgaceeeaa eetgeeeeg geegtgagea gtgagagage 2041 tgaegggeag agcagaetge atgacageaa agggeagaee aaacteetge agetgetgae 2101 caccaaatet gateagatgg agceetegee ettageeage tetttgtegg atacaaacae 2161 agaeteeaca ggtagettge etggttetgg gtetacacat ggaacetege teaaggagage 2221 geataaaatt ttgeacagae tettgeagga cageagttee eetggteetgae 2281 aacageagaa geeacaggea aagaeetgga eagagetee eetggteetggaet tggeeaagte 2281 agaagtgaet attaaacaag ageeggtgag eeceaagaag aaagagaatg cactaettee 2401 etattgeta gataaagatg atactaaaga tattggttta eeagaaataa eeceeaaace 2461 tgaggagetg gacagtaaga eagateetge eagtaacaca aaattaatag eaatgaaaa 2521 tgagaaggag gagatgaget ttgageetgg tgaceageet ggeagtgage tggacaacet 2581 ggaggagatt ttggatgatt tgeagaatag teaattacea eagetttee eagacacgae 2641 geeaggegee eetgetggat eagttgacaa geaageeate atcaatgaee teatgeaace 2701 cacagetgaa aacageeetg teaacacetg tggageecag aaaacagaac tgegaattee 2701 cacagetgaa aacageeetg teaacacetg tggageee tggaattee 2701 cacagetgaa aacageeetg teaacacetg tggageecag aaaacageae tgegaattee 2701 cacagetgaa aacageeetg teaacacetg tggageecag aaaacageae tgggaattee 2701 cacagetgaa aacageeetg teaacacetg tggageecag aaaacageae tgegaattee 2701 cacagetgaa aacageeetg teaacacetg tggageecag aaaacageae tgggaattee	1741 ccatagttat accaacagct ccctcaatgc acttcaggcc ctcagcgagg g	gcacggggt
1921 tggactatat ggggagccct ctgaaggtac aactggacaa gcagagagca gctgccatca 1981 tggagagcaa aaggaaacaa atgaccccaa cctgccccg gccgtgagca gtgagagag 2041 tgacgggcag agcagactgc atgacagcaa agggcagacc aaactcctgc agctgctgac 2101 caccaaatct gatcagatgg agccctcgcc cttagccagc tctttgtcgg atacaaaca 2161 agactccaca ggtagcttgc ctggttctgg gtctacacat ggaacctcgc tcaaggagac 2221 gcataaaatt ttgcacagac tcttgcagga cagcagttcc cctgtggact tggccaagt 2281 aacagcagaa gccacaggca aagacctgag ccaggagtcc agcagcacag ctcctggatc 2341 agaagtgact attaaacaag agccggtgag ccccaagaag aaagagaatg cactacttc 2401 ctatttgcta gataaagatg atactaaaga tattggttta ccagaaataa cccccaaacc 2461 tgagaagctg gacagtaaga cagatcctgc cagtaacaca aaattaatag caatgaaaac 2521 tgagaaggag gagatgagct ttgagcctgg tgaccagcct ggcagtgagc tggacaact 2581 ggaggagatt ttggatgatt tgcagaatag tcaattacca cagcttttcc cagacacga 2641 gccaggcgcc cctgctggat cagttgacaa gcaagccatc atcaatgacc tcatgcaacc 2701 cacagctgaa aacagccctg tcacacctgt tggagcccag aaaacagcac tgcgaattt	1801 ctcattaggg tcatcgttgg cttcaccaga cctaaaaatg ggcaatttgc a	aaactcccc
1981 tggagagcaa aaggaaacaa atgaccccaa cctgcccccg gccgtgagca gtgagagagc 2041 tgacgggcag agcagactgc atgacagcaa agggcagacc aaactcctgc agctgctgac 2101 caccaaatct gatcagatgg agccctcgcc cttagccagc tctttgtcgg atacaaacac 2161 agactccaca ggtagcttgc ctggttctgg gtctacacat ggaacctcgc tcaaggagac 2221 gcataaaatt ttgcacagac tcttgcagga cagcagttcc cctgtggact tggccaagt 2281 aacagcagaa gccacaggca aagacctgag ccaggagtcc agcagcacag ctcctggatc 2341 agaagtgact attaaacaag agccggtgag ccccaagaag aaagagaatg cactacttc 2401 ctatttgcta gataaagatg atactaaaga tattggttta ccagaaataa cccccaaacc 2461 tgagagactg gacagtaaga cagatcctgc cagtaacaca aaattaatag caatgaaaac 2521 tgagaaggag gagatgagct ttgagcctgg tgaccagcct ggcagtgagc tggacaactc 2581 ggaggagatt ttggatgatt tgcagaatag tcaattacca cagcttttcc cagacacgac 2641 gccaggcgcc cctgctggat cagttgacaa gcaagccatc atcaatgacc tcatgcaacc 2701 cacagctgaa aacagccctg tcacacctgt tggagcccag aaaacagcac tgcgaattt	1861 agttaatatg aatceteece caeteageaa gatgggaage ttggaeteaa a	agactgttt
2041 tgacgggcag agcagactgc atgacagcaa agggcagacc aaactcctgc agctgctgace 2101 caccaaatct gatcagatgg agccctcgcc cttagccagc tctttgtcgg atacaaacacae 2161 agactccaca ggtagcttgc ctggttctgg gtctacacat ggaacctcgc tcaaggagaa 2221 gcataaaatt ttgcacagac tcttgcagga cagcagttcc cctgtggact tggccaagtt 2281 aacagcagaa gccacaggca aagacctgag ccaggagtcc agcagcacag ctcctggate 2341 agaagtgact attaaacaag agccggtgag ccccaagaag aaaggaaatg cactacttc 2401 ctatttgcta gataaagatg atactaaaga tattggttta ccagaaataa cccccaaace 2461 tgagaagactg gacagtaaga cagatcctgc cagtaacaca aaattaatag caatgaaaa 2521 tgagaaggag gagatgagct ttgagcctgg tgaccagcct ggcagtgagc tggacaactt 2581 ggaggagatt ttggatgatt tgcagaatag tcaattacca cagcttttcc cagacacga 2641 gccaggcgcc cctgctggat cagttgacaa gcaagccatc atcaatgacc tcatgcaacc 2701 caccagctgaa aacagccctg tcacacctgt tggagcccag aaaacagcac tgcgaatta	1921 tggactatat ggggagccct ctgaaggtac aactggacaa gcagagagca g	gctgccatcc
2101 caccaaatct gatcagatgg agccctcgcc cttagccagc tctttgtcgg atacaaacacacacacacacacacacacacacacacaca	1981 tggagagcaa aaggaaacaa atgaccccaa cctgcccccg gccgtgagca g	gtgagagagc
2161 agactecaca ggtagettge etggttetgg gtetacacat ggaacetege teaaggagaa 2221 geataaaatt ttgeacagae tettgeagga eageagttee eetgtggaet tggeeaagte 2281 aacageagaa geeacaggea aagacetgag eegagagtee ageaggaacag eteetggate 2341 agaagtgaet attaaacaag ageeggtgag eeceaagaag aaaggaaatg eactaette 2401 etatttgeta gataaagatg atactaaaga tattggttta eeagaaataa eeceeaaace 2461 tgagagaetg gacagtaaga eagateetge eagtaacaca aaattaatag eaatgaaaace 2521 tgagaaggag gagatgaget ttgageetgg tgaceageet ggeagtgage tggacaacet 2581 ggaggagatt ttggatgatt tgeagaatag teaattacea eagettttee eagacaega 2641 geeaggegee eetgetggat eagttgaeaa geaageeate ateaatgaee teatgeaace 2701 cacagetgaa aacageeetg teacacetgt tggageecag aaaacageae tgegaatte		
2221 gcataaaatt ttgcacagac tcttgcagga cagcagttcc cctgtggact tggccaagttc2281 aacagcagaa gccacaggca aagacctgag ccaggagtcc agcagcacag ctcctggatc2341 agaagtgact attaaacaag agccggtgag ccccaagaag aaagagaatg cactacttc2401 ctatttgcta gataaagatg atactaaaga tattggttta ccagaaataa cccccaaacc2461 tgagagactg gacagtaaga cagatcctgc cagtaacaca aaattaatag caatgaaaac2521 tgagaaggag gagatgagct ttgagcctgg tgaccagcct ggcagtgagc tggacaact2581 ggaggagatt ttggatgatt tgcagaatag tcaattacca cagcttttcc cagacacga2641 gccaggcgcc cctgctggat cagttgacaa gcaagccatc atcaatgacc tcatgcaac2701 cacagctgaa aacagccctg tcacacctgt tggagcccag aaaacagcac tgcgaatta	2101 caccaaatct gatcagatgg agecetegee ettagecage tetttgtegg a	atacaaacaa
2281 aacagcagaa gccacaggca aagacctgag ccaggagtcc agcagcacag ctcctggatc 2341 agaagtgact attaaacaag agccggtgag ccccaagaag aaagagaatg cactacttc 2401 ctatttgcta gataaagatg atactaaaga tattggttta ccagaaataa cccccaaacc 2461 tgagaagactg gacagtaaga cagatcctgc cagtaacaca aaattaatag caatgaaaac 2521 tgagaaggag gagatgagct ttgagcctgg tgaccagcct ggcagtgagc tggacaact 2581 ggaggagatt ttggatgatt tgcagaatag tcaattacca cagcttttcc cagacacgag 2641 gccaggcgcc cctgctggat cagttgacaa gcaagccatc atcaatgacc tcatgcaacc 2701 cacagctgaa aacagccctg tcacacctgt tggagcccag aaaacagcac tgcgaatta	2161 agactecaca ggtagettge etggttetgg gtetacacat ggaacetege t	caaggagaa
2341 agaagtgact attaaacaag agccggtgag ccccaagaag aaagagaatg cactactto 2401 ctatttgcta gataaagatg atactaaaga tattggttta ccagaaataa cccccaaac 2461 tgagagactg gacagtaaga cagatcctgc cagtaacaca aaattaatag caatgaaaac 2521 tgagaaggag gagatgagct ttgagcctgg tgaccagcct ggcagtgagc tggacaact 2581 ggaggagatt ttggatgatt tgcagaatag tcaattacca cagcttttcc cagacacga 2641 gccaggcgcc cctgctggat cagttgacaa gcaagccatc atcaatgacc tcatgcaac 2701 cacagctgaa aacagccctg tcacacctgt tggagcccag aaaacagcac tgcgaatta	2221 gcataaaatt ttgcacagac tcttgcagga cagcagttcc cctgtggact t	ggccaagtt
2401 ctatttgcta gataaagatg atactaaaga tattggttta ccagaaataa cccccaaacc 2461 tgagaagactg gacagtaaga cagatcctgc cagtaacaca aaattaatag caatgaaaa 2521 tgagaaggag gagatgagct ttgagcctgg tgaccagcct ggcagtgagc tggacaact 2581 ggaggagatt ttggatgatt tgcagaatag tcaattacca cagcttttcc cagacacga 2641 gccaggcgcc cctgctggat cagttgacaa gcaagccatc atcaatgacc tcatgcaacc 2701 cacagctgaa aacagccctg tcacacctgt tggagcccag aaaacagcac tgcgaatta	2281 aacagcagaa gccacaggca aagacctgag ccaggagtcc agcagcacag c	ctcctggatc
2461 tgagagactg gacagtaaga cagatcetge cagtaacaca aaattaatag caatgaaaaa 2521 tgagaaggag gagatgaget ttgageetgg tgaccageet ggeagtgage tggacaact 2581 ggaggagatt ttggatgatt tgeagaatag teaattacea cagettttee cagacacgag 2641 gecaggegee cetgetggat cagttgacaa geaageeate atcaatgace teatgeaacc 2701 cacagetgaa aacageeetg teacacetgt tggageecag aaaacageae tgegaatte	2341 agaagtgact attaaacaag agccggtgag ccccaagaag aaagagaatg c	cactacttcg
2521 tgagaaggag gagatgagct ttgagcctgg tgaccagcct ggcagtgagc tggacaact 2581 ggaggagatt ttggatgatt tgcagaatag tcaattacca cagcttttcc cagacacgag 2641 gccaggcgcc cctgctggat cagttgacaa gcaagccatc atcaatgacc tcatgcaacc 2701 cacagctgaa aacagccctg tcacacctgt tggagcccag aaaacagcac tgcgaatte	2401 ctatttgcta gataaagatg atactaaaga tattggttta ccagaaataa c	ccccaaact
2581 ggaggagatt ttggatgatt tgcagaatag tcaattacca cagcttttcc cagacacgag 2641 gccaggcgcc cctgctggat cagttgacaa gcaagccatc atcaatgacc tcatgcaac 2701 cacagctgaa aacagccctg tcacacctgt tggagcccag aaaacagcac tgcgaattt	2461 tgagagactg gacagtaaga cagatcctgc cagtaacaca aaattaatag c	caatgaaaac
2641 gccaggcgcc cctgctggat cagttgacaa gcaagccatc atcaatgacc tcatgcaacc 2701 cacagctgaa aacagccctg tcacacctgt tggagcccag aaaacagcac tgcgaattt	2521 tgagaaggag gagatgagct ttgagcctgg tgaccagcct ggcagtgagc t	ggacaactt
2701 cacagetgaa aacageeetg teacacetgt tggageecag aaaacageac tgegaattt	2581 ggaggagatt ttggatgatt tgcagaatag tcaattacca cagcttttcc c	cagacacgag
	2641 gccaggcgcc cctgctggat cagttgacaa gcaagccatc atcaatgacc t	catgcaact
2761 acagagcact tttaataacc cacgaccagg gcaactgggc aggttattgc caaaccaga	2761 acagagcact tttaataacc cacgaccagg gcaactgggc aggttattgc c	caaaccagaa

2821	tttaccactt	gacatcacat	tgcaaagccc	aactggtgct	ggacctttcc	caccaatcag
2881	aaacagtagt	ccctactcag	tgatacctca	gccaggaatg	atgggtaatc	aagggatgat
2941	aggaaaccaa	ggaaatttag	ggaacagtag	cacaggaatg	attggtaaca	gtgcttctcg
3001	gcctactatg	ccatctggag	aatgggcacc	gcagagttcg	gctgtgagag	tcacctgtgc
3061	tgctaccacc	agtgccatga	accggccagt	ccaaggaggt	atgattcgga	acccagcagc
3121	cagcatcccc	atgaggccca	gcagccagcc	tggccaaaga	cagacgcttc	agtctcaggt
3181	catgaatata	gggccatctg	aattagagat	gaacatgggg	ggacctcagt	atagccaaca
3241	acaagctcct	ccaaatcaga	ctgccccatg	gcctgaaagc	atcctgccta	tagaccaggc
3301	gtcttttgcc	agccaaaaca	ggcagccatt	tggcagttct	ccagatgact	tgctatgtcc
3361	acatcctgca	gctgagtctc	cgagtgatga	gggagctctc	ctggaccagc	tgtatctggc
3421	cttgcggaat	tttgatggcc	tggaggagat	tgatagagcc	ttaggaatac	ccgaactggt
3481	cagccagagc	caagcagtag	atccagaaca	gttctcaagt	caggattcca	acatcatgct
3541	ggagcagaag	gcgcccgttt	tcccacagca	gtatgcatct	caggcacaaa	tggcccaggg
3601	tagctattct	cccatgcaag	atccaaactt	tcacaccatg	ggacagcggc	ctagttatgc
3661	cacactccgt	atgcagccca	gaccgggcct	caggcccacg	ggcctagtgc	agaaccagcc
3721	aaatcaacta	agacttcaac	ttcagcatcg	cctccaagca	cagcagaatc	gccagccact
3781	tatgaatcaa	atcagcaatg	tttccaatgt	gaacttgact	ctgaggcctg	gagtaccaac
3841	acaggcacct	attaatgcac	agatgctggc	ccagagacag	agggaaatcc	tgaaccagca
3901	tcttcgacag	agacaaatgc	atcagcaaca	gcaagttcag	caacgaactt	tgatgatgag
3961	aggacaaggg	ttgaatatga	caccaagcat	ggtggctcct	agtggtatgc	cagcaactat
4021	gagcaaccct	cggattcccc	aggcaaatgc	acagcagttt	ccatttcctc	caaactacgg
4081	aataagtcag	caacctgatc	caggctttac	tggggctacg	actccccaga	gcccacttat
4141	gtcaccccga	atggcacata	cacagagtcc	catgatgcaa	cagtctcagg	ccaacccagc
4201	ctatcaggcc	ccctccgaca	taaatggatg	ggcgcagggg	aacatgggcg	gaaacagcat
4261	gttttcccag	cagtccccac	cacactttgg	gcagcaagca	aacaccagca	tgtacagtaa
4321	caacatgaac	atcaatgtgt	ccatggcgac	caacacaggt	ggcatgagca	gcatgaacca
4381	gatgacagga	cagatcagca	tgacctcagt	gacctccgtg	cctacgtcag	ggctgtcctc
4441	catgggtccc	gagcaggtta	atgatcctgc	tctgagggga	ggcaacctgt	tcccaaacca
4501	gctgcctgga	atggatatga	ttaagcagga	gggagacaca	acacggaaat	attgctgaca
4561	ctgctgaagc	cagttgcttc	ttcagctgac	cgggctcact	tgctcaaaac	acttccagtc
4621	tggagagctg	tgtctatttg	tttcaaccca	actgacctgc	cagccggttc	tgctagagca
4681	gacaggcctg	gccctggttc	ccagggtggc	gtccactcgg	ctgtggcagg	aggagctgcc
4741	tcttctcttg	acagtctgaa	gctcgcatcc	agacagtcgc	tcagtctgtt	cactgcattc
4801	accttagtgc	aacttagatc	tctcctgcaa	aagtaaatgt	tgacaggcaa	atttcatacc
	catgtcagat		_			
4921	tcctgttcgg	ttccagacac	tggtttcttg	ctttgttttc	cctggctaac	agtctagtgc
4981	aaaagattaa	gattttatct	gggggaaaga	aaagaatttt	ttaaaaaatt	aaactaaaga
5041	tgttttaagc	taaagcctga	atttgggatg	gaagcaggac	agacaccgtg	gacagcgctg
5101	tatttacaga	cacacccagt	gcgtgaagac	caacaaagtc	acagtcgtat	ctctagaaag
5161	ctctaaagac	catgttggaa	agagtctcca	gttactgaac	agatgaaaag	gagcctgtga
	gagggctgtt	_		_	_	~
5281	gttcacctga	atcatgaatt	gagaagaaat	aattttcatt	tctaaattaa	gtccctttta
	gtttgatcag		-	_	-	
	ccctctctca					
	cagaaaccag				_	_
	agtattgaat	_	_			
	cttgatacca					
	aactctggcc					_
	gacagattta					
	gaagttttgt			-	_	-
5821	gaacttatgt	gttttaattg	tataattttt	gtgaggtata	catattgtgg	aattgactca

5881 aaaatgaggt acttcagtat taaattagat atcttcatag caatgtctc taaaggtgtt 5941 ttgtaaagga tatcaatgcc ttgattagac ctaatttgta gacttaagac tttttattt 6001 ctaaaccttg tgattctgct tataagtcat ttatctaatc tatatgatat gcagccgctg 6061 taggaaccaa ttcttgattt ttatatgtt atatctttc ttaatgacc ttagaaagac 6121 tacatgttac taagcaggcc acttttatgg ttgtttt

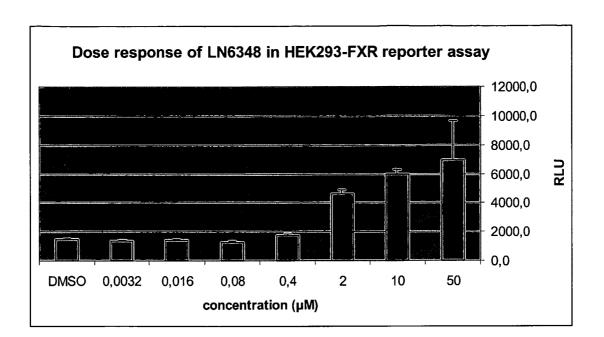


FIG 5 B

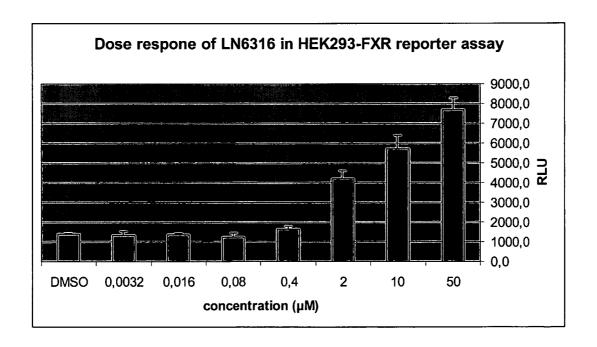


Fig 5 C

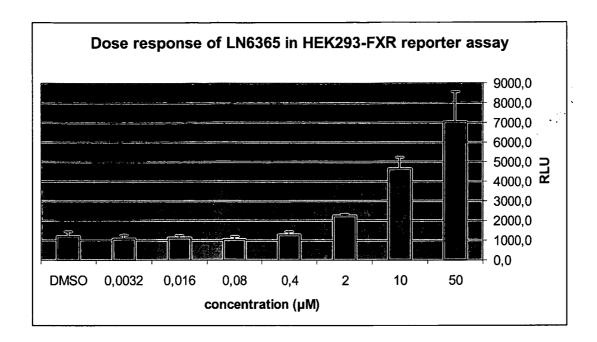
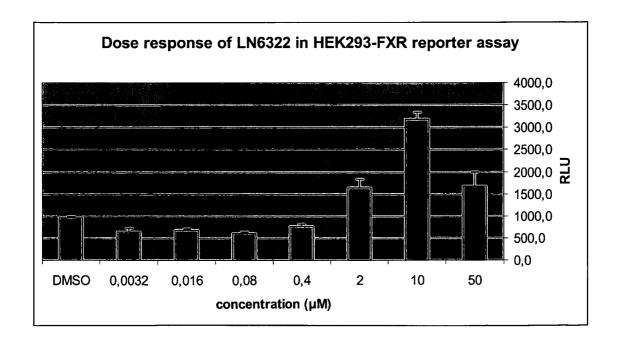


Fig 5 D



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/25438

	SSIFICATION OF SUBJECT MATTER A61K 31/426; C07D 277/40, 277/42, 277/44, 277/	46	
US CL :	:514/370; 548/194, 190		
	to International Patent Classification (IPC) or to bot	h national classification and IPC	
	DS SEARCHED		
Minimum d	ocumentation searched (classification system followe	d by classification symbols)	
U.S. :	<i>5</i> 14/ 3 70; <i>5</i> 48/194, 190		
Documentat	tion searched other than minimum documentation to	o the extent that such documents are in	ncluded in the fields
searched			
Electronic d	lata base consulted during the international search (1	name of data base and, where practicable	e, search terms used)
STN CAS	ONLINE		
C. DOC	UMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where ap	ppropriate, of the relevant passages	Relevant to claim No.
X	US 5,856,347 A (HASHIGUCHI I	ET AL.) 05 January 1999	1, 3, 6, 7, 11, 14,
	(05/01/99), see entire document, e	· · · · · · · · · · · · · · · · · · ·	15, 17-19
	compounds 1, 76, 112, 113, 171, etc.	in Table 1.	
		17.00 100 0 1 100	1 10 10
X	JP 09-235278 A (TOYAMA CHEMIC.	-	1, 18, 19
	(09/09/97), see entire document, espec	cially page 2 and compounds	
	46 and 72.		
X	US 5,602,132 A (ROGER ET AL.) 1	1 February 1997 (11/02/97).	1, 18-21, 24
71	see entire document, especially column	• '	1, 10 21, 21
	columns 45, 46, etc.		
	, , ,		
	·		
X Furtl	her documents are listed in the continuation of Box	C. See patent family annex.	
•	ecial categories of cited documents:	"I" later document published after the inte date and not in conflict with the appl	
	cument defining the general state of the art which is not considered be of particular relevance	the principle or theory underlying the	invention
"E" ear	rlier document published on or after the international filing date	"X" document of particular relevance; the considered novel or cannot be consider	
	cument which may throw doubts on priority claim(s) or which is ed to establish the publication date of another citation or other	when the document is taken alone	
вре	ocial reason (as specified)	"Y" document of particular relevance; the considered to involve an inventive step	when the document is combined
	cument referring to an oral disclosure, use, exhibition or other cans	with one or more other such docum obvious to a person skilled in the art	enus, such combination being
	cument published prior to the international filing date but later an the priority date claimed	"&" document member of the same patent	family
Date of the	actual completion of the international search	Date of mailing of the international se	arch report
09 SEPT	EMBER 2002	16 DEC 2002	
Name and r	nailing address of the ISA/US	Authorized officer	
Commission Box PCT	ner of Patents and Trademarks	Mouther Junioric	Las l
Washingto	n, D.C. 20231	LAURA L. STOCKTON	/- \
Facsimile N	lo. (703) 305-3230	Telephone No. (703) 308-1235	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/25438

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	US 5,464,847 A (COURTEMANCHE ET AL.) 07 November 1995 (07/11/95), see entire document, especially columns 2-3; and compounds in columns 45, 46, etc.	1, 18-21, 24