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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 August 2003 (14.08.2003)

PCT

(10) International Publication Number
WO 03/066858 A1

(51) International Patent Classification[7]: C12N 15/11, C07K 14/435, A01K 67/027

(21) International Application Number: PCT/GB03/00475

(22) International Filing Date: 4 February 2003 (04.02.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0202543.5 4 February 2002 (04.02.2002) GB

(71) Applicant *(for all designated States except US)*: **ISIS INNOVATION LIMITED** [GB/GB]; Ewert House, Ewert Place, Summertown, Oxford, Oxfordshire OX2 7SG (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **WOOD, Matthew, John, Andrew** [IE/GB]; University Offices, Wellington Square, Oxford OX1 2JD (GB). **EVERATT, Sarah, Louise** [GB/GB]; University Offices, Wellington Square, Oxford OX1 2JD (GB).

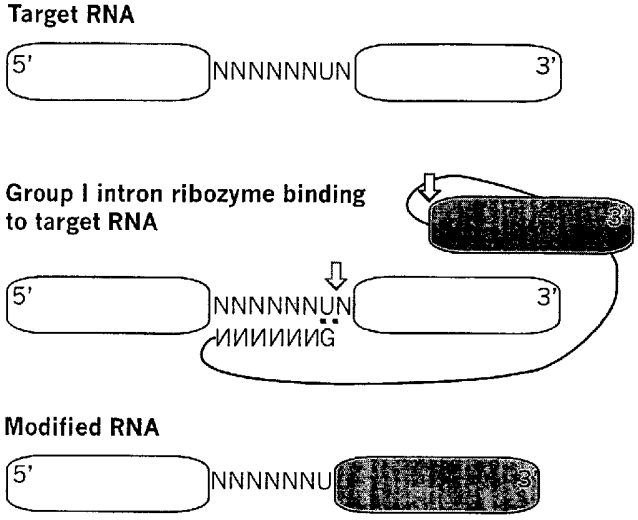
(74) Agent: **LORD, Hilton, David**; Marks & Clerk, 57-60 Lincoln's Inn Fields, London WC2A 3LS (GB).

(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, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, 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),

[Continued on next page]

(54) Title: TRANS-SPLICING RIBOZYMES TRANSFERRING ELEMENTS DERIVED FROM SEQUENCES ENCODING FLUORESCENT PROTEINS



G = G-U wobble in P1 substrate helix

⇩ = 5' splice site

⇩ = 3' splice site

(57) Abstract: A catalytic polynucleic acid molecule is capable of catalysing the cleavage of a selected mRNA into 5' and 3' cleavage products, and replacing the 3' cleavage product with a coding sequence encoding all or part of a fluorescent protein, the resulting mRNA being translatable to express a fluorescent protein, thereby to establish the success or otherwise of methods of delivering ribozymes to sites expressing target mRNA.

WO 03/066858 A1

European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, 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.

TRANS-SPLICING RIBOZYMES TRANSFERRING ELEMENTS DERIVED FROM SEQUENCES ENCODING FLUORESCENT PROTEINS

The present invention relates to model systems, and more particularly to model systems for transgenic work in animals.

Since the discovery of ribozymes there has been significant progress in the field of RNA therapeutics. Ribozymes, antisense strategies and, more recently, deoxyribozymes, or DNAzymes (Santoro *et al.* (1997) Proc Nat. Acad. Sci. USA,. **94**(9), 4262-6) have all been employed as powerful tools to interfere with gene expression at the RNA level.

Ribozymes are RNA molecules capable of catalysis, while DNAzymes are similar to ribozymes but man-made and comprise deoxyribonucleotides. In antisense strategies, antisense RNA binds to mRNA in a sequence dependent manner, the cell destroys the resulting double stranded RNA and, therefore, the expression of the gene is reduced. All of these strategies exploit complementary Watson-Crick base pairing in order to target the RNA in a sequence specific manner and, in this way, these technologies can be used to target genes for down-regulation or repair.

Hammerhead ribozymes are small catalytic RNA molecules that have the ability to cleave RNA. Hammerhead ribozymes were originally found in plant viroids and virusoids where they form a hammerhead structure to self-cleave during the replication process (Forster *et al.,* (1987) Cold Spring Harbor Symp. Quant. Biol., **52,** 249-59). Hammerhead ribozymes have been engineered to cleave in *trans*, *i.e.* to cleave other RNA molecules, in a sequence specific manner.

The hammerhead ribozyme has a three-stem structure. Stem II is the conserved catalytic core of the ribozyme and is composed of 13 nucleotides, while stems I and III can be designed to bind any nucleotide sequence. The only sequence requirement for cleavage to occur is an NUX motif in the target molecule, where N can be any nucleotide and X can be any nucleotide except for guanine. The ribozyme binds the target *via* its two binding arms and cleaves the substrate. After dissociation, it is free to react with further substrates. Mg^2+ plays a direct role in the cleavage

reaction and must be present for cleavage to occur. Accompanying Figure 3 is a schematic diagram of a hammerhead ribozyme cleaving in *trans*.

Hammerhead ribozymes can be delivered using plasmid or viral based vectors or, more recently, as chemically stabilised molecules, and can, therefore, be delivered *in vivo*. However, although hammerhead ribozymes have been shown to down-regulate gene expression *in vivo*, the down-regulation seen is often quite mild and does not persist for a length of time suitable for useful therapy.

The group I intron ribozyme from *Tetrahymena thermophilia* was the first RNA to be shown to exhibit catalytic activity (Cech *et al.* (1981) Cell, **27**(3 Pt 2), 487-96). The group I intron ribozyme is more complex than the self-cleaving hammerhead ribozymes and is significantly larger. Group I intron ribozymes are found in a variety of organisms and act to splice in *cis*, that is to perform a self splicing reaction. This reaction occurs in two stages whereby 3' and 5' exons flanking the ribozyme sequence are ligated, and a mature RNA transcript is generated. The group I intron has been used in the laboratory to splice in *trans*, that is to perform a splicing reaction on an RNA molecule distinct from itself. The accompanying Figure 4 is a schematic representation of a group I intron ribozyme-mediated modification of a trans-RNA target.

The group I intron ribozyme from *Tetrahymena thermophilia* is 421nt long with a universally conserved central catalytic core of approximately 200nt surrounded by several less conserved peripheral elements. The group I intron ribozyme has been shown to fold into a small globular structure. The splicing reaction is dependent upon an endogenous guanosine molecule and magnesium ions (*c.f.* Doherty *et al.* (2001) Ann. Rev. Biophys. Biomol. Struct., **30,** 457-75).

Although experiments have shown that up to 50% of target RNA's can be *trans*-spliced in mammalian cells (Jones *et al.*(1997) Nat. Biotechnol., **15**(9), 902-5), the ability to repair RNA's in a clinically relevant setting by *trans*-splicing appears to meet with considerably lower success rates. However, combining group I intron ribozymes that recognise different RNA target sites in the target RNA appears to enhance the efficiency of the procedure (Lan *et al.*, (2000) Mol Ther., **2**(3), 245-55).

In the context of the present invention, the term "repair" is taken to mean altering an mRNA to result in the expression of a different polypeptide from that encoded by the original. "Silencing" and its associated terms are taken to indicate that an mRNA has been down-regulated or prevented from expressing its polypeptide.

DNAzymes are catalytic DNA molecules that cleave RNA in a sequence specific manner. DNAzymes are not naturally occurring molecules, and are usually generated by *in vitro* selection procedures to produce active molecules which generally exceed the catalytic efficiency of ribozymes.

The DNAzyme binds the target *via* the two target binding arms and cleaves between an unpaired purine (A, G) and a paired pyrimidine (C, U) residue. Through the sequence of the target binding arms, the DNAzyme can be individually designed to cleave virtually any target RNA. Effective DNAzymes with binding arms ranging from 7 to 15 nucleotides have been reported (Kurreck *et al.* (2001) J Biol. Chem., **20,** 20). Magnesium ions play a direct role in the cleavage reaction and must be present for cleavage to occur.

DNAzymes are more stable than ribozymes, and stability can be further increased by chemical modification of bases. DNAzymes have a much higher catalytic activity than ribozymes. Delivery of chemically synthesised DNAzymes is generally by a transfection-based system, which has tended to limit the use of these molecules therapeutically, but other methods for their delivery are under investigation, and a DNA expression vector has been reported. As the time from design to synthesis of usable DNAzymes is usually only a matter of days, these molecules are useful for *in vitro* and cell culture work, unlike hammerhead ribozymes, which must be designed, cloned and expressed.

Investigations into the mechanisms of these molecules and the optimisation of their expression, delivery, specificity and action are hampered by the lack of an easily quantifiable model system. In particular, expression of therapeutic nucleic acids is problematic, and there are several obstacles that have to be overcome in order to be able to use ribozymes and DNAzymes as a tool for gene therapy:

the delivery of the ribozymes and DNAzymes to the appropriate tissues;

the expression of ribozymes and DNAzymes in these tissues at therapeutically relevant levels;

the specificity of the ribozyme and DNAzyme;

the persistence of expression of the ribozyme or activity of the DNAzyme; and

the co-localisation of the catalytic nucleic acid and the target RNA.

Green Fluorescent Protein has been widely exploited by scientists as an expression system (Tsien (1998) Ann. Rev. Biochem., **67,** 509-44) as it is easy to quantify and requires no additional steps to visualise. Mutagenesis of the original sequence from the jellyfish *Aequorea victoria* has resulted in a more stable and brighter protein EGFP (enhanced green fluorescent protein). Further mutagenesis of the EGFP gene has produced the colour variants, EBFP, ECFP and EYFP (Blue, Cyan and Yellow fluorescent proteins respectively – *c.f.* Cormack *et al., (1996)* Gene, **173**(1), 33-8).

We have now found that it is possible to change mRNA encoding one fluorescent protein into mRNA encoding a different fluorescent protein by the use of *trans*-splicing ribozymes, thereby providing a useful marker, for example, to confirm the presence of particular ribozymes.

Thus, in a first aspect, there is provided a catalytic polynucleic acid molecule, characterised in that the molecule is capable of catalysing the cleavage of a selected mRNA and replacing the 3' cleavage product with a coding sequence encoding all or part of a fluorescent protein, the resulting mRNA being translatable to express a fluorescent protein.

In a preferred aspect, the mRNA encodes a first fluorescent protein and the catalytic molecule encodes a second fluorescent protein, whereby replacement of the 3' cleavage product with a coding sequence encoding part of a second fluorescent protein results in mRNA translatable to express a fluorescent protein different from that encoded by the uncleaved mRNA.

Thus, in a preferred aspect, there is provided a catalytic polynucleic acid molecule, characterised in that the molecule is capable of catalysing the cleavage of

an mRNA encoding a fluorescent protein and replacing the 3' cleavage product with a coding sequence encoding part of a second fluorescent protein, the resulting mRNA being translatable to express a fluorescent protein different from that encoded by the uncleaved mRNA.

It will be appreciated that the polynucleic acid molecule will generally be a ribozyme or a DNAzyme, and references herein to ribozymes or DNAzymes include references to all such molecules, unless otherwise apparent. Typical ribozymes include hammerhead ribozymes and group I intron ribozymes. More particularly, the catalytic molecule should be able to introduce a coding sequence into an mRNA.

Preferred ribozymes of the present invention are capable of *trans*-splicing target mRNA molecules, and this is generally effected by the ribozyme of the invention comprising a 3' intron encoding the C terminus of a fluorescent protein. This fluorescent protein will be different from the fluorescent protein that is the target of the ribozyme, and will, effectively, be exchanged with the 3' cleavage product of the target mRNA to generate a different fluorescent protein.

The fluorescent protein encoded by the intron may be all or part of the protein, but must generally include at least the C terminus thereof, as the intron will generally replace the 3' end of the mRNA. It is not essential that the replaced portion be the same length as that replacing it, provided that the resulting mRNA encodes a protein that is detectable by fluorescence analysis, such as by instrument or by the naked eye.

It is generally preferred that the protein encoded by the target mRNA is also a fluorescent protein, and that both proteins belong to the eGFP family. The resulting protein preferably belongs to this family, also, and the family is discussed more, below. However, especially where one protein differs from another by more than a couple of bases in the mRNA, it is also possible that the resulting protein is a fluorescing hybrid.

In general, it will be appreciated that ribozymes of the present invention may be highly specific for a selected mRNA, or may be selective for two or more fluorescent protein mRNA's, for example. Unless otherwise indicated, reference to

the target mRNA herein as encoding a fluorescent protein includes the possibility that the target encodes another protein. It is generally preferred that the target encodes a fluorescent protein, but it may be desired to evaluate whether a target mRNA can be successfully repaired by a fluorescent protein, prior to using a therapeutic repair, for example.

Ribozymes of the present invention may be used to induce colour changes in selected cells, and it is possible that target cells will express more than one mRNA as fluorescent proteins, in which case, it may be desirable to provide ribozymes specific for one particular fluorescent protein, such as the yellow fluorescent protein (EYFP).

However, where target cells only express one fluorescent protein, then there will generally be no requirement to provide ribozymes specific for one particular mRNA.

In general, it is preferred that the ribozyme be able to target all fluorescent protein mRNA's, as those ribozymes which are specific for selected fluorescent protein mRNA's must target that site on the mRNA which is specific for that mRNA and, if this is on the 5' side of the cleavage site, then the sequence will remain in the final mRNA. This is not important, provided that the resulting mRNA encodes a different coloured fluorescent protein from the original.

It is not essential that the mRNA intron provide the sequence encoding part of a specific fluorescent protein, provided that the resulting mRNA can be transcribed into a fluorescent protein having different properties from the original. In practice, the intron will generally provide the latter portion of the mRNA that differs from one fluorescent protein to the next. In the case of EBFP, the difference between this and EGFP is only two bases and, provided that the intron comprised in the ribozyme serves to replace these bases in the original mRNA, then the resulting fluorescent protein will be changed from blue to green, or *vice versa*.

Similar considerations apply to the other fluorescent proteins, such as yellow fluorescent protein and cyan fluorescent protein, although the number of base changes is slightly greater, by comparison with EGFP. In these cases, it is generally preferred

that all of the different bases be encoded by the intron, but it may be sufficient that only some of the bases be encoded in order to achieve a change in the fluorescent properties of the expressed protein.

The fluorescent proteins are especially useful in the present invention as, although the differences in emission and excitation of these proteins are quite marked, as shown in the accompanying Figure 1, the differences between the sequences of the different coloured proteins, at the genetic level, are very small, and usually only amount to a few bases, as demonstrated in Figure 2.

For ease of reference, although the sequence variations are shown in bold and are underlined, also, the sequence differences, by comparison with the EGFP sequence, are easily summarised:

Blue	Cyan	Yellow
204	205–6	197
444	447	199-201
	468-9	210
	498-9	222-223
		617-618

Another coloured protein has been isolated from *Discosoma sp.* called dsRed, but has little sequence similarity to the other coloured proteins, also as shown in Figure 2.

The ease of discriminating between the coloured proteins, combined with the similarities between the coding sequences provides a convenient system for investigating and optimising function, delivery, specificity and expression of catalytic nucleic acids. The availability of the EGFP transgenic mouse (Pratt *et al.*, (2000) Dev Biol., **228**(1), 19-28) provides a useful starting test subject, and allows experiments to be carried out *in vivo*.

SEQ ID NO. 1 is the nucleotide sequence of EGFP mRNA, while the protein encoded by this sequence is SEQ ID NO. 2. Similarly, SEQ ID NOS. 3 and 4 are the nucleotide and protein sequences of EBFP, SEQ ID NOS. 5 and 6 are the nucleotide

and protein sequences of ECFP, and SEQ ID NOS. 7 and 8 are the nucleotide and protein sequences of EYFP. SEQ ID NO. 9 is the nucleotide sequence of dsRED.

SEQ ID NO. 10 is the mRNA sequence and SEQ ID NO. 11 is the ribozyme sequence of Figure 3. SEQ ID NOS. 12 and 13 are the mRNA sequence and DNAzyme sequence, respectively, of Figure 5.

SEQ ID NOS. 14, 15 and 16 are the nucleotide sequences of Figures 7, 8 and 9, respectively.

SEQ ID NO. 17 is the GCFPRZ1 9 base pair internal guide sequence (IGS). SEQ ID NO. 18 is the 9 base pair internal guide sequence (IGS) of GCFPRZ2. SEQ ID NOS. 19 to 21 are as defined in Example 2.

The present invention provides means by which the delivery of ribozymes to a target tissue may be established. In a preferred embodiment, a target tissue expresses a selected fluorescent protein, a ribozyme of the invention is then delivered to the tissue and the success of delivery established by ascertaining the presence of a modified fluorescent protein. For example, the EGFP mouse provides an excellent model for testing different forms of ribozyme delivery. There is no restriction on the form of delivery tested. The only consideration is to whether the tissue begins to express the fluorescent protein encoded by the test ribozyme, and which can only be expressed if delivery of ribozyme to the tissue was successful.

It will be appreciated that any suitable method may be employed to enable tissue to express a fluorescent protein, or other protein, *via* appropriate mRNA. Suitable transformation methods are well known in the art, and may involve cloning into cell lines using electroporation or by using appropriate vectors, including plasmids and other vehicles, such as viruses. The resulting cells, if eggs, for example, may be grown into whole organisms, or may be replicated into cell lines, for example, if they are another form of tissue. In any event, suitable techniques for obtaining cell lines or tissues expressing the desired mRNA are well known in the art.

Further, there is no restriction as to the specificity sought for the delivery system under test. On a larger scale, it may be desired to deliver ribozyme to the entire body, in which case an animal constitutively expressing the target mRNA in all tissues is desirable for testing. Such an animal may also be used for anything more specific, such as targeting one organ but not another, in which case the absence of the transformed mRNA can also be tested by absence of the fluorescent protein that would necessarily result from such a transformant. Such specificity can descend all the way down to organelles, as discussed below.

Thus, the present invention allows the determination of success or failure of selected delivery systems for ribozymes and DNAzymes, by positive identification of the results. It is readily feasible to test any number of parameters, including different promoters and delivery methods, such as tissue specific promoters, strong promoters, weak promoters, systemic delivery, retrograde transport and the like.

By using the system of the present invention, the active presence of a ribozyme of the invention can be measured by the presence of the second fluorescent protein. For example, a tissue may express EGFP and the ribozyme carries the intron for changing EGFP into EBFP. If the ribozyme is successfully delivered to the tissue, then it will be possible to measure the presence of expressed EBFP.

Whilst it is possible that successful delivery of ribozymes of the invention will result in substantially 100% conversion of a first mRNA into a second mRNA, it is also likely that not all of the mRNA present in the cell will be converted, so that a mixture of fluorescent proteins will be apparent in the affected tissue. However, by using appropriate fluorescent measuring means, this presents no problems to the person skilled in the art. Such results also provide valuable information regarding the system under test, as it may also provide information as to the level of gene silencing that might be expected by using the delivery system in question, or even the IGS being used, if the target mRNA is also the end target for repair. In such case, the target mRNA may also temporarily carry a fluorescent marker, which can be transformed by a ribozyme of the invention. Resultant fluorescent patterns will then provide substantial levels of information on the delivery system tested, and to the efficacy of the IGS.

The present invention envisages the use of ribozymes to change tissue colour, but it is generally preferred that the method of the present invention be employed to assay the effectiveness of means for delivering ribozymes to tissues *in vivo*. The art is replete with delivery systems that work in theory but which, often for unknown reasons, do not work in practice. The advantage of the present invention is that, by determining the success of delivery of ribozymes of the invention, such delivery systems may then be used to deliver any other ribozymes that it may be desired to deliver to the site in question.

The present invention also envisages the use of ribozymes of the invention as markers for the success of the delivery of other ribozymes. For example, a tissue may express a protein that does not normally fluoresce but that, when subjected to ribozymes of the present invention, expresses fluorescent proteins that may be detected. Once it has been confirmed that ribozymes of the invention delivered in parallel with the desired ribozymes are indeed reaching the target tissue, then delivery of the ribozymes of the invention can be discontinued and re-continued at any time that it is desired to confirm the success of delivery of ribozyme to the tissue. A suitable target protein to establish successful delivery of a ribozyme is albumin, for example. In general, where a protein is to be labelled in such a manner, it is preferred that it be relatively abundant and not critical from moment to moment, so that it is generally preferred to avoid sparsely expressed enzymes, for example.

The present invention further provides a system or method for determining the likely success of delivering a polynucleic acid molecule to a target tissue such that the polynucleic acid molecule is expressed in said tissue, wherein the target tissue expresses a selected mRNA, said system comprising delivering a catalytic polynucleic acid molecule to said tissue in a manner in which it is desired to deliver said polynucleic acid molecule, and wherein the catalytic polynucleic acid molecule is capable of catalysing the cleavage of the selected mRNA into 5' and 3' cleavage products, and replacing the 3' cleavage product with a coding sequence encoding all or part of a fluorescent protein, the resulting mRNA being translatable to express a fluorescent protein, success being determined by assaying the presence of fluorescent protein encoded by said resulting mRNA. Thus, if a ribozyme can successfully be

delivered to a cell or tissue and be expressed therein, then it is extremely likely that another polynucleotide sequence, and especially another ribozyme, such as a repair or silencing ribozyme, can also be successfully delivered and expressed. This allows the skilled person to establish the desirability of using a given delivery method quickly and easily.

Thus, the issues afflicting the use of ribozymes and DNAzymes can readily be investigated using the fluorescent protein, or FP, system described herein both generally, using the system to improve delivery vectors, testing out the kinetics of modified hammerhead species, and specifically, using the system to perfect delivery to a particular cell type, tissue or organ before using the therapeutic ribozyme of interest.

The system of the present invention, such as using a *trans*-splicing group I intron ribozyme to splice EBFP sequence onto EGFP mRNA, causing an easily quantifiable colour change from green to blue, for example, enables researchers to improve the efficiency, specificity and delivery of ribozymes and deoxyribozymes both generally and specifically.

The systems of the present invention permit the optimisation of delivery of these molecules using non-viral methods, for example, and will greatly facilitate the work being undertaken in this area, especially given that the delivery and persistence of expression of DNAzymes are currently considered to be the only major obstacles to their therapeutic use.

Stably transfected cell lines expressing EGFP, EYFP, EBFP and ECFP form a preferred embodiment of the present invention, as do host animals expressing these proteins constitutively. It is generally preferred that any line or animal express only one fluorescent protein. It will also be appreciated that the invention extends to such animals and cell lines expressing other fluorescent variants of these proteins. While EGFP, EYFP, EBFP and ECFP are the preferred proteins of the present invention, the invention also extends to other fluorescent members of this family that may result from any future genetic alterations to any one of the family. The family is generally characterised by a substantial homology of the first 195 bases from the 5' end of the

mRNA, and the invention extends to any fluorescent protein encoded by an mRNA having the first 195 bases of its base sequence of identical to bases 1 to 195 of SEQ ID NO. 1, or differing therefrom by no more than 5%, and preferably no more than 2%, particularly preferably 1% or less.

The effect of ribozymes and DNAzymes on these stably transfected cell lines can be analysed in a quantifiable manner using the FACS machine (Fluorescence Activated Cell Sorter), for example. Using such a machine, the actual number of fluorescent cells and the intensity of the fluorescence is readily determined (Cormack *et al., (1996)* Gene, **173**(1), 33-8).

RT-PCR analysis of the FP RNA extracted from cells and tissues and, similarly, Western blot analysis of the repaired FP protein may be used, for example, to provide confirmation of results obtained by such processes.

Real-time PCR using the TaqMan™ machine, for example, may be employed to provide absolute quantification and analysis of the action of the catalytic nucleic acid molecules of the present invention.

Group I intron ribozymes may be expressed at the desired location by suitable mammalian expression vectors, such as $PCDNA_3$, into which they have been cloned, together with suitable control and expression sequences, such as the SV40 or adenovirus promoters, together with an initiation sequence, for example.

Particular examples of suitable expression vectors include inactivated or attenuated retroviruses, or retroviruses having no significant pathological effect on the recipient, but which are capable of infecting cells in the recipient, adenovirus, adeno-associated virus (AAV) and lentivirus, which may be used to express hammerhead and group I intron ribozymes of the invention, for example. These molecules can be cloned into the expression vectors, such as adenovirus, under the control of a mammalian promoter specific for the synthesis of large amounts of small RNA, such as the U2 snRNA promoter (Ares *et al.* (1985) Mol Cell Biol., **5**(7), 1560-70).

DNAzymes may be stabilised by the use of internal phosphothioate bonds. Such molecules are less susceptible to degradation, and can be delivered to the target site with less likelihood of degradation, and will continue to be effective for longer. A suitable DNAzyme is the 10-23 DNAzyme (Santoro *et al.* (1997) Proc Nat. Acad. Sci. U S A,. **94**(9), 4262-6). The 10-23 DNAzyme consists of a catalytic domain of 15 nucleotides with flanking target-binding arms, as illustrated in Figure 3 with respect to hammerhead ribozymes.

Localisation of expression of the fluorescent proteins is possible using Living Colours ™ subcellular localisation vectors, for example, which localise fluorescence to specific organelles or structures within living cells in order to visualise biological processes as they occur. Using such vectors and expression systems enables investigation into the ability of delivery systems to co-localise the ribozymes and DNAzymes with these localised proteins. Delivery of catalytic nucleic acids to specific compartments of the cell in order to enhance the therapeutic activity can then be optimised (Sullenger (1995) Appl Biochem Biotechnol., **54**(1-3), 57-61).

It will be appreciated that the present invention enables the detailed construction of such catalytic molecules to be investigated and analysed in detail. Using the colour specific catalytic nucleic acids on the other colour variants, a panel of DNAzymes and ribozymes with mismatches on one of the binding arms, on both of the binding arms and mismatches in different positions along the binding arms can be created. Such a panel can then be used to provide a guide to required specificity of sequence for ribozymes to be as selective as required.

It will be appreciated that multiple ribozymes targeted to different sites on the target RNA can be more efficacious, and that this can equally apply to ribozymes of the invention. Such a use can also be exploited to help where multiple therapeutic ribozymes are desired or required, to help determine the efficacy of delivery.

Ribozymes which can benefit from the system of the present invention may be any that may be used for therapy, or otherwise, in the human or animal body. In general, such ribozymes may be used to either repair or silence gene expression products, and are applicable, for example, to genetic diseases, such as Huntingdon's,

or sickle cell anaemia, and cancer, such as leukaemia, as well as to infectious diseases, such as HIV, for example. In general, *trans*-splicing ribozymes are applicable where any other ribozyme might be used, so that any disease is a potential target.

Progress in the field of using group I intron ribozymes as tools for gene therapy has been very slow considering their potential for the treatment of dominant negative genetic disease. The system of the present invention enables these molecules to target any organ, tissue or cell type within an FP mouse, for example, and allows optimisation of delivery of the group I intron ribozyme. Even very small amounts of splicing can be visualised using antibodies that are readily available against EGFP and EBFP, for example.

Vectors are available that allow an EGFP tag to be inserted at either the N-terminus or C-terminus of a protein of choice. Making such an EGFP/target molecule fusion protein (and hence EGFP/target molecule fusion RNA) and targeting this RNA using a catalytic nucleic acid of the present invention can be used to efficiently co-localise ribozyme/DNAzyme and target, as the resulting hybrid would be readily detectable.

In a preferred embodiment, the present invention provides a system or method for enabling the selection of the best IGS, or combination of IGS's for use in repairing a given RNA. The system comprises a panel of ribozymes capable of splicing a full length FP into an RNA, the panel having a ribozyme having all possible variations, or a majority of all possible variations, of the IGS. Thus, for a 6bp IGS this would be 4^6, or 4096. While each could be trialled independently, it would be easier to test a pool of each, such as on a 96 well plate, each well containing 45 different group I intron ribozymes of known IGS sequence, all capable of splicing full length EGFP, for example. Target RNA, transfection reagent and cells can then all be added to the wells and allowed to incubate overnight. FACS analysis can be used to determine which wells contain EGFP; these wells must therefore contain a ribozyme capable of splicing EGFP onto the target RNA in a sequence specific manner. A second 96 well plate containing only those 45 ribozymes present in that well of the initial plate that contained fluorescence, could then be plated out individually (and in duplicate), and

treated in the same way as the first, by addition of target RNA, transfection reagent and cells. The plate can be incubated overnight. Any well containing the individual ribozyme capable of splicing EGFP onto the target RNA will contain fluorescence when analysed subsequently analysed using FACS or simple visual observation under the microscope, for example. An optimal IGS can thus be readily detected.

It will be appreciated that the fluorescent proteins referred to herein need not, unless otherwise specifically required, belong to the eGFP family. For example, this includes both dsRED and the original GFP.

The catalytic molecules of the invention are also useful in many other situations. For example, they find particular use in confocal microscopy, which can be used to track the movements of proteins within a cell, for example. whilst fluorescent protein markers have been used for this process previously, this has been done at the genetic level. The advantage of the present invention is that it can be used at the transcriptional level, thereby bypassing problems associated with genetic modification.

The catalytic molecules of the invention are also useful in visualising the activity of two or more proteins within a cell. Tagging a protein with a dedicated ribozyme or, at least, a ribozyme not capable of tagging a second protein of interest, and then tagging the second protein with a second ribozyme not capable of tagging the first, and wherein the two ribozymes provide for the encoding of different fluorescent proteins, enables the activities and interactions of the two, or more, proteins to be tracked, *in vivo*, by a confocal microscope, for example, without affecting anything at the genetic level.

It will also be appreciated that the invention extends to hosts, such as animals, tissues and cell lines, transformed with genes encoding one or more of the ribozymes of the invention. Such hosts will not express the fluorescent protein encoded by the ribozyme, in the absence of a target mRNA, and can be used to establish methods of obtaining target mRNA expression, or co-expression in cells expressing the ribozymes of the invention, such as by breeding, fusion, or transformation.

Hosts transformed with such genes may have them under suitable control, so that they may be controllably expressed. Such controls may take the form of inducible promoters, such as the tetracycline on/off promoter, or tissue specific promoters, for example. In addition, it may be desirable to provide a marker, such as a resistance marker, in order to be able to select for transformed hosts without having to introduce target mRNA.

Typically, the nucleic acids of this invention have catalytic activity and in a preferred aspect, the target RNA encodes a fluorescent protein, and the catalytic nucleic acid cleaves the RNA to alter the fluorescence. Usually the target RNA is cleaved and the RNA fragments cannot then be used to encode the fluorescent protein. Examples of preferred nucleic acids of this invention include hammerhead ribozymes and DNAzymes. The sequence complementary to the target RNA binds through Watson-Crick base pairing and the catalytic core cleaves the target RNA molecule, thus down-regulating the protein, in this case reducing the amount of fluorescent protein (e.g. eGFP) seen in the cells/animal in an easily quantifiable manner.

The present invention will now be further illustrated by the following, non-limiting Examples and the accompanying Figures, in which:
Figure 1 shows the excitation and emission spectra for the proteins EGFP, EBFP, ECFP, EYFP and dsRED using the Living Colours ™ detection system;
Figure 2 is a sequence alignment of the proteins of Figure 1;
Figure 3 shows the hammerhead ribozyme-mediated *trans* cleavage of an mRNA molecule;
Figure 4 is a schematic representation of group I intron ribozyme mediated modification of a trans-RNA target;
Figure 5 shows the *trans* cleavage of an mRNA molecule by a 10-23 DNAzyme;
Figure 6 shows the areas of EGFP mRNA (bases 292-364) calculated to display high or low secondary structure using the M-Fold server;
Figures 7, 8 and 9 show the EGFP RNA sequence and the position of Hammerhead ribozymes, the position of Group I intron ribozyme IGS, and the position of DNAzymes within this sequence, respectively;
Figure 10 shows the results of a cleavage assay of target RNA against increasing concentration of hammerhead ribozyme;

Figure 11 shows the optimum magnesium concentration for DNAzyme cleavage of EGFP target RNA;

Figure 12 shows a time course cleavage assay of DNAzymes; and

Figures 13a and 13b show the effects of the addition of varying quantities of DNAzymes to EGFP transfected 293 HEK cells.

EXAMPLE 1

In order for any catalytic RNA or DNA molecule to cleave or repair a target RNA molecule, it must have access to the target site. RNA is single stranded and forms complex secondary structures that may sometimes render a particular region of the RNA inaccessible. The selection of target sites for any ribozyme has always been a problem for researchers. Computer programs such as Michael Zuker's M-Fold server (Zuker (1989) Methods Enzymol., **180,** 262-88) can predict, with a degree of accuracy, regions within the RNA sequence that have relatively low secondary structure and, therefore, would theoretically be accessible to therapeutic ribozymes and DNAzymes. Accordingly, this method was employed for EGFP, and the resulting graph is shown in Figure 6, which shows the actual M-Fold output for bases 292-364 of EGFP.

Hammerhead Ribozymes

Having predicted the structure of the EGFP RNA using the M-Fold server, all the possible NUX sites were rated for accessibility, taking into account that not only the NUX site needs to be accessible, but also the regions where the arms of ribozyme bind to the target RNA.

Having done this, 3 NUX sites were selected and three hammerhead ribozymes designed:

1. GFPHH1
2. GFPHH2
3. GFPHH3

These ribozymes are specific for any of the coloured proteins (except dsRED), as the ribozymes were all designed to target regions where the sequences of the coloured proteins are identical. Inactive versions of these ribozymes were also designed, identical to the active ribozyme, other than a C → G substitution at base 23 of the conserved catalytic core. Accompanying Figure 7 illustrates the EGFP RNA sequence showing the position of the Hammerhead ribozymes GFPHH1, GFPHH2 and GFPHH3.

Group I Intron Ribozymes

Two Group I intron ribozymes were designed to splice EBFP sequence onto EGFP RNA, which in cell culture and *in vivo* would result in a colour change from green to blue:

1. G-BFP1
2. G-BFP2

Accompanying Figure 8 illustrates the EGFP RNA sequence, showing the position of the Group I intron ribozyme internal guide sequences (IGS).

DNAzymes

Two DNAzymes were designed to target regions where the sequences of the coloured proteins are identical:

1. GFP1DZ
2. GFP2DZ

GFP1DZ was designed to the same sequence as the hammerhead ribozyme GFPHH2, although the cleavage site differed.

Another DNAzyme was designed to specifically cleave EGFP RNA but not EYFP RNA:

3. GFPYDZ

Another DNAzyme was designed to specifically cleave EBFP RNA but not EGFP RNA.

4. GFPB2DZ

A scrambled DNAzyme was designed as a control

5. DZ-S

This DNAzyme has no catalytic sequence or similarity to the EGFP sequence.

Accompanying Figure 9 illustrates the EGFP RNA sequence showing the position of the DNAzymes described above.

Cleavage of EGFP mRNA *in vitro* by Hammerhead Ribozymes and DNAzymes

Cleavage assays were performed *in vitro* using full length EGFP RNA (*in vitro* transcribed using Ambion MAXIscript™ kit) and either hammerhead ribozyme RNA (*in vitro* transcribed using Ambion MEGASHORTscript™ kit) or DNAzymes (commercially synthesised by SigmaGenosys).

Typically cleavage assays were performed in the presence of magnesium ions (0-50mM) and 50mM Tris-HCl pH 8.0 in a total volume of 10µl for 2 hours at 37°C.

The reactions were then run on a 6% urea-polyacrylamide gel, the cleavage products seen are the result of successful cleavage of the target RNA by the ribozyme or DNAzyme. The results are shown in Figure 10, which shows the cleavage assay with increasing concentration of hammerhead ribozyme. The values shown are µg of ribozyme added to 10µl reaction mix containing 1µg of target RNA.

The results show that GFPHH2 cleaves EGFP RNA *in vitro*, and that the more ribozyme is added to the reaction mix, the more cleavage occurs. The inactive version of the ribozyme does not cleave the EGFP target RNA, regardless of how much ribozyme is added.

Accompanying Figure 11 determines the optimum magnesium concentration for the DNAzymes, where the values are final magnesium concentration of the reaction in mM, and shows that the optimum magnesium concentration for DNAzyme cleavage of EGFP target RNA is, in all cases, 5mM which is the physiological concentration found in cells and *in vivo*.

Accompanying Figure 12 shows the time course cleavage assay of DNAzymes, where the values shown are the reaction time in minutes. The asterisks denote the time at which the cleavage products first become visible, and it can be seen that there is some variability as cleavage products were not seen using GFPY-DZ until 60 minutes had elapsed. This does not seem to be a factor in the usefulness of these molecules as can be seen in Figures 13a and 13b, where GFPY-DZ can be seen to be particularly effective in cell culture.

Cleavage assays were also performed using the DNAzyme GFPB2DZ using EGFP RNA as target. The DNAzyme did not cleave this target because this DNAzyme was designed to cleave EBFP RNA (data not shown).

Similarly, cleavage assays performed with GFPHH1 and GFPHH3 ribozymes showed that these ribozymes cleaved as expected (data not shown).

Down-Regulation of EGFP expression by DNAzymes in Cell Culture

293 HEK (Human Embryonic Kidney) cells were transiently transfected with 1μg EGFP plasmid under control of either the CMV (Cytomegalovirus) promoter or the SV-40 (Simian Virus 40) promoter.

Quantities of DNAzyme (0-20μg) or scrambled DNAzyme were either added to the media immediately post-transfection and replenished every 12 hours, or transfected along with the EFGP plasmid.

All transfections were performed using FuGene™ from Roche. For Figures *13a-13d* all fields were also photographed in bright field to ensure that decreasing fluorescence was not due to death of cells – data not shown.

The results shown in Figure 13 show that DNAzymes can effectively down-regulate EGFP expression in cell culture, and that better down-regulation is seen with increasing amounts of DNAzyme transfected. Transfection of DNAzymes appears to enhance the down-regulation seen, compared with adding the DNAzymes to the media post-transfection.

EXAMPLE 2

Two group I intron ribozymes were designed, both targeting the EGFP gene and both mediating the splicing of ECFP RNA sequence onto EGFP RNA, thereby effecting a "repair" of green to cyan.

GCFPRZ1 has a 9 base pair internal guide sequence (IGS), SEQ ID NO. 17, which is complementary to bases 87 – 96 of the EGFP gene (the "A" residue of the ATG start codon being taken as base 1). The exon of this ribozyme comprises 698 base pairs of the ECFP sequence.

GCFPRZ2 has a 9 base pair internal guide sequence (IGS), SEQ ID NO. 18, which is complementary to bases 102 –111 of the EGFP gene (the "A" residue of the ATG start codon being taken as base 1). The exon of this ribozyme comprises 683 base pairs of the ECFP sequence.

Full length EGFP RNA was synthesised using Ambion Megascript SP6 *in vitro* transcription kit while the ribozymes were synthesised using Ambion Maxiscript T7 *in-vitro* transcription kit. The RNA was precipitated and quantified.

The GCFPRZ1 or GCFPRZ2 ribozyme, as appropriate, (1mM) was preheated in reaction buffer (50mM HEPES (pH 7.0), 150mM NaCl and 5mM $MgCl_2$) at 95°C for 2 minutes. EGFP target RNA (100nM) was preheated at 95°C for 2 minutes, in the presence (or absence) of 100μM GTP, and then added to the ribozymes and incubated at 37°C for 3 hours.

The reactions were precipitated by addition of 30μl of RNAse-free water, 5μl of 5M sodium acetate and 150μl of ice cold ethanol, and incubated at –20°C overnight. The RNA was centrifuged, washed, and resuspended in 30μl of RNAse-free water.

The RNA from the splicing reactions was subjected to reverse transcription using either random hexamer primers (which transcribes all RNA species present) or GFPrev primer (which transcribes all EGFP RNA's present as well as those of the colour derivatives).

The cDNA was resuspended in 20µl clean water. 1µl of the cDNA was taken to PCR using EGFP forward primer (which amplifies any of the colour derivatives) and a specific ECFP reverse primer which is designed to be specific for the ECFP sequence.

EGFP sequence: ctatatca*tg*gccgacaagc SEQ ID NO. 19
ECFP sequence: ctatatca*cc*gccgacaagc SEQ ID NO. 20

Thus, the ECFP reverse primer is as follows:

5' GCTTGTCGGCGGTGATATAG 3' SEQ ID NO. 21

The melting temperature (Tm) of this primer is 64.3°C, therefore the annealing temperature for the PCR reaction should be 60°C (Tm -5°C) , however to increase specificity of the primer and to completely ensure absence of binding to the eGFP cDNA, the annealing temperature used was 65°C.

The lanes contained the following ingredients:

1 =	RZ1	Reverse transcribed with Random Hexamers
2 =	RZ2	Reverse transcribed with Random Hexamers
3 =	Target only	Reverse transcribed with Random Hexamers
4 =	RZ1 no GFP	Reverse transcribed with Random Hexamers
5 =	RZ2 no GTP	Reverse transcribed with Random Hexamers
6 =	RZ1	Reverse transcribed with GFPrev
7 =	RZ2	Reverse transcribed with GFPrev
8 =	Target only	Reverse transcribed with GFPrev
9 =	RZ1 no GFP	Reverse transcribed with GFPrev
10 =	RZ2 no GTP	Reverse transcribed with GFPrev

H_2O = negative control

PCR products were observed in lanes 2, 5, 7 and 10. No PCR products were observed for any RZ1 lanes.

Thus, RZ2 *trans*-splices ECFP sequence onto EGFP RNA – as a PCR product will only be seen if ECFP sequence is present. ECFP sequence can only be present if the splice reaction occurred.

RZ2 appeared to work in the absence of GTP. This unexpected result can be explained by the carryover of unincorporated GTP from both of the *in vitro* transcription reactions, as these would not be removed during the precipitation step. As the amount of GTP required to catalyse the reaction is very small, any carry over could easily cause the reaction to proceed.

The appearance of primer dimers indicates that the PCR reaction should be performed at even higher annealing temperatures.

In order to ascertain whether RZ1 could be used successfully, a second *in vitro* transcription of RZ1 RNA was performed. This time, the plasmid template for the transcription reaction was cut with a restriction enzyme further downstream of the stop codon of the ECFP exon, in order to enable the RT reaction to be performed using a reverse primer specific for the *trans*-spliced products only. This RT-PCR should be more sensitive and allow lower levels of repair to be detected. The PCR reaction was performed this time at 68°C to further increase specificity.

The lanes on the gel contained the following ingredients:

M = molecular weight marker

1 =	RZ1	Reverse transcribed with Random Hexamers
2 =	RZ2	Reverse transcribed with Random Hexamers
3 =	Target only	Reverse transcribed with Random Hexamers
4 =	RZ1	Reverse transcribed with GFPrev

5 =	RZ2	Reverse transcribed with GFPrev
6 =	Target only	Reverse transcribed with GFPrev
7 =	RZ 1	Reverse transcribed with splice product specific primer
8 =	Target only	Reverse transcribed with splice product specific primer

H_2O = negative control

Lanes 2, 5 and 7 contained PCR products. Thus, RZ2 mediated *trans*-splicing at levels that could be detected via RT-PCR using random hexamers and GFPrev primer, while RZ1 mediated *trans*-splicing at levels that could be detected via RT-PCR using a more sensitive splice product specific reverse primer. No primer dimers were observed.

A further PCR reaction was performed as a positive control, this time using EGFP forward AND reverse primers.

The lanes of the gel contained the following ingredients:

M = molecular weight marker

1 =	RZ1	Reverse transcribed with Random Hexamers
2 =	RZ2	Reverse transcribed with Random Hexamers
3 =	Target only	Reverse transcribed with Random Hexamers
4 =	RZ1	Reverse transcribed with GFPrev
5 =	RZ2	Reverse transcribed with GFPrev
6 =	Target only	Reverse transcribed with GFPrev
7 =	RZ 1	Reverse transcribed with splice product specific primer
8 =	Target only	Reverse transcribed with splice product specific primer

H_2O = negative control

Lanes 1 to 7 showed the presence of cDNA in all reactions, proving that absence of a PCR product in the "target only" RT-PCR reactions was indeed due to the specificity of the reverse PCR primer for the CFP sequence, not because there was no cDNA present in the reaction. No band should have been, nor was, present in lane 8 (Target only Reverse transcribed with splice specific primer) as the full length

EGFP RNA present in this reaction does not have complimentarity to the splice product specific primer and therefore should not undergo transcription into cDNA

Claims:

1. A catalytic polynucleic acid molecule, characterised in that the molecule is capable of catalysing the cleavage of a selected mRNA into 5' and 3' cleavage products, and replacing the 3' cleavage product with a coding sequence encoding all or part of a fluorescent protein, the resulting mRNA being translatable to express a fluorescent protein.

2. A catalytic polynucleic acid molecule, characterised in that the molecule is capable of catalysing the cleavage of an mRNA, encoding a fluorescent protein, into 5' and 3' cleavage products, and replacing the 3' cleavage product with a coding sequence encoding part of a second fluorescent protein, the resulting mRNA being translatable to express a fluorescent protein different from that encoded by the uncleaved mRNA.

3. A molecule according to claim 1 or 2, which is a ribozyme or a DNAzyme comprising a 3' intron encoding the C terminus of a fluorescent protein.

4. A molecule according to claim 3, which is a hammerhead ribozyme or a group I intron ribozyme.

5. A molecule according to any preceding claim, wherein the fluorescent proteins belong to the GFP family.

6. A molecule according to claim 5, wherein the fluorescent proteins are selected from the group consisting of EGFP, ECFP, EBFP and EYFP.

7. A molecule according to any preceding claim, wherein the expressed fluorescent protein is encoded by mRNA having the first 195 bases of its base sequence identical to bases 1 to 195 of SEQ ID NO. 1, or differing therefrom by no more than 5%.

8. A method for determining the success of delivering a catalytic polynucleic acid molecule to a target tissue, wherein the target tissue expresses a selected mRNA,

and wherein the catalytic polynucleic acid molecule is capable of catalysing the cleavage of the selected mRNA into 5' and 3' cleavage products, and replacing the 3' cleavage product with a coding sequence encoding all or part of a fluorescent protein, the resulting mRNA being translatable to express a fluorescent protein, success being determined by assaying the presence of fluorescent protein encoded by said resulting mRNA.

9. A method according to claim 8, wherein the tissue is located in an EGFP transgenic mouse.

10. A method according to claim 8, wherein the tissue is a stably transfected cell line expressing a marker selected from the group consisting of EGFP, EYFP, EBFP and ECFP.

11. A method according to any of claims 8 to 10, wherein the target mRNA encodes a cellular protein tagged with a fluorescent protein.

12. A method for the selection of a suitable IGS, or combination of IGS's, comprising a panel of ribozymes capable of splicing a full length fluorescent protein into a selected mRNA, the panel having a ribozyme having all possible variations, or a majority of all possible variations, of the IGS.

13. A method for determining the likely success of delivering a polynucleic acid molecule to a target tissue such that the polynucleic acid molecule is expressed in said tissue, wherein the target tissue expresses a selected mRNA, said method comprising delivering a catalytic polynucleic acid molecule to said tissue in a manner in which it is desired to deliver said polynucleic acid molecule, and wherein the catalytic polynucleic acid molecule is capable of catalysing the cleavage of the selected mRNA into 5' and 3' cleavage products, and replacing the 3' cleavage product with a coding sequence encoding all or part of a fluorescent protein, the resulting mRNA being translatable to express a fluorescent protein, success being determined by assaying the presence of fluorescent protein encoded by said resulting mRNA.

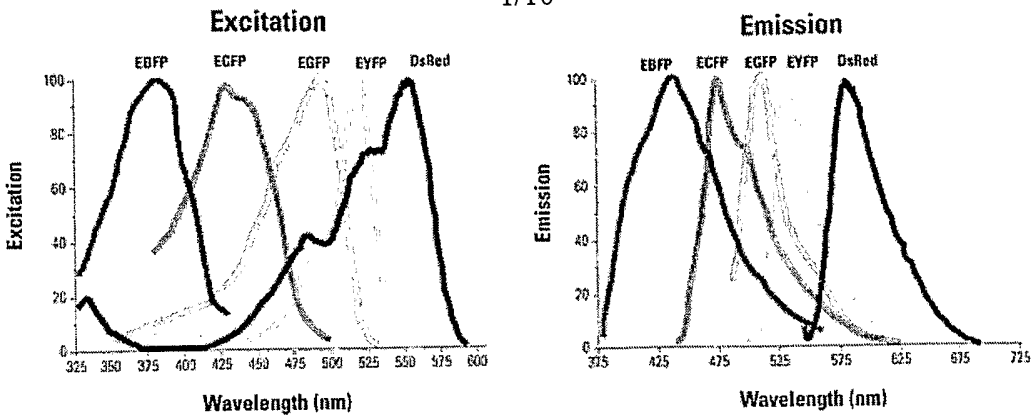


Fig.1

```
          1         10         20         30         40         50         60
          |         |          |          |          |          |          |
EGFP      ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGAC
EBFP      ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGAC
ECFP      ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGAC
EYFP      ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGAC
DRED      tc--------------CAAGAATGTTATCAAGGAGTTCATGAGGTTTAAGGTTCGCATGGAA

                    70         80         90         100        110        120
                    |          |          |          |          |          |
EGFP      GGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCA.CCTACGGCA
EBFP      GGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCA.CCTACGGCA
ECFP      GGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCA.CCTACGGCA
EYFP      GGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCA.CCTACGGCA
DRED      GGAACGGTCAATGGGCACGAGTTTGAAATAGAAGGCGAAGGAGAGGG-GAGGCCAtACGAAGGCC

          130        140        150        160        170        180        190
          |          |          |          |          |          |          |
EGFP      AGCTGACCCTGAAGTTCATCTGCACCACCGG.CAAGC..TGCCCGTGCCCTGGCCCACCCTCGTG
EBFP      AGCTGACCCTGAAGTTCATCTGCACCACCGG.CAAGC..TGCCCGTGCCCTGGCCCACCCTCGTG
ECFP      AGCTGACCCTGAAGTTCATCTGCACCACCGG.CAAGC..TGCCCGTGCCCTGGCCCACCCTCGTG
EYFP      AGCTGACCCTGAAGTTCATCTGCACCACCGG.CAAGC..TGCCCGTGCCCTGGCCCACCCTCGTG
DRED      ACAATACCGTAAAGCTTAAGGTAACCAAGGGgGGACCttTGCCATTTGCTTGGGATA-TTTTGTC

                    200        210        220        230        240        250
                    |          |          |          |          |          |
EGFP      ACCACCCTGACC.TACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACT
EBFP      ACCACCCTGACC.CACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACT
ECFP      ACCACCCTGACC.TGGGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACT
EYFP      ACCACCTTCGGC.TACGGCCTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACT
DRED      ACCACAATTTCAgTATGGAAGCAAGGTATATGTCAAGCACCCTGCCGACATA-------CCAGACT

          260        270        280        290        300        310        320
          |          |          |          |          |          |          |
EGFP      TCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGC
EBFP      TCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGC
ECFP      TCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGC
EYFP      TCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGC
DRED      ATAAAAAGCTGTCATTTCCTGAAGGATTTAAATGGGAAAGGGTCATGAACTTTGAAGACGGTGGC
```

```
                330        340        350        360        370        380
                |          |          |          |          |          |
EGFP AAC..TACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTG
EBFP AAC..TACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTG
ECFP AAC..TACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTG
EYFP AAC..TACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTG
DRED GTCgtTACTGTAACC--CAGGATTCCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTC

          390        400        410        420        430        440        450
          |          |          |          |          |          |          |
EGFP AAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAG
EBFP AAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTTCAACAG
ECFP AAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACATCAG
EYFP AAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAG
DRED ATTGGCGTGAACTTTCCTTCCGATGGACCTGTTATGCAAAA----------------------

                460        470        480        490        500        510
                |          |          |          |          |          |
EGFP CCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCC
EBFP CCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCC
ECFP CCACAACGTCTATATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCC
EYFP CCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCC
DRED ------------------------------------------------------------

          520        530        540        550        560        570        580
          |          |          |          |          |          |          |
EGFP ACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGAC
EBFP ACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGAC
ECFP ACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGAC
EYFP ACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGAC
DRED ------------------------------------------------------------

                590        600        610        620        630        640
                |          |          |          |          |          |
EGFP GGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAA
EBFP GGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAA
ECFP GGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAA
EYFP GGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCTACCAGTCCGCCCTGAGCAAAGACCCCAA
DRED ------------------------------------------------------------

          650        660        670        680        690        700        710
          |          |          |          |          |          |          |
EGFP CGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG
EBFP CGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG
ECFP CGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG
EYFP CGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG
DRED ------------------------------------------------------------

                720
                |
EGFP ACGAGCTGTACAAGTAA
EBFP ACGAGCTGTACAAGTAA
ECFP ACGAGCTGTACAAGTAA
EYFP ACGAGCTGTACAAGTAA
DRED ----------------
```

Fig. 2

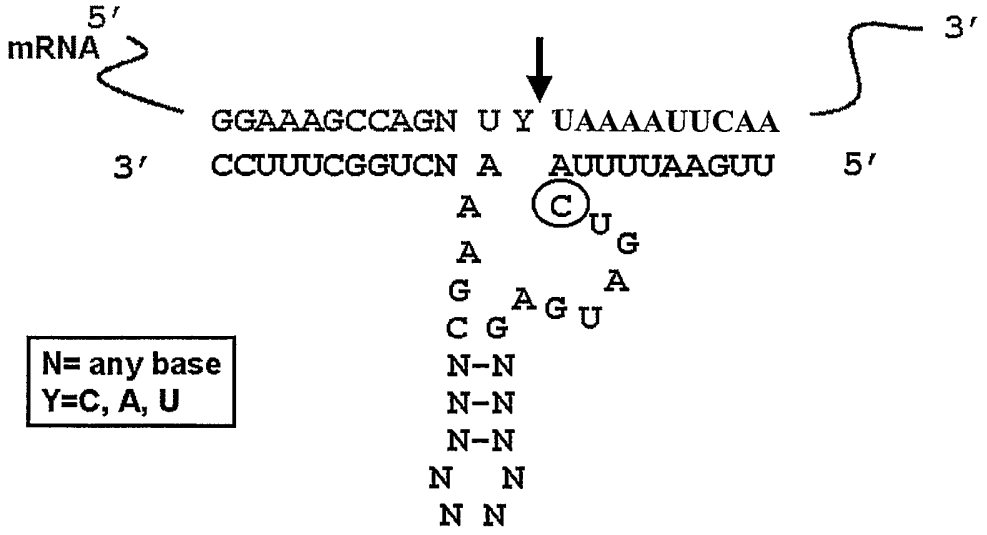
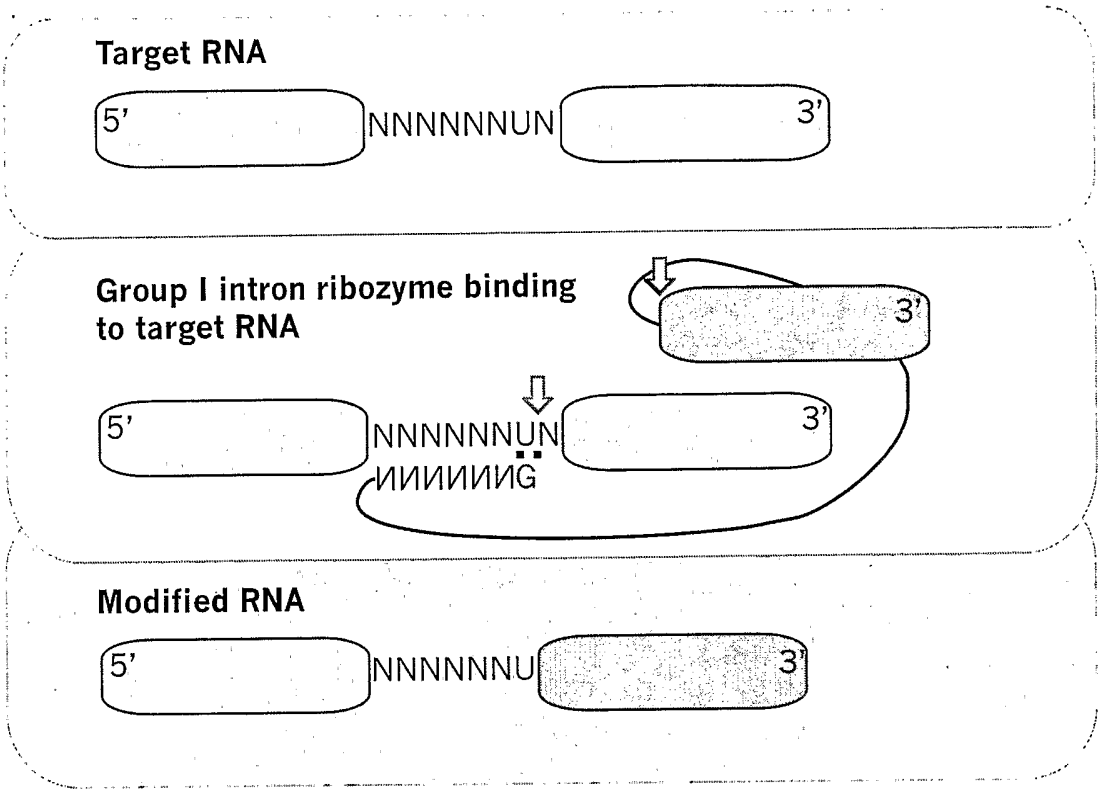

5′
mRNA
3′
GGAAAGCCAGN U Y UAAAAUUCAA
3′ CCUUUCGGUCN A AUUUUAAGUU 5′
A C U G
A A
G A G U A
C G
N-N
N-N
N-N
N N
N N
N= any base
Y=C, A, U

Fig. 3



G = G-U wobble in P1 substrate helix

⇩ = 5' splice site

⇩ = 3' splice site

Fig. 4

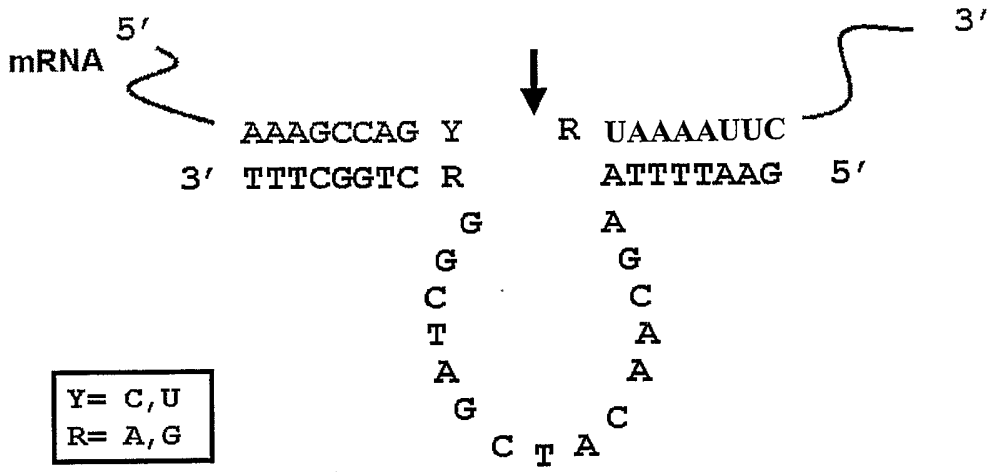

mRNA
5′
3′
AAAGCCAG Y R UAAAAUUC
3′ TTTCGGTC R ATTTTAAG 5′
G A
G G
C C
T A
A A
G C
C T A
Y= C,U
R= A,G

Fig. 5

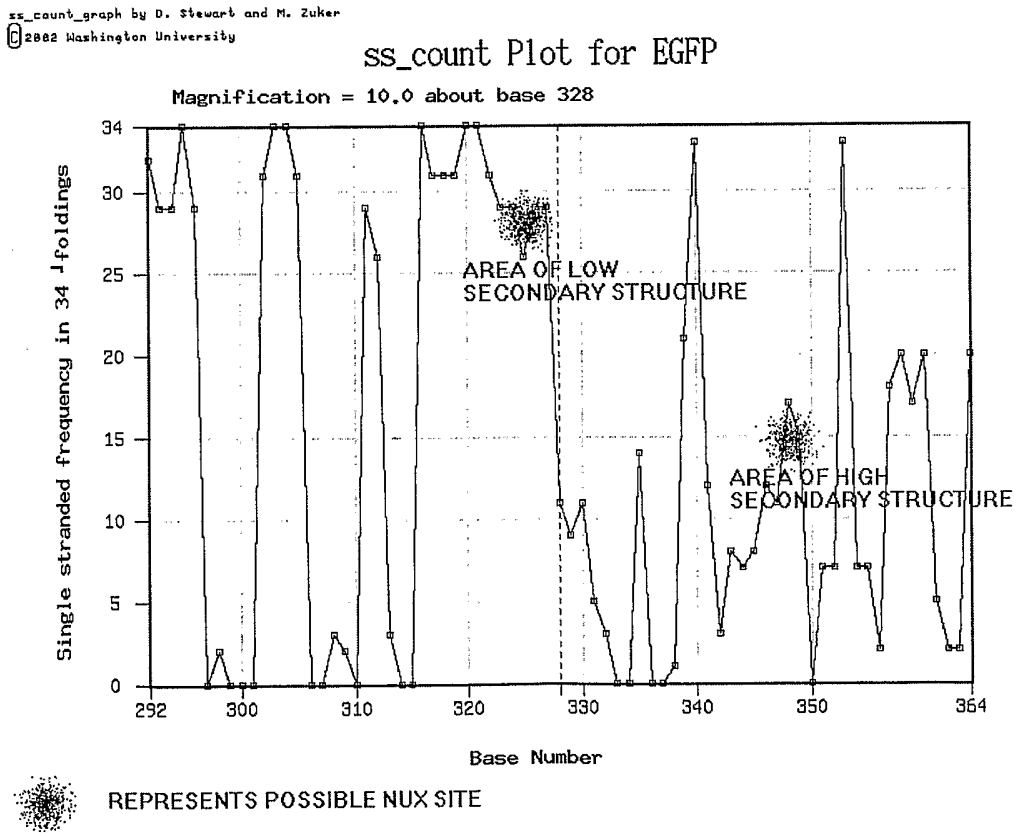

ss_count_graph by D. Stewart and M. Zuker
©2002 Washington University
ss_count Plot for EGFP
Magnification = 10.0 about base 328
Single stranded frequency in 34 foldings
0
5
10
15
20
25
30
34
292
300
310
320
330
340
350
364
Base Number
AREA OF LOW SECONDARY STRUCTURE
AREA OF HIGH SECONDARY STRUCTURE
REPRESENTS POSSIBLE NUX SITE

Fig. 6

```
          10         20         30         40         50
AUGGUGAGCA AGGGCGAGGA GCUGUUCACC GGGGUGGUGC CCAUCCUGGU        GFPHH1
          60         70         80         90        100
CGAGCUGGAC GGCGACGUAA ACGGCCACAA GUUCAGCGUG UCCGGCGAGG
         110        120        130        140        150
GCGAGGGCGA UGCCACCUAC GGCAAGCUGA CCCUGAAGUU CAUCUGCACC
         160        170        180        190        200
ACCGGCAAGC UGCCCGUGCC CUGGCCCACC CUCGUGACCA CCCUGACCUA
         210        220        230        240        250
CGGCGUGCAG UGCUUCAGCC GCUACCCCGA CCACAUGAAG CAGCACGACU        GFPHH2
         260        270        280        290        300
UCUUCAAGUC CGCCAUGCCC GAAGGCUACG UCCAGGAGCG CACCAUCUUC
         310        320        330        340        350
UUCAAGGACG ACGGCAACUA CAAGACCCGC GCCGAGGUGA AGUUCGAGGG
         360        370        380        390        400
CGACACCCUG GUGAACCGCA UCGAGCUGAA GGGCAUCGAC UUCAAGGAGG
         410        420        430        440        450
ACGGCAACAU CCUGGGGCAC AAGCUGGAGU ACAACUACAA CAGCCACAAC        GFPHH3
         460        470        480        490        500
GUCUAUAUCA UGGCCGACAA GCAGAAGAAC GGCAUCAAGG UGAACUUCAA
         510        520        530        540        550
GAUCCGCCAC AACAUCGAGG ACGGCAGCGU GCAGCUCGCC GACCACUACC
         560        570        580        590        600
AGCAGAACAC CCCCAUCGGC GACGGCCCCG UGCUGCUGCC CGACAACCAC
         610        620        630        640        650
UACCUGAGCA CCCAGUCCGC CCUGAGCAAA GACCCCAACG AGAAGCGCGA
         660        670        680        690        700
UCACAUGGUC CUGCUGGAGU UCGUGACCGC CGCCGGGAUC ACUCUCGGCA
         710        720
UGGACGAGCU GUACAAGUAA
```

Fig. 7

```
          10         20         30         40         50
AUGGUGAGCA AGGGCGAGGA GCUGUUCACC GGGGUGGUGC CCAUCCUGGU        IGS for:
                G-BFP1          and          G-BFP2
                                                             SEQUENCE OF EGFP
          60         70         80         90        100     AND VARIANTS
CGAGCUGGAC GGCGACGUAA ACGGCCACAA GUUCAGCGUG UCCGGCGAGG       IDENTICAL
         110        120        130        140        150
GCGAGGGCGA UGCCACCUAC GGCAAGCUGA CCCUGAAGUU CAUCUGCACC
         160        170        180        190        200
ACCGGCAAGC UGCCCGUGCC CUGGCCCACC CUCGUGACCA CCCUGACCUA       SEQUENCE OF EGFP
         210        220        230        240        250     AND VARIANTS NOT
CGGCGUGCAG UGCUUCAGCC GCUACCCCGA CCACAUGAAG CAGCACGACU       IDENTICAL
         260        270        280        290        300
UCUUCAAGUC CGCCAUGCCC GAAGGCUACG UCCAGGAGCG CACCAUCUUC
         310        320        330        340        350
UUCAAGGACG ACGGCAACUA CAAGACCCGC GCCGAGGUGA AGUUCGAGGG
         360        370        380        390        400
CGACACCCUG GUGAACCGCA UCGAGCUGAA GGGCAUCGAC UUCAAGGAGG
         410        420        430        440        450
ACGGCAACAU CCUGGGGCAC AAGCUGGAGU ACAACUACAA CAGCCACAAC
         460        470        480        490        500
GUCUAUAUCA UGGCCGACAA GCAGAAGAAC GGCAUCAAGG UGAACUUCAA
         510        520        530        540        550
GAUCCGCCAC AACAUCGAGG ACGGCAGCGU GCAGCUCGCC GACCACUACC
         560        570        580        590        600
AGCAGAACAC CCCCAUCGGC GACGGCCCCG UGCUGCUGCC CGACAACCAC
         610        620        630        640        650
UACCUGAGCA CCCAGUCCGC CCUGAGCAAA GACCCCAACG AGAAGCGCGA
         660        670        680        690        700
UCACAUGGUC CUGCUGGAGU UCGUGACCGC CGCCGGGAUC ACUCUCGGCA
         710        720
UGGACGAGCU GUACAAGUAA
```

Fig. 8

```
          10         20         30         40         50
AUGGUGAGCA AGGGCGAGGA GCUGUUCACC GGGGUGGUGC CCAUCCUGGU
          60         70         80         90        100
CGAGCUGGAC GGCGACGUAA ACGGCCACAA GUUCAGCGUG UCCGGCGAGG
         110        120        130        140        150
GCGAGGGCGA UGCCACCUAC GGCAAGCUGA CCCUGAAGUU CAUCUGCACC
         160        170        180        190        200
ACCGGCAAGC UGCCCGUGCC CUGGCCCACC CUCGUGACCA CCCUGACCUA
         210        220        230        240        250
CGGCGUGCAG UGCUUCAGCC GCUACCCCGA CCACAUGAAG CAGCACGACU
         260        270        280        290        300
UCUUCAAGUC CGCCAUGCCC GAAGGCUACG UCCAGGAGCG CACCAUCUUC      GFP1DZ
         310        320        330        340        350
UUCAAGGACG ACGGCAACUA CAAGACCCGC GCCGAGGUGA AGUUCGAGGG
         360        370        380        390        400
CGACACCCUG GUGAACCGCA UCGAGCUGAA GGGCAUCGAC UUCAAGGAGG
         410        420        430        440        450
ACGGCAACAU CCUGGGGCAC AAGCUGGAGU ACAACUACAA CAGCCACAAC      GFPB2DZ (A =
                                                            mismatch in EGFP)
         460        470        480        490        500
GUCUAUAUCA UGGCCGACAA GCAGAAGAAC GGCAUCAAGG UGAACUUCAA
         510        520        530        540        550
GAUCCGCCAC AACAUCGAGG ACGGCAGCGU GCAGCUCGCC GACCACUACC
         560        570        580        590        600
AGCAGAACAC CCCCAUCGGC GACGGCCCCG UGCUGCUGCC CGACAACCAC
         610        620        630        640        650
UACCUGAGCA CCCAGUCCGC CCUGAGCAAA GACCCCAACG AGAAGCGCGA      GFPYDZ
         660        670        680        690        700
UCACAUGGUC CUGCUGGAGU UCGUGACCGC CGCCGGGAUC ACUCUCGGCA      GFP2DZ
         710        720
UGGACGAGCU GUACAAGUAA
```

Fig. 9

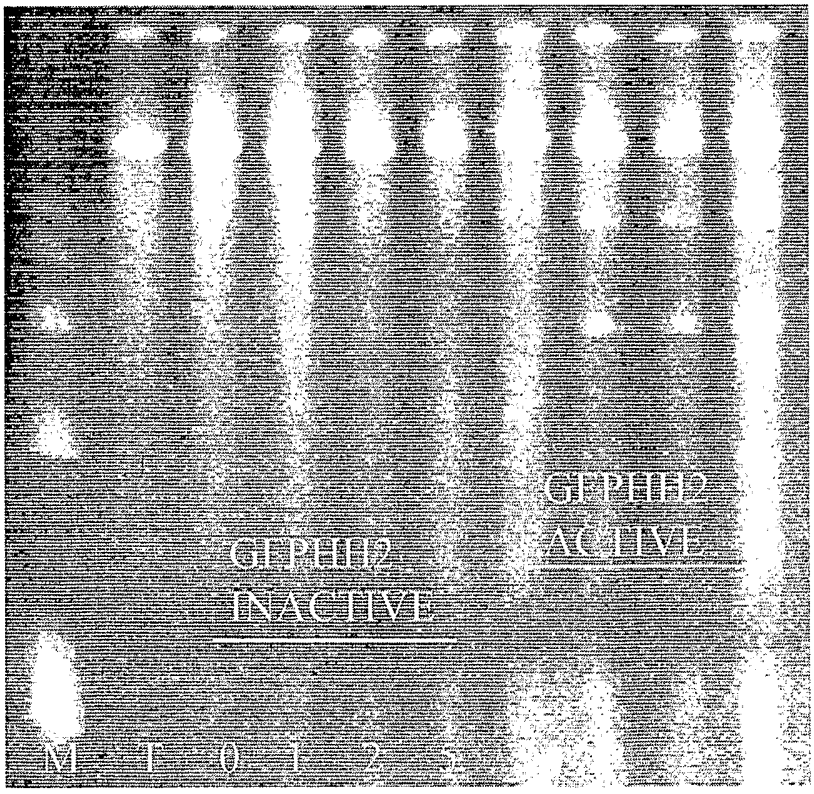


Fig. 10

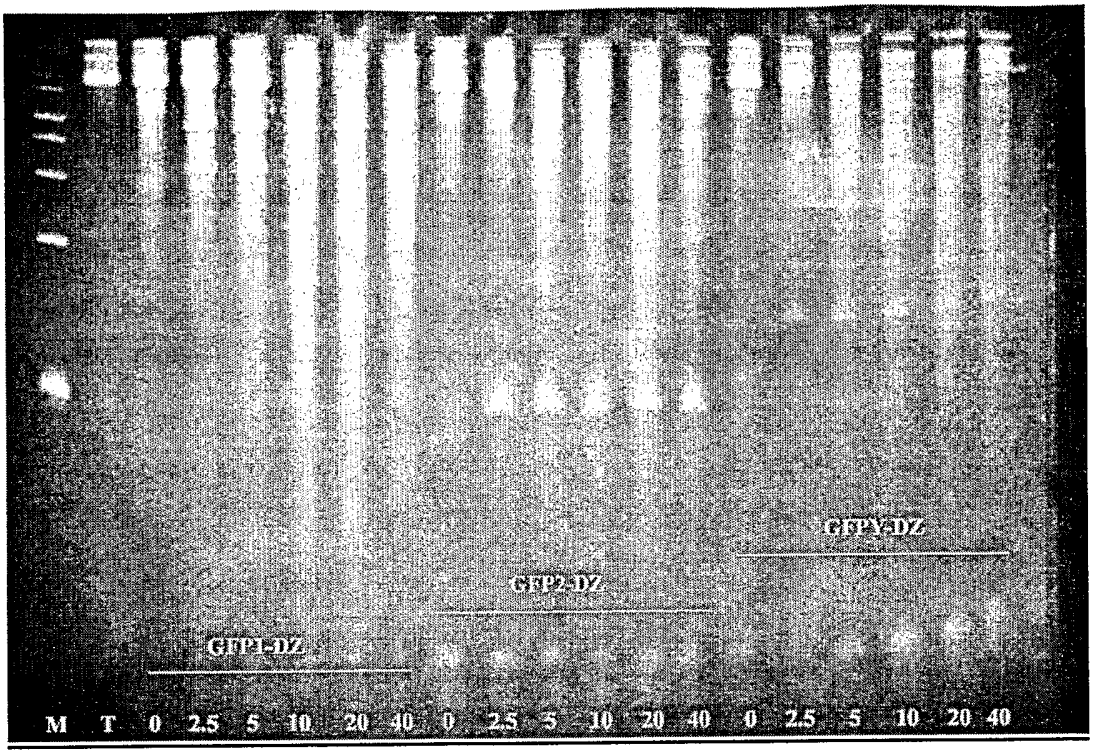

GFPY-DZ
GFP2-DZ
GFP1-DZ
M T 0 2.5 5 10 20 40 0 2.5 5 10 20 40 0 2.5 5 10 20 40

Fig. 11

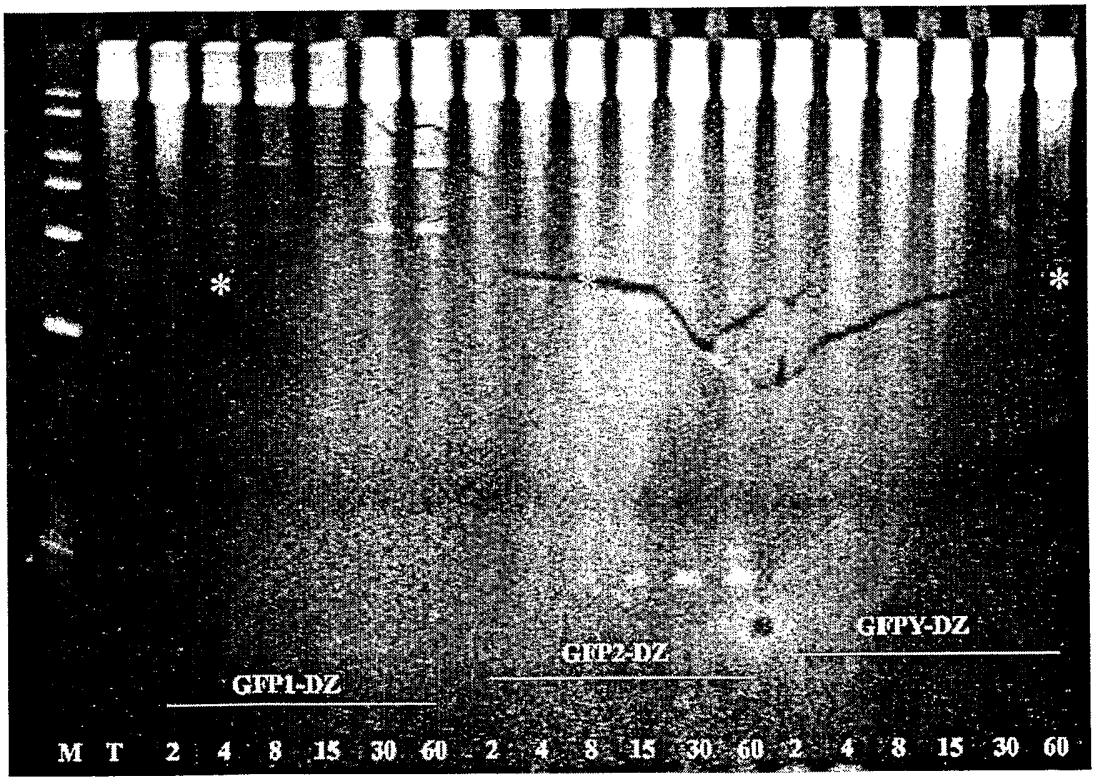
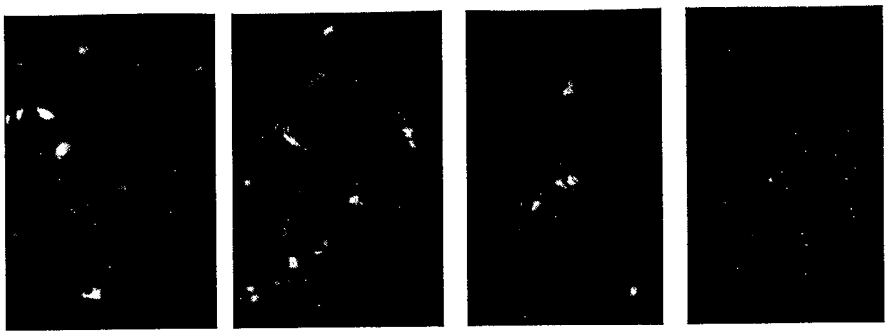

GFPY-DZ
GFP2-DZ
GFP1-DZ
M T 2 4 8 15 30 60 2 4 8 15 30 60 2 4 8 15 30 60

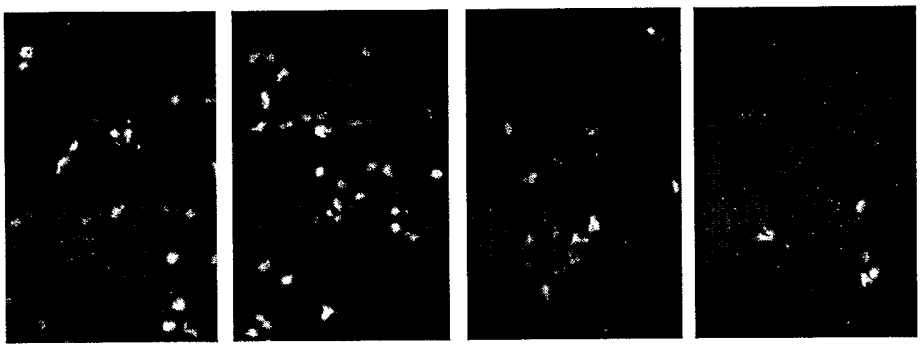
Fig. 12



CONTROL	5μg GFPY-DZ added to media	10μg GFPY-DZ added to media	20μg GFPY-DZ added to media

ALL eGFP in PCDNA3 Vector with CMV Promotor

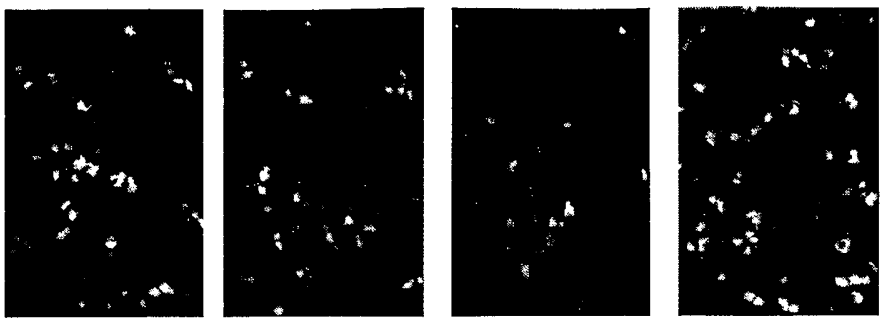
Fig. 13a



CONTROL	5μg GFPY-DZ added to media	10μg GFPY-DZ added to media	20μg GFPY-DZ added to media

ALL eGFP in SV40 Vector with SV40 Promotor

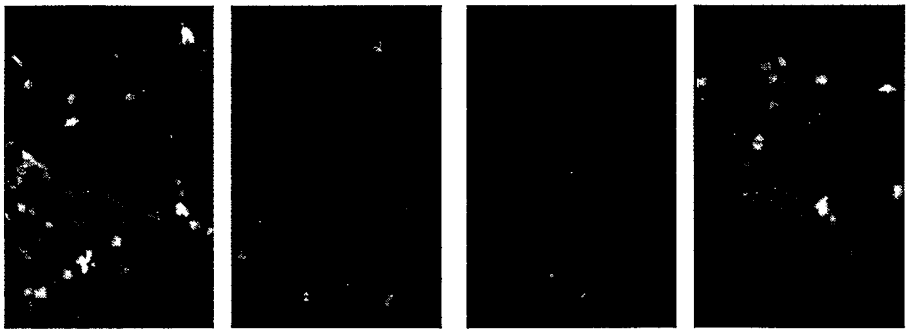
Fig. 13b



CONTROL	5μg GFP1-DZ added to media	10μg GFP1-DZ added to media	10μg DNAzyme S added to media

ALL eGFP in SV40 Vector with SV40 Promotor

Fig. 13c



CONTROL	5μg GFP1-DZ transfected	10μg GFP1-DZ transfected	10μg DNAzyme S transfected

ALL eGFP in SV40 Vector with SV40 Promotor

Fig. 13d

SEQUENCE LISTING

<110> Isis Innovation Limited

<120> Model systems

<130> WPP85148

<150> GB 0202543.5
<151> 2002-02-04

<160> 21

<170> PatentIn version 3.2

<210> 1
<211> 720
<212> DNA
<213> Aequorea victoria

<220>
<221> CDS
<222> (1)..(720)

<400> 1

```
atg gtg agc aag ggc gag gag ctg ttc acc ggg gtg gtg ccc atc ctg       48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                   10                  15

gtc gag ctg gac ggc gac gta aac ggc cac aag ttc agc gtg tcc ggc       96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20                  25                  30

gag ggc gag ggc gat gcc acc tac ggc aag ctg acc ctg aag ttc atc      144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35                  40                  45

tgc acc acc ggc aag ctg ccc gtg ccc tgg ccc acc ctc gtg acc acc      192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
    50                  55                  60

ctg acc tac ggc gtg cag tgc ttc agc cgc tac ccc gac cac atg aag      240
Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65                  70                  75                  80

cag cac gac ttc ttc aag tcc gcc atg ccc gaa ggc tac gtc cag gag      288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                  90                  95

cgc acc atc ttc ttc aag gac gac ggc aac tac aag acc cgc gcc gag      336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
            100                 105                 110

gtg aag ttc gag ggc gac acc ctg gtg aac cgc atc gag ctg aag ggc      384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
        115                 120                 125

atc gac ttc aag gag gac ggc aac atc ctg ggg cac aag ctg gag tac      432
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
    130                 135                 140

aac tac aac agc cac aac gtc tat atc atg gcc gac aag cag aag aac      480
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145                 150                 155                 160

ggc atc aag gtg aac ttc aag atc cgc cac aac atc gag gac ggc agc      528
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
```

```
                165                 170                 175

gtg cag ctc gcc gac cac tac cag cag aac acc ccc atc ggc gac ggc      576
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
            180                 185                 190

ccc gtg ctg ctg ccc gac aac cac tac ctg agc acc cag tcc gcc ctg      624
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
        195                 200                 205

agc aaa gac ccc aac gag aag cgc gat cac atg gtc ctg ctg gag ttc      672
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
    210                 215                 220

gtg acc gcc gcc ggg atc act ctc ggc atg gac gag ctg tac aag taa      720
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

<210> 2
<211> 239
<212> PRT
<213> Aequorea victoria

<400> 2

```
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                   10                  15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20                  25                  30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35                  40                  45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
    50                  55                  60

Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65                  70                  75                  80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                  90                  95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
            100                 105                 110

Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
        115                 120                 125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
    130                 135                 140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145                 150                 155                 160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165                 170                 175
```

```
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
            180                 185                 190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
        195                 200                 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
    210                 215                 220

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

<210> 3
<211> 720
<212> DNA
<213> Aequorea victoria

<220>
<221> CDS
<222> (1)..(720)

<400> 3

```
atg gtg agc aag ggc gag gag ctg ttc acc ggg gtg gtg ccc atc ctg      48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                   10                  15

gtc gag ctg gac ggc gac gta aac ggc cac aag ttc agc gtg tcc ggc      96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20                  25                  30

gag ggc gag ggc gat gcc acc tac ggc aag ctg acc ctg aag ttc atc     144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35                  40                  45

tgc acc acc ggc aag ctg ccc gtg ccc tgg ccc acc ctc gtg acc acc     192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
    50                  55                  60

ctg acc cac ggc gtg cag tgc ttc agc cgc tac ccc gac cac atg aag     240
Leu Thr His Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65                  70                  75                  80

cag cac gac ttc ttc aag tcc gcc atg ccc gaa ggc tac gtc cag gag     288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                  90                  95

cgc acc atc ttc ttc aag gac gac ggc aac tac aag acc cgc gcc gag     336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
            100                 105                 110

gtg aag ttc gag ggc gac acc ctg gtg aac cgc atc gag ctg aag ggc     384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
        115                 120                 125

atc gac ttc aag gag gac ggc aac atc ctg ggg cac aag ctg gag tac     432
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
    130                 135                 140

aac ttc aac agc cac aac gtc tat atc atg gcc gac aag cag aag aac     480
Asn Phe Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145                 150                 155                 160
```

```
ggc atc aag gtg aac ttc aag atc cgc cac aac atc gag gac ggc agc      528
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165                 170                 175

gtg cag ctc gcc gac cac tac cag cag aac acc ccc atc ggc gac ggc      576
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
            180                 185                 190

ccc gtg ctg ctg ccc gac aac cac tac ctg agc acc cag tcc gcc ctg      624
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
        195                 200                 205

agc aaa gac ccc aac gag aag cgc gat cac atg gtc ctg ctg gag ttc      672
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
    210                 215                 220

gtg acc gcc gcc ggg atc act ctc ggc atg gac gag ctg tac aag taa      720
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

<210> 4
<211> 239
<212> PRT
<213> Aequorea victoria

<400> 4

```
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                   10                  15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20                  25                  30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35                  40                  45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
    50                  55                  60

Leu Thr His Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65                  70                  75                  80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                  90                  95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
            100                 105                 110

Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
        115                 120                 125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
    130                 135                 140

Asn Phe Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145                 150                 155                 160
```

```
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165                 170                 175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
            180                 185                 190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
        195                 200                 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
    210                 215                 220

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

<210> 5
<211> 720
<212> DNA
<213> Aequorea victoria

<220>
<221> CDS
<222> (1)..(720)

<400> 5

```
atg gtg agc aag ggc gag gag ctg ttc acc ggg gtg gtg ccc atc ctg       48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                   10                  15

gtc gag ctg gac ggc gac gta aac ggc cac aag ttc agc gtg tcc ggc       96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20                  25                  30

gag ggc gag ggc gat gcc acc tac ggc aag ctg acc ctg aag ttc atc      144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35                  40                  45

tgc acc acc ggc aag ctg ccc gtg ccc tgg ccc acc ctc gtg acc acc      192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
    50                  55                  60

ctg acc tgg ggc gtg cag tgc ttc agc cgc tac ccc gac cac atg aag      240
Leu Thr Trp Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65                  70                  75                  80

cag cac gac ttc ttc aag tcc gcc atg ccc gaa ggc tac gtc cag gag      288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                  90                  95

cgc acc atc ttc ttc aag gac gac ggc aac tac aag acc cgc gcc gag      336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
            100                 105                 110

gtg aag ttc gag ggc gac acc ctg gtg aac cgc atc gag ctg aag ggc      384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
        115                 120                 125

atc gac ttc aag gag gac ggc aac atc ctg ggg cac aag ctg gag tac      432
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
    130                 135                 140
```

```
aac tac atc agc cac aac gtc tat atc acc gcc gac aag cag aag aac      480
Asn Tyr Ile Ser His Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys Asn
145                 150                 155                 160

ggc atc aag gcc aac ttc aag atc cgc cac aac atc gag gac ggc agc      528
Gly Ile Lys Ala Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165                 170                 175

gtg cag ctc gcc gac cac tac cag cag aac acc ccc atc ggc gac ggc      576
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
            180                 185                 190

ccc gtg ctg ctg ccc gac aac cac tac ctg agc acc cag tcc gcc ctg      624
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
        195                 200                 205

agc aaa gac ccc aac gag aag cgc gat cac atg gtc ctg ctg gag ttc      672
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
    210                 215                 220

gtg acc gcc gcc ggg atc act ctc ggc atg gac gag ctg tac aag taa      720
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

<210> 6
<211> 239
<212> PRT
<213> Aequorea victoria

<400> 6

```
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                   10                  15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20                  25                  30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35                  40                  45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
    50                  55                  60

Leu Thr Trp Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65                  70                  75                  80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                  90                  95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
            100                 105                 110

Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
        115                 120                 125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
    130                 135                 140
```

```
Asn Tyr Ile Ser His Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys Asn
145                 150                 155                 160

Gly Ile Lys Ala Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165                 170                 175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
            180                 185                 190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
        195                 200                 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
    210                 215                 220

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

<210> 7
<211> 720
<212> DNA
<213> Aequorea victoria

<220>
<221> CDS
<222> (1)..(720)

<400> 7

```
atg gtg agc aag ggc gag gag ctg ttc acc ggg gtg gtg ccc atc ctg       48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                   10                  15

gtc gag ctg gac ggc gac gta aac ggc cac aag ttc agc gtg tcc ggc       96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20                  25                  30

gag ggc gag ggc gat gcc acc tac ggc aag ctg acc ctg aag ttc atc      144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35                  40                  45

tgc acc acc ggc aag ctg ccc gtg ccc tgg ccc acc ctc gtg acc acc      192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
    50                  55                  60

ttc ggc tac ggc ctg cag tgc ttc gcc cgc tac ccc gac cac atg aag      240
Phe Gly Tyr Gly Leu Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys
65                  70                  75                  80

cag cac gac ttc ttc aag tcc gcc atg ccc gaa ggc tac gtc cag gag      288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                  90                  95

cgc acc atc ttc ttc aag gac gac ggc aac tac aag acc cgc gcc gag      336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
            100                 105                 110

gtg aag ttc gag ggc gac acc ctg gtg aac cgc atc gag ctg aag ggc      384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
        115                 120                 125

atc gac ttc aag gag gac ggc aac atc ctg ggg cac aag ctg gag tac      432
```

```
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
    130                 135                 140

aac tac aac agc cac aac gtc tat atc atg gcc gac aag cag aag aac      480
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145                 150                 155                 160

ggc atc aag gtg aac ttc aag atc cgc cac aac atc gag gac ggc agc      528
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165                 170                 175

gtg cag ctc gcc gac cac tac cag cag aac acc ccc atc ggc gac ggc      576
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
            180                 185                 190

ccc gtg ctg ctg ccc gac aac cac tac ctg agc tac cag tcc gcc ctg      624
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Ala Leu
        195                 200                 205

agc aaa gac ccc aac gag aag cgc gat cac atg gtc ctg ctg gag ttc      672
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
    210                 215                 220

gtg acc gcc gcc ggg atc act ctc ggc atg gac gag ctg tac aag taa      720
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

<210> 8
<211> 239
<212> PRT
<213> Aequorea victoria

<400> 8

```
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                   10                  15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20                  25                  30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35                  40                  45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
    50                  55                  60

Phe Gly Tyr Gly Leu Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys
65                  70                  75                  80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                  90                  95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
            100                 105                 110

Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
        115                 120                 125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
```

130 135 140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145 150 155 160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
165 170 175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
180 185 190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Ala Leu
195 200 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
210 215 220

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225 230 235

<210> 9
<211> 404
<212> DNA
<213> Discosoma sp.

<400> 9
tccaagaatg ttatcaagga gttcatgagg tttaaggttc gcatggaagg aacggtcaat 60

gggcacgagt ttgaaataga aggcgaagga gaggggaggc catacgaagg ccacaatacc 120

gtaaagctta aggtaaccaa ggggggacct ttgccatttg cttgggatat tttgtcacca 180

caatttcagt atggaagcaa ggtatatgtc aagcaccctg ccgacatacc agactataaa 240

aagctgtcat ttcctgaagg atttaaatgg gaaagggtca tgaactttga agacggtggc 300

gtcgttactg taacccagga ttccagtttg caggatggct gtttcatcta caaggtcaag 360

ttcattggcg tgaactttcc ttccgatgga cctgttatgc aaaa 404

<210> 10
<211> 23
<212> RNA
<213> Artificial Sequence

<220>
<223> mRNA sequence from Fig 3

<220>
<221> misc_feature
<222> (11)..(11)
<223> n is an unknown nucleotide

<400> 10
ggaaagccag nuyuaaaauu caa 23

<210> 11
<211> 42

<212> RNA
<213> Artificial Sequence

<220>
<223> Ribozyme from Fig 3

<220>
<221> misc_feature
<222> (19)..(27)
<223> n is an unknown nucleotide

<220>
<221> misc_feature
<222> (32)..(32)
<223> n is an unknown nucleotide

<400> 11
uugaauuuua cugaugagnn nnnnnncgaa ancuggcuuu cc 42

<210> 12
<211> 18
<212> RNA
<213> Artificial Sequence

<220>
<223> mRNA seq from Fig 5

<400> 12
aaagccagyr uaaaauuc 18

<210> 13
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> DNAzyme from Fig 5

<400> 13
gaattttaag caacatcgat cggrctggct tt 32

<210> 14
<211> 720
<212> RNA
<213> Aequorea victoria

<400> 14
auggugagca agggcgagga gcuguucacc gggguggugc ccauccuggu cgagcuggac 60

ggcgacguaa acggccacaa guucagcgug uccggcgagg gcgagggcga ugccaccuac 120

ggcaagcuga cccugaaguu caucugcacc accggcaagc ugcccgugcc cuggcccacc 180

cucgugacca cccugaccua cggcgugcag ugcuucagcc gcuaccccga ccacaugaag 240

cagcacgacu ucuucaaguc cgccaugccc gaaggcuacg uccaggagcg caccaucuuc 300

uucaaggacg acggcaacua caagacccgc gccgagguga aguucgaggg cgacacccug 360

gugaaccgca ucgagcugaa gggcaucgac uucaaggagg acggcaacau ccuggggcac 420

aagcuggagu acaacuacaa cagccacaac gucuauauca uggccgacaa gcagaagaac 480

ggcaucaagg ugaacuucaa gauccgccac aacaucgagg acggcagcgu gcagcucgcc 540

gaccacuacc agcagaacac ccccaucggc gacggccccg ugcugcugcc cgacaaccac 600

uaccugagca cccaguccgc ccugagcaaa gaccccaacg agaagcgcga ucacaugguc 660

cugcuggagu ucgugaccgc cgccgggauc acucucggca uggacgagcu guacaaguaa 720

<210> 15
<211> 720
<212> RNA
<213> Aequorea victoria

<400> 15
auggugagca agggcgagga gcuguucacc gggguggugc ccauccuggu cgagcuggac 60

ggcgacguaa acggccacaa guucagcgug uccggcgagg gcgagggcga ugccaccuac 120

ggcaagcuga cccugaaguu caucugcacc accggcaagc ugcccgugcc cuggcccacc 180

cucgugacca cccugaccua cggcgugcag ugcuucagcc gcuaccccga ccacaugaag 240

cagcacgacu ucuucaaguc cgccaugccc gaaggcuacg uccaggagcg caccaucuuc 300

uucaaggacg acggcaacua caagacccgc gccgagguga aguucgaggg cgacacccug 360

gugaaccgca ucgagcugaa gggcaucgac uucaaggagg acggcaacau ccuggggcac 420

aagcuggagu acaacuacaa cagccacaac gucuauauca uggccgacaa gcagaagaac 480

ggcaucaagg ugaacuucaa gauccgccac aacaucgagg acggcagcgu gcagcucgcc 540

gaccacuacc agcagaacac ccccaucggc gacggccccg ugcugcugcc cgacaaccac 600

uaccugagca cccaguccgc ccugagcaaa gaccccaacg agaagcgcga ucacaugguc 660

cugcuggagu ucgugaccgc cgccgggauc acucucggca uggacgagcu guacaaguaa 720

<210> 16
<211> 720
<212> RNA
<213> Aequorea victoria

<400> 16
auggugagca agggcgagga gcuguucacc gggguggugc ccauccuggu cgagcuggac 60

ggcgacguaa acggccacaa guucagcgug uccggcgagg gcgagggcga ugccaccuac 120

ggcaagcuga cccugaaguu caucugcacc accggcaagc ugcccgugcc cuggcccacc 180

cucgugacca cccugaccua cggcgugcag ugcuucagcc gcuaccccga ccacaugaag 240

cagcacgacu ucuucaaguc cgccaugccc gaaggcuacg uccaggagcg caccaucuuc 300

uucaaggacg acggcaacua caagacccgc gccgagguga aguucgaggg cgacacccug 360

gugaaccgca ucgagcugaa gggcaucgac uucaaggagg acggcaacau ccuggggcac 420

aagcuggagu acaacuacaa cagccacaac gucuauauca uggccgacaa gcagaagaac 480

ggcaucaagg ugaacuucaa gauccgccac aacaucgagg acggcagcgu gcagcucgcc 540

gaccacuacc agcagaacac ccccaucggc gacggccccg ugcugcugcc cgacaaccac 600

uaccugagca cccaguccgc ccugagcaaa gaccccaacg agaagcgcga ucacaugguc 660

cugcuggagu ucgugaccgc cgccgggauc acucucggca uggacgagcu guacaaguaa 720

<210> 17
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> GCF PRZ1 IGS

<400> 17
gccggacacg 10

<210> 18
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> GCF PRZ 2 IGS sequence

<400> 18
atcgccctcg 10

<210> 19
<211> 20
<212> DNA
<213> Aequorea victoria

<400> 19
ctatatcatg gccgacaagc 20

<210> 20
<211> 20
<212> DNA
<213> Aequorea victoria

<400> 20
ctatatcacc gccgacaagc 20

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

<220>
<223> ECFP reverse primer

<400> 21
gcttgtcggc ggtgatatag 20

INTERNATIONAL SEARCH REPORT

PCT/GB 03/00475

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/11 C07K14/435 A01K67/027

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K C12N A01K

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

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, MEDLINE, BIOSIS, CHEM ABS Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KOEHLER UWE ET AL: "Trans-splicing ribozymes for targeted gene delivery." JOURNAL OF MOLECULAR BIOLOGY, vol. 285, no. 5, 5 February 1999 (1999-02-05), pages 1935-1950, XP002240969 ISSN: 0022-2836 page 1939, left-hand column, paragraph 4 -right-hand column, paragraph 2 page 1943, right-hand column, line 10 - line 14 page 1944, right-hand column, paragraph 1 -page 1945, left-hand column, paragraph 1 page 1946, left-hand column, line 43 - line 47 page 1947, right-hand column, line 37 - line 50 --- -/--	1,3-11, 13

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
14 May 2003	06/06/2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Mandl, B

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	SONNTAG K C ET AL: "Blocking of IT15 gene expression as an approach for a treatment of Huntington's disease." SOCIETY FOR NEUROSCIENCE ABSTRACTS, vol. 27, no. 2, 2001, page 2576 XP009010813 31st Annual Meeting of the Society for Neuroscience;San Diego, California, USA; November 10-15, 2001 ISSN: 0190-5295 abstract	1,3-11, 13
Y	WO 99 41371 A (KECK JAMES G ;WONG JUSTIN G P (US); STRATA BIOSCIENCES INC (US)) 19 August 1999 (1999-08-19) example 4	12
Y	ROSSI J J: "Ribozymes to the rescue: repairing genetically defective mRNAs" TRENDS IN GENETICS, vol. 14, no. 8, 1 August 1998 (1998-08-01), pages 295-298, XP004129329 ISSN: 0168-9525 figure 2 page 297, left-hand column, paragraph 2 -middle column, paragraph 1	12
A	AYRE B G ET AL: "Design of highly specific cytotoxins by using trans-splicing ribozymes" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 96, no. 7, 30 March 1999 (1999-03-30), pages 3507-3512, XP002227596 ISSN: 0027-8424 the whole document	1-11,13
A	MIKHEEVA ET AL: "Use of engineered ribozymes to catalyze chimeric gene assembly" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 93, July 1996 (1996-07), pages 7486-7490, XP002101538 ISSN: 0027-8424 abstract	1-11,13
A	WO 00 71701 A (NEW ENGLAND BIOLABS INC ;COMB DONALD G (US); SUN LUO (US); XU MING) 30 November 2000 (2000-11-30) figure 21	1-11,13

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	SIEMERING K R ET AL: "MUTATIONS THAT SUPPRESS THE THERMOSENSITIVITY OF GREEN FLUORESCENT PROTEIN" CURRENT BIOLOGY, vol. 6, no. 12, December 1996 (1996-12), pages 1653-1663, XP009004294 ISSN: 0960-9822 page 1653, left-hand column, line 15 - line 20	1-13

INTERNATIONAL SEARCH REPORT

PCT/GB 03/00475

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 8,9,11 and 13, as far as they relate to a method of treatment and an in vivo application, are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [X] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-11,13

   Catalytic polynucleic acid molecules capable of catalysing the cleavage of an mRNA and replacing the 3'-cleavage product with a coding sequence encoding all or part of a fluorescent protein; and methods employing said catalytic polynucleic acid molecules.

2. Claim : 12

   A method for selection of a suitable IGS.

INTERNATIONAL SEARCH REPORT

PCT/GB 03/00475

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9941371 A	19-08-1999	AU 2678999 A	30-08-1999
		CA 2319112 A1	19-08-1999
		EP 1053317 A1	22-11-2000
		WO 9941371 A1	19-08-1999
WO 0071701 A	30-11-2000	AU 5039500 A	12-12-2000
		CA 2374497 A1	30-11-2000
		CN 1350582 T	22-05-2002
		EP 1183346 A1	06-03-2002
		JP 2003505012 T	12-02-2003
		WO 0071701 A1	30-11-2000