

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
9 February 2006 (09.02.2006)

PCT

(10) International Publication Number  
**WO 2006/013110 A1**

(51) International Patent Classification:  
*C12Q 1/68* (2006.01)

(21) International Application Number:  
PCT/EP2005/008511

(22) International Filing Date: 5 August 2005 (05.08.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
10 2004 038 359.6 6 August 2004 (06.08.2004) DE  
60/662,798 18 March 2005 (18.03.2005) US

(71) Applicant and

(72) Inventor: **RIGLER, Rudolf** [AT/CH]; 115, rue du Centre,  
CH-1025 St-Sulpice (CH).

(74) Agent: **WEICKMANN & WEICKMANN**; Postfach 860  
820, 81635 München (DE).

(81) Designated States (*unless otherwise indicated, for every  
kind of national protection available*): AE, AG, AL, AM,  
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,

CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,  
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,  
KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA,  
MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ,  
OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL,  
SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC,  
VN, YU, ZA, ZM, ZW.

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

**Declaration under Rule 4.17:**

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

**Published:**

— *with international search report*

*For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*

(54) Title: PARALLEL HIGH THROUGHPUT SINGLE MOLECULE SEQUENCING PROCESS

(57) Abstract: The invention relates to a process and to an apparatus for parallel high throughput sequencing of nucleic acid molecules, in particular in the single molecule format. The present invention further relates to a process and an apparatus for a parallel high throughput study of enzyme-substrate interactions, in particular in the single molecule format.



WO 2006/013110 A1

- 1 -

## Parallel high throughput single molecule sequencing process

### Description

5 The invention relates to a process and to an apparatus for parallel high throughput sequencing of nucleic acid molecules, in particular in the single molecule format. The present invention further relates to a process and an apparatus for a parallel high throughput study of enzyme-substrate interactions, in particular in the single molecule format.

10

Sequencing of the human genome which consists of approx.  $3 \times 10^9$  bases, or of the genome of other organisms and the determination and comparison of individual sequence variants requires the provision of sequencing methods which firstly are fast and secondly can be employed routinely and cost-effectively. Although large efforts have been made in order to accelerate familiar sequencing methods, for example the enzymic chain termination method according to Sanger et al. (Proc. Natl. Acad. Sci. USA 74 (1977) 5463), in particular by automation (Adams et al., Automated DNA Sequencing and Analysis (1994), New York, Academic Press), currently only up to 2000 bases per day can be determined using an automated sequencer.

15

20

25

30

Over recent years, new approaches to overcome the limitations of conventional sequencing methods have been developed, inter alia sequencing by scanning tunneling microscopy (Lindsay and Phillip, Gen. Anal. Tech Appl. 8 (1991), 8-13), by highly parallel capillary electrophoresis (Huang et al., Anal. Chem. 64 (1992), 2149-2154; Kambara and Takahashi, Nature 361 (1993), 565-566), by oligonucleotide hybridization (Drmanac et al., Genomics 4 (1989), 114-128; Khrapko et al., FEBS Let. 256 (1989), 118-122; Maskos and Southern, Nucleic Acids Res. 20 (1992), 1675-1678 and 1679-1684) and by matrix-assisted laser desorption/ ionization mass spectroscopy (Hillenkamp et al., Anal. Chem. 63 (1991), 1193A-1203A).

- 2 -

Another approach is single-molecule sequencing (Dörre et al., Bioimaging 5 (1997), 139-152), in which the sequence of nucleic acids is carried out by progressive enzymic degradation of fluorescently labeled single-stranded DNA molecules and by detection of the sequentially released monomeric molecules in a microstructure channel. This process has the advantage of only a single molecule of the target nucleic acid being sufficient for carrying out a sequence determination.

Although application of the abovementioned methods has already resulted in considerable progress, there is a great need for further improvements. The object on which the present invention is based was therefore to provide a process for sequencing nucleic acids, which represents a further improvement compared to the prior art and which makes possible parallel determination of individual nucleic acid molecules in a multiplex format.

PCT/EP01/07462 proposes a multiplex sequencing process which comprises providing, in an immobilized form, nucleic acid molecules carrying a plurality of fluorescent labeling groups on a support and determining simultaneously the base sequence of a plurality of nucleic acid molecules on the basis of the time-dependent change, caused when nucleotide building blocks are cleaved off, in the fluorescence of said nucleic acid molecules or/and said cleaved-off nucleotide building blocks. According to WO 2003/052137, the sequence is determined by irradiating light into the support and generating an evanescent excitation field by way of internal reflection on the support surface in the region of the immobilized nucleic acid molecules.

The present application relates to a process for sequencing nucleic acids, comprising the following steps:

- (a) providing a support having a multiplicity of nucleic acid-degrading enzyme molecules immobilized thereon,
- (b) contacting said support with free nucleic acid molecules carrying a plurality of fluorescent labeling groups,

- 3 -

(c) progressively cleaving off individual nucleotide building blocks from said free nucleic acid molecules by said immobilized enzyme molecules, and

5 (d) simultaneously determining the base sequence of a plurality of nucleic acid molecules on the basis of the time-dependent change, caused when nucleotide building blocks are cleaved off, in the fluorescence of said nucleic acid molecules or/and of said cleaved-off nucleotide building blocks.

10 The process of the invention is a support-based multiplex sequencing method which enables a multiplicity of individual free nucleic acid molecules to be studied in parallel. This is achieved by providing a support having a plurality of nucleic acid-degrading enzyme molecules immobilized thereon and determining in parallel the fluorescence of a plurality of degradation  
15 reactions. The process is preferably carried out in the form of a parallel high throughput single molecule analysis.

The support used for said process may be any planar or structured support which is suitable for immobilizing enzyme molecules. Examples of suitable  
20 support materials are glass, quartz, plastic, metals, semimetals such as silicon, for example, metal oxides such as silicon dioxide, for example, or composites comprising said materials. Said support may, at least in the region of the immobilized enzyme molecules, have sufficient optical transparency and suitable surface properties for irradiation with fluorescence  
25 excitation light or/and backscatter of fluorescence emission light through the support or for evanescent-based fluorescence detection. In principle, the support may also have any design, as long as a reaction space can be formed which enables individual nucleotide building blocks to be cleaved off progressively from the nucleic acids contacted with said support in a liquid  
30 reaction mixture.

The enzyme molecules may bind to this support via covalent or noncovalent interactions. For example, high-affinity interactions between the partners of a

- 4 -

specific binding pair, for example biotin/streptavidin or avidin, haptene/anti-haptene antibody, sugar/lectin, etc., can mediate binding of the polypeptides to the support. Thus it is possible to couple biotinylated enzyme molecules to streptavidin-coated supports. Alternatively, the enzyme molecules may also  
5 be bound to the support via adsorption. Thus enzyme molecules modified by incorporation of alkane thiol groups, may bind to metallic supports, e.g. supports made of gold. Yet another alternative is covalent immobilization in which it is possible to mediate enzyme molecule binding via reactive silane groups on a silica surface.

10 A plurality of enzyme molecules is bound to a single support. The enzyme molecules immobilized on the support and the sample liquid in contact therewith, which contains the nucleic acid molecules to be degraded, define one or more reaction spaces. Preferably at least 100, particularly preferably  
15 at least 1000, and particularly preferably at least 10 000, and up to more than  $10^6$ , enzyme molecules are bound to said support. Preference is given to said enzyme molecules binding to said support in such a way that a diluted enzyme layer is produced on said support, preferably containing from 0.01 to 2, preferably 0.1 to 1, enzyme molecules per  $\mu\text{m}^2$  of support surface  
20 area. The enzyme molecules may be applied randomly, for example by contacting a diluted solution of biotinylated enzyme molecules with a support two-dimensionally coated with streptavidin. Alternatively, the enzyme molecules may also be bound to specific regions of the support surface, for example by contacting a diluted solution of biotinylated enzyme molecules  
25 with a structured support, only particular regions of which are coated with streptavidin. The nucleic acid molecules to be sequenced may be in single-stranded or in double-stranded form. They are preferably from 50 to 2000 nucleotides, particularly preferably from 200 to 1000 nucleotides, in length. The nucleic acid molecules to be sequenced, for example DNA molecules or  
30 RNA molecules, contain a plurality of fluorescent labeling groups, with preferably at least 50%, particularly preferably at least 70%, and most preferably essentially all, for example at least 90%, of the nucleotide building blocks of one or more, for example two, three or four, base types carrying a

- 5 -

fluorescent labeling group, each base type favorably carrying a different fluorescent labeling group. It is not necessary to label completely all nucleotide building blocks of a single base type, since it is possible, when determining the sequence of a single nucleic acid molecule, to fill in possible gaps by multiple parallel determinations.

Nucleic acids labeled in this way may be generated by enzymic primer extension on a nucleic acid template by using a suitable polymerase, for example a DNA polymerase such as, for example, Taq polymerase, a thermostable DNA polymerase from *Thermococcus gorgonarius* or other thermostable organisms (Hopfner et al., PNAS USA 96 (1999), 3600-3605), or a mutated Taq polymerase (Patel and Loeb, PNAS USA 97 (2000), 5095-5100), by using fluorescently labeled nucleotide building blocks. Preference is given to polymerases without exonuclease activity, such as Vent exo<sup>-</sup> or Tgo exo<sup>-</sup>, for example. Particularly preferred methods for incorporating fluorescent labeling groups are described in Tasara et al. (Nucleic Acids Res. 31 (2003), 2636-2646) or Giller et al. (Nucleic Acids Res. 31 (2003), 2630-2635).

It is also possible to prepare the labeled nucleic acid molecules by amplification reactions, for example PCR. Thus an asymmetric PCR produces amplification products in which only one strand contains fluorescent labels. Such asymmetric amplification products can be sequenced in double-stranded form. Symmetric PCR produces nucleic acid fragments in which both strands are fluorescently labeled. These two fluorescently labeled strands can be separated and contacted separately in single-stranded form with the immobilized enzyme molecules so that it is possible to determine the sequence of one or both complementary strands separately. Alternatively, one of the two strands can be modified on the 3' end, for example by incorporating a PNA link, such that monomeric building blocks can no longer be removed by cleavage. In this case, double-strand sequencing is possible.

- 6 -

Preferably, essentially all nucleotide building blocks of at least two base types, for example two, three or four base types, carry a fluorescent label, each base type conveniently carrying a different fluorescent labeling group. If the nucleic acid molecules have not been labeled completely, it is  
5 nevertheless possible to determine the sequence completely by parallel sequencing of a plurality of molecules.

The nucleic acid template whose sequence is to be determined may be selected, for example, from DNA templates such as genomic DNA  
10 fragments, cDNA molecules, plasmids, etc., or else from RNA templates such as mRNA molecules.

The fluorescent labeling groups may be selected from known fluorescent labeling groups used for labeling biopolymers, for example nucleic acids, such as, for example, fluoresceins, rhodamines, oxazines, for example  
15 Evoblue or Gnothis Blue, pycoerythrin, Cy3, Cy5, IR dyes or derivatives thereof, etc.

The method of the invention is based on fluorescent labeling groups incorporated in nucleic acid strands interacting with neighboring groups, for  
20 example with chemical groups of the nucleic acids, in particular nucleobases such as, for example, G, or/and neighboring fluorescent labeling groups, and these interactions leading to a change in fluorescence, in particular in fluorescence intensity, compared to the fluorescent labeling groups in  
25 "isolated" form, owing to quenching processes or/and energy transfer processes. The removal by cleavage of individual nucleotide building blocks alters the overall fluorescence, for example the fluorescence intensity of an immobilized nucleic acid strand, and this change is a function of the removal by cleavage of individual nucleotide building blocks, i.e. a function of time.  
30 This time-dependent change in fluorescence may be recorded in parallel for a multiplicity of nucleic acid molecules and correlated with the base sequence of the individual nucleic acid strands. Preference is given to using those fluorescent labeling groups which, when incorporated in the nucleic

- 7 -

acid strand, are, at least partially, quenched so that the fluorescence intensity is increased after the nucleotide building block containing the labeling group or a neighboring building block causing quenching has been removed by cleavage.

5

The sequencing reaction of the process of the invention comprises progressively cleaving off individual nucleotide building blocks from the nucleic acid molecules to be sequenced by immobilized, nucleic acid-degrading enzyme molecules. Preference is given to using as enzyme molecules exonucleases, it being possible to use single-strand or double-strand exonucleases which degrade in the 5'→3' direction or 3'→5' direction. Exonucleases which are particularly preferably used are T7 DNA polymerase, E.coli exonuclease I and E.coli exonuclease III.

10

During the progressive removal by cleavage of individual nucleotide building blocks, it is possible to measure a change in fluorescence intensity of the nucleic acid strand or/and the cleaved-off nucleotide building block, owing to quenching processes or energy transfer processes. This change in fluorescence intensity with time depends on the base sequence of the nucleic acid strand studied and can therefore be correlated with the sequence. The complete sequence of a nucleic acid strand is usually determined by generating, preferably by enzymic primer extension as described above, and sequencing a plurality of nucleic acid strands, labeled on different bases, for example A, G, C and T, or combinations of two different bases. It is possible, where appropriate, to attach to the nucleic acid strand to be studied also a "sequence identifier", i.e. a labeled nucleic acid of known sequence, for example by enzymic reaction using ligase or/and terminal transferase, so that at the start of sequencing initially a known fluorescence pattern and only thereafter the fluorescence pattern corresponding to the unknown sequence to be studied is obtained.

20

25

30

In order to accelerate the removal of cleaved nucleotide building blocks from the immobilized nucleotide strands, preferably a convectional flow away from



- 8 -

the support is generated in the reaction space. The flow rate may be in the range from 1 to 10 mm/s.

5 The detection comprises irradiating light into the support, preferably by means of a laser, in order to cause excitation of the fluorescent labeling groups. It is possible, in this connection, to use one or more laser beams, for example an expanded laser beam, having a cross section of approx. 1-20 mm, or/and multiple laser beams. The detection preferably comprises a multipoint fluorescence excitation by lasers, for example a dot matrix of laser  
10 dots generated via diffraction optics or a quantum well laser.

Alternatively, fluorescence may also be detected by irradiating light into the support and generating an evanescent excitation field by way of internal reflection at the support surface in the region of the immobilized enzyme  
15 molecules. Internal reflection at one or more positions of the support surface in the region of immobilized nucleic acid molecules generates an evanescent excitation field which causes excitation of the fluorescent labeling groups of the nucleic acid molecules to be sequenced. Said reflection at the support surface is preferably total internal reflection.

20 Fluorescence emission of a plurality of nucleic acid strands may be detected in parallel using a detector matrix which comprises, for example, an electronic detector matrix, for example a CCD camera, a CMOS detector matrix, e.g. a CMOS camera, or an avalanche photodiode matrix. The  
25 detection may be carried out in such a way that fluorescence excitation and detection are carried out in parallel on all nucleic acid strands studied. A possible alternative to this is to study in several steps in each case a portion of the nucleic acid strands. Preference is given to carrying out the detection on fluorescence light which is emitted essentially orthogonally from the  
30 support surface through the reaction space or through the support body.

The detection may be carried out, for example, by means of confocal single molecule detection, for example by fluorescence correlation spectroscopy,

- 9 -

which involves exposing a very small, preferably confocal, volume element, for example from  $0.1 \times 10^{-15}$  to  $20 \times 10^{-12}$  l, to the excitation light of a laser, which light excites the receptors present in this measuring volume so that the latter emit fluorescence light, the fluorescence light emitted from said measuring volume being measured by means of a photodetector and the change in the measured emission with time being correlated with the concentration of the analyte, so that it is possible to identify, at an appropriately high dilution, individual molecules in said measuring volume. Details of the procedure and of the apparatus used for detection can be found in the disclosure of the European patent 0 679 251. The confocal determination of single molecules is furthermore described in Rigler and Mets (Soc. Photo-Opt. Instrum. Eng. 1921 (1993), 239 ff.) and Mets and Rigler (J. Fluoresc. 4 (1994) 259-264).

Alternatively or additionally, detection may also be carried out by way of time-resolved decay measurement, called "time gating", as described, for example, by Rigler et al., "Picosecond Single Photon Fluorescence Spectroscopy of Nucleic Acids", in: "Ultrafast Phenomenes", D.H. Auston, Ed., Springer 1984. Here, the fluorescent molecules are excited in a measuring volume followed by, preferably at a time interval of  $\geq 100$  ps, opening a detection interval on the photodetector. In this way it is possible to keep background signals generated by Raman effects sufficiently low so as to enable single molecules to be detected in an essentially interference-free manner.

As indicated above, the process of the present invention requires a contact of immobilized nucleic acid-degrading enzyme molecules with free nucleic acid molecules carrying plurality of fluorescent labeling groups. In a preferred embodiment of the invention, a support may be provided having a multiplicity of nucleic acid-degrading enzyme molecules and a multiplicity of nucleic acid-polymerizing enzyme molecules co-immobilized thereon. In this embodiment, the free nucleic acid molecules carrying a plurality of fluorescent labeling groups may be directly synthesized on the support, e.g.

- 10 -

by providing a nucleic acid template molecule, e.g. a single stranded DNA or RNA molecule, to which a primer molecule is bound which can be extended by incorporating labeled nucleotide building blocks *in situ* on the support. By carrying out the nucleic acid polymerization under conditions of limited processivity, e.g. by choosing enzymes having limited processivity or adjusting reaction conditions leading to limited processivity, the nucleic acid template molecule and the extended primer are liberated from the immobilized polymerase molecule and will contact a co-immobilized nucleic acid-degrading enzyme, e.g. a 3'→5'-exonuclease which cleaves off individual nucleotide building blocks leading to a time dependent change in the fluorescence indicated above.

In a further aspect, the invention relates to a process for studying enzyme-substrate interactions, comprising the following steps:

- (a) providing a support having a multiplicity of enzyme molecules immobilized thereon,
- (b) contacting said support with free substrate molecules carrying fluorescent labeling groups,
- (c) reacting said substrate molecules with said immobilized enzyme molecules, and
- (d) studying simultaneously the enzyme-substrate interaction of a plurality of individual enzyme or/and substrate molecules on the basis of the time-dependent change, caused during the enzymic reaction of substrate molecules, in the fluorescence of said substrate molecules or/and cleaved-off substrate parts or/and reaction products formed.

The process of the invention is a support-based multiplex examination method which enables a multiplicity of individual enzyme molecules or/and free substrate molecules to be studied in parallel. This is achieved by providing a support, as described above, having a plurality of enzyme molecules immobilized thereon and by determining the fluorescence of a plurality of enzyme molecules in parallel. The process is preferably carried out in the form of a parallel high throughput single molecule analysis.

- 11 -

The support used for said process corresponds to the support described in connection with the previously described sequencing method. In principle, the support may have any design, as long as it is possible to form a reaction space which enables the substrate molecules contacted with said support to be progressively reacted in a liquid reaction mixture.

The enzyme molecules may bind to the support via covalent or noncovalent interactions, as illustrated above.

In principle, the process of the invention can employ any enzymes. Examples of suitable enzymes comprise polymerases, dehydrogenases, oxidoreductases, etc. Preference is given to immobilizing on a support identical enzyme molecules, but it is also possible to immobilize different enzyme molecules, for example variants of mutagenized enzyme libraries.

As previously described, a plurality of enzyme molecules is bound to a single support. The enzyme molecules immobilized on said support and the sample fluid in contact therewith, which contains the substrate molecules, define one or more reaction spaces. The substrate molecules to be studied contain one or more fluorescent labeling groups. These mean substrates which are fluorescent themselves, substrates which result in a fluorescent product after reaction with an enzyme, or/and substrates containing a fluorescent labeling group which is not involved in the enzymic reaction.

Substrates of this kind are known from the prior art. It is possible to study one or more different substrates on a single support.

The process of the invention is based on substrate molecules into which fluorescent labeling groups have been incorporated reacting with the enzyme molecules, which reactions result in a change in fluorescence, in particular in fluorescence intensity or/and in the fluorescence wavelength, compared to the fluorescence of fluorescent labeling groups in the substrate

- 12 -

as starting material. In the course of said enzyme-substrate interaction, the overall fluorescence, for example the fluorescence intensity of a substrate molecule, changes as a function of the progressing interaction with the enzyme, i.e. as a function of time. This fluorescence change with time may  
5 be recorded in parallel for a multiplicity of enzyme or/and substrate molecules and correlated to the enzyme-substrate interaction.

Examples of enzyme-substrate interactions which may be studied using the process of the invention are hydrolysis, polymerization, transfer of functional  
10 groups, oxidation, reduction, dehydrogenation and isomerization.

The process of the present invention makes possible the direct resolution of conformation dynamics, binding events or cleaving reactions of biological molecules, which may result, for example, in elucidating protein functions.  
15 Information at the molecular level about the system to be studied may be obtained via fluorescence signal fluctuation.

The detection is carried out as described previously in connection with the sequencing. The detection is preferably carried out by means of confocal  
20 single molecule detection.

The invention still further relates to a process for sequencing nucleic acids, comprising the following steps:

- 25 (a) providing a support having a multiplicity of nucleic acid-polymerising enzyme molecules immobilized thereon,  
(b) contacting said support with (i) free nucleic acid molecules to which a primer molecule is bound and (ii) nucleotide building blocks carrying a fluorescent labeling group,  
30 (c) incorporating said nucleotide building blocks into said primer by said immobilized enzyme molecules, and  
(d) simultaneously determining the base sequence of a plurality of nucleic acid molecules on the basis of the time-dependent change,

- 13 -

caused when nucleotide building blocks are incorporated, in the fluorescence of said extended primer molecule or/and of said nucleotide building blocks.

5 The support used for this process corresponds to the supports described in connection with the previously described embodiments. The binding of the enzyme molecules to the support may also be carried out as described above.

10 The nucleic acid-polymerising enzyme is preferably selected from DNA polymerases and RNA polymerases, particularly from polymerases capable of a template specific nucleic acid polymerization.

The enzyme molecules immobilized on the support are contacted with free  
15 nucleic acid molecules, e.g. single-stranded DNA or RNA molecules, to which a primer molecule is bound. The primer molecule is preferably a nucleic acid or nucleic acid analogue molecule having a free 3'-end which can be extended by an enzymatic reaction. Further, the presence of nucleotide building blocks carrying a fluorescent labeling group is required.  
20 Preferably, each different nucleotide building block (A, G, C, T/U) contains a different fluorescent labeling group. When a nucleotide building block is incorporated into the primer molecule, a time dependent change in the fluorescence occurs which can be detected as indicated above. Preferably, the incorporation of the nucleotide building blocks into the primer is  
25 associated with a detectable increase in the fluorescence, preferably with a transient increase in the fluorescence. For example, nucleotide building blocks may be used which carry a fluorescent labeling group on the portion of the molecule which is cleaved off when the building block is incorporated into the primer, e.g. on the  $\gamma$ -phosphate group, and a fluorescence quencher  
30 group on the portion of the molecule, which is incorporated into the primer, e.g. on the  $\alpha$ -phosphate group, on the sugar and/or on the nucleobase group.

- 14 -

The invention still further relates to an apparatus for sequencing nucleic acids, comprising

- 5 (a) a support comprising a multiplicity of enzyme molecules immobilized thereon,
- (b) means for (i) delivering free nucleic acid molecules carrying a plurality of fluorescent labeling groups to said support, or (ii) generating free nucleic acid molecules carrying a plurality of  
10 fluorescent labeling groups on said support,
- (c) a reaction space for progressively cleaving off individual nucleotide building blocks from said nucleic acid molecules, and
- (d) means for determining simultaneously the base sequence of a plurality of nucleic acid molecules on the basis of the time-  
15 dependent change, caused when nucleotide building blocks are cleaved off, in the fluorescence of said nucleic acid molecules or/and said cleaved-off nucleotide building blocks.

The invention further relates to an apparatus for studying enzyme-substrate  
20 interactions, comprising

- (a) a support comprising a multiplicity of enzyme molecules immobilized thereon.
- (b) means for (i) delivering free substrate molecules carrying  
25 fluorescent labeling groups to said support, or (ii) generating free substrate molecules carrying fluorescent labeling groups on said support,
- (c) a reaction space for reacting said substrate molecules with said immobilized enzyme molecules, and
- 30 (d) means for studying simultaneously a plurality of enzyme-substrate interactions of a plurality of enzyme or/and substrate molecules on the basis of the time-dependent change, caused during the enzymic reaction, in the fluorescence of said substrate molecules or/and the

- 15 -

cleaved-off substrate parts or/and reaction products formed.

The invention further relates to an apparatus for sequencing nucleic acids comprising:

- (a) a support comprising a multiplicity of nucleic acid-polymerising enzyme molecules immobilized thereon,
- (b) means for providing (i) free nucleic acid molecules to which a primer molecule is bound and (ii) nucleotide building blocks carrying a fluorescent labeling group on the support,
- (c) a reaction space for incorporating said nucleotide building blocks into said primer, and
- (d) means for determining simultaneously the base sequence of a plurality of nucleic acid molecules on the basis of the time-dependent change, caused when nucleotide building blocks are incorporated, in the fluorescence of said extended primer molecule or/and of said nucleotide building blocks.

The apparatus is preferably a micro- or nanoscale structure suitable for determining single molecules, for example an at least partially transparent structure having channels or/and depressions. A preferred nano structure is described in PCT/EP02/02582.

The process of the invention and the apparatus of the invention may be employed, for example, in the analysis of genomes and transcriptomes or in differential analyses, for example studies regarding the difference in the genome or transcriptome of individual species or organisms within a species.

Furthermore, the figures below are intended to illustrate the present invention.

**Figure 1** is a diagrammatic representation of a support (2) of the invention



- 16 -

having a multiplicity of enzyme molecules (4) immobilized thereto. A support having a surface area of from 1 to 2 cm<sup>2</sup> may contain up to 10<sup>6</sup> enzyme molecules, for example.

5 **Figure 2** depicts how the enzyme molecules (4) immobilized to the support (2) progressively cleave off individual nucleotide building blocks (10) from a free nucleic acid (6) having a plurality of fluorescent labeling groups (8). While the nucleotide building blocks incorporated into the nucleic acid strand exhibit only low fluorescence or no fluorescence at all, owing to quenching  
10 processes, said fluorescence increases after the removal by cleavage. Said increased fluorescence may be detected by means of a detector.

**Figure 3A** depicts a first embodiment of the invention. Enzyme molecules (4) which degrade nucleic acid molecules (6) in the sample fluid are  
15 immobilized on the support (2). A laser (12) irradiates excitation light (14) through the sample fluid. Said excitation light (14) may be directed, for example by a diffraction-optical element (16), confocally to the positions of the individual immobilized enzyme molecules. The fluorescence emission light (18) emitted by the sample fluid is recorded by a detector matrix (20),  
20 for example a CCD or a CMOS camera.

The embodiment depicted in **Figure 3B** involves the enzyme molecules (4) being irradiated with excitation light (14), originating from the laser (12), through an optically transparent support (2). The emission light (18) which is  
25 likewise emitted through the support (2) is recorded by the detector matrix (20).

**Figure 4** depicts a further embodiment of the invention, wherein excitation light (14) of an expanded laser is irradiated into the optically transparent  
30 support (2) having enzyme molecules (4) immobilized thereto and nucleic acid molecules (6) to be sequenced and leaves the support (2) again after refraction at the support surface in the region of the immobilized enzyme molecules (4). In this manner, an evanescent excitation field is generated by

- 17 -

which the fluorescently labeled nucleic acid molecules or nucleotide building blocks are excited and fluoresce. The emission light (18) is directed through an optical element (22) to a detector (20).

5 The embodiment depicted in **Figure 5** involves multiple refractions (24a, 24b, 24c) generating evanescent excitation fields in the optically transparent support (2). The evanescent excitation fields may be strip- or point-shaped, for example.

10 Alternatively, it is also possible to irradiate a plurality of laser light foci upon or into the support by using a diffractive optical element, as disclosed, for example, in DE 101 26 083.0.

As shown in **Figure 6**, the enzyme molecules immobilized on the support (2)  
15 may be arranged in a random or disordered manner on the support surface (**Figure 6A**) or may be present in an ordered manner or on predefined structural areas (**Figure 6B**).

**Figure 7** depicts a further embodiment of the invention. This embodiment  
20 involves the use of a nucleic acid-polymerizing enzyme molecule (30) immobilized on a solid support (32). The immobilized enzyme (30) is contacted with a single-stranded nucleic acid template molecule (34) to which a primer molecule (36) having a 3'-end is bound. Further, nucleotide building blocks (A, G, T, C) each carrying a different fluorescence labeling  
25 group are present. Primer extension by incorporating the nucleotide building blocks generates a time-dependant change in the fluorescence which can be detected.

### Claims

1. A process for sequencing nucleic acids, comprising the following steps:
  - 5 (a) providing a support having a multiplicity of nucleic acid-degrading enzyme molecules immobilized thereon,
  - (b) contacting said support with free nucleic acid molecules carrying a plurality of fluorescent labeling groups,
  - (c) progressively cleaving off individual nucleotide building blocks from  
10 said free nucleic acid molecules by said immobilized enzyme molecules, and
  - (d) simultaneously determining the base sequence of a plurality of nucleic acid molecules on the basis of the time-dependent change, caused when nucleotide building blocks are cleaved off, in the  
15 fluorescence of said nucleic acid molecules or/and of said cleaved-off nucleotide building blocks.
2. The process as claimed in claim 1, wherein a support made of glass, plastics, quartz, metals, semimetals, metal oxides or of a composite  
20 comprising a plurality of said materials is used.
3. The process as claimed in either of claims 1 and 2, wherein the enzyme molecules are selected from single-strand and double-strand exonucleases.
4. The process as claimed in claim 3, wherein T7 DNA polymerase, E.coli exonuclease I or E.coli exonuclease III is used.
5. The process as claimed in any of claims 1 to 4, wherein the enzyme  
30 molecules are immobilized on the support in a disordered manner.
6. The process as claimed in any of claims 1 to 4, wherein the enzyme molecules are immobilized in predefined regions on the support in an

- 19 -

ordered manner.

7. The process as claimed in any of claims 1 to 6, wherein the nucleic acid molecules are labeled such that at least 50% of all nucleotide building blocks of one or more base types carry a fluorescent labeling group.
8. The process as claimed in claim 7, wherein essentially all nucleotide building blocks of one base type carry a fluorescent labeling group.
9. The process as claimed in any of claims 1 to 8, wherein determining the base sequence comprises fluorescence excitation, preferably a multipoint fluorescence excitation, by lasers.
10. The process as claimed in claim 9, wherein the immobilized enzyme molecules are irradiated through the reaction space with laser light.
11. The process as claimed in claim 10, wherein fluorescence emission light emitted through the reaction space is determined.
12. The process as claimed in claim 9, wherein the immobilized enzyme molecules are irradiated with laser light through an at least partially optically transparent support.
13. The process as claimed in claim 12, wherein fluorescence emission light emitted through the support is determined.
14. The process as claimed in claim 9, wherein laser light is irradiated into the support and the formation of an evanescent excitation field is caused by internal reflection on the support surface in the region of the immobilized enzyme molecules.
15. The process as claimed in any of the preceding claims, wherein determining the base sequence comprises detecting the fluorescence

- 20 -

emission of a plurality of nucleic acid strands by means of a detection matrix.

- 5 16. The process as claimed in claim 15, wherein the detection matrix used is a CMOS camera, a CCD camera or an avalanche photodiode matrix.
17. The process as claimed in any of the preceding claims, wherein fluorescence excitation and fluorescence detection are carried out in parallel on all nucleic acid strands studied.
18. The process as claimed in any of claims 1 to 16, wherein fluorescence excitation and fluorescence detection are carried out in several steps, in each case on a portion of the nucleic acid strands studied.
- 15 19. The process as claimed in any of the preceding claims, wherein a convectional flow away from the support is generated during the determination.
- 20 20. The process as claimed in any of the preceding claims, wherein the fluorescent labeling groups are, at least partially, quenched when incorporated into the nucleic acid strands and the fluorescence intensity is increased after removal by cleavage.
- 25 21. The process as claimed in any of the preceding claims, wherein a support is provided which further has a multiplicity of nucleic acid-polymerizing enzyme molecules immobilized thereon and the free nucleic acid molecules carrying a plurality of fluorescent labeling groups are directly synthesized on the support.
- 30 22. An apparatus for sequencing nucleic acids, comprising
- a) a support comprising a multiplicity of enzyme molecules immobilized thereon,

- 21 -

- b) means for (i) delivering free nucleic acid molecules carrying a plurality of fluorescent labeling groups to said support or (ii) generating free nucleic acid molecules carrying a plurality of fluorescent labeling groups on said support,
- 5 c) a reaction space for progressively cleaving off individual nucleotide building blocks from said nucleic acid molecules, and
- d) means for determining simultaneously the base sequence of a plurality of nucleic acid molecules on the basis of the time-dependent change, caused when nucleotide building blocks are
- 10 cleaved off, in the fluorescence of said nucleic acid molecules or/and said cleaved-off nucleotide building blocks.
23. The apparatus of claim 22 wherein the enzyme molecules are selected from nucleic acid degrading enzymes optionally in combination with
- 15 nucleic acid polymerizing enzymes.
24. The use of the apparatus as claimed in claim 22 or 23 in a process as claimed in any of claims 1 to 21.
- 20 25. A process for studying enzyme-substrate interactions, comprising the following steps:
- (a) providing a support having a multiplicity of enzyme molecules immobilized thereon,
- 25 (b) contacting said support with free substrate molecules carrying fluorescent labeling groups,
- (c) reacting said substrate molecules with said immobilized enzyme molecules, and
- (d) studying simultaneously the enzyme-substrate interaction of a
- 30 plurality of individual enzyme or/and substrate molecules on the basis of the time-dependent change, caused during the enzymic reaction of substrate molecules, in the fluorescence of said substrate molecules or/and cleaved-off substrate parts or/and reaction

- 22 -

products formed.

- 5 26. The process as claimed in claim 25, wherein a support made of glass, plastics, quartz, metals, semimetals, metal oxides or of a composite comprising a plurality of said materials is used.
27. The process as claimed in any of claims 25 or 26, wherein the enzyme molecules are selected from polymerases, dehydrogenases and oxidoreductases.
28. The process as claimed in claim 27, wherein the enzyme-substrate interactions studied are hydrolysis, polymerization, transfer of functional groups, oxidation, reduction, dehydrogenation and/or isomerization.
- 15 29. The process as claimed in any of claims 25 to 27, wherein the enzyme molecules are immobilized on the support in a disordered manner.
- 20 30. The process as claimed in any of claims 25 to 27, wherein the enzyme molecules are immobilized in predefined regions on the support in an ordered manner.
31. The process as claimed in any of claims 25 to 30, wherein the study of enzyme-substrate interactions comprises a fluorescence excitation, preferably a multipoint fluorescence excitation, by lasers.
32. The process as claimed in claim 31, wherein the immobilized enzyme molecules are irradiated through the reaction space with laser light.
- 30 33. The process as claimed in claim 32, wherein fluorescence emission light emitted through the reaction space is determined.
34. The process as claimed in claim 33, wherein the immobilized enzyme molecules are irradiated with laser light through an at least partially

- 23 -

optically transparent support.

35. The process as claimed in claim 34, wherein fluorescence emission light emitted through the support is determined.

36. The process as claimed in claim 31, wherein laser light is irradiated into the support and the formation of an evanescent excitation field is caused by internal reflection on the support surface in the region of the immobilized enzyme molecules.

37. The process as claimed in any of claims 25 to 36, wherein the study of enzyme-substrate interactions comprises detecting the fluorescence emission of a plurality of substrate molecules by means of a detection matrix.

38. The process as claimed in claim 37, wherein the detection matrix used is a CMOS camera, a CCD camera or an avalanche photodiode matrix

20 39. The process as claimed in any of the preceding claims, wherein fluorescence excitation and fluorescence detection are carried out in parallel on all substrate molecules studied.

25 40. The process as claimed in any of claims 25 to 38, wherein fluorescence excitation and fluorescence detection are carried out in several steps, in each case on a portion of the substrate molecules studied.

41. The process as claimed in any of claims 25 to 40, wherein a convectional flow away from the support is generated during the determination.

42. An apparatus for studying enzyme-substrate interactions, comprising

(a) a support comprising a multiplicity of enzyme molecules immobilized



- 24 -

thereon,

(b) means for (i) delivering free substrate molecules carrying fluorescent labeling groups to said support or (ii) generating free substrate molecules carrying fluorescent labeling groups on said support,

5

(c) a reaction space for reacting said substrate molecules with said immobilized enzyme molecules, and

(d) means for studying simultaneously a plurality of enzyme-substrate interactions of a plurality of enzyme or/and substrate molecules on the basis of the time-dependent change, caused during the enzymic reaction, in the fluorescence of said substrate molecules or/and the cleaved-off substrate parts or/and reaction products formed.

10

43. The use of the apparatus as claimed in claim 42 in a process as claimed in any of claims 25 to 41.

15

44. A process for sequencing nucleic acids, comprising the following steps:

(a) providing a support having a multiplicity of nucleic acid-polymerising enzyme molecules immobilized thereon,

20

(b) contacting said support with (i) free nucleic acid molecules to which a primer molecule is bound and (ii) nucleotide building blocks carrying a fluorescent labeling group,

(c) incorporating said nucleotide building blocks into said primer by said immobilized enzyme molecules, and

25

(d) simultaneously determining the base sequence of a plurality of nucleic acid molecules on the basis of the time-dependent change, caused when nucleotide building blocks are incorporated, in the fluorescence of said extended primer molecule or/and of said nucleotide building blocks.

30

45. The process of claim 44, wherein incorporating of the nucleotide building blocks into the primer is associated with a detectable increase in the

- 25 -

fluorescence.

5 46. The process of claim 44 or 45, wherein nucleotide building blocks are used which carry a labeling group on the  $\gamma$ -phosphate group and a fluorescence quencher group on the portion of the molecule which is incorporated into the primer.

47. An apparatus for sequencing nucleic acids comprising:

- 10 (a) a support comprising a multiplicity of nucleic acid-polymerising enzyme molecules immobilized thereon,
- (b) means for providing (i) free nucleic acid molecules to which a primer molecule is bound and (ii) nucleotide building blocks carrying a fluorescent labeling group on the support,
- 15 (c) a reaction space for incorporating said nucleotide building blocks into said primer, and
- (d) means for determining simultaneously the base sequence of a plurality of nucleic acid molecules on the basis of the time-dependent change, caused when nucleotide building blocks are
- 20 incorporated, in the fluorescence of said extended primer molecule or/and of said nucleotide building blocks.

48. The use of the apparatus of claim 47 in a process as claimed in any of claims 44-46.

25

30

1/5

Figure 1

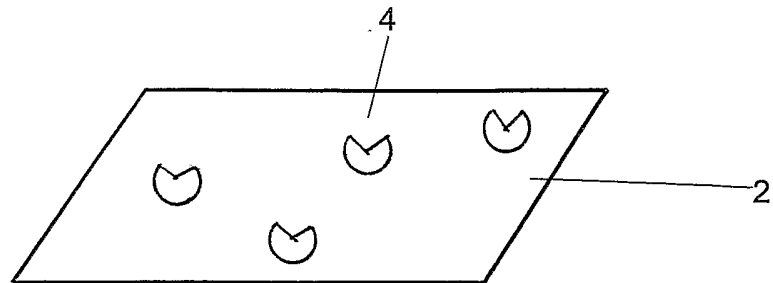
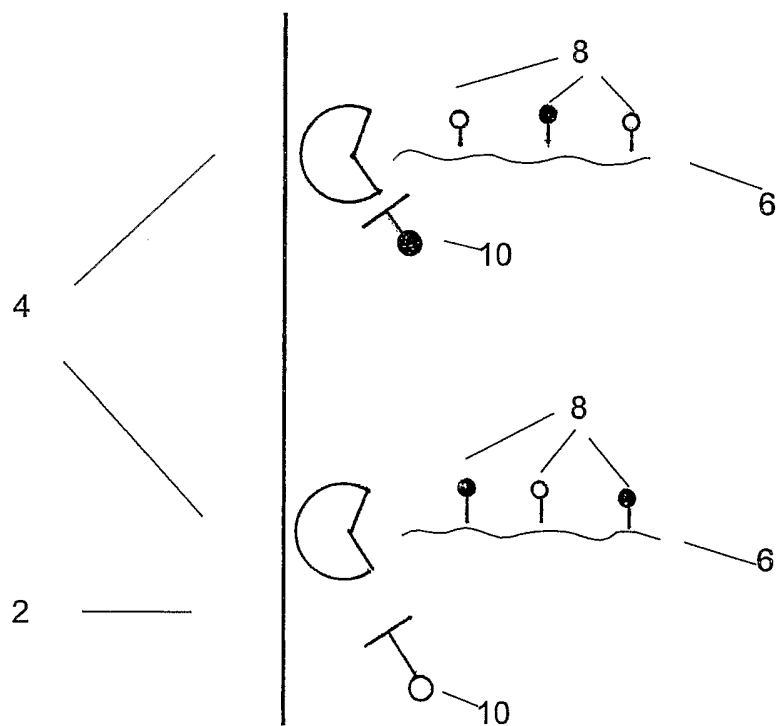


Figure 2



2/5

Figure 3A

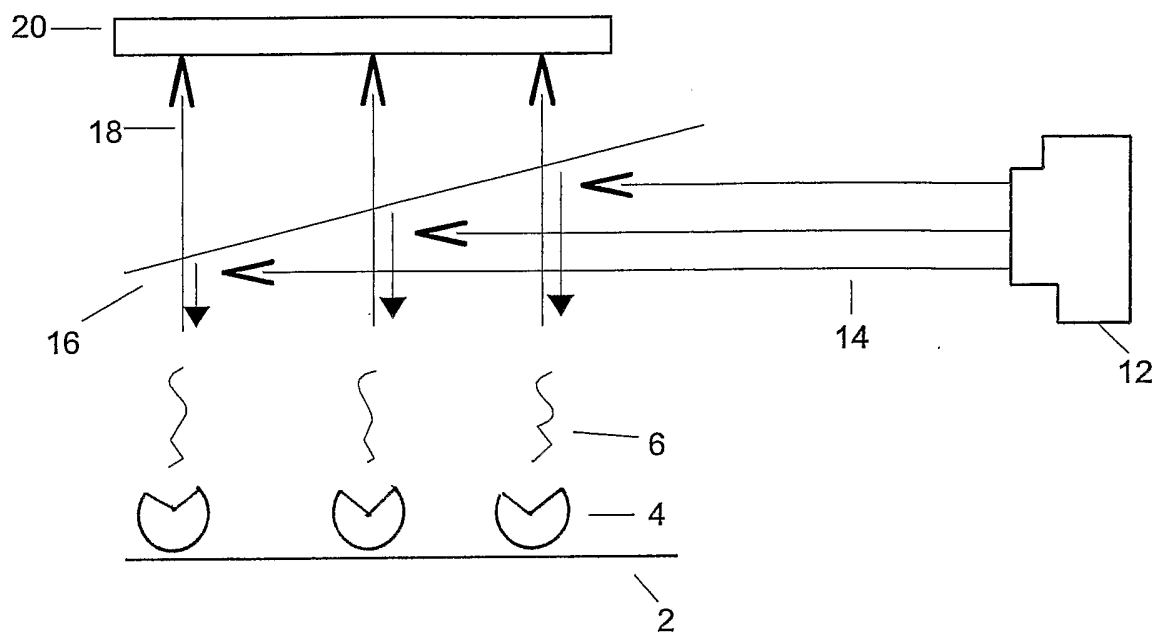


Figure 3B

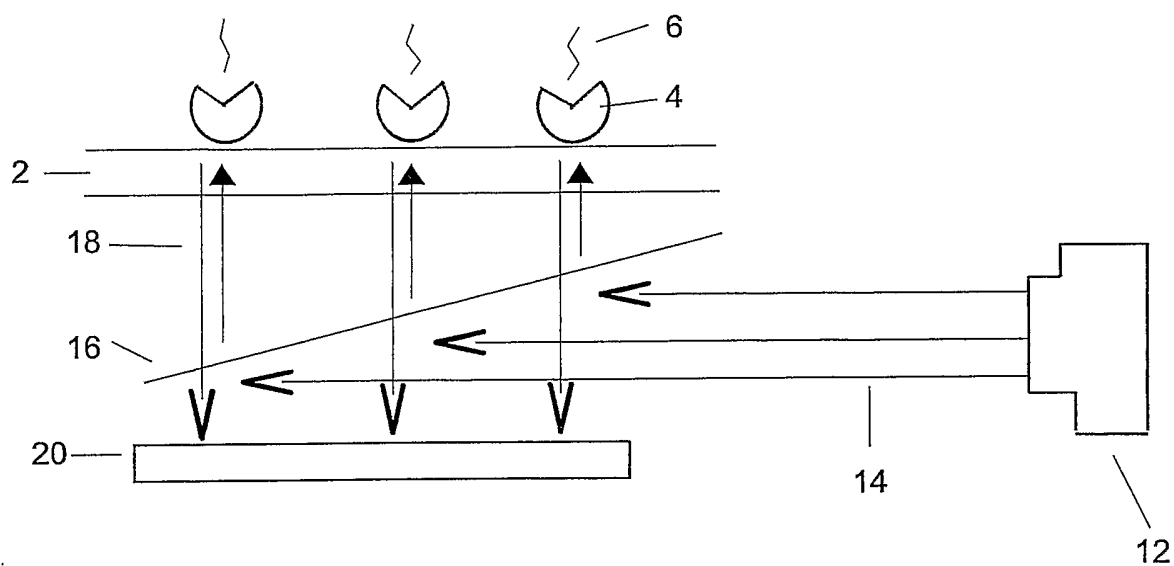


Figure 4

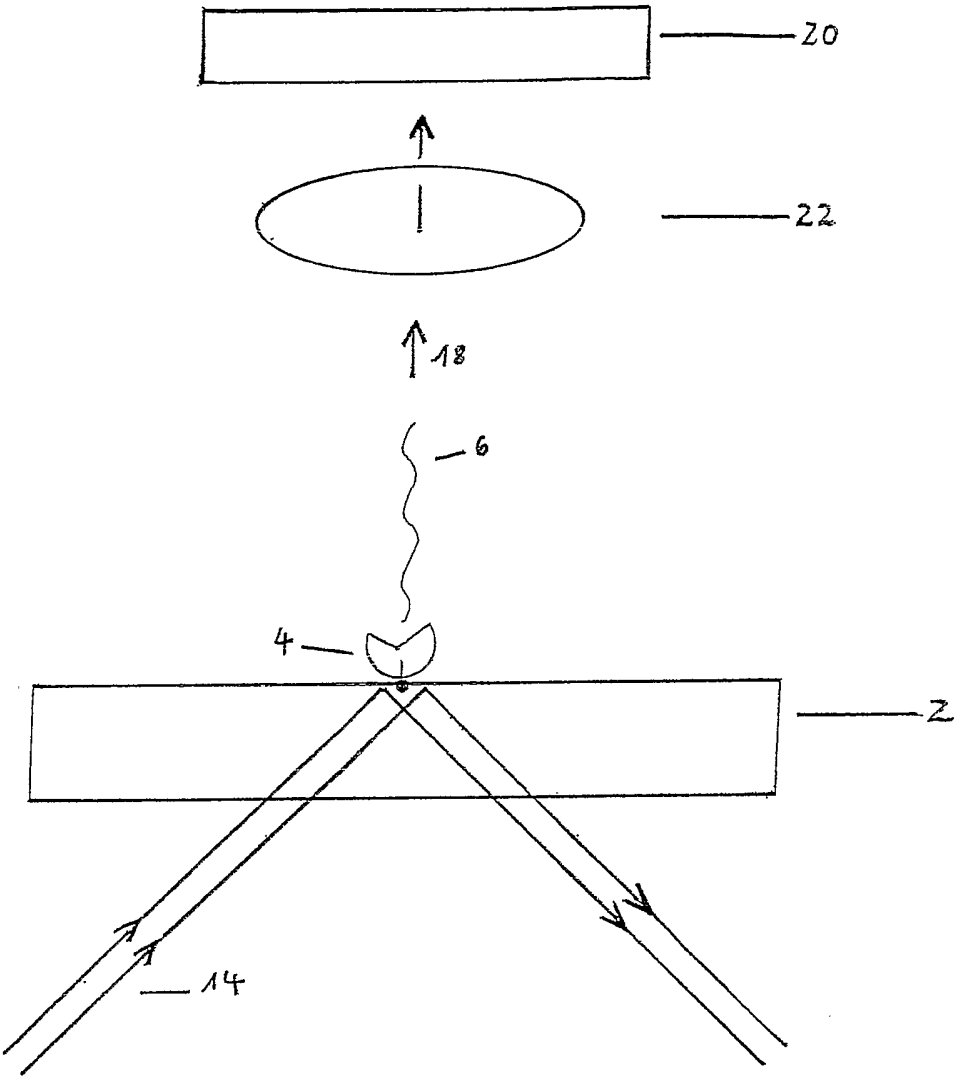


Figure 5

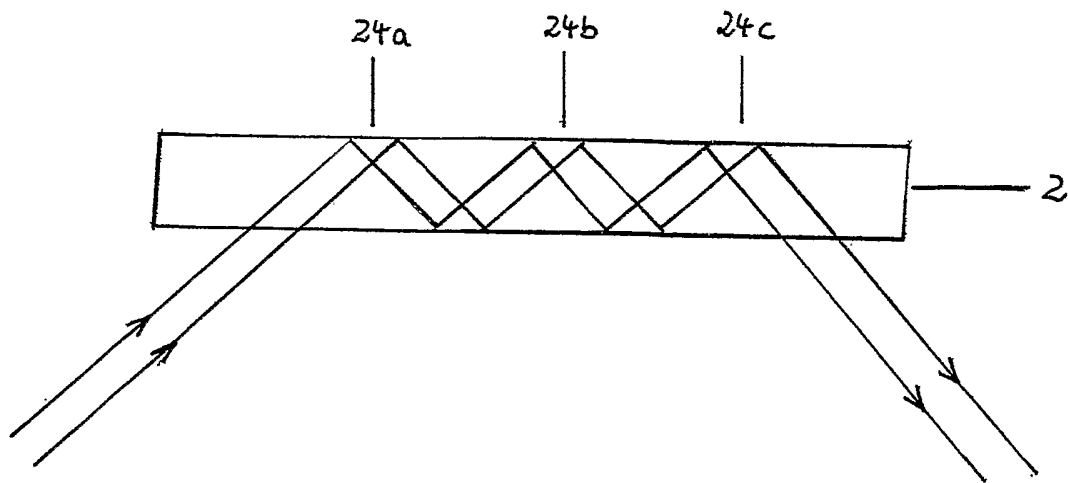


Figure 6A

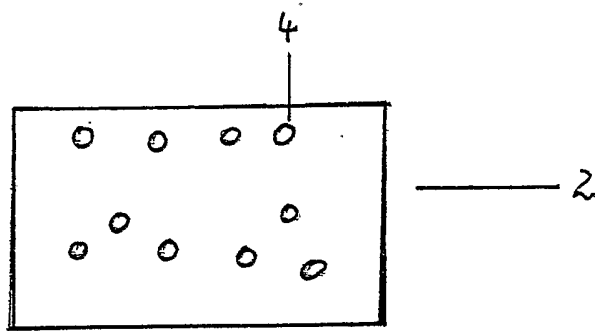


Figure 6B

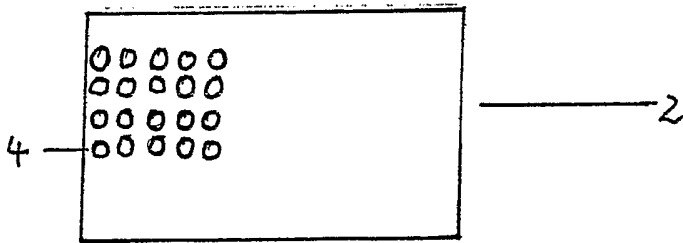
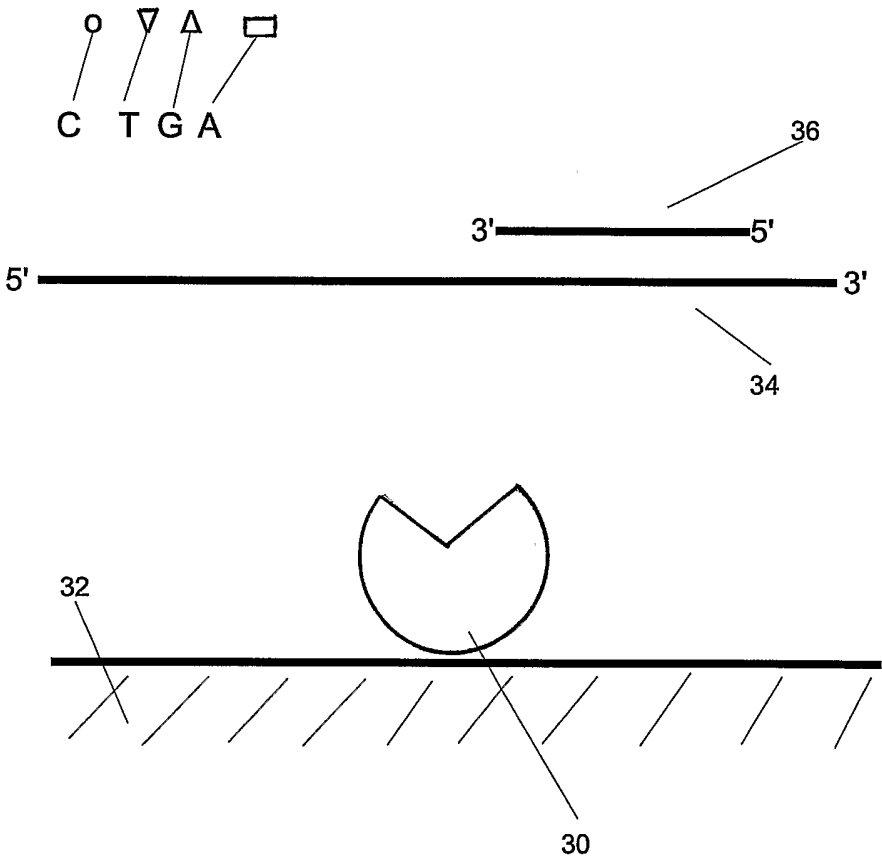


Figure 7



# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP2005/008511

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> C12Q1/68		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) C12Q		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, PAJ, BIOSIS		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 03/052136 A (GNOTHIS HOLDING SA; RIGLER, RUDOLF; VOGEL, HORST) 26 June 2003 (2003-06-26) the whole document -----	1-24
X	US 6 210 896 B1 (CHAN EUGENE Y) 3 April 2001 (2001-04-03) esp figure 1 the whole document -----	22-43
X	WO 01/16375 A (THE GOVERNMENT OF THE UNITED STATES OF AMERICA, AS REPRESENTED BY THE) 8 March 2001 (2001-03-08) the whole document -----	22-24, 44-48
-/--		
<div style="display: flex; justify-content: space-between;"> <div> <input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. </div> <div> <input checked="" type="checkbox"/> Patent family members are listed in annex. </div> </div>		
° Special categories of cited documents :		
<div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>*A* document defining the general state of the art which is not considered to be of particular relevance</p> <p>*E* earlier document but published on or after the international filing date</p> <p>*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>*O* document referring to an oral disclosure, use, exhibition or other means</p> <p>*P* document published prior to the international filing date but later than the priority date claimed</p> </div> <div style="width: 45%;"> <p>*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>*&amp;* document member of the same patent family</p> </div> </div>		
Date of the actual completion of the international search  18 November 2005		Date of mailing of the international search report  24/11/2005
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer  Cornelis, K



# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP2005/008511

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 00/70073 A (CORNELL RESEARCH FOUNDATION, INC) 23 November 2000 (2000-11-23) abstract; claims 1,40 -----	22-24, 44-48
A	WO 03/052137 A (GNOTHIS HOLDING SA; RIGLER, RUDOLF) 26 June 2003 (2003-06-26) the whole document -----	1-48

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP2005/008511

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 03052136	A	26-06-2003	AU 2002358756 A1 DE 10162535 A1 EP 1458890 A2 US 2005130159 A1	30-06-2003 03-07-2003 22-09-2004 16-06-2005
US 6210896	B1	03-04-2001	NONE	
WO 0116375	A	08-03-2001	AU 7086800 A	26-03-2001
WO 0070073	A	23-11-2000	AU 5027600 A CA 2373385 A1 EP 1179085 A1 JP 2002543847 T	05-12-2000 23-11-2000 13-02-2002 24-12-2002
WO 03052137	A	26-06-2003	AU 2002361162 A1 DE 10162536 A1 EP 1458892 A2	30-06-2003 17-07-2003 22-09-2004