



(86) Date de dépôt PCT/PCT Filing Date: 2002/10/18

(87) Date publication PCT/PCT Publication Date: 2003/05/01

(45) Date de délivrance/Issue Date: 2015/07/07

(85) Entrée phase nationale/National Entry: 2004/03/23

(86) N° demande PCT/PCT Application No.: GB 2002/004739

(87) N° publication PCT/PCT Publication No.: 2003/035678

(30) Priorités/Priorities: 2001/10/24 (GB0125498.6);
2002/06/28 (GB0214945.8)

(51) Cl.Int./Int.Cl. *C12N 15/55* (2006.01),
A61K 31/7088 (2006.01), *A61K 38/46* (2006.01),
A61K 48/00 (2006.01), *C07H 21/00* (2006.01),
C07K 14/47 (2006.01), *C07K 16/40* (2006.01),
C07K 7/00 (2006.01), *C12N 9/16* (2006.01),
C12P 21/02 (2006.01), *C12Q 1/44* (2006.01),
C12Q 1/68 (2006.01), *G01N 33/53* (2006.01)

(72) Inventeur/Inventor:
LAI, TONY, GB

(73) Propriétaire/Owner:
UNIVERSITY COLLEGE CARDIFF CONSULTANTS
LIMITED, GB

(74) Agent: GOWLING LAFLEUR HENDERSON LLP

(54) Titre : SEQUENCE DE FACTEUR SPERMATIQUE

(54) Title: SPERM FACTOR SEQUENCES

(57) **Abrégé/Abstract:**

A nucleic acid sequence, including an isolated, purified or recombinant nucleic acid sequence, comprises: (a) a nucleic acid sequence encoding a polypeptide encompassed by the present invention, namely, a PLC zeta; PLC ζ amino acid sequence, capable of triggering calcium oscillations in oocytes; (b) a sequence substantially homologous to or that hybridizes to sequence (a) under stringent conditions; (c) a sequence substantially homologous to or that hybridizes to the sequences (a) or (b) but for degeneracy of the genetic code; and (d) an oligonucleotide specific for any of the sequences (a), (b) or (c) above.

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
1 May 2003 (01.05.2003)

PCT

(10) International Publication Number
WO 03/035678 A3

(51) International Patent Classification⁷: **C07K 14/00**

(21) International Application Number: PCT/GB02/04739

(22) International Filing Date: 18 October 2002 (18.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0125498.6 24 October 2001 (24.10.2001) GB
0214945.8 28 June 2002 (28.06.2002) GB

(71) Applicant (*for all designated States except US*): **UNIVERSITY OF WALES COLLEGE OF MEDICINE** [GB/GB]; Heath Park, Cardiff CF14 4XN (GB).

(72) Inventor; and

(75) Inventor/Applicant (*for US only*): **LAI, Tony** [GB/GB]; University of Wales College of Medicine, Heath Park, Cardiff CF14 4XN (GB).

(74) Agents: **WYNNE-JONES, LAINE & JAMES** et al.; Morgan Arcade Chambers, 33 St Mary Street, Cardiff CF10 1AF (GB).

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

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

Published:

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

(88) Date of publication of the international search report:
26 June 2003

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: SPERM FACTOR SEQUENCES

(57) Abstract: A nucleic acid sequence, including an isolated, purified or recombinant nucleic acid sequence, comprises: (a) a nucleic acid sequence encoding a polypeptide encompassed by the present invention, namely, a PLC zeta; PLC ζ amino acid sequence, capable of triggering calcium oscillations in oocytes; (b) a sequence substantially homologous to or that hybridizes to sequence (a) under stringent conditions; (c) a sequence substantially homologous to or that hybridizes to the sequences (a) or (b) but for degeneracy of the genetic code; and (d) an oligonucleotide specific for any of the sequences (a), (b) or (c) above.



WO 03/035678 A3

SPERM FACTOR SEQUENCES

This invention relates to the regulation and control of intracellular free calcium ion concentrations and more particularly to the control of cytoplasmic calcium

5 oscillations (CCOs) in mammalian oocytes (eggs). In particular, it relates to phospho-
inositide-specific phospholipase C proteins that trigger CCOs that are
indistinguishable from those occurring at oocyte fertilization and nucleic acid
sequences encoding such proteins, and the use thereof in biotechnology, diagnosis or
medicine.

10

Transient changes in the intracellular calcium (Ca^{2+}) concentration are known to be
responsible for activating numerous physiological processes, including memory
formation, muscle contraction, hormone secretion, fertilization, gene transcription and
apoptosis. One striking phenomenon observed in numerous cell types, including
15 cardiac myocytes, endothelial cells and eggs, is the generation of a series of regular
calcium transients or oscillations in response to cellular stimuli. The best-studied
example of this phenomenon is during mammalian fertilization, where calcium levels
in the egg begin to oscillate in a regular fashion following fusion with the sperm.

These calcium oscillations occurring at fertilisation, sometimes referred to as the
20 “calcium waves”, are believed to be the trigger for egg activation and consequent
embryo development. Studies carried out over many years have attempted to discover
and isolate the causative agent of this phenomenon with a view to using it for research
and for a variety of practical applications, including diagnosis.

25 This striking Ca^{2+} signalling phenomenon in fertilized mammalian eggs arises from
increases in inositol 1,4,5-trisphosphate (IP_3) levels, which activates IP_3 receptor-
mediated Ca^{2+} release from intracellular stores in the egg. However, the basic
mechanism involved in stimulation of phospho-inositide metabolism following
sperm-egg interaction has not been determined in any species.

30

The 'sperm factor hypothesis' of signalling at fertilization proposes that spermatozoa contain a soluble Ca^{2+} -releasing factor that enters the egg after the gamete membranes fuse together and generates Ca^{2+} oscillations. This is consistent with the finding that cytoplasmic fusion of sperm and egg is a prelude to Ca^{2+} release. Direct support for this hypothesis comes from experiments where micro-injection into eggs of either single spermatozoa or soluble sperm extracts triggers Ca^{2+} oscillations similar to those at fertilization in mammalian - and some non-mammalian - eggs. The mammalian sperm factor that generates Ca^{2+} oscillations is protein-based; acts across species; and can cause Ca^{2+} release in somatic cells as well as in cell-free systems, such as sea urchin egg homogenates. Sperm specifically express a Ca^{2+} oscillation-inducing protein, because micro-injecting messenger RNA (mRNA) isolated from spermatogenic cells, but not mRNA from other tissues, elicits fertilization-like Ca^{2+} oscillations in mouse eggs.

In intact eggs and egg homogenates, mammalian sperm extracts trigger Ca^{2+} release via stimulating IP_3 production, indicating involvement of a phospho-inositide-specific phospholipase C (*ie* PI-PLC, usually referred to in short as PLC) in the signal transduction mechanism. The high level of PLC enzyme activity measured biochemically in sperm extracts has led some researchers to suggest that the sperm factor may itself be a PLC. However, the PLC-beta, gamma and delta (β , γ and δ) isoforms that exist in sperm are not detected in the chromatographic fractions of sperm extract that specifically cause Ca^{2+} oscillations. Also, when purified, recombinant PLC β 2, γ 1 or δ 1 proteins are added to egg homogenates, they fail to cause Ca^{2+} release. A PLC δ 4 splice variant expressed in sperm has been shown to be involved in the acrosome reaction, rather than Ca^{2+} release in eggs at fertilization. Previous research in this field has been described in international patent specification no. WO 96/25495, to which reference should be made for a full understanding of, and as background to, the present application.

Patent specification no. WO 96/25945 assigned the cause of the above-mentioned calcium oscillations to a substance (a sperm factor) present in the equatorial segment of sperm, which was believed to diffuse into the egg after fusion therewith. This substance was identified as a 33kD (approx.) protein of specified amino acid sequence. The nucleic acid coding for this protein was also specified. However, after cloning the gene and undertaking subsequent expression studies, it was concluded that this sperm factor candidate was unable to reconstitute calcium oscillations. A truncated form of the *c-kit* receptor, has also previously been a sperm factor candidate. However, neither these two, nor any other sperm proteins, have been shown to generate Ca^{2+} oscillations in eggs, the single-most distinctive feature of mammalian fertilization.

These observations have led some workers in the field to conclude, "sperm-derived PLC is not responsible for initiating Ca^{2+} release at fertilization" (Mehlmann *et al* in Dev Biol 236 492-501 (2001)), whilst others have stated: "the identification of this protein remains a problem for the next century of fertilization research" (Runft *et al* in Dev Biol 245 237-54 (2002)).

On the contrary, these observations led us to investigate the possible existence of a distinct, uncharacterised sperm PLC isoform. The present invention relates to the presence of a new PLC isoform specifically expressed in mammalian sperm (hereinafter called PLC-zeta; PLC ζ), which uniquely possesses all the essential properties of the sperm factor. The results of our studies are consistent with sperm PLC being the physiological trigger of egg activation, and thus an essential protein for mammalian fertilization and embryo development.

The amino acid sequences of both the human and mouse proteins are given hereinafter as SEQ ID NOS: 1 and 2, respectively, and their nucleic acid coding sequences as SEQ ID NOS: 3 and 4, respectively. Also given is the rat protein as SEQ ID NO: 11, and its nucleic acid coding sequence SEQ ID NO: 10.

Recently, the Genbank database disclosed various nucleic acid sequences of human and mouse testes, without attributing any function thereto and predicting an open reading frame (ORF; protein or polypeptide sequence) having a start position
5 corresponding to a position being at least 100 amino acids from the start position of the SEQ ID NOS: 3 and 4. In particular, Genbank Accession No AK006672 (deposited 05-JUL-2001) comprises 2227 base pairs of mouse testis sequence but predicts an ORF encoding 537 amino acids with a start position corresponding to position aa 111 (MEIDH) of the mouse sequence [SEQ ID NO: 4] (*ie* missing the first
10 110aa (amino acids).);

Genbank Accession No XM029802 (deposited 16-OCT-2001) comprises 2113 base pairs of human testis sequence, not identical to and predictive of an ORF encoding 504 amino acids with a start position corresponding to position aa 105
15 (MSKAI) of the human sequence [SEQ ID NO: 3] (*ie* missing the first 104aa);

Genbank Accession No NM033123 (deposited 21-AUG-2001) comprises 2132 base pairs of human testis sequence in database, but predicts an ORF encoding 504 amino acids with a start position corresponding to position aa 105 (MSKAI) of the human
20 sequence [SEQ ID NO: 3] (*ie* missing the first 104aa); and

Genbank Accession No AY035866 (deposited 22-JUN-2001) comprises 2132 base pairs of human testis sequence in database, but predicts an ORF encoding 504 amino acids with a start position corresponding to position aa 105 (MSKAI) of the human
25 sequence [SEQ ID NO: 3] (*ie* missing the first 104aa).

Genbank Accession No. AB070108 (deposited 16-AUG-2001) comprises 2219 base pairs of monkey testis sequence with an ORF of 1923 base pairs (nucleotides 220-2142) encoding 641 amino acids, without attributing any function thereto or
30 connection with a putative sperm factor. [SEQ ID NOS: 6 and 7, respectively].

Similarly, Genbank Accession No. AB070109 (deposited 16-AUG-2001) comprises 2218 base pairs of monkey testis sequence with an ORF of 1920 base pairs (nucleotides 220-2139) encoding 640 amino acids, without attributing any function thereto or connection with a putative sperm factor. [SEQ ID NOs: 8 and 9,
5 respectively].

The differences in the protein sequence between AB070108 and AB070109 are shown below:

AB070108 (aa334-343):	E E E	E K	F K E S E
10 AB070109 (aa334-342):	E E E	- R	F K E S E

Accordingly, the present invention provides a PLC-zeta protein, characterised by exhibiting one or more of the following properties:

- 15 (a) An amino acid sequence comprising in the range of from 600 to 720, preferably 600 to 699, more preferably 600 to 650, amino acid residues;
- (b) A domain sequence comprising the EF hand, X, Y, and C2 domains but absent the PH domain; and
- 20 (c) At least five consecutive amino acid residues from a conserved region, which region is selected from:
- | | | | |
|-----------|-----------|----------------|----------------------|
| (i) | QDDFRGGKI | (11-19); | |
| (ii) | LLEKLD | (27-32); and | |
| 25 (iii) | QGRIT | (52-56) | in the EF1 domain; |
| (iv) | ENRKIL | (82-87); and | |
| (v) | FLTQEY | (95-101) | in the EF2 domain; |
| (vi) | YQQFNE | (403-408) | in the Y domain; and |
| (vii) | TLTIR | (516-520); | |
| 30 (viii) | ISGIQLP | (522-528); and | |
| (ix) | LCMNKGYRR | (609-617) | in the C2 domain, |

wherein the residues are denoted by their conventional single letter codes and the numbers in parentheses refer to the sequence co-ordinates within the 641 amino acid ORF of the monkey AB070108 (monkey A) sequence .

- 5 Below is a table (Table 1) showing a comparison between lengths of various PLCs (criterion (a)); Figure 3 shows the domain comparison between the various PLCs (criterion (b)); and [SEQ ID NO: 12] illustrates the conserved regions of PLC-zeta cross-species, compared to other PLCs (criterion (c)). Comparison between the PLC sequence types was made using the Clustal W analysis program, available at
- 10 <http://www.clustalw.genome.ad.jp>, using the default settings.

Table 1: PLCs – Sequence Lengths

PLC Types	Species	No. of amino acid residues
zeta	human	608
zeta	monkeyA	641
zeta	mouse	647
zeta	rat	646
beta 1	human	1211
beta 2	human	1181
beta 3	mouse	1234
beta 4	mouse	1175
gamma 1	human	1290
gamma 2	human	1252
delta 1	human	756
delta 2	bovine	764
delta 3	human	736
delta 4	rat	772
epsilon	human	2302

PLC Types	Species	No. of amino acid residues
1	potato	596
2	potato	565
3	potato	585

Accordingly, the present invention provides an isolated, purified or recombinant nucleic acid molecule comprising:

a nucleic acid molecule encoding a PLC-zeta; PLC ζ , polypeptide, capable of

5 triggering
calcium oscillations in oocytes.

The nucleic acid molecule of the invention is identified by the virtue of the sequences disclosed herein and further includes sequences substantially homologous thereto or
10 sequences that hybridize thereto under stringent conditions.

In a further aspect of the invention there is provided at least one oligonucleotide specific for a part of the aforementioned sequences. Preferably, said oligonucleotide includes the primers described herein and more specifically the following:

15

Forward human primer: 5' CAG CGA GCT CTT ATC TGA CGT ACC AAA C 3'
(28mer).

20

Reverse TriplEx primer: 5' CTC GGG AAG CGC GCC ATT GTG TTG GT 3'
(26mer).

Forward mouse primer: 5' GCT AAC GCG TCA GTT ACA TGC GTC ACT C 3'
(28mer)

25

Reverse T7 primer: 5' GTA ATA CGA CTC ACT ATA GGG C 3' (22mer)

Forward human primer: 5' CAG CGA GCT CTT ATC TGA CGT ACC AAA C 3'
(28mer)

5 Reverse human primer: 5' ATG AAA CTA TGG AAA TGA GAT GGT 3' (24mer)

Forward mouse primer: 5' GCT AAC GCG TCA GTT ACA TGC GTC ACT C 3'
(28mer)

10 Reverse mouse primer: 5' ATC ATG GAA AGC CAA CTT C 3' (19mer)

By "substantially homologous" herein is meant that the nucleic acid sequence has at least 70% identity of its nucleotide bases with those of sequence (a), in matching positions in the sequence. A further 10% of its nucleotide bases may comprise
15 conservative substitutions (with similar bases), and therefore the sequence has at least 80% overall homology. More preferred are sequences having at least 80% identity with the sequence (a) and about 90% overall homology. Such homologous sequences encode a protein having substantially the same biological activity as the proteins of the invention.

20

Oligonucleotides "specific for" any of these nucleic acid sequences (a) to (c) above are useful for identifying and isolating the biologically active peptides of this invention, and comprise a unique sequence encoding a unique fragment of the amino acid sequence of the peptide.

25

In particular, the present invention provides a nucleic acid sequence as defined above, wherein the sequence is a DNA or RNA sequence, such as cDNA, cRNA or mRNA. More particularly, the present invention provides:

a DNA sequence identified herein by [SEQ ID NO: 3], which sequence (being the human PLC-zeta; PLC ζ nucleotide sequence, 1827 nucleotides) corresponds with the polypeptide identified herein as [SEQ ID NO: 1];

5 a DNA sequence identified herein by [SEQ ID NO: 4], which sequence (being the mouse PLC-zeta; PLC ζ nucleotide sequence, 1944 nucleotides) corresponds with the polypeptide identified herein as [SEQ ID NO: 2]; and

10 a DNA sequence identified herein by [SEQ ID NO: 10], which sequence (being the rat PLC-zeta; PLC ζ nucleotide sequence, 1938 nucleotides) corresponds with the polypeptide identified herein as [SEQ ID NO: 11].

The mouse sequence has been deposited under Genbank Accession No AF 435950, which comprises 1941 nucleotides of the protein-coding region plus the stop codon (3
15 nucleotides) (these, together, consist of the [SEQ ID NO: 4]) plus the untranslated region (totalling 2187 nucleotides) identified herein as [SEQ ID NO: 5].

Therefore, the present invention further provides a polypeptide of:

[SEQ ID NO: 1], being the human PLC-zeta; PLC ζ amino acid sequence, 608
20 residues;
[SEQ ID NO: 2], being the mouse PLC-zeta; PLC ζ amino acid sequence, total 647 residues; and
[SEQ ID NO: 11], being the rat PLC-zeta; PLC ζ amino acid sequence, total 646 residues, in which amino acids are represented by their conventional single letter
25 codes.

Furthermore, the invention provides for the use of certain known sequences to which a function has not previously been assigned as a PLC-zeta, PLC ζ or sperm factor. In particular, the invention provides for such use of the monkey proteins [SEQ ID NOs:
30 7 and 9].

The deduced human and mouse proteins of SEQ ID NOS: 1 and 2 differ by 39 amino acids in length and their cDNA sequences differ correspondingly. It will be appreciated that similarly active proteins and corresponding nucleic acid sequences encoding them will be present in the sperm of other mammalian species, including
5 species of farm animals *eg* sheep and pigs, and other animal species *eg* fish. All such proteins and nucleic acid sequences have a high degree of sequence homology with one another, and can be readily isolated using the newly discovered DNA sequences or parts thereof to probe the appropriate cDNA libraries of other species. It is expected that the molecular weight of the proteins will be in the range of from 65 to
10 80 kD, preferably in the range of from 70 to 75 kD, especially about 70kD, as determined by mass spectrometry.

Derivatives of the proteins disclosed herein (*ie* of [SEQ ID NOS: 1, 2 and 11], and homologous sequences) having substantially similar biological activity are also
15 encompassed. For example, one or more of which derivatives may comprise post-translational modifications, such as glycosylation at asparagine, serine or threonine; and/or sulphato- or phospho- groups on tyrosine, such as are commonly found in polypeptides; polymorphisms, such as single nucleotide polymorphisms (SNPs); and those further comprising a leader/signal sequence.

20 The invention further provides a tagged derivative of a PLC-zeta, such as a tagged derivative of any polypeptide sequence specifically identified herein, including [SEQ ID NOS: 1, 2, 7, 9 and 11], for use in identifying the PLC-zeta in diagnostic tests, other assays or otherwise as a research or clinical tool. Suitably, the PLC-zeta is
25 tagged with c-Myc as described in Example 6 hereinbelow, antibodies to which are commercially available (*eg* from Santa Cruz Biotechnology).

A polypeptide encompassed by this invention can also be prepared by providing or culturing a host, transformed with an expression vector comprising a DNA sequence
30 encoding the polypeptide under such conditions that the polypeptide is expressed therein, and optionally isolating the polypeptide thus obtained. This approach is

typically based on obtaining a nucleotide sequence encoding the polypeptide it is wished to express, and expressing the polypeptide in a recombinant organism. The cultivation of the genetically modified organism leads to the production of the desired product displaying full biological activity. The present invention therefore also
5 comprises a polypeptide produced by a recombinant DNA technique, which polypeptide is one encompassed above. The invention further comprises a synthetic, or protein-engineered, polypeptide encompassed above.

The present invention therefore further provides: a recombinant construct comprising
10 any nucleic acid sequence according to the invention; a vector comprising such a construct; and a host transformed or transfected by such a vector.

The present invention therefore still further provides a cultured or non-human cell, plasmid, virus, live organism or other vehicle that has been genetically- or protein-
15 engineered to produce a polypeptide according to the present invention, said cell, plasmid, virus, live organism or other vehicle having incorporated expressibly therein a sequence as disclosed herein. Such cells may include animal, such as mammal, for example human or humanised cells, for use in gene therapy to treat or prevent conditions such as those mentioned herein. Such cells particularly include stem cells
20 derived by cell nuclear transfer in accordance with the present invention. The present invention therefore also further provides animal clones derived from nuclear transfer techniques enhanced by using the PLC-zeta of this invention.

Therefore, the present invention further provides a method for the preparation of a
25 polypeptide according to the present invention, which method comprises:
(a) isolation and/or purification from mammalian sperm; or
(b) expression of a nucleic acid sequence encoding the polypeptide and, optionally, isolation and/or purification of the resulting polypeptide.

30 The present invention therefore comprises *inter alia* the human, mouse, rat or other mammalian protein PLC-zeta, or non-mammalian (eg fish) PLC-zeta, the nucleic acid

sequence coding therefor, cells transfected with the nucleic acid sequence, and a process for producing PLC-zeta by cultivation of the transfected cells and recovery of the expressed product.

- 5 The recombinant proteins, especially the mouse (including the c-Myc-tagged mouse), monkey (both AB 070108 and AB070109) and human PLC-zeta, have been shown to generate cytoplasmic calcium oscillations (CCOs) when introduced into mammalian cells. Furthermore, the injection of complementary RNA (cRNA) encoding PLC-zeta into mouse eggs also generates identical CCOs to those observed when they are
10 fertilized by sperm. It has also been found that PLC-zeta is capable of producing embryo development to the blastocyst stage (*ie* the stage at which stem cells are found).

Accordingly, the invention also provides a variety of applications and/or uses of the
15 proteins and nucleic acid sequences of this invention, including the following:

1. *Treatment of mammalian infertility:*

The human PLC-zeta; PLC ζ protein we have identified may be used in treating human male infertility. This PLC-zeta; PLC ζ protein triggers calcium changes upon
20 sperm fusion with egg, the physiological process which results in egg activation and consequent embryo development. Absence or significant reduction of the level of active PLC-zeta; PLC ζ in sperm would be expected to result in infertile males. That the PLC-zeta; PLC ζ protein is highly expressed in mammalian testis is supported by the following:

- 25 (a) the cDNA has been isolated from testis cDNA libraries (human testis and mouse spermatid); and
(b) search of the EST database using our PLC-zeta; PLC ζ sequences human and mouse reveals sequence matches found in testis-derived cDNA libraries.
- 30 Assay of the PLC-zeta; PLC ζ protein in human sperm samples may therefore be used to identify males who have less than normal levels of the active protein (*ie* protein

having the ability to cause cell calcium oscillations) and are infertile for this reason. This assay may be achieved by the use of antibodies to the protein prepared by methods well known to those skilled in the art.

- 5 To correct such deficiencies, the addition of active PLC-zeta; PLC ζ to sperm lacking an active PLC-zeta; PLC ζ can be carried out in conjunction with the clinical IVF (*in vitro* fertilization) technique of intra-cytoplasmic sperm injection, ICSI (Intra-Cytoplasmic Sperm Injection, comprising introduction of a single sperm directly into the egg). The ICSI procedure has been successfully used by major IVF clinics to
10 produce thousands of live human births.

2. *Improvement in stem cell production:*

- The ethical use of 'spare' embryo-derived stem cells in therapy of human degenerative diseases has been of great public debate recently. The generation of stem cells
15 directly from a patient would remove the need for use of donated embryos. Cloning of cells, tissues and animals (*eg* 'Dolly', the sheep) have been achieved by fusing a somatic cell with an enucleated egg. Activation of the fused egg to trigger development of the hybrid cell to form a blastocyst, from which stem cells can be harvested, is a very low efficiency process with <1% success rate. Thus, the
20 application of a native protein with a physiologically relevant activity, that is, to trigger egg activation, following the fusion process occurring between the somatic cell and the egg cell, would enhance the success rate of fused cells in proceeding to develop further.

- 25 Stem cells derived from nuclear transfer techniques enhanced by using PLC-zeta have potential application to a variety of human diseases and conditions, including Parkinsonism, Alzheimer's disease, heart failure and diabetes, to which stem cell therapy could be applied.

30 3. *Animal cloning:*

An extension of the application 2., above, is to implant the successfully developing blastocyst into a pregnant female host to produce full development to term and live birth of clones derived from a single adult animal cell. This process is currently being developed for the production of biomedicines in transgenic animals, e.g. sheep and
5 pigs, as well as for the potential use of animal cells and organs for transplantation into humans but the current success rate for this procedure, as mentioned above is very low, <1%, due to the difficulties in achieving viable hybrid cells upon fusion.

In another aspect, the present invention provides a method for the treatment or
10 prevention of a condition or disorder mentioned herein, wherein the polypeptide is administered by means of being expressed in the cells of the patient, which cells have incorporated expressibly therein a nucleic acid sequence coding for the polypeptide. Alternative to gene therapy, the polypeptides of the invention may be administered as a pharmaceutical formulation.

15 Accordingly, the present invention provides the use of a polypeptide described herein or a nucleic acid sequence coding for the polypeptide in medicine, including gene therapy; and also the use of such a polypeptide in the manufacture of a medicament.

20 Therefore, according to a further aspect of the present invention, there is provided a pharmaceutical formulation comprising a polypeptide according to the invention (as described above) and a pharmaceutically acceptable carrier therefor. The term "pharmaceutically acceptable carrier" as used herein should be taken to mean any inert, non-toxic, solid or liquid filler, diluent or encapsulating material, or other
25 excipient, which does not react adversely with the active ingredient(s) or with a patient.

Such formulations and carriers are well known in the art and include pharmaceutical formulations that may be, for example, administered to a patient systemically, such as
30 parenterally, or orally or topically.

The term 'parenteral' as used here includes subcutaneous, intravenous, intramuscular, intra-arterial and intra-tracheal injection, and infusion techniques. Parenteral formulations are preferably administered intravenously, either in bolus form or as a continuous infusion, or subcutaneously, according to known procedures. Preferred
5 liquid carriers, which are well known for parenteral use, include sterile water, saline, aqueous dextrose, sugar solutions, ethanol, glycols and oils.

Tablets and capsules for oral administration may contain conventional excipients such as binding agents, fillers, lubricants, wetting agents, and the like. Oral liquid
10 preparations may be in the form of aqueous or oily suspensions, solutions, emulsions, syrups, elixirs or the like, or may be presented as a dry product for reconstitution with water or other suitable vehicle for use. Such liquid preparations may contain conventional additives, such as suspending agents, emulsifying agents, non-aqueous vehicles and preservatives.

15 Formulations suitable for topical application may be in the form of aqueous or oily suspensions, solutions, emulsions, gels or, preferably, emulsion-based ointments.

Unit doses of the pharmaceutical formulations according to the invention may contain
20 daily-required amounts of the polypeptides, or sub-multiples thereof to make a desired dose. The optimum therapeutically-acceptable dosage and dose rate for a given patient (which may be a mammal, such as a human) depend on a variety of factors, such as the potency of the active ingredient(s); the age, body weight, general health, sex and diet of the patient; the time and route of administration; rate of clearance; the
25 object of the treatment (for example, treatment or prophylaxis); and the nature of the disease to be treated.

It is expected that systemic doses in the range of from 0.005 to 50 mg/kg body weight, preferably of from 0.005 to 10 mg/kg and more preferably 0.01 to 1 mg/kg, will be
30 effective. According to the nature of the disease being treated, one single dose may

comprise in the range of from 0.005 to 10 mg/kg body weight active ingredient, whether applied systemically or topically.

The present invention therefore further provides:

- 5 (a) the use of a polypeptide of this invention in therapy;
 - (b) the use of a polypeptide of this invention in the preparation of a medicament;
 - (c) a method for the treatment or prevention of a condition in a patient, which condition involves suppression, inhibition or inactivation of the generation of CCOs, which method comprises administration to said patient of a non-toxic, inhibitory
10 amount of a polypeptide of the invention;
 - (d) the use of a polypeptide of this invention in the generation of CCOs in mammalian cells;
 - (e) a method of treating male infertility in a mammal, which method comprises adding the polypeptide of this invention to the sperm of the mammal; and
15 (f) a method of improving oocyte-somatic cell nuclear transfer efficiency in cell cloning, which method comprises adding a polypeptide according to this invention, or nucleic acid encoding the polypeptide, to an oocyte before or after fusion with the contents of a somatic cell.
- 20 Furthermore, the protein or nucleic acid sequence coding therefor according to this invention may be used in a diagnostic method to determine the state of fertility (*eg* whether fertile or infertile) of a respective mammal, such as a human.

Accordingly, the present invention further provides a diagnostic method for
25 determining the fertility status of a mammal, which method comprises determining the amount of a protein according to this invention, or nucleic acid sequence coding therefore, present or absent in a test sample obtained from the mammal, which amount is indicative of the level of fertility of the mammal.

30 A further diagnostic or screening method comprises:

- (a) obtaining a test sample comprising a nucleotide sequence of the mammalian PLC ζ gene from the individual; and
- (b) comparing a region of the sequence obtained from the test sample with the corresponding region of a wild type mammalian PLC ζ sequence, such as [SEQ ID NO: 3, 4, 5, 6, 8 or 10]

whereby a variation in the sample sequence relative to the predetermined sequence is indicative of a condition, such as lowered fertility or infertility, associated with disruption in calcium oscillation patterns that are a prerequisite to normal biological function absent the condition.

Preferably, the test sample comprises genomic DNA.

A particularly preferred screening method is one for screening an individual suspected of a fertility problem, which screening method comprises the steps of:

- (a) obtaining a test sample comprising a nucleotide sequence of the human PLC ζ gene or an amino acid sequence encoded thereby from the individual; and
- (b) analysing the test sample for the presence of a variant of the human PLC ζ gene or an amino acid sequence encoded thereby or for the presence of one or more surrogate markers that are indicative of or correlated to the presence of a variant of the human PLC ζ gene or an amino acid sequence encoded thereby, wherein the variant of the human PLC ζ gene or an amino acid sequence encoded thereby exhibits at least one variation when compared to the wild type PLC ζ sequence.

It will be evident to the person skilled in the art that the above methods apply equally to other mammals than humans and to other animals than mammals.

The analysis step (b) may be selected from one or more of: conventional protein sequencing methods (such as mass spectroscopy, micro-array analysis, pyrosequencing, *etc*), and/or antibody-based methods of detection (*eg* ELISA). In any of the methods according to the invention, antibodies to the protein may be raised.

Therefore, in a method of testing for male infertility, which method comprises assaying the protein PLC ζ in a sperm sample, the method could be carried out using an antibody to the protein, in particular, a monoclonal antibody to the protein PLC ζ . Alternatively, the PLC-zeta gene sequence may be determined in a sample comprising
5 genomic DNA, using methods known to those skilled in the art, such as PCR amplification, restriction enzyme analysis and DNA sequencing.

Accordingly, the present invention still further provides an antibody raised to a polypeptide according to the invention, particularly a monoclonal antibody thereto.

10

The screening method may comprise the use of simultaneous screens for multiple, known variations or for all possible variations by hybridization of a labelled sample of DNA (cDNA or genomic DNA derived from the individual) to micro-arrays of variation-specific oligonucleotide probes immobilised on a solid support. For
15 example, chip technology may be used, wherein the chip is a miniature parallel analytical device.

The methods of the invention may be carried out using a kit, which kit may comprise:
(a) an oligonucleotide comprising a nucleic acid sequence corresponding to a region
20 of a PLC ζ variant, which region incorporates at least one variation from the corresponding wild-type PLC ζ gene sequence; and/or
(b) an oligonucleotide comprising a nucleic acid sequence corresponding to the wild-type PLC ζ gene sequence in the region specified in (a); and/or
(c) an oligonucleotide comprising a nucleic acid sequence corresponding to a specific
25 region of the wild-type PLC ζ gene sequence, which specific region comprises a sequence not otherwise present in the genomic DNA of the mammal; and/or
(d) antibodies, such as monoclonal antibodies, raised to any peptide sequence corresponding to an oligonucleotide specific to any one of (a) to (c) above; and, optionally,
30 (e) one or more reagent(s) suitable for amplifying (eg by carrying out PCR) desired regions of the individual's DNA.

Preferably, any of kit components (a) to (c) comprise(s) a plurality of said oligonucleotides immobilised on a solid support.

- 5 In a further aspect, the present invention provides an inhibitor or antagonist of PLC-zeta for use in reducing, suppressing or preventing cytoplasmic calcium oscillations in oocytes and/or for reducing or inhibiting fertility. Such PLC-zeta inhibitors or antagonists may comprise known chemical compounds, biological material or other agents, or may comprise new active agents. Accordingly, the invention further
- 10 provides an active agent suitable for reducing, suppressing or preventing cytoplasmic calcium oscillations in oocytes and/or for reducing or inhibiting fertility, which active agent is an inhibitor or antagonist of PLC-zeta. Such active agents may be provided in the form of a pharmaceutical formulation in association with a pharmaceutically acceptable carrier therefore, as described above, and may be suitable for use as a male
- 15 contraceptive.

20

25

30

In accordance with an aspect of the present invention, there is provided a diagnostic method for determining the fertility status of a mammal. The method involves determining in a test sample obtained from the mammal a presence or absence of:

a polypeptide having the sequence of SEQ ID NO: 1, being human PLC ζ polypeptide, or a sequence 70% identical thereto; or

a nucleic acid having a sequence encoding human PLC ζ polypeptide capable of triggering calcium oscillations in oocytes and identified herein by SEQ ID NO: 3, or a sequence 70% identical thereto, or a nucleic acid which hybridises to a nucleic acid having the sequence of SEQ ID NO: 3 under stringent conditions,

wherein the absence of the polypeptide or the nucleic acid indicates that the mammal is infertile. The method may involve the test sample including genomic DNA.

In accordance with another aspect of the invention, there is provided a diagnostic or screening method. The method involves:

- (a) sequencing the genetic code of a nucleic acid molecule of human PLC ζ gene in a test sample isolated from a human individual; and
- (b) comparing a region of the code obtained from the test sample with the corresponding region of a wild type human PLC ζ nucleic acid sequence of SEQ ID NO: 3 to determine if there are any variations,

whereby a variation in the sample code relative to the predetermined sequence is indicative of a condition, corresponding to lowered fertility or infertility, associated with disruption in calcium oscillation patterns that are a prerequisite to normal biological function absent in the condition. The method may involve the test sample including genomic DNA.

In accordance with another aspect of the invention, there is provided a screening method for screening a human individual suspected of a fertility problem. The screening method involves the following steps:

- (a) sequencing a nucleic acid of human PLC ζ gene or a polypeptide encoded thereby to determine a nucleic acid sequence or an amino acid sequence, respectively, the nucleic acid or the polypeptide being in a test sample obtained from the human individual; and
- (b) comparing the nucleic acid sequence with the wild type PLC ζ gene sequence of SEQ ID NO:3 or comparing the amino acid sequence with the wild-type PLC ζ amino acid sequence of SEQ ID NO:1, to detect a variant,

wherein detection of a variant indicates a fertility problem. The method may involve the test sample including genomic DNA.

In accordance with another aspect of the invention, there is provided a fertility diagnostic or screening kit for a mammal. The kit involves, in the alternative:

- (a) an oligonucleotide having a nucleic acid sequence corresponding to a region of a PLC ζ variant, the region incorporating at least one variation from a corresponding wild-type PLC ζ gene sequence of SEQ ID NO:3;
- (b) an oligonucleotide having a nucleic acid sequence corresponding to the wild-type PLC ζ gene sequence in the region specified in (a);
- (c) an oligonucleotide having a nucleic acid sequence corresponding to a specific region of the wild-type PLC ζ gene sequence of SEQ ID NO:3, the specific region having a sequence not otherwise present in the genomic DNA of the mammal;
- (d) antibodies, or monoclonal antibodies, raised to any oligonucleotide specific to any one of (a) to (c) above; or
- (e) any combination of components (a) to (d).

The fertility kit may involve one or more reagents suitable for carrying out PCR for amplifying desired regions of the mammal's DNA. The kit components (a) to (c) may involve a plurality of the oligonucleotides immobilised on a solid support.

The invention will now be further described in the following, non-limiting, Examples, with reference to the accompanying Figures 1 to 5, in which:

Figure 1: is a plot of calcium concentration (nM; ordinate) with time (secs; abscissa), showing expression of mouse PLC-zeta plasmid DNA by transfection in CHO cells;

Figure 2: is a plot of calcium concentration (nM; ordinate) with time (secs; abscissa), showing expression of mouse PLC-zeta complementary RNA by micro-injection into mouse eggs;

Figure 3: is a schematic alignment of PLC regions, showing similarities and differences between PLC-zeta and other PLCs;

Figure 4a: is a graph of the percentage of mouse eggs reaching 2-cell stage after 24 hours and morula/blastocyst stage after 96 hours, following micro-injection with PLC-zeta cRNA (0.02 mg/ml) or pathogenically activated with strontium (5 mM, 4 hours) or fertilised with sperm in vivo and placing in a 5% CO₂ incubator at 37C;

5

Figure 4b: comprises two micrographs illustrating mouse embryos at the 2-cell stage and blastocyst stage, respectively, following the treatment illustrated in Figure 4a;

10 Figure 5: is a micrograph illustrating mouse egg 24 hours following micro-injection with ^{D210R}PLC-zeta, illustrating lack of development to 2-cell stage.

Figure 6a: shows dose-dependent calcium oscillations in fura-red loaded mouse eggs, triggered by micro-injection of cRNA encoding mouse sperm PLC-zeta (2 and 15 0.002 mg/ml, top and middle traces, respectively) and after pre-incubation with 10 uM cycloheximide (0.02 mg/ml, bottom trace); and

Figure 6b: illustrates the mean interspike interval of calcium oscillations in eggs, following micro-injection of various PLC-zeta cRNA concentrations. Compared with 20 the interval observed upon *in vitro* fertilisation (IVF). * indicates statistically significant (Student's unpaired t-test) from IVF at the 5% level.

Figure 7: Structure of the human *plc-zeta* gene.

The genomic organisation of the fifteen *plc-zeta* exons identified within the 179456 25 bp contig (Accession number AC023940) are shown aligned to a 54.8 kb region of chromosome 12 (12p12.3). Exons are labelled E1 to E15. The start and stop codons for hPLC α are located within E2 and E15, respectively. Solid line between exons represent the introns (see Table 2).

30 Figure 8: Ca²⁺ oscillations in mouse oocytes microinjected with human PLC-zeta cRNA. A. Dose-dependent Ca²⁺ oscillations in MII-arrested mouse oocytes after

microinjection of hPLC-zeta cRNA. The four traces show the cytoplasmic Ca^{2+} oscillations observed upon microinjection with cRNA at the indicated pipette concentration, from 20 to 0.02 $\mu\text{g/ml}$. B. Mean interspike interval of Ca^{2+} oscillations in mouse oocytes triggered by the various hPLC-zeta cRNA concentrations. The number of microinjected oocytes is shown above each dose. The mean interspike interval at each dose is statistically different from each other using a students paired t-test, $p = <0.0001$ (20 $\mu\text{g/ml}$, 4.21 ± 1.79 ; 2.0 $\mu\text{g/ml}$, 9.26 ± 7.14 ; 0.2 $\mu\text{g/ml}$, 16.0 ± 6.40 ; 0.02 $\mu\text{g/ml}$, 24.34 ± 7.68).

Figure 9: Embryonic development of mouse oocytes microinjected with human PLC-zeta cRNA.

A. Mouse oocytes were microinjected with different hPLC-zeta cRNA concentrations (20-0.2 $\mu\text{g/ml}$). The percentage of oocytes reaching the 2-cell stage after 24 hours and morula/blastocyst after 96 hours were recorded.

B. Micrographs showing development of mouse embryos at the 2-cell stage (left) and blastocyst stage (right) following microinjection of unfertilized oocytes with hPLC-zeta cRNA (0.2 $\mu\text{g/ml}$).

Figure 10: Ca^{2+} oscillations in mouse oocytes with simian PLC-zeta cRNA.

A. Dose-dependent Ca^{2+} oscillations in MII-arrested mouse oocytes after microinjection of sPLC-zeta cRNA. The three traces show the cytoplasmic Ca^{2+} oscillations observed upon microinjection with cRNA at the indicated pipette concentration, from 200 to 2 $\mu\text{g/ml}$.

B. Mean interspike interval of Ca^{2+} oscillations in mouse oocytes triggered by the various sPLC-zeta cRNA concentrations. The number of microinjected oocytes is shown above each dose. The mean interspike interval at each dose is statistically different from each other using a students paired t-test, $p = <0.0001$ (200 $\mu\text{g/ml}$, 3.18 ± 0.55 ; 20 $\mu\text{g/ml}$, 7.35 ± 2.69 ; 2.0 $\mu\text{g/ml}$, 15.77 ± 5.20).

Figure 11: Mean interspike intervals observed with human, simian and mouse PLC-zeta cRNA.

Comparison of the mean interspike interval of Ca^{2+} oscillations in mouse oocytes triggered by the three species of PLC-zeta cRNA. Human, simian and mouse PLC-zeta cRNAs each triggered Ca^{2+} oscillations within 2 hours of microinjection of 200-2.0 $\mu\text{g/ml}$ PLC-zeta cRNA. Only hPLC-zeta was effective at the lower doses of 0.2
5 and 0.02 $\mu\text{g/ml}$. The number of oocytes microinjected is shown above each dose. The mean interspike interval at each dose for human, simian and mouse is statistically different from each other using a students paired t-test ($p = <0.005$).

Table 2 shows the genomic organization of the human PLC-zeta gene. The gene is
10 localized to chromosome 12p12.3

EXAMPLE 1: Isolation of the nucleic acid/protein by PCR cloning (Human)

The human expressed sequence tag (EST) database at NCBI (National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20891, U.S.A.) was searched using the BLAST algorithm for phospho-inositide-specific phospholipase C-related sequences using the published sequence of the rat phospholipase C delta 4 isoform (NCBI accession number U16655 –). Of the numerous positive 'hits' that were obtained, a class of novel ESTs was observed to be derived from human testis cDNA (*eg* accession numbers AI217888; AA707583; AA861064; AA609626).

Using the same approach for database searching as above, the mouse EST database at NCBI gave a related class of novel ESTs derived from mouse testis cDNA (*eg* accession numbers AV257260, AV277909, AV273316, and AV277562).

All these ESTs represent partial testis cDNA sequences (comprising fewer than 400 base pairs), as a complete open reading frame (ORF) was not identified in any of them.

Using polymerase chain reaction (PCR) cloning techniques with specific oligonucleotides designed to amplify sequences related to those ESTs described above, the complete protein coding sequence of the human and mouse phospholipase C-zeta; PLC ζ were obtained as follows:

The primers used for PCR from a human testis cDNA library (Clontech Laboratories 1020 East Meadow Circle, Palo Alto, CA94303-4230, U.S.A. #HL5503u) were:

Forward human primer: 5' CAG CGA GCT CTT ATC TGA CGT ACC AAA C 3'
(28mer)

Reverse TriplEx primer: 5' CTC GGG AAG CGC GCC ATT GTG TTG GT 3'
(26mer)

5 The forward primer was derived from the human EST sequences and included the
predicted stop codon TGA, underlined. The reverse primer encoded the Clontech
lambda TriplEx2 vector sequence. PCR was performed in a 50uL reaction volume
with initial denaturation at 96 ° C for 3 minutes, followed by 30 cycles of 94 ° C for 30
seconds, 60 ° C for 30 seconds and 72 ° C for 3 minutes, and a final extension at 72 ° C
for 5 minutes. The single ~2 kilobase product amplified using these primers with Pfu
10 DNA polymerase, according to manufacturer's instructions (Promega Corporation
catalogue # M7745, Promega UK Ltd, Delta House, Chilworth Research Centre,
Southampton SO16 7NS, U.K.), was cloned into the commercial vector pTOPO-Blunt
and plasmids transformed into competent *E. coli* for plasmid DNA preparation
according to manufacturer's instructions (Invitrogen Inc. catalogue no. K2800-20,
15 Invitrogen BV, PO Box 2312, 9704 CH Groningen, The Netherlands). Plasmid DNA
was isolated from *E. coli* cultures using Qiagen miniprep purification columns
according to manufacturer's instructions (Qiagen cat. no. 12125, QIAGEN Ltd. – UK,
Boundary Court, Gatwick Road, Crawley, West Sussex, RH10 9AX, U.K.).

20 The primers used for PCR from a mouse spermatid cDNA library (made using the
lambda ZAP II vector (available from Stratagene Inc. 11011 North Torrey Pines
Road, La Jolla, CA 92037, U.S.A.), provided by Dr. Paul Burgoyne, National Institute
for Medical Research, London) were:

25 Forward mouse primer: 5' GCT AAC GCG TCA GTT ACA TGC GTC ACT C 3'
(28mer)

Reverse T7 primer: 5' GTA ATA CGA CTC ACT ATA GGG C 3' (22mer)

30 The forward primer was derived from the mouse EST sequences and included the
predicted stop codon TCA, underlined. The reverse primer encoded Stratagene

lambda ZAP II vector sequence (T7 sequence). PCR was performed in a 50uL reaction volume with initial denaturation at 96 °C for 3 minutes followed by 30 cycles of 94 °C for 30 seconds, 60 °C for 30 seconds and 72 °C for 3 minutes, and a final extension at 72 °C for 5 minutes. The single ~2 kilobase product amplified using
5 these primers with Pfu DNA polymerase, according to manufacturer's instructions (Promega Corp.), was cloned into the commercial vector pTOPO-Blunt and plasmids transformed into competent *E.coli* for plasmid DNA preparation according to manufacturer's instructions (Invitrogen Inc.). Plasmid DNA was isolated from *E.coli* cultures using Qiagen Miniprep™ purification columns according to manufacturer's
10 instructions (Qiagen).

Nucleotide sequence analysis of the amplified and cloned human and mouse DNAs was determined by standard dideoxy sequencing performed on an Applied Biosystems ABI377 automated DNA sequencer using the dRhodamine dye terminator kit (PE
15 Applied Biosystems, Kelvin Close, Birchwood Science Park North, Warrington, WA3 7PB, U.K.). Open reading frame (ORF) analysis of the complete human and mouse nucleotide sequences using MacVector sequence analysis software (Oxford Molecular, The Medawar Centre, Oxford Science Park, Oxford, OX4 4GA, U.K.) revealed the complete protein coding sequence of the human and mouse PLC-zeta;
20 PLC proteins. The human sequence revealed an ORF of 1824 base pairs encoding a 608 amino acid sequence (SEQ ID NO: 1). The mouse sequence revealed an ORF of 1941 base pairs encoding a 647 amino acid sequence (SEQ ID NO: 2).

Identification and cloning of simian PLC-zeta

25 A cynomolgus monkey cDNA library was prepared from size-selected, adult *Macaca fascicularis* testes cDNAs of >1.5 kb, and a number of novel, full-length insert DNA sequences were determined. Blast searching with the hPLC-zeta sequence revealed two homologous simian sequences derived from the adult *M. fascicularis* testis cDNA library (Accession numbers, AB070108 and AB070109). The ORF
30 within these two cynomolgusmonkey cDNA clones were amplified by PCR with *Pfu* DNA polymerase, as described above, cloned into pcDNA3.1-V5-His-TOPO

(Invitrogen) (pcDNA-zeta) and the insert DNA sequenced along both strands, as described above. Homology sequence analysis and alignment was performed using ClustalW (available via the internet) and domain structure by RPS-Blast.

5

EXAMPLE 2: Preparation of recombinant vectors for expression in mammalian cells

The complete ORF of both human and mouse PLC-zeta; PLC ζ sequences were sub-
 10 cloned into the mammalian expression vector, pTarget (Promega, Delta House, Chilworth Research Centre, Southampton SO16 7NS, U.K.). The full-length sequences were amplified by PCR with Pfu polymerase (Promega) as described above, using specific oligonucleotides designed to the start and stop codons as follows:

15 The human primers used were:

Forward human primer: 5' CAG CGA GCT CTT ATC TGA CGT ACC AAA C 3'
 (28mer)

20 Reverse human primer: 5' ATG AAA CTA TGG AAA TGA GAT GGT 3' (24mer)

The reverse human primer included the start codon, ATG, underlined, and the forward human primer included the stop codon as used in the original PCR cloning steps described above. PCR was performed as described above. The ~1.8 kilobase product
 25 was cloned into pTOPO-Blunt and the DNA insert was sequenced as described above. The ~1.8 kilobase human DNA insert was excised from the pTOPO-Blunt vector by digestion with the restriction enzyme *EcoRI*, the restricted fragment was separated by agarose gel electrophoresis, purified using the Qiagen DNA gel extraction kit and ligated into the *EcoRI* pre-digested mammalian vector, pTarget. Ligation was
 30 performed at 12 °C overnight in the presence of 10 units of T4 DNA ligase (Promega), and ligated plasmid was transformed into competent *E.coli* XL-1 Blue

(Stratagene), and plasmid DNA purified using Qiagen columns as described above. Restriction enzyme digestion of plasmid DNA revealed the clones containing the correct orientation of the human PLC-zeta; PLC insert.

5 The mouse primers used were:

Forward mouse primer: 5' GCT AAC GCG TCA GTT ACA TGC GTC ACT C 3'
(28mer)

10 Reverse mouse primer: 5' ATC ATG GAA AGC CAA CTT C 3' (19mer)

The reverse mouse primer included the start codon, ATG, underlined, and the forward mouse primer included the stop codon as used in the original PCR cloning steps described above. PCR was performed as described above. The ~1.9 kilobase product
15 was cloned into pTOPO-Blunt and the DNA insert was sequenced as described above. The ~1.9 kilobase mouse DNA insert was excised from the pTOPO-Blunt vector by digestion with the restriction enzyme *EcoRI*, the restricted fragment was separated by agarose gel electrophoresis, purified using the Qiagen DNA gel extraction kit and ligated into the *EcoRI* pre-digested mammalian vector, pTarget. Ligation was
20 performed at 12 °C overnight in the presence of 10 units of T4 DNA ligase (Promega), and ligated plasmid was transformed into competent *E.coli* XL-1 Blue (Stratagene), and plasmid DNA purified using Qiagen columns as described above. Restriction enzyme digestion of plasmid DNA revealed the clones containing the correct orientation of the mouse PLC-zeta; PLCζ insert.

25

EXAMPLE 3: Transfection of human and mouse expression plasmids into CHO cells

30 The human and mouse pTarget/ PLC expression plasmid DNAs prepared as described in Example 2 were separately introduced, by a lipid-mediated transfection

procedure, into the Chinese hamster ovary (CHO) cell line grown in tissue culture. CHO cells cultured in serum-containing media, DMEM, (Dulbecco's Modified Eagle Medium) to a density of 500,000 cells per culture dish, were transfected with 40µg plasmid DNA plus 40uL of Lipofectamine2000 (Life Technologies Ltd, 3 Fountain Drive, Inchinnan Business Park, Paisley, U.K.) in serum-free DMEM. After 15 hours, the CHO cells were returned to serum-containing DMEM.

In parallel, control experiments, identical CHO cells were treated in the same way with Lipofectamine but in the absence of plasmid DNA.

EXAMPLE 4: Demonstration of effectiveness – PLC in CHO Cells

Transfected cells prepared according to Example 3 were washed with culture medium 30 minutes after transfection, then incubated with the calcium-sensitive fluorescent indicator, fura-2-AM for 60 minutes. After further washing with medium, the cells were then placed on a microscope stage and the changes in cell calcium levels, as detected by the fluorescence of the fura-2, were monitored. Only in cells transfected with the PLC ζ expression plasmid, the cell calcium level was observed to change periodically. This specific temporal behaviour of cell calcium, *ie* to produce calcium oscillations, is the same as that observed in eggs when fused with sperm at fertilization, and when soluble sperm proteins are injected directly into eggs. Figure 1 demonstrates this with respect to mouse PLC ζ . This indicates that the novel PLC ζ proteins we have identified in human and mouse testis may be used to specifically control cell calcium levels in mammalian cells.

EXAMPLE 5: Demonstration of effectiveness – PLC in Oocytes

The open reading frame of human and mouse PLC ζ cloned into the pTargetT vector as described in Example 2, were linearised by restriction, and complementary RNA

(cRNA) encoding PLC ζ was synthesized with a Ribomax RNA synthesis kit (Promega) then re-suspended in 120mM KCl, 20mM HEPES, pH 7.4. Mouse oocytes arrested at MII stage were harvested from female mice and loaded with fura 2-AM for 10 minutes, washed in H-KSOM and placed on a Nikon Diaphot stage. cRNA was
 5 micro-injected to 3-5% of egg volume and calcium was monitored as described by Swann, K in *Development* 110 1295-1302 (1990).

Figure 2 demonstrates that mouse PLC ζ in mouse eggs causes calcium oscillations. The data from the experiments of Examples 4 and 5 clearly show that PLC ζ should
 10 have the effect of causing calcium oscillations in any cell type.

EXAMPLE 6: Further Evidence that PLC-Zeta is Sperm Factor

15 Complementary RNA synthesis and *in vitro* translation

The 1941bp open reading frame of mouse PLC ζ was cloned into pCR-Blunt II-TOPO, sequenced, and subcloned (pTarget, Promega) to generate pTarget-mPLC ζ . Complementary RNA (cRNA) was synthesized from linearised pTarget-mPLC ζ (Ribomax RNA synthesis, Promega) in the presence of 3 mM m⁷G(5')ppp(5')G, isopropanol precipitated and resuspended in DEPC-treated water containing 4 U/ul
 20 RNasin (Promega). Mutagenesis of ²¹⁰Asp to ²¹⁰Arg in PLC ζ to produce D210R_{PLC ζ} was achieved using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). Constructs and cRNAs for rat PLC δ 1 and Δ PH_{PLC δ 1}, which encoded the full-length (756 amino acids) and PH domain-deleted PLC δ 1 (Δ 1-132),
 25 respectively, and D210R_{PLC ζ} were produced in pTarget as above. cRNA (2 ug) was expressed *in vitro* (Reticulocyte lysate system, Promega) in the presence of [³⁵S]methionine (Amersham Pharmacia). Radiolabelled protein, analysed by SDS-PAGE and autoradiography, was displayed using QuantityOne software (BioRad).

c-Myc-Epitope tagging, bacterial expression and PLC ζ quantitation

The 1941 bp open reading frame of mouse PLC ζ was subcloned into pGBK-T7 (Clontech) with an in-frame c-Myc epitope tag at the 5'-end (Lopez *et al* J Biol Chem 276 2758-2765 (2001)). The c-Myc-PLC ζ was further subcloned into pcDNA3.1 and
5 sequence-verified before cRNA synthesis from the T7 site (Ribomax) for egg micro-injection, as described above. For bacterial expression, c-Myc-PLC ζ was subcloned into pBAD (Invitrogen) with an in-frame hexahistidine tag at the 3' end. The c-Myc-PLC ζ -Histag protein was produced in 0.2% w/v arabinose-induced, BL21(DE3)pLysS *E.coli*, after extraction of the pelleted bacteria by five freeze-thaw
10 and ultrasonication cycles, then purified by nickel affinity chromatography (ProBond, Invitrogen). Protein quantitation was performed using the BCA protein assay (Pierce) Densitometric analysis of the c-Myc-PLC ζ band expressed in eggs micro-injected with different cRNA concentrations, c-Myc-PLC ζ -Histag protein purified from *E.coli*, and calibrated sperm extract PLC ζ derived from 10^4 - 10^6 mouse sperm,
15 employed a c-Myc monoclonal antibody (1:2000, Santa Cruz Biotechnology) and rabbit anti-PLC ζ antiserum (1:1000), respectively, using QuantityOne software (BioRad). A calibration standard plot, from analysis by immunoblot densitometry (Malek *et al* Biotechniques 6 1150-1153 (1997)) using the c-Myc antibody, was constructed using defined amounts of c-Myc-PLC ζ -Histag protein, purified from
20 *E.coli*, to enable the calculation of the relative c-Myc-PLC ζ content in batches of 100 micro-injected eggs. For the quantitation analysis, expression of the c-Myc-PLC ζ protein was assumed to be linear with time after cRNA micro-injection, as has been shown for micro-injected EGFP cRNA expressed in mouse eggs. This assumption was necessary because the c-Myc-PLC ζ protein was below the detection limit within
25 3 hours of cRNA micro-injection. Hence, for a single mouse egg, the calculated 440-750 fg of c-Myc-PLC ζ protein expressed 5 hours after micro-injection with 0.02 mg/ml cRNA, was equivalent to 44-75 fg expressed at 0.5 hours) (the time when the first Ca²⁺ transient is normally observed). A separate calibration plot using the anti-PLC ζ antibody was constructed with different c-Myc-PLC ζ -Histag protein

concentrations to enable estimation of the relative PLC ζ content in defined numbers of mouse sperm.

Results are given in sections (a) to (c) below.

5

Immunodepletion of PLC ζ from sperm extracts

Soluble extracts (Parrington *et al* Biochem J 341 1-4 (1999)) prepared from hamster sperm were incubated for 1 hour at 4°C with control IgG or anti-PLC ζ antibody that had been covalently attached to Protein G beads (1 mg/ml, Seize X Kit, Pierce). The PLC ζ content of the supernatant and precipitated beads was determined by immunoblot analysis with anti-PLC ζ antibody. Antibody-treated sperm supernatants were also analysed for Ca²⁺ release activity by fluo-3 fluorometry with sea urchin egg homogenates, monitored using a Perkin-Elmer LS50B fluorimeter (as described by Jones *et al* in FEBS Letts 437 297-300 (1998)). They were also analysed for ability to generate CCOs by micro-injection into mouse eggs, as described below. Maximal immunodepletion of the sperm PLC ζ protein was achieved by using an optimised ratio of antibody beads to sperm extract for each experiment (n=4). The optimal ratio was empirically determined for each sperm extract preparation as the minimum concentration of sperm extract (0.3-0.8 mg/ml) that still retains Ca²⁺ release activity after treatment with the control IgG beads.

Results are given in section (d), below.

Preparation and handling of gametes

Mouse egg procedures were carried out either in HEPES-buffered KSOM or amino acid supplemented KSOM (Summers *et al* Human Reprod 15 1791-1801 (2000)). Female MF1 mice were super-ovulated by injection with 5 IU of PMSG followed 48 hours later by HCG (Intervet). Eggs were collected 13.5-14.5 hours after HCG, maintained in 100 μ l droplets of H-KSOM under mineral oil at 37°C and cRNA micro-injections performed within 1 hour. Expression of c-Myc-PLC ζ in eggs was examined 5 hours after cRNA micro-injection, by adding SDS sample buffer to

pelleted eggs and incubating at 95°C for 5 minutes prior to SDS-PAGE, immunoblot
 then densitometric analysis with the c-Myc monoclonal antibody, as described above.
 Calibrated mouse sperm pellets were re-suspended in 10 mM Tris-HCl pH 7.5, 15
 mM dithiothreitol (Perry *et al* Biol Reprod 60 747-755 (1999)), then subjected to 5
 5 freeze-thaw cycles in liquid N₂ and centrifuged at 20,000 x g at 4°C for 10 minutes,
 before densitometric analysis of the soluble extract with PLCζ antibody, as described
 above. For *in vitro* fertilization studies, sperm were capacitated for 2-3 hours before
 adding to eggs. Egg activation and development studies were in H-KSOM containing
 2 μM cytochalasin D for 4 hours. Further development to 2-cell stage, morula and
 10 blastocyst stage was carried out in 50 μl droplets of KSOM under mineral oil at 37°C
 in a 5% CO₂ incubator.

Measurement of intracellular Ca²⁺ in MII-arrested mouse eggs

Eggs loaded with 4 μM Fura red-AM (Molecular Probes) for 10 minutes were washed
 15 in H-KSOM and placed on a Nikon Diaphot stage. Loading media included
 sulfonpyrazone to prevent dye compartmentalisation and extrusion (Lawrence *et al*
 Development 124 223-241 (1997)). cRNA solutions in 120 mM KCl, 20 mM HEPES,
 pH 7.4, were micro-injected to 3-5% of egg volume as previously described (Swann
 1990, *ibid*, Example 5). Protein synthesis was inhibited in control experiments where
 20 eggs were pre-incubated in solution containing 10 μM cycloheximide for 30 minutes
 prior to micro-injection with PLCζ cRNA (0.02 mg/ml; n=9). Injection volume was
 estimated from the displacement caused by bolus injection. Ca²⁺ measurements were
 performed on a CCD-based imaging system as previously described (Lawrence *et al*,
 1997; *ibid*), or a Zeiss Axiovert 100 with illumination from a monochromator
 25 (Photonics) controlled by MetaFluor v4.0 (Universal Imaging Corp).

Results

(a) PLCζ triggers Ca²⁺ oscillations in eggs

The defining character of the mammalian sperm factor is the ability to elicit CCOs
 30 that mimic the fertilization-associated transients displayed by mammalian eggs. To
 examine whether sperm PLCζ could trigger such CCOs, we introduced PLCζ

complementary RNA (cRNA) by micro-injection into MII-arrested mouse eggs, as described previously for spermatogenic cell mRNA. Eggs micro-injected with a pipette concentration of 2 mg/ml PLC ζ cRNA, corresponding to <0.1 mg/ml in the egg after a 3-5% injection volume, underwent a prolonged series of CCOs, similar to those shown in Figure 2, that commence within 15-20 minutes. The high oscillation frequency is similar to that observed upon micro-injection of concentrated sperm extracts into mouse eggs. CCOs of similar amplitude, but lower frequency, were obtained with a 1000-fold dilution to 0.002 mg/ml PLC ζ cRNA (Figure 6a, middle trace; 0.0001 mg/ml in egg). None of the eggs treated with cycloheximide to block protein synthesis showed any Ca²⁺ transients after PLC ζ cRNA-micro-injection (0.02 mg/ml, n=9; Fig. 6a, bottom trace). Robust CCOs were observed in 100% of the eggs micro-injected with the four different PLC ζ cRNA concentrations tested, ranging from 0.002-2 mg/ml (Figure 6b). Importantly, the frequency, but not the amplitude, of CCOs varied with PLC ζ cRNA concentration, directly matching the same phenomenon observed with different concentrations of sperm extract. The highest pipette concentration used, 2 mg/ml, produced CCOs with a mean interspike interval of 7.3 $\pm$ 3.2 minutes (Figure 6b). The lowest pipette concentration of PLC ζ cRNA that gives oscillations within 2 hours of injection (0.002 mg/ml), displayed a mean interspike interval of 20.1 $\pm$ 5.4 minutes (Figure 6b). Both of these values are significantly different to the mean interspike interval produced with in vitro fertilization (IVF) of mouse eggs (12.1 $\pm$ 5.8 minutes). However, the interspike intervals for 0.2 and 0.02 mg/ml PLC ζ cRNA (13.6 $\pm$ 3.2 and 12.7 $\pm$ 6.0 minutes, respectively) are not significantly different from IVF.

25 (b) *Fertilization-like Ca²⁺ signals via PLC ζ*

The CCOs at fertilization display some unique features. The first Ca²⁺ transient invariably lasts longer than subsequent oscillations, and exhibits a set of intriguing, smaller sinusoidal increases on top of the main peak. Micro-injection of a pipette concentration of PLC ζ cRNA that produces an interspike interval matching IVF (i.e. 0.02 mg/ml; Figure 6b), resulted not only in the same, longer initial Ca²⁺ transient,

but also displayed a similar pattern of smaller sinusoidal increases. The first Ca^{2+} increase after 0.02 mg/ml PLC ζ cRNA micro-injection matches the first IVF transient in both average duration (PLC ζ 2.8 ± 0.6 minutes, n=39 versus IVF 3.0 ± 0.7 minutes, n=16), and also in reproducibly producing the cluster of smaller Ca^{2+} increases superimposed on the first transient. A concentration of 0.02 mg/ml PLC ζ cRNA was used for subsequent micro-injection experiments, unless stated otherwise, to provide the precise Ca^{2+} signaling conditions that are stereotypical of fertilization.

(c) *Physiological level of PLC ζ in a single sperm*

In order to quantitate the PLC ζ expressed in micro-injected eggs, a c-Myc epitope tag was introduced at the N-terminus of PLC ζ , as described above. Micro-injected c-Myc-PLC ζ cRNA at different concentrations was as effective at generating Ca^{2+} oscillations in eggs as the untagged PLC ζ , indicating that the N-terminal attachment of the c-Myc tag is not deleterious to PLC ζ activity, as was shown for c-Myc-PLC.

Furthermore, the c-Myc-PLC ζ protein expressed in eggs was readily detected in immunoblots using an anti-c-Myc monoclonal antibody, as a single band with the predicted mass of 78 kDa, whereas uninjected eggs exhibited no immunoreactivity. Comparison of the relative mobility of native mouse sperm PLC ζ (74 kDa) and recombinant c-Myc-PLC ζ protein (78 kDa [74 kDa PLC ζ + 4 kDa c-Myc tag]) indicated that the deduced ORF of the PLC ζ cDNA clone ([SEQ ID NO: 2], 74 kDa) represents the complete sperm PLC ζ sequence. Densitometric analysis of the immunoreactive 78 kDa c-Myc-PLC ζ protein expressed in eggs, compared with calibrated amounts of purified recombinant c-Myc-PLC ζ protein produced in bacteria, enabled the determination of 44-75 fg/egg (n=4) as the amount of PLC ζ protein that triggers Ca^{2+} oscillations using 0.02 mg/ml cRNA. This cRNA concentration is the one that most closely mimics the IVF response, though ten-fold lower levels (i.e. 4-8 fg PLC ζ protein/egg using 0.002 mg/ml cRNA) are also able to cause Ca^{2+} oscillations (Figure 6).

The PLC ζ content of sperm was also determined by densitometry with a PLC ζ polyclonal antibody using a defined number of mouse sperm and compared with calibrated amounts of recombinant PLC ζ protein. Using densitometric values within the recombinant PLC ζ protein calibration plot, obtained from samples comprising
5 10^4 - 10^6 mouse sperm, a single mouse sperm was calculated to contain 20-50 fg PLC ζ protein (n=4). The level of PLC ζ able to produce Ca $^{2+}$ oscillations in a single egg similar to fertilization (4-75 fg, i.e. with 0.002-0.02 mg/ml cRNA) is therefore in the same range as the single sperm content of PLC ζ (20-50 fg). The observed quantitative correlation indicates that the PLC ζ from a single sperm is sufficient to produce the
10 Ca $^{2+}$ oscillations observed upon sperm-egg fusion.

(d) *Sperm PLC ζ depletion abrogates Ca $^{2+}$ oscillations*

To address whether the PLC ζ in sperm is uniquely responsible for Ca $^{2+}$ mobilisation in eggs, the PLC ζ content of sperm extracts was specifically depleted using an anti-
15 PLC ζ antibody, as described above. Immunoblot analysis indicated that sperm extract supernatant retains the PLC ζ protein after control antibody treatment, in contrast to PLC ζ antibody-treated supernatant where the PLC ζ is absent. Analysis of the corresponding precipitated antibody samples revealed that the sperm PLC ζ is effectively removed by PLC ζ antibody, but not by the control antibody. Assessment
20 of Ca $^{2+}$ release activity in antibody-treated sperm extracts using sea urchin egg homogenate assays showed that PLC ζ -depleted samples lack any Ca $^{2+}$ mobilising activity, whereas a robust Ca $^{2+}$ release was observed with the control antibody-treated sperm extract containing PLC ζ protein. Moreover, micro-injection of antibody-treated sperm extracts into mouse eggs illustrated that the ability of
25 untreated samples to generate IVF-like Ca $^{2+}$ oscillations is fully preserved in control antibody-treated samples, while PLC ζ -depletion effectively abrogates Ca $^{2+}$ release activity.

These PLC ζ antibody depletion experiments (n=4) suggest that PLC ζ is the sole component of sperm extracts possessing the ability to cause Ca²⁺ release in mouse eggs. Taken together with evidence that the PLC ζ level in a single mouse sperm is sufficient to trigger IVF-like Ca²⁺ oscillations in a single mouse egg, the
5 immunodepletion data provides compelling evidence that PLC ζ is synonymous with the previously described mammalian sperm factor.

(e) *PLC ζ activates normal embryo development*

Since eggs that were micro-injected with PLC ζ cRNA (0.02mg/ml) displayed all the
10 properties of Ca²⁺ oscillations indistinguishable from those of IVF (Results (a) and (b) above) and is equivalent to the PLC ζ content of a single sperm ((c) above), their ongoing development was monitored for several days after PLC ζ -micro-injection. PLC ζ -micro-injected eggs underwent activation (Figure 4a) because normal development proceeded to the 2-cell stage within 24 hours (78%, n=147), and many
15 reached the morula or blastocyst stages by 4-5 days (62%, n=76). None of the eggs micro-injected with buffer control reached the 2-cell stage, indicating activation as an artefact of micro-injection procedure did not occur. The proportion of PLC ζ -induced embryos that developed to either the 2-cell, or morula and blastocyst stages, was the same as for eggs that are either parthenogenetically activated by strontium ions
20 (n=75), or when embryos are collected at the 1-cell stage from female mice after in vivo fertilization (n=101) upon mating with males (Figure 4a).

Photomicrographs taken at 24 hours and 5 days after PLC ζ -micro-injection into mouse eggs show the appearance of normal embryo development to the 2-cell stage
25 and blastocyst stage (left and right panel, respectively, Figure 4b). There were no morphological differences to embryos obtained after fertilization with sperm. Thus, after inducing Ca²⁺ oscillations in the egg, sperm PLC ζ -micro-injection also triggered the entire cascade of events required for activation and embryo development, in the same manner as sperm at fertilization.

30

The possibility remained that a novel action of PLC ζ other than PIP₂ hydrolysis is responsible for egg activation, such as a protein-protein interaction with a distinct egg molecule. To test whether an enzymatically active PLC ζ is required for egg activation and embryo development, the D210RPLC ζ cRNA (0.02mg/ml), which was shown to
 5 be defective in triggering Ca²⁺ oscillations, was micro-injected, and egg activation assessed after 24 hours. None of the D210RPLC ζ cRNA-micro-injected eggs were found to proceed to the pronuclear or 2-cell stage (Figure 5, n=20), suggesting that the enzymatic function of sperm PLC ζ is critical for egg activation.

10 *Human PLC-zeta triggers Ca oscillations in mouse oocytes*

To examine the ability of hPLC-zeta to cause Ca²⁺ changes, cRNA for hPLC-zeta was microinjected into MII-arrested mouse oocytes with a pipette concentration of 20 μ g/ml hPLC-zeta cRNA, which corresponds to 0.001 mg/ml in the oocyte after a 3-5% injection volume. Figure 8A shows a representative example Ca²⁺ recording for
 15 each of the four different concentrations of hPLC-zeta cRNA that were microinjected. At 20 μ g/ml hPLC-zeta cRNA triggered high frequency Ca²⁺ oscillations within 10-15 minutes of microinjection (mean interspike interval: 4.21 ± 1.79 mins). As was observed with mouse PLC-zeta cRNA and hamster sperm extract microinjection, Ca²⁺ oscillations of lower frequency were obtained with lower concentrations of
 20 stimulus (Swann, 1990; Saunders et al, 2002). It was notable that even at pipette concentrations of 0.02 μ g/ml, hPLC-zeta cRNA could still induce Ca²⁺ oscillations within two hours of microinjection. Although a wide range of cRNA concentrations from 20-0.02 μ g/ml were used, the Ca²⁺ oscillations observed at each concentration lasted for a similar period of 3-4 hours (Fig. 8A). The mean interspike interval data
 25 showing the dose-response relationship with hPLC-zeta cRNA is summarised in the histogram in Figure 8B.

Embryo development with hPLC -zeta

The microinjection of 20 μ g/ml mPLC-zeta into mouse oocytes was previously
 30 demonstrated to induce Ca²⁺ oscillations and development to the blastocyst stage at rates comparable to that of *in vitro* fertilization . To examine if hPLC-zeta is also able

to support development, and what effect the oscillation frequency might have on embryo development, MII-arrested oocytes were injected with 20, 2.0 and 0.2 $\mu\text{g/ml}$ hPLC-zeta cRNA and monitored after 24h and 96h. All three concentrations were effective at activating the oocytes and enabling development to the 2-cell stage (Fig. 9). Using 2.0 and 0.2 $\mu\text{g/ml}$ hPLC-zeta cRNA, mouse embryo development to morula/blastocyst was 33.3 and 38.9%, respectively (Fig. 9A). This compares with developmental rates with *in vivo* fertilization and parthenogenetic activation of 55-60% under our conditions using outbred mouse strains. It was conspicuous, however, that the high Ca^{2+} oscillation frequency (low mean interspike interval) produced with 20 $\mu\text{g/ml}$ was ineffective at supporting development to morula/blastocyst stages (1.8% of oocytes reaching morula/blastocyst) and most of these embryos arrested at the 2-cell stage.

Micrographs of the mouse embryos produced by hPLC-zeta cRNA microinjection show they are morphologically similar to those following *in vitro* fertilization (Fig. 9B), analogous to the observations with mPLC-zeta, though blastocyst cell numbers have not been determined. These data suggest that microinjection of hPLC-zeta cRNA into unfertilized eggs alone can trigger early embryonic development to blastocyst stages in mouse embryos, but it appears that the high frequency of Ca^{2+} oscillations caused by the higher doses of hPLC-zeta is detrimental to development beyond the 2-cell stage.

Simian PLC-zeta triggers Ca^{2+} oscillations in mouse oocytes.

The observations described above (Figs. 8 and 9), show that the human and mouse PLC-zeta can cause fertilization-like Ca^{2+} oscillations that initiates activation and development of mouse oocytes. The identification of two related, testis-specific cDNA sequences of 2.3 kb from *M.*

fascicularis, and the high degree of similarity of their ORF with the human and mouse PLC-zeta, enabled the prediction that these were simian PLC-zeta homologues. We therefore compared the ability of cRNA prepared from the two forms of sPLC-zeta, designated s1PLC-zeta and s2PLC-zeta (AB070108 and AB070109, respectively), to generate Ca^{2+} oscillations in mouse oocytes. Both forms were able to trigger Ca^{2+} oscillations and no functional difference was detected upon microinjecting either

s1PLC-zeta or s2PLC-zeta cRNA (data not shown). For all subsequent experiments s1PLC-zeta was used (AB070108). Figure 10A shows that s1PLC-zeta cRNA triggered dose-dependent Ca^{2+} oscillations in mouse oocytes comparable to those seen with human and mouse PLC-zeta, at each of the three doses tested (0.2, 0.02, 0.002 mg/ml). Similar to the data with human PLC-zeta, (Fig. 8A), the period over which Ca^{2+} oscillations occurred was 3-4 hours for each of the three s1PLC-zeta cRNA concentrations microinjected. However, the frequency of Ca^{2+} spikes was different for each cRNA concentration, with the mean interspike interval decreasing with higher level of the stimulus (Fig. 10B). This data suggests that PLC-zeta, derived from the sperm/testis of various mammals lacks any species-specificity and, once introduced by microinjection, is able to trigger Ca^{2+} oscillations in heterologous mammalian oocytes. This finding is fully consistent with earlier observations of sperm extracts derived from various sources, including non-mammalian species, each causing Ca^{2+} oscillations in different mammalian oocytes. Figure 11 compares the mean interspike intervals for the three different mammalian forms of PLC-zeta at various pipette cRNA concentrations. Microinjecting cRNA for mPLC-zeta, hPLC-zeta and sPLC-zeta all gave rise to Ca^{2+} oscillations over a range of concentrations from 200 to 2 $\mu\text{g/ml}$. However, hPLC-zeta was distinct in being able to cause Ca^{2+} oscillations at the lower concentrations of 0.2-0.02 $\mu\text{g/ml}$ (Fig 11). This suggests that under the same experimental conditions, the human form of PLC-zeta is more effective at generating Ca^{2+} oscillations in mouse oocytes than the PLC-zeta from mouse and monkey.

In addition to demonstrating that hPLC-zeta and sPLC-zeta are able to cause Ca^{2+} oscillations in mouse oocytes (Figs. 8 and 10), we obtained empirical evidence that hPLC-zeta is more effective at causing Ca^{2+} oscillations than sPLC-zeta and mPLC-zeta (Fig. 11). The minimal amount of hPLC-zeta cRNA required to trigger Ca^{2+} oscillations was 1-2 orders of magnitude lower (0.2-0.02 $\mu\text{g/ml}$) than the minimally effective dose of mouse or simian PLC-zeta cRNA (2 $\mu\text{g/ml}$). These differences were observed as a consistent feature with different batches of cRNA that were each tested for expression *in vitro* (data not shown). The superior potency of

hPLC-zeta cRNA is therefore likely to represent a genuine feature of the hPLC-zeta protein. Thus, we could predict that there is at least an order of magnitude difference in the sensitivity of mouse oocytes to hPLC-zeta compared with mPLC-zeta. It is not clear why hPLC-zeta exhibits greater virility than mPLC-zeta or sPLC-zeta.

5 Subsequent to stimulating Ca^{2+} oscillations in mouse oocytes, the human PLC-zeta was also able to trigger development of embryos to the blastocyst stage (Fig. 8). This suggests that hPLC-zeta is able to produce all of the normal events of oocyte activation. However, one feature of the greater efficacy of hPLC-zeta is that high cRNA levels caused very high frequency Ca^{2+} oscillations in mouse oocytes (Fig.
10 8A, top trace). At concentrations of cRNA that resulted in Ca^{2+} oscillations of ~ 1 spike every 5 minutes, hPLC-zeta was able to effect oocyte activation, but the embryos arrested at the 2-cell stage (Fig. 9A). Previous studies have shown that high frequency Ca^{2+} oscillations may either lead to apoptosis of oocytes, or to developmental changes in postimplantation embryos. Our data provides the first
15 indication that high frequency Ca^{2+} oscillations can also activate an oocyte, but this non-physiological stimulus leads to arrest during the early cleavage stages.

In conclusion, we have herein disclosed and characterized the zeta isoform of phospholipase C and so elucidated a key trigger in the fertilization process.

20

Table 2.Genomic organisation of the human *plc-zeta* gene

Sequence coordinates and length of exons and introns comprising the human *plc-zeta* gene localised to chromosome 12p12.3

Chromosome 12 coordinates	Exon Number	Intron Number	Length (basepairs)
4443338 - 4443286	1		53
4443285 - 4442862		1	424
4442861 - 4442712	2		150
4442711 - 4441697		2	1015
4441696 - 4441564	3		133
4441563 - 4428887		3	12677
4428886 - 4428667	4		220
4428666 - 4424987		4	3680
4424986 - 4424784	5		203
4424783 - 4418341		5	6443
4418340 - 4418195	6		146
4418194 - 4410669		6	7526
4410668 - 4410519	7		150
4410518 - 4407131		7	3388
4407130 - 4407045	8		86
4407044 - 4406922		8	123
4406921 - 4406853	9		69
4406852 - 4405305		9	1548
4405304 - 4405151	10		154
4405150 - 4401621		10	3530
4401620 - 4401503	11		118
4401502 - 4400434		11	1069
4400433 - 4400263	12		171
4400262 - 4393572		12	6691
4393571 - 4393443	13		129
4393442 - 4389634		13	3809
4389633 - 4389484	14		150
4389483 - 4388679		14	805
4388678 - 4388535	15		144

JF WCM 83 PCT Sequences.

SEQ ID NO:1 - Human PLC-zeta; PLCζ amino acid sequence, 608 residues

MEMRWFLSKIQDDFRGGKINLEKTQRLLEKLDIRCSYIHVKQIFKDNDRLKQGRITIEEFRA
IYRIITHREEIIEIFNTYSENKILLASNLAQFLTQEQYAAEMSKAIAFEIIQKYEPIEEVRKAH
QMSLEGFTRYMDSRECLLFKNECRKVYQDMTHPLNDYFISSSHNTYL VSDQLLGPSDLWG
YVSALVKGCRCLEIDCWDGAQNEPVVYHGYTLTSKLLFKTVIQAIHKYAFMTSDYPVVLS
LENHCSTAQQEVMADNLQATFGESLLSDMLDDFPDTLP SPEALKFKILVKNKKIGTLKETH
ERKGSDKRGDNQDKETGVKKLPGVMLFKKKKTRKLLKIALALSDLVIYTKAEKFKSFQHSR
LYQQFNENNSIGETQARKLSKLRVHEFIFHTRKFITRIYPKATRADSSNFNPQEFWNIGCQM
VALNFQTPGLPMDLQNGKFLDNGGSGYILKPHFLRESKSYFNPSNIKEGMPITLTIRLISGIQ
LPLTHSSSNKGDSLVIIEVFGVPNDQMKQQTRVIKKNAFSPRWNETFTFIIHVPELALIRFVV
EGQGLIAGNEFLGQYTLPLLCMNKGYRRIPLFSRMGESLEPASLFVYVWYVR

JF WCM 83 PCT Sequences.

SEQ ID NO:2 - Mouse PLC-zeta; PLCζ amino acid sequence, total 647 residues

MESQLHELAEARWFLSKVQDDFRGGKINVEITHKLLEKLDFPCHFAHVKHIFKENDRQNQ
GRITIEEFRAIYRCIVHREEITEIFNTYTENRKILSENSLIEFLTQEQYEMEIDHSDSVEIINKYE
PIEEVKGERQMSIEGFARYMFSSSECLLFKENCKTVYQDMNHPLSDYFISSSHNTYLISDQIL
GPSDIWGYVSALVKGCRCLEIDCWDGSQLNEPIVYHGYTFTSKLLFKTVVQAINKYAFVTSD
YPVVLLENHCSPGQQEVMASILQSTFGDFLLSDMLEEFPDTLPSPEALKFKILVKNRKVGT
LSEETHERIGTDKSGQVLEWKEVTYEDGDEDSGMDPETWDVFLSRIKEEREADPSTLSGIAG
VKKRKRKMKIAMALSDLVIYTKAEKFRNFQYSRVYQQFNETNSIGESRARKLSKLRVHEFI
FHTAAFITRVYPKMMRADSSNFNPQEFWNVGCQMVALNFQTPGLPMDLQNGKFLDNNGGS
GYILKPDILRDTTLGFNPNEPEYDDHPVTLTIRIISGIQLPVSSSSNTPDIVVIEVYGVPNDHV
KQQTRVVKNNAFSPKWNETFTFLIQVPELALIRFVVETQQGLLSGNELLGQYTLPLVLCMNK
GYRRVPLFSKSGANLEPSSLFIYVWYFRE

SEQ ID NO:3 - Human PLC-zeta; PLCζ nucleotide sequence, 1827 nucleotides

ATGGAAATGAGATGGTTTTTGTCAAAGATTCAGGATGACTTCAGAGGTGGAAAAATTA
ACCTAGAAAAAACTCAGAGGTTACTTGAAAAATTAGATATTCGGTGCAGTTATATTCA
TGTGAAACAGATTTTTTAAGGACAATGACAGGCTGAAACAAGGAAGAATCACCATAGA
AGAATTTAGAGCAATTTATCGAATTATCACGCACAGAGAAGAAATTATTGAGATTTTC
AACACATATTCTGAAAACCGGAAAATTCTTTTAGCAAGTAATCTGGCTCAATTTCTGAC
ACAAGAACAATATGCAGCTGAGATGAGTAAAGCTATTGCTTTTGAGATCATTCAGAAA
TACGAGCCTATCGAAGAAGTTAGGAAAGCACACCAAATGTCATTAGAAGGTTTTACAA
GATACATGGATTCACGTGAATGTCTACTGTTTAAAAATGAATGTAGAAAAGTTTATCA
AGATATGACTCATCCATTAAATGATTATTTTATTTTCATCTTCACATAACACATATTTGGT
ATCTGATCAATTATTGGGACCAAGTGACCTTTGGGGATATGTAAGTGCCCTTGTGAAA
GGATGCCGTTGTTTTGGAGATTGACTGCTGGGATGGAGCACAAAATGAACCTGTTGTAT
ATCATGGCTACACACTCACAAGCAAACCTTCTGTTTAAAACTGTTATCCAAGCTATACAC
AAGTATGCATTCATGACATCTGACTACCCAGTGGTGCTCTCTTTAGAAAATCACTGCTC
CACTGCCCAACAAGAAGTAATGGCAGACAATTTGCAGGCTACTTTTGGAGAGTCCTTG
CTTTCTGATATGCTTGATGATTTTCCTGATACTCTACCATCACCAGAGGCACTAAAATT
CAAAATATTAGTTAAAAATAAGAAAATAGGAACCTTAAAGGAAACCCATGAAAGAAA
AGGTTCTGATAAGCGTGGAGACAATCAAGACAAGGAAACAGGGGTAAAAAAGTTACC
TGGAGTAATGCTTTTCAAGAAAAAGAAGACCAGGAAGCTAAAAATTGCTCTGGCCTTA
TCTGATCTTGTCATTTATACGAAAGCTGAGAAATTCAAAGCTTTCAACATTCAAGATT
ATATCAGCAATTTAATGAAAATAATTCTATTGGGGAGACACAAGCCCGAAAACCTTTCA
AAATTGCGAGTCCATGAGTTTATTTTTTCACACCAGGAAGTTCATTACCAGAATATATCC
CAAAGCAACAAGAGCAGACTCTTCTAATTTTAATCCCCAAGAATTTTGGAATATAGGT
TGTCAAATGGTGGCTTTAAATTTCCAGACCCCTGGTCTGCCCATGGATCTGCAAAATGG
GAAATTTTTGGATAATGGTGGTTCTGGATATATTTTGAAACCACATTTCTTAAGAGAGA
GTAAATCATACTTTAACCCAAGTAACATAAAAGAGGGTATGCCAATTACACTTACAAT
AAGGCTCATCAGTGGTATCCAGTTGCCTCTTACTCATTCATCTAACAAAGGTGATT
CATTAGTAATTATAGAAGTTTTTGGTGTTCCAAATGATCAAATGAAGCAGCAGACTCG
TGTAATTAAAAAAAATGCTTTTAGTCCAAGATGGAATGAAACATTCACATTTATTATTC
ATGTCCCAGAATTGGCATTGATACGTTTTGTTGTTGAAGGTCAAGGTTTAATAGCAGGA
AATGAATTTCTTGGGCAATATACTTTGCCACTTCTATGCATGAACAAAGGTTATCGTCG
TATTCCTCTGTTTTCCAGAATGGGTGAGAGCCTTGAGCCTGCTTCACTGTTTGTTTATGT
TTGGTACGTCAGATAA

SEQ ID NO: 4 - Mouse - PLC-zeta; PLCζ nucleotide sequence, 1944 nucleotides

ATGGAAAGCCAACTTCATGAGCTCGCAGAAGCAAGATGGTTTTTGTCAAAGGTTTCAGG
ATGATTTTAGAGGTGGAAAAATCAACGTTGAAATTACTCACAACTGCTTGAGAACT
TGATTTCCCATGCCACTTTGCTCATGTGAAACATATTTTAAAGGAAAATGACAGACAGA
ACCAAGGAAGAATCACCATTGAAGAGTTTAGAGCCATTTACCGGTGTATTGTACATAG
AGAAGAAATCACGGAGATTTTCAACACGTATACTGAAAATAGGAAAATTCTTTCTGAG
AACAGTCTGATTGAGTTTCTAACCCAAGAGCAGTATGAAATGGAGATCGATCACTCTG
ATTCAGTAGAGATCATCAATAAGTATGAGCCTATTGAAGAAGTAAAGGGTGAGCGAC
AGATGTCAATTGAAGGTTTCGCAAGATACATGTTTTTCATCAGAATGTCTACTGTTTAAA
GAGAACTGTAAAACCGTGTACCAAGATATGAATCATCCATTAAGTGATTATTTTATTTC
ATCATCTCACAAACACATATTTGATATCCGATCAAATATTGGGACCGAGTGACATTTGG
GGATATCTAAGTGCTCTTGTGAAAGGCTGCCGCTGTCTGCAATTGACTGCTGGGATG
GATCCCAAATGAGCCCATTGTGTACCATGGTTACACATTCACCAGCAAGCTTCTCTTC
AAAACGTGTGGTCCAAGCAATAAACAAGTATGCCTTTGTGACATCTGATTACCCAGTAG
TGCTGTCCTTAGAAAATCACTGCTCCCCTGGTCAGCAGGAAGTGATGGCTAGCATTCTG
CAGAGCACCTTTGGGAGACTTCCTGCTTTCGGACATGCTTGAGGAGTTTCCAGATACACT
ACCGTCTCCAGAGGCACTAAAATTCAAAATATTAGTGAAAAACAGGAAAGTGGGAAC
CTTATCTGAAACCCACGAGAGGATAGGAACCGACAAAAGTGGCCAAGTGCTAGAATG
GAAAGAAGTCATCTATGAAGATGGTGATGAAGACTCAGGAATGGATCCAGAAACATG
GGATGTCTTCCTATCACGGATCAAGGAGGAGAGGGAAGCAGATCCCTCGACATTGAGT
GGAATAGCAGGCGTCAAGAAAAGGAAGAGGAAGATGAAAATAGCCATGGCCTTATCT
GATCTTGTCATTTATACTAAGGCTGAGAAGTTCCGAAACTTCCAATATTCAAGAGTCTA
TCAGCAATTTAATGAGACCAATTCGATTGGAGAGTCTCGAGCTCGAAAACCTTCCAAA
TTGAGAGTCCATGAGTTTATTTTCCACACCGCGGCATTCATCACCAGAGTATACCCCAA
AATGATGAGAGCAGACTCTTCTAACTTTAACCCTCAAGAGTTTTGGAATGTAGGATGT
CAGATGGTGGCCTTGAACCTTCAAACCCCTGGACTGCCTATGGATTTGCAAAACGGGA
AATTTTTGGATAATGGAGGCTCTGGATATATTTTGAAGCCAGACATCCTTAGAGATAC
AACCCTGGGCTTTAACCCTAAATGAACCAGAATATGACGACCATCCAGTTACCCTCACA
ATCCGAATCATCAGTGGGATCCAGTTGCCTGTTAGCTCATCCTCTAACACGCCTGACAT
AGTAGTGATCATAGAAGTCTACGGTGTTCCAAACGACCACGTGAAGCAGCAGACTCGT
GTTGTTAAGAATAATGCTTTTAGTCCAAAGTGGAATGAAACATTTACATTTCTTATTCA
AGTGCCAGAACTGGCATTGATACGTTTTGTTGTTGAAACTCAACAAGGCTTATTATCAG
GAAATGAATTACTCGGGCAGTACACTTTACCAGTTCTTTGCATGAACAAAGGTTATCGT

CGTGTTCCCTCTGTTTTCCAAATCCGGTGCGAACCTTGAACCTTCCTCACTGTTTATTTAT
GTTTGGTACTTCAGAGAGTGA

SEQ ID NO: 5 - Mouse PLC-zeta nucleotide sequence, including untranslated region
(Genbank Accession No AF 35950) (2187 basepairs)

CCCCCCCCGGCAAGCCATGCCAACTGTGAAGGTTTTAAGAACTTAGCTTCTGGACAAGT
TTTCTTCGAAAGTGAAAAGCAGTAGCAGCGAGAACAGCTGATGACGGTCACAAAAG
ACAGTGTTACTTCTAAGACAAGTGACACCTTAGACGAAGAGCCCTCTATGGGAGGACA
AGCGGCCCCAGATCATGAAATCATGGAAAGCCAACTTCATGAGCTCGCAGAAGCAAGA
TGGTTTTTGTCAAAGGTTTCAGGATGATTTTAGAGGTGGAAAAATCAACGTTGAAATTA
CTCACAAACTGCTTGAGAACTTGATTTCCCATGCCACTTTGCTCATGTGAAACATATT
TTTAAGGAAAATGACAGACAGAACCAAGGAAGAATCACCATTGAAGAGTTTAGAGCC
ATTTACCGGTGTATTGTACATAGAGAAGAAATCACGGAGATTTTCAACACGTATACTG
AAAATAGGAAAATTCTTTCTGAGAACAGTCTGATTGAGTTTCTAACCCAAGAGCAGTA
TGAAATGGAGATCGATCACTCTGATTCAGTAGAGATCATCAATAAGTATGAGCCTATT
GAAGAAGTAAAGGGTGAGCGACAGATGTCAATTGAAGGTTTCGCAAGATACATGTTTT
CATCAGAATGTCTACTGTTTAAAGAGAACTGTAAAACCGTGTACCAAGATATGAATCA
TCCATTAAGTGATTATTTTATTTTCATCATCTCACACACATATTTGATATCCGATCAAAT
ATTGGGACCGAGTGACATTTGGGGATATGTAAGTGCTCTTGTGAAAGGCTGCCGCTGT
CTGGAAATGACTGCTGGGATGGATCCCAAAATGAGCCCATTTGTGTACCATGGTTACA
CATTACACAGCAAGCTTCTCTTCAAACTGTGGTCCAAGCAATAAACAAGTATGCCTTT
GTGACATCTGATTACCCAGTAGTGCTGTCCTTAGAAAATCACTGCTCCCCTGGTCAGCA
GGAAGTGATGGCTAGCATTCTGCAGAGCACCTTTGGAGACTTCCTGCTTTCGGACATG
CTTGAGGAGTTTCCAGATACACTACCGTCTCCAGAGGCACTAAAATTCAAAATATTAG
TGAAAAACAGGAAAGTGGGAAACCTTATCTGAAACCCACGAGAGGATAGGAACCGACA
AAAGTGGCCAAGTGCTAGAATGGAAAGAAGTCATCTATGAAGATGGTGATGAAGACT
CAGGAATGGATCCAGAAACATGGGATGTCTTCCTATCACGGATCAAGGAGGAGAGGG
AAGCAGATCCCTCGACATTGAGTGGAATAGCAGGCGTCAAGAAAAGGAAGAGGAAGA
TGAAAATAGCCATGGCCTTATCTGATCTTGTCATTTATACTAAGGCTGAGAAGTTCCGA
AACTTCCAATATTCAAGAGTCTATCAGCAATTTAATGAGACCAATTCGATTGGAGAGT
CTCGAGCTCGAAAACCTTTCCAAATTGAGAGTCCATGAGTTTATTTTCCACACCGCGGCA
TTCATCACCAGAGTATACCCCAAAATGATGAGAGCAGACTCTTCTAACTTTAACCCTCA
AGAGTTTTTGGAATGTAGGATGTCAGATGGTGGCCTTGAACCTTTCAAACCCCTGGACTG
CCTATGGATTTGCAAAACGGGAAATTTTTGGATAATGGAGGCTCTGGATATATTTTGA
AGCCAGACATCCTTAGAGATACAACCCTGGGCTTTAACCCTAAATGAACCAGAATATGA
CGACCATCCAGTTACCCTCACAAATCCGAATCATCAGTGGGATCCAGTTGCCTGTTAGCT
CATCCTCTAACACGCCTGACATAGTAGTGATCATAGAAGTCTACGGTGTTCCAAACGA
CCACGTGAAGCAGCAGACTCGTGTTGTTAAGAATAATGCTTTTAGTCCAAAGTGGAAT
GAAACATTTACATTTCTTATTCAAGTGCCAGAAGTGGCATTGATACGTTTTGTTGTTGA
AACTCAACAAGGCTTATTATCAGGAAATGAATTACTCGGGCAGTACACTTTACCAGTT
CTTTGCATGAACAAAGGTTATCGTCGTGTTCTCTGTTTTTCCAAATCCGGTGCGAACCT
TGAACCTTCCTCACTGTTTATTTATGTTTGGTACTTCAGAGAGTGACGCATGTAAGTGA
CGCGTTAGCTACACATCACAGTAAACAGTCCAAAG

SEQ ID NO: 6 - Monkey PLC-zeta Accession number AB070108 Nucleotide sequence of 1923 bases encoding 641 amino acids

ATGGAAATGAAATGGTTTTTGTCTGAAGATTCAGGATGACTTTAGAGGTGGAAAAATTA
ATCTAGAAAAAACTCAGAGGTTACTTGAAAAATTAGATATTCGGTGTCAGTTATATTCA
TGTGAAACGGATATTTAAGGACAATGACAGGCTGAAACAAGGAAGAATCACCATAGA
AGAATTTAGAGCAATTTATCGAATTCTCACGCACAGAGAAGAAATTGTTGAGATTTTC
AACGCATATTCTGAAAACCGGAAAATTCTTTTAGAAAAATAATCTGGTTCAATTTCTGAC
ACAAGAACAATATACAACCTGAGATGAGTAAACTATTGCTTTTGAGATCATTCAGAAA
TATGAACCTATCGAAGAAGTTAGGAAAGCACGCCAAATGTCATTAGAAGGTTTTACAA
GATATATGGATTACCGTGAATGTCAACTATTTAAAAATGAATGTAGAAAAGTTTATCA
AGATATGACTCATCCATTAAATGATTATTTTATTTTCATCTTCACATAACACATATTTGGT
ATCTGATCAATTAGTGGGACCAAGTGACCTTTGGGGATATGTAAGTGCCCTTGTGAAA
GGATGCCGTTGTTTTGGAGATTGATTGCTGGGATGGAGCACAAAATGAACCTGTTGTAT
ATCATGGCTACACACTCACCAGCAAACCTTCTGTTTAAACTGTTATCCAAGCTATACAC
AAGTACGCATTCATGACATCTGACTACCCAGTGGTGCTCTCTTTGGAAAATCACTGCTC
CCCTGCCCAGCAAGAAATAATGGCAGACAATTTGCAGACGACTTTTGGAGAGTCCTTG
CTTTCTGATATGCTTGCTGATTTTCCTGATACTCTACCATCACCAGAGGCACTAAAATT
CAAAGTATTAGTTAAAAATAAGAAAATAGGAACTTAAAGGAAACCCATGAAAGAA
AGGTTCTGATAAGCGTGGTAAGGTGGAGGAATGGGAAGAAGAAGTGGCAGATCTGGA
GGAGGAGGAGGAGGAGGAGATTCAAAGAATCAGAAATATTCGAATCTGTTTTAGG
AGAAAATCAAGACAAGGAAACAGGGGTAAAAAAGTTATCTGGAGTAACGCTTTTCAA
GAAAAAGAAGACCAGGAAGCTAAAAATTGCTCTGGCCTTATCTGATCTTGTCATTTAT
ACTAAAGCTGAGAAGTTCAAAAGCTTTCAACATTCAAGATTATATCAGCAATTTAATG
AAAATAATTCTATTGGGGAGACACAAGCCCGAAAACCTTTCAAATTGAGAGCCCATGA
GTTTATTTTTTCACACCAGGAAGTTCATTACCAGAATATATCCCAAAGCAACAAGAGCA
GACTCTTCTAATTTTAATCCCCAAGAATTTTGGGAATATAGGTTGTCAAATGGTGGCTTT
AAATTTCCAGACCCCTGGTCTGCCTATGGATCTGCAAAATGGGAAATTTTTGGATAATG
GTGGTTCTGGATATATTTTGAAACCACATTTCTTAAGAGAGAGTGAATCATACTTTAAC
CCAAGTGACATAAAAGACAGTATGCCAATTACACTTACAATAAGGCTCATCAGTGGTA
TCCAGTTGCCTCTTACTCATTTCATCATCTAACAAAGGTGATACATTAGTAATTATAGAA
GTTTTTGGTGTTCCAAATGATCAAATGAAGCAGCAGACTCGTGTAATTAATAAAAAAATG
CTTTTAGTCCAAGATGGAATGAAACATTTACATTTATTATTCATGTCCCAGAATTGGCA
TTGATACGTTTTTGTGTTGAAAGTCAAGGTTTAATAGCAGGAAATGAATTTCTTGGGCA
ATATACTTTGCCACTTCTATGCATGAACAAAGGCTATCGTCGTGTTCTCTGTTTTCCCG
AATGGGTGAGAGCCTTGAGCCTGCTTCACTGTTTGTTTATGTTTGGTACGTCAGA

SEQ ID NO: 7 - Monkey PLC-zeta Accession number AB070108 Protein sequence of 641 amino acids

MEMKWFLSKIQDDFRGGKINLEKTQRLLEKLDIRCSYIHVKRIFKDNDRLKQGRITIEEFRA
IYRILTHREEIVEIFNAYSENKILLENLNVQFLTQEYTTTMSKTIAFEIQKYEPIEEEVRKAR
QMSLEGFTRYMDSRECQLFKNECRKVYQDMTHPLNDYFISSSHNTYLVSDQLVGPSDLW
GYVSALVKGCRCLCIDCWGAQNEPVVYHGYTLTSKLLFKTVIQAIHKYAFMTSDYPVVL
SLENHCSPAQQEIMADNLQTTFGESLLSDMLADFPDTLPSPEALKFKVLVKNKKIGTLKET
HERKGSDKRGKVEEWEEVADLEEEEEEEEEKFKESEIFESVLGENQDKETGVKKLSGVTLF
KKKKTRKLIKIALALS DLVIYTKAEKFSFQHSRLYQQFNENNSIGETQARKLSKLRAHEFIF
HTRKFITRIYPKATRADSSNFPNQEFWNIGCQMVALNFQTPGLPMDLQNGKFLDNGGSGYI
LKPHFLRESESYFNPSDIKDSMPITLTIRLISGIQLPLTHSSSNKGD TLVIIEVFGVPNDQMKQ
QTRVIKKNAFSPRWNETFTFIIHVPELALIRFVVESQGLIAGNEFLGQYTLPLLCMNKGYRR
VPLFSRMGESLEPASLFVYVWYVR

SEQ ID NO: 8 - Monkey PLC-zeta Accession number AB070109 Nucleotide sequence of 1917 bases encoding 640 amino acids

ATGGAAATGAAATGGTTTTTGTCTGAAGATTCAGGATGACTTTAGAGGTGGAAAAATTA
 ATCTAGAAAAAACTCAGAGGTTACTTGAAAAATTAGATATTCGGTGCAGTTATATTCA
 TGTGAAACGGATATTTAAGGACAATGACAGGCTGAAACAAGGAAGAATCACCATAGA
 AGAATTTAGAGCAATTTATCGAATTCTCACGCACAGAGAAGAAATTGTTGAGATTTTC
 AACGCATATTCTGAAAACCGGAAAATTCTTTTAGAAAATAATCTGGTTCAATTTCTGAC
 ACAAGAACAATATACAACCTGAGATGAGTAAACTATTGCTTTTGAGATCATTTCAGAAA
 TATGAACCTATCGAAGAAGTTAGGAAAGCACGCCAAATGTCATTAGAAGGTTTTACAA
 GATATATGGATTACGTGAATGTCAACTATTTAAAAATGAATGTAGAAAAGTTTATCA
 AGATATGACTCATCCATTAAATGATTATTTTATTTTCATCTTCACATAACACATATTTGGT
 ATCTGATCAATTAGTGGGACCAAGTGACCTTTGGGGATATGTAAGTGCCCTTGTGAAA
 GGATGCCGTTGTTTGGAGATTGATTGCTGGGATGGAGCACAAAATGAACCTGTTGTAT
 ATCATGGCTACACACTCACCAGCAAACCTTCTGTTTAAACTGTTATCCAAGCTATACAC
 AAGTACGCATTCATGACATCTGACTACCCAGTGGTGCTCTCTTTAGAAAATCACTGCTC
 CCTGCCCAGCAAGAAATAATGGCAGACAATTTGCAGACGACTTTTGGAGAGTCCTTG
 CTTTCTGATATGCTTGCTGATTTTCTGATACTCTACCATCACCAGAGGCACTAAAATT
 CAAAGTATTAGTTAAAAATAAGAAAATAGGAACCTTAAAGGAAACCCATGAAAGAAA
 AGGTTCTGATAAGCGTGGTAAGGTGGAGGAATGGGAAGAAGAAGTGGCAGATCTGGA
 GGAGGAGGAGGAGGAGGAGGAGAAATTCAAAGAATCAGAAATATTCGAATCTGTTTT
 AGGAGAAAATCAAGACAAGGAAACAGGGGTAAAAAAGTTATCTGGAGTAACGCTTTT
 CAAGAAAAAGAAGACCAGGAAGCTAAAAATTGCTCTGGCCTTATCTGATCTTGTCATT
 TATACTAAAGCTGAGAAGTTCAAAAGCTTTCAACATTCAAGATTATATCAGCAATTTA
 ATGAAAATAATTCTATTGGGGAGACACAAGCCCGAAAACCTTTCAAATTGAGAGCCCA
 TGAGTTTATTTTTTCACACCAGGAAGTTCATTACCAGAATATATCCCAAAGCAACAAGA
 GCAGACTCTTCTAATTTTAAATCCCCAAGAATTTTGGGAATATAGGTTGTCAAATGGTGGC
 TTAAATTTCCAGACCCCTGGTCTGCCTATGGATCTGCAAAATGGGAAATTTTTGGATA
 ATGGTGGTTCTGGATATATTTTGAAACCACATTTCTTAAGAGAGAGTGAATCATACTTT
 AACCCAAGTGACATAAAAGACAGTATGCCAATTACACTTACAATAAGGCTCATCAGTG
 GTATCCAGTTGCCTCTTACTCATTTCATCATCTAACAAAGGTGATACATTAGTAATTATA
 GAAGTTTTTGGTGTTCCAAATGATCAAATGAAGCAGCAGACTCGTGTAATTAAAAAAA
 ATGCTTTTAGTCCAAGATGGAATGAAACATTTACATTTATTATTTCATGTCCCAGAATTG
 GCATTGATACGTTTTGTTGTTGAAAGTCAAGGTTTAATAGCAGGAAATGAATTTCTTGG
 GCAATATACTTTGCCACTTCTATGCATGAACAAAGGCTATCGTCGTGTTCTCTGTTTT
 CCCGAATGGGTGAGAGCCTTGAGCCTGCTTCACTGTTTGTATTATGTTTGGTACGTCAGA

SEQ ID NO: 9 - Monkey PLC-zeta Accession number AB010109 Protein sequence of 370 amino acids

MEMKWFLSKIQDDFRGGKINLEKTQRLLEKLDIRCSYIHVKRIFKDNDRLKQGRITIEEFRA
IYRILTHREEIVEIFNAYSENKRILLENNLVQFLTQEYTTTEMSKTIAFEIIQKYEPIEEVRKAR
QMSLEGFTRYMDSRECQLFKNECRKVYQDMTHPLNDYFISSSHNTYLVSDQLVGPSDLW
GYVSALVKGCRCLCIDCWGAQNEPVVYHGYTLTSKLLFKTVIQAIHKYAFMTSDYPVVL
SLENHCSPAQQEIMADNLQTTFGESLLSDMLADFPDTLPSPEALKFKVLVKNKKIGTLKET
HERKGSDKRGKVEEWEEVADLEEEEEERFKESEIFESVLGENQDKETGVKKLSGVTLFK
KKKTRKCLKIALALSDLVIYTKAEKFKSFQHSRLYQQFNENNSIGETQARKLSKLRAHEFIFH
TRKFITRIYPKATRADSSNFPQEFWNIGCQMVALNFQTPGLPMDLQNGKFLDNGGSGYIL
KPHFLRESESYFNPSDIKDSMPITLTIRLISGIQLPLTHSSSNKGDTLVIEVFGVPNDQMKQQ
TRVIKKNAFSPRWNETFTFIHVPELALIRFVVESQGLIAGNEFLGQYTLPLLCMNKGYRRV
PLFSRMGESLEPASLFVYVWYVR

SEQ ID NO: 10 - Rat PLC-zeta Nucleotide sequence of 1938 bases encoding 646 amino acids

ATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCGCCCTTTTCAGG
ATGATTTTCAGAGGTGGAAAAATCAGCGCTGGAATTACTCACAAACTGCTCGAGAACT
TGATTTCCCATGCCATTTTGCTCATGTGAAACGTATTTTAAAGGAAAATGACAGACATA
ACCAAGGAAGAATCACCACCGAAGATTTTAGAACCATCTATCGGTGTATTGTACATAG
AGAAGAGATCGTTGAGATTTTCAACACGTATACTGAAAACAGGAAAATTCTCCCCGAG
GACAGTCTGATTGAATTTCTAACCCAAGAGCAGTATGAAATGGAGATGGATGAGTCCA
GTTTCAGTGGAGATCATCCAGAAGTACGAGCCCATTGCAGAAGTAAAGAACGAGCGGC
AGATGTCAATTGAAGGTTTTGCAAGATACATGTTTTCTTCAGAATGTCTACTGTTTAA
GAGACGTGTAACACAGTGTACCAAGATATGAATAAGCCACTAAATGATTACTATATTT
CATCGTCTCACAAACACATATTTGATATCTGATCAAATATTGGGACCAAGTGACATTTGG
GGATATATAAGTGCCCTCGTGAAAGGTTGCCGCTGTCTGGAAATTGACTGCTGGGATG
GAGCACAAAATGAACCCATTGTGTACCATGGCTACACTCTCACCAGCAAGCTTCTCTTC
AAAACCGTTATCCAAGCAATAAACAAGTACGCCTTCGTGACGTCTGATTACCCAGTGG
TGCTGTCCTTAGAGAATCACTGCTCCCCTGGTCAACAGGAAGTGATGACCGACATTCT
GCAGAGTACCTTTGGAGACTTTCTGCTCTCAGACATACTTGACGAGTTTCCAGACAGTT
TGCCATCTCCAGAGGCACTGAAATTTAAAATATTAGTGAAAAATAAGAAAGTTGGAAC
CTTATCTGAAACCCGCGAGAGGCTGGGGACTGACAAAAGGGGGCATAGCGCTAGACTT
GGAAGAAGAAATCTATGAAAATGAAGACGAAGACTCAGGAAAGGAGCCAGAAACGT
GGGATGATTTCTGTACGGGTTAAGGAGGAGCAGGAGGCAGACCCCTCAACGTTGAG
CGGAATAGCAGATGCCAAGAAAAAGATCAGGAAGCTAAGAGTAGCTCTGGCCTTATC
TGATCTTGTCATTTATACCAAGCTGAGAAGTTCCGAAACTTCCAATATTCAAGAGTCT
ATCAGCAGTTTAATGAGACCACTTCTATGGGAGAGTCTCGAGCTCGAAAACTTTCAA
ATTGAGAGCCCATGAGTTTATTTTCCCACTGCAGCGTTCATCACCAGAGTGTACCCCA
AGTTCACGAGAGCAGACTCTTCTAATTTTAATCCTCAAGAGTTTTTGGAATGTGGGCTGT
CAGATGGTGGCCTTGAATTTTCAAACCCCTGGACTGCCTATGGATTTGCAAAACGGGA
AATTTTTGGATAATGGAGGCTCTGGATATGTTTTGAAGCCAGACTTTCTTAGAGACACA
ACTTTGGGCTTTAACCCTAATGAACCAGAAGGAGATGGCCATCCGGTTACCCTCACGA
TCCGACTCATCAGTGGGATCCAGTTGCCTGTAAACGTGCCCTCAAATACATCTGACATA
ATAGTGATCATAGAAGTCTACGGTGTCCCAAACGACCACATGAAGCAGCAGAGTCGTG
CCGTTAAGAACAATGCTTTTAGTCCAAGGTGGAATGAAACATTCACATTTCTTATTCAA
GTGCCAGAATTGGCACTGATACGTTTCGTTGTTGAAACTCAAGGCTTCCTATCGGGAA
ATGAATTACTTGGGCAGTACACTTTACCCGTTCTTTGCATGAACAAAGGTTATCGTCGT
GTTCTCTGTTTTTCCAAATCCGGTGCGAACCTTGAACCTTCCTCTCTGTTTATTACGTT
TGGTACTACAGAGAG

SEQ ID NO: 11 - Rat PLC-zeta Protein sequence of 646 amino acids

MHARAAASVMDICRIRPFQDDFRGGKISAGITHKLLLEKLDFPCHFAHVKRIFKENDRHNQG
RITTEDFRTIYRCIVHREEIVEIFNTY TENRKILPEDSLIEFLTQEQYEMEMDESSSVEIIQKYE
PLAEVKNERQMSIEGFARYMFSSSECLLFKETCNTVYQDMNKPLNDYYISSSHNTYLISDQIL
GPSDIWGYISALVKGCRCLCIDCWGDAQNEPIVYHGYTLTSKLLFKTVIQAINKYAFVTSD
YPVVLSLENHCSPGQQEVMTDILQSTFGDFLLSDILDEFDPSLPSPEALKFKILVKNKKVGT
LSETRERLGTDKRGIALDLEEEIYENEDEDSGKEPETWDDFLSRVKEEQEADPSTLSGIADA
KKKIRKLRVALALSDLVIYTKAEKFRNFQYSRVYQQFNETTSMGESRARKLSKLRAHEFIF
HTAAFITRVYPKFTRADSSNFNPQEFWNVGCQMVALNFQTPGLPMDLQNGKFLDNGGSG
YVLKPDFLRDITTLGFNPNEPEGDGHPVTLTIRLISGIQLPVNVPSNTSDIIVIEVYGVPNDHM
KQQSRAVKNNAFSPRWNETFTFLIQVPELALIRFVVETQGFLSGNELLGQYTLPLVLCMNKG
YRRVPLFSKSGANLEPSSLFIYVWYYRE

SEQ ID NO: 12 – PLC Sequences Alignment

zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	1	MTSEEMTASVLIPVTQRKVVSAQSAADESSEKVS DINISKAHTVRRSGET
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	51	SHTISQLNKLKEEPSGSNLPKILSIAREKIVSDENSNEKCWEKIMPDSAK
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	101	NLNINCNNILRNHQHGLPQRQFYEMYNSVAEEDLCLETGIPSPLEKRVFP
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	151	GIQLELDRPSMGISPLGNQSVIIETGRAHPDSRRAVFHFHYEVDRRMSDT
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----

delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	201	FCTLSENLIILDDCGNCVPLPGGEEKQKKNYVAYTCKLMELAKNCDNKNEQ
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	251	LQCDHCDTLNDKYFCFECSCEKVDVMVYSGDSFCRKDFDTSQAAKTFLSHF
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	301	EDFPDNCDDVEEDAFKSKKERSTLLVRRFCKNDREVKKSVYTGTRAIVRT
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	351	LPSGHIGLTAWSYIDQKRNGPLLPCGRVMEPPSTVEIRQDGSQRLSEAQW
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----

delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	401	.YPIYNAVRREETENTVGSLLHFLTKLPASETAHGRI SVGPCCLKQCVRDTV
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	451	CEYRATLQRTSISQYITGSLLEATTSLGARSGLLSTFGGSTGRMMLKERQ
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	501	PGPSVANSNALPSSSAGISKELIDLQPLIQFPPEEVASILMEQEQTYYRRV
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	551	LPVDYLCFLTRDLGTPECQSSLPCLKASISASILTTONGEHNALEDLVMR
delta1-human	1	-----
delta3-human	1	-----

delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	601	FNEVSSWVTWLILTAGSMEEKREVF SYLVHVAKCCWNMGNYNAVMEFLAG
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	651	LRSRKVLKMWQFMDQSDIETMRSLKDAMAQHESSEYRKVVTRALHIPGC
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	701	KVVPFCGVFLKELCEVLGDGASGLMKLCPRYNSQEETLEFVADYSGQDNFL
delta1-human	1	-----
delta3-human	1	-----
delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----
beta3-mouse	1	-----
beta2-human	1	-----
beta4-mouse	1	-----
epsilon-human	751	QRVGQNGLNSEKESTVNSIFQVIRSCNRSLETDEEDSPSEGNSSRKSSL
delta1-human	1	-----
delta3-human	1	-----

delta2-bovine	1	-----
delta4-rat	1	-----
gamma1-human	1	-----
gamma2-human	1	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	1	-----MGSLOGIATK
beta3-mouse	1	-----MAGARPGVHAL
beta2-human	1	-----MSLLNP
beta4-mouse	1	-----MAKPYE
epsilon-human	801	KDKSRWQFIIGDLLSDNDIFEQSKEYDSHGSEDSQKAEDHGTETLIPWYV
delta1-human	1	-----MDSGRDFLTLLH
delta3-human	1	-----M
delta2-bovine	1	-----MAYLLQG
delta4-rat	1	-----MASQIQK
gamma1-human	1	-----MAGAASPCANGCGPGAPS
gamma2-human	1	-----MSTTVNVDSL
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	11	ILIRILSDALTRKET---DLKSSSTIVTP-IIILRTDPQGFEFYWTDQNK
beta3-mouse	12	QLEPSTVVKTLRRGSKFIKWDEEASSRNL-VTLRVDPNNGFELYWTGPNME
beta2-human	7	VLLPPKVKAYLSQGERFIKWDDETTIVASP-VILRVDPKGYLYWYQSKE
beta4-mouse	7	FNWQKEVPSELOEGAVFDREYEESEFVFEPNCLFKVDEFGFELTWKSEGKE
epsilon-human	851	LSIQADVHQFLLOQATVIHYDQDTHLSAR-CFLQLOPDNSTLTWVKPTTA
delta1-human	12	GLQDDEDLQALLKGSQLLKVKSSSWRRERFYKIQELCKTIWQ-----
delta3-human	2	GLTEDEDVRAMLGRSLRKIRSRTHWHERLYRIQELGLSVWF-----
delta2-bovine	8	RLPINQDLLLMQKGTMMRKVRSKSWKKLRFFRLQDDGMTVWH-----
delta4-rat	8	LLTTNQDLLLMQKGTMMRKVRTKSWKKLRYFRLQDDGMTVWH-----
gamma1-human	19	DAEVLHLCRSLVSTVMTLFYSKKSQRPERKTFQVKLETRQITWS-----
gamma2-human	12	EYEKSQIKRALELSTVMTVFSFRKS-TPERRTVQVIMETROQVWS-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	56	-----
beta3-mouse	61	-----
beta2-human	56	-----
beta4-mouse	57	-----
epsilon-human	900	SPASSKAKLGVLNNTAEPGKFPLLGNAGLSSLTEGVLDLFAVKAVYMGRP
delta1-human	54	-----
delta3-human	44	-----
delta2-bovine	50	-----
delta4-rat	50	-----
gamma1-human	64	-----
gamma2-human	56	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	56	-----
beta3-mouse	61	-----
beta2-human	56	-----
beta4-mouse	57	-----
epsilon-human	950	GIDIHTVCVQNKLGSMFLSETGVTLTYGLQTTDNRLLEHFAVKHTAKMLF
delta1-human	54	-----
delta3-human	44	-----

delta2-bovine	50	-----
delta4-rat	50	-----
gamma1-human	64	-----
gamma2-human	56	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	56	-----
beta3-mouse	61	-----
beta2-human	56	-----
beta4-mouse	57	-----
epsilon-human	1000	SGLLELTRAVRKMRKFPDQRQQLRKQYVSLYQEDGRYEGPTLAHAVELF
delta1-human	54	-----
delta3-human	44	-----
delta2-bovine	50	-----
delta4-rat	50	-----
gamma1-human	64	-----
gamma2-human	56	-----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	56	-----TELLDLSL
beta3-mouse	61	-----VDTLDISS
beta2-human	56	-----MEFLDITS
beta4-mouse	57	-----GOVLECSL
epsilon-human	1050	GGRRWSARNPSPGTSAKNAEKPNMQRNNTLGISTTKKKKKILMRGESGEV
delta1-human	54	-----ESRKVMRTP
delta3-human	44	-----QRR-IPRAP
delta2-bovine	50	-----ARQAGGRAK
delta4-rat	50	-----GRHLESISK
gamma1-human	64	-----RGADKIEG
gamma2-human	56	-----KTADKIEG
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	64	VKDARCGRHAKAPKDPKRLRELLDVGN-IGRLEQRMITVYVCPDLVN----
beta3-mouse	69	IRDTRTGRYARLPKDPNIREVLGFGGPDTRLLEKLMTVVAGPDPVN----
beta2-human	64	IRDTRFGKFAPKPKSQKLRDVFNMDFPDNSFLLKTLTVVSGPDMVD----
beta4-mouse	65	INSIRQAAPKDPK--ILAALEAVGKSENDLEGRILCWCSQTDLVN----
epsilon-human	1100	TDDEMATRKAKMHIECRSRSGSDPDINEQEESEVNALANPPNPLPSRRA
delta1-human	63	ESQLFSIEDIQEVEMGHRTGLEKFARDVPEDRCFESIVFKDQ-----
delta3-human	52	SQHIFVQHIEAVREGHQSEGLRRFGGAFAPARCLTIAFKGR-----
delta2-bovine	59	PS--FSISDVTVRREGHESELLRNLAEEFPLEQGFITVFNHGR-----
delta4-rat	59	PT--FSISDVERIRKQDSSELLRYLVVEEFPLEQGFITVFNHGR-----
gamma1-human	72	AIDIREIKEIRPGKTSRDFRYQEDPAFRPDQSHCFVILYGMFRLK----
gamma2-human	64	FLDIMEIKEIRPGKNSKDEERAK--AVRQKEDCCFTILYGTQFVLS----
zeta-monkeyA	1	-----
zeta-human	1	-----
zeta-mouse	1	-----
zeta-rat	1	-----
beta1-human	109	-----ISHLNIVAFQEEVAKETNEVFSL
beta3-mouse	115	-----TTFINFMVQDDTVKVWSEELFKL
beta2-human	110	-----LTFHNFVSYKENVGKAWAEDVLAL
beta4-mouse	109	-----IGFTYMVAENPEVTQOWVEGLRSI
epsilon-human	1150	HSLTTAGSPNLAAGTSSPIRPVSSPVLSSSNKSPSSAWSSSSWHGRIKGG
delta1-human	105	-----RNTLDLIAFSPADAQHWVLGLHKL
delta3-human	94	-----RKNLDLAAPTAAEAQRWVRGLTKL

RECTIFIED SHEET (RULE 91)

delta2-bovine 249 IDRYEFSSESGNLRHV
 delta4-rat 249 IDRYEFSSENGRLLRVLSKDGFLSYLCSAD-GNIFNPDCLPTIYQDMTOPLS
 gamma1-human 278 LSFLRDPLREIEEPYFFLDEFVTELESKENS VWNSQLDAVCPDTMNNPLS
 gamma2-human 270 TKFIDDTMRETAEPFLFVDEFLLTYLESRENSIWDEKYDAVDMQDMNNPLS

zeta-monkeyA 163 DYFISSSHNTYLVSDQLVGPDLW YV SALVKGCRCLCIDCWGDG--AQNE
 zeta-human 163 DYFISSSHNTYLVSDQLVGPDLW YV SALVKGCRCLCIDCWGDG--AQNE
 zeta-mouse 171 DYFISSSHNTYLVSDQLVGPDLW YV SALVKGCRCLCIDCWGDG--SQNE
 zeta-rat 171 DYYISSSHNTYLVSDQLVGPDLW YV SALVKGCRCLCIDCWGDG--AQNE
 beta1-human 319 HYFINSSHNTYLTAGQIAGNSSVEMYRQALLSGCRCVELDCWRGRTAEED
 beta3-mouse 325 AYFINSSHNTYLTAGQIAGNSSVEMYRQALLSGCRCVELDCWRGRTAEED
 beta2-human 320 HYFINSSHNTYLTAGQIAGNSSVEMYRQALLSGCRCVELDCWRGRTAEED
 beta4-mouse 321 HYFINSSHNTYLTAGQIAGNSSVEMYRQALLSGCRCVELDCWRGRTAEED
 epsilon-human 1400 YYYIESSSHNTYLTGHOKESSVELYSQVLLQGCRCVELDCWDG--DDGM
 delta1-human 304 HYLVSSSHNTYLLDQLAGPSSSTEAYIRALCKGCRCLCIDCWGDG--PNQE
 delta3-human 292 HYFISSSHNTYLTDSQIGPSSSTEAYVRAFAQGCRCVELDCWDG--PGGE
 delta2-bovine 298 HYYINSSHNTYLVGDQLCGQSSVEGYIRALKRGCRCLCIDCWGDG--PSGE
 delta4-rat 298 HYYINSSHNTYLLGDQFCGQSSVEGYIRALKRGCRCLCIDCWGDG--PDGE
 gamma1-human 328 HYWISSSHNTYLTGDQFSSESLFAYARCLRMGCRCLCIDCWGDG--PDGM
 gamma2-human 320 HYWISSSHNTYLTGDQFSSESLFAYARCLRMGCRCLCIDCWGDG--PDGM

zeta-monkeyA 211 PV YHG TLTSK LFKTVIOAIHKYAFMTSDYPVVLLENHCS-PAQQEI
 zeta-human 211 PV YHG TLTSK LFKTVIOAIHKYAFMTSDYPVVLLENHCS-TAQQEV
 zeta-mouse 219 PI YHG TETSK LFKTVIOAINKYAFVTSYDYPVVLLENHCS-PGQQEV
 zeta-rat 219 PI YHG TLTSK LFKTVIOAINKYAFVTSYDYPVVLLENHCS-PGQQEV
 beta1-human 369 PVIHGHGTMTTEISEKEVKAIAECAFKTSPEPILLSFENHVDSPKQOAK
 beta3-mouse 375 PEITHGHGTMTTEVPLRDVLEAIAEAFKTSPPYVILSFENHVDSPKQOAK
 beta2-human 370 PIITHGHGTMTTDIEFKEAIAEAFKTSPPYVILSFENHVDSPKQOAK
 beta4-mouse 371 PIITHGHGTMTTDIEFKEAIAEAFKTSPPYVILSFENHVDSPKQOAK
 epsilon-human 1448 PIIYHGHGTPTTKIPFKEVVEAIDRSFINSDLPIIISTENHCS-LPQORR
 delta1-human 352 PIIYHGHGTPTTKIPFKEVVEAIDRSFINSDLPIIISTENHCS-LPQORR
 delta3-human 340 PVIYHGHGTPTTKIPFKEVVEAIDRSFINSDLPIIISTENHCS-LPQORR
 delta2-bovine 346 PIVYHGHGTPTTKIPFKEVVEAIDRSFINSDLPIIISTENHCS-LPQORR
 delta4-rat 346 PVIYHGHGTPTTKIPFKEVVEAIDRSFINSDLPIIISTENHCS-LPQORR
 gamma1-human 376 PVIYHGHGTPTTKIPFKEVVEAIDRSFINSDLPIIISTENHCS-LPQORR
 gamma2-human 368 PVIYHGHGTPTTKIPFKEVVEAIDRSFINSDLPIIISTENHCS-LPQORR

zeta-monkeyA 260 MADNLQSTFGESLL DMLADFEEDT----LPSPEALKFKVLVKNKKITGLK
 zeta-human 260 MADNLQSTFGESLL DMLADFEEDT----LPSPEALKFKVLVKNKKITGLK
 zeta-mouse 268 MASILQSTFGESLL DMLADFEEDT----LPSPEALKFKVLVKNKKITGLK
 zeta-rat 268 MTDILQSTFGESLL DMLADFEEDT----LPSPEALKFKVLVKNKKITGLK
 beta1-human 419 MAEYCRLIFGDALIMEPLEKYPLESGVPLPSFMDLMYKILVKNKKKSHKS
 beta3-mouse 425 MAEYCRLIFGDALIMEPLEKYPLESGVPLPSFMDLMYKILVKNKKKSHKS
 beta2-human 420 MAEYCRLIFGDALIMEPLEKYPLESGVPLPSFMDLMYKILVKNKKKSHKS
 beta4-mouse 420 MSKYCEDLFGDALLKQALESELEPGRPLPSFMDLMYKILVKNKKKSHKS
 epsilon-human 1497 MAEIFKTVFGEKLVTKLFETDFSDPMLPSFMDLMYKILVKNKKKSHKS
 delta1-human 401 MARHLHAILGPMMLNRPLDG-VTNS---LPSPEQLNGKILVKNKKKSHKS
 delta3-human 389 MARHLCTILGPMMLNRPLDG-VTNS---LPSPEQLNGKILVKNKKKSHKS
 delta2-bovine 395 IVRHLTEILGPMMLNRPLDG-VTNS---LPSPEQLNGKILVKNKKKSHKS
 delta4-rat 395 MAHHLIAILEGEQLLSTTLEEQLIDI---LPSPEQLNGKILVKNKKKSHKS
 gamma1-human 425 MAQYFKKVLGSLTLTKPVEISADG---LPSPEQLNGKILVKNKKKSHKS
 gamma2-human 417 MAKAFKEVFGDLTLTKPTEASADQ---LPSPEQLNGKILVKNKKKSHKS

zeta-monkeyA 306 ETHERKGS KRGKVEEWEEVADLEEEEEEEEFKESEI-----
 zeta-human 306 ETHERKGS KRGKVEEWEEVADLEEEEEEEEFKESEI-----
 zeta-mouse 314 ETHERIGT KSGQVLEWKEVIYE----DGDEDSGMDPET-----
 zeta-rat 314 ETHERIGT KSGQVLEWKEVIYE----DGDEDSGMDPET-----
 beta1-human 469 S---EGSGKKKLSEQASNTYSDSSSMFEPSSP---GAG-----
 beta3-mouse 475 TGVPDSSVRKRPLEQNSALSSESAATEPSSP---QLGSPSSDSCPGLSN
 beta2-human 470 PTSSSKDTGGEAEGSSPPSAPAVWAGEEGTELEEE-----
 beta4-mouse 470 EKKQLEALKSMMEAGESAAPASILEDDNEEEIESADQEEEAHPEYKFG--
 epsilon-human 1547 TPVDILKQKAHQQLASMQVQAYNGGNANPRPANNEEEDEEEDYDYDYESL
 delta1-human 447 PPGGEGGPEATVVSDEDEAAEMEDEA-----
 delta3-human 436 SEDGRALS DREEEEDDEEEEEEVEA-----

delta2-bovine	442	EEEEPEAELEAEQEA	
delta4-rat	442	VVESDKEEELEKDEGSDDLDPASAE	
gamma1-human	471	AYEEVPTSMYSENDISNSIKNGILYLEDPVNHEWYPHYFVLTSKIIYS	
gamma2-human	461	----RGDVLVNMEDKKDEHKQOGELYMWDSIDQKWTRHYCAIADAKLSFS	
zeta-monkeyA	345	-----FESVLGE	
zeta-human	318	-----D	
zeta-mouse	349	-----WDVFLSR	
zeta-rat	349	-----WDDFLSR	
beta1-human	501	-----EADT	
beta3-mouse	522	GEEVGL-----EKTSLEPQKSLGEESLSREPNVMPDRDREDEE	
beta2-human	505	-----EVEEEEE	
beta4-mouse	518	-----NELSADDYSHKEAVANS	
epsilon-human	1597	SDDNIL-----EDRPENKSCN-DKLQFEYNEEI PKRIKKADNSA	
delta1-human	473	-----VRSRV	
delta3-human	462	-----AAQRR	
delta2-bovine	468	-----PQDLS	
delta4-rat	468	-----DMQSQ	
gamma1-human	521	EETSSDQGNEDDEEPEKEVSSSTELHSNEKWFHGKLGAGRDGRHIAERLLT	
gamma2-human	507	DDIEQT----MEEVFPQDIPPTLHFGEKWFHKKVEKRTS----AEKLLQ	
zeta-monkeyA	352	NQDKETGVKKLSGVTLFKKKKTRKLTALALSDLVYT A	
zeta-human	319	NQDKETGVKKLPVLMFLKKKKTRKLTALALSDLVYT A	
zeta-mouse	356	IKEEREADPSTLSGIAGVKKRKRKMTAMALSDLVYT A	
zeta-rat	356	VKEEQEADPSTLSGIADAKKIRKLRVALALSDLVYT A	
beta1-human	505	ESDDDDDDDDCKKSSMDEGTAGSEAMATEEMSNLVNYIQP	
beta3-mouse	561	EDEEEETTPDKKPTTDEGTASSEVNATEEMSTLVNYVEP	
beta2-human	512	EESGNLDEEEIKKMQSDEGTAGLEVTAEEEMSLVNYIQP	
beta4-mouse	535	VKKGLVTVEDEQAWMASYKYVGATTNIHPYLSMTMINYAQP	
epsilon-human	1635	CNKGKVDMEELGEEFYLDQNKKESTRQIAPELSDLVIYRQA	
delta1-human	478	QHKPKED-----KLRLAQELSDMVIYCKS	
delta3-human	467	LAK-----QISPELSAIAVYCHA	
delta2-bovine	473	PRSEDKKKPK-----AILCPALSALVVYLKA	
delta4-rat	473	PESQEQASGNKSKNKKFLLQSSTTILCPDLSALVVYLRT	
gamma1-human	571	EYCIETGAPDGSFLVRESETFVGDYTLSEFWRNGKVQHCRHSRQDAGTPK	
gamma2-human	549	EYCMETGGKDGTFLVRESETFPNDYTLSEFWRNGKVQHCRIRSTMEGGTLK	
zeta-monkeyA	392	-----EKFKSFOHS	
zeta-human	359	-----EKFKSFOHS	
zeta-mouse	396	-----EKFRNFOYS	
zeta-rat	396	-----EKFRNFOYS	
beta1-human	545	-----VKEESHEIS	
beta3-mouse	601	-----VKEKSFEAS	
beta2-human	552	-----TKFVSFEFS	
beta4-mouse	575	-----VKEQGEHVA	
epsilon-human	1675	-----VKFPGLSTLNASGSSRGKERKSRKSI FGNNPGRMSPGETASFNK	
delta1-human	502	-----VHFGGESPPG	
delta3-human	485	-----TELRLTLHPA	
delta2-bovine	500	-----VTTFYSETHS	
delta4-rat	513	-----APECSETHS	
gamma1-human	621	FELTDN-----	
gamma2-human	599	YYLTDN-----	
zeta-monkeyA	401	-----RLYQQFNENNSIGETQA KL	
zeta-human	368	-----RLYQQFNENNSIGETQA KL	
zeta-mouse	405	-----RVYQQFNENNSIGESRA KL	
zeta-rat	405	-----RVYQQFNENNSIGESRA KL	
beta1-human	554	-----KKRNKSFEMSSFEVETKGLEQ	
beta3-mouse	610	-----KKRNKCFEMSSFEVETKAMEQ	
beta2-human	561	-----AQKNRSYVSSFTLKA YDL	
beta4-mouse	584	-----EERNIHYNMSSFNESVGLGY	
epsilon-human	1719	TSGKSSCEGIRQTWEESSSPLNPTTSLSAIIRTPKCHISSLNENNAAKRL	
delta1-human	512	-----TPGQAFYEMASFSENRALRL	
delta3-human	494	-----PNAPQPCOVSSLSERKAKKL	

delta2-bovine	509	-----KENVHYDSSFSKAKNL
delta4-rat	522	-----LVFDSLVDLITHYQQVPLRCNEF
gamma1-human	627	-----LRFRMYALICHYRETHLPCEAF
gamma2-human	605	
zeta-monkeyA	421	S LRAHEE FHTRKE RIYPKATRADSSN NPQEFWNIGCOMVALNFOT
zeta-human	388	S LRVHEE FHTRKE RIYPKATRADSSN NPQEFWNIGCOMVALNFOT
zeta-mouse	425	S LRVHEE FHTRKE RVYPKMMRADSSN NPQEFWNIGCOMVALNFOT
zeta-rat	425	S LRAHEE FHTRKE RVYPKFTRADSSN NPQEFWNIGCOMVALNFOT
beta1-human	574	LTKSPVEFVEYNKMSRIYPKGTVDSSNMPQLEFWNAGCOMVALNFOT
beta3-mouse	630	LTKSPMEFVEYNKQOLSRIYPKGTVDSSNMPQLEFWNAGCOMVALNFOT
beta2-human	581	LSKASVQFVDYNKQMSRIYPKGTVDSSNMPQLEFWNAGCOMVALNFOT
beta4-mouse	604	LKTHAIEFVNYNKQMSRIYPKGTVDSSNMPQLEFWNAGCOMVALNFOT
epsilon-human	1769	CRYSQKLIQHTACQLLRTYPAATRIDSSNPNPLMFWLHGQIOLVALNFOT
delta1-human	532	LOESGNGFVRHNVGHLRSRIYPAGWRTDSSNYSFVEMWNGGCQIOLVALNFOT
delta3-human	514	IREAGNSEFVRHNAQLTRVYPLGLRMNSANYSPQEMWNSGCQIOLVALNFOT
delta2-bovine	529	IKEAGDEFVQHNAWQLSRVYPSGLRTDSSNYPQEFWNAGCOMVAMNMOT
delta4-rat	542	IREAGNEFVQHNAWQLSRVYPSGLRTDSSNYPQEFWNAGCOMVAMNMOT
gamma1-human	650	EMRLSEVPVQTHAHESKEWYHASLTRAQAEHMLMRVPRDGAFLVRKRNEP
gamma2-human	628	ELRLTDPVPNPPEHESKPWYDLSLRGEAEDMLMRIPRDGAFLVRKRNEP
zeta-monkeyA	471	PGLPMDL NGKFLDNGGSGYLLKPHFLRESESYFNESE-IKDSMPI TLTI
zeta-human	438	PGLPMDL NGKFLDNGGSGYLLKPHFLRESKSYFNESE-IKEGMPITLTI
zeta-mouse	475	PGLPMDL NGKFLDNGGSGYLLKPDILRDTTLGFPNEFEYDDHPVTLTI
zeta-rat	475	PGLPMDL NGKFLDNGGSGYLLKPDILRDTTLGFPNEFEYDDHPVTLTI
beta1-human	624	MDLAMQIIMGMVEYNGKSGYRLKPEFMRR---PDKHFDPFTEGIVDGVIA
beta3-mouse	680	LDLPMQIIMGMVEYNGKSGYRLKPEFMRR---PDKSEDPFTEVIVDGVIA
beta2-human	631	MDLPMQIIMGMVEYNGKSGYRLKPEFMRR---PDKQENPFESVDRIDVVA
beta4-mouse	654	PDLAMQIIMGMVEYNGKSGYRLKPEFMRR---PDRTEPFESVDRIDVVA
epsilon-human	1819	DDLPLHLNAAFMFEANGGCGYVLKPPVLWDKNCMPYQKFSPLERDLDSMDP
delta1-human	582	PCPEMDVYQDRQDNGACGYVLKPAFLRDPNGTFNER---ALAQGPWWAR
delta3-human	564	PCPEMDVYQDRQDNGACGYVLKPAFLRDPNGTFNER---EYPGPPR
delta2-bovine	579	AGLEMDLCDGLFRQNGACGYVLKPAFLRDAQSSFHPERFISPFKAQTLII
delta4-rat	592	AGSAMDCDGLFRQNGGSGYVLKPEFLRDTQSSFNEMKIVSLYKAQILVV
gamma1-human	700	NSYALSFRAEGKIKHCRVQOEGQTVMLGN-SEFDSLVDLISYEEKHPLYR
gamma2-human	678	DSYALTFRARGKVKHCRINRDGPHFVLGTSAYFESLVELVSYEEKHSLYR
zeta-monkeyA	520	R GIG PLT-----HSSSNKGDTLV VEG P---NDQM QQT VIKK
zeta-human	487	R GIG PLT-----HSSSNKGDTLV VEG P---NDQM QQT VIKK
zeta-mouse	525	R GIG PVS-----SSS-NTPLIV VYG P---NDHV QQT VIKK
zeta-rat	525	R GIG PVS-----VPS-NTPLIV VYG P---NDHM QQS AVKN
beta1-human	671	NTLSVKIISGQ-----FLSDKKVGTVEVDMEGLPVDTRRKAFKTKTSQG
beta3-mouse	727	NALRVKVISGQ-----FLSDKKVGTVEVDMEGLPVDTR-RKYRTKTSQG
beta2-human	678	TTLSTIVISGQ-----FLSERSVRTVEVELEGLPDPKRRYRTKLSPT
beta4-mouse	701	ATCSVQVISGQ-----FLSDKKIGTYVEVDMEGLPDTIRKEFRTRMVMN
epsilon-human	1869	AVYSLTIVISGQ-----NVCPSNSMGSPCIEVDVLGMPLEDSCHFRTPHR
delta1-human	629	KRLNIRVISGQQLPKVNKNK-NSIVDPKVTVEIHGVSRODVASQAVITN
delta3-human	607	TTLSTIVISGQQLPKVNKNK-NSIVDPKVTVEIHGVSRODVASQAVITN
delta2-bovine	629	QEPWLQVISGQQLPKVDNTKEQSIQVDPVLRVEIFGVRLPDTTQOETSIVEN
delta4-rat	642	QV-----ISGQRLPKVDNTKETTVVDPVLRVELYGVPELTKQOETSIVEN
gamma1-human	749	KMKLRYPINEEALEK--IGTAEPDYGALYEGRNPGFYVEANPMPTFKCAV
gamma2-human	728	KMKLRYPVTEPELLER--YNTER-DINSLYDVSRYVDPSEINPSPMPORTV
zeta-monkeyA	563	NAFSRWNETTFTI HVPELAL RF VESQG-LIAGNEFLGQYTLPLLCM
zeta-human	530	NAFSRWNETTFTI HVPELAL RF VEGQG-LIAGNEFLGQYTLPLLCM
zeta-mouse	567	NAFSRWNETTFTI QVPELAL RF VETQGLLSGNELLGQYTLPLVLCM
zeta-rat	567	NAFSRWNETTFTI QVPELAL RF VETQG-FLSGNELLGQYTLPLVLCM
beta1-human	716	NAVNFVWEFEPIVFKKVLPTLACLRIAYVE---EGGKFVGHRIPLVSAI
beta3-mouse	771	NSFNFWDEEFPDFPKVVLPTLASLRIAAFE---EGGKFVGHRIPLVSAI
beta2-human	723	NSFNFWDEEFPDFPKVVLPTLASLRIAAFE---EGGKFVGHRIPLVSAI
beta4-mouse	746	NGLNFWNEESFVRKVLPLDLAVLRIAYVD---DNNKLGQRIPLPLDGL
epsilon-human	1914	NTLNPMWN-EQFLFRVHFEEDLVFLRFVVEN---NSSAVTAQRIPLPLKAL
delta1-human	678	NGFNFWDEEFAFEVVPDLALIRFLVEDYD-ASSKNDFTIGOSTIPLNSL
delta3-human	657	NGFNFWDEEFAFEVVPDLALIRFLVEDYD-ATSPNDFVGOFTIPLPLSL

delta2-bovine	679	NGFNHYWGQTLCERL	
delta4-rat	687	NGINPYWGETFYEQIQVPELAMLRFFVVKDYS-RTSRNNFEGQYTLPTWTCM	
gamma1-human	797	KALFDYKADREDELTFIKSAIIQNVEKQEGGWWRGDYGGKKQLWFPSNYV	
gamma2-human	775	KALYDYKAKRSELSFCRGALIHNVSKPEGGWWKGDYGTRIQQYFPSNYV	
zeta-monkeyA	612	NKGYR-VPLFSRMGESLEPA-LFV-WYVR-----	
zeta-human	579	NKGYR-IFLFSRMGESLEPA-LFV-WYVR-----	
zeta-mouse	617	NKGYR-VPLFSKSGANLEPS-LFI-WYFRE-----	
zeta-rat	616	NKGYR-VPLFSKSGANLEPS-LFI-WYFRE-----	
beta1-human	763	RPGYHYICLRNERNQPLTLPVAVFYIEVKDYVPDTPYADVIEALSNPRIYV	
beta3-mouse	818	RSGYHYVCLRNEANQPLCLPALLIYTEASDYIPDDHQDYAEALINPIKHV	
beta2-human	770	NSGYHHLCLHSESNMPLTMPALFIIFLEMKDYPGAWADLTVALANPIKEF	
beta4-mouse	793	QAGYRHISLRNEGNKPLSLPTIFCNIVLKYVDPDGFGLDIVDALSDPKKFL	
epsilon-human	1960	KRGYRHLQLRNLHNEVLEISSLFINSRRMEENSNGNTMSASSMFN-----	
delta1-human	727	KQGYRHVHIMSKNGDQHPSATLFFVKISLQD-----	
delta3-human	706	KQGYRHHILSKDGASLSPATLFIQIRIORS-----	
delta2-bovine	728	QQGYRHHILLSKDGLSLHPASIFVHICTQEVSEEAES-----	
delta4-rat	736	KHGYRHVSLLEDGTSLHPASIFVYTCMQEDLDMDEP-----	
gamma1-human	847	EEMVNPVVALEPEREHLDENSPIGDLLRGVLDVPACQIAIRPEGKNNRLFV	
gamma2-human	825	EDISTADFEELEKQIIEDN-PIGSLCRGILDNTYNVVKAPQGNKQKSFV	
zeta-monkeyA		-----	
zeta-human		-----	
zeta-mouse		-----	
zeta-rat		-----	
beta1-human	813	NLMEQRAKQLAALTLEDEEEVKKEADPGETPSEAPSEARTTPAENGVNHT	
beta3-mouse	868	SLMDQRAKQLAALIGESQAQASTETYQETPCQPGSQLPSNPTPNPLDAS	
beta2-human	820	SAHDTKSVKLKEAMG-----GLPEKPFPLASPVASQ	
beta4-mouse	843	SITEKRADQMRAMGIETSDIADVPSDTS----KNDKKGKANPAKANVTPQ	
epsilon-human	2005	-----TEERKCLQTHRVTVH	
delta1-human		-----	
delta3-human		-----	
delta2-bovine		-----	
delta4-rat		-----	
gamma1-human	897	FSISMASVAHWSLDVAADSQEELQDWVKKIREVAQTADARLTEGKIMERR	
gamma2-human	874	FILEPKKEQGDPPVEFATDRVEELFEWFQSIREITWKIDSKENNMKYWEKN	
zeta-monkeyA		-----	
zeta-human		-----	
zeta-mouse		-----	
zeta-rat		-----	
beta1-human	863	TTLTPKPPSQALHSQPAPGSVKAPAKTEDLIQSVLTEVEAQTIIEELKQOK	
beta3-mouse	918	PRWPPGPTTSSTSSSLSS-----PGQRDDLIASILSEVTPTPLEELRSHK	
beta2-human	851	VNGALAPTSNGSPAARAG-----AREEAMKEAAEPRTASLEELRELK	
beta4-mouse	889	SSSELRPPTTAAALGSGQE-----AKKGIELIPQVRIEDLKQMK	
epsilon-human	2020	GVPGPPEPFTVFTINGGTK-----AKQLLQQLILTNEQDIK	
delta1-human		-----	
delta3-human		-----	
delta2-bovine		-----	
delta4-rat		-----	
gamma1-human	947	KKIALELSELVVYCRVPVFDEEKIGTERACYRDMSSFPETKAEKYVNKAK	
gamma2-human	924	QSIAIELSDLVVYCKPT--SKTKDNLENPDFREIRSFVETKADSIIRQKP	
zeta-monkeyA		-----	
zeta-human		-----	
zeta-mouse		-----	
zeta-rat		-----	
beta1-human	913	SFVKLQKKHYKEMKDLVKRHHKKTDLIKEHTTKYNEIQNDYLRRRAALE	
beta3-mouse	963	AMVKLRSRQDRDLRELHKKHQRKAVALTRRLDGLAQARAEGKCRPSPSA	
beta2-human	893	GVVKLQRRHEKELRELERRGARRWEELLQRGAAQLAELGPPGPGVGACK	
beta4-mouse	927	AYLKLHKKQKELNSLKKKHAKHEHSTMQKLHCTQVDKIVAQYDKEKSTHE	
epsilon-human	2054	PVTTDYFLMEEKYFISKEKNECRKQPFQRAIGPEEEIMQILSSWFPEEGY	
delta1-human		-----	
delta3-human		-----	

```

delta2-bovine -----
delta4-rat -----
gamma1-human 997 GKKFLQYNRLQLSRIYPKGQRLDSSNYDPLPMWICGSQ LVALNFQTPDKP
gamma2-human 972 -VDLLKYNQKGLTRVYPKGQ RVDSSNYDPFRLWL CGSQMVALNFQTADKY

zeta-monkeyA -----
zeta-human -----
zeta-mouse -----
zeta-rat -----
beta1-human 963 KSAKKDSKKKSEPSSPDHGSSTIEQDLAALDAEMTQK LIDLKDKQQQQLL
beta3-mouse 1013 LGKATNSEDVKEE-----EEAKQYREFQNRQVQSLL
beta2-human 943 LGPGKGSRRKRS LPREESAG----AAPGEGPEGVDGRVRELKDRLELELL
beta4-mouse 977 KILEKAMKKKGGS-----NCLEIKKET
epsilon-human 2104 MG-----RIVLK
delta1-human -----
delta3-human -----
delta2-bovine -----
delta4-rat -----
gamma1-human 1047 MQMNQALEMTGRHCGYVLQ PSTMRDEAFDPDKSSLRGL EPCAISIEVLG
gamma2-human 1021 MQMNHALFSLNGRTGYVLQ PESMRTEKYDMPPE SQR-KILMTLTVKVLG

zeta-monkeyA -----
zeta-human -----
zeta-mouse -----
zeta-rat -----
beta1-human 1013 NLRQEQYYSEKYQKREHIKLLIQK L TDVAEECONNQLKKLKEICEKEKKE
beta3-mouse 1044 ELREAQADVETKRKLEHLRQA HQRLKEVVLDTHTTQFKRLKELNEREKKE
beta2-human 989 RQGEEQYECVLKRKEQHVAEQISKMMELAREKQA AELKALKETSENDTKE
beta4-mouse 999 EIKIQTLTDDHKS KVKEIVAQHTKEWSEMINTHSAEEQEIRDLHLSQQCE
epsilon-human 2111 TQQENLEEKNI VQDDKEVILSSEESFFVQVHDVSPEQPRTVIKAPRVST
delta1-human -----
delta3-human -----
delta2-bovine -----
delta4-rat -----
gamma1-human 1097 ARHLPKNGRGIVCPFEVEIEVAGAEYDSTKQKTEFVVDNGLNPVWP--AKP
gamma2-human 1070 ARHLPKLGRSIACPFVEVEICGA EYGNNKEKTTVVNDNGLSPIWAPTQEK

zeta-monkeyA -----
zeta-human -----
zeta-mouse -----
zeta-rat -----
beta1-human 1063 LKKKMDKKRQEKIT--EAKSKDKS QMEEEKTEMIRSYIQEVVQYIKRLEE
beta3-mouse 1094 LQKILDRKRNN SIS--EAKTREKHKKEVELTEINRRHITESVNSIRLEE
beta2-human 1039 MKKKLETKRLE RIQGMTKVTTDKMAQERL KREINNSHIQEVVQVIKQMT E
beta4-mouse 1049 LLRKLLINAHEQQTQQLKLSHDRESKEMRAHQAKISMENSKAISQDKSIK
epsilon-human 2161 AQDVIQQTLC KAKYSYSILSNPNPSDYVLL EEVVKD TTNKKTTPKSSQR
delta1-human -----
delta3-human -----
delta2-bovine -----
delta4-rat -----
gamma1-human 1145 FHFQISNPEFAFLRFVVEEDMFSDQNF LAQATFPVKGLKTGYRAVPLKN
gamma2-human 1120 VTFEIYDPNLAFLRFVVEEDMFSDPNFLAHATYPIKAVKSGFRSVPLKN

zeta-monkeyA -----
zeta-human -----
zeta-mouse -----
zeta-rat -----
beta1-human 1111 AQSKRQEKLV EKHKEIRQQILDEKPKLQVELEQEYQDKFKRLPLEILEFV
beta3-mouse 1142 AQKQRHERLVAGQQQVLQQLEEEEPKLLAQLTQECQEQRERLPQEIRRCL
beta2-human 1089 NLERHQEKLEEKQAACLEQIREMEKQFQKEALAEYEARMKGLEAEVKESV
beta4-mouse 1099 NKAERERRVRELNSSNTKKFLEERKRLAMKQSKEMDQLKKVQLEHLEFLE
epsilon-human 2211 VLLDQECVFQAQSKWKGAGKFILKLKEQVQASREDKKKGISFAS ELKKLT
delta1-human -----
delta3-human -----

```

```

delta2-bovine -----
delta4-rat -----
gamma1-human 1195 NYSEDLELASLLIKIDIFPAKENGDLSPFSGTSLRERGSDASGQLFHGRA
gamma2-human 1170 GYSEDIELASLLVFCEMRPVLESEEEELYSSCRQLRRRQEELNNQLFLYDT

zeta-monkeyA -----
zeta-human -----
zeta-mouse -----
zeta-rat -----
beta1-human 1161 QEAMKGKISED SNHGSAPLSLSSDPGKVNHKTPSSEELGGDIPGKEFDTP
beta3-mouse 1192 LGETAEG LGDGPLVACASNGHAPGSGGHLSSADSESQEENTQL-----
beta2-human 1139 RACLRTC FFPSEAKDKPERACECPPELCEQDPLIAKADAQESRL-----
beta4-mouse 1149 KQNEQAKEMQQMVKLEAEMDRRPATVV-----
epsilon-human 2261 KSTKQPRGLTSPSOLLTSESIQTKEEKPVGGLSPVTQWITDSD-----
delta1-human -----
delta3-human -----
delta2-bovine -----
delta4-rat -----
gamma1-human 1245 REGSFESRYQQPFEDFRISQEH LADHFDSRERRAPRRTRVNGDNRL----
gamma2-human 1220 HQNLRNANRDALVKEFSVNENHSSCTRRNATRG-----

zeta-monkeyA -
zeta-human -
zeta-mouse -
zeta-rat -
beta1-human 1211 L
beta3-mouse -
beta2-human -
beta4-mouse -
epsilon-human -
delta1-human -
delta3-human -
delta2-bovine -
delta4-rat -
gamma1-human -
gamma2-human -

```

SEQUENCE LISTING

<110> University of Wales College of Medicine

<120> Sperm factor sequences

<130> PCT/GB2002/004739

<140> PCT/GB2002/004739

<141> 2002-10-18

<160> 11

<170> PatentIn Ver. 2.1

<210> 1

<211> 608

<212> PRT

<213> Homo sapiens

<400> 1

Met Glu Met Arg Trp Phe Leu Ser Lys Ile Gln Asp Asp Phe Arg Gly
1 5 10 15

Gly Lys Ile Asn Leu Glu Lys Thr Gln Arg Leu Leu Glu Lys Leu Asp
20 25 30

Ile Arg Cys Ser Tyr Ile His Val Lys Gln Ile Phe Lys Asp Asn Asp
35 40 45

Arg Leu Lys Gln Gly Arg Ile Thr Ile Glu Glu Phe Arg Ala Ile Tyr
50 55 60

Arg Ile Ile Thr His Arg Glu Glu Ile Ile Glu Ile Phe Asn Thr Tyr
65 70 75 80

Ser Glu Asn Arg Lys Ile Leu Leu Ala Ser Asn Leu Ala Gln Phe Leu
85 90 95

Thr Gln Glu Gln Tyr Ala Ala Glu Met Ser Lys Ala Ile Ala Phe Glu
100 105 110

Ile Ile Gln Lys Tyr Glu Pro Ile Glu Glu Val Arg Lys Ala His Gln
115 120 125

Met Ser Leu Glu Gly Phe Thr Arg Tyr Met Asp Ser Arg Glu Cys Leu
130 135 140

Leu Phe Lys Asn Glu Cys Arg Lys Val Tyr Gln Asp Met Thr His Pro
 145 150 155 160

Leu Asn Asp Tyr Phe Ile Ser Ser Ser His Asn Thr Tyr Leu Val Ser
 165 170 175

Asp Gln Leu Leu Gly Pro Ser Asp Leu Trp Gly Tyr Val Ser Ala Leu
 180 185 190

Val Lys Gly Cys Arg Cys Leu Glu Ile Asp Cys Trp Asp Gly Ala Gln
 195 200 205

Asn Glu Pro Val Val Tyr His Gly Tyr Thr Leu Thr Ser Lys Leu Leu
 210 215 220

Phe Lys Thr Val Ile Gln Ala Ile His Lys Tyr Ala Phe Met Thr Ser
 225 230 235 240

Asp Tyr Pro Val Val Leu Ser Leu Glu Asn His Cys Ser Thr Ala Gln
 245 250 255

Gln Glu Val Met Ala Asp Asn Leu Gln Ala Thr Phe Gly Glu Ser Leu
 260 265 270

Leu Ser Asp Met Leu Asp Asp Phe Pro Asp Thr Leu Pro Ser Pro Glu
 275 280 285

Ala Leu Lys Phe Lys Ile Leu Val Lys Asn Lys Lys Ile Gly Thr Leu
 290 295 300

Lys Glu Thr His Glu Arg Lys Gly Ser Asp Lys Arg Gly Asp Asn Gln
 305 310 315 320

Asp Lys Glu Thr Gly Val Lys Lys Leu Pro Gly Val Met Leu Phe Lys
 325 330 335

Lys Lys Lys Thr Arg Lys Leu Lys Ile Ala Leu Ala Leu Ser Asp Leu
 340 345 350

Val Ile Tyr Thr Lys Ala Glu Lys Phe Lys Ser Phe Gln His Ser Arg
 355 360 365

Leu Tyr Gln Gln Phe Asn Glu Asn Asn Ser Ile Gly Glu Thr Gln Ala
 370 375 380

Arg Lys Leu Ser Lys Leu Arg Val His Glu Phe Ile Phe His Thr Arg
 385 390 395 400

Lys Phe Ile Thr Arg Ile Tyr Pro Lys Ala Thr Arg Ala Asp Ser Ser
 405 410 415
 Asn Phe Asn Pro Gln Glu Phe Trp Asn Ile Gly Cys Gln Met Val Ala
 420 425 430
 Leu Asn Phe Gln Thr Pro Gly Leu Pro Met Asp Leu Gln Asn Gly Lys
 435 440 445
 Phe Leu Asp Asn Gly Gly Ser Gly Tyr Ile Leu Lys Pro His Phe Leu
 450 455 460
 Arg Glu Ser Lys Ser Tyr Phe Asn Pro Ser Asn Ile Lys Glu Gly Met
 465 470 475 480
 Pro Ile Thr Leu Thr Ile Arg Leu Ile Ser Gly Ile Gln Leu Pro Leu
 485 490 495
 Thr His Ser Ser Ser Asn Lys Gly Asp Ser Leu Val Ile Ile Glu Val
 500 505 510
 Phe Gly Val Pro Asn Asp Gln Met Lys Gln Gln Thr Arg Val Ile Lys
 515 520 525
 Lys Asn Ala Phe Ser Pro Arg Trp Asn Glu Thr Phe Thr Phe Ile Ile
 530 535 540
 His Val Pro Glu Leu Ala Leu Ile Arg Phe Val Val Glu Gly Gln Gly
 545 550 555 560
 Leu Ile Ala Gly Asn Glu Phe Leu Gly Gln Tyr Thr Leu Pro Leu Leu
 565 570 575
 Cys Met Asn Lys Gly Tyr Arg Arg Ile Pro Leu Phe Ser Arg Met Gly
 580 585 590
 Glu Ser Leu Glu Pro Ala Ser Leu Phe Val Tyr Val Trp Tyr Val Arg
 595 600 605

<210> 2

<211> 647

<212> PRT

<213> Mouse

<400> 2

Met Glu Ser Gln Leu His Glu Leu Ala Glu Ala Arg Trp Phe Leu Ser
1 5 10 15

Lys Val Gln Asp Asp Phe Arg Gly Gly Lys Ile Asn Val Glu Ile Thr
20 25 30

His Lys Leu Leu Glu Lys Leu Asp Phe Pro Cys His Phe Ala His Val
35 40 45

Lys His Ile Phe Lys Glu Asn Asp Arg Gln Asn Gln Gly Arg Ile Thr
50 55 60

Ile Glu Glu Phe Arg Ala Ile Tyr Arg Cys Ile Val His Arg Glu Glu
65 70 75 80

Ile Thr Glu Ile Phe Asn Thr Tyr Thr Glu Asn Arg Lys Ile Leu Ser
85 90 95

Glu Asn Ser Leu Ile Glu Phe Leu Thr Gln Glu Gln Tyr Glu Met Glu
100 105 110

Ile Asp His Ser Asp Ser Val Glu Ile Ile Asn Lys Tyr Glu Pro Ile
115 120 125

Glu Glu Val Lys Gly Glu Arg Gln Met Ser Ile Glu Gly Phe Ala Arg
130 135 140

Tyr Met Phe Ser Ser Glu Cys Leu Leu Phe Lys Glu Asn Cys Lys Thr
145 150 155 160

Val Tyr Gln Asp Met Asn His Pro Leu Ser Asp Tyr Phe Ile Ser Ser
165 170 175

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

Ile Trp Gly Tyr Val Ser Ala Leu Val Lys Gly Cys Arg Cys Leu Glu
195 200 205

Ile Asp Cys Trp Asp Gly Ser Gln Asn Glu Pro Ile Val Tyr His Gly
210 215 220

Tyr Thr Phe Thr Ser Lys Leu Leu Phe Lys Thr Val Val Gln Ala Ile
225 230 235 240

Asn Lys Tyr Ala Phe Val Thr Ser Asp Tyr Pro Val Val Leu Ser Leu
245 250 255

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

Pro Glu Tyr Asp Asp His Pro Val Thr Leu Thr Ile Arg Ile Ile Ser
515 520 525

Gly Ile Gln Leu Pro Val Ser Ser Ser Ser Asn Thr Pro Asp Ile Val
530 535 540

Val Ile Ile Glu Val Tyr Gly Val Pro Asn Asp His Val Lys Gln Gln
545 550 555 560

Thr Arg Val Val Lys Asn Asn Ala Phe Ser Pro Lys Trp Asn Glu Thr
565 570 575

Phe Thr Phe Leu Ile Gln Val Pro Glu Leu Ala Leu Ile Arg Phe Val
580 585 590

Val Glu Thr Gln Gln Gly Leu Leu Ser Gly Asn Glu Leu Leu Gly Gln
595 600 605

Tyr Thr Leu Pro Val Leu Cys Met Asn Lys Gly Tyr Arg Arg Val Pro
610 615 620

Leu Phe Ser Lys Ser Gly Ala Asn Leu Glu Pro Ser Ser Leu Phe Ile
625 630 635 640

Tyr Val Trp Tyr Phe Arg Glu
645

<210> 3

<211> 1827

<212> DNA

<213> Homo sapiens

<400> 3

atggaaatga gatgggttttt gtcaaagatt caggatgact tcagaggtgg aaaaattaac 60
ctagaaaaaa ctcagagggtt acttgaaaaa ttagatattc ggtgcagtta tattcatgtg 120
aaacagattt ttaaggacaa tgacaggctg aaacaaggaa gaatcaccat agaagaattt 180
agagcaattt atcgaattat cagcacaga gaagaaatta ttgagatttt caacacatat 240
tctgaaaacc ggaaaattct tttagcaagt aatctggctc aatttctgac acaagaacaa 300
tatgcagctg agatgagtaa agctattgct tttgagatca ttcagaaata cgagcctatc 360
gaagaagtta ggaaagcaca ccaaatgtca ttagaagggtt ttacaagata catggattca 420
cgtgaatgtc tactgtttta aaatgaatgt agaaaagttt atcaagatat gactcatcca 480
ttaaatgatt attttatttc atcttcacat aacacatatt tggtatctga tcaattattg 540
ggaccaagtg acctttgggg atatgtaagt gcccttgtga aaggatgccg ttgtttggag 600
attgactgct gggatggagc acaaaatgaa cctgttgtat atcatggcta cacactcaca 660
agcaaacttc tgtttaaaac tgttatccaa gctatacaca agtatgcatt catgacatct 720
gactaccag tgggtgctctc tttagaaaat cactgctcca ctgccaaca agaagtaatg 780

```

gcagacaatt tgcaggctac ttttggagag tccttgcttt ctgatatgct tgatgatattt 840
cctgatactc taccatcacc agaggcacta aaattcaaaa tattagttaa aaataagaaa 900
ataggaacct taaaggaaac ccatgaaaga aaagggttctg ataagcgtgg agacaatcaa 960
gacaaggaaa caggggtaaa aaagttacct ggagtaatgc ttttcaagaa aaagaagacc 1020
aggaagctaa aaattgctct ggccttatct gatcttgtca tttatacgaa agctgagaaa 1080
ttcaaaagct ttcaacattc aagattatat cagcaattta atgaaaataa ttctattggg 1140
gagacacaag cccgaaaact ttcaaaattg cgagtccatg agtttatttt tcacaccagg 1200
aagttcatta ccagaatata tcccaaagca acaagagcag actcttctaa ttttaatccc 1260
caagaattttt ggaatatagg ttgtcaaagt gtggctttta atttccagac ccctgggtctg 1320
cccatggatc tgcaaaatgg gaaatttttg gataatgggtg gttctggata tattttgaaa 1380
ccacatttct taagagagag taaatcatal ttttaaccaa gtaacataaa agagggtatg 1440
ccaattacac ttacaataag gctcatcagt ggtatccagt tgcctcttac tcattcatca 1500
tctaacaaag gtgattcatt agtaattata gaagtttttg gtgttccaaa tgatcaaatg 1560
aagcagcaga ctctgttaat taaaaaaaaat gcttttagtc caagatggaa tgaaacattc 1620
acatttatta ttcatgtccc agaattggca ttgatacggt ttgttggtga aggtcaaggt 1680
ttaatagcag gaaatgaatt tcttgggcaa tatactttgc cacttctatg catgaacaaa 1740
ggttatcgtc gtattcctct gttttccaga atgggtgaga gccttgagcc tgcttcactg 1800
tttgtttatg tttggtacgt cagataa 1827

```

<210> 4

<211> 1944

<212> DNA

<213> Mouse

<400> 4

```

atggaaagcc aacttcatga gctcgcagaa gcaagatggg ttttgtcaaa ggttcaggat 60
gatttttagag gtggaaaaat caacgttgaa attactcaca aactgcttga gaaacttgat 120
ttcccatgcc actttgctca tgtgaaacat atttttaagg aaaatgacag acagaaccaa 180
ggaagaatca ccattgaaga gtttagagcc atttaccggg gtattgtaca tagagaagaa 240
atcacggaga ttttcaacac gtatactgaa aataggaaaa ttctttctga gaacagtctg 300
attgagtttc taaccaaga gcagtatgaa atggagatcg atcactctga ttcagtagag 360
atcatcaata agtatgagcc tattgaagaa gttaaagggtg agcgacagat gtcaattgaa 420
ggtttcgcaa gatacatgtt ttcattcagaa tgtctactgt ttaaagagaa ctgtaaaacc 480
gtgtaccaag atatgaatca tccattaagt gattatttta tttcatcatc tcacaacaca 540
tatttgatat ccgatcaaat attgggaccg agtgacattt ggggatatgt aagtgtctct 600
gtgaaaggct gccgctgtct ggaaattgac tgctgggatg gatcccaaaa tgagccatt 660
gtgtaccatg gttacacatt caccagcaag cttctcttca aaactgtggg ccaagcaata 720
aacaagtatg cctttgtgac atctgattac ccagtagtgc tgtccttaga aaatcactgc 780
tcccctggtc agcaggaagt gatggctagc attctgcaga gcacctttgg agacttcctg 840
ctttcggaca tgcttgagga gtttccagat acactaccgt ctccagaggc actaaaattc 900
aaaatattag tgaaaaacag gaaagtggga accttatctg aaaccacga gaggatagga 960
accgacaaaa gtggccaagt gctagaatgg aaagaagtca tctatgaaga tggatgatgaa 1020
gactcaggaa tggatccaga aacatgggat gtcttctat cacggatcaa ggaggagagg 1080
gaagcagatc cctcgacatt gagtggaata gcaggcgtca agaaaaggaa gaggaagatg 1140
aaaatagcca tggccttata tgatcttgtc atttatacta aggctgagaa gttccgaaac 1200
ttccaatatt caagagtcta tcagcaattt aatgagacca attcgattgg agagtctcga 1260
gctcgaaaac tttccaaatt gagagtccat gagtttattt tccacaccgc ggcattcatc 1320

```

```

accagagtat accccaaaat gatgagagca gactcttcta actttaaccc tcaagagttt 1380
tggaatgtag gatgtcagat ggtggccttg aactttcaaa cccctggact gcctatggat 1440
ttgcaaaacg ggaaattttt ggataatgga ggctctggat atattttgaa gccagacatc 1500
cttagagata caaccctggg ctttaaccca aatgaaccag aatatgacga ccatccagtt 1560
accctcaciaa tccgaatcat cagtgggatc cagttgcctg ttagctcatc ctctaacacg 1620
cctgacatag tagtgatcat agaagtctac ggtgttccaa acgaccacgt gaagcagcag 1680
actcgtgttg ttaagaataa tgcttttagt ccaaagtgga atgaaacatt tacatttctt 1740
attcaagtgc cagaactggc attgatacgt tttgttggtg aaactcaaca aggcttatta 1800
tcaggaaatg aattactcgg gcagtacact ttaccagttc tttgcatgaa caaaggttat 1860
cgtcgtgttc ctctgttttc caaatccggt gcgaaccttg aaccttcctc actgtttatt 1920
tatgtttggt acttcagaga gtga 1944

```

<210> 5

<211> 2187

<212> DNA

<213> Mouse

<400> 5

```

cccccccggc aagccatgcc aactgtgaag gttttaagaa cttagcttct ggacaagttt 60
tcttcgaaag tgaaaagcag tagcagcgag aacagctgat gacggtcaca aaaagacagt 120
gttacttcta agacaagtga caccttagac gaagagccct ctatgggagg acaagcggcc 180
cagatcatga aatcatggaa agccaacttc atgagctcgc agaagcaaga tggtttttgt 240
caaaggttca ggatgatttt agaggtggaa aaatcaacgt tgaaattact cacaactgc 300
ttgagaaact tgatttccca tgccactttg ctcatgtgaa acatattttt aaggaaaatg 360
acagacagaa ccaaggaaga atcaccattg aagagttag agccatttac cggtgtattg 420
tacatagaga agaaatcacg gagattttca acacgtatac tgaaaatagg aaaattcttt 480
ctgagaacag tctgattgag tttctaaccc aagagcagta tgaaatggag atcgatcact 540
ctgattcagt agagatcatc aataagtatg agcctattga agaagtaaag ggtgagcgac 600
agatgtcaat tgaaggtttc gcaagataca tgttttcatc agaatgtcta ctgtttaaag 660
agaactgtaa aaccgtgtac caagatatga atcatccatt aagtgattat tttatttcat 720
catctcacia cacatatttg atatccgatc aaatattggg accgagtgc atttggggat 780
atgtaagtgc tcttgtgaaa ggctgccgct gtctggaaat tgactgctgg gatggatccc 840
aaaatgagcc cattgtgtac catggttaca cattcaccag caagcttctc ttcaaaactg 900
tggtccaagc aataaacaag tatgcctttg tgacatctga ttaccagta gtgctgtcct 960
tagaaaatca ctgctcccct ggtcagcagg aagtgatggc tagcattctg cagagcacct 1020
ttggagactt cctgcttttcg gacatgcttg aggagtttcc agatacacta ccgtctccag 1080
aggcactaaa attcaaaata ttagtgaaaa acaggaaagt gggaacctta tctgaaaccc 1140
acgagaggat aggaaccgac aaaagtggcc aagtgctaga atggaaagaa gtcattctatg 1200
aagatggtga tgaagactca ggaatggatc cagaaacatg ggatgtcttc ctatcacgga 1260
tcaaggagga gaggaagca gatccctcga cattgagtgg aatagcaggc gtcaagaaaa 1320
ggaagaggaa gatgaaaata gccatggcct tatctgatct tgtcatttat actaaggctg 1380
agaagtccg aaacttccaa tattcaagag tctatcagca atttaatgag accaattcga 1440
ttggagagtc tcgagctcga aaactttcca aattgagagt ccatgagttt attttccaca 1500
ccgcggcatt catcaccaga gtatacccca aaatgatgag agcagactct tctaacttta 1560
accctcaaga gttttggaat gtaggatgtc agatgggtggc cttgaacttt caaaccctg 1620
gactgcctat ggatttgcaa aacgggaaat ttttggaata tggaggctct ggatatattt 1680
tgaagccaga catccttaga gatacaaccc tgggctttta ccaaatagaa ccagaatatg 1740

```

```

acgaccatcc agttaccctc acaatccgaa tcatcagtgg gatccagttg cctgttagct 1800
catcctctaa cacgcctgac atagtagtga tcatagaagt ctacggtggt ccaaacgacc 1860
acgtgaagca gcagactcgt gttgttaaga ataatgcttt tagtccaaag tggaatgaaa 1920
catttacatt tcttattcaa gtgccagaac tggcattgat acgttttgtt gttgaaactc 1980
aacaaggctt attatcagga aatgaattac tcgggcagta cactttacca gttctttgca 2040
tgaacaaagg ttatcgtcgt gttcctctgt tttccaaatc cggtgcgaa cttgaacctt 2100
cctcactggt tatttatggt tggtaactca gagagtgcgc catgtaactg acgcgttagc 2160
tacacatcac agtaaacagt ccaaagg                                     2187

```

<210> 6

<211> 1923

<212> DNA

<213> Monkey

<400> 6

```

atggaaatga aatgggttttt gtcgaagatt caggatgact ttagagggtg aaaaattaat 60
ctagaaaaaa ctcagagggtt acttgaaaaa ttagatattc ggtgcagtta tattcatgtg 120
aaacggatat ttaaggacaa tgacaggctg aaacaaggaa gaatcaccat agaagaattt 180
agagcaattt atcgaattct cacgcacaga gaagaaattg ttgagatttt caacgcatat 240
tctgaaaacc ggaaaattct tttagaaaat aatctggttc aattttctgac acaagaacaa 300
tatacaactg agatgagtaa aactattgct tttgagatca ttcagaaata tgaacctatc 360
gaagaagtta ggaaagcacg ccaaattgtc ttagaagggt ttacaagata tatggattca 420
cgtgaatgtc aactatttaa aaatgaatgt agaaaagttt atcaagatat gactcatcca 480
ttaaatgatt attttatttc atcttcacat aacacatatt tggatatctg tcaattagt 540
ggaccaagtg acctttgggg atatgtaagt gcccttgtga aaggatgccg ttgtttggag 600
attgattgct gggatggagc acaaaatgaa cctgttgtat atcatggcta cacactcacc 660
agcaaacttc tgtttaaaac tggtatccaa gctatacaca agtacgcatt catgacatct 720
gactaccagc tgggtgctctc tttagaaaat cactgctccc ctgccagca agaaataatg 780
gcagacaatt tgcagacgac ttttgagagag tccttgcttt ctgatatgct tgctgatttt 840
cctgatactc taccatcacc agaggcacta aaattcaaag tattagttaa aaataagaaa 900
ataggaacct taaaggaaac ccattgaaaga aaagggtctg ataagcgtgg taagggtggag 960
gaatgggaag aagaagtggc agatctggag gaggaggagg aggaggagga gaaattcaaa 1020
gaatcagaaa tattcgaatc tgttttagga gaaaatcaag acaaggaaac aggggtaaaa 1080
aagttatctg gagtaacgct tttcaagaaa aagaagacca ggaagctaaa aattgctctg 1140
gccttatctg atcttgtcat ttatactaaa gctgagaagt tcaaaagctt tcaacattca 1200
agattatata agcaatttaa tgaataaat tctattgggg agacacaagc ccgaaaactt 1260
tcaaaattga gagcccatga gtttattttt cacaccagga agttcattac cagaatatat 1320
cccaaagcaa caagagcaga ctcttctaatt tttaatcccc agaatttttg gaatataggt 1380
tgtcaaatgg tggcttttaa tttccagacc cctgggtctgc ctatggatct gcaaaatggg 1440
aaatttttgg ataatggtgg ttctggatat attttgaaac cacatttctt aagagagagt 1500
gaatcatact ttaacccaag tgacataaaa gacagtatgc caattacact tacaataagg 1560
ctcatcagtg gtatccagtt gcctcttact cattcatcat ctaacaaagg tgatacat 1620
gtaattatag aagtttttgg tgttccaaat gatcaaatga agcagcagac tcgtgtaatt 1680
aaaaaaaaatg ctttttagtcc aagatggaat gaaacattta catttattat tcatgtccca 1740
gaattggcat tgatacgttt tgttggtgaa agtcaagggt taatagcagg aaatgaattt 1800
cttgggcaat atactttgcc acttctatgc atgaacaaag gctatcgtcg tgttcctctg 1860
ttttcccgaa tgggtgagag ccttgagcct gcttcaactgt ttgtttatgt ttggtacgtc 1920

```

aga

1923

<210> 7

<211> 641

<212> PRT

<213> Monkey

<400> 7

Met	Glu	Met	Lys	Trp	Phe	Leu	Ser	Lys	Ile	Gln	Asp	Asp	Phe	Arg	Gly
1				5					10					15	

Gly	Lys	Ile	Asn	Leu	Glu	Lys	Thr	Gln	Arg	Leu	Leu	Glu	Lys	Leu	Asp
			20					25					30		

Ile	Arg	Cys	Ser	Tyr	Ile	His	Val	Lys	Arg	Ile	Phe	Lys	Asp	Asn	Asp
		35					40					45			

Arg	Leu	Lys	Gln	Gly	Arg	Ile	Thr	Ile	Glu	Glu	Phe	Arg	Ala	Ile	Tyr
	50					55					60				

Arg	Ile	Leu	Thr	His	Arg	Glu	Glu	Ile	Val	Glu	Ile	Phe	Asn	Ala	Tyr
65					70					75					80

Ser	Glu	Asn	Arg	Lys	Ile	Leu	Leu	Glu	Asn	Asn	Leu	Val	Gln	Phe	Leu
				85					90					95	

Thr	Gln	Glu	Gln	Tyr	Thr	Thr	Glu	Met	Ser	Lys	Thr	Ile	Ala	Phe	Glu
			100					105					110		

Ile	Ile	Gln	Lys	Tyr	Glu	Pro	Ile	Glu	Glu	Val	Arg	Lys	Ala	Arg	Gln
		115					120					125			

Met	Ser	Leu	Glu	Gly	Phe	Thr	Arg	Tyr	Met	Asp	Ser	Arg	Glu	Cys	Gln
	130					135					140				

Leu	Phe	Lys	Asn	Glu	Cys	Arg	Lys	Val	Tyr	Gln	Asp	Met	Thr	His	Pro
145					150					155					160

Leu	Asn	Asp	Tyr	Phe	Ile	Ser	Ser	Ser	His	Asn	Thr	Tyr	Leu	Val	Ser
			165						170					175	

Asp	Gln	Leu	Val	Gly	Pro	Ser	Asp	Leu	Trp	Gly	Tyr	Val	Ser	Ala	Leu
		180						185					190		

Val	Lys	Gly	Cys	Arg	Cys	Leu	Glu	Ile	Asp	Cys	Trp	Asp	Gly	Ala	Gln
		195				200					205				

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

WO 03/035678

PCT/GB02/04739

Ala Leu Asn Phe Gln Thr Pro Gly Leu Pro Met Asp Leu Gln Asn Gly
 465 470 475 480

Lys Phe Leu Asp Asn Gly Gly Ser Gly Tyr Ile Leu Lys Pro His Phe
 485 490 495

Leu Arg Glu Ser Glu Ser Tyr Phe Asn Pro Ser Asp Ile Lys Asp Ser
 500 505 510

Met Pro Ile Thr Leu Thr Ile Arg Leu Ile Ser Gly Ile Gln Leu Pro
 515 520 525

Leu Thr His Ser Ser Ser Asn Lys Gly Asp Thr Leu Val Ile Ile Glu
 530 535 540

Val Phe Gly Val Pro Asn Asp Gln Met Lys Gln Gln Thr Arg Val Ile
 545 550 555 560

Lys Lys Asn Ala Phe Ser Pro Arg Trp Asn Glu Thr Phe Thr Phe Ile
 565 570 575

Ile His Val Pro Glu Leu Ala Leu Ile Arg Phe Val Val Glu Ser Gln
 580 585 590

Gly Leu Ile Ala Gly Asn Glu Phe Leu Gly Gln Tyr Thr Leu Pro Leu
 595 600 605

Leu Cys Met Asn Lys Gly Tyr Arg Arg Val Pro Leu Phe Ser Arg Met
 610 615 620

Gly Glu Ser Leu Glu Pro Ala Ser Leu Phe Val Tyr Val Trp Tyr Val
 625 630 635 640

Arg

<210> 8

<211> 1920

<212> DNA

<213> Monkey

<400> 8

atggaaatga aatgggttttt gtcgaagatt caggatgact ttagaggttg aaaaattaat 60
 ctgaaaaaaa ctcagaggtt acttgaaaaa ttagatattc ggtgcagtta tattcatgtg 120
 aaacggatat ttaaggacaa tgacaggctg aaacaaggaa gaatcaccat agaagaattt 180
 agagcaattt atcgaattct cacgcacaga gaagaaattg ttgagatttt caacgcatat 240
 tctgaaaacc ggaaaattct tttagaaaat aatctggttc aatttctgac acaagaacaa 300

tatacaactg agatgagtaa aactattgct tttagatca ttcagaaata tgaacctatc 360
 gaagaagtta ggaaagcacg ccaaagtca ttagaagggt ttacaagata tatggattca 420
 cgtgaatgtc aactatttaa aaatgaatgt agaaaagttt atcaagatat gactcatcca 480
 ttaaatgatt atttttatttc atcttcacat aacacatatt tggatatctga tcaattagtg 540
 ggaccaagtg acctttgggg atatgtaagt gcccttgtga aaggatgccg ttgtttggag 600
 attgattgct gggatggagc acaaaatgaa cctgttgtat atcatggcta cacactcacc 660
 agcaaacttc tgtttaaaac tgttatccaa gctatacaca agtacgcatt catgacatct 720
 gactacccag tggtgctctc tttggaaaat cactgctccc ctgcccagca agaaataatg 780
 gcagacaatt tgcagacgac ttttggagag tccttgcttt ctgatatgct tgctgatttt 840
 cctgatactc taccatcacc agaggcacta aaattcaaag tattagttaa aaataagaaa 900
 ataggaacct taaaggaaac ccatgaaaga aaagggtctg ataagcgtgg taagggtggag 960
 gaatgggaag aagaagtggc agatctggag gaggaggagg aggaggagag attcaaagaa 1020
 tcagaaatat tcgaatctgt tttaggagaa aatcaagaca aggaaacagg ggtaaaaaag 1080
 ttatctggag taacgctttt caagaaaaag aagaccagga agctaaaaat tgctctggcc 1140
 ttatctgate ttgtcattta tactaaagct gagaagttca aaagctttca acattcaaga 1200
 ttatatcagc aatttaataa aaataattct attggggaga cacaagcccg aaaactttca 1260
 aaattgagag cccatgagtt tattttttcac accaggaagt tcattaccag aatatatccc 1320
 aaagcaacaa gagcagactc ttctaatttt aatccccaag aattttggaa tataggttgt 1380
 caaatggtgg ctttaaattt ccagaccctt ggtctgccta tggatctgca aaatgggaaa 1440
 tttttggata atgggtggtc tggatatatt ttgaaaccac atttcttaag agagagtga 1500
 tcatacttta acccaagtga cataaaagac agtatgccaa ttacacttac aataaggctc 1560
 atcagtggta tccagttgcc tcttactcat tcatcatcta acaaagggtga tacattagta 1620
 attatagaag tttttggtgt tccaaatgat caaatgaagc agcagactcg tgtaattaaa 1680
 aaaaatgctt ttagtccaag atggaatgaa acatttacat ttattattca tgtcccagaa 1740
 ttggcattga tacgttttgt tggtgaaagt caaggtttaa tagcaggaaa tgaatttctt 1800
 gggcaatata ctttgccact tctatgcatg aacaaaggct atcgctcgtgt tcctctgttt 1860
 tcccgaatgg gtgagagcct tgagcctgct tctctgtttg tttatgtttg gtacgtcaga 1920

<210> 9

<211> 640

<212> PRT

<213> Monkey

<400> 9

Met Glu Met Lys Trp Phe Leu Ser Lys Ile Gln Asp Asp Phe Arg Gly
 1 5 10 15

Gly Lys Ile Asn Leu Glu Lys Thr Gln Arg Leu Leu Glu Lys Leu Asp
 20 25 30

Ile Arg Cys Ser Tyr Ile His Val Lys Arg Ile Phe Lys Asp Asn Asp
 35 40 45

Arg Leu Lys Gln Gly Arg Ile Thr Ile Glu Glu Phe Arg Ala Ile Tyr
 50 55 60

Arg Ile Leu Thr His Arg Glu Glu Ile Val Glu Ile Phe Asn Ala Tyr

65						70										75						80
Ser	Glu	Asn	Arg	Lys	Ile	Leu	Leu	Glu	Asn	Asn	Leu	Val	Gln	Phe	Leu							
				85					90					95								
Thr	Gln	Glu	Gln	Tyr	Thr	Thr	Glu	Met	Ser	Lys	Thr	Ile	Ala	Phe	Glu							
			100					105					110									
Ile	Ile	Gln	Lys	Tyr	Glu	Pro	Ile	Glu	Glu	Val	Arg	Lys	Ala	Arg	Gln							
		115					120					125										
Met	Ser	Leu	Glu	Gly	Phe	Thr	Arg	Tyr	Met	Asp	Ser	Arg	Glu	Cys	Gln							
	130					135					140											
Leu	Phe	Lys	Asn	Glu	Cys	Arg	Lys	Val	Tyr	Gln	Asp	Met	Thr	His	Pro							
145					150					155					160							
Leu	Asn	Asp	Tyr	Phe	Ile	Ser	Ser	Ser	His	Asn	Thr	Tyr	Leu	Val	Ser							
				165					170					175								
Asp	Gln	Leu	Val	Gly	Pro	Ser	Asp	Leu	Trp	Gly	Tyr	Val	Ser	Ala	Leu							
			180					185					190									
Val	Lys	Gly	Cys	Arg	Cys	Leu	Glu	Ile	Asp	Cys	Trp	Asp	Gly	Ala	Gln							
		195					200					205										
Asn	Glu	Pro	Val	Val	Tyr	His	Gly	Tyr	Thr	Leu	Thr	Ser	Lys	Leu	Leu							
		210				215						220										
Phe	Lys	Thr	Val	Ile	Gln	Ala	Ile	His	Lys	Tyr	Ala	Phe	Met	Thr	Ser							
225					230					235				240								
Asp	Tyr	Pro	Val	Val	Leu	Ser	Leu	Glu	Asn	His	Cys	Ser	Pro	Ala	Gln							
				245					250					255								
Gln	Glu	Ile	Met	Ala	Asp	Asn	Leu	Gln	Thr	Thr	Phe	Gly	Glu	Ser	Leu							
			260					265					270									
Leu	Ser	Asp	Met	Leu	Ala	Asp	Phe	Pro	Asp	Thr	Leu	Pro	Ser	Pro	Glu							
		275					280					285										
Ala	Leu	Lys	Phe	Lys	Val	Leu	Val	Lys	Asn	Lys	Lys	Ile	Gly	Thr	Leu							
		290				295					300											
Lys	Glu	Thr	His	Glu	Arg	Lys	Gly	Ser	Asp	Lys	Arg	Gly	Lys	Val	Glu							
305					310					315					320							
Glu	Trp	Glu	Glu	Glu	Val	Ala	Asp	Leu	Glu	Glu	Glu	Glu	Glu	Glu	Glu							

	325		330		335
Arg Phe Lys Glu Ser Glu Ile Phe Glu Ser Val Leu Gly Glu Asn Gln					
	340		345		350
Asp Lys Glu Thr Gly Val Lys Lys Leu Ser Gly Val Thr Leu Phe Lys					
	355		360		365
Lys Lys Lys Thr Arg Lys Leu Lys Ile Ala Leu Ala Leu Ser Asp Leu					
	370		375		380
Val Ile Tyr Thr Lys Ala Glu Lys Phe Lys Ser Phe Gln His Ser Arg					
	385		390		400
Leu Tyr Gln Gln Phe Asn Glu Asn Asn Ser Ile Gly Glu Thr Gln Ala					
		405	410		415
Arg Lys Leu Ser Lys Leu Arg Ala His Glu Phe Ile Phe His Thr Arg					
		420	425		430
Lys Phe Ile Thr Arg Ile Tyr Pro Lys Ala Thr Arg Ala Asp Ser Ser					
		435	440		445
Asn Phe Asn Pro Gln Glu Phe Trp Asn Ile Gly Cys Gln Met Val Ala					
		450	455		460
Leu Asn Phe Gln Thr Pro Gly Leu Pro Met Asp Leu Gln Asn Gly Lys					
		465	470		475
Phe Leu Asp Asn Gly Gly Ser Gly Tyr Ile Leu Lys Pro His Phe Leu					
		485	490		495
Arg Glu Ser Glu Ser Tyr Phe Asn Pro Ser Asp Ile Lys Asp Ser Met					
		500	505		510
Pro Ile Thr Leu Thr Ile Arg Leu Ile Ser Gly Ile Gln Leu Pro Leu					
		515	520		525
Thr His Ser Ser Ser Asn Lys Gly Asp Thr Leu Val Ile Ile Glu Val					
		530	535		540
Phe Gly Val Pro Asn Asp Gln Met Lys Gln Gln Thr Arg Val Ile Lys					
		545	550		555
Lys Asn Ala Phe Ser Pro Arg Trp Asn Glu Thr Phe Thr Phe Ile Ile					
		565	570		575
His Val Pro Glu Leu Ala Leu Ile Arg Phe Val Val Glu Ser Gln Gly					

	580		585		590										
Leu	Ile	Ala	Gly	Asn	Glu	Phe	Leu	Gly	Gln	Tyr	Thr	Leu	Pro	Leu	Leu
	595						600					605			
Cys	Met	Asn	Lys	Gly	Tyr	Arg	Arg	Val	Pro	Leu	Phe	Ser	Arg	Met	Gly
	610					615					620				
Glu	Ser	Leu	Glu	Pro	Ala	Ser	Leu	Phe	Val	Tyr	Val	Trp	Tyr	Val	Arg
625					630					635					640

<210> 10

<211> 1938

<212> DNA

<213> Rat

<400> 10

```

atgcatgctc gagcggccgc cagtgtgatg gatatctgca gaattcgccc ttttcaggat 60
gatttcagag gtggaaaaat cagcgtgga attactcaca aactgctcga gaaacttgat 120
ttcccatgcc attttgctca tgtgaaacgt atttttaagg aaaatgacag acataaccaa 180
ggaagaatca ccaccgaaga ttttagaacc atctatcggg gtattgtaca tagagaagag 240
atcgttgaga ttttcaacac gtatactgaa aacaggaaaa ttctccccga ggacagtctg 300
attgaatttc taaccaaga gcagtatgaa atggagatgg atgagtccag ttcagtggag 360
atcatccaga agtacgagcc cattgcagaa gtaaagaacg agcggcagat gtcaattgaa 420
ggttttgcaa gatacatgtt ttcttcagaa tgtctactgt ttaaagagac gtgtaacaca 480
gtgtaccaag atatgaataa gccactaat gattactata tttcatcgtc tcacaacaca 540
tatttgatat ctgatcaaatt attgggacca agtgacattt ggggatatat aagtgccctc 600
gtgaaagggt gccgctgtct ggaaattgac tgctgggatg gagcacaaaa tgaaccatt 660
gtgtaccatg gctacactct caccagcaag cttctcttca aaaccgttat ccaagcaata 720
aacaagtacg ctttcgtgac gtctgattac ccagtgggtg tgccttaga gaatcactgc 780
tcccctggtc aacaggaagt gatgaccgac attctgcaga gtacctttgg agactttctg 840
ctctcagaca tacttgacga gtttccagac agtttgccat ctccagaggc actgaaattt 900
aaaatattag tgaanaataa gaaagttgga acctatctg aaaccgcga gaggctggg 960
actgacaaaa ggggcatagc gctagacttg gaagaagaaa tctatgaaaa tgaagacgaa 1020
gactcaggaa aggagccaga aacgtgggat gatttcctgt cacgggttaa ggaggagcag 1080
gaggcagacc cctcaacggt gagcgggaata gcagatgcca agaaaaagat caggaagcta 1140
agagtagctc tggccttata tgatcttgct atttatacca agctgagaa gttccgaaac 1200
ttccaatatt caagagtcta tcagcagttt aatgagacca cttctatggg agagtctcga 1260
gctcgaaaac tttcaaaatt gagagcccat gagtttattt tccacactgc agcgttcatc 1320
accagagtgt accccaagtt cacgagagca gactcttcta attttaatcc tcaagagttt 1380
tggaatgtgg gctgtcagat ggtggccttg aattttcaaa cccctggact gcctatggat 1440
ttgcaaaacg ggaaattttt ggataatgga ggctctggat atgttttgaa gccagacttt 1500
cttagagaca caactttggg ctttaaccca aatgaaccag aaggagatgg ccatccggtt 1560
accctcacga tccgactcat cagtgggatc cagttgcctg ttaacgtgcc ctcaaataca 1620

```

WO 03/035678

PCT/GB02/04739

tctgacataa tagtgatcat agaagtctac ggtgtcccaa acgaccacat gaagcagcag 1680
 agtcgtgccg ttaagaacaa tgcttttagt ccaaggtgga atgaaacatt cacatttctt 1740
 attcaagtgc cagaattggc actgatacgt ttcgttggtg aaactcaagg cttcctatcg 1800
 ggaaatgaat tacttgggca gtacacttta cccgttcttt gcatgaacaa aggttatcgt 1860
 cgtgttcctc tgttttccaa atccggtgcg aaccttgaac cttcctctct gtttatattac 1920
 gtttggtact acagagag 1938

<210> 11

<211> 646

<212> PRT

<213> Rat

<400> 11

Met His Ala Arg Ala Ala Ala Ser Val Met Asp Ile Cys Arg Ile Arg
 1 5 10 15

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

His Lys Leu Leu Glu Lys Leu Asp Phe Pro Cys His Phe Ala His Val
 35 40 45

Lys Arg Ile Phe Lys Glu Asn Asp Arg His Asn Gln Gly Arg Ile Thr
 50 55 60

Thr Glu Asp Phe Arg Thr Ile Tyr Arg Cys Ile Val His Arg Glu Glu
 65 70 75 80

Ile Val Glu Ile Phe Asn Thr Tyr Thr Glu Asn Arg Lys Ile Leu Pro
 85 90 95

Glu Asp Ser Leu Ile Glu Phe Leu Thr Gln Glu Gln Tyr Glu Met Glu
 100 105 110

Met Asp Glu Ser Ser Ser Val Glu Ile Ile Gln Lys Tyr Glu Pro Ile
 115 120 125

Ala Glu Val Lys Asn Glu Arg Gln Met Ser Ile Glu Gly Phe Ala Arg
 130 135 140

Tyr Met Phe Ser Ser Glu Cys Leu Leu Phe Lys Glu Thr Cys Asn Thr
 145 150 155 160

Val Tyr Gln Asp Met Asn Lys Pro Leu Asn Asp Tyr Tyr Ile Ser Ser
 165 170 175

Ser His Asn Thr Tyr Leu Ile Ser Asp Gln Ile Leu Gly Pro Ser Asp

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

WO 03/035678

PCT/GB02/04739

435	440	445
Arg Ala Asp Ser Ser Asn Phe Asn Pro Gln Glu Phe Trp Asn Val Gly		
450	455	460
Cys Gln Met Val Ala Leu Asn Phe Gln Thr Pro Gly Leu Pro Met Asp		
465	470	475 480
Leu Gln Asn Gly Lys Phe Leu Asp Asn Gly Gly Ser Gly Tyr Val Leu		
	485	490 495
Lys Pro Asp Phe Leu Arg Asp Thr Thr Leu Gly Phe Asn Pro Asn Glu		
	500	505 510
Pro Glu Gly Asp Gly His Pro Val Thr Leu Thr Ile Arg Leu Ile Ser		
	515	520 525
Gly Ile Gln Leu Pro Val Asn Val Pro Ser Asn Thr Ser Asp Ile Ile		
	530	535 540
Val Ile Ile Glu Val Tyr Gly Val Pro Asn Asp His Met Lys Gln Gln		
545	550	555 560
Ser Arg Ala Val Lys Asn Asn Ala Phe Ser Pro Arg Trp Asn Glu Thr		
	565	570 575
Phe Thr Phe Leu Ile Gln Val Pro Glu Leu Ala Leu Ile Arg Phe Val		
	580	585 590
Val Glu Thr Gln Gly Phe Leu Ser Gly Asn Glu Leu Leu Gly Gln Tyr		
	595	600 605
Thr Leu Pro Val Leu Cys Met Asn Lys Gly Tyr Arg Arg Val Pro Leu		
	610	615 620
Phe Ser Lys Ser Gly Ala Asn Leu Glu Pro Ser Ser Leu Phe Ile Tyr		
625	630	635 640
Val Trp Tyr Tyr Arg Glu		
	645	

CLAIMS

1. A diagnostic method for determining the fertility status of a mammal, the method comprising determining in a test sample obtained from the mammal a presence or absence of:

a polypeptide comprising the sequence of SEQ ID NO: 1, being human PLC ζ polypeptide, or a sequence 70% identical thereto; or

a nucleic acid comprising a sequence encoding human PLC ζ polypeptide capable of triggering calcium oscillations in oocytes and identified herein by SEQ ID NO: 3, or a sequence 70% identical thereto, or a nucleic acid which hybridises to a nucleic acid comprising the sequence of SEQ ID NO: 3 under stringent conditions,

wherein the absence of said polypeptide or said nucleic acid indicates that said mammal is infertile.

2. A diagnostic or screening method comprising:

- (a) sequencing the genetic code of a nucleic acid molecule of human PLC ζ gene comprised in a test sample isolated from a human individual; and
- (b) comparing a region of the code obtained from the test sample with the corresponding region of a wild type human PLC ζ nucleic acid sequence of SEQ ID NO: 3 to determine if there are any variations,

whereby a variation in the sample code relative to the predetermined sequence is indicative of a condition, corresponding to lowered fertility or infertility, associated with disruption in calcium oscillation patterns that are a prerequisite to normal biological function absent in the condition.

3. A screening method for screening a human individual suspected of a fertility problem, the screening method comprising the following steps:

- (a) sequencing a nucleic acid of human PLC ζ gene or a polypeptide encoded thereby to determine a nucleic acid sequence or an amino acid sequence, respectively, said nucleic acid or said polypeptide being comprised in a test sample obtained from the human individual; and
- (b) comparing the nucleic acid sequence with the wild type PLC ζ gene sequence of SEQ ID NO:3 or comparing the amino acid sequence with the wild-type PLC ζ amino acid sequence of SEQ ID NO:1, to detect a variant,

wherein detection of a variant indicates a fertility problem.

4. The method of any one of claims 1 to 3, wherein the test sample comprises genomic DNA.
5. A fertility diagnostic or screening product for a mammal, the product comprising, in the alternative:
 - (a) an oligonucleotide having a nucleic acid sequence corresponding to a region of a PLC ζ variant, the region incorporating at least one variation from a corresponding wild-type PLC ζ gene sequence of SEQ ID NO:3;
 - (b) an oligonucleotide having a nucleic acid sequence corresponding to the wild-type PLC ζ gene sequence in the region specified in (a);
 - (c) an oligonucleotide having a nucleic acid sequence corresponding to a specific region of the wild-type PLC ζ gene sequence of SEQ ID NO:3, the specific region comprising a sequence not otherwise present in the genomic DNA of the mammal;
 - (d) antibodies, or monoclonal antibodies, raised to any peptide sequence encoded by any oligonucleotide specific to any one of (a) to (c) above; or
 - (e) any combination of components (a) to (d).
6. The fertility product of claim 5, further comprising one or more reagents suitable for carrying out PCR for amplifying desired regions of the mammal's DNA.
7. The fertility product of claim 5 or 6, wherein any of the components (a) to (c) comprise(s) a plurality of said oligonucleotides immobilised on a solid support.

FIGURE 1

Expression of mouse PLC-zeta in CHO cells

Plot of calcium concentration (nM; ordinate) with time (secs; abscissa)

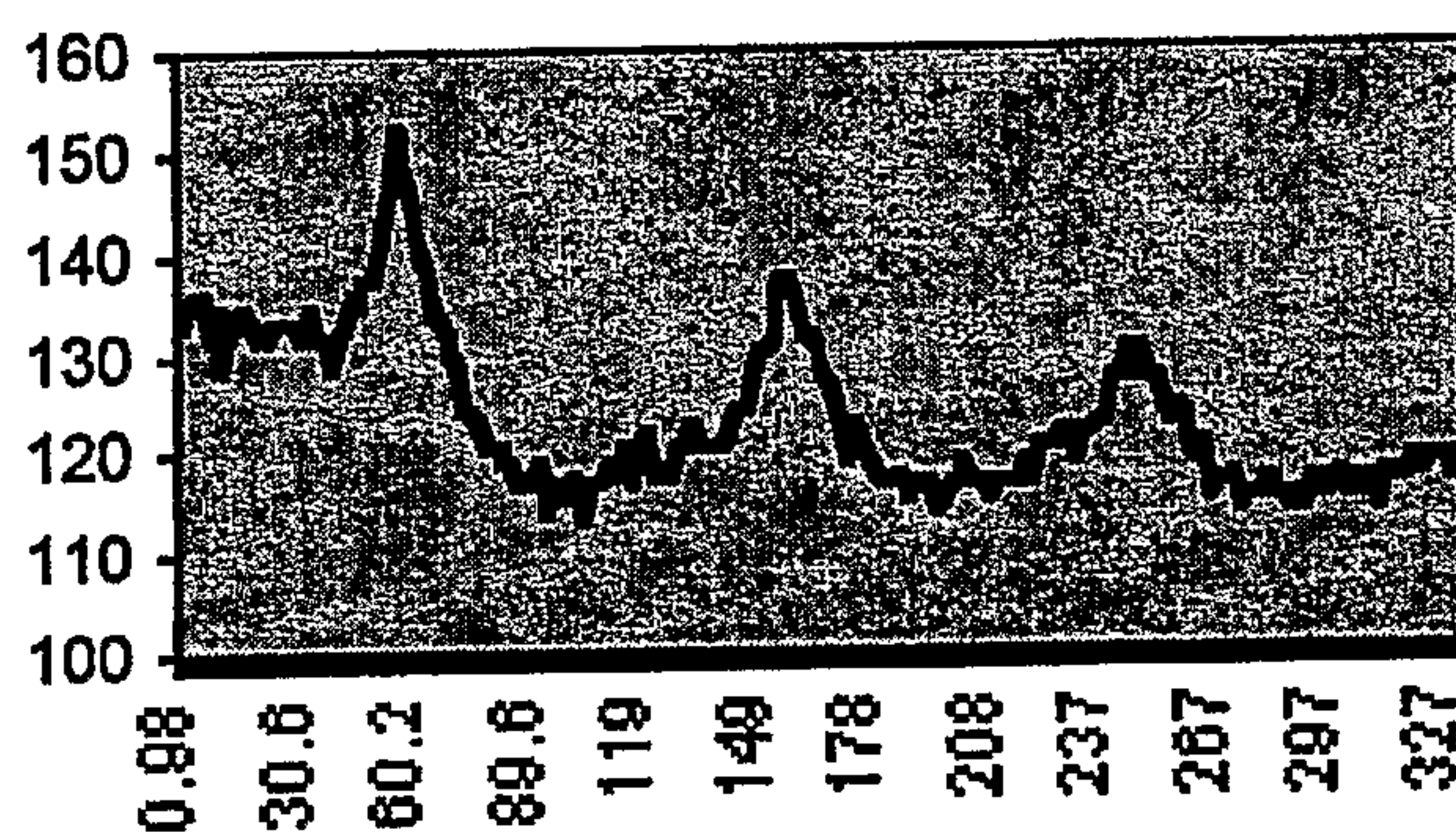


FIGURE 2

Expression of mouse PLC-zeta complementary RNA by
micro-injection into Mouse Eggs

Plot of calcium concentration (ordinate) with time (abscissa)

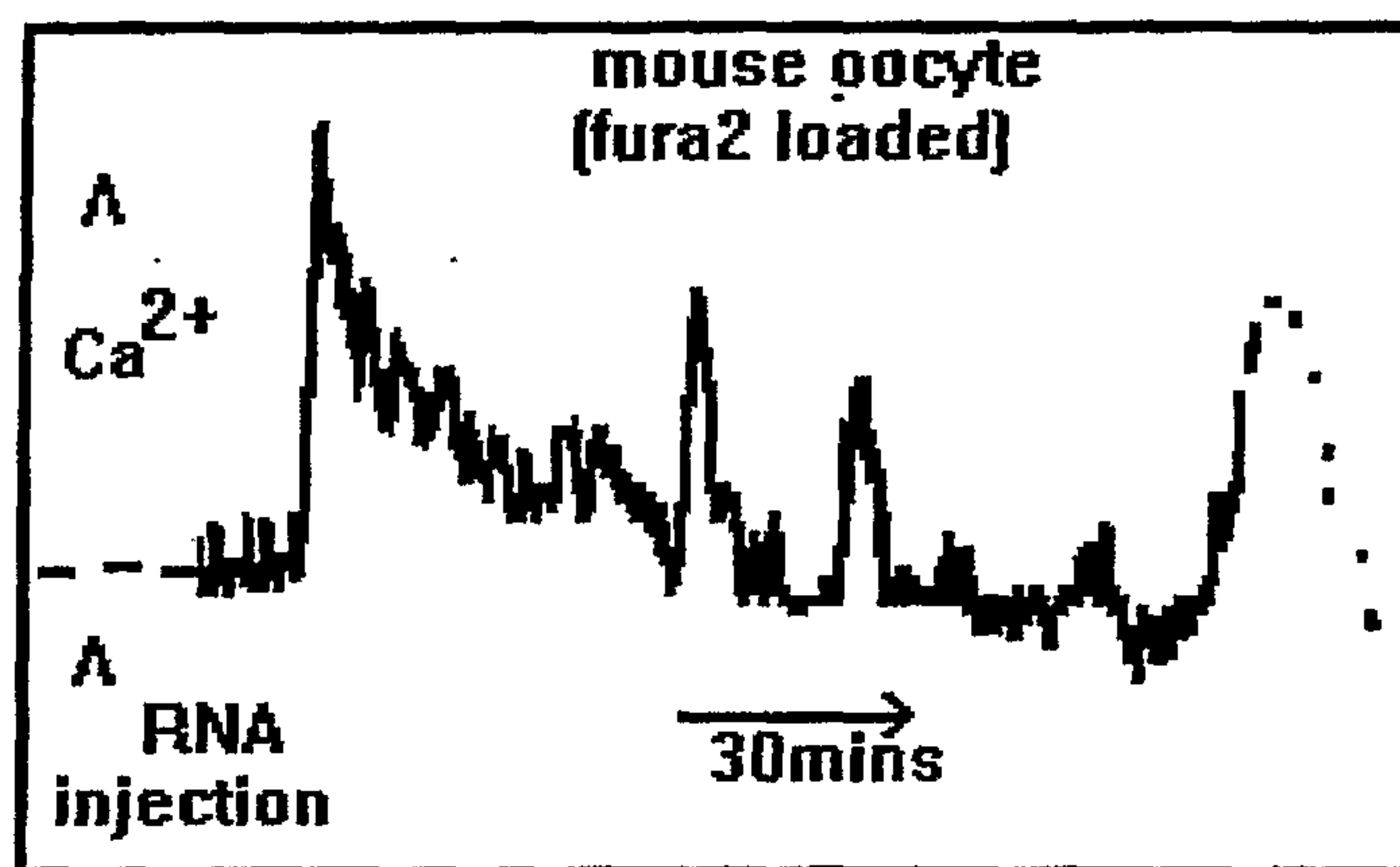


FIGURE 3

Schematic Alignment of PLC Regions

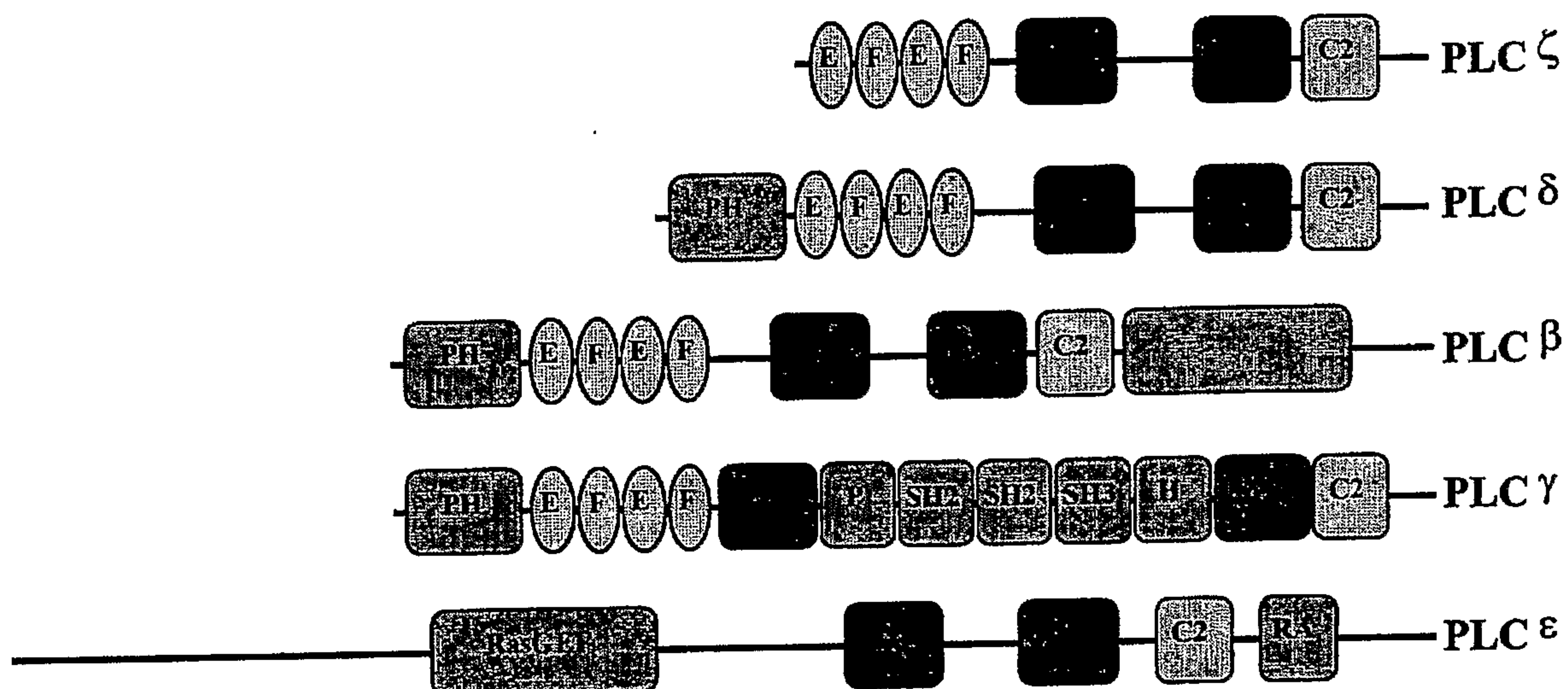
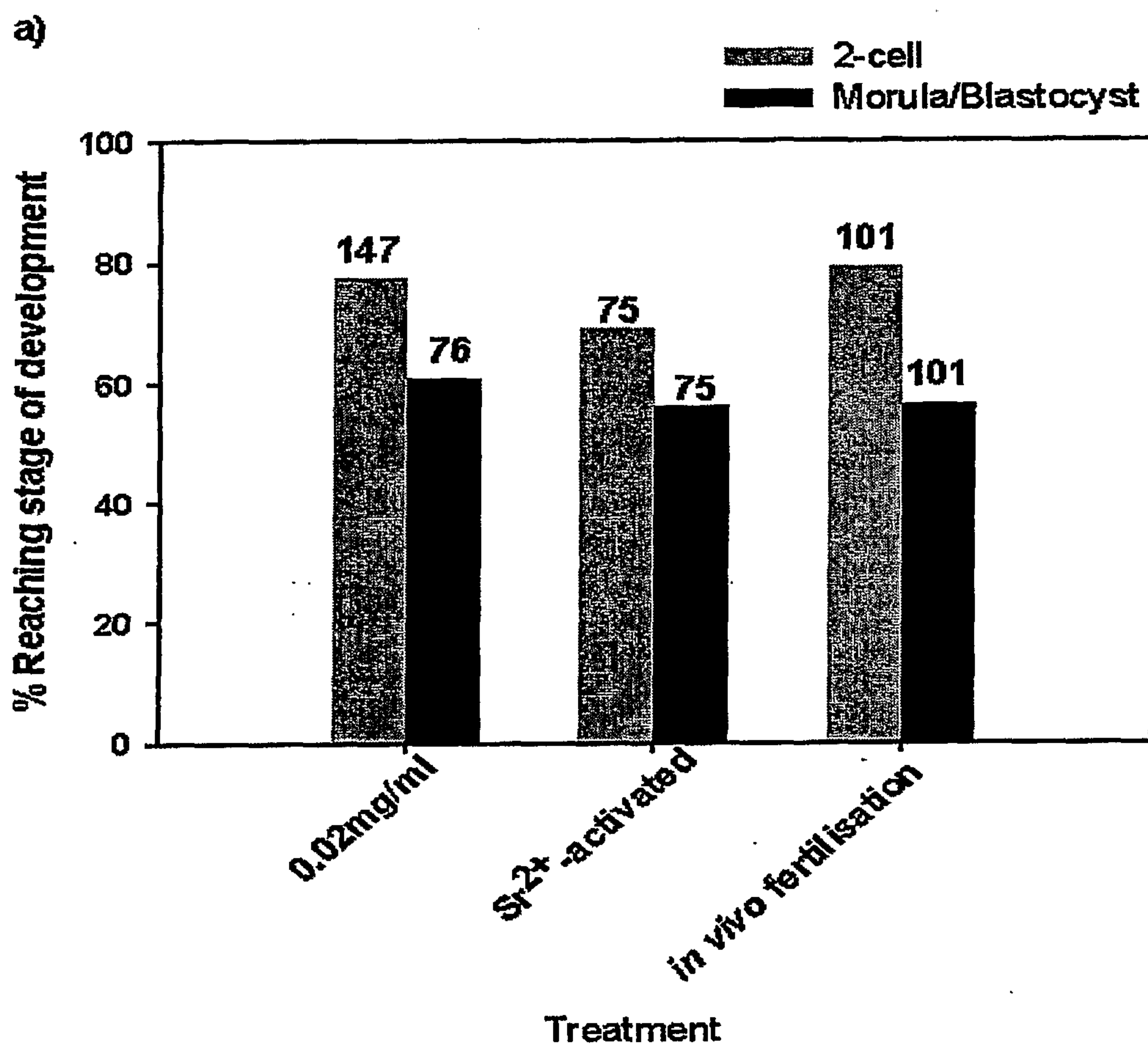


FIGURE 4

Activation and Embryo Development to Blastocyst in
PL C-zeta-injected Mouse Eggs



b)

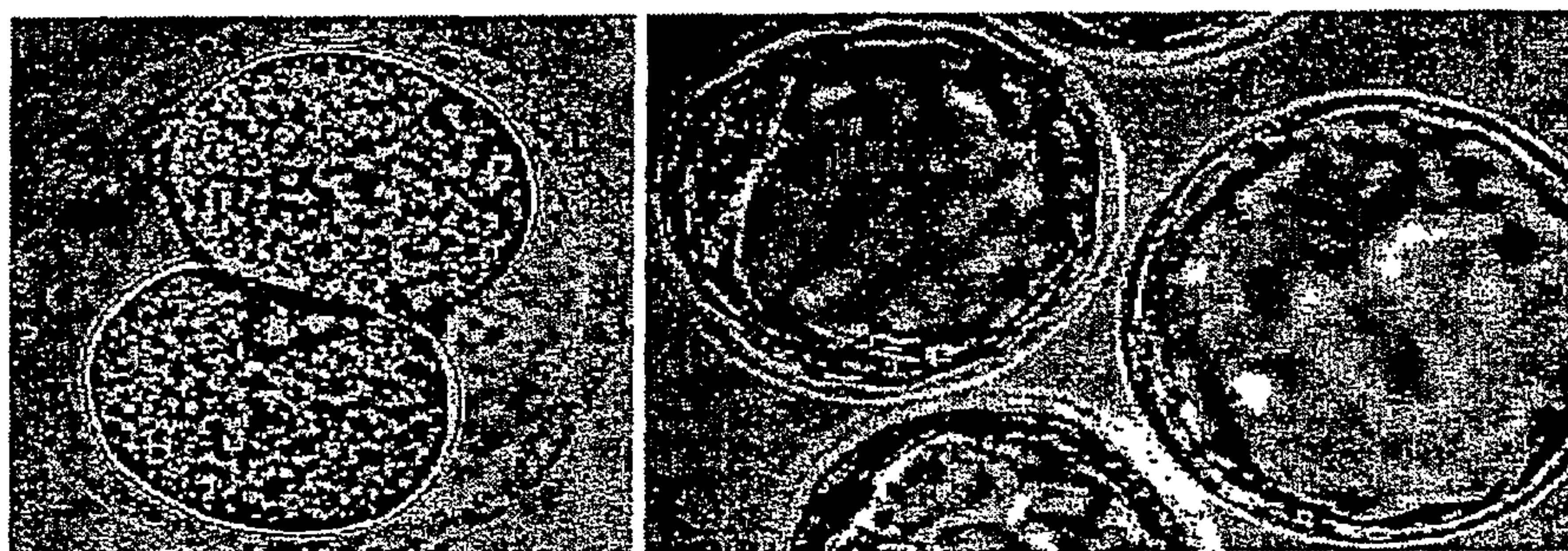


FIGURE 5

Activation and Embryo Development to Blastocyst in
D210R PLC-zeta-injected Mouse Eggs

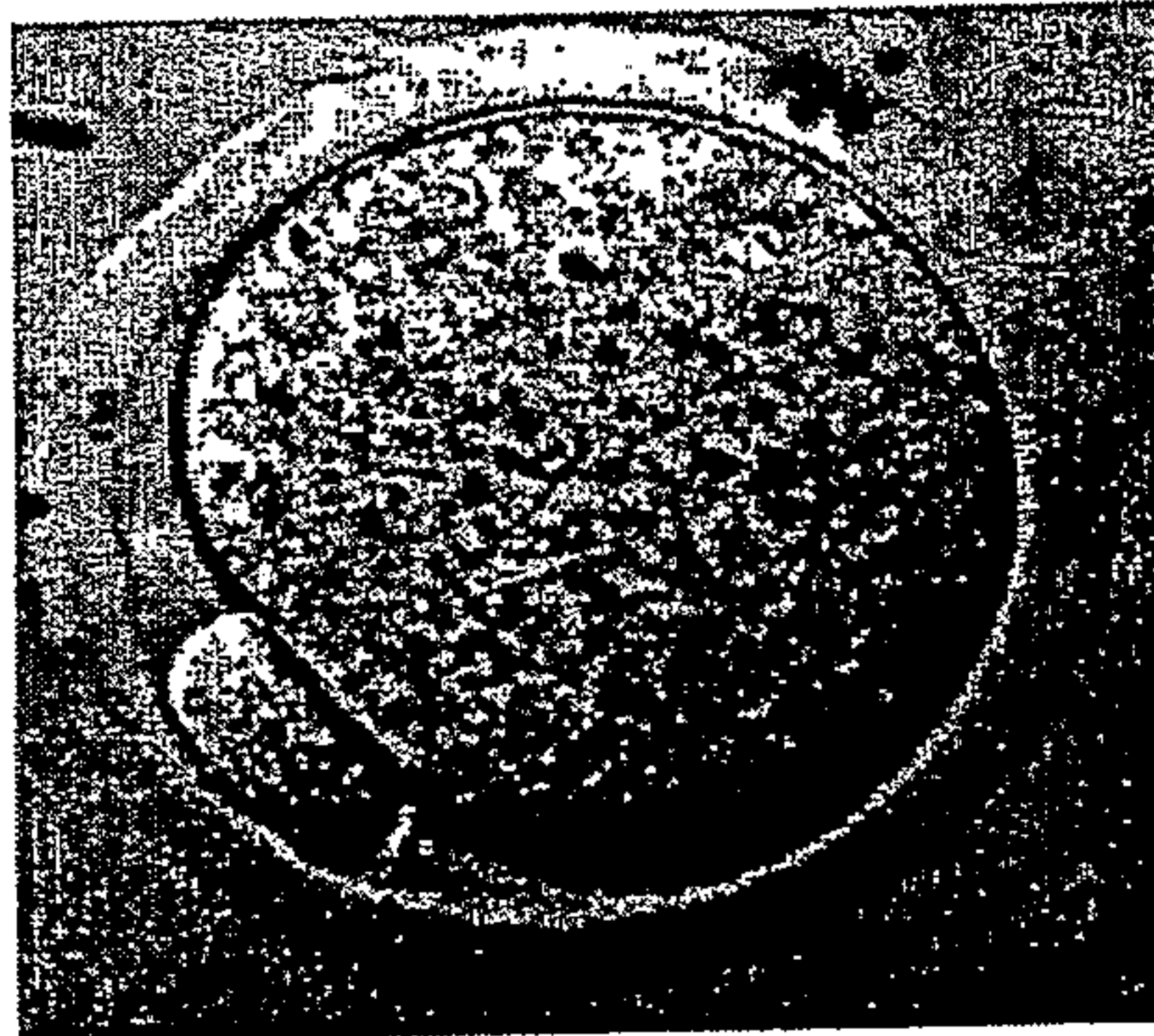


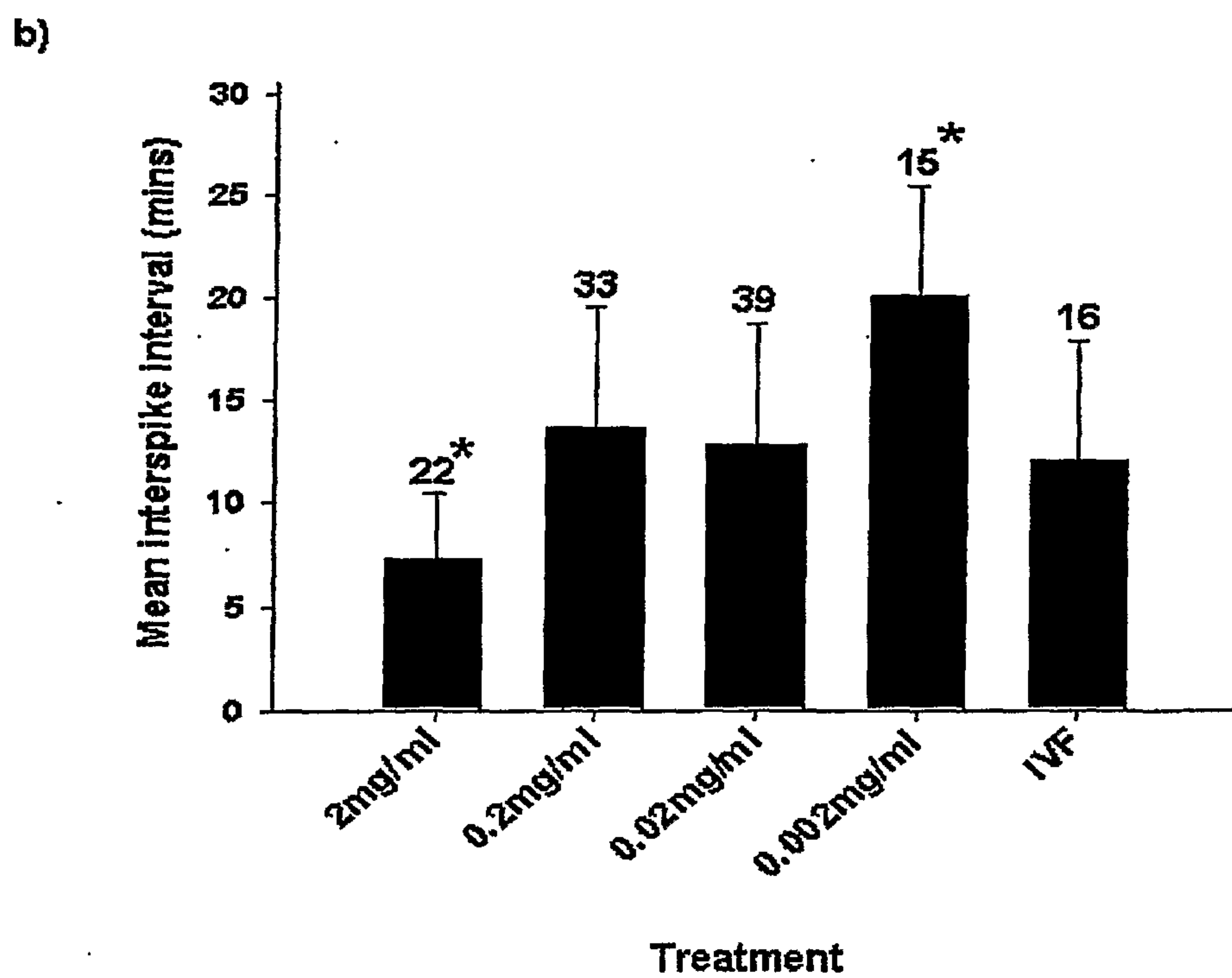
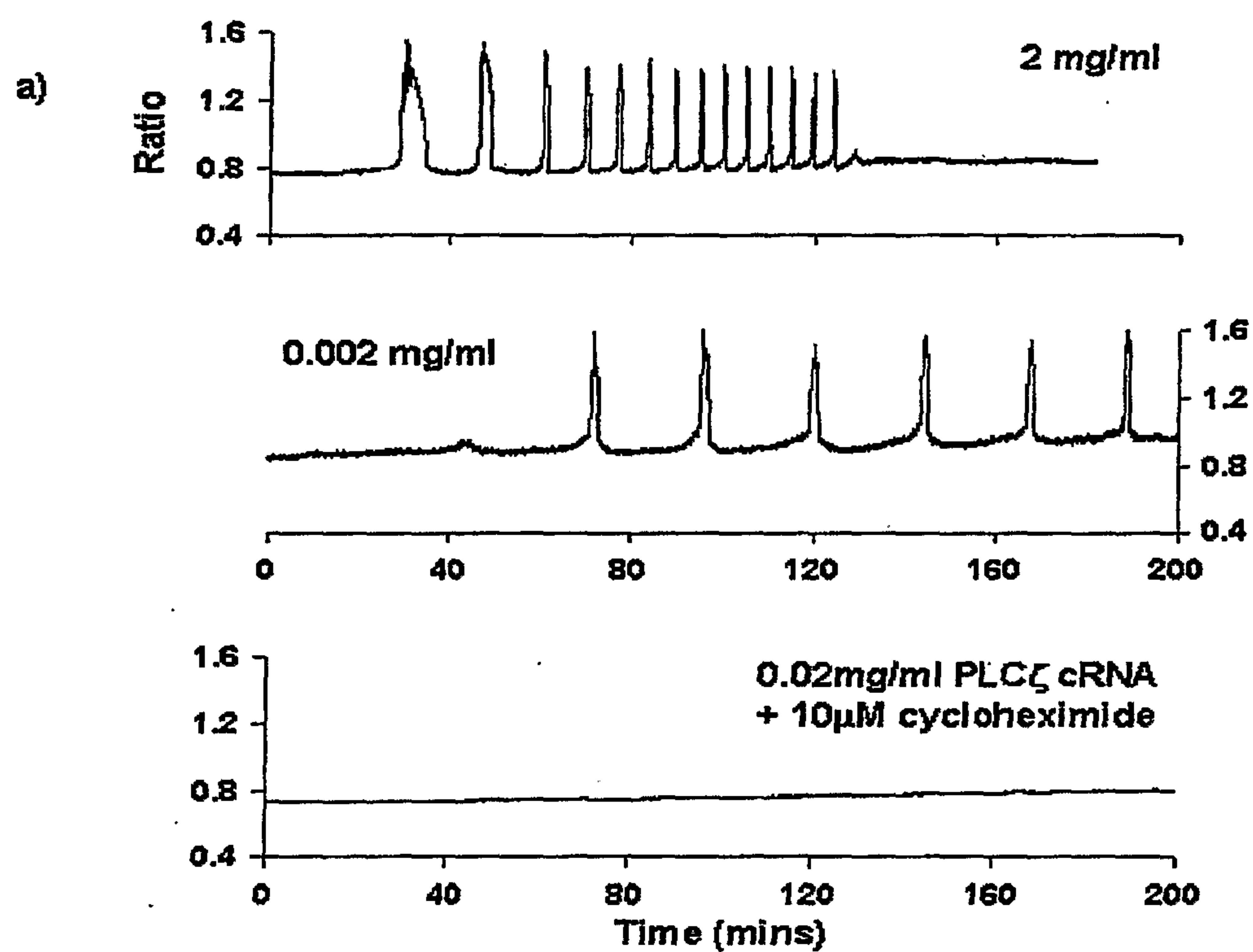
FIGURE 6

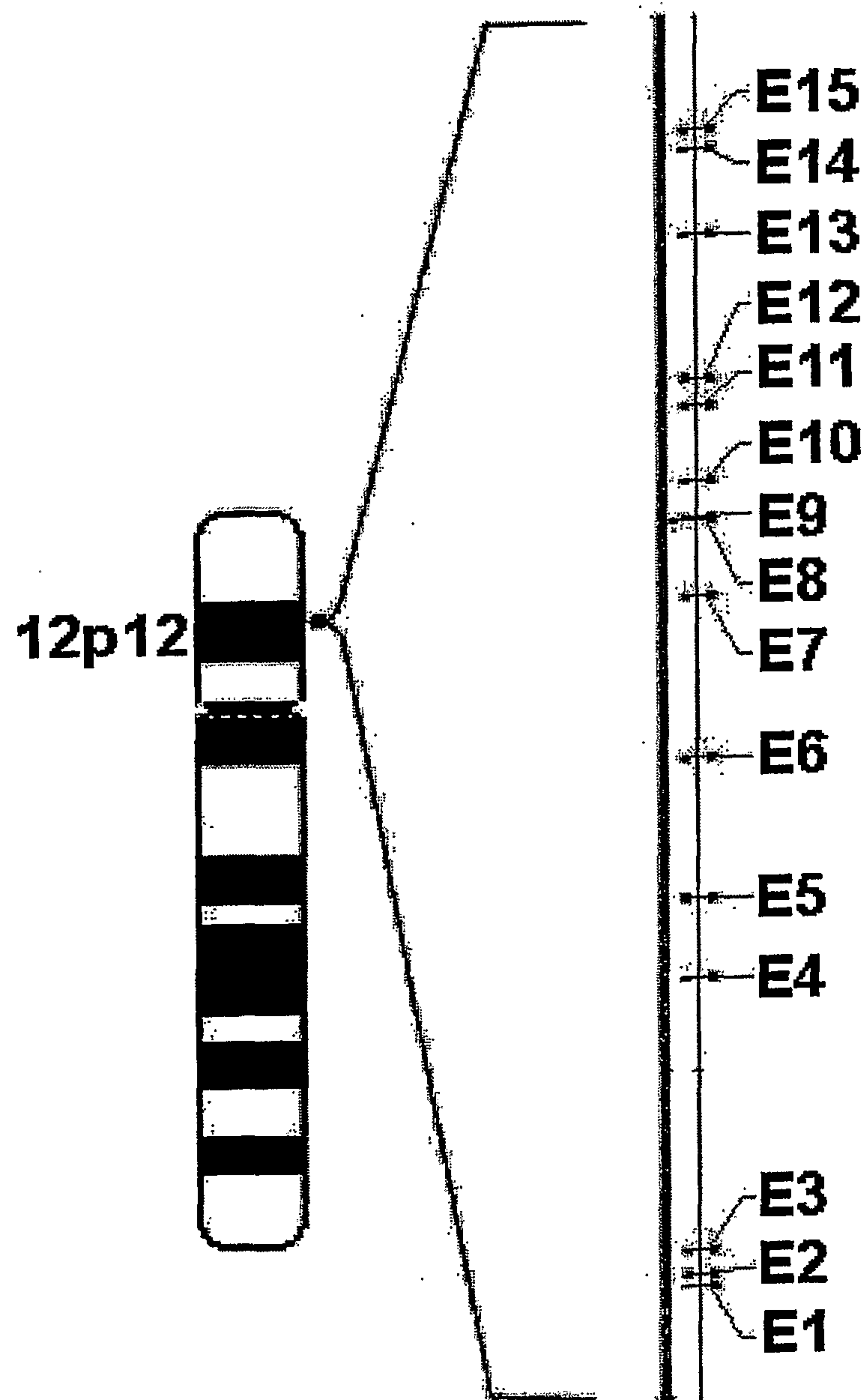
FIGURE 7

FIGURE 8a

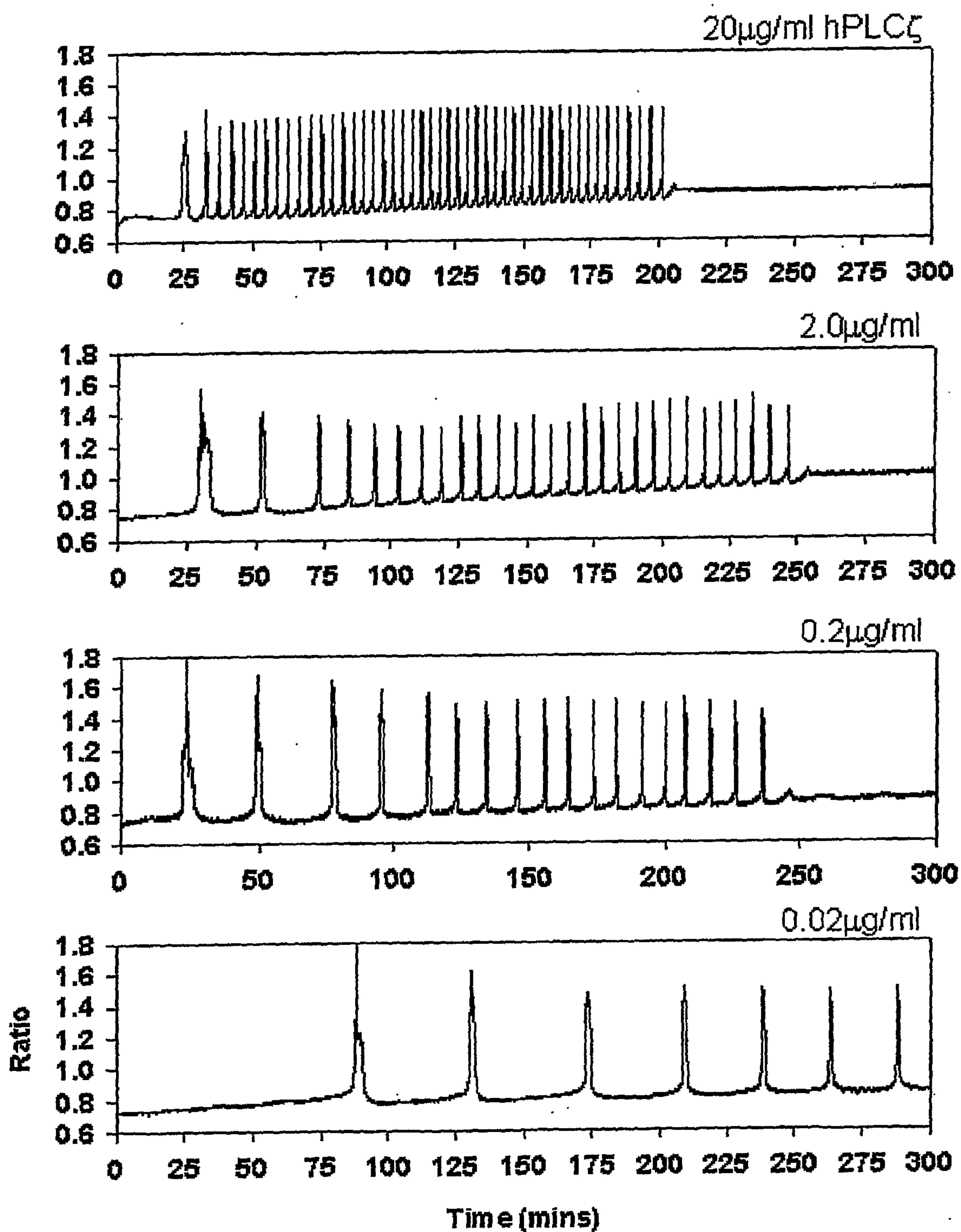


FIGURE 8b

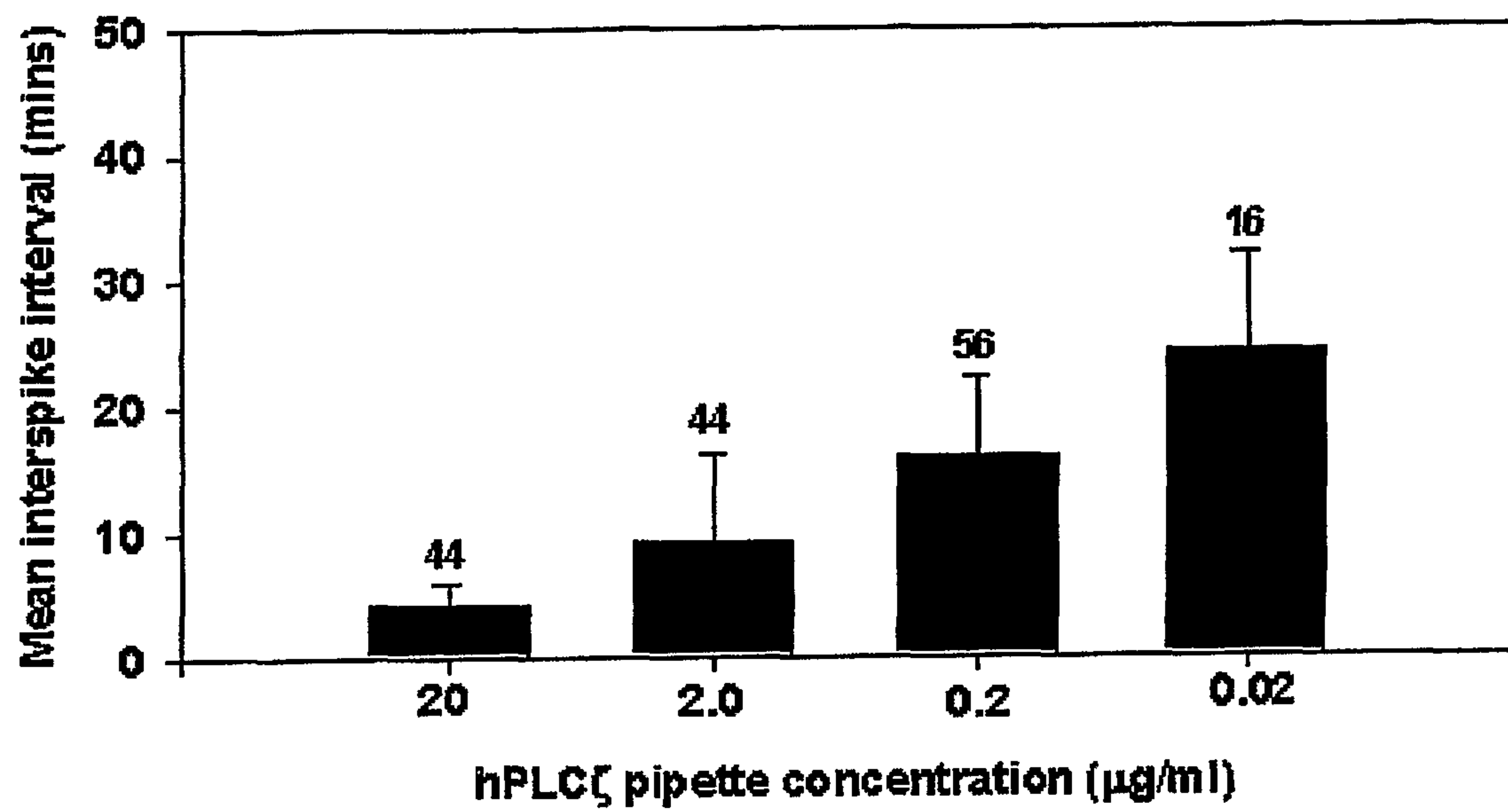


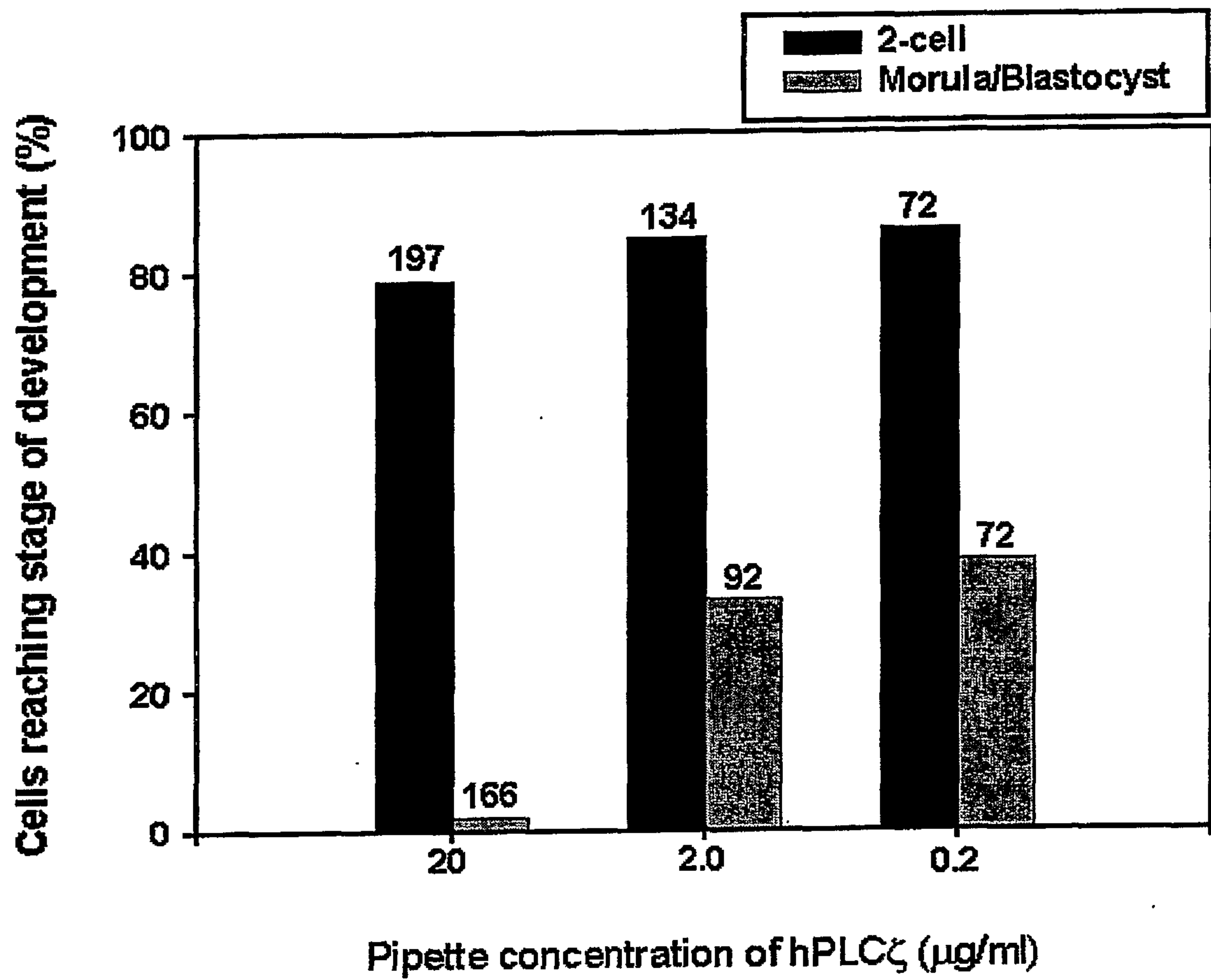
FIGURE 9a

FIGURE 9b

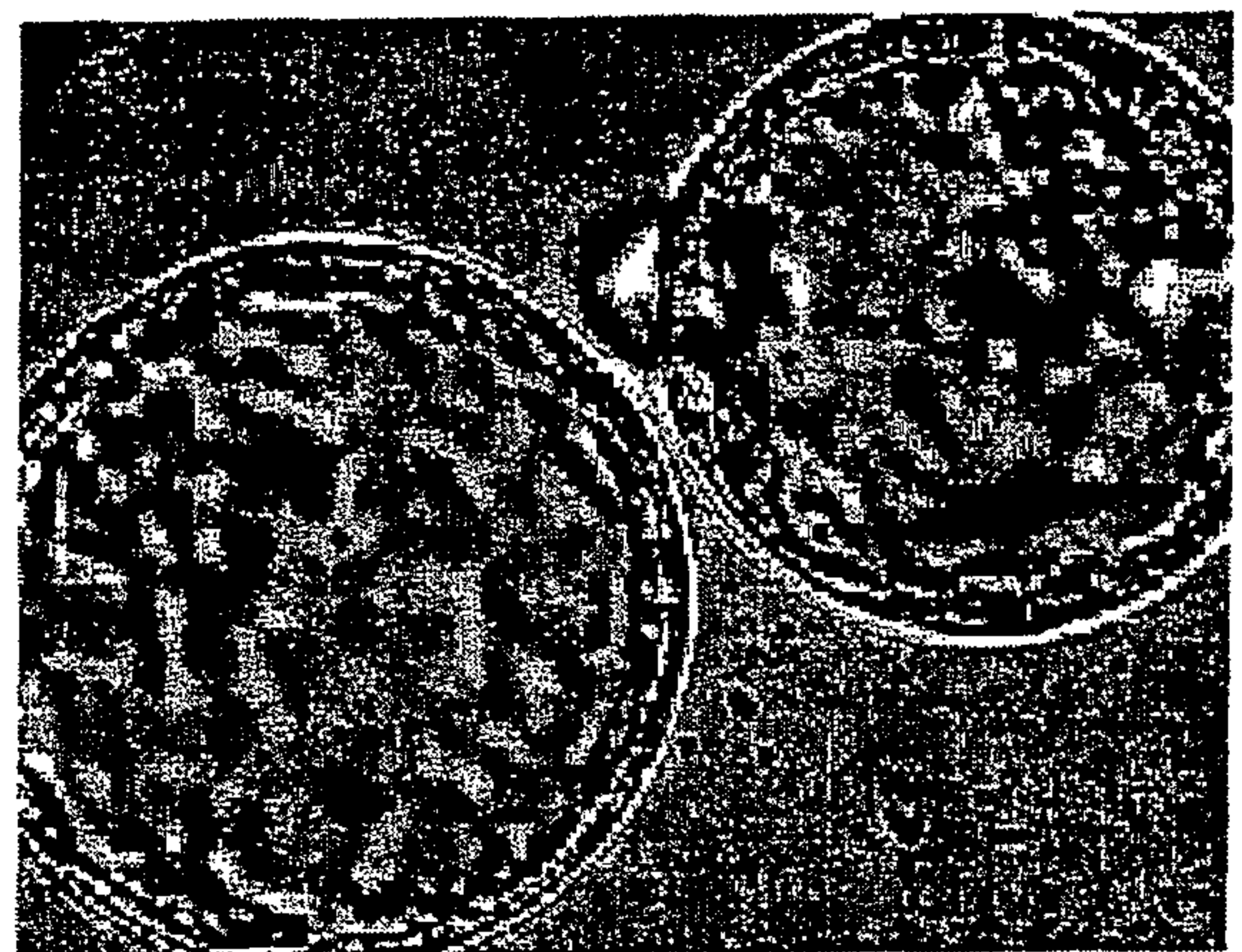
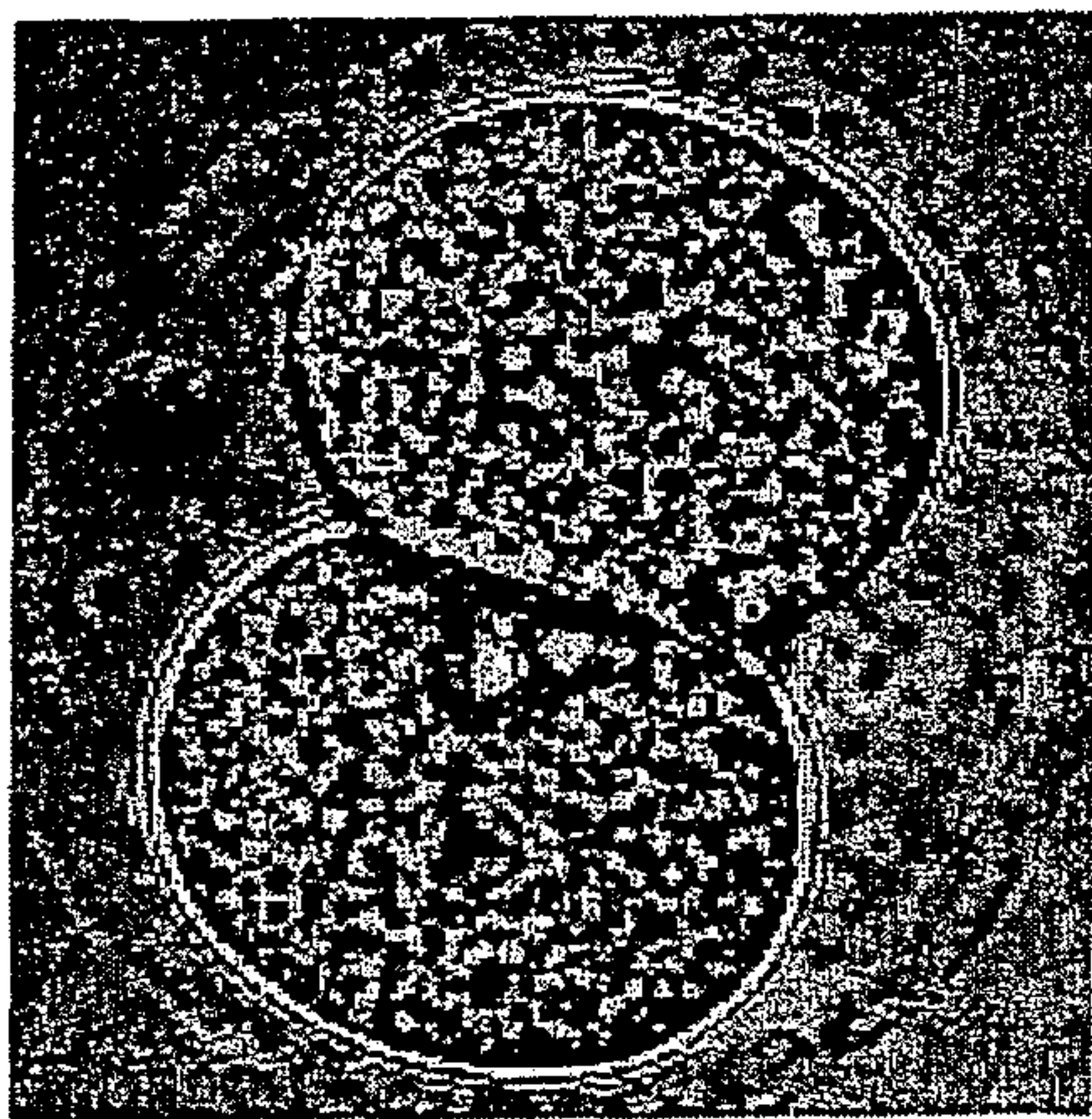


FIGURE 10a

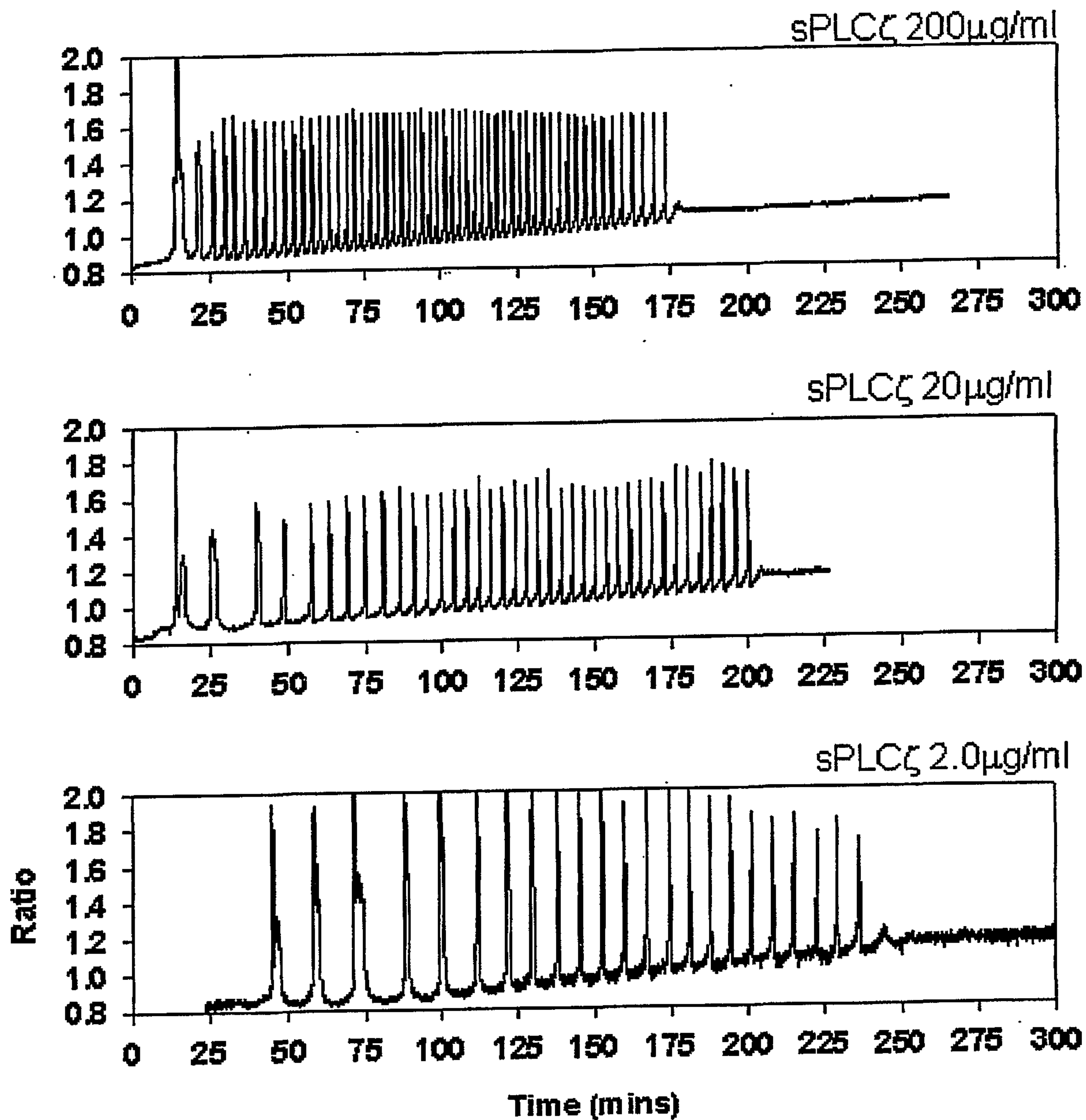


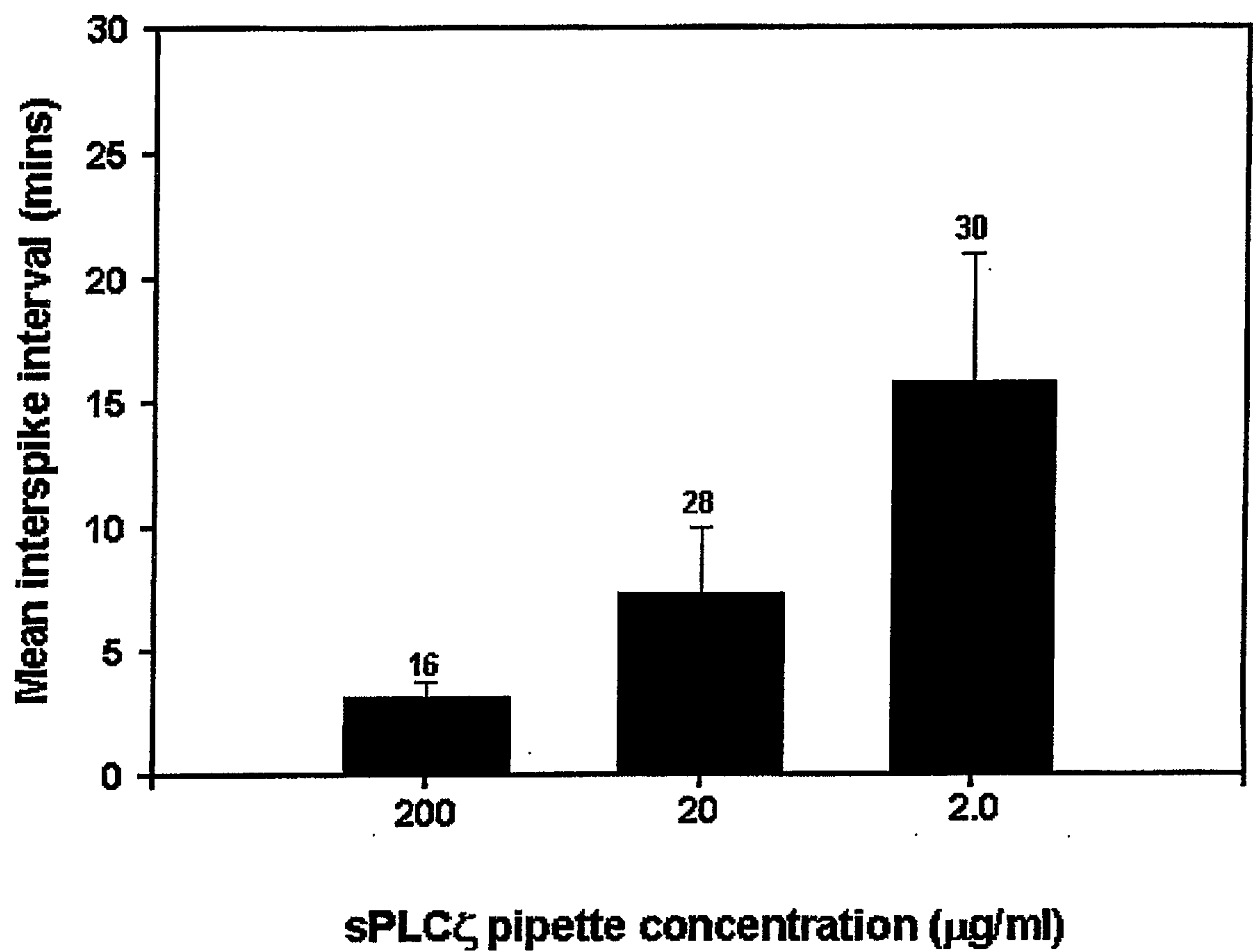
FIGURE 10b

FIGURE 11