



Office de la Propriété
Intellectuelle
du Canada

Un organisme
d'Industrie Canada

Canadian
Intellectual Property
Office

An agency of
Industry Canada

CA 2489174 C 2013/02/05

(11)(21) **2 489 174**

(12) **BREVET CANADIEN
CANADIAN PATENT**

(13) **C**

(86) Date de dépôt PCT/PCT Filing Date: 2003/07/10
(87) Date publication PCT/PCT Publication Date: 2004/01/22
(45) Date de délivrance/Issue Date: 2013/02/05
(85) Entrée phase nationale/National Entry: 2004/12/09
(86) N° demande PCT/PCT Application No.: EP 2003/007516
(87) N° publication PCT/PCT Publication No.: 2004/007718
(30) Priorités/Priorities: 2002/07/10 (EP02015532.1);
2002/08/23 (EP02018906.4)

(51) Cl.Int./Int.Cl. *C12N 15/11* (2006.01),
A61K 48/00 (2006.01), *C07K 14/47* (2006.01),
C12Q 1/68 (2006.01), *A61K 38/00* (2006.01)

(72) Inventeurs/Inventors:
TUSCHL, THOMAS, US;
MARTINEZ, JAVIER, US;
PATKANIOWSKA, AGNIESZKA, US;
URLAUB, HENNING, DE;
LUEHRMANN, REINHARD, DE

(73) Propriétaire/Owner:
MAX-PLANCK GESELLSCHAFT ZUR FOERDERUNG
DER WISSENSCHAFTEN E.V., DE

(74) Agent: NORTON ROSE CANADA
S.E.N.C.R.L., S.R.L./LLP

(54) Titre : INTERFERENCE DE L'ARN (ARNI) PAR DES MOLECULES D'ARN SIMPLE BRIN
(54) Title: RNA-INTERFERENCE BY SINGLE-STRANDED RNA MOLECULES

(57) **Abrégé/Abstract:**

The present invention relates to sequence and structural features of single-stranded (ss)RNA molecules required to mediate target-specific nucleic acid modifications by RNA-interference (RNAi), such as target mRNA degradation and/or DNA methylation.



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

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
22 January 2004 (22.01.2004)

PCT

(10) International Publication Number
WO 2004/007718 A3

(51) International Patent Classification⁷: C12N 15/11,
C07K 14/47, C12Q 1/68, A61K 48/00

(74) Agents: WEISS, Wolfgang et al.; Weickmann & Weickmann, Postfach 860 820, 81635 München (DE).

(21) International Application Number:
PCT/EP2003/007516

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(22) International Filing Date: 10 July 2003 (10.07.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
02015532.1 10 July 2002 (10.07.2002) EP
02018906.4 23 August 2002 (23.08.2002) EP

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): MAX-PLANCK-GESELLSCHAFT ZUR FÖRDERUNG DER WISSENSCHAFTEN E.V. [DE/DE]; Hofgartenstrasse 8, 80539 München (DE).

Declaration under Rule 4.17:

— of inventorship (Rule 4.17(iv)) for US only

(72) Inventors; and

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

(75) Inventors/Applicants (*for US only*): TUSCHL, Thomas [DE/US]; 504 East 63rd Street #37R, New York, NY 10021 (US). MARTINEZ, Javier [ES/US]; 504 East 63rd Street, Apt. 15 M, New York, NY 10021 (US). PATKAN-IOWSKA, Agnieszka [PL/US]; Rockefeller University, 1230 York Avenue, App. 306, Box 186, New York, NY 10021 (US). URLAUB, Henning [DE/DE]; Reinhäuser Landstr. 41, 37083 Göttingen (DE). LÜHRMANN, Reinhard [DE/DE]; Ewiges Tal 16b, 35041 Marburg-Michelbach (DE).

(88) Date of publication of the international search report:
25 March 2004

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

(54) Title: RNA-INTERFERENCE BY SINGLE-STRANDED RNA MOLECULES

(57) Abstract: The present invention relates to sequence and structural features of single-stranded (ss)RNA molecules required to mediate target-specific nucleic acid modifications by RNA-interference (RNAi), such as target mRNA degradation and/or DNA methylation.



WO 2004/007718 A3

- 1 -

RNA-interference by single-stranded RNA molecules

Description

5

The present invention relates to sequence and structural features of single-stranded (ss)RNA molecules required to mediate target-specific nucleic acid modifications by RNA-interference (RNAi), such as target mRNA degradation and/or DNA methylation.

10

Most eukaryotes possess a cellular defense system protecting their genomes against invading foreign genetic elements. Insertion of foreign elements is believed to be generally accompanied by formation of dsRNA that is interpreted by the cell as a signal for unwanted gene activity (e.g. Ahlquist, Science 296 (2002), 1270 - 1273; Fire et al., Nature 391 (1998), 806 - 811). Dicer RNase III rapidly processes dsRNA to small dsRNA fragments of distinct size and structure (e.g. Bernstein et al., Nature 409 (2001), 363 - 366), the small interfering RNAs (siRNAs) (Elbashir et al., Genes & Dev. 15 (2001 b), 188 - 200), which direct the sequence-specific degradation of the single-stranded mRNAs of the invading genes. siRNA duplexes have 2- to 3-nt 3' overhanging ends and contain 5' phosphate and free 3' hydroxyl termini (WO 02/44321). The process of posttranscriptional dsRNA-dependent gene silencing is commonly referred to as RNA interference (RNAi), and in some instances is also linked to transcriptional silencing.

25

Experimental introduction of siRNA duplexes into mammalian cells is now widely used to disrupt the activity of cellular genes homologous in sequence to the introduced dsRNA. Used as a reverse genetic approach, siRNA-induced gene silencing accelerates linking of gene sequence to biological function. siRNA duplexes are short enough to bypass general dsRNA-induced unspecific effects in vertebrate animal and mammalian

30

- 2 -

cells. siRNAs may also be expressed intracellularly from introduced expression plasmids or viral vectors providing an alternative to chemical RNA synthesis. Therefore, an understanding of how siRNAs act in mammalian systems is important for refining this gene silencing technology and for producing gene-specific therapeutic agents.

Biochemical studies have begun to unravel the mechanistic details of RNAi. The first cell-free systems were developed using *D. melanogaster* cell or embryo extracts, and were followed by the development of *in vitro* systems from *C. elegans* embryo and mouse embryonal carcinoma cells. While the *D. melanogaster* lysates support the steps of dsRNA processing and sequence-specific mRNA targeting, the latter two systems only recapitulate the first step.

RNAi in *D. melanogaster* extracts is initiated by ATP-dependent processing of long dsRNA to siRNAs by Dicer RNase III (e.g. Bernstein et al., (2001), supra). Thereafter, siRNA duplexes are assembled into a multi-component complex, which guides the sequence-specific recognition of the target mRNA and catalyzes its cleavage (e.g. Elbashir (2001 b), supra). This complex is referred to as RNA-induced silencing complex (RISC) (Hammond et al., Nature 404 (2000), 293 - 296). siRNAs in *D. melanogaster* are predominantly 21- and 22-nt, and when paired in a manner to contain a 2-nt 3' overhanging structure effectively enter RISC (Elbashir et al., EMBO J. 20 (2001 c), 6877 - 6888). Mammalian systems have siRNAs of similar size, and siRNAs of 21- and 22-nt also represent the most effective sizes for silencing genes expressed in mammalian cells (e.g. Elbashir et al., Nature 411 (2001 a), 494 - 498, Elbashir et al., Methods 26 (2002), 199 - 213).

RISC assembled on siRNA duplexes in *D. melanogaster* embryo lysate targets homologous sense as well as antisense single-stranded RNAs for degradation. The cleavage sites for sense and antisense target RNAs are

- 3 -

located in the middle of the region spanned by the siRNA duplex. Importantly, the 5'-end, and not the 3'-end, of the guide siRNA sets the ruler for the position of the target RNA cleavage. Furthermore, a 5' phosphate is required at the target-complementary strand of a siRNA duplex for RISC activity, and ATP is used to maintain the 5' phosphates of the siRNAs (Nykänen et al., Cell 107 (2001), 309 - 321). Synthetic siRNA duplexes with free 5' hydroxyls and 2-nt 3' overhangs are so readily phosphorylated in *D. melanogaster* embryo lysate that the RNAi efficiencies of 5'-phosphorylated and non-phosphorylated siRNAs are not significantly different (Elbashir et al. (2001 c), supra).

Unwinding of the siRNA duplex must occur prior to target RNA recognition. Analysis of ATP requirements revealed that the formation of RISC on siRNA duplexes required ATP in lysates of *D. melanogaster*. Once formed, RISC cleaves the target RNA in the absence of ATP. The need for ATP probably reflects the unwinding step and/or other conformational rearrangements. However, it is currently unknown if the unwound strands of an siRNA duplex remain associated with RISC or whether RISC only contains a single-stranded siRNA.

A component associated with RISC was identified as Argonaute2 from *D. melanogaster Schneider 2* (S2) cells (Hammond et al., Science 293 (2001 a), 1146 - 1150), and is a member of a large family of proteins. The family is referred to as Argonaute or PPD family and is characterized by the presence of a PAZ domain and a C-terminal Piwi domain, both of unknown function (Cerutti et al., Trends Biochem. Sci. (2000), 481 - 482); Schwarz and Zamore, Genes & Dev. 16 (2002), 1025 - 1031). The PAZ domain is also found in Dicer. Because Dicer and Argonaute2 interact in S2 cells, PAZ may function as a protein-protein interaction motif. Possibly, the interaction between Dicer and Argonaute2 facilitates siRNA incorporation into RISC. In *D. melanogaster*, the Argonaute family has five members, most of which were shown to be involved in gene silencing and

- 4 -

development. The mammalian members of the Argonaute family are poorly characterized, and some of them have been implicated in translational control, microRNA processing and development. The biochemical function of Argonaute proteins remains to be established and the development of
5 more biochemical systems is crucial.

Here we report on the analysis of human RISC in extracts prepared from HeLa cells. The reconstitution of RISC and the mRNA targeting step revealed that RISC is a ribonucleoprotein complex that is composed of a
10 single-stranded siRNA. Once RISC is formed the incorporated siRNA can no longer exchange with free siRNAs. Surprisingly, RISC can be reconstituted in HeLa S100 extracts providing single-stranded siRNAs. Introducing 5' phosphorylated single-stranded antisense siRNAs into HeLa cells potently silences an endogenous gene with similar efficiency than duplex siRNA.

15 The object underlying the present invention is to provide novel agents capable of mediating target-specific RNAi.

The solution of this problems is provided by the use of a single-stranded
20 RNA molecule for the manufacture of an agent for inhibiting the expression of said target transcript. Surprisingly, it was found that single-stranded RNA molecules are capable of inhibiting the expression of target transcripts by RNA-interference (RNAi).

25 The length of the single-stranded RNA molecules is preferably from 14-50 nt, wherein at least the 14 to 20 5'-most nucleotides are substantially complementary to the target RNA transcript. The RNA oligonucleotides may have a free 5' hydroxyl moiety, or a moiety which is 5' phosphorylated (by means of chemical synthesis or enzymatic reactions) or
30 which is modified by 5'-monophosphate analogues.

- 5 -

The inhibition of target transcript expression may occur in vitro, e.g. in eucaryotic, particularly mammalian cell cultures or cell extracts. On the other hand, the inhibition may also occur in vivo i.e. in eucaryotic, particularly mammalian organisms including human beings.

5

Preferably, the single-stranded RNA molecule has a length from 15 - 29 nucleotides. The RNA-strand may have a 3' hydroxyl group. In some cases, however, it may be preferable to modify the 3' end to make it resistant against 3' to 5' exonucleases. Tolerated 3'-modifications are for example
 10 terminal 2'-deoxy nucleotides, 3' phosphate, 2',3'-cyclic phosphate, C3 (or C6, C7, C12) aminolinker, thiol linkers, carboxyl linkers, non-nucleotidic spacers (C3, C6, C9, C12, abasic, triethylene glycol, hexaethylene glycol), biotin, fluoresceine, etc.

15 The 5'-terminus comprises an OH group, a phosphate group or an analogue thereof. Preferred 5' phosphate modifications are 5'-monophosphate ((HO)₂(O)P-O-5'), 5'-diphosphate ((HO)₂(O)P-O-P(HO)(O)-O-5'), 5'-triphosphate ((HO)₂(O)P-O-(HO)(O)P-O-P(HO)(O)-O-5'), 5'-guanosine cap (7-methylated or non-methylated) (7m-G-O-5'-(HO)(O)P-O-(HO)(O)P-O-
 20 P(HO)(O)-O-5'), 5'-adenosine cap (Appp), and any modified or unmodified nucleotide cap structure (N-O-5'-(HO)(O)P-O-(HO)(O)P-O-P(HO)(O)-O-5'), 5'-monothiophosphate (phosphorothioate; (HO)₂(S)P-O-5'), 5'-monodithiophosphate (phosphorodithioate; (HO)(HS)(S)P-O-5'), 5'-phosphorothiolate ((HO)₂(O)P-S-5'); any additional combination of
 25 oxygen/sulfur replaced monophosphate, diphosphate and triphosphates (e.g. 5'-alpha-thiotriphosphate, 5'-gamma-thiotriphosphate, etc.), 5'-phosphoramidates ((HO)₂(O)P-NH-5', (HO)(NH₂)(O)P-O-5'), 5'-alkylphosphonates (R = alkyl = methyl, ethyl, isopropyl, propyl, etc., e.g. RP(OH)(O)-O-5'-, (OH)₂(O)P-5'-CH₂-), 5'-alkyletherphosphonates (R = alkylether = methoxymethyl (MeOCH₂-), ethoxymethyl, etc., e.g. RP(OH)(O)-O-5'-).

30

- 6 -

The sequence of the RNA molecule of the present invention has to have a sufficient identity to a nucleic acid target molecule in order to mediate target-specific RNAi. Thus the single-stranded RNA molecule of the present invention is substantially complementary to the target transcript.

The target RNA cleavage reaction guided by the single-stranded RNA molecules of the present invention is highly sequence-specific. However, not all positions of the RNA molecule contribute equally to target recognition. Mismatches, particularly at the 3'-terminus of the single-stranded RNA molecule, more particularly the residues 3' to the first 20 nt of the single-stranded RNA molecule are tolerated. Especially preferred are single-stranded RNA molecules having at the 5'-terminus at least 15 and preferably at least 20 nucleotides which are completely complementary to a predetermined target transcript or have at only mismatch and optionally up to 35 nucleotides at the 3'-terminus which may contain 1 or several, e.g. 2, 3 or more mismatches.

In order to enhance the stability of the single-stranded RNA molecules, the 3'-ends may be stabilized against degradation, e.g. they may be selected such that they consist of purine nucleotides, particularly adenosine or guanosine nucleotides. Alternatively or additionally, 3' nucleotides may be substituted by modified nucleotide analogues, including backbone modifications of ribose and/or phosphate residues.

In an especially preferred embodiment of the present invention the RNA molecule may contain at least one modified nucleotide analogue. The nucleotide analogues may be located at positions where the target-specific activity, e.g. the RNAi mediating activity is not substantially affected, e.g. in a region at the 5'-end and/or the 3'-end of the RNA molecule. Particularly, the 3'-terminus may be stabilized by incorporating modified nucleotide analogues, such as non-nucleotidic chemical derivatives such as

- 7 -

C3 (or C6, C7, C12) aminolinker, thiol linkers, carboxyl linkers, non-nucleotidic spacers (C3, C6, C9, C12, abasic, triethylene glycol, hexaethylene glycol), biotin, fluoresceine, etc. A further modification, by which the nuclease resistance of the RNA molecule may be increased, is by covalent coupling of inverted nucleotides, e.g. 2'-deoxyribonucleotides or ribonucleotides to the 3'-end of the RNA molecule. A preferred RNA molecule structure comprises: 5'-single-stranded siRNA-3'-O-P(O)(OH)-O-3'-N, wherein N is a nucleotide, e.g. a 2'-deoxyribonucleotide or ribonucleotide, typically an inverted thymidine residue, or an inverted oligonucleotide structure, e.g. containing up to 5 nucleotides.

Preferred nucleotide analogues are selected from sugar- or backbone-modified ribonucleotides. It should be noted, however, that also nucleobase-modified ribonucleotides, i.e. ribonucleotides, containing a non-naturally occurring nucleobase instead of a naturally occurring nucleobase such as uridines or cytidines modified at the 5-position, e.g. 5-(2-amino)propyl uridine, 5-bromo uridine; 5-methyl-cytidine; adenosines and guanosines modified at the 8-position, e.g. 8-bromo guanosine; deaza nucleotides, e.g. 7-deaza-adenosine; O- and N-alkylated nucleotides, e.g. N6-methyl adenosine are suitable. In preferred sugar-modified ribonucleotides the 2' OH-group is replaced by a group selected from H, OR, R, halo, SH, SR, NH₂, NHR, NR₂ or CN, wherein R is C₁-C₆ alkyl, alkenyl, alkynyl or methoxyethoxy, and halo is F, Cl, Br or I. In preferred backbone-modified ribonucleotides the phosphoester group connecting to adjacent ribonucleotides is replaced by a modified group, e.g. a phosphorothioate, phosphorodithioate, N3'-O5'- and/or N5'-O3' phosphoramidate group. It should be noted that the above modifications may be combined. For example, complementary or non-complementary nucleotides at the 3'-terminus, particularly after at least 15, more particularly after at least 20 5'-terminal nucleotides may be modified without significant loss of activity.

- 8 -

The single-stranded RNA molecule of the invention may be prepared by chemical synthesis. Methods of synthesizing RNA molecules are known in the art.

5 The single-stranded RNAs can also be prepared by enzymatic transcription from synthetic DNA templates or from DNA plasmids isolated from recombinant bacteria and subsequent 5'-terminal modification. Typically, phage RNA polymerases are used such as T7, T3 or SP6 RNA polymerase.

10 A further aspect of the present invention relates to a method of mediating RNA interference in a cell or an organism comprising the steps:

(a) contacting the cell or organism with the single-stranded RNA molecule of the invention under conditions wherein target-specific nucleic acid modifications may occur and

15

(b) mediating a target-specific nucleic acid modification effected by the single-stranded RNA towards a target nucleic acid having a sequence portion substantially complementary to the single-stranded RNA.

20

Preferably the contacting step (a) comprises introducing the single-stranded RNA molecule into a target cell, e.g. an isolated target cell, e.g. in cell culture, a unicellular microorganism or a target cell or a plurality of target cells within a multicellular organism. More preferably, the introducing step

25

comprises a carrier-mediated delivery, e.g. by liposomal carriers and/or by injection. Further suitable delivery systems include Oligofectamine (Invitrogen) and Transit-TKO siRNA Transfection reagent (Mirus)

30 The method of the invention may be used for determining the function of a gene in a cell or an organism or even for modulating the function of a gene in a cell or an organism, being capable of mediating RNA interference.

- 9 -

The cell is preferably a eukaryotic cell or a cell line, e.g. a plant cell or an animal cell, such as a mammalian cell, e.g. an embryonic cell, a pluripotent stem cell, a tumor cell, e.g. a teratocarcinoma cell or a virus-infected cell. The organism is preferably a eukaryotic organism, e.g. a plant or an animal,
5 such as a mammal, particularly a human.

The target gene to which the RNA molecule of the invention is directed may be associated with a pathological condition. For example, the gene may be a pathogen-associated gene, e.g. a viral gene, a tumor-associated
10 gene or an autoimmune disease-associated gene. The target gene may also be a heterologous gene expressed in a recombinant cell or a genetically altered organism. By determining or modulating, particularly, inhibiting the function of such a gene valuable information and therapeutic benefits in the agricultural field or in the medicine or veterinary medicine field may
15 be obtained.

The ssRNA is usually administered as a pharmaceutical composition. The administration may be carried out by known methods, wherein a nucleic acid is introduced into a desired target cell in vitro or in vivo. Commonly
20 used gene transfer techniques include calcium phosphate, DEAE-dextran, electroporation and microinjection and viral methods (Graham, F.L. and van der Eb, A.J. (1973) Virol. 52, 456; McCutchan, J.H. and Pagano, J.S. (1968), J. Natl. Cancer Inst. 41, 351; Chu, G. et al (1987), Nucl. Acids Res. 15, 1311; Fraley, R. et al. (1980), J. Biol. Chem. 255, 10431; Capecchi, M.R. (1980), Cell 22, 479). A recent addition to this arsenal of techni-
25 ques for the introduction of nucleic acids into cells is the use of cationic liposomes (Felgner, P.L. et al. (1987), Proc. Natl. Acad. Sci USA 84, 7413). Commercially available cationic lipid formulations are e.g. Tfx 50 (Promega) or Lipofectamin2000 (Life Technologies). A further preferred
30 method for the introduction of RNA into a target organism, particularly into a mouse, is the high-pressure tail vein injection (Lewis, D.L. et al. (2002), Nat.Genet.29, 29; McCaffrey, A.P. et al. (2002), Nature 418, 38-39).

- 10 -

Herein, a buffered solution comprising the single-stranded RNA (e.g. about 2 ml) is injected into the tail vein of the mouse within 10 s.

Thus, the invention also relates to a pharmaceutical composition containing
5 as an active agent at least one single-stranded RNA molecule as described above and a pharmaceutical carrier. The composition may be used for diagnostic and for therapeutic applications in human medicine or in veterinary medicine.

10 For diagnostic or therapeutic applications, the composition may be in form of a solution, e.g. an injectable solution, a cream, ointment, tablet, suspension or the like. The composition may be administered in any suitable way, e.g. by injection, by oral, topical, nasal, rectal application etc. The carrier
15 may be any suitable pharmaceutical carrier. Preferably, a carrier is used, which is capable of increasing the efficacy of the RNA molecules to enter the target-cells. Suitable examples of such carriers are liposomes, particularly cationic liposomes. A further preferred administration method is injection.

20 A further preferred application of the RNAi method is a functional analysis of eukaryotic cells, or eukaryotic non-human organisms, preferably mammalian cells or organisms and most preferably human cells, e.g. cell lines such as HeLa or 293 or rodents, e.g. rats and mice. By transfection with
25 suitable single-stranded RNA molecules which are homologous to a predetermined target gene or DNA molecules encoding a suitable single-stranded RNA molecule a specific knockout phenotype can be obtained in a target cell, e.g. in cell culture or in a target organism. The presence of short single-stranded RNA molecules does not result in an interferon response from the host cell or host organism.

30 In an especially preferred embodiment, the RNA molecule is administered associated with biodegradable polymers, e.g. polypeptides, poly(d,l-lactic-

- 11 -

co-glycolic acid) (PLGA), polylysine or polylysine conjugates, e.g. polylysine-graft-imidazole acetic acid, or poly(beta-amino ester) or microparticles, such as microspheres, nanoparticles or nanospheres. More preferably the RNA molecule is covalently coupled to the polymer or
5 microparticle, wherein the covalent coupling particularly is effected via the 3'-terminus of the RNA molecule.

Further, the invention relates to a pharmaceutical composition for inhibiting the expression of a target transcript by RNAi comprising as an active agent
10 a single-stranded RNA molecule having a length from 14 - 50, preferably 15 - 29 nucleotides wherein at least the 14 - 20 5'-most nucleotides are substantially complementary to said target transcript.

Furthermore, the invention relates to a method for the prevention or
15 treatment of a disease associated with overexpression of at least one target gene comprising administering a subject in need thereof a single-stranded RNA molecule having a length from 14 - 50, preferably 15-29 nucleotides wherein at least the 14 - 20 5'-most nucleotides are substantially complementary to a target transcript in an amount which is
20 therapeutically effective for RNAi.

Still, a further subject matter of the invention is a eukaryotic cell or a eukaryotic non-human organism exhibiting a target gene-specific knockout phenotype comprising an at least partially deficient expression of at least
25 one endogenous target gene wherein said cell or organism is transfected with at least one single-stranded RNA molecule capable of inhibiting the expression of at least one endogenous target gene. It should be noted that the present invention allows the simultaneous delivery of several antisense RNAs of different sequences, which are either cognate to a
30 different or the same target gene.

- 12 -

Gene-specific knockout phenotypes of cells or non-human organisms, particularly of human cells or non-human mammals may be used in analytic procedures, e.g. in the functional and/or phenotypical analysis of complex physiological processes such as analysis of gene expression profiles and/or proteomes. For example, one may prepare the knock-out phenotypes of human genes in cultured cells which are assumed to be regulators of alternative splicing processes. Among these genes are particularly the members of the SR splicing factor family, e.g. ASF/SF2, SC35, SRp20, SRp40 or SRp55. Further, the effect of SR proteins on the mRNA profiles of predetermined alternatively spliced genes such as CD44 may be analysed. Preferably the analysis is carried out by high-throughput methods using oligonucleotide based chips.

Using RNAi based knockout technologies, the expression of an endogenous target gene may be inhibited in a target cell or a target organism. The endogeneous gene may be complemented by an exogeneous target nucleic acid coding for the target protein or a variant or mutated form of the target protein, e.g. a gene or a cDNA, which may optionally be fused to a further nucleic acid sequence encoding a detectable peptide or polypeptide, e.g. an affinity tag, particularly a multiple affinity tag. Variants or mutated forms of the target gene differ from the endogeneous target gene in that they encode a gene product which differs from the endogeneous gene product on the amino acid level by substitutions, insertions and/or deletions of single or multiple amino acids. The variants or mutated forms may have the same biological activity as the endogeneous target gene. On the other hand, the variant or mutated target gene may also have a biological activity, which differs from the biological activity of the endogeneous target gene, e.g. a partially deleted activity, a completely deleted activity, an enhanced activity etc.

30

The complementation may be accomplished by coexpressing the polypeptide encoded by the exogeneous nucleic acid, e.g. a fusion protein com-

- 13 -

prising the target protein and the affinity tag and the double stranded RNA molecule for knocking out the endogeneous gene in the target cell. This coexpression may be accomplished by using a suitable expression vector expressing both the polypeptide encoded by the exogeneous nucleic acid, e.g. the tag-modified target protein and the single-stranded RNA molecule or alternatively by using a combination of expression vectors. Proteins and protein complexes which are synthesized de novo in the target cell will contain the exogeneous gene product, e.g. the modified fusion protein. In order to avoid suppression of the exogeneous gene product expression by the RNAi molecule, the nucleotide sequence encoding the exogeneous nucleic acid may be altered on the DNA level (with or without causing mutations on the amino acid level) in the part of the sequence which is homologous to the single-stranded RNA molecule. Alternatively, the endogeneous target gene may be complemented by corresponding nucleotide sequences from other species, e.g. from mouse.

Preferred applications for the cell or organism of the invention is the analysis of gene expression profiles and/or proteomes. In an especially preferred embodiment an analysis of a variant or mutant form of one or several target proteins is carried out, wherein said variant or mutant forms are reintroduced into the cell or organism by an exogeneous target nucleic acid as described above. The combination of knockout of an endogeneous gene and rescue by using mutated, e.g. partially deleted exogeneous target has advantages compared to the use of a knockout cell. Further, this method is particularly suitable for identifying functional domains of the target protein. In a further preferred embodiment a comparison, e.g. of gene expression profiles and/or proteomes and/or phenotypic characteristics of at least two cells or organisms is carried out. These organisms are selected from:

- (i) a control cell or control organism without target gene inhibition,
- (ii) a cell or organism with target gene inhibition and

- 14 -

- (iii) a cell or organism with target gene inhibition plus target gene complementation by an exogenous target nucleic acid.

5 The method and cell of the invention may also be used in a procedure for identifying and/or characterizing pharmacological agents, e.g. identifying new pharmacological agents from a collection of test substances and/or characterizing mechanisms of action and/or side effects of known pharmacological agents.

10 Thus, the present invention also relates to a system for identifying and/or characterizing pharmacological agents acting on at least one target protein comprising:

- (a) a eukaryotic cell or a eukaryotic non-human organism capable of expressing at least one endogenous target gene coding for said target protein,
- 15 (b) at least one single-stranded RNA molecule capable of inhibiting the expression of said at least one endogenous target gene by RNAi and
- (c) a test substance or a collection of test substances wherein pharmacological properties of said test substance or said collection are to
- 20 be identified and/or characterized.

Further, the system as described above preferably comprises:

- (d) at least one exogenous target nucleic acid coding for the target protein or a variant or mutated form of the target protein wherein
- 25 said exogenous target nucleic acid differs from the endogenous target gene on the nucleic acid level such that the expression of the exogenous target nucleic acid is substantially less inhibited by the single-stranded RNA molecule than the expression of the endogenous target gene.
- 30

- 15 -

Furthermore, the RNA knockout complementation method may be used for preparative purposes, e.g. for the affinity purification of proteins or protein complexes from eukaryotic cells, particularly mammalian cells and more particularly human cells. In this embodiment of the invention, the exogenous target nucleic acid preferably codes for a target protein which is fused to an affinity tag.

The preparative method may be employed for the purification of high molecular weight protein complexes which preferably have a mass of $\geq$ 150 kD and more preferably of $\geq$ 500 kD and which optionally may contain nucleic acids such as RNA. Specific examples are the heterotrimeric protein complex consisting of the 20 kD, 60 kD and 90 kD proteins of the U4/U6 snRNP particle, the splicing factor SF3b from the 17S U2 snRNP consisting of 5 proteins having molecular weights of 14, 49, 120, 145 and 155 kD and the 25S U4/U6/U5 tri-snRNP particle containing the U4, U5 and U6 snRNA molecules and about 30 proteins, which has a molecular weight of about 1.7 MD.

This method is suitable for functional proteome analysis in mammalian cells, particularly human cells.

Finally, the invention relates to a purified and isolated mammalian, particularly human RNA-induced silencing complex (RISC) having an apparent molecular weight of less than about 150-160 kDa, e.g. about 120 to 150-160 kDa. The RISC comprises polypeptide and optionally nucleic acid components, particularly single-stranded RNA molecules as described above. The RISC may be used as a target for diagnosis and/or therapy, as a diagnostic and/or therapeutic agent itself, as a molecular-biological reagent or as component in a screening procedure for the identification and/or characterization of pharmaceutical agents.

- 15a -

The present description provides the use of single-stranded RNA molecule having a length from 14-50 nucleotides wherein at least the 14-20 5' most nucleotides are complementary to a target transcript for the manufacture of an agent for inhibiting the expression of the target transcript in mammalian cells, the RNA molecule having a phosphate analogue at its 5' terminus, wherein the expression is inhibited by RNA-interference.

The present description also provides the use as disclosed herein for diagnosis of a disease or condition which is amenable to the use of a single-stranded RNA molecule having a length of 14-50 nucleotides, wherein the 5' most nucleotides are complementary to a target transcript, wherein the RNA molecule has a phosphate analog at its 5' terminus to inhibit the expression of a target transcript by RNA interference.

The present description further provides the use as disclosed herein for therapy of a disease or condition which is amenable to the use of a single-stranded RNA molecule having a length of 14-50 nucleotides, wherein the 5' most nucleotides are complementary to a target transcript, wherein the RNA molecule has a phosphate analog at its 5' terminus to inhibit the expression of a target transcript by RNA interference.

- 16 -

Polypeptide components of RISC preferably comprise members of the Argonaute family of proteins, and contain eIF2C1 and/or eIF2C2, and possibly at least one other expressed eIF2C family member, particularly selected from eIF2C3, eIF2C4, HILI and HIWI.

5

Expression or overexpression of one or several proteins present in RISC in suitable host cells, e.g. eukaryotic cells, particularly mammalian cells, is useful to assist an RNAi response. These proteins may also be expressed or overexpressed in transgenic animals, e.g. vertebrates, particularly mammals, to produce animals particularly sensitive to injected single-stranded or double-stranded siRNAs. Further, the genes encoding the proteins may be administered for therapeutic purposes, e.g. by viral or non-viral gene delivery vectors.

15 It is also conceivable to administer a siRNA/eIF2C1 or 2 complex directly by the assistance of protein transfection reagents (e.g. Amphoteric Protein Transfection Reagents, ProVectin protein (Imgenex), or similar products) rather than RNA/DNA transfection. This may have technical advantages over siRNA transfection that are limited to nucleic acid transfection.

20

Alternatively to the application of siRNAs as synthetic double-stranded or single-stranded siRNAs, it is conceivable to also administer an antisense siRNA precursor molecule in the form of a hairpin stem-loop structure comprising 19 to 29 base pairs in the stem with or without 5' or 3' overhanging ends on one side of the duplex and a nucleotide or non-nucleotide loop on the other end. Preferably, the hairpin structure has a 3' overhang of from 1-5 nucleotides. Further, the precursor may contain modified nucleotides as described above, particularly in the loop and/or in the 3' portion, particularly in the overhang. The siRNA or precursors of siRNAs may also be introduced by viral vectors or RNA expression systems into a RISC compound, e.g. eIF2C1 and/or 2 overexpressing organism or cell line. The siRNA precursors may also be generated by direct expression

25
30

- 17 -

within an organism or cell line. This may be achieved by transformation with a suitable expression vector carrying a nucleic acid template operatively linked to an expression control sequence to express the siRNA precursor.

5

Further, the present invention is explained in more detail in the following figures and examples.

10 Figure legends

Figure 1. HeLa cytoplasmic S100 extracts show siRNA-dependent target RNA cleavage.

(A) Representation of the 177-nt ³²P-cap-labeled target RNA with the targeting siRNA duplex. Target RNA cleavage site and the length of the expected cleavage products is also shown. The fat black line positioned under the antisense siRNA is used in the following figures as symbol to indicate the region of the target RNA, which is complementary to the antisense siRNA sequence. (B) Comparison of the siRNA mediated target RNA cleavage using the previously established *D. melanogaster* embryo in vitro system and HeLa cell S100 cytoplasmic extract. 10 nM cap-labeled target RNA was incubated with 100 nM siRNA as described in materials. Reaction products were resolved on a 6% sequencing gel. Position markers were generated by partial RNase T1 digestion (T1) and partial alkaline hydrolysis (OH) of the cap-labeled target RNA. The arrow indicates the 5' cleavage product, the 3' fragment is unlabeled and therefore invisible.

25

30

Figure 2. Chemical modification of the 5' end of the antisense but not the sense siRNAs prevents sense target RNA cleavage in HeLa S100 extracts.

(A) Illustration of the possible 5' and 3' aminolinker modifications of the sense and antisense strands of a siRNA duplex. L5 represents a 6-carbon chain aminolinker connected via a 5'-phosphodiester linkage, L3 represents

- 18 -

a 7-carbon aminolinker connected via a phosphodiester bond to the terminal 3' phosphate. s, sense; as, antisense. (B) Target RNA cleavage testing various combinations of 5' and 3' aminolinker-modified siRNA duplexes. NC (negative control) shows an incubation reaction of the target RNA in the absence of siRNA duplex. T1, RNase T1 ladder; OH, partial alkaline hydrolysis ladder.

Figure 3. siRNA containing 3'-terminal phosphates are subjected to ligation as well as dephosphorylation reactions.

10 (A) Sequence of the radiolabeled siRNA duplex. The labeled nucleotide was joined to synthetic 20-nt antisense siRNA by T4 RNA ligation of ³²pCp. The various combinations of 5' and 3' hydroxyl/phosphate were prepared as described in materials. X and Y indicate 5' and 3' modifications of the antisense siRNA. (B) Fate of the antisense siRNA during incubation of the modified siRNA duplexes in HeLa S100 extract in the presence of non-radiolabeled target RNA. The different phosphorylated forms of the antisense siRNA were distinguished based on their gel mobility. Identical results were obtained when using 5' phosphorylated sense siRNA or when leaving out the target RNA during incubation. Ligation products are only observed when 3' phosphates were present on the labeled antisense siRNA.

Figure 4. RISC is a stable complex that does not rapidly exchange bound siRNA.

25 Increasing concentrations of non-specific siRNA compete with target-specific RISC formation when added simultaneously to HeLa S100 extracts (lanes 4 to 7). However, when the unspecific siRNA duplex is added 15 min after pre-incubation with the specific siRNA duplex, no more competition was observed (3 lanes to the right). T1, RNase T1 ladder.

30

Figure 5. Partial purification of human RISC.

- 19 -

(A) Graphical representation of the structure of the biotinylated siRNA duplex used for affinity purification of siRNA-associated factors. L3 indicates a C7-aminolinker that was conjugated to a photo-cleavable biotin N-hydroxysuccinimidyl ester; UV indicates photocleavage of the UV-sensitive linkage to release affinity selected complexes under native conditions. (B) Superdex-200 gel filtration analysis of siRNA-protein complexes (siRNPs) recovered by UV treatment/elution (UV elu) from the streptavidin affinity column. Fractions were assayed for their ability to sequence-specifically cleave the cap-labeled target RNA. The number of the collected fractions and the relative positions of the aldolase (158 kDa) and BSA (66 kDa) size markers are indicated. (C) Glycerol gradient (5%-20%) sedimentation of siRNPs recovered by UV treatment/elution from the streptavidin affinity column. For legend, see (B). When monitoring the precise size of target RNA cleavage fragments using internally ³²P-UTP-labeled, capped mRNA, the sum is equal to the full-length transcript, thus indicating that target RNA is indeed only cleaved once in the middle of the region spanned by the siRNA.

Figure 6. RISC contains a single-stranded siRNA. siRNPs were subjected to affinity selection after incubation using siRNA duplexes with one or both strands biotinylated. The eluate recovered after UV treatment or the unbound fraction after streptavidin affinity selection (flow-through) was assayed for target RNA degradation. If the antisense strand was biotinylated, all sense target RNA-cleaving RISC was bound to the streptavidin beads, while sense siRNA biotinylation resulted in RISC activity of the flow-through. The cleavage reaction in the flow-through fraction was less efficient than in the UV eluate, because affinity-selected RISC was more concentrated.

Figure 7. Single-stranded antisense siRNAs reconstitute RISC in HeLa S100 extracts.

- 20 -

Analysis of RISC reconstitution using single-stranded or duplex siRNAs comparing HeLa S100 extracts (A) and the previously described *D. melanogaster* embryo lysate (B). Different concentrations of single-stranded siRNAs (s, sense; as, antisense) and duplex siRNA (ds) were tested for specific targeting of cap-labeled substrate RNA. 100 nM concentrations of the antisense siRNA reconstituted RISC in HeLa S100 extract, although at reduced levels in comparison to the duplex siRNA. Reconstitution with single-stranded siRNAs was almost undetectable in *D. melanogaster* lysate, presumably because of the higher nuclease activity in this lysate causing rapid degradation of uncapped single-stranded RNAs .

Figure 8. Single-stranded antisense siRNAs mediate gene silencing in HeLa cells.

(A) Silencing of nuclear envelope protein lamin A/C. Fluorescence staining of cells transfected with lamin A/C-specific siRNAs and GL2 luciferase (control) siRNAs. Top row, staining with lamin A/C specific antibody; middle row, Hoechst staining of nuclear chromatin; bottom row, phase contrast images of fixed cells. (B) Quantification of lamin A/C knockdown after Western blot analysis. The blot was stripped after lamin A/C probing and reprobbed with vimentin antibody. Quantification was performed using a Lumi-Imager (Roche) and LumiAnalyst software to quantitate the ECL signals (Amersham Biosciences), differences in gel loading were corrected relative to non-targeted vimentin protein levels. The levels of lamin A/C protein were normalized to the non-specific GL2 siRNA duplex.

Figure 9. Antisense siRNAs of different length direct target RNA cleavage in HeLa S100 extracts.

(A) Graphical representation of the experiment. Antisense siRNAs were extended towards the 5' side (series 1, 20 to 25-nt) or the 3' side (series 2, 20 to 23-nt).

(B) Target RNA cleavage using the antisense siRNAs described in (A). HeLa S100 extract was incubated with 10 nM cap-labeled target RNA and 100

- 21 -

nM antisense siRNAs at 30 °C for 2.5 h. Reaction products were resolved on a 6 % sequencing gel. Position markers were generated by partial RNase T1 digestion (T1) and partial alkaline hydrolysis (OH) of the cap-labeled target RNA. Arrows indicate the position of the 5' cleavage products generated by the different antisense siRNAs. The fat black lines on the left (series 1) and the right (series 2) indicate the region of the target RNA, which is complementary to the antisense siRNA sequences.

Figure 10. Length dependence of antisense siRNAs and effect of terminal modifications for targeting RNA cleavage in HeLa S100 extracts.

HeLa S100 extract was incubated with 10 nM cap-labeled target RNA and 100 nM antisense siRNAs at 30°C for 2.5h. Reaction products were resolved on a 6% sequencing gel. Position markers were generated by partial RNase T1 digestion (T1) of the cap-labeled target RNA. The fat black line on the left indicates the region of the target RNA, which is complementary to the 21-nt antisense siRNA sequence. The siRNA sequences used in each experiment are listed below (sense and antisense siRNAs are listed together, they were pre-annealed to form duplex siRNAs). p, phosphate; t, 2'-deoxythymidine, c, 2'-deoxycytidine, g, 2'-deoxycytidine, g, 2'-deoxyguanosine; L, aminolinker, B, photocleavable biotin; A,C,G,U, ribonucleotides.

Lane	Sense siRNA (5' – 3')	Antisense siRNA (5' – 3')
1		pUCGAAGUAUCCG CG
2		pUCGAAGUAUCCG CGUACGUG

- 22 -

5

10

3		pUCGAAGUAUUCCG CGUACGUGAUGU
4		pUCGAAGUAUUCCG CGUACGUGAUGUUC
5		pUCGAAGUAUUCCG CGUACGUGAUGUUC AC
6		pUCGAAGUAUUCCG CG
7		pUCGAAGUAUUCCG CGUACGUG
8		pUCGAAGUAUUCCG CGUACGUGAUGU
9		pUCGAAGUAUUCCG CGUACGUGAUGUUC
10		pUCGAAGUAUUCCG CGUACGUGAUGUUC AC
11		pUCGAAGUAUUCCG CGUACGUG
12		pUCGAAGUAUUCCG CGUACGtg
13		pUCGAAGUAUUCCG CGUACGUU
14		pUCGAAGUAUUCCG CGUACGtt

- 23 -

5

10

15		pUCGAAGUAUUCCG CGUACGUG
16		pUCGAAGUAUUCCG CGUACGtg
17		pUCGAAGUAUUCCG CGUACGUU
18		pUCGAAGUAUUCCG CGUACGtt
19	CGUACGCGGAAUACUUCG AAA	pUCGAAGUAUUCCG CGUACGUG
20	CGUACGCGGAAUACUUCG AAA	pUCGAAGUAUUCCG CGUACGtg
21	CGUACGCGGAAUACUUCG AAA	pUCGAAGUAUUCCG CGUACGUU
22	CGUACGCGGAAUACUUCG AAA	pUCGAAGUAUUCCG CGUACGtt
23		tCGAAGUAUUCCGC GUACGUULB
24	cGUACGCGGAAUACUUCG AUULB	tCGAAGUAUUCCGC GUACGUULB
25		ptCGAAGUAUUCCGC GUACGttLB
26	cGUACGCGGAAUACUUCG AttLB	ptCGAAGUAUUCCGC GUACGttLB

- 24 -

27		ptCGAAGUAUCCGC GUACGttL
----	--	----------------------------

Figure 11: Single-stranded antisense siRNAs mediate gene silencing in HeLa cells.

5 Quantification of lamin A/C knockdown after Western blot analysis. The blot was stripped after lamin A/C probing and reprobed with vimentin antibody. Quantification was performed using a Lumi-Imager (Roche) and LumiAnalyst software to quantitate the ECL signals (Amersham Biosciences), differences in gel loading were corrected relative to
10 non-targeted vimentin protein levels. The levels of lamin A/C protein were normalized to the non-specific GL2 siRNA duplex.

Figure 12. Protein composition of affinity purified RISC.

(A) Silver-stained SDS-PAGE gel of affinity-selected ribonucleoprotein
15 complexes after glycerol gradient (5%-20%) sedimentation. The arrow indicates the band containing eIF2C1 and eIF2C2. Molecular size markers are indicated on the left. The asterisk indicates a fraction for which the protein pellet was lost after precipitation. (B) Target RNA cleavage assay of the collected fractions. RISC activity peaked in fraction 7 and 8; bu, buffer.

20

Figure 13. Mass spectrometric characterization of eIF2C1 and eIF2C2.

The 100 kDa band was analysed by mass spectrometry. Mass spectrum indicating the peptide peaks corresponding to eIF2C2 (A) and eIF2C1 (B).
(C) Alignment of eIF2C2 and eIF2C1 amino-acid sequences indicating the
25 position of the identified peptides. Sequence differences are indicated by yellow boxes.

Figure 14. Predicted amino-acid sequences of the six human Argonaute protein family members.

30

- 25 -

Figure 15. Alignment of the sequences of the six human Argonaute protein family members.

Predicted sequences of human eIF2C1-4, HILI and HIWI have been aligned using ClustalX program.

5

Figure 16. Predicted cDNA sequences of the six human Argonaute protein family members.

10

Figure 17. All members of the Argonaute family but HIWI are expressed in HeLa cells.

15

RT-PCR analysis on polyA RNA from HeLa cells. (A) Primers (forward and reverse) used for nested and semi-nested PCR amplification of the different Argonautes and expected length of the PCR products. (B) Agarose gel electrophoresis of the obtained PCR products, confirming the expected length. Left lanes, 100 bp DNA ladder.

20

Example

1. Material and methods

25

1.1 siRNA synthesis and biotin conjugation

30

siRNAs were chemically synthesized using RNA phosphoramidites (Proligo, Hamburg, Germany) and deprotected and gel-purified as described previously. 5' aminolinkers were introduced by coupling MMT-C6-aminolinker phosphoramidite (Proligo, Hamburg), 3' C7-aminolinkers were introduced by assembling the oligoribonucleotide chain on 3'-aminomodifier (TFA) C7 Icaa control pore glass support (Chemgenes, MA, USA). The sequences for GL2 luciferase siRNAs were as

- 26 -

described (Elbashir et al., 2001a, supra). If 5'-phosphates were to be introduced, 50 to 100 nmoles of synthetic siRNAs were treated with T4 polynucleotide kinase (300 μ l reaction, 2.5 mM ATP, 70 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM DTT, 30 U T4 PNK, New England Biolabs, 45 min, 37 °C) followed by ethanol precipitation.

3' Terminal ^{32}P Cp labeling (Figure 3) was performed in a 30 μ l reaction (17 μ M siRNA, 0.5 μ M ^{32}P Cp (110 TBq/mmol), 15% DMSO, 20 U T4 RNA ligase, NEB, and 1x NEB-supplied reaction buffer) for 1.5 h at 37 °C, and gel-purified. One half of the pCp-labeled RNA was dephosphorylated (25 μ l reaction, 500 U alkaline phosphatase, Roche, and Roche-supplied buffer, 30 min, 50 °C), followed by phenol/chloroform extraction and ethanol precipitation. Half of this reaction was 5' phosphorylated (20 μ l reaction, 2 units T4 polynucleotide kinase, NEB, 10 mM ATP, NEB-supplied buffer, 60 min, 37 °C). A quarter of the initial pCp-labeled siRNA was also 5' phosphorylated (10 μ l reaction, 10 units 3' phosphatase-free T4 polynucleotide kinase, Roche, 10 mM ATP, Roche-supplied buffer, 3 min, 37 °C).

For conjugation to biotin, 20 to 65 nmoles of fully deprotected aminolinker-modified siRNA were dissolved in 100 μ l of 100 mM sodium borate buffer (pH 8.5) and mixed with a solution of 1 mg of EZ-Link NHS-PC-LC-Biotin (Pierce, IL, USA) in 100 μ l of anhydrous dimethylformamide. The solution was incubated for 17 h at 25 °C in the dark. Subsequently, siRNAs were precipitated by the addition of 60 μ l 2 M sodium acetate (pH 6.0) and 1 ml ethanol. The RNA pellet was collected by centrifugation and biotin-conjugated siRNA was separated from non-reacted siRNA on a preparative denaturing 18% acrylamide gel (40 cm length) in the dark. The RNA bands were visualized by 254 nm UV shadowing and minimized exposure time. The bands were excised, and the RNA was eluted overnight in 0.3 M NaCl at 4 °C and recovered by ethanol

- 27 -

precipitation. siRNA duplexes were formed as previously described (Elbashir et al., Methods 26 (2002), 199 - 213).

1.2 Preparation of S100 extracts from HeLa cells

5 Cytoplasm from HeLa cells adapted to grow at high density was prepared following the Dignam protocol for isolation of HeLa cell nuclei (Dignam et al., Nucleic Acids Res. 11 (1983), 1475 - 1489). The cytoplasmic fraction was supplemented with KCl, MgCl₂ and glycerol to final concentrations of 100 mM, 2 mM and 10%, respectively. At this stage, the extracts can be
10 stored frozen at -70 °C after quick-freezing in liquid nitrogen without loss of activity. S100 extracts were prepared by ultracentrifugation at 31.500 rpm for 60 minutes at 4 °C using a Sorvall T-865 rotor. The protein concentration of HeLa S100 extract varied between 4 to 5 mg/ml as determined by Bradford assay.

15

1.3 Affinity purification of RISC with 3' biotinylated siRNA duplexes

For affinity purification of siRNA-associated protein complexes from HeLa S100 extracts, 10 nM of a 3' double-biotinylated siRNA duplex were incubated in 0.2 mM ATP, 0.04 mM GTP, 10 U/ml RNasin, 6 µg/ml
20 creatine kinase, and 5 mM creatine phosphate in 60% S100 extract at 30 °C for 30 to 60 min and gentle rotation. Thereafter, 1 ml slurry of Immobilized Neutravidin Biotin Binding Protein (Pierce, IL, USA) was added per 50 ml of reaction solution and the incubation was continued for another 60 to 120 min at 30 °C with gentle rotation. The Neutravidin
25 beads were then collected at 2000 rpm for 2 minutes at 4 °C in a Heraeus Megafuge 1.0 R centrifuge using a swinging bucket rotor type 2704. Effective capturing of RISC components after affinity selection was confirmed by assaying the supernatant for residual RISC activity with and without supplementing fresh siRNA duplexes. The collected Neutravidin
30 beads were washed with 10 volumes of buffer A relative to the bead volume (30 mM HEPES, pH 7.4, 100 mM KCl, 2 mM MgCl₂, 0.5 mM DTT, 10% glycerol) followed by washing with 5 volumes of buffer B (same as

- 28 -

buffer A with only 3% glycerol content). The beads were transferred to a 0.8 x 4 cm Poly-Prep chromatography column (BioRad; CA, USA) by resuspending in 3 volumes of buffer B at 4 °C, followed by 10 volumes of washing with buffer B. Washing of the beads was continued by 10
5 volumes of buffer B increased to 300 mM KCl. The column was then reequilibrated with regular buffer B. To recover native siRNA-associated complexes, the column was irradiated in the cold room by placing it at a 2 cm distance surrounded by four 312 nm UV lamps (UV-B tube, 8 W, Herolab, Germany) for 30 minutes. To recover the photocleaved siRNP
10 solution, the column was placed into a 50 ml Falcon tube and centrifuged at 2000 rpm for 1 minute at 4 °C using again the 2704 rotor. For full recovery of siRNPs, the beads were once again resuspended in buffer B followed by a second round of UV treatment for 15 minutes. Both eluates were pooled and assayed for target RNA degradation.

15

1.4 Target RNA cleavage assays

Cap-labeled target RNA of 177 nt was generated as described (Elbashir et al., EMBO J. 20 (2001 c), 6877 - 6888) except that his-tagged guanylyl transferase was expressed in *E. coli* from a plasmid generously provided by
20 J. Wilusz and purified to homogeneity. If not otherwise indicated, 5' phosphorylated siRNA or siRNA duplex was pre-incubated in supplemented HeLa S100 extract at 30 °C for 15 min prior to addition of cap-labeled target RNA. After addition of all components, final concentrations were 100 nM siRNA, 10 nM target RNA, 1 mM ATP, 0.2 mM GTP, 10 U/ml
25 RNasin, 30 µg/ml creatine kinase, 25 mM creatine phosphate, 50% S100 extract. Incubation was continued for 2.5 h. siRNA-mediated target RNA cleavage in *D. melanogaster* embryo lysate was performed as described (Zamore et al., Cell 101 (2000), 25 - 33). Affinity-purified RISC in buffer B was assayed for target RNA cleavage without preincubation nor addition of
30 extra siRNA (10 nM target RNA, 1 mM ATP, 0.2 mM GTP, 10 U/ml RNasin, 30 µg/ml creatine kinase, 25 mM creatine phosphate, 50% RISC in buffer B). Cleavage reactions were stopped by the addition of 8 vols of

- 29 -

proteinase K buffer (200 mM Tris-HCl pH 7.5, 25 mM EDTA, 300 mM NaCl, 2% w/v SDS). Proteinase K, dissolved in 50 mM Tris-HCl pH 8.0, 5 mM CaCl₂, 50% glycerol, was added to a final concentration of 0.6 mg/ml and processed as described (Zamore et al. (2000), supra). Samples were
5 separated on 6% sequencing gels.

1.5 Analytical Gel Filtration

UV-eluates in buffer B were fractionated by gel filtration using a Superdex 200 PC 3.2/30 column (Amersham Biosciences) equilibrated with buffer A
10 on a SMART system (Amersham Biosciences). Fractionation was performed by using a flow rate of 40 μ l/minute and collecting 100 μ l fractions. Fractions were assayed for specific target RNA cleavage. Size calibration was performed using molecular size markers thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa) and BSA (66
15 kDa) (Amersham Biosciences).

1.6 Glycerol gradient sedimentation

UV-eluates were layered on top of 4 ml linear 5% to 20% (w/w) glycerol gradient adjusted to 30 mM HEPES, pH 7.4, 100 mM KCl, 2 mM MgCl₂,
20 0.5 mM DTT. Centrifugation was performed at 35000 rpm for 14.5 h at 4 °C using a Sorvall SW 60 rotor. Twenty fractions of 0.2 ml volume were removed sequentially from the top and 15 μ l aliquots were used to assay for target RNA cleavage.

25

2. Results

2.1 A human biochemical system for siRNA functional analysis

We were interested in assaying siRNA-mediated target RNA degradation in
30 human cell extracts, because siRNAs are powerful reagents to knockdown gene expression in human cells but the action of siRNAs in human cells was uncertain. To investigate whether siRNAs guide target RNA

- 30 -

degradation in human cells with a similar mechanism to the one observed in *D. melanogaster* (e.g. Elbashir et al. (2001 b), supra), we prepared substrates for targeted mRNA degradation as described previously (Elbashir et al. (2001 c), supra). A 5'-³²P-cap-labeled, 177-nt RNA transcript, derived from a segment of the firefly luciferase gene, was incubated in HeLa cell S100 or *D. melanogaster* embryo extracts with a 21-nt siRNA duplex in the presence of an ATP regeneration system (Figure 1A, B). siRNA cleavage assays were performed at 25 °C in *D. melanogaster* lysate and at 30 °C in HeLa S100 extracts for 2.5 h. After deproteinization using proteinase K, the reaction products were separated on a 6% sequencing gel.

Similar to the previous observation in *D. melanogaster* lysate, we observed the appearance of a cleavage product in HeLa S100 extract at exactly the same position, thus indicating that the siRNA duplex guides target RNA cleavage in the human system with the same specificity and mechanism. The cleavage reaction appeared less efficient when compared to the *D. melanogaster* system, but this could be explained by the 5-fold lower total protein concentration of HeLa S100 extracts (25 mg/ml vs. 5 mg/ml). Similar to *D. melanogaster* lysates, siRNA duplexes without 5' phosphate were rapidly 5' phosphorylated in HeLa S100 extracts (see below) and the ability to cleave the target RNA was independent of the presence of a 5' phosphate on the synthetic siRNA duplexes.

Comparative analysis of the efficiency of siRNA duplexes of different length in *D. melanogaster* lysate and in transfected mammalian cells indicated that the differences in silencing efficiencies between 20- to 25-nt siRNA duplexes were less pronounced in mammalian cells than in *D. melanogaster* (Elbashir et al. (2002), supra). Duplexes of 24- and 25-nt siRNAs were inactive in *D. melanogaster* lysate, whereas the same duplexes were quite effective for silencing when introduced by transfection into HeLa cells. We therefore asked whether siRNA duplexes of 20- to 25-nt are able to reconstitute RISC also with approximately equal

- 31 -

efficiency. Indeed, we observed no large differences in our biochemical assay, and the position of target RNA cleavage was as predicted according to the cleavage guiding rules established in *D. melanogaster* lysate (data not shown). Our biochemical results therefore support the in vivo observations.

2.2 5' modification of the guide siRNA inhibits RISC activity

Modification of siRNAs at their termini is important for developing siRNA-based affinity purification schemes or for conjugating reporter tags for biophysical measurements. The most common method for introducing reactive side chains into nucleic acids is by chemical synthesis using aminolinker derivatives (Eckstein (1991), Oligonucleotides and analogues, 2nd Ed., Oxford UK, Oxford University Press). After complete deprotection of the oligonucleotide, the primary amine is typically reacted with the N-hydroxysuccinimidyl ester of the desired compound. We have introduced 5' and 3' aminolinkers with six and seven methylene groups as spacers, respectively. The linker-modified siRNA duplexes were tested for mediating target RNA degradation in HeLa S100 extract (Figure 2A, B). Modification of the 5'- end of the antisense guide siRNA abolished target RNA cleavage, while modification of neither the sense 5'-end nor of both 3'-ends showed any inhibitory effect. In an identical experiment using *D. melanogaster* embryo lysate, we observed a similar pattern of RISC activity although the duplex carrying the 5' aminolinker-modified antisense siRNA showed some residual activity (data not shown). Presumably, introduction of additional atoms or the change in terminal phosphate electric charge at the 5'-end of the antisense siRNA interfered with its ability to function as guide RNA. The critical function of the guide siRNAs 5' end was previously documented (Elbashir et al. (2001 c), supra).

The ability to modify siRNAs at their 3'-end suggests that siRNAs do not play a major role for priming dsRNA synthesis and do not act as primers for degenerative PCR. The fate of a siRNA in HeLa S100 extracts was

- 32 -

followed directly by incubation of an internally ^{32}pCp -radiolabeled siRNA duplexes. The radiolabeled antisense siRNA strand was also prepared with different 5' and 3' phosphate modifications (Figure 3A). All described combinations of siRNA duplexes were fully competent for RISC-dependent target RNA degradation (data not shown). As previously observed for *D. melanogaster* lysates (Nykänen et al. (2001), supra), rapid 5' phosphorylation of siRNA duplexes with free 5' hydroxyl termini was apparent. To our surprise, we noted that a small fraction of the 3' phosphorylated antisense siRNA could be ligated to the opposing 5' hydroxyl of the sense siRNA producing a lower mobility band. The inter-strand ligation was confirmed by changing the length of the unlabeled sense siRNA, which resulted in the expected mobility changes of the ligation product (data not shown). RNA ligase activity was previously observed in HeLa S100 extracts and it is mediated by two enzymatic activities (e.g. Vicente and Filipowicz, Eur. J. Biochem., 176 (1988), 431 - 439). The 3' terminal phosphate is first converted to a 2',3'-cyclic phosphate requiring ATP and 3' terminal phosphate cyclase. Thereafter, the opposing 5' hydroxyl is ligated to the cyclic phosphate end by an as yet uncharacterized RNA ligase. We chemically synthesized the predicted 5' phosphorylated, 42-nt ligation product and found that it is unable to mediate target RNA cleavage, presumably because it can not form activated RISC. The majority of the 3' phosphorylated duplexes siRNA was gradually dephosphorylated at its 3' end and emerged chemically similar to naturally generated siRNA. Together, these observations indicate that the cell has a mechanism to preserve the integrity of siRNAs. We were unable to detect a proposed siRNA-primed polymerization product (Figure 3B), suggesting that siRNAs do not function as primers for template-dependent dsRNA synthesis in our system. However, we acknowledge that a proposed RNA-dependent polymerase activity may have been inactivated during preparation of our extracts.

- 33 -

2.3 siRNAs incorporated into RISC do not compete with a pool of free siRNAs

In order to analyze RISC assembly and stability, we tested whether
5 target-unspecific siRNA duplexes were able to compete with target-specific
siRNA duplexes. When specific and non-specific siRNA duplexes were
co-incubated in HeLa S100 extracts, increasing concentrations of
unspecific siRNA duplex competed with the formation of target-specific
RISC (Figure 4, left lanes). However, when target-specific siRNAs were
10 pre-incubated in HeLa S100 extract for 15 min in the absence of
competitor siRNA duplex, the assembled siRNA in the target-specific RISC
could no longer be competed with the target-unspecific siRNA duplex
(Figure 4, right lanes). This result suggests that RISC is formed during the
first 15 minutes of incubation and that siRNAs were irreversibly associated
15 with the protein components of RISC during the 2.5 h time window of the
experiment.

2.4 Purification of human RISC

20 After having the 3' termini of siRNAs defined as the most suitable position
for chemical modification, a photo-cleavable biotin derivative was
conjugated to the 3' aminolinker-modified siRNAs. A photo-cleavable biotin
derivative was selected because of the advantage of recovering RISC under
non-denaturing conditions after capturing complexes on streptavidin-coated
25 affinity supports. 3' Conjugation of biotin to the sense, antisense or to
both of the strands did not affect target RNA cleavage when compared to
non-biotinylated siRNAs (data not shown). siRNA duplexes with biotin
residues on both 3' ends were therefore used for affinity purification
(Figure 5A). The double biotinylated siRNA duplex was incubated in HeLa
30 S100 extracts in the presence of ATP, GTP, creatine phosphate, and
creatine kinase for ATP regeneration. Thereafter, streptavidin-conjugated
agarose beads were added to capture the biotinylated siRNA

- 34 -

ribonucleoprotein complexes (siRNPs) including RISC. After extensive washing of the collected beads, the siRNPs were released by UV irradiation at 312 nm. The eluate cleaved target RNA sequence-specifically, thus indicating that RISC was recovered in its native state from the resin (Figure 5B, C, lane UV elu). The flow-through from the affinity selection showed no detectable RISC activity indicating complete binding of RISC by the beads (Figure 6). The affinity eluate was further analyzed by applying it onto a Superdex 200 gel filtration column (Figure 5B) as well as a 5%-20% glycerol gradient ultra-centrifugation (Figure 5C). Individual fractions were collected and assayed for target RNA cleavage without the addition of any further siRNA. RISC activity appeared between the molecular size markers aldolase (158 kDa) and BSA (66 kDa) after gel filtration or glycerol gradient centrifugation (Figure 5B, C). The molecular size of human RISC is therefore estimated to be between 90 and 160 kDa, significantly smaller than the complex previously analyzed in *D. melanogaster* lysates (Hammond et al. (2000), supra; Nykänen et al. (2001), supra). The small size of RISC suggests that Dicer (210 kDa) is not contained in RISC and that the formation of RISC from synthetic siRNAs may occur independently of Dicer. While these results do not rule out a role for Dicer during assembly of RISC, they emphasize the absence of Dicer in RISC.

2.5 RISC contains a single siRNA strand and can be reconstituted using single-stranded siRNAs

Two models are currently discussed concerning the siRNA strand composition of RISC. The first model suggests that both strands of the initially added siRNA duplex are physically present in RISC, but in an unwound conformation. The second model proposes that RISC carries only a single siRNA strand, implying loss of one of the siRNA strands during assembly. The latter model has been favored based on the analogy to miRNA precursor processing, where only one 21-nt strand accumulated from a dsRNA hairpin precursor. The molecular basis for the asymmetry of

- 35 -

the miRNA precursor processing reaction is not yet understood. Because siRNAs have symmetric 2-nt 3'-overhangs it is assumed that siRNA duplexes enter RISC with equal probability for both orientations, thus giving rise to distinct sense and antisense targeting RISCs.

5

To address the constitution of siRNAs in RISC, we affinity selected the assembled complexes with siRNA duplexes that were biotinylated at only one of the two constituting strands or both (Figure 6). If both strands were present together in RISC, the cleavage activity should be affinity selected on Neutravidin independently of the position of the biotin residue. In contrast, we observed target RNA cleavage from UV eluates after streptavidin selection only for siRNA duplexes with biotin conjugated to the antisense strand, but not the sense strand (Figure 6). RISC activity, assembled on siRNA duplexes with only the sense siRNA biotinylated, remained in the flow-through. These data suggest that RISC contains only a single-stranded RNA molecule.

To assess whether single-stranded siRNAs may be able to reconstitute RISC, single-stranded 5' phosphorylated siRNAs as well as the siRNA duplex were incubated at concentrations between 1 to 100 nM with cap-labeled target RNA in HeLa S100 extract (Figure 7A). At 100 nM single-stranded antisense siRNA, we detected RISC-specific target RNA cleavage, thus confirming that single-stranded siRNAs are present in RISC. At lower concentrations of single-stranded siRNAs, RISC formation remained undetectable while duplex siRNAs were effectively forming RISC even at 1 nM concentration. Therefore, a specific pathway exists which converts double-stranded siRNA into single-stranded siRNA containing RISC. Using *D. melanogaster* embryo lysate, we were unable to detect RISC activity from antisense siRNA (Figure 7B), presumably because of the high load of single-strand specific ribonucleases (Elbashir et al. (2001 b), supra). Furthermore, 5' phosphorylated 20- to 25-nt antisense siRNAs were able to mediate RISC-specific target RNA degradation in HeLa S100

30

- 36 -

extract producing the same target RNA cleavage sites as duplex siRNAs of this length (data not shown).

Finally, we tested single-stranded and duplex siRNAs for targeting of an
5 endogenous gene in HeLa cells following our standard protocol previously
established for silencing of lamin A/C. 200 nM concentrations of
single-stranded siRNAs with and without 5' phosphate and 100 nM
concentrations of duplex siRNAs were transfected into HeLa cells. Lamin
A/C levels were monitored 48 h later using immunofluorescence (Figure
10 8A) and quantitative luminescence-based Western blot analysis (Figure 8B).
Non-phosphorylated antisense siRNA caused a substantial knockdown of
lamin A/C to about 25% of its normal level while 5' phosphorylated siRNAs
reduced the lamin A/C content to less than 5%, similar to the reduction
observed with the lamin A/C 5' phosphorylated (data not shown) or
15 non-phosphorylated duplex siRNA (Figure 8). Sense siRNA and GL2
unspecific siRNA did not affect lamin A/C levels. The levels of non-targeted
vimentin protein were monitored and used for normalizing of the loading of
the lanes of the lamin A/C Western blots.

20 Gene silencing was also observed with phosphorylated as well as
non-phosphorylated antisense siRNAs ranging in size between 19 to 29 nt.
The phosphorylated antisense siRNAs were consistently better performing
than the non-phosphorylated antisense, and their silencing efficiencies
were comparable to that of the conventional duplex siRNA (Figure 11).

25

2.6 Protein composition of RISC

In order to identify the protein components of the RNA-induced silencing
complex (RISC) in HeLa S100 extract, the specific affinity selection
previously outlined was used. UV eluates were fractionated on a 5-20%
30 glycerol gradient, fractions were recovered from the gradient and analysed
for protein composition and target RNA endonucleolytic activity.

- 37 -

Two proteins of approximately 100 kDa were identified by mass spectrometry in the peak fraction of the endonucleolytic activity (Figure 12, fractions 7 and 8), corresponding to eIF2C1 and eIF2C2/GERp95 (Figure 13A and B). These proteins are 82% similar and are both members of the Argonaute family (Figure 13C). The first evidence that Argonaute proteins are part of RISC was provided by classical biochemical fractionation studies using dsRNA-transfected *D. melanogaster* S2 cells (Hammond et al., 2001, supra). The closest relative to *D. melanogaster* Argonaute2, *D. melanogaster* Argonaute1, was recently shown to be required for RNAi (Williams and Rubin, PNAS USA 99 (2002), 6889-6894).

Mass spectrometry analysis also revealed the presence of three peptides belonging exclusively to the HILI member of the Argonaute family of proteins. The sequences of those peptides are: NKQDFMDLSICTR, corresponding to positions 17-29 of the protein; TEYVAESFLNCLRR, corresponding to positions 436-449 of the protein, and; YNHDLPARIIVYR, corresponding to positions 591-603 of the protein. This finding suggests that the protein HILI may also be part of RISC.

In human, the Argonaute family is composed of 6 members, eIF2C1, eIF2C2, eIF2C3, eIF2C4, HILI and HIWI (Figure 14). The alignment of the six predicted amino-acid sequences show a high conservation, in particular between the eIF2C members, and HILI and HIWI (Figure 15). Predicted cDNA sequences encoding the Argonaute proteins are also shown (Figure 16).

The expression of the human Argonaute proteins was also investigated in HeLa cells by RT-PCR analysis using total and poly (A) selected RNA. All members of the family but HIWI were detected (Figure 17).

- 38 -

3. Discussion

The development of a human biochemical system for analysis of the mechanism of RNAi is important given the recent success of siRNA
5 duplexes for silencing genes expressed in human cultured cells and the potential for becoming a sequence-specific therapeutic agent. Biochemical systems are useful for defining the individual steps of the RNAi process and for evaluating the constitution and molecular requirements of the participating macromolecular complexes. For the analysis of RNAi, several
10 systems were developed, with the *D. melanogaster* systems being the most comprehensive as they enable to reconstitute dsRNA processing as well as the mRNA targeting. For mammalian systems, reconstitution of the mRNA targeting reaction has not yet been accomplished. Here, we describe the development and application of a biochemical system prepared
15 from the cytoplasmic fraction of human HeLa cells, which is able to reconstitute the human mRNA-targeting RNA-induced silencing complex (RISC). Formation of RISC was accomplished using either 5' phosphorylated or non-phosphorylated siRNA duplexes; as well as single-stranded antisense siRNAs; non-phosphorylated siRNA duplexes and
20 presumably also single-stranded antisense siRNAs are rapidly 5' phosphorylated in HeLa cell extracts (Figure 3).

Biochemical characterization of siRNA function

Reconstitution of RISC activity was only observed using cytoplasmic HeLa
25 extracts. HeLa nuclear extracts assayed under the same conditions did not support siRNA-specific target RNA cleavage, thus suggesting that RISC components are located predominantly in the cytoplasm (data not shown).

Modifications of the 5' and 3' termini of siRNAs were tested in order to
30 assess the importance of the siRNA termini for the targeting step. It was found that the 5' end modification of the guide siRNA was more inhibitory for target RNA cleavage than 3' end modification. Introduction of the 3'

- 39 -

biotin affinity tag into the target-complementary guide siRNA enabled us to affinity select sense-RNA-targeting RISC, whereas 3' biotinylation of the sense siRNA strand resulted in RISC activity in the flowthrough. Furthermore, the single RNA strand composition of RISC was confirmed by reconstituting the sequence-specific endonuclease complex using 5'-phosphorylated single-stranded guide siRNA. The reconstitution of RISC from single-stranded siRNA was however less effective and required 10- to 100-fold higher concentrations compared to duplex siRNA. Reconstitution of RISC from single-stranded siRNA was undetectable using *D. melanogaster* embryo lysate, which is most likely explained by the high content of 5' to 3' exonucleases in embryo lysate.

The size of RISC in HeLa lysate was determined by gel filtration as well as glycerol gradient ultracentrifugation after streptavidin affinity purification with 3' biotinylated siRNA duplexes. Sizes for RISC in *D. melanogaster* systems have been reported within a range of less than 230 to 500 kDa, however size determinations were conducted without having affinity purified RISC. Our affinity-purified RISC sediments in a narrow range between the size makers of 66 and 158 kDa. The differences to the reported sizes for RISC are not species-specific as we observed a similar size for RISC in *D. melanogaster* S2 cell cytoplasmic extracts after affinity purification (data not shown).

It has previously been proposed that siRNAs act as primers for target RNA-templated dsRNA synthesis (Lipardi et al., Cell 107 (2001), 297 - 307) although homologs for such RNA-dependent RNA polymerases known to participate in gene silencing in other systems are not identified in *D. melanogaster* or mammalian genomes. Analysis of the fate of siRNA duplexes in the HeLa cell system did not provide evidence for such a siRNA-primed activity (Figure 3), but indicates that the predominant pathway for siRNA-mediated gene silencing is sequence-specific endonucleolytic target RNA degradation.

- 40 -

Single-stranded 5' phosphorylated antisense siRNAs as triggers of mammalian gene silencing

It was previously noted that introduction of sense and antisense RNAs of several hundred nucleotides in length into *C. elegans* was able to
5 sequence-specifically silence homologous genes (Guo and Kemphues, Cell 81 (1995), 611 - 620). Later, it was suggested that the sense and antisense RNA preparation were contaminated with a small amount of dsRNA, which was responsible for the silencing effect and is a much more potent inducer of gene silencing (Fire et al. (1998), supra). It is however
10 conceivable that antisense RNA directly contributed to initiation of silencing. Indeed, most recently it was shown that antisense RNAs between 22 and 40 nt, but not sense RNAs were able to activate gene silencing in *C. elegans* (Tijsterman et al., Science 295 (2002), 694 - 697). The authors, however, favored the hypothesis of siRNA-primed dsRNA
15 synthesis.

We have shown that modification of the 3' ends of antisense siRNA did not interfere with reconstitution of RISC in the human system. Together, these observations suggest that the driving forces for gene silencing in *C.*
20 *elegans* may be predominantly dsRNA synthesis followed by Dicer cleavage, while in human and possibly also in *D. melanogaster* RISC-specific target mRNA degradation predominates.

Targeting of endogenously expressed lamin A/C by transfection of duplex
25 siRNA into HeLa cells was the first reported example of siRNA-induced gene silencing. Lamin A/C protein was drastically reduced by a lamin A/C-specific siRNA duplex within two days post transfection, while unspecific siRNA duplexes showed no effect. At the time, transfection of non-phosphorylated sense or antisense siRNA did not reveal a substantial
30 effect on lamin A/C levels, although more recently a minor reduction upon antisense siRNA transfection was noticed when similar concentrations of antisense siRNA were delivered as described in this study. However, the

- 41 -

effect was not interpreted as RISC-specific effect. Assaying 5'-phosphorylated antisense siRNAs revealed a substantial increase in lamin A/C silencing. Probably, 5' phosphorylated siRNAs are more stable or enter RISC more rapidly. Alternatively, the 5' end of transfected
5 single-stranded siRNA may be less rapidly phosphorylated in the cell in comparison to duplex siRNAs.

Finally, it should be noted that HeLa cells are generally poor in nucleases and represent one of the preferred mammalian systems for studying
10 RNA-processing or transcription reactions in vivo and in vitro. However, it can be expected that 5' phosphorylated single-stranded antisense siRNAs are suitable to knockdown gene expression in other cell types or tissues with a different content of nucleases, since chemical strategies to improve
nuclease resistance of single stranded RNA are available. The general
15 silencing ability of various cell types may also depend on the relative levels of siRNA/miRNA-free eIF2C1 and eIF2C2 proteins capable of associating with exogenously delivered siRNAs.

In summary, single-stranded 5'-phosphorylated antisense siRNAs of 19- to
20 29-nt in size broaden the use of RNA molecules for gene silencing because they can enter the mammalian RNAi pathway in vitro as well as in vivo through reconstitution of RISC. Human eIF2C1 and/or eIF2C2 seem to play a critical role in this process. Considering the feasibility of modulating the stability and uptake properties of single-stranded RNAs, 5'-phosphorylated
25 single-stranded antisense siRNAs may further expand the utility of RNAi-based gene silencing technology as tool for functional genomics as well as therapeutic applications.

Argonaute proteins are a distinct class of proteins, containing a PAZ and
30 Piwi domain (Cerutti et al., 2000, supra) and have been implicated in many processes previously linked to post-transcriptional silencing, however only limited biochemical information is available.

- 42 -

Human eIF2C2 is the ortholog of rat GERp95, which was identified as a component of the Golgi complex or the endoplasmic reticulum and copurified with intracellular membranes (Cikaluk et al., Mol.Biol.Cell 10 (1999), 3357-3722). More recently, HeLa cell eIF2C2 was shown to be
5 associated with microRNAs and components of the SMN complex, a regulator of ribonucleoprotein assembly, suggesting that eIF2C2 plays a role in miRNA precursor processing or miRNA function (Mourelatos et al., Genes & Dev.16 (2002), 720-728). A more provocative hypothesis is that miRNAs are also in a RISC-like complex, which could potentially mediate
10 target RNA degradation, if only perfectly matched miRNA target mRNAs existed. Sequence analysis using cloned human and mouse, however, did not reveal the presence of such perfectly complementary sequences in the genomes (Lagos-Quintana et al., Science 294 (2001), 853-858). Therefore, miRNPs may only function as translational regulators of partially
15 mismatched target mRNAs, probably by recruiting additional factors that prevent dissociation from mismatched target mRNAs.

Human eIF2C1 has not been linked to gene silencing previously, but it is more than 80% similar in sequence to eIF2C2 (Koesters et al., Genomics
20 61 (1999), 210-218). This similarity may indicate functional redundancy, but it is also conceivable that functional RISC may contain eIF2C1 and eIF2C2 heterodimers. The predicted molecular weight of this heterodimeric complex would be slightly larger than the observed size of 90-160 kDa, but because size fractionation is based on globular shape, we can not disregard
25 this possibility at this time.

Due to the high conservation between the members of the Argonaute family, it is possible that peptides that derive from regions 100% conserved in the 6 predicted proteins, may belong to members others than
30 eIF2C1 and eIF2C2. In this respect, three peptides were identified with masses corresponding to HILI, meaning that this protein might be also a component of RISC.

- 43 -

To precisely assess the protein composition of RISC, reconstitution of the siRNA-mediated target RNA cleavage must be achieved by using recombinant proteins which may be obtained by cloning and expression in suitable bacterial or eukaryotic systems.

5

We expect that the biochemical characterization of the siRNA-mediated target RNA degradation process will have immediate applications, such as the development of cell lines or transgenic animals overexpressing RISC components. The efficiency in targeting endogenous genes in those lines or
10 organisms will be enhanced. Furthermore, a reconstituted in vitro system for RNAi will allow the design of more potent and specific siRNA to achieve gene silencing.

43a

SEQUENCE LISTING

<110> Max-Planck-Gesellschaft zur Forderung der Wissenschaften E.V.

<120> RNA-interference by single-stranded RNA molecules

<130> 4659-524CA

<140> 2,489,174

<141> 2003-07-10

<150> EP 02015532.1

<151> 2002-07-10

<150> EP 02018906.4

<151> 2002-08-23

<160> 108

<170> PatentIn Ver. 2.1

<210> 1

<211> 15

<212> RNA

<213> HeLa

<220>

<223> antisense siRNA (5' - 3')

<400> 1

ucgaaguauu ccgcg

15

<210> 2

<211> 21

<212> RNA

<213> HeLa

<220>

<223> antisense siRNA (5' - 3')

<400> 2

ucgaaguauu ccgcguacgu g

21

<210> 3

<211> 25

<212> RNA

<213> HeLa

43b

<220>
<223> antisense siRNA (5' - 3')

<400> 3
ucgaaguauu ccgcguacgu gaugu

25

<210> 4
<211> 27
<212> RNA
<213> HeLa

<220>
<223> antisense siRNA (5' - 3')

<400> 4
ucgaaguauu ccgcguacgu gauguuc

27

<210> 5
<211> 29
<212> RNA
<213> HeLa

<220>
<223> antisense siRNA (5' - 3')

<400> 5
ucgaaguauu ccgcguacgu gauguucac

29

<210> 6
<211> 15
<212> RNA
<213> HeLa

<220>
<223> antisense siRNA (5' - 3')

<400> 6
ucgaaguauu ccgcg

15

<210> 7
<211> 21
<212> RNA
<213> HeLa

43c

<220>
<223> antisense siRNA (5' - 3')

<400> 7
ucgaaguauu ccgcguacgu g

21

<210> 8
<211> 25
<212> RNA
<213> HeLa

<220>
<223> antisense siRNA (5' - 3')

<400> 8
ucgaaguauu ccgcguacgu gaugu

25

<210> 9
<211> 27
<212> RNA
<213> HeLa

<220>
<223> antisense siRNA (5' - 3')

<400> 9
ucgaaguauu ccgcguacgu gauguuc

27

<210> 10
<211> 29
<212> RNA
<213> HeLa

<220>
<223> antisense siRNA (5' - 3')

<400> 10
ucgaaguauu ccgcguacgu gauguucac

29

<210> 11
<211> 21
<212> RNA
<213> HeLa

43d

<220>

<223> Description of Combined DNA/RNA Molecule: t =
2'-deoxythymidine

<220>

<223> antisense siRNA (5' - 3')

<400> 11

ucgaaguauu ccgcguacgu g

21

<210> 12

<211> 21

<212> DNA

<213> HeLa

<220>

<223> antisense siRNA (5' - 3')

<400> 12

ucgaaguauu ccgcguacgt g

21

<210> 13

<211> 21

<212> RNA

<213> HeLa

<220>

<223> Description of Combined DNA/RNA Molecule: t =
2'-deoxythymidine

<220>

<223> antisense siRNA (5' - 3')

<400> 13

ucgaaguauu ccgcguacgu u

21

<210> 14

<211> 21

<212> DNA

<213> HeLa

<220>

<223> antisense siRNA (5' - 3')

43e

<400> 14
ucgaaguauu ccgcguacgt t

21

<210> 15
<211> 21
<212> RNA
<213> HeLa

<220>
<223> Description of Combined DNA/RNA Molecule: t =
2'-deoxythymidine

<220>
<223> antisense siRNA (5' - 3')

<400> 15
ucgaaguauu ccgcguacgu g

21

<210> 16
<211> 21
<212> DNA
<213> HeLa

<220>
<223> antisense siRNA (5' - 3')

<400> 16
ucgaaguauu ccgcguacgt g

21

<210> 17
<211> 21
<212> RNA
<213> HeLa

<220>
<223> Description of Combined DNA/RNA Molecule: t =
2'-deoxythymidine

<220>
<223> antisense siRNA (5' - 3')

<400> 17
ucgaaguauu ccgcguacgu u

21

43f

<210> 18
<211> 21
<212> DNA
<213> HeLa

<220>
<223> antisense siRNA (5' - 3')

<400> 18
ucgaaguauu ccgcguacgt t

21

<210> 19
<211> 21
<212> RNA
<213> HeLa

<220>
<223> sense siRNA (5' - 3')

<400> 19
cguacgcgga auacuucgaa a

21

<210> 20
<211> 21
<212> RNA
<213> HeLa

<220>
<223> antisense siRNA (5' - 3')

<400> 20
ucgaaguauu ccgcguacgu g

21

<210> 21
<211> 21
<212> RNA
<213> HeLa

<220>
<223> Description of Combined DNA/RNA Molecule: t =
2'-deoxythymidine

<220>
<223> sense siRNA (5' - 3')

43g

<400> 21
cguacgcgga auacuucgaa a 21

<210> 22
<211> 21
<212> DNA
<213> HeLa

<220>
<223> antisense siRNA (5' - 3')

<400> 22
ucgaaguauu ccgcguacgt g 21

<210> 23
<211> 21
<212> RNA
<213> HeLa

<220>
<223> sense siRNA (5' - 3')

<400> 23
cguacgcgga auacuucgaa a 21

<210> 24
<211> 21
<212> RNA
<213> HeLa

<220>
<223> antisense siRNA (5' - 3')

<400> 24
ucgaaguauu ccgcguacgu u 21

<210> 25
<211> 21
<212> RNA
<213> HeLa

<220>
<223> Description of Combined DNA/RNA Molecule: t =
2'-deoxythymidine

43h

<220>

<223> sense siRNA (5' - 3')

<400> 25

cguacgcgga auacuucgaa a

21

<210> 26

<211> 21

<212> DNA

<213> HeLa

<220>

<223> Description of Combined DNA/RNA Molecule: t =
2'-deoxythymidine

<220>

<223> antisense siRNA (5' - 3')

<400> 26

ucgaaguauu ccgcguacgt t

21

<210> 27

<211> 21

<212> DNA

<213> HeLa

<220>

<223> antisense siRNA (5' - 3')

<400> 27

tcgaaguauu ccgcguacgu u

21

<210> 28

<211> 21

<212> RNA

<213> HeLa

<220>

<223> Description of Combined DNA/RNA Molecule: t =
2'-deoxythymidine

<220>

<223> sense siRNA (5' - 3')

43i

<400> 28
cguacgcgga auacuucgau u

21

<210> 29
<211> 21
<212> DNA
<213> HeLa

<220>
<223> Description of Combined DNA/RNA Molecule: t =
2'-deoxythymidine

<220>
<223> antisense siRNA (5' - 3')

<400> 29
tcgaaguauu ccgcguacgu u

21

<210> 30
<211> 21
<212> DNA
<213> HeLa

<220>
<223> Description of Combined DNA/RNA Molecule: t =
2'-deoxythymidine

<220>
<223> antisense siRNA (5' - 3')

<400> 30
tcgaaguauu ccgcguacgt t

21

<210> 31
<211> 21
<212> DNA
<213> HeLa

<220>
<223> Description of Combined DNA/RNA Molecule: t =
2'-deoxythymidine

<220>
<223> sense siRNA (5' - 3')

43j

<400> 31
cguacgcgga auacuucgat t

21

<210> 32
<211> 21
<212> DNA
<213> HeLa

<220>
<223> Description of Combined DNA/RNA Molecule: t =
2'-deoxythymidine

<220>
<223> antisense siRNA (5' - 3')

<400> 32
tcgaaguauu ccgcguacgt t

21

<210> 33
<211> 21
<212> DNA
<213> HeLa

<220>
<223> antisense siRNA (5' - 3')

<400> 33
tcgaaguauu ccgcguacgt t

21

<210> 34
<211> 13
<212> PRT
<213> HeLa

<220>
<223> peptide fragment of HILI, which belongs to the
Argonaute family of proteins; corresponding to
position 17 - 29

<400> 34
Asn Lys Gln Asp Phe Met Asp Leu Ser Ile Cys Thr Arg
1 5 10

43k

<210> 35
 <211> 14
 <212> PRT
 <213> HeLa

<220>
 <223> peptide fragment of HILI, which belongs to the
 Argonaute family of proteins; corresponding to
 position 436-449

<220>
 <223>

<400> 35
 Thr Glu Tyr Val Ala Glu Ser Phe Leu Asn Cys Leu Arg Arg
 1 5 10

<210> 36
 <211> 13
 <212> PRT
 <213> HeLa

<220>
 <223> peptide fragment of HILI, which belongs to the
 Argonaute family of proteins; corresponding to
 position 591-603

<400> 36
 Tyr Asn His Asp Leu Pro Ala Arg Ile Ile Val Tyr Arg
 1 5 10

<210> 37
 <211> 35
 <212> RNA
 <213> HeLa

<220>
 <223> target RNA

<400> 37
 aacaucacgu acgcggaaua cuucgaaaug uccgu

35

<210> 38
 <211> 21
 <212> RNA
 <213> HeLa

<220>
<223> strand of siRNA duplex

<400> 38
cguacgcgga auacuucgau u

21

<210> 39
<211> 21
<212> RNA
<213> HeLa

<220>
<223> strand of siRNA duplex

<400> 39
uugcaugcgc cuaugaagc u

21

<210> 40
<211> 21
<212> RNA
<213> HeLa

<220>
<223> strand of siRNA duplex

<400> 40
cguacgcgga auacuucgaa a

21

<210> 41
<211> 20
<212> RNA
<213> HeLa

<220>
<223> strand of siRNA duplex

<400> 41
ugcaugcgcc uuaugaagcu

20

<210> 42
<211> 12
<212> PRT
<213> HeLa

43m

<220>

<223> peptide fragments of eIF2C2 obtained by mass
spektrum

<400> 42

Val Leu Gln Pro Pro Ser Ile Leu Tyr Gly Gly Arg
1 5 10

<210> 43

<211> 12

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of eIF2C2 obtained by mass
spektrum

<400> 43

Gln Glu Ile Ile Gln Asp Leu Ala Ala Met Val Arg
1 5 10

<210> 44

<211> 12

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of eIF2C2 obtained by mass
spektrum

<400> 44

His Leu Pro Ser Met Arg Tyr Thr Pro Val Gly Arg
1 5 10

<210> 45

<211> 12

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of eIF2C2 obtained by mass
spektrum

<400> 45

Lys Leu Thr Asp Asn Gln Thr Ser Thr Met Ile Arg
1 5 10

43n

<210> 46
<211> 13
<212> PRT
<213> HeLa

<220>
<223> peptide fragments of eIF2C2 obtained by mass
spektrum

<400> 46
Tyr Ala Gln Gly Ala Asp Ser Val Glu Pro Met Phe Arg
1 5 10

<210> 47
<211> 14
<212> PRT
<213> HeLa

<220>
<223> peptide fragments of eIF2C2 obtained by mass
spektrum

<400> 47
Asp Lys Val Glu Leu Glu Val Thr Leu Pro Gly Glu Gly Lys
1 5 10

<210> 48
<211> 14
<212> PRT
<213> HeLa

<220>
<223> peptide fragments of eIF2C2 obtained by mass
spektrum

<400> 48
Asp Ala Gly Met Pro Ile Gln Gly Gln Pro Cys Phe Cys Lys
1 5 10

<210> 49
<211> 14
<212> PRT
<213> HeLa

<220>
<223> peptide fragments of eIF2C2 obtained by mass
spektrum

430

<400> 49

Thr	Gln	Ile	Phe	Gly	Asp	Arg	Lys	Pro	Val	Phe	Asp	Gly	Arg
1				5					10				

<210> 50

<211> 15

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of eIF2C2 obtained by mass
spektrum

<400> 50

Ala	Thr	Ala	Arg	Ser	Ala	Pro	Asp	Arg	Gln	Glu	Glu	Ile	Ser	Lys
1				5					10					15

<210> 51

<211> 14

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of eIF2C2 obtained by mass
spektrum

<400> 51

Asp	Tyr	Gln	Pro	Gly	Ile	Thr	Phe	Ile	Val	Val	Gln	Lys	Arg
1				5					10				

<210> 52

<211> 14

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of eIF2C2 obtained by mass
spektrum

<400> 52

Ser	Ala	Pro	Asp	Arg	Gln	Glu	Glu	Ile	Ser	Lys	Leu	Met	Arg
1				5					10				

<210> 53

<211> 14

<212> PRT

<213> HeLa

43p

<220>

<223> peptide fragments of elF2C2 obtained by mass
spektrum

<400> 53

Tyr Pro His Leu Pro Cys Leu Gln Val Gly Gln Glu Gln Lys
1 5 10

<210> 54

<211> 17

<212> PRT

<213> HeLa

<220>

<223>

<400> 54

Ser Phe Phe Thr Ala Ser Glu Gly Cys Ser Asn Pro Leu Gly Gly Gly
1 5 10 15

Arg

<210> 55

<211> 23

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of elF2C2 obtained by mass
spektrum

<400> 55

Tyr His Leu Val Asp Lys Glu His Asp Ser Ala Glu Gly Ser His Thr
1 5 10 15

Ser Gly Gln Ser Asn Gly Arg
20

<210> 56

<211> 12

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of elF2C1 obtained by mass
spektrum

43q

<400> 56

Val Leu Pro Ala Pro Ile Leu Gln Tyr Gly Gly Arg
1 5 10

<210> 57

<211> 12

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of eIF2C1 obtained by mass
spektrum

<400> 57

Ser Val Ser Ile Pro Ala Pro Ala Tyr Tyr Ala Arg
1 5 10

<210> 58

<211> 12

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of eIF2C1 obtained by mass
spektrum

<400> 58

Thr Ser Pro Gln Thr Leu Ser Asn Leu Cys Leu Lys
1 5 10

<210> 59

<211> 13

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of eIF2C1 obtained by mass
spektrum

<400> 59

Tyr Ala Gln Gly Ala Asp Ser Val Glu Pro Met Phe Arg
1 5 10

<210> 60

<211> 14

<212> PRT

<213> HeLa

43r

<220>

<223> peptide fragments of eIF2C1 obtained by mass
spektrum

<400> 60

Asn Ile Tyr Thr Val Thr Ala Leu Pro Ile Gly Asn Glu Arg
1 5 10

<210> 61

<211> 14

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of eIF2C1 obtained by mass
spektrum

<400> 61

Val Asp Phe Glu Val Thr Ile Pro Gly Glu Gly Lys Asp Arg
1 5 10

<210> 62

<211> 14

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of eIF2C1 obtained by mass
spektrum

<400> 62

Asp Ala Gly Met Pro Ile Gln Gly Gln Pro Cys Phe Cys Lys
1 5 10

<210> 63

<211> 14

<212> PRT

<213> HeLa

<220>

<223> peptide fragments of eIF2C1 obtained by mass
spektrum

<400> 63

Asn Ile Asp Glu Gln Pro Lys Pro Leu Thr Asp Ser Gln Arg
1 5 10

43s

<210> 64
<211> 14
<212> PRT
<213> HeLa

<220>
<223> peptide fragments of eIF2C1 obtained by mass
spektrum

<400> 64
Ser Ala Pro Asp Arg Gln Glu Glu Ile Ser Arg Leu Met Lys
1 5 10

<210> 65
<211> 14
<212> PRT
<213> HeLa

<220>
<223> peptide fragments of eIF2C1 obtained by mass
spektrum

<400> 65
Asp Tyr Gln Pro Gly Ile Thr Tyr Ile Val Val Gln Lys Arg
1 5 10

<210> 66
<211> 14
<212> PRT
<213> HeLa

<220>
<223>

<400> 66
Tyr Pro His Leu Pro Cys Leu Gln Val Gly Gln Glu Gln Lys
1 5 10

<210> 67
<211> 17
<212> PRT
<213> HeLa

<220>
<223> peptide fragments of eIF2C1 obtained by mass
spektrum

43t

<400> 67

Ser Phe Phe Ser Pro Pro Glu Gly Tyr Tyr His Pro Leu Gly Gly Gly
 1 5 10 15

Arg

<210> 68

<211> 857

<212> PRT

<213> human

<220>

<223> eIF2C1, predicted protein sequence

<400> 68

Met Glu Ala Gly Pro Ser Gly Ala Ala Ala Gly Ala Tyr Leu Pro Pro
 1 5 10 15

Leu Gln Gln Val Phe Gln Ala Pro Arg Arg Pro Gly Ile Gly Thr Val
 20 25 30

Gly Lys Pro Ile Lys Leu Leu Ala Asn Tyr Phe Glu Val Asp Ile Pro
 35 40 45

Lys Ile Asp Val Tyr His Tyr Glu Val Asp Ile Lys Pro Asp Lys Cys
 50 55 60

Pro Arg Arg Val Asn Arg Glu Val Val Glu Tyr Met Val Gln His Phe
 65 70 75 80

Lys Pro Gln Ile Phe Gly Asp Arg Lys Pro Val Tyr Asp Gly Lys Lys
 85 90 95

Asn Ile Tyr Thr Val Thr Ala Leu Pro Ile Gly Asn Glu Arg Val Asp
 100 105 110

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

Ser Ile Lys Trp Leu Ala Ile Val Ser Trp Arg Met Leu His Glu Ala
 130 135 140

Leu Val Ser Gly Gln Ile Pro Val Pro Leu Glu Ser Val Gln Ala Leu
 145 150 155 160

Asp Val Ala Met Arg His Leu Ala Ser Met Arg Tyr Thr Pro Val Gly
 165 170 175

Arg Ser Phe Phe Ser Pro Pro Glu Gly Tyr Tyr His Pro Leu Gly Gly
 180 185 190

Gly Arg Glu Val Trp Phe Gly Phe His Gln Ser Val Arg Pro Ala Met
 195 200 205

Trp	Lys	Met	Met	Leu	Asn	Ile	Asp	Val	Ser	Ala	Thr	Ala	Phe	Tyr	Lys
	210					215					220				
Ala	Gln	Pro	Val	Ile	Glu	Phe	Met	Cys	Glu	Val	Leu	Asp	Ile	Arg	Asn
225					230					235					240
Ile	Asp	Glu	Gln	Pro	Lys	Pro	Leu	Thr	Asp	Ser	Gln	Arg	Val	Arg	Phe
				245					250					255	
Thr	Lys	Glu	Ile	Lys	Gly	Leu	Lys	Val	Glu	Val	Thr	His	Cys	Gly	Gln
			260					265					270		
Met	Lys	Arg	Lys	Tyr	Arg	Val	Cys	Asn	Val	Thr	Arg	Arg	Pro	Ala	Ser
		275					280					285			
His	Gln	Thr	Phe	Pro	Leu	Gln	Leu	Glu	Ser	Gly	Gln	Thr	Val	Glu	Cys
	290					295					300				
Thr	Val	Ala	Gln	Tyr	Phe	Lys	Gln	Lys	Tyr	Asn	Leu	Gln	Leu	Lys	Tyr
305					310					315					320
Pro	His	Leu	Pro	Cys	Leu	Gln	Val	Gly	Gln	Glu	Gln	Lys	His	Thr	Tyr
				325					330					335	
Leu	Pro	Leu	Glu	Val	Cys	Asn	Ile	Val	Ala	Gly	Gln	Arg	Cys	Ile	Lys
			340					345					350		
Lys	Leu	Thr	Asp	Asn	Gln	Thr	Ser	Thr	Met	Ile	Lys	Ala	Thr	Ala	Arg
		355					360					365			
Ser	Ala	Pro	Asp	Arg	Gln	Glu	Glu	Ile	Ser	Arg	Leu	Met	Lys	Asn	Ala
	370					375					380				
Ser	Tyr	Asn	Leu	Asp	Pro	Tyr	Ile	Gln	Glu	Phe	Gly	Ile	Lys	Val	Lys
385					390					395					400
Asp	Asp	Met	Thr	Glu	Val	Thr	Gly	Arg	Val	Leu	Pro	Ala	Pro	Ile	Leu
				405					410					415	
Gln	Tyr	Gly	Gly	Arg	Asn	Arg	Ala	Ile	Ala	Thr	Pro	Asn	Gln	Gly	Val
			420					425					430		
Trp	Asp	Met	Arg	Gly	Lys	Gln	Phe	Tyr	Asn	Gly	Ile	Glu	Ile	Lys	Val
		435					440					445			
Trp	Ala	Ile	Ala	Cys	Phe	Ala	Pro	Gln	Lys	Gln	Cys	Arg	Glu	Glu	Val
	450					455					460				
Leu	Lys	Asn	Phe	Thr	Asp	Gln	Leu	Arg	Lys	Ile	Ser	Lys	Asp	Ala	Gly
465					470					475					480
Met	Pro	Ile	Gln	Gly	Gln	Pro	Cys	Phe	Cys	Lys	Tyr	Ala	Gln	Gly	Ala
				485					490					495	
Asp	Ser	Val	Glu	Pro	Met	Phe	Arg	His	Leu	Lys	Asn	Thr	Tyr	Ser	Gly
			500					505					510		

43v

Leu	Gln	Leu	Ile	Ile	Val	Ile	Leu	Pro	Gly	Lys	Thr	Pro	Val	Tyr	Ala	
		515					520					525				
Glu	Val	Lys	Arg	Val	Gly	Asp	Thr	Leu	Leu	Gly	Met	Ala	Thr	Gln	Cys	
	530					535					540					
Val	Gln	Val	Lys	Asn	Val	Val	Lys	Thr	Ser	Pro	Gln	Thr	Leu	Ser	Asn	
545					550					555					560	
Leu	Cys	Leu	Lys	Ile	Asn	Val	Lys	Leu	Gly	Gly	Ile	Asn	Asn	Ile	Leu	
				565					570					575		
Val	Pro	His	Gln	Arg	Ser	Ala	Val	Phe	Gln	Gln	Pro	Val	Ile	Phe	Leu	
			580					585					590			
Gly	Ala	Asp	Val	Thr	His	Pro	Pro	Ala	Gly	Asp	Gly	Lys	Lys	Pro	Ser	
		595					600					605				
Ile	Thr	Ala	Val	Val	Gly	Ser	Met	Asp	Ala	His	Pro	Ser	Arg	Tyr	Cys	
	610					615					620					
Ala	Thr	Val	Arg	Val	Gln	Arg	Pro	Arg	Gln	Glu	Ile	Ile	Glu	Asp	Leu	
625					630					635					640	
Ser	Tyr	Met	Val	Arg	Glu	Leu	Leu	Ile	Gln	Phe	Tyr	Lys	Ser	Thr	Arg	
				645					650					655		
Phe	Lys	Pro	Thr	Arg	Ile	Ile	Phe	Tyr	Arg	Asp	Gly	Val	Pro	Glu	Gly	
			660					665					670			
Gln	Leu	Pro	Gln	Ile	Leu	His	Tyr	Glu	Leu	Leu	Ala	Ile	Arg	Asp	Ala	
		675					680					685				
Cys	Ile	Lys	Leu	Glu	Lys	Asp	Tyr	Gln	Pro	Gly	Ile	Thr	Tyr	Ile	Val	
	690					695					700					
Val	Gln	Lys	Arg	His	His	Thr	Arg	Leu	Phe	Cys	Ala	Asp	Lys	Asn	Glu	
705					710					715					720	
Arg	Ile	Gly	Lys	Ser	Gly	Asn	Ile	Pro	Ala	Gly	Thr	Thr	Val	Asp	Thr	
				725					730					735		
Asn	Ile	Thr	His	Pro	Phe	Glu	Phe	Asp	Phe	Tyr	Leu	Cys	Ser	His	Ala	
			740					745					750			
Gly	Ile	Gln	Gly	Thr	Ser	Arg	Pro	Ser	His	Tyr	Tyr	Val	Leu	Trp	Asp	
		755					760					765				
Asp	Asn	Arg	Phe	Thr	Ala	Asp	Glu	Leu	Gln	Ile	Leu	Thr	Tyr	Gln	Leu	
	770					775					780					
Cys	His	Thr	Tyr	Val	Arg	Cys	Thr	Arg	Ser	Val	Ser	Ile	Pro	Ala	Pro	
785					790					795					800	
Ala	Tyr	Tyr	Ala	Arg	Leu	Val	Ala	Phe	Arg	Ala	Arg	Tyr	His	Leu	Val	
				805					810					815		

43w

Asp Lys Glu His Asp Ser Gly Glu Gly Ser His Ile Ser Gly Gln Ser
 820 825 830

Asn Gly Arg Asp Pro Gln Ala Leu Ala Lys Ala Val Gln Val His Gln
 835 840 845

Asp Thr Leu Arg Thr Met Tyr Phe Ala
 850 855

<210> 69
 <211> 860
 <212> PRT
 <213> human

<220>
 <223> eIF2C2, predicted protein sequence

<400> 69

Met Gly Val Leu Ser Ala Ile Pro Ala Leu Ala Pro Pro Ala Pro Pro
 1 5 10 15

Pro Pro Ile Gln Gly Tyr Ala Phe Lys Pro Pro Pro Arg Pro Asp Phe
 20 25 30

Gly Thr Ser Gly Arg Thr Ile Lys Leu Gln Ala Asn Phe Phe Glu Met
 35 40 45

Asp Ile Pro Lys Ile Asp Ile Tyr His Tyr Glu Leu Asp Ile Lys Pro
 50 55 60

Glu Lys Cys Pro Arg Arg Val Asn Arg Glu Ile Val Glu His Met Val
 65 70 75 80

Gln His Phe Lys Thr Gln Ile Phe Gly Asp Arg Lys Pro Val Phe Asp
 85 90 95

Gly Arg Lys Asn Leu Tyr Thr Ala Met Pro Leu Pro Ile Gly Arg Asp
 100 105 110

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

Phe Lys Val Ser Ile Lys Trp Val Ser Cys Val Ser Leu Gln Ala Leu
 130 135 140

His Asp Ala Leu Ser Gly Arg Leu Pro Ser Val Pro Phe Glu Thr Ile
 145 150 155 160

Gln Ala Leu Asp Val Val Met Arg His Leu Pro Ser Met Arg Tyr Thr
 165 170 175

Pro Val Gly Arg Ser Phe Phe Thr Ala Ser Glu Gly Cys Ser Asn Pro
 180 185 190

43x

Leu	Gly	Gly	Gly	Arg	Glu	Val	Trp	Phe	Gly	Phe	His	Gln	Ser	Val	Arg		
		195					200					205					
Pro	Ser	Leu	Trp	Lys	Met	Met	Leu	Asn	Ile	Asp	Val	Ser	Ala	Thr	Ala		
	210					215					220						
Phe	Tyr	Lys	Ala	Gln	Pro	Val	Ile	Glu	Phe	Val	Cys	Glu	Val	Leu	Asp		
225					230					235					240		
Phe	Lys	Ser	Ile	Glu	Glu	Gln	Gln	Lys	Pro	Leu	Thr	Asp	Ser	Gln	Arg		
				245					250					255			
Val	Lys	Phe	Thr	Lys	Glu	Ile	Lys	Gly	Leu	Lys	Val	Glu	Ile	Thr	His		
			260					265					270				
Cys	Gly	Gln	Met	Lys	Arg	Lys	Tyr	Arg	Val	Cys	Asn	Val	Thr	Arg	Arg		
		275					280					285					
Pro	Ala	Ser	His	Gln	Thr	Phe	Pro	Leu	Gln	Gln	Glu	Ser	Gly	Gln	Thr		
	290					295					300						
Val	Glu	Cys	Thr	Val	Ala	Gln	Tyr	Phe	Lys	Asp	Arg	His	Lys	Leu	Val		
305					310					315					320		
Leu	Arg	Tyr	Pro	His	Leu	Pro	Cys	Leu	Gln	Val	Gly	Gln	Glu	Gln	Lys		
				325					330						335		
His	Thr	Tyr	Leu	Pro	Leu	Glu	Val	Cys	Asn	Ile	Val	Ala	Gly	Gln	Arg		
			340					345					350				
Cys	Ile	Lys	Lys	Leu	Thr	Asp	Asn	Gln	Thr	Ser	Thr	Met	Ile	Arg	Ala		
		355					360					365					
Thr	Ala	Arg	Ser	Ala	Pro	Asp	Arg	Gln	Glu	Glu	Ile	Ser	Lys	Leu	Met		
	370					375					380						
Arg	Ser	Ala	Ser	Phe	Asn	Thr	Asp	Pro	Tyr	Val	Arg	Glu	Phe	Gly	Ile		
385					390					395					400		
Met	Val	Lys	Asp	Glu	Met	Thr	Asp	Val	Thr	Gly	Arg	Val	Leu	Gln	Pro		
				405					410					415			
Pro	Ser	Ile	Leu	Tyr	Gly	Gly	Arg	Asn	Lys	Ala	Ile	Ala	Thr	Pro	Val		
			420					425					430				
Gln	Gly	Val	Trp	Asp	Met	Arg	Asn	Lys	Gln	Phe	His	Thr	Gly	Ile	Glu		
		435					440					445					
Ile	Lys	Val	Trp	Ala	Ile	Ala	Cys	Phe	Ala	Pro	Gln	Arg	Gln	Cys	Thr		
	450					455					460						
Glu	Val	His	Leu	Lys	Ser	Phe	Thr	Glu	Gln	Leu	Arg	Lys	Ile	Ser	Arg		
465					470					475					480		
Asp	Ala	Gly	Met	Pro	Ile	Gln	Gly	Gln	Pro	Cys	Phe	Cys	Lys	Tyr	Ala		
				485					490						495		

43y

Gln	Gly	Ala	Asp	Ser	Val	Glu	Pro	Met	Phe	Arg	His	Leu	Lys	Asn	Thr		
			500					505					510				
Tyr	Ala	Gly	Leu	Gln	Leu	Val	Val	Val	Ile	Leu	Pro	Gly	Lys	Thr	Pro		
		515					520					525					
Val	Tyr	Ala	Glu	Val	Lys	Arg	Val	Gly	Asp	Thr	Val	Leu	Gly	Met	Ala		
	530					535					540						
Thr	Gln	Cys	Val	Gln	Met	Lys	Asn	Val	Gln	Arg	Thr	Thr	Pro	Gln	Thr		
545					550					555					560		
Leu	Ser	Asn	Leu	Cys	Leu	Lys	Ile	Asn	Val	Lys	Leu	Gly	Gly	Val	Asn		
				565					570					575			
Asn	Ile	Leu	Leu	Pro	Gln	Gly	Arg	Pro	Pro	Val	Phe	Gln	Gln	Pro	Val		
			580					585					590				
Ile	Phe	Leu	Gly	Ala	Asp	Val	Thr	His	Pro	Pro	Ala	Gly	Asp	Gly	Lys		
		595					600					605					
Lys	Pro	Ser	Ile	Ala	Ala	Val	Val	Gly	Ser	Met	Asp	Ala	His	Pro	Asn		
	610					615					620						
Arg	Tyr	Cys	Ala	Thr	Val	Arg	Val	Gln	Gln	His	Arg	Gln	Glu	Ile	Ile		
625					630					635					640		
Gln	Asp	Leu	Ala	Ala	Met	Val	Arg	Glu	Leu	Leu	Ile	Gln	Phe	Tyr	Lys		
				645					650					655			
Ser	Thr	Arg	Phe	Lys	Pro	Thr	Arg	Ile	Ile	Phe	Tyr	Arg	Asp	Gly	Val		
			660					665					670				
Ser	Glu	Gly	Gln	Phe	Gln	Gln	Val	Leu	His	His	Glu	Leu	Leu	Ala	Ile		
		675					680					685					
Arg	Glu	Ala	Cys	Ile	Lys	Leu	Glu	Lys	Asp	Tyr	Gln	Pro	Gly	Ile	Thr		
	690					695					700						
Phe	Ile	Val	Val	Gln	Lys	Arg	His	His	Thr	Arg	Leu	Phe	Cys	Thr	Asp		
705					710					715					720		
Lys	Asn	Glu	Arg	Val	Gly	Lys	Ser	Gly	Asn	Ile	Pro	Ala	Gly	Thr	Thr		
				725					730					735			
Val	Asp	Thr	Lys	Ile	Thr	His	Pro	Thr	Glu	Phe	Asp	Phe	Tyr	Leu	Cys		
			740					745					750				
Ser	His	Ala	Gly	Ile	Gln	Gly	Thr	Ser	Arg	Pro	Ser	His	Tyr	His	Val		
		755					760					765					
Leu	Trp	Asp	Asp	Asn	Arg	Phe	Ser	Ser	Asp	Glu	Leu	Gln	Ile	Leu	Thr		
	770					775					780						
Tyr	Gln	Leu	Cys	His	Thr	Tyr	Val	Arg	Cys	Thr	Arg	Ser	Val	Ser	Ile		
785					790					795					800		

43z

Pro Ala Pro Ala Tyr Tyr Ala His Leu Val Ala Phe Arg Ala Arg Tyr
 805 810 815

His Leu Val Asp Lys Glu His Asp Ser Ala Glu Gly Ser His Thr Ser
 820 825 830

Gly Gln Ser Asn Gly Arg Asp His Gln Ala Leu Ala Lys Ala Val Gln
 835 840 845

Val His Gln Asp Thr Leu Arg Thr Met Tyr Phe Ala
 850 855 860

<210> 70
 <211> 924
 <212> PRT
 <213> human

<220>
 <223> eIF2C3, predicted protein sequence

<400> 70
 Ser Arg Ser Arg Val Pro Val Pro Gly Pro Gly Ala Ala Ala Ala Pro
 1 5 10 15

Cys Pro Ala Pro Ala Ser Pro Arg Arg His Pro Ser Ala Asn Ile Pro
 20 25 30

Glu Ile Lys Arg Tyr Ala Ala Ala Ala Ala Ala Ala Gly Pro Gly
 35 40 45

Ala Gly Gly Ala Gly Asp Arg Gly Glu Ala Ala Pro Ala Ala Ala Met
 50 55 60

Glu Ala Leu Gly Pro Gly Pro Pro Ala Ser Leu Phe Gln Pro Pro Arg
 65 70 75 80

Arg Pro Gly Leu Gly Thr Val Gly Lys Pro Ile Arg Leu Leu Ala Asn
 85 90 95

His Phe Gln Val Gln Ile Pro Lys Ile Asp Val Tyr His Tyr Asp Val
 100 105 110

Asp Ile Lys Pro Glu Lys Arg Pro Arg Arg Val Asn Arg Glu Val Val
 115 120 125

Asp Thr Met Val Arg His Phe Lys Met Gln Ile Phe Gly Asp Arg Gln
 130 135 140

Pro Gly Tyr Asp Gly Lys Arg Asn Met Tyr Thr Ala His Pro Leu Pro
 145 150 155 160

Ile Gly Arg Asp Arg Val Asp Met Glu Val Thr Leu Pro Gly Glu Gly
 165 170 175

43aa

Lys	Asp	Gln	Thr	Phe	Lys	Val	Ser	Val	Gln	Trp	Val	Ser	Val	Val	Ser	
			180					185					190			
Leu	Gln	Leu	Leu	Leu	Glu	Ala	Leu	Ala	Gly	His	Leu	Asn	Glu	Val	Pro	
		195					200					205				
Asp	Asp	Ser	Val	Gln	Ala	Leu	Asp	Val	Ile	Thr	Arg	His	Leu	Pro	Ser	
	210					215					220					
Met	Arg	Tyr	Thr	Pro	Val	Gly	Arg	Ser	Phe	Phe	Ser	Pro	Pro	Glu	Gly	
225					230					235					240	
Tyr	Tyr	His	Pro	Leu	Gly	Gly	Gly	Arg	Glu	Val	Trp	Phe	Gly	Phe	His	
				245					250					255		
Gln	Ser	Val	Arg	Pro	Ala	Met	Trp	Asn	Met	Met	Leu	Asn	Ile	Asp	Val	
			260					265					270			
Ser	Ala	Thr	Ala	Phe	Tyr	Arg	Ala	Gln	Pro	Ile	Ile	Glu	Phe	Met	Cys	
		275					280					285				
Glu	Val	Leu	Asp	Ile	Gln	Asn	Ile	Asn	Glu	Gln	Thr	Lys	Pro	Leu	Thr	
	290					295					300					
Asp	Ser	Gln	Arg	Val	Lys	Phe	Thr	Lys	Glu	Ile	Arg	Gly	Leu	Lys	Val	
305					310					315					320	
Glu	Val	Thr	His	Cys	Gly	Gln	Met	Lys	Arg	Lys	Tyr	Arg	Val	Cys	Asn	
				325					330					335		
Val	Thr	Arg	Arg	Pro	Ala	Ser	His	Gln	Thr	Phe	Pro	Leu	Gln	Leu	Glu	
			340					345					350			
Asn	Gly	Gln	Ala	Met	Glu	Cys	Thr	Val	Ala	Gln	Tyr	Phe	Lys	Gln	Lys	
		355					360					365				
Tyr	Ser	Leu	Gln	Leu	Lys	Tyr	Pro	His	Leu	Pro	Cys	Leu	Gln	Val	Gly	
	370					375					380					
Gln	Glu	Gln	Lys	His	Thr	Tyr	Leu	Pro	Leu	Glu	Val	Cys	Asn	Ile	Val	
385					390					395					400	
Ala	Gly	Gln	Arg	Cys	Ile	Lys	Lys	Leu	Thr	Asp	Asn	Gln	Thr	Ser	Thr	
				405					410					415		
Met	Ile	Lys	Ala	Thr	Ala	Arg	Ser	Ala	Pro	Asp	Arg	Gln	Glu	Glu	Ile	
			420					425					430			
Ser	Arg	Leu	Val	Lys	Ser	Asn	Ser	Met	Val	Gly	Gly	Pro	Asp	Pro	Tyr	
		435					440					445				
Leu	Lys	Glu	Phe	Gly	Ile	Val	Val	His	Asn	Glu	Met	Thr	Glu	Leu	Thr	
	450					455					460					
Gly	Arg	Val	Leu	Pro	Ala	Pro	Met	Leu	Gln	Tyr	Gly	Gly	Arg	Asn	Lys	
465					470					475					480	

43bb

Thr	Val	Ala	Thr	Pro	Asn	Gln	Gly	Val	Trp	Asp	Met	Arg	Gly	Lys	Gln
				485					490					495	
Phe	Tyr	Ala	Gly	Ile	Glu	Ile	Lys	Val	Trp	Ala	Val	Ala	Cys	Phe	Ala
			500					505					510		
Pro	Gln	Lys	Gln	Cys	Arg	Glu	Asp	Leu	Leu	Lys	Ser	Phe	Thr	Asp	Gln
		515					520					525			
Leu	Arg	Lys	Ile	Ser	Lys	Asp	Ala	Gly	Met	Pro	Ile	Gln	Gly	Gln	Pro
	530					535					540				
Cys	Phe	Cys	Lys	Tyr	Ala	Gln	Gly	Ala	Asp	Ser	Val	Glu	Pro	Met	Phe
545					550					555					560
Lys	His	Leu	Lys	Met	Thr	Tyr	Val	Gly	Leu	Gln	Leu	Ile	Val	Val	Ile
				565					570					575	
Leu	Pro	Gly	Lys	Thr	Pro	Val	Tyr	Ala	Glu	Val	Lys	Arg	Val	Gly	Asp
			580					585					590		
Thr	Leu	Leu	Gly	Met	Ala	Thr	Gln	Cys	Val	Gln	Val	Lys	Asn	Val	Val
		595					600					605			
Lys	Thr	Ser	Pro	Gln	Thr	Leu	Ser	Asn	Leu	Cys	Leu	Lys	Ile	Asn	Ala
	610					615					620				
Lys	Leu	Gly	Gly	Ile	Asn	Asn	Val	Leu	Val	Pro	His	Gln	Arg	Pro	Ser
625					630					635					640
Val	Phe	Gln	Gln	Pro	Val	Ile	Phe	Leu	Gly	Ala	Asp	Val	Thr	His	Pro
				645					650					655	
Pro	Ala	Gly	Asp	Gly	Lys	Lys	Pro	Ser	Ile	Ala	Ala	Val	Val	Gly	Ser
			660					665					670		
Met	Asp	Gly	His	Pro	Ser	Arg	Tyr	Cys	Ala	Thr	Val	Arg	Val	Gln	Thr
		675					680					685			
Ser	Arg	Gln	Glu	Ile	Ser	Gln	Glu	Leu	Leu	Tyr	Ser	Gln	Glu	Val	Ile
	690					695					700				
Gln	Asp	Leu	Thr	Asn	Met	Val	Arg	Glu	Leu	Leu	Ile	Gln	Phe	Tyr	Lys
705					710					715					720
Ser	Thr	Arg	Phe	Lys	Pro	Thr	Arg	Ile	Ile	Tyr	Tyr	Arg	Gly	Gly	Val
				725					730					735	
Ser	Glu	Gly	Gln	Met	Lys	Gln	Val	Ala	Trp	Pro	Glu	Leu	Ile	Ala	Ile
			740					745					750		
Arg	Lys	Ala	Cys	Ile	Ser	Leu	Glu	Glu	Asp	Tyr	Arg	Pro	Gly	Ile	Thr
		755					760					765			
Tyr	Ile	Val	Val	Gln	Lys	Arg	His	His	Thr	Arg	Leu	Phe	Cys	Ala	Asp
	770					775					780				

43cc

Lys Thr Glu Arg Val Gly Lys Ser Gly Asn Val Pro Ala Gly Thr Thr
 785 790 795 800
 Val Asp Ser Thr Ile Thr His Pro Ser Glu Phe Asp Phe Tyr Leu Cys
 805 810 815
 Ser His Ala Gly Ile Gln Gly Thr Ser Arg Pro Ser His Tyr Gln Val
 820 825 830
 Leu Trp Asp Asp Asn Cys Phe Thr Ala Asp Glu Leu Gln Leu Leu Thr
 835 840 845
 Tyr Gln Leu Cys His Thr Tyr Val Arg Cys Thr Arg Ser Val Ser Ile
 850 855 860
 Pro Ala Pro Ala Tyr Tyr Ala Arg Leu Val Ala Phe Arg Ala Arg Tyr
 865 870 875 880
 His Leu Val Asp Lys Asp His Asp Ser Ala Glu Gly Ser His Val Ser
 885 890 895
 Gly Gln Ser Asn Gly Arg Asp Pro Gln Ala Leu Ala Lys Ala Val Gln
 900 905 910
 Ile His His Asp Thr Gln His Thr Met Tyr Phe Ala
 915 920

<210> 71
 <211> 855
 <212> PRT
 <213> human

<220>
 <223> eIF2C4, predicted protein sequence

<400> 71
 Ala Gly Pro Ala Gly Ala Gln Pro Leu Leu Met Val Pro Arg Arg Pro
 1 5 10 15
 Gly Tyr Gly Thr Met Gly Lys Pro Ile Lys Leu Leu Ala Asn Cys Phe
 20 25 30
 Gln Val Glu Ile Pro Lys Ile Asp Val Tyr Leu Tyr Glu Val Asp Ile
 35 40 45
 Lys Pro Asp Lys Cys Pro Arg Arg Val Asn Arg Glu Val Val Asp Ser
 50 55 60
 Met Val Gln His Phe Lys Val Thr Ile Phe Gly Asp Arg Arg Pro Val
 65 70 75 80
 Tyr Asp Gly Lys Arg Ser Leu Tyr Thr Ala Asn Pro Leu Pro Val Ala
 85 90 95

43dd

Thr	Thr	Gly	Val	Asp	Leu	Asp	Val	Thr	Leu	Pro	Gly	Glu	Gly	Gly	Lys
			100					105					110		
Asp	Arg	Pro	Phe	Lys	Val	Ser	Ile	Lys	Phe	Val	Ser	Arg	Val	Ser	Trp
		115					120					125			
His	Leu	Leu	His	Glu	Val	Leu	Thr	Gly	Arg	Thr	Leu	Pro	Glu	Pro	Leu
	130					135					140				
Glu	Leu	Asp	Lys	Pro	Ile	Ser	Thr	Asn	Pro	Val	His	Ala	Val	Asp	Val
145					150					155					160
Val	Leu	Arg	His	Leu	Pro	Ser	Met	Lys	Tyr	Thr	Pro	Val	Gly	Arg	Ser
				165					170					175	
Phe	Phe	Ser	Ala	Pro	Glu	Gly	Tyr	Asp	His	Pro	Leu	Gly	Gly	Gly	Arg
			180					185					190		
Glu	Val	Trp	Phe	Gly	Phe	His	Gln	Ser	Val	Arg	Pro	Ala	Met	Trp	Lys
		195					200					205			
Met	Met	Leu	Asn	Ile	Asp	Val	Ser	Ala	Thr	Ala	Phe	Tyr	Lys	Ala	Gln
	210					215					220				
Pro	Val	Ile	Gln	Phe	Met	Cys	Glu	Val	Leu	Asp	Ile	His	Asn	Ile	Asp
225					230					235					240
Glu	Gln	Pro	Arg	Pro	Leu	Thr	Asp	Ser	His	Arg	Val	Lys	Phe	Thr	Lys
				245					250					255	
Glu	Ile	Lys	Gly	Leu	Lys	Val	Glu	Val	Thr	His	Cys	Gly	Thr	Met	Arg
			260					265					270		
Arg	Lys	Tyr	Arg	Val	Cys	Asn	Val	Thr	Arg	Arg	Pro	Ala	Ser	His	Gln
		275					280					285			
Thr	Phe	Pro	Leu	Gln	Leu	Glu	Asn	Gly	Gln	Thr	Val	Glu	Arg	Thr	Val
	290					295					300				
Ala	Gln	Tyr	Phe	Arg	Glu	Lys	Tyr	Thr	Leu	Gln	Leu	Lys	Tyr	Pro	His
305					310					315					320
Leu	Pro	Cys	Leu	Gln	Val	Gly	Gln	Glu	Gln	Lys	His	Thr	Tyr	Leu	Pro
				325					330					335	
Leu	Glu	Val	Cys	Asn	Ile	Val	Ala	Gly	Gln	Arg	Cys	Ile	Lys	Lys	Leu
			340					345					350		
Thr	Asp	Asn	Gln	Thr	Ser	Thr	Met	Ile	Lys	Ala	Thr	Ala	Arg	Ser	Ala
		355					360					365			
Pro	Asp	Arg	Gln	Glu	Glu	Ile	Ser	Arg	Leu	Val	Arg	Ser	Ala	Asn	Tyr
	370					375					380				
Glu	Thr	Asp	Pro	Phe	Val	Gln	Glu	Phe	Gln	Phe	Lys	Val	Arg	Asp	Glu
385					390					395					400

43ee

Met	Ala	His	Val	Thr	Gly	Arg	Val	Leu	Pro	Ala	Pro	Met	Leu	Gln	Tyr		
				405					410					415			
Gly	Gly	Arg	Asn	Arg	Thr	Val	Ala	Thr	Pro	Ser	His	Gly	Val	Trp	Asp		
			420					425					430				
Met	Arg	Gly	Lys	Gln	Phe	His	Thr	Gly	Val	Glu	Ile	Lys	Met	Trp	Ala		
		435					440					445					
Ile	Ala	Cys	Phe	Ala	Thr	Gln	Arg	Gln	Cys	Arg	Glu	Glu	Ile	Leu	Lys		
	450					455					460						
Gly	Phe	Thr	Asp	Gln	Leu	Arg	Lys	Ile	Ser	Lys	Asp	Ala	Gly	Met	Pro		
465					470					475					480		
Ile	Gln	Gly	Gln	Pro	Cys	Phe	Cys	Lys	Tyr	Ala	Gln	Gly	Ala	Asp	Ser		
				485					490					495			
Val	Glu	Pro	Met	Phe	Arg	His	Leu	Lys	Asn	Thr	Tyr	Ser	Gly	Leu	Gln		
			500					505					510				
Leu	Ile	Ile	Val	Ile	Leu	Pro	Gly	Lys	Thr	Pro	Val	Tyr	Ala	Glu	Val		
		515					520					525					
Lys	Arg	Val	Gly	Asp	Thr	Leu	Leu	Gly	Met	Ala	Thr	Gln	Cys	Val	Gln		
	530					535					540						
Val	Lys	Asn	Val	Ile	Lys	Thr	Ser	Pro	Gln	Thr	Leu	Ser	Asn	Leu	Cys		
545					550					555					560		
Leu	Lys	Ile	Asn	Val	Lys	Leu	Gly	Gly	Ile	Asn	Asn	Ile	Leu	Val	Pro		
				565					570					575			
His	Gln	Arg	Pro	Ser	Val	Phe	Gln	Gln	Pro	Val	Ile	Phe	Leu	Gly	Ala		
			580					585					590				
Asp	Val	Thr	His	Pro	Pro	Ala	Gly	Asp	Gly	Lys	Lys	Pro	Ser	Ile	Ala		
		595					600					605					
Ala	Val	Val	Gly	Ser	Met	Asp	Ala	His	Pro	Ser	Arg	Tyr	Cys	Ala	Thr		
	610					615					620						
Val	Arg	Val	Gln	Arg	Pro	Arg	Gln	Glu	Ile	Ile	Gln	Asp	Leu	Ala	Ser		
625					630					635					640		
Met	Val	Arg	Glu	Leu	Leu	Ile	Gln	Phe	Tyr	Lys	Ser	Thr	Arg	Phe	Lys		
				645					650					655			
Pro	Thr	Arg	Ile	Ile	Phe	Tyr	Arg	Asp	Gly	Val	Ser	Glu	Gly	Gln	Phe		
			660					665					670				
Arg	Gln	Val	Leu	Tyr	Tyr	Glu	Leu	Leu	Ala	Ile	Arg	Glu	Ala	Cys	Ile		
		675					680					685					
Ser	Leu	Glu	Lys	Asp	Tyr	Gln	Pro	Gly	Ile	Thr	Tyr	Ile	Val	Val	Gln		
	690					695					700						

43ff

Lys Arg His His Thr Arg Leu Phe Cys Ala Asp Arg Thr Glu Arg Val
 705 710 715 720
 Gly Arg Ser Gly Asn Ile Pro Ala Gly Thr Thr Val Asp Thr Asp Ile
 725 730 735
 Thr His Pro Tyr Glu Phe Asp Phe Tyr Leu Cys Ser His Ala Gly Ile
 740 745 750
 Gln Gly Thr Ser Arg Pro Ser His Tyr His Val Leu Trp Asp Asp Asn
 755 760 765
 Cys Phe Thr Ala Asp Glu Leu Gln Leu Leu Thr Tyr Gln Leu Cys His
 770 775 780
 Thr Tyr Val Arg Cys Thr Arg Ser Val Ser Ile Pro Ala Pro Ala Tyr
 785 790 795 800
 Tyr Ala His Leu Val Ala Phe Arg Ala Arg Tyr His Leu Val Asp Lys
 805 810 815
 Glu His Asp Ser Ala Glu Gly Ser His Val Ser Gly Gln Ser Asn Gly
 820 825 830
 Arg Asp Pro Gln Ala Leu Ala Lys Ala Val Gln Ile His Gln Asp Thr
 835 840 845
 Leu Arg Thr Met Tyr Phe Ala
 850 855

<210> 72
 <211> 764
 <212> PRT
 <213> human

<220>
 <223> HILI, predicted protein sequence

<400> 72
 Ile Ser Ser Gly Asp Ala Gly Ser Thr Phe Met Glu Arg Gly Val Lys
 1 5 10 15
 Asn Lys Gln Asp Phe Met Asp Leu Ser Ile Cys Thr Arg Glu Lys Leu
 20 25 30
 Ala His Val Arg Asn Cys Lys Thr Gly Ser Ser Gly Ile Pro Val Lys
 35 40 45
 Leu Val Thr Asn Leu Phe Asn Leu Asp Phe Pro Gln Asp Trp Gln Leu
 50 55 60
 Tyr Gln Tyr His Val Thr Tyr Ile Pro Asp Leu Ala Ser Arg Arg Leu
 65 70 75 80

43gg

Arg	Ile	Ala	Leu	Leu	Tyr	Ser	His	Ser	Glu	Leu	Ser	Asn	Lys	Ala	Lys
				85					90					95	
Ala	Phe	Asp	Gly	Ala	Ile	Leu	Phe	Leu	Ser	Gln	Lys	Leu	Glu	Glu	Lys
			100					105					110		
Val	Thr	Glu	Leu	Ser	Ser	Glu	Thr	Gln	Arg	Gly	Glu	Thr	Ile	Lys	Met
		115					120					125			
Thr	Ile	Thr	Leu	Lys	Arg	Glu	Leu	Pro	Ser	Ser	Ser	Pro	Val	Cys	Ile
	130					135					140				
Gln	Val	Phe	Asn	Ile	Ile	Phe	Arg	Lys	Ile	Leu	Lys	Lys	Leu	Ser	Met
145					150					155					160
Tyr	Gln	Ile	Gly	Arg	Asn	Phe	Tyr	Asn	Pro	Ser	Glu	Pro	Met	Glu	Ile
				165				170						175	
Pro	Gln	His	Lys	Leu	Ser	Leu	Trp	Pro	Gly	Phe	Ala	Ile	Ser	Val	Ser
			180					185					190		
Tyr	Phe	Glu	Arg	Lys	Leu	Leu	Phe	Ser	Ala	Asp	Val	Ser	Tyr	Lys	Val
		195					200				205				
Leu	Arg	Asn	Glu	Thr	Val	Leu	Glu	Phe	Met	Thr	Ala	Leu	Cys	Gln	Arg
	210					215					220				
Thr	Gly	Leu	Ser	Cys	Phe	Thr	Gln	Thr	Cys	Glu	Lys	Gln	Leu	Ile	Gly
225					230					235					240
Leu	Ile	Val	Leu	Thr	Arg	Tyr	Asn	Asn	Arg	Thr	Tyr	Ser	Ile	Asp	Asp
				245					250					255	
Ile	Asp	Trp	Ser	Val	Lys	Pro	Thr	His	Thr	Phe	Gln	Lys	Arg	Asp	Gly
			260					265					270		
Thr	Glu	Ile	Thr	Tyr	Val	Asp	Tyr	Tyr	Lys	Gln	Gln	Tyr	Asp	Ile	Thr
		275				280						285			
Val	Ser	Asp	Leu	Asn	Gln	Pro	Met	Leu	Val	Ser	Leu	Leu	Lys	Lys	Lys
	290					295					300				
Arg	Asn	Asp	Asn	Ser	Glu	Ala	Gln	Leu	Ala	His	Leu	Ile	Pro	Glu	Leu
305					310					315					320
Cys	Phe	Leu	Thr	Gly	Leu	Thr	Asp	Gln	Ala	Thr	Ser	Asp	Phe	Gln	Leu
				325					330					335	
Met	Lys	Ala	Val	Ala	Glu	Lys	Thr	Arg	Leu	Ser	Pro	Ser	Gly	Arg	Gln
			340					345					350		
Gln	Arg	Leu	Ala	Arg	Leu	Val	Asp	Asn	Ile	Gln	Arg	Asn	Thr	Asn	Ala
		355					360					365			
Arg	Phe	Glu	Leu	Glu	Thr	Trp	Gly	Leu	His	Phe	Gly	Ser	Gln	Ile	Ser
	370					375					380				

43hh

Leu	Thr	Gly	Arg	Ile	Val	Pro	Ser	Glu	Lys	Ile	Leu	Met	Gln	Asp	His
385					390					395					400
Ile	Cys	Gln	Pro	Val	Ser	Ala	Ala	Asp	Trp	Ser	Lys	Asp	Ile	Arg	Thr
				405					410					415	
Cys	Lys	Ile	Leu	Asn	Ala	Gln	Ser	Leu	Asn	Thr	Trp	Leu	Ile	Leu	Cys
			420					425					430		
Ser	Asp	Arg	Thr	Glu	Tyr	Val	Ala	Glu	Ser	Phe	Leu	Asn	Cys	Leu	Arg
		435					440					445			
Arg	Val	Ala	Gly	Ser	Met	Gly	Phe	Asn	Val	Met	Cys	Ile	Leu	Pro	Ser
	450					455					460				
Asn	Gln	Lys	Thr	Tyr	Tyr	Asp	Ser	Ile	Lys	Lys	Tyr	Leu	Ser	Ser	Asp
465					470					475					480
Cys	Pro	Val	Pro	Ser	Gln	Cys	Val	Leu	Ala	Arg	Thr	Leu	Asn	Lys	Gln
				485					490					495	
Gly	Met	Met	Met	Ser	Ile	Ala	Thr	Lys	Ile	Ala	Met	Gln	Met	Thr	Cys
			500					505					510		
Lys	Leu	Gly	Gly	Glu	Leu	Trp	Ala	Val	Glu	Ile	Pro	Leu	Lys	Ser	Leu
		515					520					525			
Met	Val	Val	Gly	Ile	Asp	Val	Cys	Lys	Asp	Ala	Leu	Ser	Lys	Asp	Val
	530					535					540				
Met	Val	Val	Gly	Cys	Val	Ala	Ser	Val	Asn	Pro	Arg	Ile	Thr	Arg	Trp
545					550					555					560
Phe	Ser	Arg	Cys	Ile	Leu	Gln	Arg	Thr	Met	Thr	Asp	Val	Ala	Asp	Cys
				565					570					575	
Leu	Lys	Val	Phe	Met	Thr	Gly	Ala	Leu	Asn	Lys	Trp	Tyr	Lys	Tyr	Asn
			580					585					590		
His	Asp	Leu	Pro	Ala	Arg	Ile	Ile	Val	Tyr	Arg	Ala	Gly	Val	Gly	Asp
		595					600					605			
Gly	Gln	Leu	Lys	Thr	Leu	Ile	Glu	Tyr	Glu	Val	Pro	Gln	Leu	Leu	Ser
	610					615					620				
Ser	Val	Ala	Glu	Ser	Ser	Ser	Asn	Thr	Ser	Ser	Arg	Leu	Ser	Val	Ile
625					630					635					640
Val	Val	Arg	Lys	Lys	Cys	Met	Pro	Arg	Phe	Phe	Thr	Glu	Met	Asn	Arg
				645					650					655	
Thr	Val	Gln	Asn	Pro	Pro	Leu	Gly	Thr	Val	Val	Asp	Ser	Glu	Ala	Thr
			660					665					670		
Arg	Asn	Glu	Trp	Gln	Tyr	Asp	Phe	Tyr	Leu	Ile	Ser	Gln	Val	Ala	Cys
		675					680					685			

43ii

Arg Gly Thr Val Ser Pro Thr Tyr Tyr Asn Val Ile Tyr Asp Asp Asn
 690 695 700
 Gly Leu Lys Pro Asp His Met Gln Arg Leu Thr Phe Lys Leu Cys His
 705 710 715 720
 Leu Tyr Tyr Asn Trp Pro Gly Ile Val Ser Val Pro Ala Pro Cys Gln
 725 730 735
 Tyr Ala His Lys Leu Thr Phe Leu Val Ala Gln Ser Ile His Lys Glu
 740 745 750
 Pro Ser Leu Glu Leu Ala Asn His Leu Phe Tyr Leu
 755 760

<210> 73
 <211> 861
 <212> PRT
 <213> human

<220>
 <223> HIWI, predicted protein sequence

<400> 73
 Met Thr Gly Arg Ala Arg Ala Arg Ala Arg Gly Arg Ala Arg Gly Gln
 1 5 10 15
 Glu Thr Ala Gln Leu Val Gly Ser Thr Ala Ser Gln Gln Pro Gly Tyr
 20 25 30
 Ile Gln Pro Arg Pro Gln Pro Pro Pro Ala Glu Gly Glu Leu Phe Gly
 35 40 45
 Arg Gly Arg Gln Arg Gly Thr Ala Gly Gly Thr Ala Lys Ser Gln Gly
 50 55 60
 Leu Gln Ile Ser Ala Gly Phe Gln Glu Leu Ser Leu Ala Glu Arg Gly
 65 70 75 80
 Gly Arg Arg Arg Asp Phe His Asp Leu Gly Val Asn Thr Arg Gln Asn
 85 90 95
 Leu Asp His Val Lys Glu Ser Lys Thr Gly Ser Ser Gly Ile Ile Val
 100 105 110
 Arg Leu Ser Thr Asn His Phe Arg Leu Thr Ser Arg Pro Gln Trp Ala
 115 120 125
 Leu Tyr Gln Tyr His Ile Asp Tyr Asn Pro Leu Met Glu Ala Arg Arg
 130 135 140
 Leu Arg Ser Ala Leu Leu Phe Gln His Glu Asp Leu Ile Gly Lys Cys
 145 150 155 160

His	Ala	Phe	Asp	Gly 165	Thr	Ile	Leu	Phe	Leu 170	Pro	Lys	Arg	Leu	Gln 175	Gln
Lys	Val	Thr	Glu 180	Val	Phe	Ser	Lys	Thr 185	Arg	Asn	Gly	Glu	Asp 190	Val	Arg
Ile	Thr	Ile 195	Thr	Leu	Thr	Asn	Glu 200	Leu	Pro	Pro	Thr	Ser 205	Pro	Thr	Cys
Leu	Gln 210	Phe	Tyr	Asn	Ile	Ile 215	Phe	Arg	Arg	Leu	Leu 220	Lys	Ile	Met	Asn
Leu 225	Gln	Gln	Ile	Gly	Arg 230	Asn	Tyr	Tyr	Asn 235	Pro	Asn	Asp	Pro	Ile	Asp 240
Ile	Pro	Ser	His 245	Arg	Leu	Val	Ile	Trp	Pro 250	Gly	Phe	Thr	Thr	Ser 255	Ile
Leu	Gln	Tyr	Glu 260	Asn	Ser	Ile	Met	Leu 265	Cys	Thr	Asp	Val	Ser 270	His	Lys
Val	Leu 275	Arg	Ser	Glu	Thr	Val	Leu 280	Asp	Phe	Met	Phe 285	Asn	Phe	Tyr	His
Gln	Thr 290	Glu	Glu	His	Lys	Phe 295	Gln	Glu	Gln	Val	Ser 300	Lys	Glu	Leu	Ile
Gly 305	Leu	Val	Val	Leu	Thr 310	Lys	Tyr	Asn	Asn 315	Lys	Thr	Tyr	Arg	Val	Asp 320
Asp	Ile	Asp	Trp 325	Asp	Gln	Asn	Pro	Lys	Ser 330	Thr	Phe	Lys	Lys	Ala 335	Asp
Gly	Ser	Glu 340	Val	Ser	Phe	Leu	Glu 345	Tyr	Tyr	Arg	Lys	Gln 350	Tyr	Asn	Gln
Glu	Ile 355	Thr	Asp	Leu	Lys	Gln	Pro 360	Val	Leu	Val	Ser	Gln 365	Pro	Lys	Arg
Arg 370	Arg	Gly	Pro	Gly	Gly	Thr 375	Leu	Pro	Gly	Pro	Ala 380	Met	Leu	Ile	Pro
Glu 385	Leu	Cys	Tyr	Leu	Thr 390	Gly	Leu	Thr	Asp	Lys 395	Met	Arg	Asn	Asp	Phe 400
Asn	Val	Met	Lys 405	Asp	Leu	Ala	Val	His	Thr 410	Arg	Leu	Thr	Pro	Glu 415	Gln
Arg	Gln	Arg	Glu 420	Val	Gly	Arg	Leu	Ile 425	Asp	Tyr	Ile	His	Lys 430	Asn	Asp
Asn	Val 435	Gln	Arg	Glu	Leu	Arg	Asp 440	Trp	Gly	Leu	Ser	Phe 445	Asp	Ser	Asn
Leu 450	Leu	Ser	Phe	Ser	Gly	Arg 455	Ile	Leu	Gln	Thr	Glu 460	Lys	Ile	His	Gln

43kk

Gly	Gly	Lys	Thr	Phe	Asp	Tyr	Asn	Pro	Gln	Phe	Ala	Asp	Trp	Ser	Lys	465	470	475	480
Glu	Thr	Arg	Gly	Ala	Pro	Leu	Ile	Ser	Val	Lys	Pro	Leu	Asp	Asn	Trp		485	490	495
Leu	Leu	Ile	Tyr	Thr	Arg	Arg	Asn	Tyr	Glu	Ala	Ala	Asn	Ser	Leu	Ile		500	505	510
Gln	Asn	Leu	Phe	Lys	Val	Thr	Pro	Ala	Met	Gly	Met	Gln	Met	Arg	Lys		515	520	525
Ala	Ile	Met	Ile	Glu	Val	Asp	Asp	Arg	Thr	Glu	Ala	Tyr	Leu	Arg	Val		530	535	540
Leu	Gln	Gln	Lys	Val	Thr	Ala	Asp	Thr	Gln	Ile	Val	Val	Cys	Leu	Leu		545	550	555
Ser	Ser	Asn	Arg	Lys	Asp	Lys	Tyr	Asp	Ala	Ile	Lys	Lys	Tyr	Leu	Cys		565	570	575
Thr	Asp	Cys	Pro	Thr	Pro	Ser	Gln	Cys	Val	Val	Ala	Arg	Thr	Leu	Gly		580	585	590
Lys	Gln	Gln	Thr	Val	Met	Ala	Ile	Ala	Thr	Lys	Ile	Ala	Leu	Gln	Met		595	600	605
Asn	Cys	Lys	Met	Gly	Gly	Glu	Leu	Trp	Arg	Val	Asp	Ile	Pro	Leu	Lys		610	615	620
Leu	Val	Met	Ile	Val	Gly	Ile	Asp	Cys	Tyr	His	Asp	Met	Thr	Ala	Gly		625	630	635
Arg	Arg	Ser	Ile	Ala	Gly	Phe	Val	Ala	Ser	Ile	Asn	Glu	Gly	Met	Thr		645	650	655
Arg	Trp	Phe	Ser	Arg	Cys	Ile	Phe	Gln	Asp	Arg	Gly	Gln	Glu	Leu	Val		660	665	670
Asp	Gly	Leu	Lys	Val	Cys	Leu	Gln	Ala	Ala	Leu	Arg	Ala	Trp	Asn	Ser		675	680	685
Cys	Asn	Glu	Tyr	Met	Pro	Ser	Arg	Ile	Ile	Val	Tyr	Arg	Asp	Gly	Val		690	695	700
Gly	Asp	Gly	Gln	Leu	Lys	Thr	Leu	Val	Asn	Tyr	Glu	Val	Pro	Gln	Phe		705	710	715
Leu	Asp	Cys	Leu	Lys	Ser	Ile	Gly	Arg	Gly	Tyr	Asn	Pro	Arg	Leu	Thr		725	730	735
Val	Ile	Val	Val	Lys	Lys	Arg	Val	Asn	Thr	Arg	Phe	Phe	Ala	Gln	Ser		740	745	750
Gly	Gly	Arg	Leu	Gln	Asn	Pro	Leu	Pro	Gly	Thr	Val	Ile	Asp	Val	Glu		755	760	765

4311

Val Thr Arg Pro Glu Trp Tyr Asp Phe Phe Ile Val Ser Gln Ala Val
770 775 780

Arg Ser Gly Ser Val Ser Pro Thr His Tyr Asn Val Ile Tyr Asp Asn
785 790 795 800

Ser Gly Leu Lys Pro Asp His Ile Gln Arg Leu Thr Tyr Lys Leu Cys
805 810 815

His Ile Tyr Tyr Asn Trp Pro Gly Val Ile Arg Val Pro Ala Pro Cys
820 825 830

Gln Tyr Ala His Lys Leu Ala Phe Leu Val Gly Gln Ser Ile His Arg
835 840 845

Glu Pro Asn Leu Ser Leu Ser Asn Arg Leu Tyr Tyr Leu
850 855 860

<210> 74

<211> 2571

<212> DNA

<213> human

<220>

<223> eIF2C1, cDNA of predicted ORF

<400> 74

atggaagcgg gaccctcggg agcagctgcg ggcgcttacc tgccccccct gcagcaggtg 60
ttccaggcac ctgcgcggcc tggcattggc actgtgggga aaccaatcaa gtccttggcc 120
aattactttg aggtggacat ccctaagatc gacgtgtacc actacgaggt ggacatcaag 180
ccggataagt gtccccgtag agtcaaccgg gaagtgggtg aatacatggt ccagcatttc 240
aagcctcaga tctttggtga tcgcaagcct gtgtatgatg gaaagaagaa catttacact 300
gtcacagcac tgccatttgg caacgaacgg gtgcactttg aggtgacaat ccctggggaa 360
gggaaggatc gaatctttaa ggtctccatc aagtggctag ccattgtgag ctggcgaatg 420
ctgcatgagg ccctggtcag cggccagatc cctgttcctt tggagtctgt gcaagccctg 480
gatgtggcca tgaggcacct ggcatccatg aggtacaccc ctgtgggccg ctcccttctc 540
tcaccgcctg agggctacta ccaccgcctg ggggggtgggc gcgaggtctg gttcggcttt 600
caccagtctg tgcgccctgc catgtggaag atgatgctca acattgatgt ctacgccact 660
gccttttata aggcacagcc agtgattgag ttcattgtgt aggtgctgga catcaggaac 720
atagatgagc agccaagcc cctcacggac tctcagcgcg ttcgcttcac caaggagatc 780
aagggcctga aggtggaagt caccactgt ggacagatga agaggaagta ccgcgtgtgt 840

43mm

aatgttacct gtcgcctctg tagccatcag acattcccct tacagctgga gagtggacag 900
actgtggagt gcacagtggc acagtatttc aagcagaaat ataaccttca gctcaagtat 960
ccccatctgc cctgcctaca agttggccag gaacaaaagc atacctacct tcccctagag 1020
gtctgtaaca ttgtggctgg gcagcgctgt attaaaaagc tgaccgacaa ccagacctcg 1080
accatgataa aggccacagc tagatccgct ccagacagac aggaggagat cagtcgcctg 1140
atgaagaatg ccagctacaa cttagatccc tacatccagg aatttgggat caaagtgaag 1200
gatgacatga cggaggtgac agggcgagtg ctgccggcgc ccattcttga gtacggcggc 1260
cggaaccggg ccattgccac acccaatcag ggtgtctggg acatgcgggg gaaacagttc 1320
tacaatggga ttgagatcaa agtctgggac atcgctgct tcgcacccca aaaacagtgt 1380
cgagaagagg tgctcaagaa cttcacagac cagctgcgga agatttccaa ggatgcgggg 1440
atgcctatcc aggggtcaacc ttgtttctgc aaatatgcac agggggcaga cagcgtggag 1500
cctatgttcc ggcattctca gaacacctac tcagggtctg agctcattat tgtcatcctg 1560
ccagggaaga cgccggtgta tgctgaggtg aaacgtgtcg gagatacact cttgggaatg 1620
gctacgcagt gtgtgcaggt gaagaacgtg gtcaagacct cacctcagac tctgtccaac 1680
ctctgcctca agatcaatgt caaacttggt ggcattaaca acatcctagt cccacaccag 1740
cgctctgccg tttttcaaca gccagtgata ttctggggag cagatgttac acacccccca 1800
gcaggggatg ggaaaaaacc ttctatcaca gcagtggtag gcagtatgga tgcccccccc 1860
agccgatact gtgctactgt gcgggtacag cgaccacggc aagagatcat tgaagacttg 1920
tcctacatgg tgcgtgagct cctcatccaa ttctacaagt ccaccggtt caagcctacc 1980
cgcatcatct tctaccgaga tggggtgcct gaaggccagc taccacagat actccactat 2040
gagctactgg ccattcgtga tgctgcac aaactggaaa aggactacca gcctgggatc 2100
acttatattg tgggtgcagaa acgcatcac acccgcttt tctgtgctga caagaatgag 2160
cgaattggga agagtggtaa catcccagct gggaccacag tggacaccaa catcacccac 2220
ccatttgagt ttgacttcta tctgtgcagc cacgcaggca tccagggcac cagccgacca 2280
tcccattact atgttctttg ggatgacaac cgtttcacag cagatgagct ccagatcctg 2340
acgtaccagc tgtgccacac ttacgtacga tgcacacgct ctgtctctat cccagcacct 2400
gcctactatg cccgcctggt ggctttccgg gcacgatacc acctggtgga caaggagcat 2460
gacagtggag aggggagcca catatcgggg cagagcaatg ggcgggaccc ccaggccctg 2520
gcaaagccg tgcaggttca ccaggatact ctgcgcacca tgtacttcgc t 2571

43nn

<210> 75
<211> 2580
<212> DNA
<213> human

<220>
<223> eIF2C2, cDNA of predicted ORF

<400> 75
atgggtgttc tctctgccat tcccgcactt gcacctcctg cgccgccgcc ccccatccaa 60
ggatatgcct tcaagcctcc acctagacct gactttggga cctccgggag aacaatcaaa 120
ttacaggcca atttcttcga aatggacatc cccaaaattg acatctatca ttatgaattg 180
gatatcaagc cagagaagtg cccgaggaga gttaacaggg aaatcgtgga acacatggtc 240
cagcacttta aaacacagat ctttggggat cggaagcccg tgtttgacgg caggaagaat 300
ctatacacag ccatgcccct tccgattggg agggacaagg tggagctgga ggtcacgctg 360
ccaggagaag gcaaggatcg catcttcaag gtgtccatca agtgggtgtc ctgcgtgagc 420
ttgcaggcgt tacacgatgc actttcaggg cggtgccca gcgtcccttt tgagacgac 480
caggccctgg acgtggatc gaggcacttg ccatccatga ggtacacccc cgtgggccgc 540
tccttcttca ccgcgtccga aggtgctct aacctcttg gcgggggccg agaagtgtgg 600
tttggtctcc atcagtcctg ccggccttct ctctggaaaa tgatgctgaa tattgatgtg 660
tcagcaacag cgtttttacaa ggcacagcca gtaatcgagt ttgtttgtga agttttggat 720
tttaaaagta ttgaagaaca acaaaaacct ctgacagatt cccaaagggt aaagtttacc 780
aaagaaatta aaggtctaaa ggtggagata acgcactgtg ggcagatgaa gaggaagtac 840
cgtgtctgca atgtgacccg gcggcccgcc agtcaccaaa cattcccgtc gcagcaggag 900
agcgggcaga cgggtggagt cacggtggcc cagtatttca aggacaggca caagttggtt 960
ctgcgtacc cccacctccc atgtttacaa gtcggacagg agcagaaaca cacctacctt 1020
ccctggagg tctgtaacat tgtggcagga caaagatgta ttaaaaaatt aacggacaat 1080
cagacctcaa ccatgatcag agcaactgct aggtcggcgc ccgatcggca agaagagatt 1140
agcaaattga tgcgaagtgc aagtttcaac acagatccat acgtccgtga atttggaatc 1200
atggtcaaag atgagatgac agacgtgact gggcgggtgc tgcagccgcc ctccatcctc 1260
tacgggggca ggaataaagc tattgcgacc cctgtccagg gcgtctggga catgcggaac 1320
aagcagttcc acacgggcat cgagatcaag gtgtgggcca ttgcgtgctt cgcgccccag 1380

4300

cgccagtgca cggaagtcca tctgaagtcc ttcacagagc agctcagaaa gatctcgaga 1440
 gacgctggca tgcccatcca gggccagccg tgcttctgca aatacgcgca gggggcggac 1500
 agcgtggagc ccatgttccg gcacctgaag aacacgtatg cgggcctgca gctggtggtg 1560
 gtcacctcgc ccggcaagac gcccggtgtac gccgaggtca agcgcgtggg agacacggtg 1620
 ctgggggatgg ccacgcagtg cgtgcagatg aagaacgtgc agaggaccac gccacagacc 1680
 ctgtccaacc tttgcctgaa gatcaacgtc aagctgggag gcgtgaacaa catcctgctg 1740
 ccccagggca ggccgcgggt gttccagcag cccgtcatct ttctgggagc agacgtcact 1800
 ccccccccg ccgggggatgg gaagaagccc tccattgccg ccgtggtggg cagcatggac 1860
 gccaccccca atcgctactg cgccaccgtg cgcggtgcagc agcaccggca ggagatcata 1920
 caagacctgg ccgccatggt ccgcgagctc ctcattccagt tctacaagtc cagcgcttc 1980
 aagcccaccc gcatcatctt ctaccgcgac ggtgtctctg aaggccagtt ccagcaggtt 2040
 ctccaccacg agttgctggc catccgtgag gcctgtatca agctagaaaa agactaccag 2100
 ccggggatca ccttcacgtt ggtgcagaag aggcaccaca cccggctctt ctgcactgac 2160
 aagaacgagc gggttgggaa aagtggaaac attccagcag gcacgactgt ggacacgaaa 2220
 atcaccaccc ccaccgagtt cgactttctac ctgtgtagtc acgctggcat ccaggggaca 2280
 agcaggcctt cgcactatca cgtcctctgg gacgacaatc gtttctcctc tgatgagctg 2340
 cagatcctaa cctaccagct gtgtcacacc tacgtgcgct gcacacgctc cgtgtccatc 2400
 ccagcgccag catactacgc tcacctggtg gccttccggg ccaggtagca cctggtggat 2460
 aaggaacatg acagtgtga aggaagccat acctctgggc agagtaacgg gcgagaccac 2520
 caagcactgg ccaaggcggg ccagggtcac caagacactc tgcgcaccat gtactttgct 2580

<210> 76
 <211> 2772
 <212> DNA
 <213> human

<220>
 <223> eIF2C3, cDNA of predicted ORF

<400> 76
 agccggagcc gggtccttgt ccccgggccg ggcgccgccg ccgccccctg ccacgcgcc 60
 gcgtctccgc ggcgccaccc cagcgccaat attccggaga tcaagcgtta cgcggcggcg 120
 gcggcgggcg cggcggggcc cggagcggga ggcgccgggg accggggcga ggcggccccc 180

43pp

gccgccgcca tggaggcgct gggacccgga cctccggcta gcctgtttca gccacctcgt 240
cgtcctggcc ttggaactgt tggaaaacca attcgactgt tagccaatca ttttcagggt 300
cagattccta aaatagatgt gtatcactat gatgtggata ttaagcctga aaaacggcct 360
cgtagagtca acagggaggt agtagataca atgggtgcggc acttcaagat gcaaataatt 420
ggtgatcggc agcctgggta tgatggcaaa agaaacatgt acacagcaca tccactacca 480
attggacggg atagggttga tatggaggtg actcttccag gcgagggtaa agaccaaaca 540
tttaaagtgt ctgttcagt ggtgtcagtt gtgagccttc agttgctttt agaagctttg 600
gctgggcact tgaatgaagt cccagatgac tcagtacaag cacttgatgt tatcacaaga 660
caccttcctt ccatgaggta caccacagtg ggcggttctt ttttctcacc cccggaaggt 720
tactaccacc ctctgggagg gggcagggag gtctggtttg gttttcatca gtctgtgaga 780
cctgccatgt ggaatatgat gctcaacatt gatgtatctg caactgcttt ctaccgggct 840
cagcctatca ttgagttcat gtgtgaggtt ttagacattc agaacatcaa tgaacagacc 900
aaacctctaa cagactccca gcgtgtcaaa tttaccaaaag aaatcagagg tctcaaagtt 960
gaggtgaccc actgtggaca gatgaaacga aaataccgag tttgtaatgt gactagacgg 1020
ccagccagtc atcaaaacttt tcctttgcag ctagaaaacg gtcaagctat ggaatgtaca 1080
gtagctcaat attttaagca aaagtatagt ctgcaactga aataccccca tcttccctgt 1140
ctccaagtgg gacaagaaca aaagcataca tacttgccac tcgaggtctg taatatagtg 1200
gcaggacagc gatgtatcaa gaagctcaca gacaatcaga cttccacaat gatcaaagct 1260
acagcaagat ctgctcctga cagacaggaa gagatcagta gactggtgaa gagcaacagt 1320
atgggtgggtg gacctgatcc ataccttaaa gaatttggtg ttgttggtcca caatgaaatg 1380
acagagctca caggcagggt acttccagca ccaatgctgc aatatggagg ccggaataaa 1440
acagtagcca caccacaacca ggggtgtctgg gacatgcgag gaaagcagtt ttatgctggc 1500
attgaaatta aagtttgggc agttgcttgt tttgcacctc agaaacaatg tagggaagat 1560
ttactaaaga gtttcactga ccagctgcgt aaaatctcta aggatgcagg aatgcccac 1620
cagggtcagc catgtttctg caagtatgca caaggtgcag acagtgtgga gcctatgttt 1680
aaacatctga aaatgactta tgtgggccta cagctaatag tggttatcct gcctggaaag 1740
acaccagtat atgcggaggt gaaacgtggt ggagataccc ttctaggtat ggccacacag 1800
tgtgtccagg taaaaaatgt agtgaagacc tcacctcaaa ccctttccaa tctttgcctg 1860
aagataaatg caaaacttgg aggaattaac aatgtgcttg tgccctcatca aaggccctcg 1920

43qq

gtgttccagc agcctgtcat cttcctggga gcggatgtca cacaccccc agcaggggat 1980
 gggaagaaac cttccattgc tgctgtggtt ggcagtatgg atggccaccc cagccggtac 2040
 tgtgccaccg ttcgggtgca gacttcccgg caggagatct cccaagagct cctctacagt 2100
 caagaggtea tccaggacct gactaacatg gttcgagagc tgctgattca gttctacaaa 2160
 tccacacgct tcaaaccacac tcggatcatc tattaccgtg gaggggtatc tgagggacaa 2220
 atgaaacagg tagcttggcc agaactaata gcaattcgaa aggcattgtat tagcttggaa 2280
 gaagattacc ggccaggaat aacttatatt gtggtgcaaa aaagacatca cacacgactc 2340
 ttctgtgcag ataaaacaga aagggtaggg aaaagtggca atgtaccagc aggcactaca 2400
 gtggatagta ccatcacaca tccatctgag tttgactttt acctctgtag tcatgcagga 2460
 attcagggaa ccagccgtcc ctacattac caggctctgt gggatgacaa ctgcttcaact 2520
 gcagatgaac tccagctact gacttaccag ctgtgtcaca cctatgtgag gtgcactcgc 2580
 tcagtctcta ttccagcccc tgcatattat gcccggttg tagcatttag ggcaaggat 2640
 catctggtgg ataaagatca tgacagtgcg gaaggcagtc atgtgtcagg acagagcaac 2700
 ggccgggatc ctcaggcctt ggctaaggct gtgcaaacc accatgatac ccagcacacg 2760
 atgtattttg cc 2772

<210> 77
 <211> 2568
 <212> DNA
 <213> human

<220>
 <223> eIF2C4, cDNA of predicted ORF

<400> 77
 gcaggaccgc ctggggccca gccctactc atggtgcca gaagacctgg ctatggcacc 60
 atgggcaaac ccattaaact gctggctaac tgttttcaag ttgaaatccc aaagattgat 120
 gtctacctct atgaggtaga tattaacca gacaagtgtc ctaggagagt gaacagggag 180
 gtggttgact caatgggtca gcattttaaa gtaactatat ttggagaccg tagaccagtt 240
 tatgatggaa aaagaagtct ttacaccgcc aatccacttc ctgtggcaac tacaggggta 300
 gatttagacg ttactttacc tggggaagggt ggaaaagatc gacctttcaa ggtgtcaatc 360
 aaatttgtct ctcgggtgag ttggcaccta ctgcatgaag tactgacagg acggaccttg 420
 cctgagccac tggaattaga caagccaatc agcactaacc ctgtccatgc cgttgatgtg 480

43rr

gtgctacgac atctgccctc catgaaatac acacctgtgg ggcgttcatt tttctccgct 540
 ccagaaggat atgaccaccc tctgggaggg ggcagggaag tgtgggtttgg attccatcag 600
 tctgttcggc ctgccatgtg gaaaatgatg cttaatatcg atgtttctgc cactgccttc 660
 taaaaagcac aacctgtaat tcagttcatg tgtgaagttc ttgatattca taatattgat 720
 gagcaaccaa gacctctgac tgattctcat cgggtaaaat tcaccaaaga gataaaaggt 780
 ttgaagggtg aagtgactca ttgtggaaca atgagacgga aataccgtgt ttgtaatgta 840
 acaaggaggg ctgccagtca tcaaaccctt cctttacagt tagaaaacgg ccaaactgtg 900
 gagagaacag tagcgcagta tttcagagaa aagtatactc ttcagctgaa gtacccgcac 960
 cttccctgtc tgcaagtcgg gcaggaacag aaacacacct acctgccact agaagtctgt 1020
 aatattgtgg cagggcaacg atgtatcaag aagctaacag acaatcagac ttccactatg 1080
 atcaaggcaa cagcaagatc tgcaccagat agacaagagg aaattagcag attggtaaga 1140
 agtgcaaatt atgaaacaga tccatttggt caggagtctc aatttaaagt tcgggatgaa 1200
 atggctcatg taactggacg cgtacttcca gcacctatgc tccagtatgg aggacggaat 1260
 cggacagtag caacaccgag ccatggagta tgggacatgc gagggaaaca attccacaca 1320
 ggagttgaaa tcaaaatgtg ggctatcgct tgttttgcc aacagaggca gtgcagagaa 1380
 gaaatattga agggtttcac agaccagctg cgtaagattt ctaaggatgc agggatgcc 1440
 atccagggcc agccatgctt ctgcaaatat gcacaggggg cagacagcgt agagcccatg 1500
 ttccggcatc tcaagaacac atattctggc ctacagctta ttatcgatc cctgccgggg 1560
 aagacaccag tgtatgcgga agtgaaacgt gtaggagaca cacttttggg tatggctaca 1620
 caatgtgttc aagtcaagaa tgtaataaaa acatctctc aaactctgtc aaacttgtgc 1680
 ctaaagataa atgttaaact cggagggatc aataatattc ttgtacctca tcaaagacct 1740
 tctgtgttcc agcaaccagt gatctttttg ggagccgatg tcaatcatcc acctgctggt 1800
 gatggaaaga agccttctat tgctgctgtt gtaggtagta tggatgcaca cccaagcaga 1860
 tactgtgcca cagtaagagt tcagagaccc cgacaggaga tcatccagga cttggcctcc 1920
 atggtcgggg aacttcttat tcaattttat aagtcaactc ggttcaagcc tactcgtatc 1980
 atcttttatc gggatggtgt ttcagagggg cagtttaggc aggtattata ttatgaacta 2040
 ctagcaattc gagaagcctg catcagtttg gagaaagact atcaacctgg aataacctac 2100
 attgtagttc agaagagaca tcacactcga ttattttgtg ctgataggac agaaagggtt 2160
 ggaagaagtg gcaatatccc agctggaaca acagttgata cagacattac acacccatat 2220

43ss

gagttcgatt tttacctctg tagccatgct ggaatacagg gtaccagtcg tccttcacac 2280
 tatcatgttt tatgggatga taactgcttt actgcagatg aacttcagct gctaacttac 2340
 cagctctgcc acacttacgt acgctgtaca cgatctgttt ctatacctgc accagcgtat 2400
 tatgctcacc tggtagcatt tagagccaga tatcatcttg tggacaaaga acatgacagt 2460
 gctgaaggaa gtcacgtttc aggacaaagc aatgggagag atccacaagc tcttgccaag 2520
 gctgtacaga ttcaccaaga taccttacgc acaatgtact tcgcttaa 2568

<210> 78
 <211> 2292
 <212> DNA
 <213> human

<220>
 <223> HILI, cDNA of predicted ORF

<400> 78
 atatcttctg gtgatgctgg aagtaccttc atggaaagag gtgtgaaaaa caaacaggac 60
 tttatggatt tgagtatctg taccagagaa aaattggcac atgtgagaaa ttgtaaaaca 120
 ggttccagtg gaatacctgt gaaactgggtt acaaacctct ttaacttaga ttttcccca 180
 gactggcagc tataccagta ccatgtgaca tatattccag atttagcatc tagaaggctg 240
 agaattgctt tactttatag tcatagtga ctttccaaca aagcaaaagc attcgacggg 300
 gccatccttt ttctgtcaca aaagctagaa gaaaagggtc cagagttgtc aagtgaact 360
 caaagagggtg agactataaa gatgactatc accctgaaga gggagctgcc atcaagttct 420
 ccctgtgtga tccaggctct caatatcatc ttcagaaaga tcctcaaaaa gttgtccatg 480
 taccaaattg gacggaactt ctataatcct tcagagccaa tggaaattcc ccagcacaaa 540
 ttatcccttt ggctgggtt tgccatttct gtgtcatatt ttgaaaggaa gctcctgttt 600
 agtgctgatg tgagttacaa agtcctccgg aatgagacgg ttctggaatt catgactgct 660
 ctctgtcaaa gaactggctt gtcctgtttc acccagacgt gtgagaagca gctaataagg 720
 ctcatgtcc ttacaagata caataacaga acctactcca ttgatgacat tgactgggtc 780
 gtgaagccca cacacacctt tcagaagcgg gatggcaccg agatcaccta tgtggattac 840
 tacaagcagc agtatgatat tactgtatcg gacctgaatc agcccatgct tgtagtctg 900
 ttaaagaaga agagaaatga caacagtga gctcagctcg cccacctgat acctgagctc 960
 tgctttctaa cagggtgac tgaccaggca acatctgatt tccagctgat gaaggctgtg 1020

43tt

gctgaaaaga cacgtctcag tccttcaggc cggcagcagc gcctggccag gcttgtggac 1080
 aacatccaga ggaataccaa tgctcgcttt gaactagaga cctggggact gcatttttga 1140
 agccagatat ctctgactgg ccggattgtg ccttcagaaa aaatattaat gcaagaccac 1200
 atatgtcaac ctgtgtctgc tgctgactgg tccaaggata ttcgaacttg caagatttta 1260
 aatgcacagt ctttgaatac ctggttgatt ttatgtagcg acagaactga atatgttgcc 1320
 gagagctttc tgaactgctt gagaagagtt gcaggttcca tgggatttaa tgtaatgtgc 1380
 attctgcctt ctaatcagaa gacctattat gattccatta aaaaatattt gagctcagac 1440
 tgcccagtc ccaagccaatg tgtgcttgct cggaccttga ataaacaggg catgatgatg 1500
 agtatcgcca ccaagatcgc tatgcagatg acttgcaagc tcggaggcga gctgtgggct 1560
 gtggaaatac ctttaaagtc cctgatggtg gtcgggtattg atgtctgtaa agatgcactc 1620
 agcaaggacg tgatggttgt tggatgcgtg gccagtgtta accccagaat caccaggtgg 1680
 ttttcccgt gtatccttca gagaacaatg actgatgttg cagattgctt gaaagttttc 1740
 atgactggag cactcaacaa atggtacaag tacaatcatg atttgccagc acggataatt 1800
 gtgtaccgtg ctggtgtagg ggatggtcag ctgaaaacac ttattgaata tgaagtccca 1860
 cagctgctga gcagtgtggc agaatccagc tcaaatacca gctcaagact gtcggtgatt 1920
 gtggtcagga agaagtgcac gccacgattc ttaccgaaa tgaaccgcac tgtacagaac 1980
 cccccacttg gcactgttgt ggattcagaa gcaacacgta acgaatggca gtatgacttt 2040
 tatctgatca gccaggtggc ctgccgggga actgttagtc ctacctacta taatgtcatc 2100
 tatgatgaca acggcttgaa gcccgaccat atgcagagac ttacattcaa attgtgccac 2160
 ctgtactaca actggccggg catagtcagt gtcccagcac catgtcagta tgctcacaag 2220
 ctgacctttc tgggtggcaca aagcattcat aaagaacca gtctggaatt agccaacat 2280
 ctcttctacc tg 2292

<210> 79
 <211> 2583
 <212> DNA
 <213> human

<220>
 <223> HIWI, cDNA of predicted ORF

<400> 79
 atgactggga gagcccgagc cagagccaga ggaagggccc gcggtcagga gacagcgag 60

43uu

ctggtgggct ccactgccag tcagcaacct gggttatattc agcctaggcc tcagccgcca 120
 ccagcagagg gggaattatt tggccgtgga cggcagagag gaacagcagg aggaacagcc 180
 aagtcacaag gactccagat atctgctgga tttcaggagt tatcgttagc agagagagga 240
 ggtcgtcgta gagattttca tgatcttggg gtgaatacaa ggcagaacct agaccatggt 300
 aaagaatcaa aaacagggttc ttcaggcatt atagtaagggt taagcactaa ccatttccgg 360
 ctgacatccc gtccccagtg ggccttatat cagtatcaca ttgactataa cccactgatg 420
 gaagccagaa gactccgttc agctcttctt tttcaacacg aagatctaataa tggaaagtgc 480
 catgcttttg atggaacgat attattttta cctaaaagac tacagcaaaa ggttactgaa 540
 gtttttagta agaccggaa tggagaggat gtgaggataa cgatcacttt aacaaatgaa 600
 cttccaccta catcaccaac ttgtttgcag ttctataata ttattttcag gaggcttttg 660
 aaaatcatga atttgcaaca aattggacga aattattata acccaaataa cccaattgat 720
 attccaagtc acaggttggg gatttggcct ggcttacta cttccatcct tcagtatgaa 780
 aacagcatca tgctctgcac tgacgttagc cataaagtcc ttcgaagtga gactgttttg 840
 gatttcatgt tcaactttta tcatcagaca gaagaacata aatttcaaga acaagtttcc 900
 aaagaactaa taggttttagt tgttcttacc aagtataaca ataagacata cagagtggat 960
 gatattgact gggaccagaa tccaagagc acctttaaga aagccgacgg ctctgaagtc 1020
 agcttcttag aatactacag gaagcaatac aaccaagaga tcaccgactt gaagcagcct 1080
 gtcttggtea gccagcccaa gagaaggcgg ggcctgggg ggacactgcc agggcctgcc 1140
 atgctcattc ctgagctctg ctatcttaca ggtctaactg ataaaatgcg taatgatttt 1200
 aacgtgatga aagacttagc cgttcataca agactaactc cagagcaaag gcagcgtgaa 1260
 gtgggacgac tcattgatta cattcataaa aacgataatg ttcaaaggga gcttcgagac 1320
 tggggtttga gctttgattc caacttactg tccttctcag gaagaatttt gcaaacagaa 1380
 aagattcacc aaggtggaaa aacatttgat tacaatccac aatttgcaga ttggtccaaa 1440
 gaaacaagag gtgcaccatt aattagtgtt aagccactag ataactggct gttgatctat 1500
 acgcgaagaa attatgaagc agccaattca ttgatacaaa atctatttaa agttacacca 1560
 gccatgggca tgcaaatgag aaaagcaata atgattgaag tggatgacag aactgaagcc 1620
 tacttaagag tcttacagca aaaggtcaca gcagacaccc agatagttgt ctgtctgttg 1680
 tcaagtaatc ggaaggacaa atacgatgct attaaaaaat acctgtgtac agattgccct 1740
 accccaagtc agtgtgtggg ggcccgaacc ttaggcaaac agcaaactgt catggccatt 1800

43vv

gctacaaaga ttgccctaca gatgaactgc aagatgggag gagagctctg gaggggtggac 1860
 atccccctga agctcgtgat gatcgttggc atcgattggt accatgacat gacagctggg 1920
 cggaggtcaa tcgcaggatt tggtgccagc atcaatgaag ggatgacccg ctggttctca 1980
 cgctgcatat ttcaggatag aggacaggag ctggtagatg ggctcaaagt ctgcctgcaa 2040
 gcggctctga gggcttggaa tagctgcaat gagtacatgc ccagccggat catcgtgtac 2100
 cgcgatggcg taggagacgg ccagctgaaa aactgggtga actacgaagt gccacagttt 2160
 ttggattgtc taaaatccat tggtagaggt tacaacccta gactaacggt aattgtggtg 2220
 aagaaaagag tgaacaccag attttttgct cagtctggag gaagacttca gaatccactt 2280
 cctggaacag ttattgatgt agaggttacc agaccagaat ggtatgactt ttttatcgtg 2340
 agccaggctg tgagaagtgg tagtgtttct cccacacatt acaatgtcat ctatgacaac 2400
 agcggcctga agccagacca catacagcgc ttgacctaca agctgtgcca catctattac 2460
 aactggccag gtgtcattcg tgttctgtct ccttgccagt acgcccacaa gctggctttt 2520
 cttgttggcc agagtattca cagagagcca aatctgtcac tgtcaaaccg cttttactac 2580
 ctc 2583

<210> 80
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:
 oligodeoxynucleotide - primer

<400> 80
 gaggtctgta acattgtggc 20

<210> 81
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:
 oligodeoxynucleotide - primer

<400> 81
 cggtagaaga tgatgcgggt 20

43ww

<210> 82
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 82
aagttcttga gcacctcttc tcga

24

<210> 83
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 83
gaggtctgta acattgtggc

20

<210> 84
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 84
cggtagaaga tgatgcgggt

20

<210> 85
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

43xx

<400> 85
ccacaccagc gctctgcc 18

<210> 86
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 86
ctcacgcacc atgtagga 18

<210> 87
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 87
gaggtctgta acattgtggc 20

<210> 88
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 88
cggtagaaga tgatgcgggt 20

<210> 89
<211> 18
<212> DNA
<213> Artificial Sequence

43yy

<220>

<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 89

atcctgctgc cccaaggg

18

<210> 90

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 90

gatctcctgc cggtgctg

18

<210> 91

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 91

gaggtctgta acattgtggc

20

<210> 92

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 92

cggtagaaga tgatgcgggt

20

43zz

<210> 93
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 93
gatctcctgc cggtgctg

18

<210> 94
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 94
agagcaacag tatggtgggt ggac

24

<210> 95
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 95
tggatgtgtg atggtact

18

<210> 96
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

43aaa

<400> 96
cctctacagt caagaggt

18

<210> 97
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 97
cacttgaatg aagtccca

18

<210> 98
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 98
tcctggatga cctcttgact gtag

24

<210> 99
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 99
agagcaacag tatggtgggt ggac

24

<210> 100
<211> 26
<212> DNA
<213> Artificial Sequence

43bbb

<220>

<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 100

tccggcatct caagaacaca tattct

26

<210> 101

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 101

gaactcatat ggggtgtgtaa tgtctg

26

<210> 102

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 102

atccaggact tggcctcc

18

<210> 103

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 103

cagcaciaat tatccctt

18

43ccc

<210> 104
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 104
cggcctgaag gactgagacg tgt

23

<210> 105
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 105
gtgtgtgggc ttcactga

18

<210> 106
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 106
tctctgtcaa agaactggct tgtcct

26

<210> 107
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

43ddd

<400> 107
ctgtacagtg cggttcat

18

<210> 108
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligodeoxynucleotide - primer

<400> 108
cggcctgaag gactgagacg tgt

23

- 44 -

Claims

1. Use of single-stranded RNA molecule having a length from 14-50 nucleotides wherein at least the 14-20 5'most nucleotides are complementary to a target transcript for the manufacture of an agent for inhibiting the expression of said target transcript in mammalian cells, said RNA molecule having a phosphate analogue at its 5'terminus, wherein said expression is inhibited by RNA-interference.
2. The use of claim 1, wherein said RNA molecule has a length from 15-29 nucleotides.
3. The use of claim 1 or 2, wherein said RNA molecule has a free 5'hydroxyl moiety or a moiety selected from phosphate groups or analogues thereof.
4. The use of claim 1 or 2, wherein said RNA molecule has 5'-moiety selected from 5'-monophosphate, 5'-diphosphate, 5'-triphosphate, 7-methylated or non-methylated 5'-guanosine cap, 5'-adenosine cap, and a modified or unmodified nucleotide cap structure, 5'-monothiophosphate, 5'-monodithiophosphate, 5'-phosphorothiolate; an additional combination of oxygen/sulfur replaced monophosphate, diphosphate and triphosphates, 5'-phosphoramidates, 5'-alkylphosphonates, 5'-alkyletherphosphonates, and ethoxymethyl.
5. The use of claim 4, wherein the nucleotide cap structure is (N-O-5'-(HO)(O)P-O-(HO)(O)P-O-P(HO)(O)-O-5').
6. The use of claim 4, wherein the combination of oxygen/sulfur replaced monophosphate, diphosphate and triphosphates is 5'-alpha-thiotriphosphate or 5'-gamma-thiotriphosphate.
7. The use of claim 4, wherein the 5'-alkylphosphonates is methylphosphonate, ethylphosphonate, isopropylphosphonate, or propylphosphonate.

- 45 -

8. The use of claim 4, wherein the 5'-alkyletherphosphonate is 5'-methoxymethylphosphonate or 5'-ethoxymethylphosphonate.
9. The use of any one of claims 1 to 8, wherein said RNA molecule is completely complementary to said target transcript optionally with exception of nucleotides that extend beyond position 20 being counted from the 5'terminus.
10. The use of any one of claims 1 to 9, wherein said RNA molecule comprises at least one modified nucleotide analogue.
11. The use of claim 10, wherein the modified nucleotide analogues are selected from sugar-backbone- and nucleobase-modified ribonucleotides and combinations thereof.
12. The use of any one of claims 1 to 11 for the inhibition of target gene expression *in vitro*.
13. The use of any one of claims 1 to 11 for the inhibition of target gene expression *in vivo*.
14. The use of any one of claims 1 to 13, wherein said RNA molecule is formulated as a pharmaceutical composition which contains a pharmaceutically acceptable carrier.
15. The use of claim 14, wherein said carrier is selected from cationic liposomes and cationic lipid formulations.
16. The use of any one of claims 1 to 15, wherein said RNA molecule is covalently coupled to a copolymer or a microparticle.
17. The use of claim 16, wherein said covalent coupling between said RNA molecule and said copolymer or microparticle occurs via the 3'-terminus of the RNA molecule.
18. The use of any one of claims 13 to 17 for diagnosis of a disease or

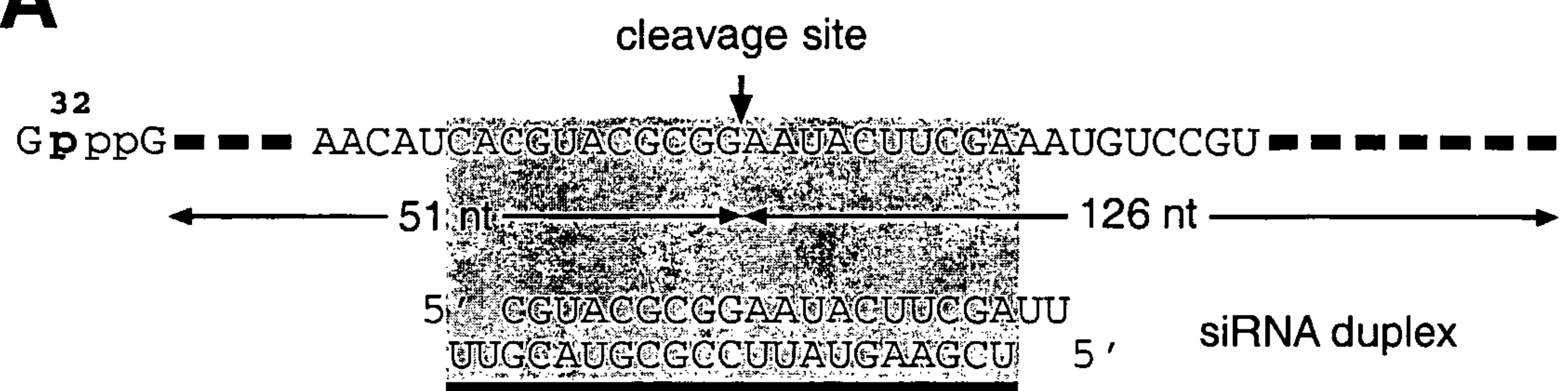
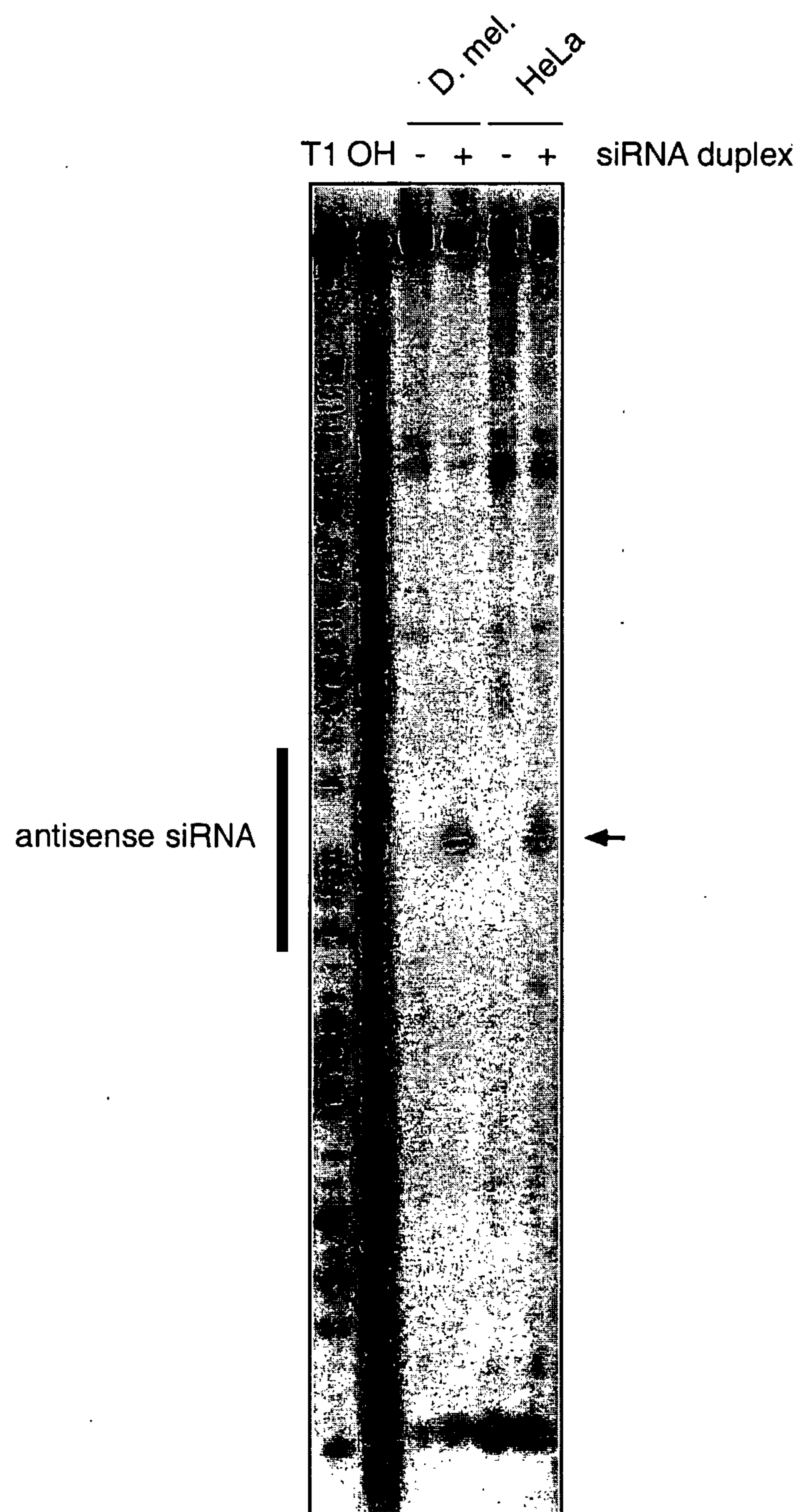
- 46 -

condition which is amenable to the use of a single-stranded RNA molecule having a length of 14-50 nucleotides, wherein the 5' most nucleotides are complementary to a target transcript, wherein said RNA molecule has a phosphate analog at its 5' terminus to inhibit the expression of a target transcript by RNA interference.

19. The use of any one of claims 13 to 17 for therapy of a disease or condition which is amenable to the use of a single-stranded RNA molecule having a length of 14-50 nucleotides, wherein the 5' most nucleotides are complementary to a target transcript, wherein said RNA molecule has a phosphate analog at its 5' terminus to inhibit the expression of a target transcript by RNA interference.
20. The use of claim 19 for the prevention or treatment of diseases associated with overexpression of at least one target transcript.
21. The use of claim 20, wherein the diseases are selected from tumor diseases, inflammatory diseases, infectious diseases, degenerative diseases and autoimmune diseases.
22. The use of claim 21, wherein the infectious diseases are viral infections.

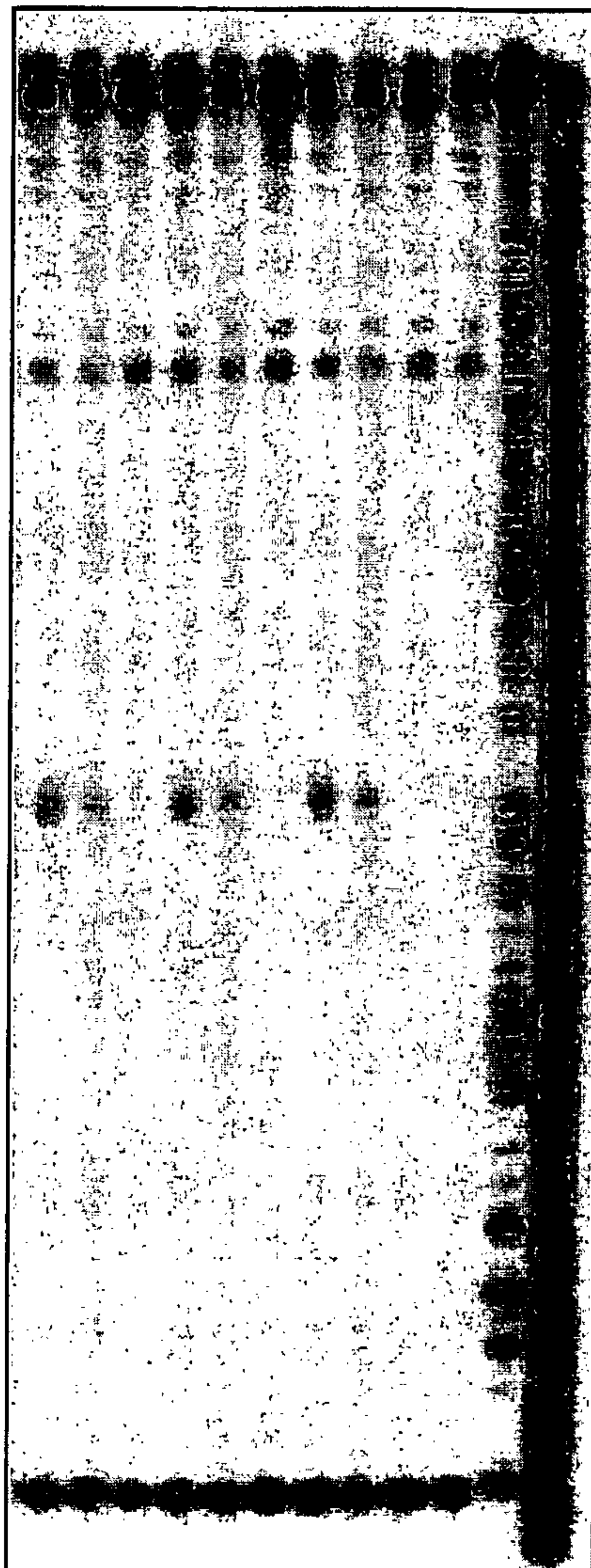
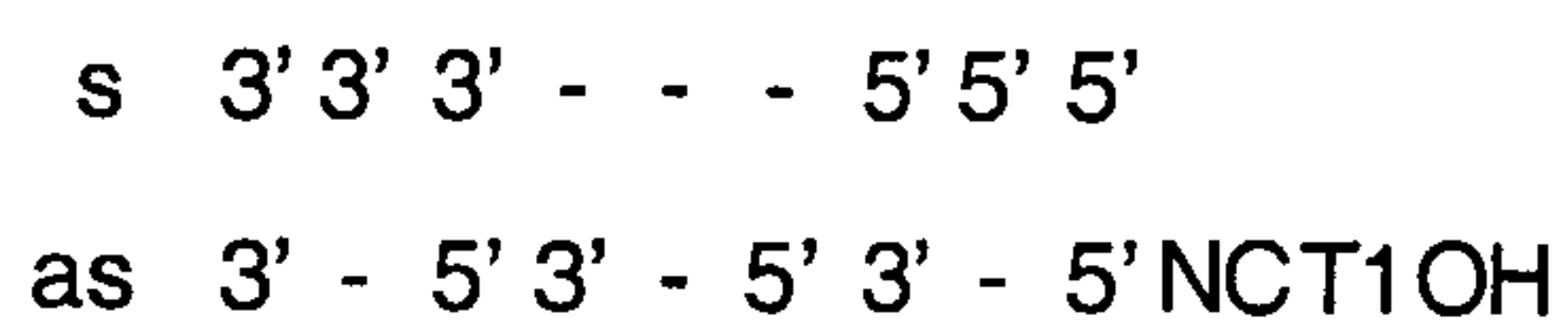
1/25

Martinez et al. Figure 1

A**B**

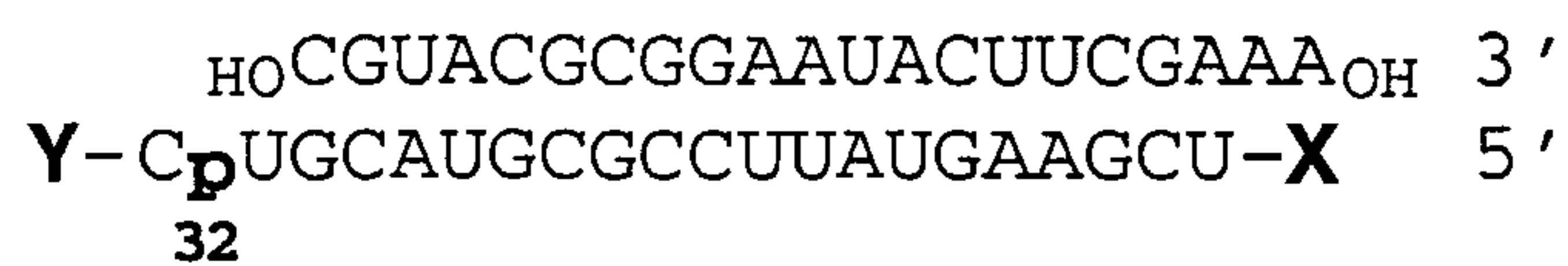
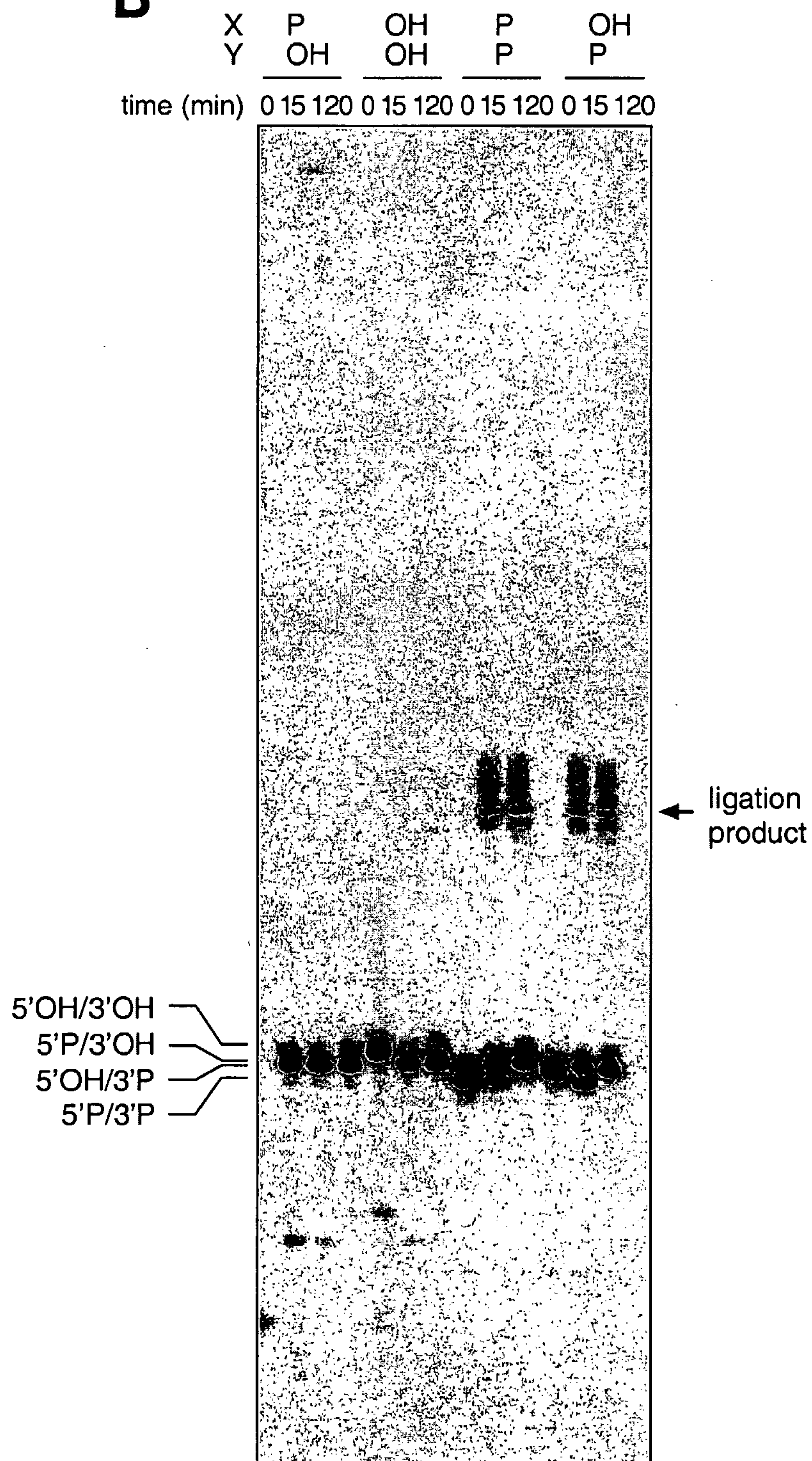
2/25

Martinez et al. Figure 2

A**B**

3/25

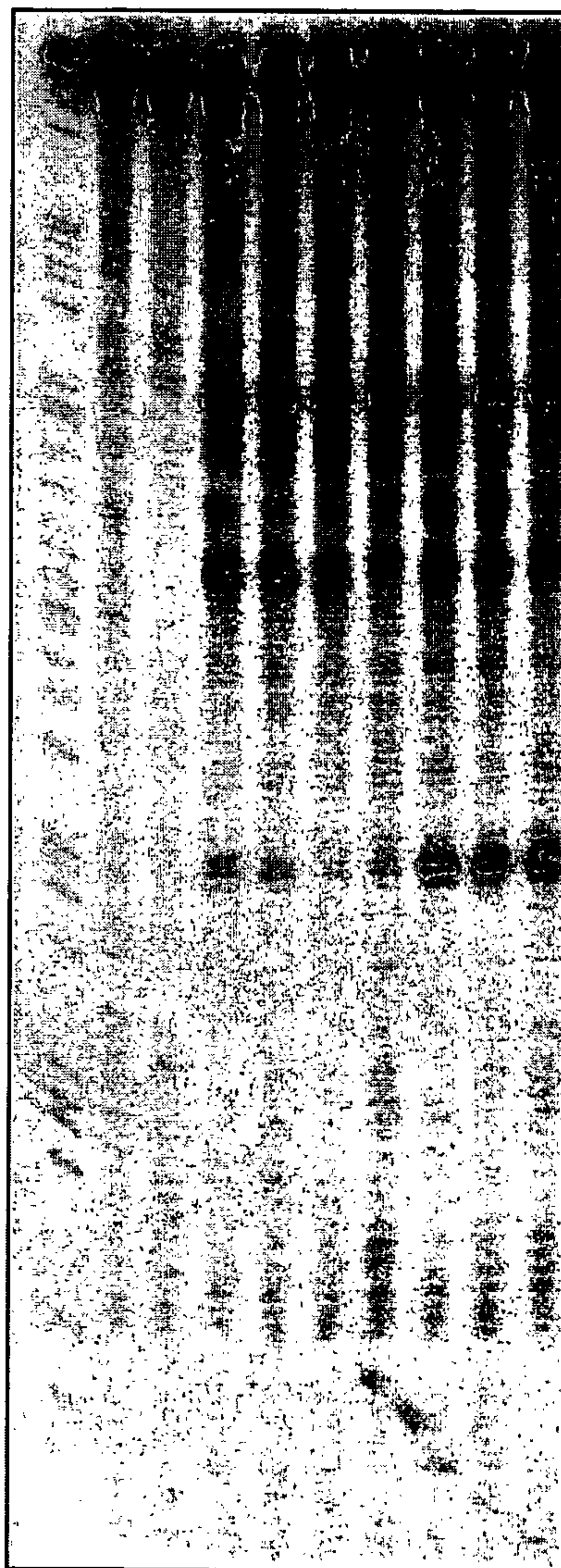
Martinez et al. Figure 3

A**B**

4/25

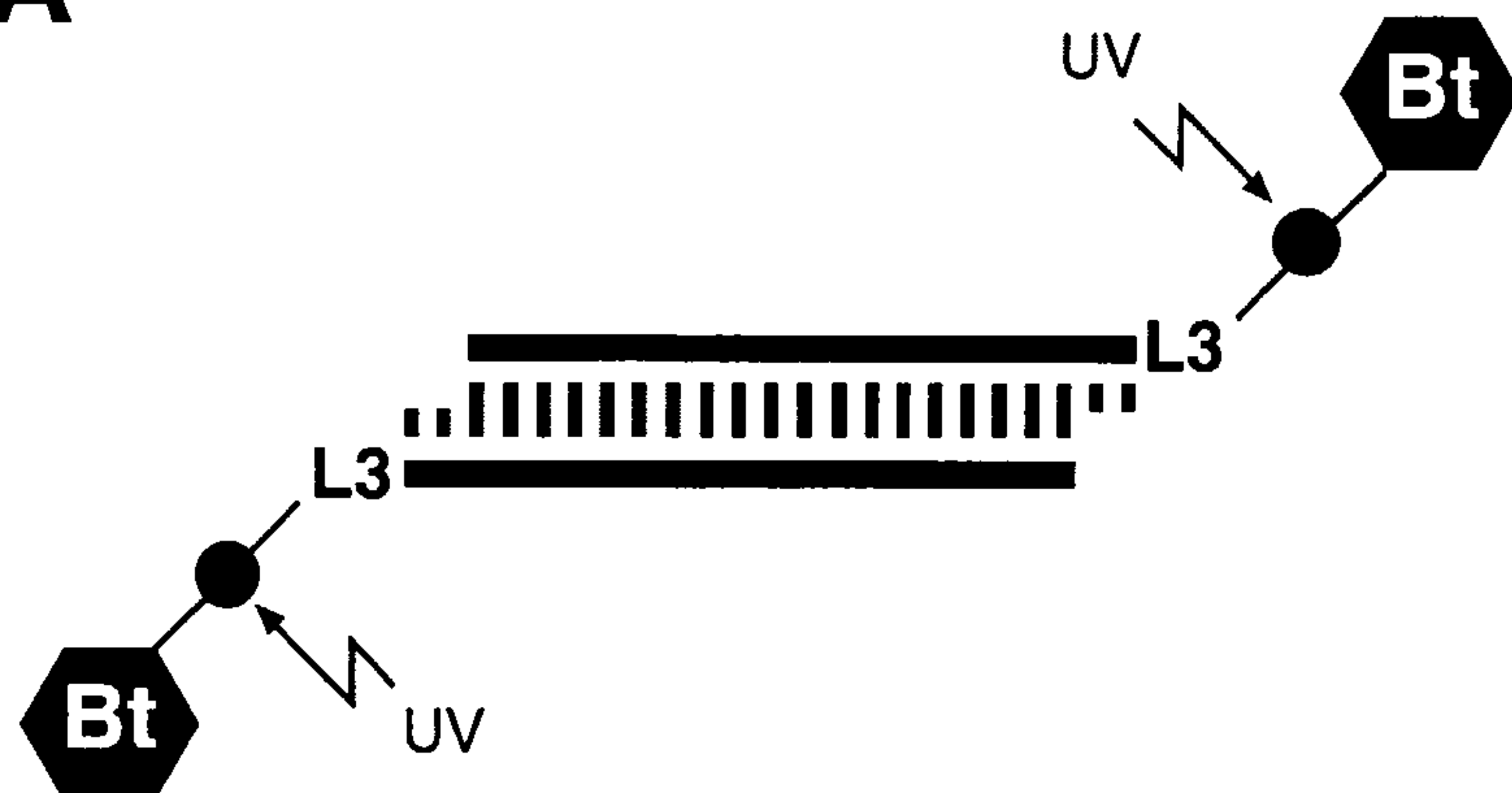
Martinez et al. Figure 4

time point of competitor siRNA addition (min)		0				15				
										
siRNA duplexes	competitor (nM)	100	0	0	1000	10	1000			
	specific (nM)	0	100							
HeLa S100		T1	-	-	+	+	+	+	+	+



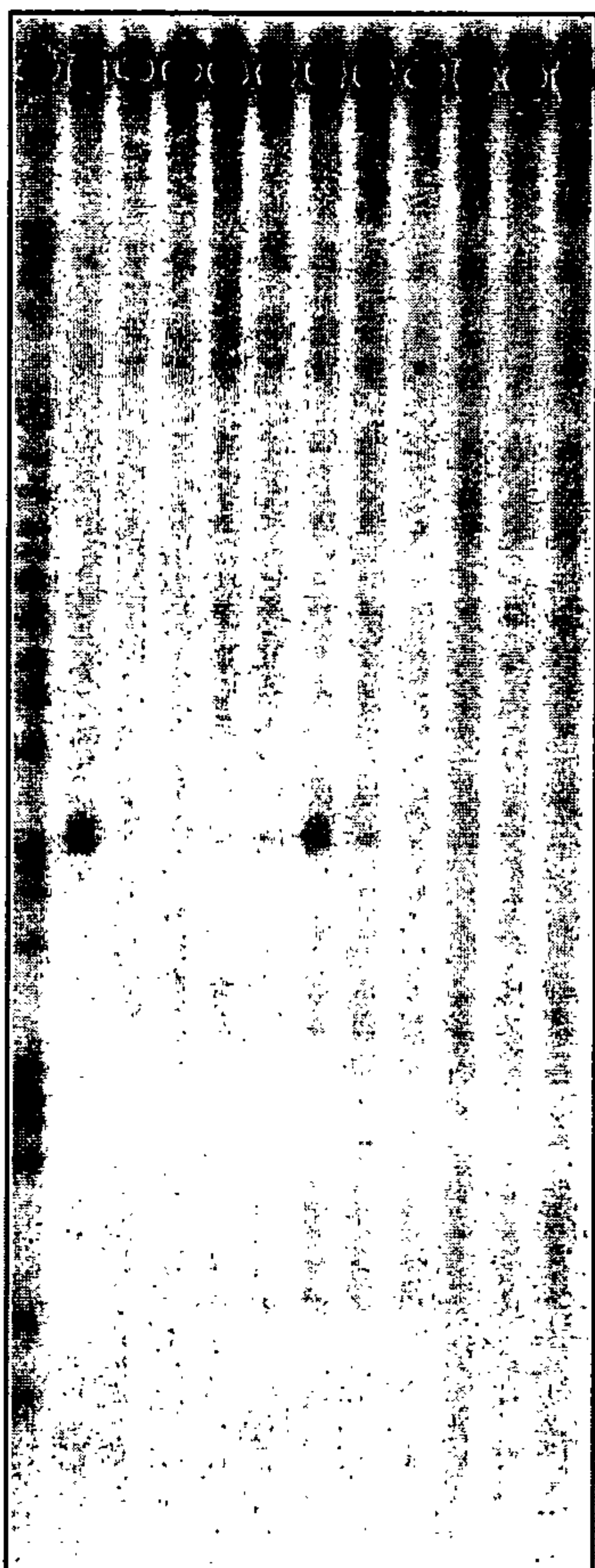
5/25

Martinez et al. Figure 5

A**B**

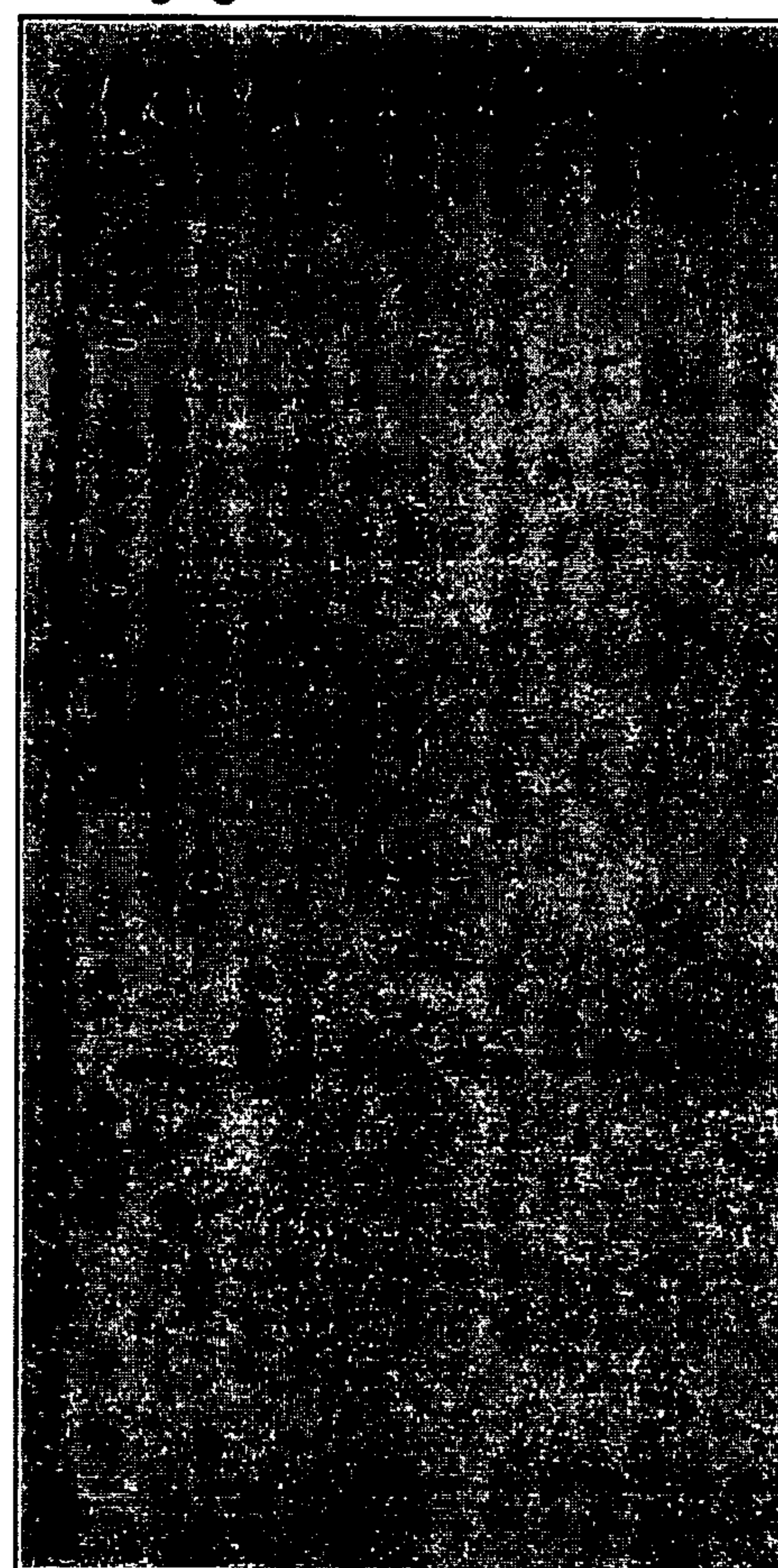
Superdex 200

aldolase
BSA
T1 UV elu 2 4 6 8 10 11 12 13 14 15

**C**

glycerol gradient 5-20%

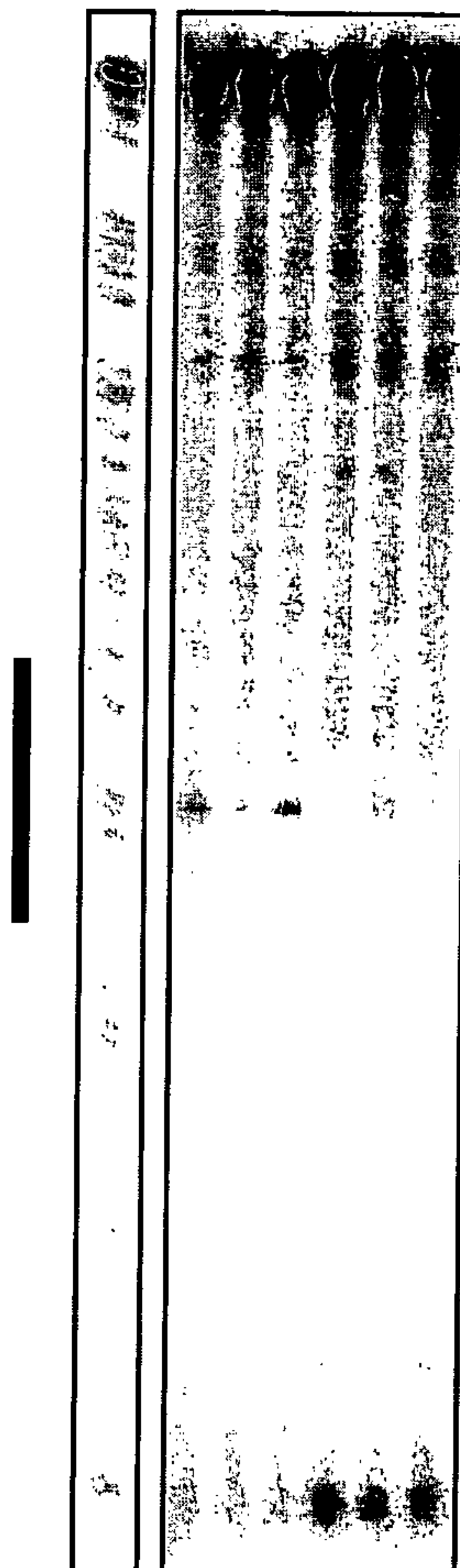
aldolase
BSA
OHT1 5 100 UV elu 1 3 5 7 9 11 13 15 17 19 20



6/25

Martinez et al. Figure 6

		UV	flow-
		eluate	through
s-Biotin		+	+
as-Biotin		+	+
T1	+	+	+



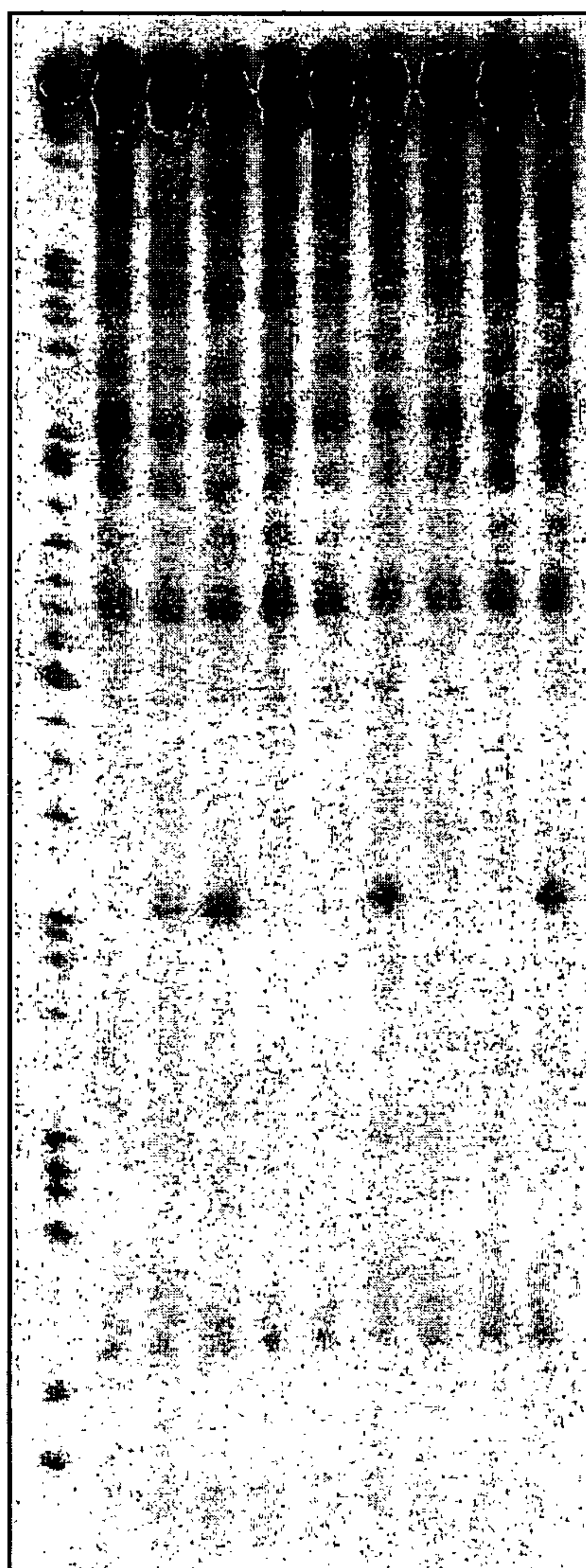
7/25

Martinez et al. Figure 7

A

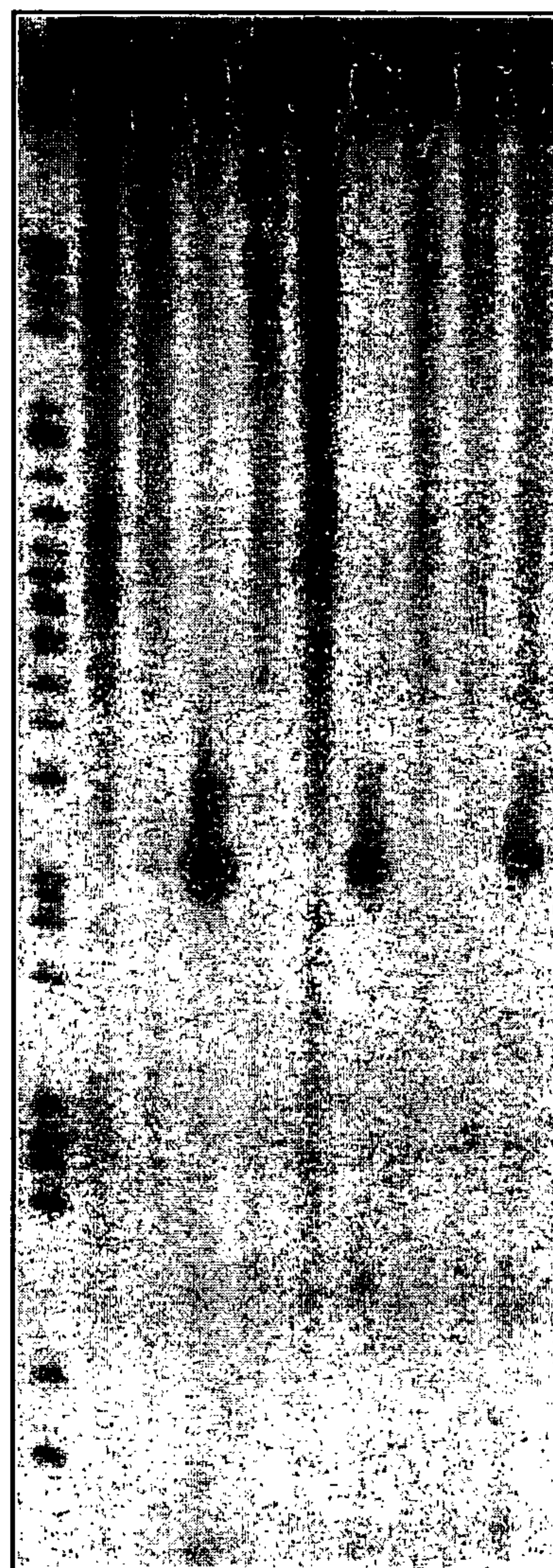
HeLa S100

c (nM)	100			10			1			
siRNA	T1	s	as	ds	s	as	ds	s	as	ds

**B**

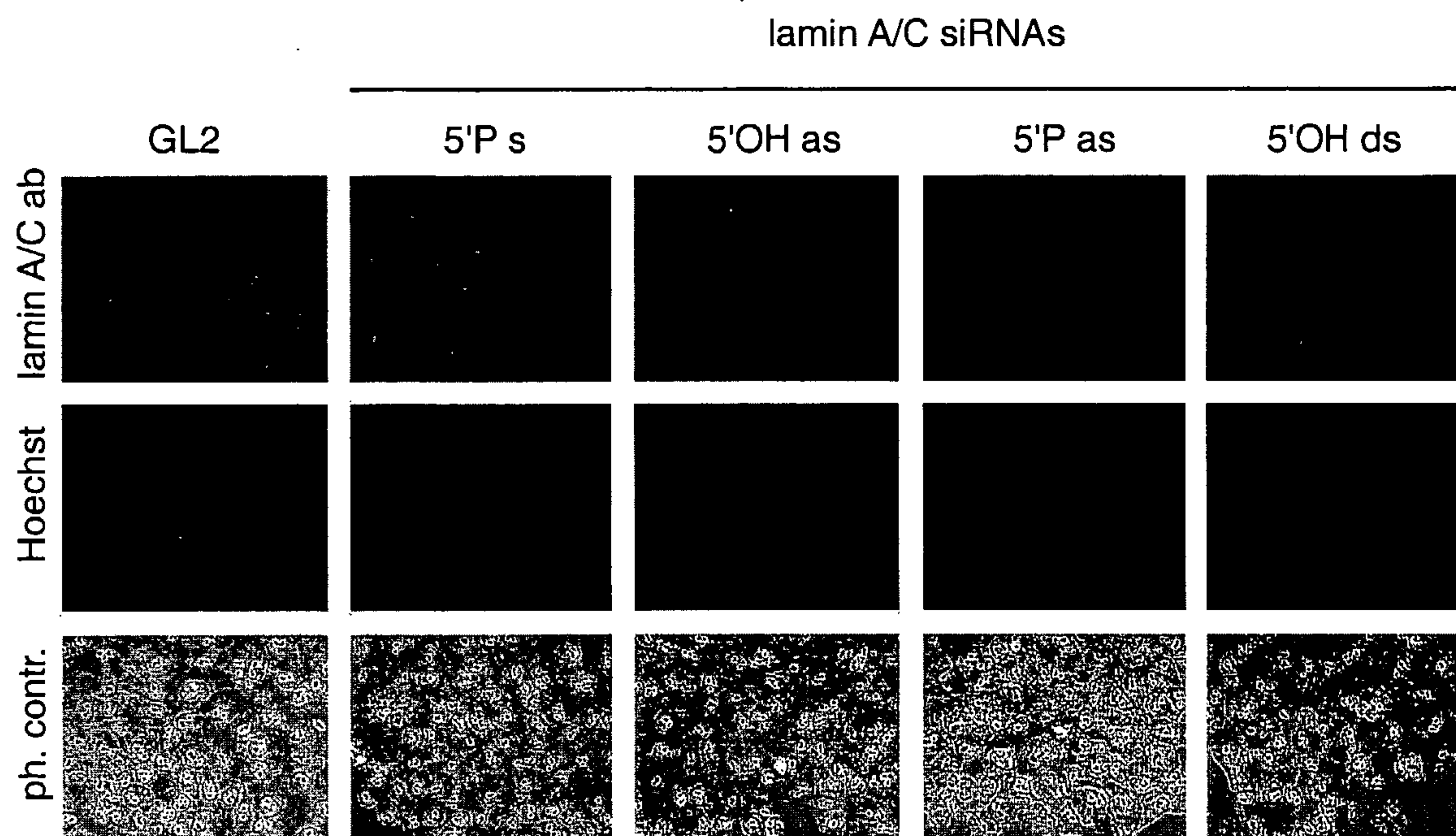
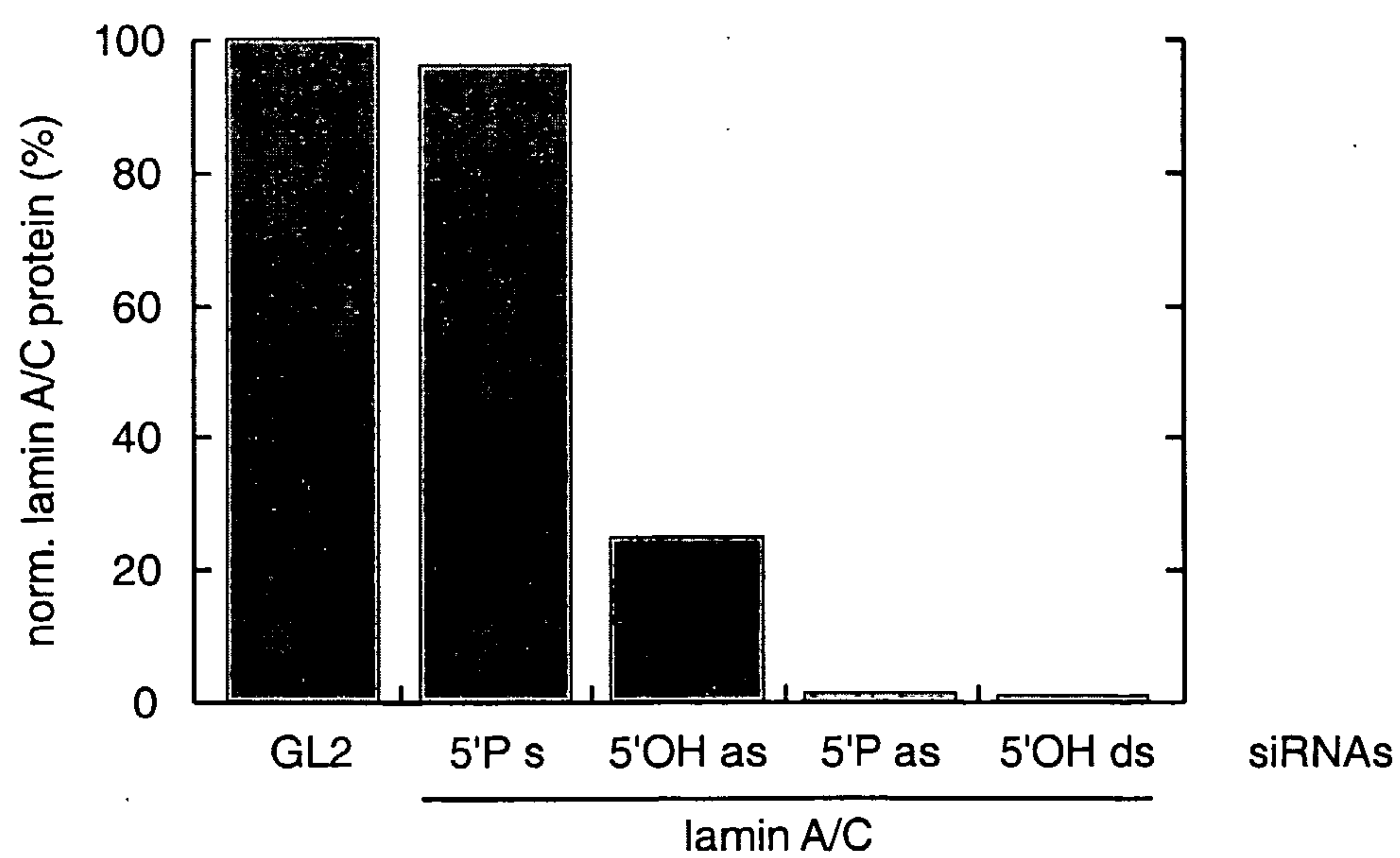
Drosophila embryo

c (nM)	100			10			1			
siRNA	T1	s	as	ds	s	as	ds	s	as	ds



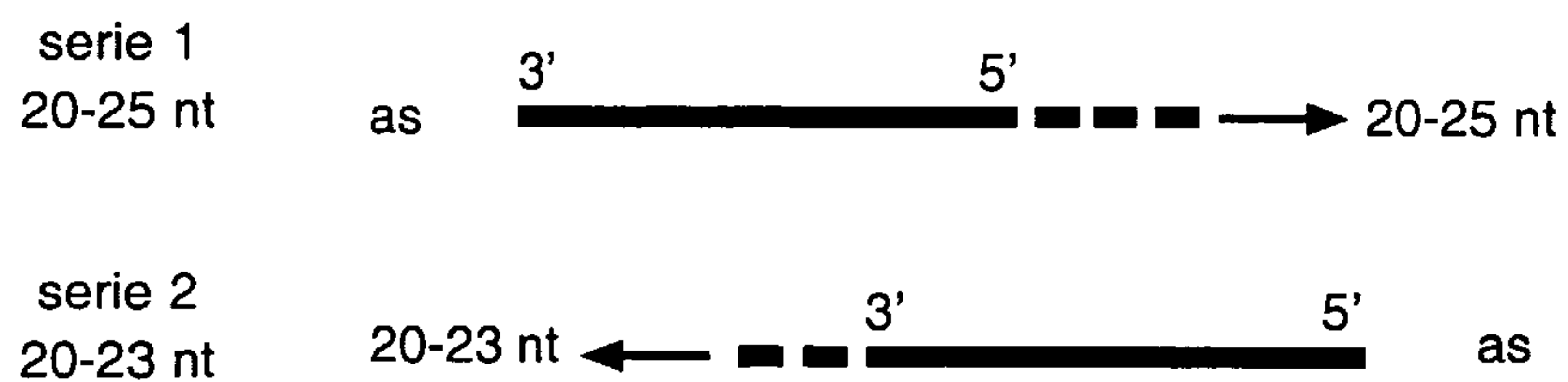
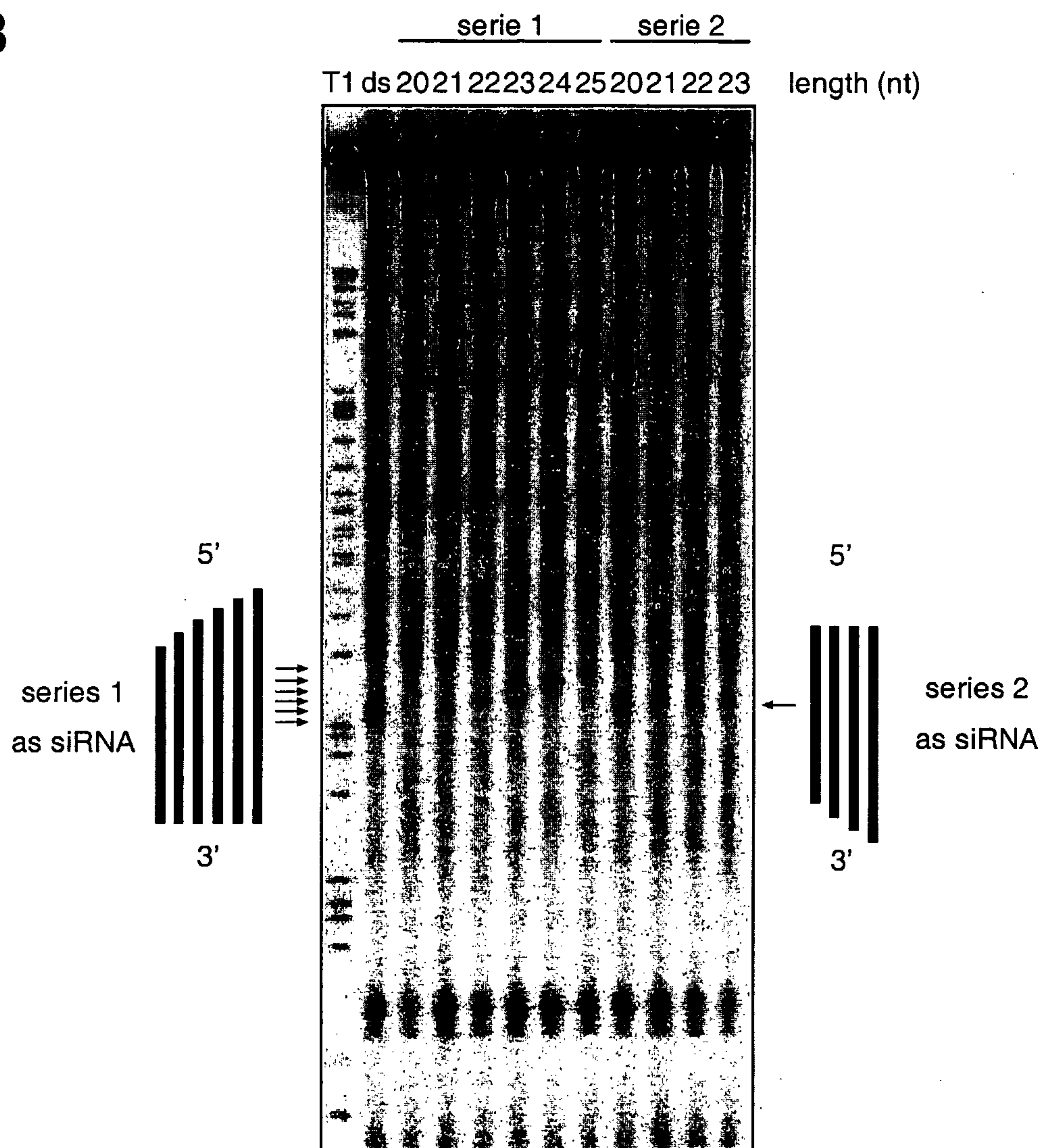
8/25

Martinez et al. Figure 8

A**B**

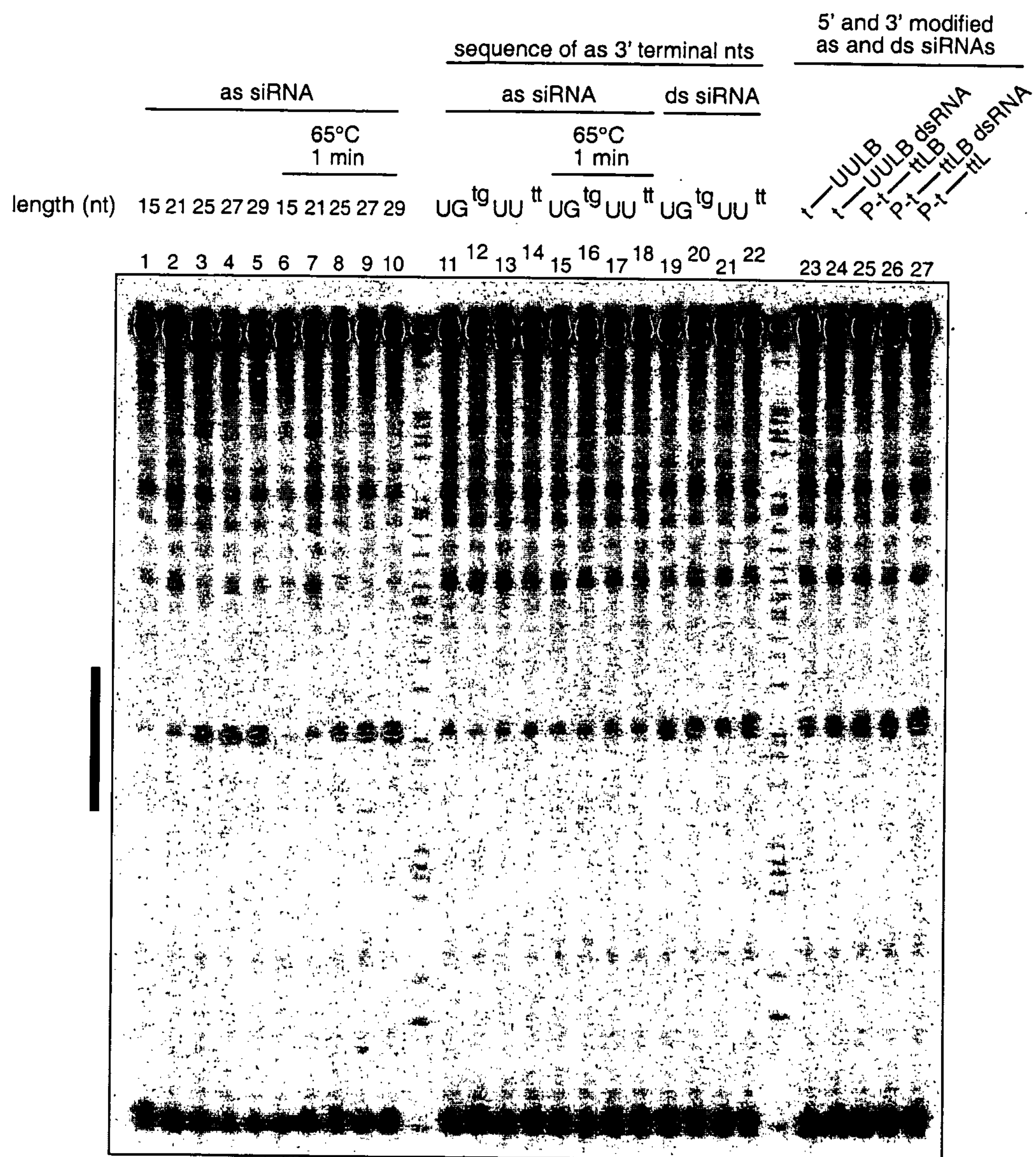
9/25

Martinez et al., Figure 9

A**B**

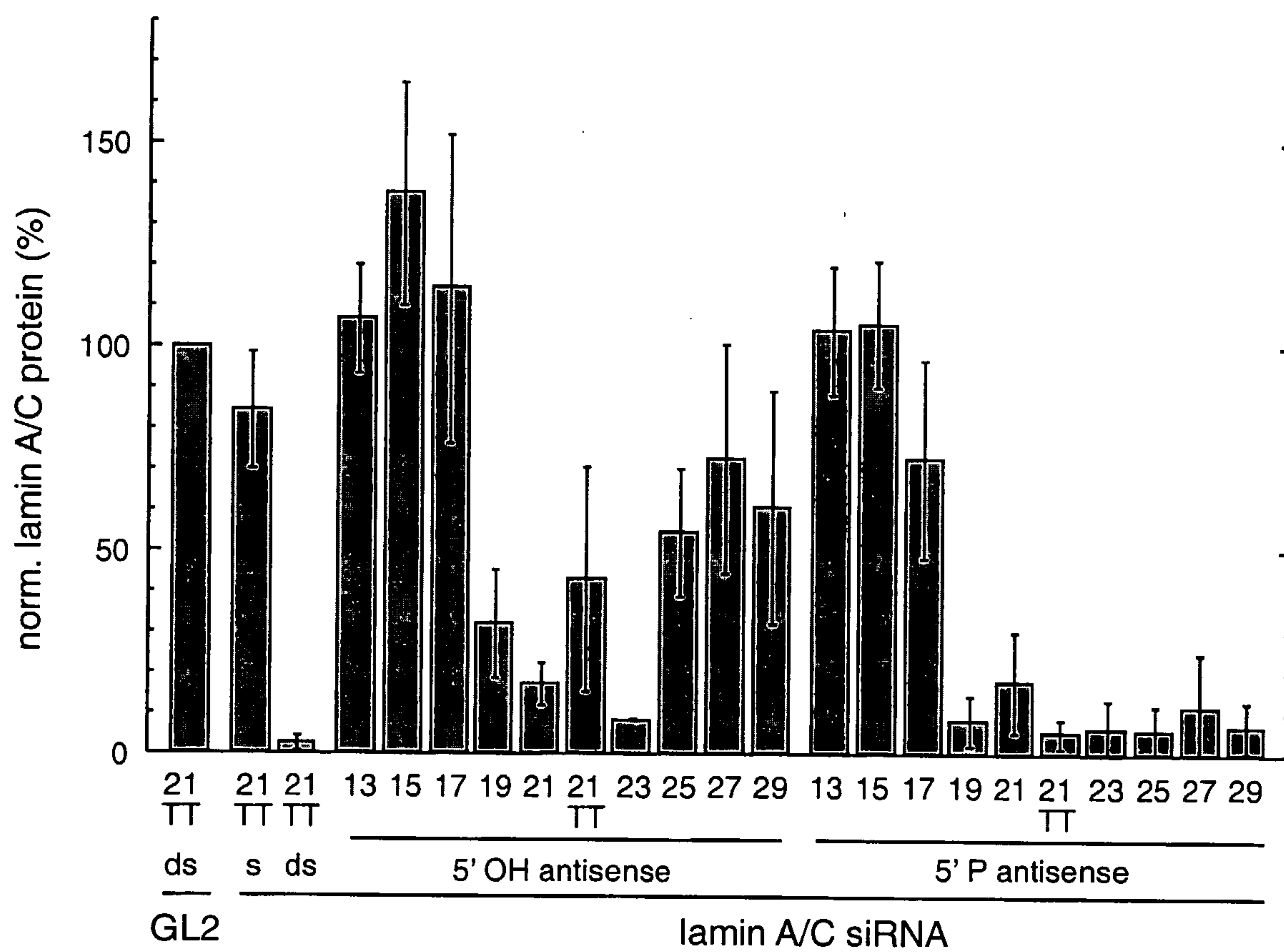
10/25

Martinez et al., Figure 10



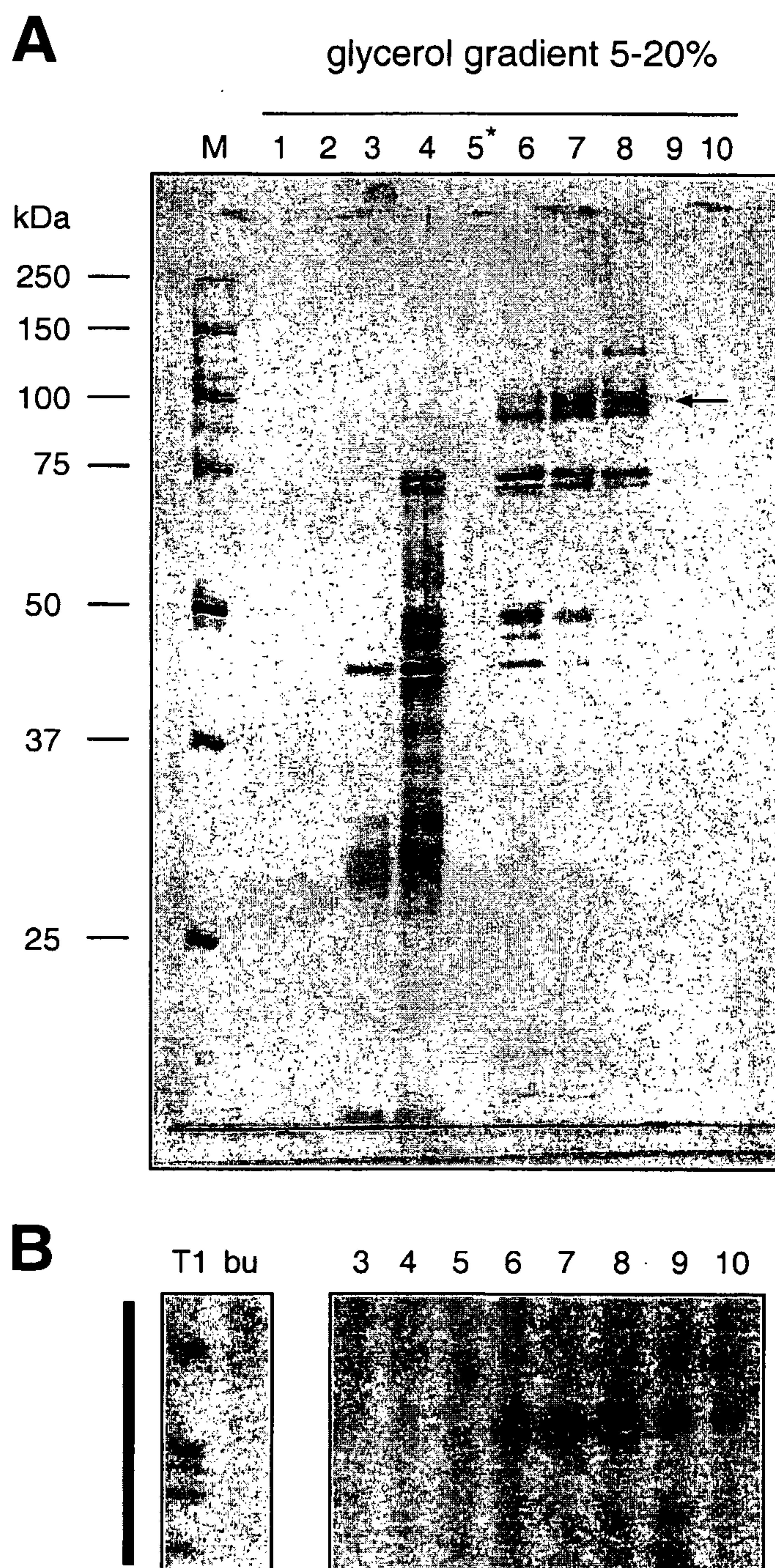
11/25

Martinez et al. Figure 11



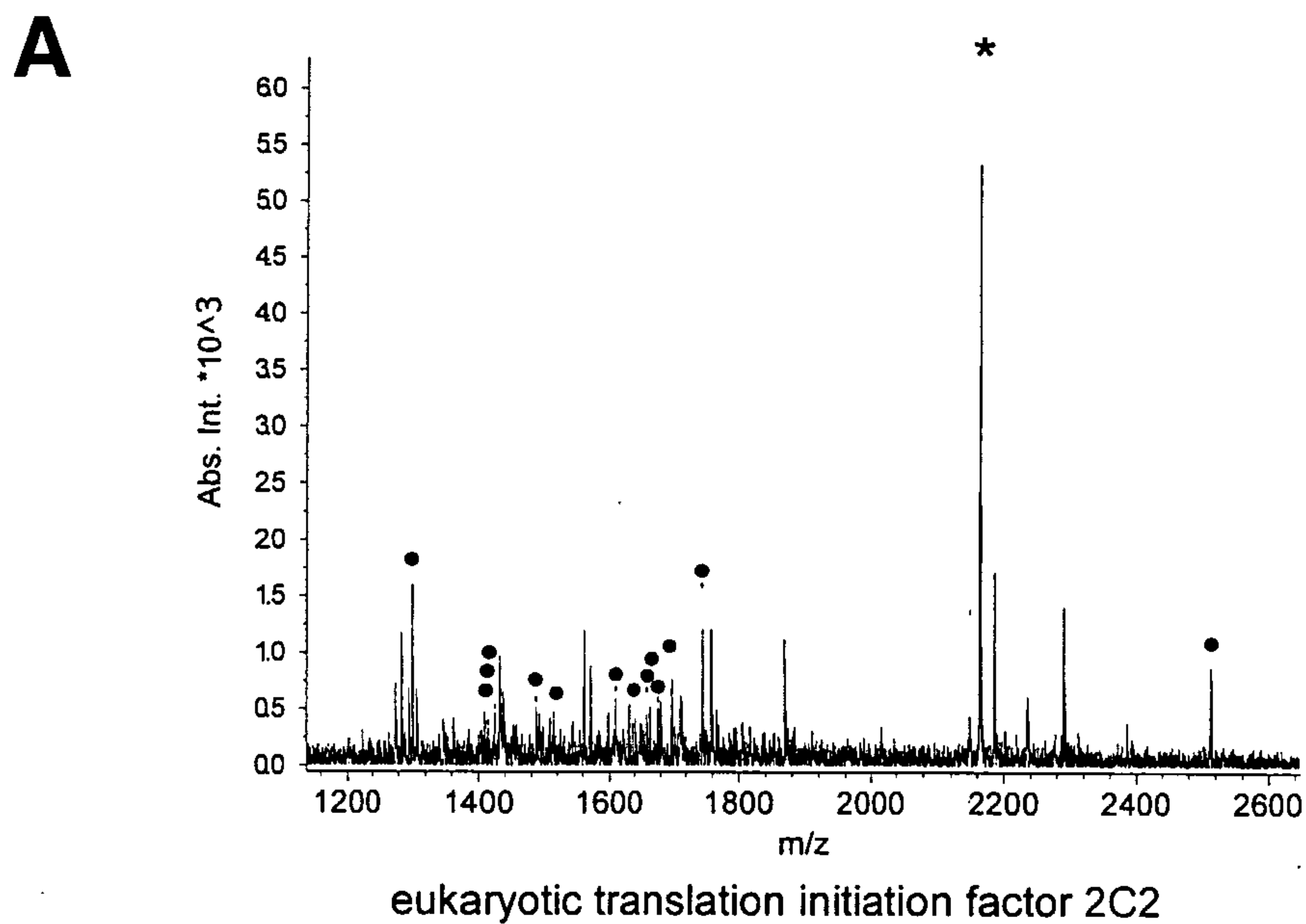
12/25

Martinez et al. Figure 12



13/25

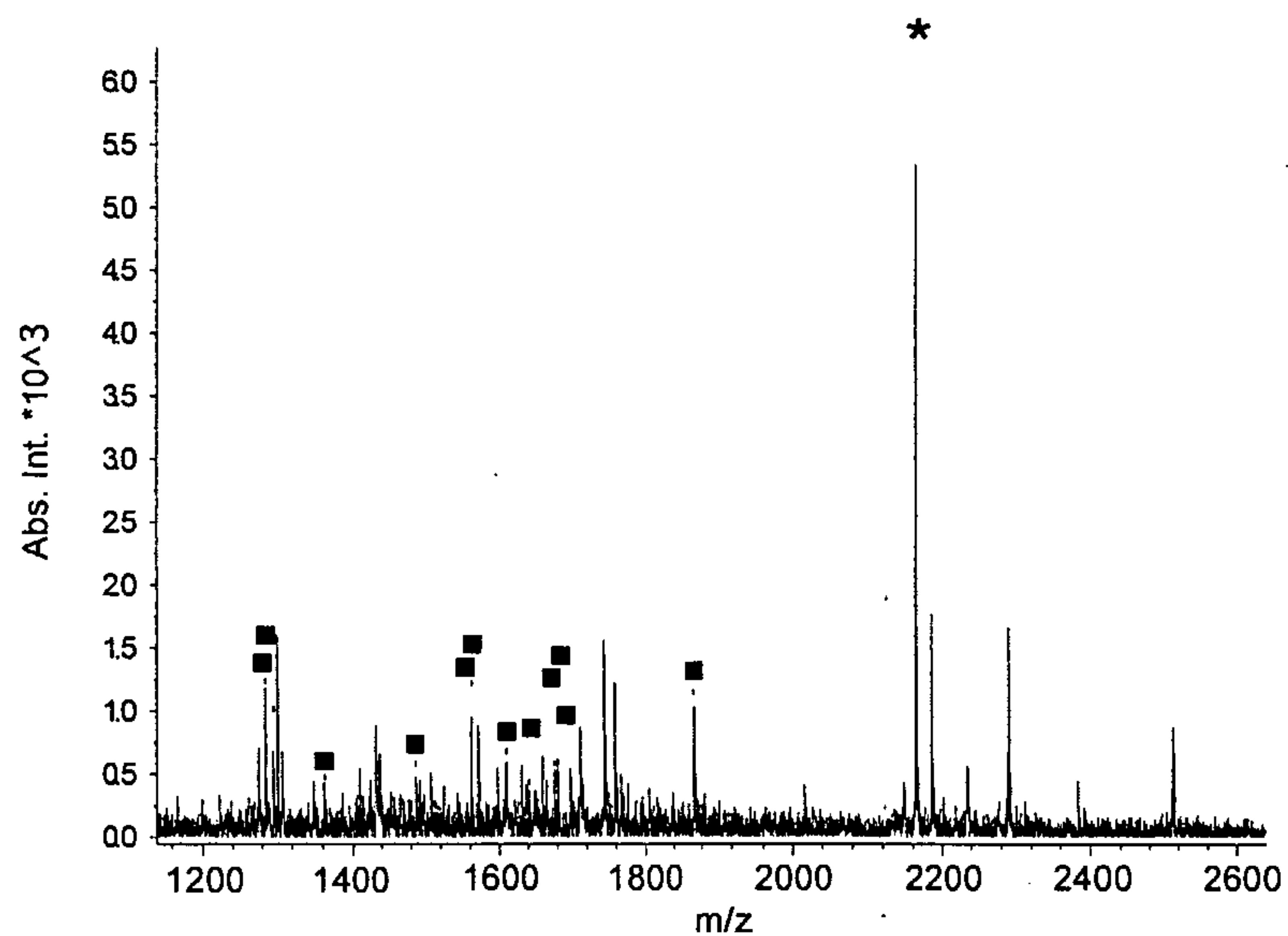
Martinez et al. Figure 13 A



Observed	Mr (expt)	Mr (calc)	Delta	Position	Miss	Peptide
1299.67	1298.67	1298.73	-0.07	413 - 424	0	VLQPPSILYGGR
1402.64	1401.64	1401.74	-0.10	637 - 648	0	QEIIQDLAAMVR Oxidation(M)
1413.62	1412.61	1412.73	-0.12	169 - 180	1	HLPSMRYTPVGR
1423.60	1422.59	1422.71	-0.12	356 - 367	1	KLTDNQTSTMIR Oxidation(M)
1486.56	1485.56	1485.66	-0.10	495 - 507	0	YAQGADSVEPMFR Oxidation(M)
1513.71	1512.70	1512.80	-0.10	112 - 125	1	DKVELEVTLPGEGK
1608.67	1607.66	1607.69	-0.03	481 - 494	0	DAGMPIQQQPCFCK
1635.84	1634.83	1634.85	-0.02	85 - 98	1	TQIFGDRKPVFDGR
1658.85	1657.85	1657.84	0.01	368 - 382	2	ATARSAPDRQEEISK
1663.85	1662.85	1662.91	-0.06	698 - 711	1	DYQPGITFIVVQKR
1675.79	1674.78	1674.84	-0.06	372 - 385	2	SAPDRQEEISKLMR Oxidation(M)
1696.77	1695.76	1695.84	-0.08	323 - 336	0	YPHLPCLQVGQEQK
1743.75	1742.74	1742.77	-0.03	181 - 197	0	SFFTASEGCSNPLGGGR
2511.07	2510.06	2510.12	-0.05	816 - 838	1	YHLVDKEHDSAEGSHTSGQSNR

14/25

Martinez et al. Figure 13 B

B

eukaryotic translation initiation factor 2C1

Observed	Mr (expt)	Mr (calc)	Delta	Position	Miss	Peptide
1283.66	1282.65	1282.74	-0.09	410 - 421	0	VLPAPILQYGGR
1294.65	1293.64	1293.67	-0.03	794 - 805	0	SVSIPAPAYYAR
1361.61	1360.60	1360.70	-0.10	553 - 564	0	TSPQTLNLCLK
1486.56	1485.56	1485.66	-0.10	492 - 504	0	YAQGADSVEPMFR Oxidation(M)
1560.76	1559.75	1559.83	-0.08	97 - 110	0	NIYTVTALPIGNER
1561.76	1560.75	1560.78	-0.02	111 - 124	1	VDFEVTIPGEGKDR
1608.67	1607.66	1607.69	-0.03	478 - 491	0	DAGMPIQGQPCFCK
1640.74	1639.73	1639.82	-0.08	240 - 253	0	NIDEQPKPLTDSQR
1675.79	1674.78	1674.84	-0.06	369 - 382	2	SAPDRQEEISRMLK Oxidation(M)
1679.86	1678.85	1678.90	-0.05	695 - 708	1	DYQPGITYIVVQKR
1696.77	1695.76	1695.84	-0.08	320 - 333	0	YPHLPCLQVGQEQK
1867.85	1866.85	1866.87	-0.02	178 - 194	0	SFFSPPEGYYHPLGGGR

15/25

Martinez et al. Figure 13 C

C

■ eIF2C2 peptides ■ eIF2C1 peptides — PAZ domain
■ oxidized ■ oxidized — PIWI domain

■ eIF2C2 MGVLSTAFALAPFAPPPPTQGYAFKPPRPDPTGSGRTIKLQANTFMDIPKIDTYHYELDIPKPCPRRVNREIVEHMQHFKTQIFGDRKPVEDGKKN 100
■ eIF2C1 --MEAGPSGAAGAYLPPFQQVFOAPRRPGIGTVGKFIKLLANTFEVDIPKIDVYHYEVDIPKPCPRRVNREIVEHMQHFKTQIFGDRKPVEDGKKN 97

1.....10.....20.....30.....40.....50.....60.....70.....80.....90.....100

■ eIF2C2 NYTAMPLPIGRDVEVEVTTPGEGKDRIKFSIKWVSCVSEQALHIALSGRIPEVPPETQALDVVMRHLPSMRYTPVGRSFTTASGQCSNPLGGGREVW 200
■ eIF2C1 NYTVTALPIGNERVDFEVTTPGEGKDRIKFSIKWVSCVSEQALHIALSGRIPEVPPETQALDVVMRHLPSMRYTPVGRSFTTASGQCSNPLGGGREVW 197

.....110.....120.....130.....140.....150.....160.....170.....180.....190.....200

■ eIF2C2 FGFHQSVRPSLWKMMLNIDVSATAFYKAQPVIEFVCEVLDIFKSIHQKPLTDSQRVETKEIKGLKVEVTHCGQMKRKYRVCHVTRRPASHQTFFPLQQE 300
■ eIF2C1 FGFHQSVRPSLWKMMLNIDVSATAFYKAQPVIEFVCEVLDIFKSIHQKPLTDSQRVETKEIKGLKVEVTHCGQMKRKYRVCHVTRRPASHQTFFPLQQE 297

.....210.....220.....230.....240.....250.....260.....270.....280.....290.....300

■ eIF2C2 SGQTVECTVAQYFKDRHKLVLRYPHLPCLQVQGEQKHTYLPLEVCNIVAGQRCIKKLTNDQTSTMIKATARSAPDRQEEISKLMKASYNTPYVREFGI 400
■ eIF2C1 SGQTVECTVAQYFKDRHKLVLRYPHLPCLQVQGEQKHTYLPLEVCNIVAGQRCIKKLTNDQTSTMIKATARSAPDRQEEISKLMKASYNTPYVREFGI 397

.....310.....320.....330.....340.....350.....360.....370.....380.....390.....400

■ eIF2C2 MVKDEMTVTGRVLQEPSTHYGGRRKAIATPVQGVWDMRKQFHYGIEIKVWAIACFAPQKQCTEVHLKSTTQQLRKISKDAGMPIQQQPCFKYAQQAD 500
■ eIF2C1 MVKDDMTVTGRVLQEPSTHYGGRRKAIATPVQGVWDMRKQFHYGIEIKVWAIACFAPQKQCTEVHLKSTTQQLRKISKDAGMPIQQQPCFKYAQQAD 497

.....410.....420.....430.....440.....450.....460.....470.....480.....490.....500

■ eIF2C2 SVFPMFRHLKNTYAGLQLVYVILPGKTPVYAEVVRVGDITLGMATQCQVQKKNVQRTTPQTLNLCCLKINVKLGQVNNILVPHQSRSAVFQQPVIFLGADVT 600
■ eIF2C1 SVFPMFRHLKNTYAGLQLVYVILPGKTPVYAEVVRVGDITLGMATQCQVQKKNVQRTTPQTLNLCCLKINVKLGQVNNILVPHQSRSAVFQQPVIFLGADVT 597

.....510.....520.....530.....540.....550.....560.....570.....580.....590.....600

■ eIF2C2 HPPAGDGKKPSIYAVVGSMDAHPNRYCATVRVQQRQEEIIDLSYVRELLIQFYKSTRFKPTRIIFYRDGVSEGOQVQLHHELLAIREACIKLEKDYQ 700
■ eIF2C1 HPPAGDGKKPSIYAVVGSMDAHPNRYCATVRVQQRQEEIIDLSYVRELLIQFYKSTRFKPTRIIFYRDGVSEGOQVQLHHELLAIREACIKLEKDYQ 697

.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700

■ eIF2C2 PGITYIVVQKRHHTRLFCADKNERVQKSGNIPAGTTVDTHITHPYFDFYLCSHAGIQGTSRPSHYVWLWDDNRFTHDELQILTYQLCHTYVRCSTRSVSI 800
■ eIF2C1 PGITYIVVQKRHHTRLFCADKNERVQKSGNIPAGTTVDTHITHPYFDFYLCSHAGIQGTSRPSHYVWLWDDNRFTHDELQILTYQLCHTYVRCSTRSVSI 797

.....710.....720.....730.....740.....750.....760.....770.....780.....790.....800

■ eIF2C2 PAPAYYALVAFRARYHLVDKEHDSLEGSHSQGSNGRDHQAALAKAVQVHQDTLRTMYFA 860
■ eIF2C1 PAPAYYALVAFRARYHLVDKEHDSLEGSHSQGSNGRDHQAALAKAVQVHQDTLRTMYFA 857

.....810.....820.....830.....840.....850.....860

16/25

Martinez et al. Figure 14

>eIF2C1, predicted protein sequence

MEAGPSGAAAGAYLPPLQQVFQAPRRPGIGTVGKPIKLLANYFEVDIPKIDVYHYEVDIKPD
 KCPRRVNREVVVEYMVQHFQKQIFGDRKPVYDGGKNIYTVTALPIGNERNVDFEVTIPGEGKDR
 IFKVSIKWLAIVSWRMLHEALVSGQIPVPLESVQALDVAMRHLASMRYTPVGRSFFSPPEGY
 YHPLGGGREVWFGFHQSVRPAMWKMLNIDVSATAFYKAQPVIEFMCEVLDIRNIDEQPKPL
 TDSQVRVFTKEIKGLKVEVTHCGQMKRKYRVCNVTRRPASHQTFPLQLESGQTVECTVAQYF
 KQKYNLQLKYPHLPCLQVGQEQKHTYLPLEVCNIVAGQRCIKKLTNDQTSTMIKATARSAPD
 RQEEISRLMKNASYNLDPYIQEFGIKVKDDMTTEVTGRVLPAPILQYGGNRNRAIATPNQGVWD
 MRGKQFYNGIEIKVWAIACFAPQKQCREEVLKNFTDQLRKISKDAGMPIQGQPCFCKYAQGA
 DSVEPMFRHLKNTYSGLQLIIVILPGKTPVYAEVKRVGDTLLGMATQCVQVKNVVKTSPTL
 SNLCLKINVKLGGINNVLPHQRSVVFQQPVIIFLGADVTHPPAGDGKKPSITAVVGSMDAHP
 SRYCATVRVQRPRQEIIEDLSYMVRELLIQFYKSTRFKPTRIIFYRDGVPEGQLPQILHYEL
 LAIRDACIKLEKDYQPGITYIVVQKRHHTRLFCADKNERIGKSGNIPAGTTVDNITHPFEF
 DFYLCSHAGIQGTSRPSHYVWLWDDNRFADDELQILTYQLCHTYVRCTRSVSIAPAPYYARL
 VAFRARYHLVDKEHDSGEGSHISGQSNGRDPQALAKAVQVHQDTLRTMYFA

>eIF2C2, predicted protein sequence

MGVLSAIPALAPPAPPPPIQGYAFKPPRPDFGTSGRTIKLQANFFEMDIPKIDIYHYELDI
 KPEKCPRRVNREIVEHVMVQHFQKQIFGDRKPVFDGRKNLYTAMPLPIGRDKVELEVTLPGEG
 KDRIFKVSIKWVSCVSLQALHDALSGRLPSVPFETIQALDVVMRHLPSMRYTPVGRSFFTAS
 EGCSNPLGGGREVWFGFHQSVRPSLWKMLNIDVSATAFYKAQPVIEFVCEVLDFKSIIEEQQ
 KPLTDSQVRVKFTKEIKGLKVEITHCGQMKRKYRVCNVTRRPASHQTFPLQQESGQTVECTVA
 QYFKDRHKLVLRYPHLPCLQVGQEQKHTYLPLEVCNIVAGQRCIKKLTNDQTSTMIRATARS
 APDRQEEISKLMRSASFNTDPYVREFGIMVKDEMTDVTGRVLQPPSILYGGNRNKAIAIPVQG
 VWDMRNKQFHTGIEIKVWAIACFAPQROCTEVHLKSFTQELRKISRDAAGMPIQGQPCFCKYA
 QGADSVPEPMFRHLKNTYAGLQLVVILPGKTPVYAEVKRVGDTVLMATQCVQMKNVQRTTP
 QTLNLCLKINVKLGGINNILLPQGRPPVFQQPVIIFLGADVTHPPAGDGKKPSIAAVVGSMD
 AHPNRYCATVRVQQHRQEIIQDLAAMVRELLIQFYKSTRFKPTRIIFYRDGVSEGQFQQVLH
 HELLAIREACIKLEKDYQPGITFIVVQKRHHTRLFCTDKNERVGKSGNIPAGTTVDTKITHP
 TEFDFYLCSHAGIQGTSRPSHYVWLWDDNRFSSDELQILTYQLCHTYVRCTRSVSIAPAPYY
 AHLVAFRARYHLVDKEHDSAEGSHTSGQSNGRDHQALAKAVQVHQDTLRTMYFA

>eIF2C3, predicted protein sequence

SRSRVPVPGPGAAAAPCPAPASPRRHPSANIPEIKRYAAAAAAGPGAGGAGDRGEAAPAA
 AMEALGPGPPASLFQPPRRPGLGTVGKPIRLLANHFQVQIPKIDVYHYDVDIKPEKRPRRVN
 REVVDTMVRHFQKQIFGDRQPGYDGKRNMYTAHPLPIGRDRVDMVTLPGEGKDQTFKVSQ
 WVSVSLQLLLEALAGHLNEVPDDSVQALDVITRHLPSMRYTPVGRSFFSPPEGYYHPLGGG
 REVWFGFHQSVRPAMWNMMLNIDVSATAFYRAQPIIEFMCEVLDIQNINEQTKPLTDSQVRK
 FTKEIRGLKVEVTHCGQMKRKYRVCNVTRRPASHQTFPLQLENGQAMECTVAQYFKQKYSQ
 LKYPHLPCLQVGQEQKHTYLPLEVCNIVAGQRCIKKLTNDQTSTMIKATARSAPDRQEEISR
 LVKSNSMVGPDYPYLKEFGIVVHNEMTELTGRVLPAPMLQYGGNRNKTVAIPNQGVDWMRGKQ
 FYAGIEIKVWAVACFAPQKQCREDLLKSFTDQLRKISKDAGMPIQGQPCFCKYAQGADSVPEP
 MFKHLKMTYVGLQLIVVILPGKTPVYAEVKRVGDTLLGMATQCVQVKNVVKTSPTLNLCL
 KINAKLGGINNVLPHQRPVSFQQPVIIFLGADVTHPPAGDGKKPSIAAVVGSMDGHPSTRYCA
 TVRVQTSRQEISQELLYSQEVIQDLTNMVRELLIQFYKSTRFKPTRIIYYRGGVSEGQMKQV
 AWPELIAIRKACISLEEDYRPGITYIVVQKRHHTRLFCADKTERVGKSGNVIPAGTTVDSTIT
 HPSEFDFYLCSHAGIQGTSRPSHYQVLWDDNCFADDELQLLTYQLCHTYVRCTRSVSIAPAPA
 YYARLVAFRARYHLVDKDHDHSAEGSHVSGQSNGRDPQALAKAVQIHHDTHQHTMYFA

Page 1 of 2

17/25

Martinez et al. Figure 14

>eIF2C4, predicted protein sequence

AGPAGAQPPLLMVPRRPGYGTMGKPIKLLANCFQVEIPKIDVYLYEVDIKPDKCPRRVNREV
 DSMVQHFKVTIFGDRRPVYDGRSLYTANPLPVATTGVDLDVTLPGEGGKDRPFKVSIFVS
 RVSWHLLHEVLTGRTLPEPLELDKPISTNPVHAVDVLRHLPSMKYTPVGRSFFSAPEGYDH
 PLGGGREVWFGFHQSVRPAMWKMLNIDVSATAFYKAQPVIQFMCEVLDIHNIDEQPRPLTD
 SHRVKFTKEIKGLKVEVTHCGTMRKRYVCNVTRRPASHQTFPLQLENGQTVERTVAQYFRE
 KYTLQLKYPHLPCLQVGQEQKHTYLPLEV CNIVAGQRCIKKLTNDQTSTMIKATARSAPDRQ
 EEISRLVRSANYETDPFVQEFQFKVRDEMAHVTGRVLPAPMLQYGGRNRTVATPSHGVWDMR
 GKQFHTGVEIKMWAIACFATQRQCREEILKGFTDQLRKISKDAGMPIQGGQPCFCKYAQGADS
 VEPMFRHLKNTYSGLQLIIVILPGKTPVYAEVKRVGDTLLGMATQCVQVKNVIKTSPQTLN
 LCLKINVKLGGINNVLPHQRPSVFFQPVIFLGADVTHPPAGDGKKPSIAAVVGSMDAHP
 YCATVRVQRPRQEIIQDLASRVRELLIQFYKSTRFKPTRIIFYRDGVSEGQFRQVLYYELLA
 IREACISLEKDYQPGITYIVVQKRHHTRLFCADRTERVGRSGNI PAGTTVDTDITHPYEFD
 YLCSHAGIQGTSRPSHYHVLWDDNCFTADELQLLTYQLCHTYVRCTRSVSI PAPAYY AHLVA
 FRARYHLVDKEHDSAEGSHVSGQSNRDPQALAKAVQIHQDTLRTMYFA

>HILI, predicted protein sequence

ISSGDAGSTFMERGVKNKQDFMDLSICTREKLAHVNRCKTGSSGIPVKLVNTLNFNLDFPQDW
 QLYQYHVTYIPDLASRLRLRIALLYSHSELSNKAFAFDGAILFLSQKLEEKVTELSSETQGE
 TIKMTITLKRELPSPPVCIQVFNIIFRKILKKLSMYQIGRNFYNPSEPMEIPQHKLSLWPG
 FAISVSYFERKLLFSADVSYKVLNRNETVLEFMTALCQRTGLSCFTQTCEKQLIGLIVLTRYN
 NRTYSIDDDIDWSVKPTHFTFQKRDGTEITYVDYKQQYDITVSDLNQPM LVSLKKRNDNSE
 AQLAHLIPELCFLTGLTDQATSDFQLMKAVAEKTRLSPSGRQQLARLVDNIQRNTNARFEL
 ETWGLHFGSQISLTGRIVPSEKILMQDHICQPVSAADWSKDIRTCKILNAQSLNTWLILCSD
 RTEYVAESFLNCLRRVAGSMGFNVMCILPSNQKTYYSIKKYLSSDCPVPSQCVLARTLNKQ
 GMMMSIATKIAMQMTCKLGELWAVEIPLKSLMVVGIDVCKDALSKDVMVGCVASVNPRIT
 RWF SRCILQRTMTDVADCLKVFM TGALNKWYKYNHDL PARIIVYRAGVGDGQLKTLIEYEV
 QLLSSVAESSNTSSRLSVIVVRKKCMRPF FTEMNRTVQNPPLGTVVDSEATRNEWQYDFYL
 ISQVACRGTVSPTYYNVIYDDNGLKPDHMQR LTFKLCHLYYNWPGIVSV PAPCQYAHKLTF
 VAQSIHKEPSLELANHLFYL

>HIWI, predicted protein sequence

MTGRARARARGRARGQETAQLVGSTASQQPGYIQPRPQPPPAEGELFGRGRQRGTAGGTAKS
 QGLQISAGFQELSLAERGGRRRDFHDLGVNTRQNL DHVKESKTGSSGIIIVRLSTNHFRLTSR
 PQWALYQYHIDYNPLMEARRLRSALLFQHEDLIGKCHA FDGTILFLPKRLQQKVTEVFSKTR
 NGEDVRITITLTNELPPTSPTCLQFYNIIFRRLKIMNLQQIGRNYNPNDPIDIPSHRLVI
 WPGFTTSILQYENSIMLCTDVSHKVL RSETVLD FMFNFYHQTEEHKFQE QVSKELIGLVVLT
 KYN NKTYRVDDIDWDQNP KSTFKKADGSEVS FLEYRKYQNQEITDLKQPV LV SQPKRRRGP
 GGTLP GPAMLIPELCYLTGLTDKMRNDFNVMKDLAVHTRLTPEQRQREVGRLIDYIHKNDNV
 QREL RDWGLSFDSNLLSFSGRILQTEKIHQGGKTFDYNPQFADWSKETRGAPLISVKPLDNW
 LLIYTRRNYEAANS LIQNLFKVTPAMGMQMRKAIMIEVDDRTEAYLRVLQQKV TADTQIVVC
 LLSSNRKDKYDAIKKYLCTDCPTPSQCVVARTLGKQQTVMATKIALQMNCKMGGELWRVD
 IPLKLVMI VGIDCYHDMTAGRRS IAGFVASINEGMTRWFSRCIFQDRGQELVDGLKVCLQAA
 LRAWNSCNEYMP SR IIVYRDGVGDGQLKTLVNYEVPQFLDCLKSIGRGYNPRLTVIVVKRV
 NTRFFAQSGGRLQNPLPGTVIDVEVTRPEWYDFFIVSQAVRSGSVSPTHYNVIYDNSGLKPD
 HIQRLTYKLCHIYYNWPGVIRVPAPCQYAHKLAF LVGQSIHREP NLSLSNRLYYL

1	100	191	282	373	474	573	671	771	870	924
2	34	126	225	325	415	514	612	702	801	855
3	45	136	227	327	417	516	614	704	803	857
4	48	139	230	330	420	519	617	707	806	860
5	24	124	213	298	396	495	591	641	737	764
6	89	189	278	363	463	557	649	739	834	861
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										
29										
30										
31										
32										
33										
34										
35										
36										
37										
38										
39										
40										
41										
42										
43										
44										
45										
46										
47										
48										
49										
50										
51										
52										
53										
54										
55										
56										
57										
58										
59										
60										

19/25

Martinez et al. Figure 16

>eIF2C1, cDNA sequence of predicted ORF

ATGGAAGCGGGACCCCTCGGGAGCAGCTGCGGGCGCTTACCTGCCCCCCTGCAGCAGGTGTT
CCAGGCACCTCGCCGGCCTGGCATTGGCACTGTGGGGAAACCAATCAAGCTCCTGGCCAATT
ACTTTGAGGTGGACATCCCTAAGATCGACGTGTACCACTACGAGGTGGACATCAAGCCGGAT
AAGTGTCCCCGTAGAGTCAACCGGGAAGTGGTGAATACATGGTCCAGCATTTCAAGCCTCA
GATCTTTGGTGATCGCAAGCCTGTGTATGATGGAAAGAAGAACATTTACACTGTGCACAGCAC
TGCCCATTTGGCAACGAACGGGTCGACTTTGAGGTGACAATCCCTGGGGAAGGGAAGGATCGA
ATCTTTAAGGTCTCCATCAAGTGGCTAGCCATTGTGAGCTGGCGAATGCTGCATGAGGCCCT
GGTCAGCGGCCAGATCCCTGTTCCCTTGGAGTCTGTGCAAGCCCTGGATGTGGCCATGAGGC
ACCTGGCATCCATGAGGTACACCCCTGTGGGCCGCTCCTTCTTCTCACC GCCTGAGGGCTAC
TACCACCCGCTGGGGGGTGGGCGCGAGGTCTGGTTTCGGCTTTTACCAGTCTGTGCGCCCTGC
CATGTGGAAGATGATGCTCAACATTGATGTCTCAGCCACTGCCTTTTATAAGGCACAGCCAG
TGATTGAGTTCATGTGTGAGGTGCTGGACATCAGGAACATAGATGAGCAGCCCAAGCCCCCTC
ACGGACTCTCAGCGCGTTTCGCTTCACCAAGGAGATCAAGGGCCTGAAGGTGGAAGTCACCCA
CTGTGGACAGATGAAGAGGAAGTACCGCGTGTGTAATGTTACCCGTCGCCCTGCTAGCCATC
AGACATTTCCCTTACAGCTGGAGAGTGGACAGACTGTGGAGTGCACAGTGGCACAGTATTTT
AAGCAGAAATATAACCTTCAGCTCAAGTATCCCCATCTGCCCTGCCTACAAGTTGGCCAGGA
ACAAAAGCATACTACCTTCCCTTAGAGGTCTGTAACATTGTGGCTGGGCAGCGCTGTATTA
AAAAGCTGACCGACAACCAGACCTCGACCATGATAAAGGCCACAGCTAGATCCGCTCCAGAC
AGACAGGAGGAGATCAGTCGCCTGATGAAGAATGCCAGCTACAACCTAGATCCCTACATCCA
GGAATTTGGGATCAAAGTGAAGGATGACATGACGGAGGTGACAGGGCGAGTGCTGCCGGCGC
CCATCTTGACAGTACGGCGGCCGGAACCGGGCCATTGCCACACCCAATCAGGGTGTCTGGGAC
ATGCGGGGGAAACAGTTCTACAATGGGATTGAGATCAAAGTCTGGGCCATCGCCTGCTTCGC
ACCCCAAAAACAGTGTGAGAAAGAGGTGCTCAAGAACTTCACAGACCAGCTGCGGAAGATTT
CCAAGGATGCGGGGATGCCTATCCAGGGTCAACCTTGTCTTCTGCAAATATGCACAGGGGGCA
GACAGCGTGGAGCCTATGTTCCGGCATCTCAAGAACACCTACTCAGGGCTGCAGCTCATAT
TGTCATCCTGCCAGGGAAGACGCCGGTGTATGCTGAGGTGAAACGTGTCCGAGATACACTCT
TGGGAATGGCTACGCAGTGTGTGCAGGTGAAGAACGTGGTCAAGACCTCACCTCAGACTCTG
TCCAACCTCTGCCTCAAGATCAATGTCAAACCTTGGTGGCATTAACAACATCCTAGTCCCACA
CCAGCGCTCTGCCGTTTTTCAACAGCCAGTGATATTCCTGGGAGCAGATGTTACACACCCCC
CAGCAGGGGATGGGAAAAAACCTTCTATCACAGCAGTGGTAGGCAGTATGGATGCCCCACCCC
AGCCGATACTGTGCTACTGTGCGGGTACAGCGACCACGGCAAGAGATCATTGAAGACTTGTC
CTACATGGTGCCTGAGCTCCTCATCCAATTCTACAAGTCCACCCGTTTCAAGCCTACCCGCA
TCATCTTCTACCGAGATGGGGTGCCTGAAGGCCAGCTACCCAGATACTCCACTATGAGCTA
CTGGCCATTCGTGATGCCTGCATCAAACCTGGAAAAGGACTACCAGCCTGGGATCACTTATAT
TGTGGTGCAGAAACGCCATCACACCCGCCTTTTCTGTGCTGACAAGAATGAGCGAATTGGGA
AGAGTGGTAACATCCCAGCTGGGACCACAGTGGACACCAACATCACCCACCCATTTGAGTTT
GACTTCTATCTGTGCAGCCACGCAGGCATCCAGGGCACCAGCCGACCATCCCATTACTATGT
TCTTTGGGATGACAACCGTTTTACAGCAGATGAGCTCCAGATCCTGACGTACCAGCTGTGCC
ACACTTACGTACGATGCACACGCTCTGTCTCTATCCCAGCACCTGCCTACTATGCCCCGCTG
GTGGCTTTCCGGGCACGATACCACCTGGTGGACAAGGAGCATGACAGTGGAGAGGGGAGCCA
CATATCGGGGCAGAGCAATGGGCGGGACCCCCAGGCCCTGGCCAAAGCCGTGCAGGTTCACC
AGGATACTCTGCGCACCATGTACTTCGCT

20/25

Martinez et al. Figure 16

>eIF2C2, cDNA sequence of predicted ORF

ATGGGTGTTCTCTCTGCCATTCCCGCACTTGCACCTCCTGCGCCGCCGCCCCCATCCAAGG
ATATGCCTTCAAGCCTCCACCTAGACCCGACTTTGGGACCTCCGGGAGAACAATCAAATTAC
AGGCCAATTTCTTCGAAATGGACATCCCCAAAATTGACATCTATCATTATGAATTGGATATC
AAGCCAGAGAAGTGCCCGAGGAGAGTTAACAGGGAAATCGTGGAACACATGGTCCAGCACTT
TAAAACACAGATCTTTGGGGATCGGAAGCCCGTGTTTGACGGCAGGAAGAATCTATACACAG
CCATGCCCCCTCCGATTGGGAGGGACAAGGTGGAGCTGGAGGTCACGCTGCCAGGAGAAGGC
AAGGATCGCATCTTCAAGGTGTCCATCAAGTGGGTGTCCTGCGTGAGCTTGCAGGCGTTACA
CGATGCACTTTCAGGGCGGCTGCCCAGCGTCCCTTTTGAGACGATCCAGGCCCTGGACGTGG
TCATGAGGCACTTGCCATCCATGAGGTACACCCCGTGGGCCGCTCCTTCTTCACCGCGTCC
GAAGGCTGCTCTAACCCTCTTGGCGGGGGCCGAGAAGTGTGGTTTGGCTTCCATCAGTCCGT
CCGGCCTTCTCTCTGGAAAATGATGCTGAATATTGATGTGTTCAGCAACAGCGTTTTACAAGG
CACAGCCAGTAATCGAGTTTGTGTGAAGTTTGGATTTTAAAAGTATTGAAGAACAACAA
AAACCTCTGACAGATTCCCAAAGGGTAAAGTTTACCAAAGAAATTAAAGGTCTAAAGGTGGA
GATAACGCACTGTGGGCAGATGAAGAGGAAGTACCGTGTCTGCAATGTGACCCGGCGGCCCG
CCAGTCACCAAACATTCCCGCTGCAGCAGGAGAGCGGGCAGACGGTGGAGTGCACGGTGGCC
CAGTATTTCAAGGACAGGCACAAGTTGGTTCTGCGCTACCCCCACCTCCCATGTTTACAAGT
CGGACAGGAGCAGAAACACACCTACCTTCCCCTGGAGGTCTGTAACATTGTGGCAGGACAAA
GATGTATTAAAAAATTAACGGACAATCAGACCTCAACCATGATCAGAGCAACTGCTAGGTCTG
GCGCCCGATCGGCAAGAAGAGATTAGCAAATTGATGCGAAGTGCAAGTTTCAACACAGATCC
ATACGTCCGTGAATTTGGAATCATGGTCAAAGATGAGATGACAGACGTGACTGGGCGGGTGC
TGCAGCCGCCCTCCATCCTCTACGGGGGCAGGAATAAAGCTATTGCGACCCCTGTCCAGGGC
GTCTGGGACATGCGGAACAAGCAGTTCCACACGGGCATCGAGATCAAGGTGTGGGCCATTGC
GTGCTTCGCCCCCAGCGCCAGTGCACGGAAGTCCATCTGAAGTCCTTCACAGAGCAGCTCA
GAAAGATCTCGAGAGACGCTGGCATGCCCATCCAGGGCCAGCCGTGCTTCTGCAAATACGCG
CAGGGGGCGGACAGCGTGGAGCCCATGTTCCGGCACCTGAAGAACACGTATGCGGGCCTGCA
GCTGGTGGTGGTCATCCTGCCCGCAAGACGCCCGTGTACGCCGAGGTCAAGCGCGTGGGAG
ACACGGTGTGTTGGGATGGCCACGCAGTGCCTGCAGATGAAGAACGTGCAGAGGACCACGCCA
CAGACCCTGTCCAACCTTTGCCTGAAGATCAACGTCAAGCTGGGAGGCGTGAACAACATCCT
GCTGCCCCAGGGCAGGCCGCCGGTGTTCAGCAGCCCGTCATCTTTCTGGGAGCAGACGTCA
CTACCCCCCGCCGGGGATGGGAAGAAGCCCTCCATTGCCGCGGTGGTGGGCAGCATGGAC
GCCACCCCCAATCGCTACTGCGCCACCGTGCCTGTGCAGCAGCACCGGCAGGAGATCATA
AGACCTGGCCGCCATGGTCCGCGAGCTCCTCATCCAGTTCTACAAGTCCACGCGCTTCAAGC
CCACCCGCATCATCTTCTACCGCGACGGTGTCTCTGAAGGCCAGTTCCAGCAGGTTCTCCAC
CACGAGTTGCTGGCCATCCGTGAGGCCTGTATCAAGCTAGAAAAAGACTACCAGCCCGGGAT
CACCTTCATCGTGGTGCAGAAGAGGCACCACACCCGGCTCTTCTGCACTGACAAGAACGAGC
GGGTTGGGAAAAGTGGAACATTCCAGCAGGCACGACTGTGGACACGAAAATCACCCACCCC
ACCGAGTTCGACTTCTACCTGTGTAGTCACGCTGGCATCCAGGGGACAAGCAGGCCTTCGCA
CTATCACGTCTCTGGGACGACAATCGTTTCTCCTCTGATGAGCTGCAGATCCTAACCTACC
AGCTGTGTACACCTACGTGCGCTGCACACGCTCCGTGTCCATCCCAGCGCCAGCATACTAC
GCTCACCTGGTGGCCTTCCGGGCCAGGTACCACCTGGTGGATAAGGAACATGACAGTGCTGA
AGGAAGCCATACCTCTGGGCAGAGTAACGGGCGAGACCACCAAGCACTGGCCAAGGCGGTCC
AGGTTACCAAGACACTCTGCGCACCATGTACTTTGCT

21/25

Martinez et al. Figure 16

>eIF2C3, cDNA sequence of predicted ORF

AGCCGGAGCCGGGTCCCTGTCCCCGGGCGGGCGCCGCGCCGCCCCCTGCCAGCGCCCGC
GTCTCCGCGGGCGCCACCCAGCGCCAATATTCGGAGATCAAGCGTTACGCGGCGGCGGCGG
CGGCGGCGGCGGGGCCCCGGAGCGGGAGGCGCCGGGGACCGGGGCGAGGCGGCCCCCGCCGCC
GCCATGGAGGCGCTGGGACCCGGACCTCCGGCTAGCCTGTTTCAGCCACCTCGTCGTCCTGG
CCTTGGAACGTGTTGAAAACCAATTCGACTGTTAGCCAATCATTTTCAGGTTTCAATTCTTA
AAATAGATGTGTATCACTATGATGTGGATATTAAGCCTGAAAAACGGCCTCGTAGAGTCAAC
AGGGAGGTAGTAGATACAATGGTGCGGCACTTCAAGATGCAAATATTTGGTGATCGGCAGCC
TGGGTATGATGGCAAAAGAAACATGTACACAGCACATCCACTACCAATTGGACGGGATAGGG
TTGATATGGAGGTGACTCTTCCAGGCGAGGGGTAAAGACCAAACATTTAAAGTGTCTGTTTCAG
TGGGTGTCAGTTGTGAGCCTTCAGTTGCTTTTAGAAGCTTTGGCTGGGCACTTGAATGAAGT
CCCAGATGACTCAGTACAAGCACTTGATGTTATCACAAGACACCTTCCCTCCATGAGGTACA
CCCCAGTGGGCGGTTCTTTTCTCACCCCGGAAGGTTACTACCACCCTCTGGGAGGGGGC
AGGGAGGTCTGGTTTGGTTTTTCATCAGTCTGTGAGACCTGCCATGTGGAATATGATGCTCAA
CATTGATGTATCTGCAACTGCTTTCTACCGGGCTCAGCCTATCATTTGAGTTCATGTGTGAGG
TTTTAGACATTCAGAACATCAATGAACAGACCAAACCTCTAACAGACTCCCAGCGTGTCAA
TTTACCAAAGAAATCAGAGGTCTCAAAGTTGAGGTGACCCACTGTGGACAGATGAAACGAAA
ATACCGAGTTTGTAAATGTGACTAGACGGCCAGCCAGTCATCAAACCTTTTCTTTGTCAGCTAG
AAAACGGTCAAGCTATGGAATGTACAGTAGCTCAATATTTTAAGCAAAGTATAGTCTGCAA
CTGAAATACCCCATCTTCCCTGTCTCCAAGTGGGACAAGAACAAGCATACTACTTGCC
ACTCGAGGTCTGTAATATAGTGGCAGGACAGCGATGTATCAAGAAGCTCACAGACAATCAGA
CTTCCACAATGATCAAAGCTACAGCAAGATCTGCTCCTGACAGACAGGAAGAGATCAGTAGA
CTGGTGAAGAGCAACAGTATGGTGGGTGGACCTGATCCATACCTTAAAGAATTTGGTATTGT
TGTCCACAATGAAATGACAGAGCTCACAGGCAGGGTACTTCCAGCACCAATGCTGCAATATG
GAGGCCGGAATAAAACAGTAGCCACACCCAACCAGGGTGTCTGGGACATGCGAGGAAAGCAG
TTTTATGCTGGCATTGAAATTAAAGTTTGGGCAGTTGCTTGTTTTGCACCTCAGAAACAATG
TAGGGAAGATTTACTAAAGAGTTTCACTGACCAGCTGCGTAAATCTCTAAGGATGCAGGAA
TGCCCATCCAGGGTCAGCCATGTTTCTGCAAGTATGCACAAGGTGCAGACAGTGTGGAGCCT
ATGTTTAAACATCTGAAAATGACTTATGTGGGCCTACAGCTAATAGTGGTTATCCTGCCTGG
AAAGACACCAGTATATGCGGAGGTGAAACGTGTTGGAGATACCCTTCTAGGTATGGCCACAC
AGTGTGTCCAGGTAAAAAATGTAGTGAAGACCTCACCTCAAACCTTTCCAATCTTTGCCTG
AAGATAAATGCAAACTTGGAGGAATTAACAATGTGCTTGTGCCTCATCAAAGGCCCTCGGT
GTTCCAGCAGCCTGTCATCTTCTGAGGCGGATGTCACACACCCCCCAGCAGGGGATGGGA
AGAAACCTTCCATTGCTGCTGTGGTTGGCAGTATGGATGGCCACCCCAGCCGGTACTGTGCC
ACCGTTCGGGTGCAGACTTCCCGGCAGGAGATCTCCAAGAGCTCCTCTACAGTCAAGAGGT
CATCCAGGACCTGACTAACATGGTTCGAGAGCTGCTGATTCAGTTCTACAAATCCACACGCT
TCAAACCCACTCGGATCATCTATTACCGTGGAGGGGTATCTGAGGGACAAATGAAACAGGTA
GCTTGGCCAGAACTAATAGCAATTCGAAAGGCATGTATTAGCTTGAAGAAGATTACCGGCC
AGGAATAACTTATATTGTGGTGCAAAAAAGACATCACACACGACTCTTCTGTGCAGATAAAA
CAGAAAGGGTAGGGAAAAGTGGCAATGTACCAGCAGGCACTACAGTGGATAGTACCATCACA
CATCCATCTGAGTTTGACTTTTACCTCTGTAGTCATGCAGGAATTCAGGGAACCAGCCGTCC
CTCACATTACCAGGTCTTGTGGGATGACAACTGCTTCACTGCAGATGAACTCCAGCTACTGA
CTTACCAGCTGTGTGCACACCTATGTGAGGTGCACTCGCTCAGTCTCTATTCCAGCCCCCTGCA
TATTATGCCCCGCTTGTAGCATTTAGGGCAAGGTATCATCTGGTGGATAAAGATCATGACAG
TGCGGAAGGCAGTCATGTGTGAGGACAGAGCAACGGCCGGGATCCTCAGGCCTTGGCTAAGG
CTGTGCAAATCCACCATGATACCCAGCACACGATGTATTTTGCC

Martinez et al. Figure 16

GCAGGACCCGCTGGGGCCAGCCCCTACTCATGGTGCCCAGAAGACCTGGCTATGGCACCAT
GGGCAAACCCATTAAACTGCTGGCTAACTGTTTTCAAGTTGAAATCCCAAAGATTGATGTCT
ACCTCTATGAGGTAGATATTAAACCAGACAAGTGTCTAGGAGAGTGAACAGGGAGGTGGTT
GACTCAATGGTTCAGCATTTTAAAGTAACTATATTTGGAGACCGTAGACCAGTTTATGATGG
AAAAAGAAGTCTTTACACCGCCAATCCACTTCCTGTGGCAACTACAGGGGTAGATTTAGACG
TTACTTTACCTGGGGAAGGTGGAAAAGATCGACCTTTCAAGGTGTCAATCAAATTTGTCTCT
CGGGTGAGTTGGCACCTACTGCATGAAGTACTGACAGGACGGACCTTGCCTGAGCCACTGGA
ATTAGACAAGCCAATCAGCACTAACCCTGTCCATGCCGTTGATGTGGTGCTACGACATCTGC
CCTCCATGAAATACACACCTGTGGGGCGTTCATTTTTCTCCGCTCCAGAAGGATATGACCAC
CCTCTGGGAGGGGGCAGGGAAGTGTGGTTTGGATTCCATCAGTCTGTTCCGGCCTGCCATGTG
GAAAATGATGCTTAATATCGATGTTTCTGCCACTGCCTTCTACAAAGCACAACTGTAATTC
AGTTCATGTGTGAAGTTCCTTGATATTCATAATATTGATGAGCAACCAAGACCTCTGACTGAT
TCTCATCGGGTAAATTCACCAAAGAGATAAAAGGTTTGAAGGTTGAAGTGAAGTCAATTGTGG
AACAAATGAGACGGAAATACCGTGTTTGTAATGTAACAAGGAGGCCTGCCAGTCATCAAACCT
TTCCTTTACAGTTAGAAAACGGCCAAACTGTGGAGAGAACAGTAGCGCAGTATTTTCAGAGAA
AAGTATACTCTTCAGCTGAAGTACCCGCACCTTCCCTGTCTGCAAGTCGGGCAGGAACAGAA
ACACACCTACCTGCCACTAGAAGTCTGTAATATTGTGGCAGGGCAACGATGTATCAAGAAGC
TAACAGACAATCAGACTTCCACTATGATCAAGGCAACAGCAAGATCTGCACCAGATAGACAA
GAGGAAATTAGCAGATTGGTAAGAAGTGCAAATTATGAAACAGATCCATTTGTTTCAGGAGTT
TCAATTTAAAGTTCGGGATGAAATGGCTCATGTAAGTGGACGCGTACTTCCAGCACCTATGC
TCCAGTATGGAGGACGGAATCGGACAGTAGCAACACCGAGCCATGGAGTATGGGACATGCGA
GGGAAACAATTCCACACAGGAGTTGAAATCAAATGTGGGCTATCGCTTGTTTGGCCACACA
GAGGCAGTGCAGAGAAGAAATATTGAAGGGTTTCACAGACCAGCTGCGTAAGATTTCTAAGG
ATGCAGGGATGCCCATCCAGGGCCAGCCATGCTTCTGCAAATATGCACAGGGGGCAGACAGC
GTAGAGCCCATGTTCCGGCATCTCAAGAACACATATTCTGGCCTACAGCTTATTATCGTCAT
CCTGCCGGGGAAGACACCAGTGTATGCGGAAGTGAAACGTGTAGGAGACACACTTTTGGGTA
TGGCTACACAATGTGTTCAAGTCAAGAATGTAATAAAAACATCTCCTCAAACCTCTGTCAAAC
TTGTGCCTAAAGATAAATGTTAACTCGGAGGGATCAATAATATTCTTGTAACCTCATCAAAG
ACCTTCTGTGTTCCAGCAACCAGTGATCTTTTTTGGGAGCCGATGTCACTCATCCACCTGCTG
GTGATGGAAAGAAGCCTTCTATTGCTGCTGTTGTAGGTAGTATGGATGCACACCCAAGCAGA
TACTGTGCCACAGTAAGAGTTCAGAGACCCCGACAGGAGATCATCCAGGACTTGGCCTCCAT
GGTCCGGGAACCTTCTTATTCAATTTTATAAGTCAACTCGGTTCAGCCTACTCGTATCATCT
TTTATCGGGATGGTGTTTCAGAGGGGCAGTTTAGGCAGGTATTATATTATGAACTACTAGCA
ATTTCGAGAAGCCTGCATCAGTTTGGAGAAAGACTATCAACCTGGAATAACCTACATTGTAGT
TCAGAAGAGACATCACACTCGATTATTTTGTGCTGATAGGACAGAAAGGGTTGGAAGAAGTG
GCAATATCCCAGCTGGAACAACAGTTGATACAGACATTACACACCCATATGAGTTTCGATTTT
TACCTCTGTAGCCATGCTGGAATACAGGGTACCAGTCGTCCTTCACACTATCATGTTTTATG
GGATGATAACTGCTTTACTGCAGATGAACTTCAGCTGCTAACTTACCAGCTCTGCCACACTT
ACGTACGCTGTACACGATCTGTTTCTATACCTGCACCAGCGTATTATGCTCACCTGGTAGCA
TTTAGAGCCAGATATCATCTTGTGGACAAAGAACATGACAGTGCTGAAGGAAGTCACGTTTC
AGGACAAAGCAATGGGCGAGATCCACAAGCTCTTGCCAAGGCTGTACAGATTCACCAAGATA
CCTTACGCACAATGTACTTCGCTTAA

23/25

Martinez et al. Figure 16

>HILI, cDNA sequence of predicted ORF

ATATCTTCTGGTGATGCTGGAAGTACCTTCATGGAAAGAGGTGTGAAAAACAAACAGGACTT
TATGGATTTGAGTATCTGTACCAGAGAAAAATTGGCACATGTGAGAAATTGTAAACAGGTT
CCAGTGGAATACCTGTGAAACTGGTTACAAACCTCTTTAACTTAGATTTTCCCAAGACTGG
CAGCTATACCAGTACCATGTGACATATATTCCAGATTTAGCATCTAGAAGGCTGAGAATTGC
TTTACTTTATAGTCATAGTGAACCTTTCCAACAAAGCAAAAGCATTCGACGGTGCCATCCTTT
TTCTGTCACAAAAGCTAGAAGAAAAGGTCACAGAGTTGTCAAGTGAAACTCAAAGAGGTGAG
ACTATAAAGATGACTATCACCTGAAGAGGGAGCTGCCATCAAGTTCTCCCGTGTGCATCCA
GGTCTTCAATATCATCTTCAGAAAGATCCTCAAAAAGTTGTCCATGTACCAAATTGGACGGA
ACTTCTATAATCCTTCAGAGCCAATGGAAATTCCCCAGCACAAATTATCCCTTTGGCCTGGG
TTTGCCATTTCTGTGTCATATTTTGAAAGGAAGCTCCTGTTTAGTGCTGATGTGAGTTACAA
AGTCCTCCGGAATGAGACGGTTCTGGAATTCATGACTGCTCTCTGTCAAAGAACTGGCTTGT
CCTGTTTCACCCAGACGTGTGAGAAGCAGCTAATAGGGCTCATTGTCTTACAAGATACAAT
AACAGAACCTACTCCATTGATGACATTGACTGGTCAGTGAAGCCCACACACACCTTTCAGAA
GCGGGATGGCACCGAGATCACCTATGTGGATTACTACAAGCAGCAGTATGATATTACTGTAT
CGGACCTGAATCAGCCCATGCTTGTAGTCTGTAAAGAAGAAGAGAAATGACAACAGTGAG
GCTCAGCTCGCCACCTGATACCTGAGCTCTGCTTTCTAACAGGGCTGACTGACCAGGCAAC
ATCTGATTTCCAGCTGATGAAGGCTGTGGCTGAAAAGACACGTCTCAGTCTTCAGGCCGGC
AGCAGCGCCTGGCCAGGCTTGTGGACAACATCCAGAGGAATACCAATGCTCGCTTTGAACTA
GAGACCTGGGGACTGCATTTTGGAAGCCAGATATCTCTGACTGGCCGGATTGTGCCTTCAGA
AAAAATATTAATGCAAGACCACATATGTCAACCTGTGTCTGCTGCTGACTGGTCCAAGGATA
TTCGAACCTTGCAAGATTTTAAATGCACAGTCTTTGAATACCTGGTTGATTTTATGTAGCGAC
AGAACTGAATATGTTGCCGAGAGCTTTCTGAACTGCTTGAGAAGAGTTGCAGGTTCCATGGG
ATTTAATGTAATGTGCATTCTGCCTTCTAATCAGAAGACCTATTATGATTCCATTAAAAAAT
ATTTGAGCTCAGACTGCCAGTCCCAAGCCAATGTGTGCTTGCTCGGACCTTGAATAAACAG
GGCATGATGATGAGTATCGCCACCAAGATCGCTATGCAGATGACTTGCAAGCTCGGAGGCGA
GCTGTGGGCTGTGGAAATACCTTTAAAGTCCCTGATGGTGGTCCGTATTGATGTCTGTAAAG
ATGCACTCAGCAAGGACGTGATGGTTGTTGGATGCGTGGCCAGTGTTAACCCCAAGATCACC
AGGTGGTTTTTCCCGCTGTATCCTTCAGAGAACAATGACTGATGTTGCAGATTGCTTGAAAGT
TTTCATGACTGGAGCACTCAACAAATGGTACAAGTACAATCATGATTTGCCAGCACGGATAA
TTGTGTACCGTGCTGGTGTAGGGGATGGTCAGCTGAAAACACTTATTGAATATGAAGTCCCA
CAGCTGCTGAGCAGTGTGGCAGAATCCAGCTCAAATACCAGCTCAAGACTGTCGGTGATTGT
GGTCAGGAAGAAGTGCATGCCACGATTCTTTACCGAAATGAACCGCACTGTACAGAACCCCC
CACTTGGCACTGTTGTGGATTTCAGAAGCAACACGTAACGAATGGCAGTATGACTTTTATCTG
ATCAGCCAGGTGGCCTGCCGGGGAAGTGTAGTCCTACCTACTATAATGTCATCTATGATGA
CAACGGCTTGAAGCCCGACCATATGCAGAGACTTACATTCAAATTGTGCCACCTGTACTACA
ACTGGCCGGGCATAGTCAGTGTCCCAGCACCATGTCAGTATGCTCACAAGCTGACCTTTCTG
GTGGCACAAAGCATTCATAAAGAACCCAGTCTGGAATTAGCCAACCATCTCTTCTACCTG

24/25

Martinez et al. Figure 16

>HIWI, cDNA sequence of predicted ORF

ATGACTGGGAGAGCCCGAGCCAGAGCCAGAGGAAGGGCCCGCGGTCAGGAGACAGCGCAGCT
GGTGGGCTCCACTGCCAGTCAGCAACCTGGTTATATTCAGCCTAGGCCTCAGCCGCCACCAG
CAGAGGGGGAATTATTTGGCCGTGGACGGCAGAGAGGAACAGCAGGAGGAACAGCCAAGTCA
CAAGGACTCCAGATATCTGCTGGATTTCAGGAGTTATCGTTAGCAGAGAGAGGAGGTCGTCTG
TAGAGATTTTCATGATCTTGGTGTGAATACAAGGCAGAACCTAGACCATGTTAAAGAATCAA
AAACAGGTTCTTCAGGCATTATAGTAAGGTTAAGCACTAACCATTTCCGGCTGACATCCCGT
CCCCAGTGGGCCTTATATCAGTATCACATTGACTATAACCCACTGATGGAAGCCAGAAGACT
CCGTTTCAGCTCTTCTTTTCAACACGAAGATCTAATTGGAAAGTGCCATGCTTTTGATGGAA
CGATATTATTTTACCTAAAAGACTACAGCAAAAGGTTACTGAAGTTTTTAGTAAGACCCGG
AATGGAGAGGATGTGAGGATAACGATCACTTTAACAATGAACTTCCACCTACATCACCAAC
TTGTTTGCAGTTCTATAATATTATTTTCAGGAGGCTTTTGAAAATCATGAATTTGCAACAAA
TTGGACGAAATTATTATAACCCAAATGACCCAATTGATATTCCAAGTCACAGGTTGGTGATT
TGGCCTGGCTTCACTACTTCCATCCTTCAGTATGAAAACAGCATCATGCTCTGCACTGACGT
TAGCCATAAAGTCCTTCGAAGTGAGACTGTTTTGGATTTCATGTTCAACTTTTATCATCAGA
CAGAAGAACATAAATTTCAAGAACAAGTTTCCAAAGAACTAATAGGTTTAGTTGTTCTTACC
AAGTATAACAATAAGACATACAGAGTGGATGATATTGACTGGGACCAGAATCCCAAGAGCAC
CTTTAAGAAAGCCGACGGCTCTGAAGTCAGCTTCTTAGAATACTACAGGAAGCAATACAACC
AAGAGATCACCGACTTGAAGCAGCCTGTCTTGGTCAGCCAGCCCAAGAGAAGGCGGGGCCCT
GGGGGGACACTGCCAGGGCCTGCCATGCTCATTCCTGAGCTCTGCTATCTTACAGGTCTAAC
TGATAAAATGCGTAATGATTTTAACGTGATGAAAGACTTAGCCGTTTCATACAAGACTAACTC
CAGAGCAAAGGCAGCGTGAAGTGGGACGACTCATTGATTACATTCAATAAAACGATAATGTT
CAAAGGGAGCTTCGAGACTGGGGTTTGAGCTTTGATTCCAACCTACTGTCCTTCTCAGGAAG
AATTTTGCAAACAGAAAAGATTCACCAAGGTGAAAAACATTTGATTACAATCCACAATTTG
CAGATTGGTCCAAAGAAACAAGAGGTGCACCATTAATTAGTGTTAAGCCACTAGATAACTGG
CTGTTGATCTATACGCGAAGAAATTATGAAGCAGCCAATTCATTGATACAAAATCTATTTAA
AGTTACACCAGCCATGGGCATGCAAATGAGAAAAGCAATAATGATTGAAGTGGATGACAGAA
CTGAAGCCTACTTAAGAGTCTTACAGCAAAAGGTCACAGCAGACACCCAGATAGTTGTCTGT
CTGTTGTCAAGTAATCGGAAGGACAAATACGATGCTATTAATAAATACTGTGTACAGATTG
CCCTACCCCAAGTCAGTGTGTGGTGGCCCGAACCTTAGGCAAACAGCAAACCTGTCATGGCCA
TTGCTACAAAGATTGCCCTACAGATGAACTGCAAGATGGGAGGAGAGCTCTGGAGGGTGGAC
ATCCCCCTGAAGCTCGTGATGATCGTTGGCATCGATTGTTACCATGACATGACAGCTGGGCG
GAGGTCAATCGCAGGATTTGTTGCCAGCATCAATGAAGGGATGACCCGCTGGTTCTCACGCT
GCATATTTTCAGGATAGAGGACAGGAGCTGGTAGATGGGCTCAAAGTCTGCCTGCAAGCGGCT
CTGAGGGCTTGAATAGCTGCAATGAGTACATGCCCAGCCGGATCATCGTGTACCGCGATGG
CGTAGGAGACGGCCAGCTGAAAACACTGGTGAACCTACGAAGTGCCACAGTTTTTGGATTGTC
TAAAATCCATTGGTAGAGGTTACAACCCTAGACTAACGGTAATTGTGGTGAAGAAAAGAGTG
AACACCAGATTTTTTGCTCAGTCTGGAGGAAGACTTCAGAATCCACTTCCTGGAACAGTTAT
TGATGTAGAGGTTACCAGACCAGAATGGTATGACTTTTTTATCGTGAGCCAGGCTGTGAGAA
GTGGTAGTGTTTCTCCACACATTACAATGTCATCTATGACAACAGCGGCCTGAAGCCAGAC
CACATACAGCGCTTGACCTACAAGCTGTGCCACATCTATTACAACCTGGCCAGGTGTCATTCG
TGTTCTGCTCCTTGCCAGTACGCCCACAAGCTGGCTTTTCTTGTGGCCAGAGTATTCACA
GAGAGCCAAATCTGTCACTGTCAAACCGCCTTTACTACCTC

25/25

Martinez et al. Figure 17

A

Gene name	1 st primer pair (5'-3')	2 nd primer pair (5'-3')	Expected length (bp)
eIF2C1	GAGGTCTGTAACATTGTGGC* CGGTAGAAGATGATGCGGGT	GAGGTCTGTAACATTGTGGC* AAGTTCTTGAGCACCTCTTCTCGA	287
	GAGGTCTGTAACATTGTGGC CGGTAGAAGATGATGCGGGT	CCACACCAGCGCTCTGCC CTCACGCACCATGTAGGA	207
eIF2C2	GAGGTCTGTAACATTGTGGC CGGTAGAAGATGATGCGGGT	ATCCTGCTGCCCCAAGGG GATCTCCTGCCGGTGCTG	186
	GAGGTCTGTAACATTGTGGC* CGGTAGAAGATGATGCGGGT	GAGGTCTGTAACATTGTGGC* GATCTCCTGCCGGTGCTG	891
eIF2C3	AGAGCAACAGTATGGTGGGTGGAC TGGATGTGTGATGGTACT*	CCTCTACAGTCAAGAGGT TGGATGTGTGATGGTACT*	334
	CACTTGAATGAAGTCCCA TCCTGGATGACCTCTTGACTGTAG*	AGAGCAACAGTATGGTGGGTGGAC TCCTGGATGACCTCTTGACTGTAG*	808
eIF2C4	TCCGGCATCTCAAGAACACATATTCT GAACTCATATGGGTGTGTAATGTCTG*	ATCCAGGACTTGGCCTCC GAACTCATATGGGTGTGTAATGTCTG*	324
HILI	CAGCACAAATTATCCCTT* CGGCCTGAAGGACTGAGACGTGT	CAGCACAAATTATCCCTT* GTGTGTGGGCTTCACTGA	264
	TCTCTGTCAAAGAACTGGCTTGTCTT* CTGTACAGTGCGGTTTCAT	TCTCTGTCAAAGAACTGGCTTGTCTT* CGGCCTGAAGGACTGAGACGTGT	393

* primers used in both reactions (semi-nested PCR)

B

Gene name	eIF2C1		eIF2C2		eIF2C3		eIF2C4	HILI	
Expected length (bp)	287	207	186	891	808	334	324	264	393
PCR products	