



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 601 27 162 T2 2008.04.24**

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 337 541 B1**

(21) Deutsches Aktenzeichen: **601 27 162.9**

(86) PCT-Aktenzeichen: **PCT/US01/31243**

(96) Europäisches Aktenzeichen: **01 977 533.7**

(87) PCT-Veröffentlichungs-Nr.: **WO 2002/029003**

(86) PCT-Anmeldetag: **05.10.2001**

(87) Veröffentlichungstag

der PCT-Anmeldung: **11.04.2002**

(97) Erstveröffentlichung durch das EPA: **27.08.2003**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **07.03.2007**

(47) Veröffentlichungstag im Patentblatt: **24.04.2008**

(51) Int Cl.⁸: **C12Q 1/68 (2006.01)**

C07H 19/10 (2006.01)

C07H 19/207 (2006.01)

H01J 49/04 (2006.01)

(30) Unionspriorität:

684670 06.10.2000 US

300894 P 26.06.2001 US

(84) Benannte Vertragsstaaten:

**AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT,
LI, LU, MC, NL, PT, SE, TR**

(73) Patentinhaber:

**The Trustees of Columbia University in the City of
New York, New York, N.Y., US**

(72) Erfinder:

**ITAGAKI, Yasuhiro, New York, NY 10027, US; JU,
Jingyue, Englewood Cliffs, NJ 07632, US; LI,
Zengmin, New York, NY 10032, US; EDWARDS,
John Robert, New York, NY 10025, US**

(74) Vertreter:

**MÜLLER FOTTNER STEINECKE Rechtsanwälte
Patentanwälte, 80335 München**

(54) Bezeichnung: **Massives Parallelverfahren zur Dekodierung von DNA und RNA**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung

[0001] Die vorliegende Erfindung beansprucht die Priorität der vorläufigen US-Anmeldung Nr. 60/300,894, eingereicht am 26. Juni 2001, und ist eine Teilfortführungsanmeldung der US-Anmeldenummer 09/684,670, eingereicht am 06. Oktober 2000.

Hintergrund der vorliegenden Erfindung

[0002] Durch die Angabe von Autor und Jahr der Veröffentlichung (in Klammern) wird im Rahmen der vorliegenden Anmeldung auf verschiedene Veröffentlichungen Bezug genommen. Vollständige Angaben zu diesen Referenzen sind am Ende der Beschreibung unmittelbar vor den Ansprüchen zu finden. Der Offenbarungsgehalt dieser Veröffentlichungen wird hiermit in seiner Gesamtheit durch Bezugnahme in die vorliegende Anmeldung einbezogen, um den Stand der Technik, welchem die vorliegende Erfindung zuzuordnen ist, umfassender zu beschreiben.

[0003] Die Fähigkeit, Desoxyribonukleinsäure (DNA) exakt und schnell zu sequenzieren revolutioniert gegenwärtig die Fachgebiete der Biologie und der Medizin. Die Konfluenz des umfangreichen Human Genome Projekts verzeichnet ein exponentielles Wachstum in der Entwicklung von genetischen Analysetechnologien mit hohem Durchsatz. Diese rapide technologische Entwicklung, die die Fachgebiete Chemie, Ingenieurwesen, Biologie und Computerwissenschaften betrifft, ermöglicht es, nicht mehr nur einzelne Gene in einem Arbeitsschritt zu studieren, sondern vielmehr komplette Genome zu analysieren und miteinander zu vergleichen.

[0004] Mit der Fertigstellung der ersten vollständigen Sequenzkarte des menschlichen Genoms werden viele in höchstem Maße polymorphe Bereiche des Genoms sowohl in Exons als auch in Introns bekannt werden. Es ist eine Herausforderung für die Pharmakogenomik, die Gene und funktionalen Polymorphismen umfassend zu identifizieren, die mit der Variabilität der Reaktion auf Wirkstoffe in Verbindung gebracht werden. (Roses, 2000). Die Re-Sequenzierung der polymorphen Bereiche im Genom, die mit der Entwicklung einer Krankheit verknüpft sind, wird in hohem Maße zum Verständnis von Krankheiten wie Krebs sowie zur therapeutischen Entwicklung beitragen. Es besteht daher ein Bedarf an exakten Hochdurchsatzverfahren zur Re-Sequenzierung der in höchstem Maße variablen Intron-/Exonregionen des Genoms, um den gesamten Umfang des Potentials der vollständigen Sequenzkarte des menschlichen Genoms zu erforschen. Die derzeitige dem Stand der Technik entsprechende Technologie zur Hochdurchsatz-Sequenzierung von DNA, so wie sie für das Human Genome Projekt verwendet wird (Pennisi, 2000) stellen Kapillar-Array-DNA-Sequencer dar, welche sich der laserinduzierten Fluoreszenzdetektion bedienen (Smith et al., 1986; Ju et al., 1995, 1996; Kheterpal et al., 1996; Salas-Solano et al., 1998). Verbesserungen in der Polymerase, die zur gleichförmigen Terminationseffizienz führen, und die Einführung von thermostabilen Polymerasen konnten die Qualität der Sequenzierungsdaten ebenfalls beträchtlich verbessern (Tabor und Richardson, 1987, 1995). Wenngleich die Kapillar-Array-DNA-Sequenzierungstechnologie bis zu einem gewissen Grad die Anforderungen von groß angelegten DNA-Sequenzierungsprojekten bezüglich des Durchsatzes und der Leselänge erfüllt, so müssen doch der Durchsatz und die Genauigkeit, die für Mutationsstudien erforderlich sind, für eine große Vielzahl von Anwendungen verbessert werden – von der Entdeckung von krankheitsverursachenden Genen bis hin zur forensischen Identifikation. So haben z. B. die auf Elektrophorese basierenden DNA-Sequenzierungsverfahren Schwierigkeiten bei der eindeutigen Detektion von Heterozygoten und sind in Regionen, die reich an Nukleotiden umfassend Guanin oder Cytosin sind, aufgrund von Kompressionen nicht zu 100% exakt (Bowling et al., 1991; Yamakawa et al., 1997). Außerdem sind die ersten paar Basen im Anschluss an die Primingstelle oft durch das hohe Fluoreszenzsignal von überschüssigen mit Farbstoff markierten Primern oder mit Farbstoff markierten Terminatoren maskiert und sind daher schwer zu identifizieren. Daher stellt die Notwendigkeit der Elektrophorese zur DNA-Sequenzierung noch immer einen Hemmschuh für DNA-Sequenzierungs- und Mutationsdetektionsprojekte mit hohem Durchsatz dar.

[0005] Das Konzept der DNA-Sequenzierung durch Synthese ohne die Verwendung der Elektrophorese wurde im Jahr 1988 erstmals offenbart (Hyman, 1988) und beinhaltet die Detektion der Identität eines jeden Nukleotids, während es im Verlauf einer Polymerasereaktion in den wachsenden DNA-Strang integriert wird. Ein solches Schema, gekoppelt mit dem Chipformat und der laserinduzierten Fluoreszenzdetektion, hat das Potential, den Durchsatz von DNA-Sequenzierungsprojekten signifikant zu erhöhen. Demzufolge erforschten mehrere Gruppen ein solches System mit der Zielsetzung, ein DNA-Sequenzierungsverfahren mit äußerst hohem Durchsatz zu entwickeln (Cheeseman, 1994; Metzker et al., 1994). Bisher wurde noch nicht von einem umfassenden Erfolg bei der Verwendung eines solchen Systems zur eindeutigen Sequenzierung von DNA berichtet. Der Ansatz der Pyrosequenzierung unter Einsatz von vier natürlichen Nukleotiden (umfassend eine Base von Adenin (A), Cytosin (C), Guanin (G) oder Thymin (T)) und verschiedenen anderen Enzymen zur Se-

quenzierung von DNA durch Synthese wird mittlerweile weithin zur Detektion von Mutationen verwendet (Rognaghi, 1998). Bei diesem Ansatz basiert die Detektion auf dem Pyrophosphat (PPi), das während der DNA-Polymerasereaktion freigesetzt wird, der quantitativen Umwandlung von Pyrophosphat in Adenosintriphosphat (ATP) durch Sulfurylase und der nachfolgenden Produktion von sichtbarem Licht durch Luziferase aus Glühwürmchen. Mit diesem Verfahren können nur bis zu 30 Basenpaare (bps) von Nukleotidsequenzen sequenziert werden und jedes der 4 Nukleotide muss separat hinzugefügt und detektiert werden. Lange Abschnitte identischer Basen können mittels des Verfahrens der Pyrosequenzierung nicht eindeutig identifiziert werden.

[0006] Aktuellere Arbeiten in der Fachliteratur über DNA-Sequenzierung durch ein Syntheseverfahren ist zu meist auf die Konstruktion und die Synthese eines photospaltbaren chemischen Restes fokussiert, der an einen Fluoreszenzfarbstoff geknüpft wird, um die 3'-OH-Gruppe der Desoxynukleosid-Triphosphate (dNTPs) zu schützen (Welch et al., 1999; WO 00/53805, WO 91/06878, WO 01/92284). Es wird von einem teilweisen Erfolg bei der Integration des 3'-modifizierten Nukleotids mittels DNA-Polymerase berichtet. Der Grund dafür ist, dass die 3'-Position auf der Desoxyribose sehr nahe an den Aminosäureresten an der aktiven Stelle der Polymerase liegt und dass die Polymerase daher gegenüber einer Modifikation in diesem Bereich des Desoxyribo-serings empfindlich ist. Auf der anderen Seite ist bekannt, dass modifizierte DNA-Polymerasen (Thermo Sequenase und Taq FS Polymerase) dazu in der Lage sind, Nukleotide mit extensiven Modifikationen mit umfangreichen Gruppen so wie Energietransfer-Farbstoffen an der 5-Position der Pyrimidine (T und C) und an der 7-Position der Purine (G und A) zu erkennen (Rosenblum et al., 1997; Zhu et al., 1994). Es wurden die ternären Komplexe von Ratten-DNA-Polymerase, einem DNA-Templateprimer und Didesoxycytidintriphosphat (ddCTP) bestimmt (Pelletier et al., 1994), was diese Tatsache belegt. Wie in [Fig. 1](#) gezeigt, zeigt die 3D-Struktur an, dass der umliegende Bereich der 3'-Position des Desoxyribo-serings in ddCTP sehr überfüllt ist, während an der 5-Position der Cytidinbase ausreichend Platz für eine Modifikation vorhanden ist.

[0007] Der in der vorliegenden Anmeldung offenbarte Ansatz sieht vor, Nukleotidanaloga herzustellen, indem man eine eindeutige Markierung, so wie einen Fluoreszenzfarbstoff oder einen Masse-Tag, durch einen spaltbaren Linker an die Nukleotidbase knüpft, so wie an die 5-Position der Pyrimidine (T und C) und an die 7-Position der Purine (G und A), sich eines kleinen abspaltbaren chemischen Restes zu bedienen, um die 3'-OH-Gruppe der Desoxyribose zu schützen und diese damit unreaktiv zu machen und die Nukleotidanaloga als Terminatoren in den wachsenden DNA-Strang zu integrieren. Die Detektion der eindeutigen Markierung wird die Sequenzidentität des Nukleotids ergeben. Bei der Entfernung der Markierung und der 3'-OH-Schutzgruppe wird die Polymerasereaktion fortschreiten, um das nächste Nukleotidanaloga zu integrieren und die nächste Base zu detektieren.

[0008] Es ist ebenfalls wünschenswert, eine photospaltbare Gruppe zu verwenden, um die 3'-OH-Gruppe zu schützen. Eine photospaltbare Gruppe ist jedoch im Allgemeinen recht umfangreich und die DNA-Polymerase wird daher Schwierigkeiten haben, die Nukleotidanaloga zu integrieren, welche den photospaltbaren Rest enthalten, der die 3'-OH-Gruppe schützt. Wenn zum Schützen der 3'-OH-Gruppe kleine chemische Reste verwendet werden können, die man mit hoher Ausbeute chemisch leicht abspalten kann, so sollten solche Nukleotidanaloga auch als Substrate für die DNA-Polymerase angesehen werden. Es wurde berichtet, dass 3'-O-Methoxy-Desoxynukleotide gute Substrate für etliche Polymerasen sind (Axelrod et al., 1978). Es wurde ebenfalls gezeigt, dass 3'-O-Allyl-dATP mittels Ventr (exo-)DNA-Polymerase in den wachsenden DNA-Strang integriert wird (Metzker et al., 1994). Das Verfahren zur chemischen Abspaltung der Methoxygruppe ist jedoch stringent und erfordert anhydrische Bedingungen. Es ist daher nicht zweckmäßig, eine Methoxygruppe zu verwenden, um die 3'-OH-Gruppe für eine DNA-Sequenzierung durch Synthese zu schützen. Es wurde ebenfalls eine Estergruppe zum Schützen der 3'-OH-Gruppe des Nukleotids erforscht, es wurde jedoch gezeigt, dass diese von den Nukleophilen an der aktiven Stelle in der DNA-Polymerase abgespalten wurde (Canard et al., 1995). Chemische Gruppen mit Elektrophilen, wo wie Ketongruppen, sind nicht dazu geeignet, das 3'-OH des Nukleotids in enzymatischen Reaktionen zu schützen, da in der Polymerase starke Nukleophile anwesend sind. Es ist bekannt, dass MOM-(-CH₂OCH₃) und Allyl-(-CH₂CH=CH₂)-gruppen zum Schützen einer -OH-Gruppe verwendet und mit hoher Ausbeute chemisch abgespalten werden können (Ireland et al., 1986; Kamal et al., 1999). Der in der vorliegenden Anmeldung offenbarte Ansatz sieht vor, Nukleotidanaloga, die mit abspaltbaren eindeutigen Markierungen so wie Fluoreszenzfarbstoffen oder Masse-Tags versehen sind und in denen die 3'-OH-Gruppe mittels eines abspaltbaren chemischen Restes so wie entweder einer MOM-Gruppe (-CH₂OCH₃) oder einer Allylgruppe (-CH₂CH=CH₂) geschützt wird, als Terminatoren in den wachsenden DNA-Strang zu integrieren. Der optimierte Nukleotidsatz (3'-RO-A_{LABEL1}, 3'-RO-C_{LABEL2}, 3'-RO-G_{LABEL3}, 3'-RO-T_{LABEL4}, worin R die chemische Gruppe bezeichnet, die zum Schützen des 3'-OH verwendet wird) kann dann zur DNA-Sequenzierung mittels des Syntheseansatzes verwendet werden.

[0009] Die Verwendung der Massenspektrometrie (MS) zur Detektion von kleinen und stabilen Molekülen bie-

tet viele Vorteile. Beispielsweise kann die Massenauflösung so hoch wie bei einem Dalton liegen. Daher bietet der in der vorliegenden Anmeldung offenbarte MS-Ansatz, im Vergleich zu gelelektrophoretischen Sequenzierungssystemen und dem Ansatz der laserinduzierten Fluoreszenzdetektion, die sich überschneidende Fluoreszenzemissionsspektren aufweisen, was zu Schwierigkeiten bei der Detektion von Heterozygoten führt, eine äußerst hohe Auflösung von Sequenzierungsdaten durch Detektion der abgespaltenen kleinen Masse-Tags anstatt des langen DNA-Fragments. Dieses Verfahren bietet ebenfalls eine extrem schnelle Trennung in einer Größenordnung von Mikrosekunden. Die hohe Auflösung gestattet eine exakte digitale Detektion von Mutationen und Heterozygoten. Es ist ein weiterer Vorteil der Sequenzierung mittels Massenspektrometrie durch Detektion der kleinen Masse-Tags, dass die mit den auf Gel basierenden Systemen in Verbindung gebrachten Kompressionen vollständig eliminiert werden.

[0010] Um ein durchgehendes hybridisiertes Primerextensionsprodukt mit der Template-DNA aufrecht zu erhalten, kann ein Primer, der eine stabile Schleife enthält, um eine Einheit zu bilden, welche zum Selbst-Priming in einer Polymerasereaktion in der Lage ist, an das 3'-Ende eines jeden einzelsträngigen DNA-Templates, das auf einer festen Oberfläche so wie einem Chip immobilisiert ist, ligiert werden. Dieser Ansatz wird das Problem lösen, dass das wachsende Extensionsprodukt in jedem Zyklus weggewaschen wird.

[0011] Saxon und Bertozzi (2000) entwickelten eine elegante und in höchstem Maße spezifische Kopplungschemie, die eine spezifische Gruppe, welche einen Phosphinrest enthält, an eine Azidogruppe an der Oberfläche einer biologischen Zelle knüpft. In der vorliegenden Anmeldung wird sich dieser Kopplungschemie bedient, um eine feste Oberfläche zu schaffen, welche mit einem kovalent verknüpften Phosphinrest beschichtet ist, und um Produkte einer Polymerase-Kettenreaktion (PCR) zu erzeugen, die am 5'-Ende eine Azidogruppe zur spezifischen Kopplung des DNA-Templates mit der festen Oberfläche enthalten. Ein Beispiel für eine feste Oberfläche sind Glaskanäle, welche eine Innenwand mit einer unebenen oder porösen Oberfläche aufweisen, um die Angriffsfläche zu vergrößern. Ein weiteres Beispiel ist ein Chip.

[0012] Die vorliegende Anmeldung offenbart ein neues und vorteilhaftes System zur DNA-Sequenzierung mittels des Syntheseansatzes, in dem ein stabiles DNA-Template verwendet wird, welches für die Polymerasereaktion zum Selbst-Priming in der Lage und kovalent an eine feste Oberfläche wie einen Chip gebunden ist, sowie 4 eindeutige Nukleotidanaloga (${}_{3\text{'-RO}}\text{-A}_{\text{LABEL1}}$, ${}_{3\text{'-RO}}\text{-C}_{\text{LABEL2}}$, ${}_{3\text{'-RO}}\text{-G}_{\text{LABEL3}}$, ${}_{3\text{'-RO}}\text{-T}_{\text{LABEL4}}$). Der Erfolg dieses neuen Systems wird die Entwicklung eines DNA-Sequenzierungssystems für Polymorphismus, pharmakogenetische Anwendungen und für die Sequenzierung von vollständigen Genomen gestatten, welches sich durch einen äußerst hohen Durchsatz sowie eine hohe Genauigkeit auszeichnet. Diese schnelle und exakte DNA-Resequenzierungssystem wird benötigt auf Gebieten wie der Detektion von Single-Nucleotide-Polymorphismen (SNPs) (Chee et al., 1996), der seriellen Analyse der Genexpression (SAGE) (Velculescu et al., 1995), der forensischen Identifikation sowie in Studien zur genetischen Krankheitsassoziation.

Zusammenfassung der vorliegenden Erfindung

[0013] Die vorliegende Erfindung betrifft ein Verfahren zur Sequenzierung einer Nukleinsäure mittels sequenzieller Bestimmung der Identität eines Nukleotidanalogs nachdem das Nukleotidanalogen in einer Polymerasereaktion in einen wachsenden DNA-Strang integriert ist, welches die folgenden Schritte umfasst:

- i. Befestigung eines 5'-Endes der Nukleinsäure an einer festen Oberfläche;
- ii. Befestigung eines Primers an die an der festen Oberfläche befestigte Nukleinsäure;
- iii. Hinzufügen einer Polymerase und eines oder mehrerer unterschiedlicher Nukleotidanaloga an die Nukleinsäure, um dadurch ein Nukleotidanalogen in den wachsenden DNA-Strang zu integrieren, wobei das integrierte Nukleotidanalogen die Polymerasereaktion terminiert und wobei jedes unterschiedliche Nukleotidanalogen umfasst: (a) eine Base ausgewählt aus der Gruppe bestehend aus Adenin, Guanin, Cytosin, Thymin und Uracil sowie deren Analoga; (b) eine mittels eines spaltbaren Linkers an der Base oder einem Analogon der Base befestigte eindeutige Markierung; (c) eine Desoxyribose; und (d) eine abspaltbare chemische Gruppe zum Schützen einer -OH-Gruppe an einer 3'-Position der Desoxyribose, wobei die abspaltbare chemische Gruppe $-\text{CH}_2\text{OCH}_3$ oder $-\text{CH}_2\text{CH}=\text{CH}_2$ ist;
- iv. Waschen der festen Oberfläche, um nicht integrierte Nukleotidanaloga zu entfernen;
- v. Bestimmung der Identität der an dem Nukleotidanalogen befestigten eindeutigen Markierung, welches in den wachsenden DNA-Strang integriert wurde, um dadurch das integrierte Nukleotidanalogen zu identifizieren;
- vi. Hinzufügen einer oder mehrerer chemischer Verbindungen, um jegliche nicht reagierte -OH-Gruppe auf dem an der Nukleinsäure befestigten Primer oder auf einem durch Hinzufügen eines oder mehrerer Nukleotide oder Nukleotidanaloga zu dem Primer gebildeten Primerextensionsstrangs permanent zu schützen;
- vii. Abspaltung des abspaltbaren Linkers zwischen dem Nukleotidanalogen, das in den wachsenden

DNA-Strang integriert wurde, und der eindeutigen Markierung;

viii. Abspaltung der abspaltbaren chemischen Gruppe, welche die -OH-Gruppe an der 3'-Position der Desoxyribose schützt, um die -OH-Gruppe zu entschützen und Waschen der festen Oberfläche, um abgespaltene Verbindungen zu entfernen; und

ix. Wiederholen der Schritte (iii) bis einschließlich (viii), um so für jede Wiederholung die Identität des neu in den wachsenden DNA-Strang integrierten Nukleotidanalogs zu bestimmen;

wobei in dem Fall, dass es sich bei der eindeutigen Markierung um einen Farbstoff handelt, die Reihenfolge der Schritte (v) bis einschließlich (vii) folgendermaßen lautet: (v), (vi) und (vii); und

wobei in dem Fall, dass es sich bei der eindeutigen Markierung um einen Masse-Tag handelt, die Reihenfolge der Schritte (v) bis einschließlich (vii) folgendermaßen lautet: (vi), (vii) und (v).

[0014] Die vorliegende Erfindung stellt ein Verfahren zur Befestigung einer Nukleinsäure an eine feste Oberfläche bereit, welches umfasst:

- i. Beschichten der festen Oberfläche mit einem Phosphinrest,
- ii. Befestigen einer Azidogruppe an ein 5'-Ende der Nukleinsäure, und
- iii. Immobilisieren des 5'-Endes der Nukleinsäure an die feste Oberfläche durch Interaktion zwischen dem Phosphinrest auf der festen Oberfläche und der Azidogruppe am 5'-Ende der Nukleinsäure.

[0015] Die vorliegende Erfindung stellt ein Nukleotidanalogs bereit, welches umfasst:

- a) eine Base ausgewählt aus der Gruppe bestehend aus Adenin oder einem Analogon von Adenin, Cytosin oder einem Analogon von Cytosin, Guanin oder einem Analogon von Guanin, Thymin oder einem Analogon von Thymin sowie Uracil oder einem Analogon von Uracil;
- b) eine eindeutige Markierung, die durch einen spaltbaren Linker an der Base oder einem Analogon der Base befestigt ist;
- c) eine Desoxyribose; und
- d) eine abspaltbare chemische Gruppe, um eine -OH-Gruppe an einer 3'-Position der Desoxyribose zu schützen, wobei die abspaltbare Gruppe $-\text{CH}_2\text{OCH}_3$ oder $-\text{CH}_2\text{CH}=\text{CH}_2$ ist.

Kurze Beschreibung der Figuren

[0016] Fig. 1: Die 3D-Struktur des ternären Komplexes von Ratten-DNA-Polymerase, ein DNA-Template-Primer und Didesoxycytidintriphosphat (ddCTP). Die linke Seite der Darstellung zeigt den Mechanismus zur Hinzufügung von ddCTP und die rechte Seite der Darstellung zeigt die aktive Seite der Polymerase. Hier ist zu beachten, dass die 3'-Position des Di-Desoxyribooserings sehr überfüllt ist, während an der 5'-Position der Cytidinbase ausreichend Platz zur Verfügung steht.

[0017] Fig. 2A-Fig. 2B: Schematische Darstellung der Sequenzierung mittels des Syntheseansatzes. A: Beispiel, in dem die eindeutigen Markierungen Farbstoffe sind und die feste Oberfläche ein Chip ist. B: Beispiel, in dem die eindeutigen Markierungen Masse-Tags sind und die feste Oberfläche aus in einen Glaschip geätzten Kanälen besteht. A, C, G, T; Nukleotidtriphosphate umfassend die Basen Adenin, Cytosin, Guanin und Thymin; d, Desoxy; dd, Didesoxy; R, abspaltbare chemische Gruppe, welche verwendet wird, um die -OH-Gruppe zu schützen; Y, abspaltbarer Linker.

[0018] Fig. 3: Das Syntheschema für die Immobilisierung eines azido (N_3)-markierten DNA-Fragments an eine feste Oberfläche, die mit einem Triarylphosphinrest beschichtet ist. Me, Methylgruppe; Phosphor; Ph, Phenyl.

[0019] Fig. 4: Die Synthese des Triarylphosphin-N-hydrosuccinimid (NHS)-Esters.

[0020] Fig. 5: Das Syntheschema zur Befestigung einer Azido (N_3)-Gruppe mittels eines Linkers an das 5'-Ende eines DNA-Fragments, welches dann verwendet wird, um mit dem Triarylphosphinrest auf einer festen Oberfläche zu koppeln. DMSO, Dimethylsulfonyloxid.

[0021] Fig. 6A-Fig. 6B: Ligieren des Primers mit Schleife (B) an das immobilisierte einzelsträngige DNA-Template, wobei ein selbst-geprimter DNA-Template-Rest auf einer festen Oberfläche gebildet wird. P (im Kreis), Phosphat.

[0022] Fig. 7: Beispiele von Strukturen für vier Nukleotidanaloga zur Verwendung bei der Sequenzierung mittels des Syntheseansatzes. Jedes Nukleotidanalogs weist einen mittels eines photospaltbaren Linkers an ei-

ner Base befestigten eindeutigen Fluoreszenzfarbstoff auf und das 3'-OH ist entweder exponiert oder geschützt mit einer MOM-Gruppe oder einer Allylgruppe. FAM, 5-Carboxyfluorescein; R6G, 6-Carboxyrhodamin-6G; TAM, N,N,N',N'-Tetramethyl-6-carboxyrhodamin; ROX, 6-Carboxy-X-rhodamin. R=H, CH₂OCH₃ (MOM) oder CH₂CH=CH₂ (Allyl).

[0023] Fig. 8: Ein repräsentatives Schema für die Synthese des Nucleotidanalogs ${}^{3'-RO-G-Tam}$. Ein ähnliches Schema kann verwendet werden, um die anderen drei modifizierten Nucleotide zu konstruieren: ${}^{3'-RO-A-Farbstoff1}$, ${}^{3'-RO-C-Farbstoff2}$, ${}^{3'-RO-T-Farbstoff4}$. (i) Tetrakis (Triphenylphosphin) Palladium (0); (ii) POCl₃, Bn₄N⁺Pyrophosphat; (iii) NH₄OH; (iv) Na₂CO₃/NaHCO₃ (pH = 9,0)/DMSO.

[0024] Fig. 9: Ein Schema zum Testen der Sequenzierung mittels des Syntheseansatzes. Jedes Nucleotid, modifiziert durch die Befestigung eines eindeutigen Fluoreszenzfarbstoffs, wird nacheinander hinzugefügt, basierend auf dem komplementären Template. Der Farbstoff wird detektiert und abgespalten, um den Ansatz zu testen. Farbstoff 1 = Fam; Farbstoff 2 = R6G; Farbstoff 3 = Tam; Farbstoff 4 = Rox.

[0025] Fig. 10: Die erwarteten Photospaltungsprodukte von DNA enthaltend einen photospaltbaren Farbstoff (Tam). Lichtabsorption (300–360 nm) durch den aromatischen 2-Nitrobenzyl-Rest bewirkt die Reduktion der 2-Nitro-Gruppe zu einer Nitroso-Gruppe und eine Sauerstoffinsertion in die an der 2-Position gelegene Kohlenstoff-Wasserstoff-Bindung, gefolgt von einer Abspaltung und Decarboxylierung (Pillai, 1980).

[0026] Fig. 11: Synthese von PC-LC-Biotin-FAM, um die Photolyseeffizienz des Fluorophors zu bewerten, das mit der photospaltbaren 2-Nitrobenzyl-Linker-Gruppe gekoppelt ist.

[0027] Fig. 12: Fluoreszenzspektra ($\lambda_{ex} = 480$ nm) von PC-LC-Biotin-FAM, das auf einem mit Streptavidin beschichteten Mikroskopslide aus Glas immobilisiert ist (a); nach 10 min Photolyse ($\lambda_{irr} = 350$ nm; $\sim 0,5$ mW/cm²) (b); und nach Waschen mit Wasser, um den photogespaltene Farbstoff zu entfernen (c).

[0028] Fig. 13A-Fig. 13B: Syntheschema für das Schützen des 3'-OH des Nucleotids.

[0029] Fig. 14: Chemische Abspaltung der MOM-Gruppe (oberste Reihe) und der Allylgruppe (unterste Reihe), um das 3'-OH in dem Nucleotid freizulegen. CITMS = Chlortrimethylsilan.

[0030] Fig. 15A-Fig. 15B: Beispiele für energietransfergekoppelte Färbesysteme, wobei Fam oder Cy2 als Lichtabsorber (Energietransferdonor) und Cl₂Fam, Cl₂R6G, Cl₂Tam oder Cl₂Rox als Energietransferakzeptor verwendet wird. Cy2, Cyanin; FAM, 5-Carboxyfluorescein; R6G, 6-Carboxy-rhodamin-6G; TAM, N,N,N',N'-Tetramethyl-6-carboxyrhodamin; ROX, 6-Carboxy-X-rhodamin.

[0031] Fig. 16: Die Synthese eines photospaltbaren mit einem Farbstoff markierten Energietransfer-Nucleotids. DMF, Dimethylformid. DEC = 1-(3-Dimethylaminopropyl)-3-Ethylcarbodiimid-hydrochlorid. R=H, CH₂OCH₃ (MOM) oder CH₂CH=CH₂ (Allyl).

[0032] Fig. 17: Strukturen von vier Masse-Tag-Vorläufern und vier photoaktiven Masse-Tags. Vorläufer: a) Acetophenon; b) 3-Fluoracetophenon; c) 3,4-Difluoracetophenon; und d) 3,4-Dimethoxyacetophenon. Es werden vier photoaktive Masse-Tags verwendet, um für die Identität eines jeden der vier Nucleotide (A, C, G, T) zu kodieren.

[0033] Fig. 18: "Atmospheric Pressure Chemical Ionization" (APCI)-Massenspektrum von in Fig. 17 gezeigten Masse-Tag-Vorläufern.

[0034] Fig. 19: Beispiele für Strukturen von vier Nucleotidanaloga zur Verwendung bei der Sequenzierung mittels des Syntheseansatzes. Jedes Nucleotidanaloga weist einen mittels eines photospaltbaren Linkers an der Base befestigten eindeutigen Masse-Tag auf und das 3'-OH ist entweder exponiert oder geschützt mit einer MOM-Gruppe oder einer Allylgruppe. Die eckigen Klammern zeigten an, dass der Masse-Tag abspaltbar ist. R=H, CH₂OCH₃ (MOM) oder CH₂CH=CH₂ (Allyl).

[0035] Fig. 20: Beispiel für die Synthese des NHS-Esters eines Masse-Tags (Tag-3). Ein ähnliches Schema wird zur Konstruktion anderer Masse-Tags verwendet.

[0036] Fig. 21: Ein repräsentatives Schema für die Synthese des Nucleotidanalogs ${}^{3'-RO-G-Tag3}$. Ein ähnliches Schema wird verwendet, um die drei anderen modifizierten Basen zu konstruieren: ${}^{3'-RO-A-Tag1}$, ${}^{3'-RO-C-Tag2}$,

$3'$ -RO⁻T-Tag⁴. (i) Tetrakis (Triphenylphosphin) Palladium (0); (ii) POCl₃, Bn₄N⁺Pyrophosphat; (iii) NH₄OH; (iv) Na₂CO₃/NaHCO₃ (pH = 9,0)/DMSO.

[0037] [Fig. 22](#): Beispiele für erwartete Photospaltungsprodukte von DNA enthaltend einen photospaltbaren Masse-Tag.

Detaillierte Beschreibung der Erfindung

[0038] Die folgenden Definitionen werden als Hilfestellung zum besseren Verständnis der vorliegenden Erfindung bereitgestellt.

[0039] Wie hierin verwendet, bezeichnet der Ausdruck, eine -OH-Gruppe zu schützen, das Ersetzen des "H" in der -OH-Gruppe durch eine chemische Gruppe. Wie hierin offenbart, ist die -OH-Gruppe des Nukleotidanalogs mit einer abspaltbaren chemischen Gruppe geschützt. Eine -OH-Gruppe zu entschützen bedeutet, die chemische Gruppe von einer geschützten -OH-Gruppe abzuspalten und die chemische Gruppe durch "H" zu ersetzen, d.h. das "R" in -OR durch "H" zu ersetzen, wobei "R" die chemische Gruppe ist, die verwendet wird, um die -OH-Gruppe zu schützen.

[0040] Die Nukleotidbasen werden wie folgt abgekürzt: Adenin (A), Cytosin (C), Guanin (G), Thymin (T) und Uracil (U).

[0041] Ein Analogon einer Nukleotidbase bezieht sich auf ein strukturelles und funktionelles Derivat der Base eines Nukleotids, welches von einer Polymerase als Substrat erkannt werden kann. Das heißt z. B., dass ein Analogon von Adenin (A) Wasserstoffbrückenbindungen mit Thymin (T), ein C-Analogon Wasserstoffbrückenbindungen mit G, ein G-Analogon Wasserstoffbrückenbindungen mit C und dass ein T-Analogon Wasserstoffbrückenbindungen mit A in einem Doppelhelixformat bilden sollte. Beispiele für Analoga von Nukleotidbasen schließen ein, sind jedoch nicht beschränkt auf, 7-Deazaadenin und 7-Deazaguanin, wobei das Stickstoffatom an der 7'-Position von Adenin oder Guanin durch ein Kohlenstoffatom substituiert wird.

[0042] Ein Nukleotidanalogen bezieht sich auf eine chemische Verbindung, die strukturell und funktionell dem Nukleotid ähnlich ist, d.h. das Nukleotidanalogen kann von Polymerase als ein Substrat erkannt werden. Das heißt z. B., dass ein Nukleotidanalogen umfassend Adenin oder ein Analogon von Adenin Wasserstoffbrückenbindungen mit Thymin, ein Nukleotidanalogen umfassend C oder ein Analogon von C Wasserstoffbrückenbindungen mit G, ein Nukleotidanalogen umfassend G oder ein Analogon von G Wasserstoffbrückenbindungen mit C und dass ein Nukleotidanalogen umfassend T oder ein Analogon von T Wasserstoffbrückenbindungen mit A in einem Doppelhelixformat bilden sollte. Beispiele für hierin offenbarte Nukleotidanaloga schließen Analoga ein, welche ein Analogon der Nukleotidbase umfassen, so wie 7-Deazaadenin oder 7-Deazaguanin, wobei das Stickstoffatom an der 7-Position von Adenin oder Guanin durch ein Kohlenstoffatom substituiert wird. Weitere Beispiele schließen Analoga ein, in denen mittels eines spaltbaren Linkers eine Markierung an der 5-Position von Cytosin oder Thymin oder an der 7-Position von Deazaadenin oder Deazaguanin befestigt ist. Andere Beispiele schließen Analoga ein, in denen ein kleiner chemischer Rest, so wie -CH₂OCH₃ oder -CH₂CH=CH₂, verwendet wird, um die -OH-Gruppe an der 3'-Position von Desoxyribose zu schützen. Analoga von Didesoxynukleotiden können in ähnlicher Weise hergestellt werden.

[0043] Wie hierin verwendet, ist eine poröse Oberfläche eine Oberfläche, die Poren aufweist oder anderweitig uneben ist, so dass die Angriffsfläche der porösen Oberfläche in Relation zur Angriffsfläche einer glatten Oberfläche größer ist.

[0044] Die vorliegende Erfindung betrifft ein Verfahren zur Sequenzierung einer Nukleinsäure mittels der sequentiellen Bestimmung der Identität eines Nukleotidanalogs nachdem das Nukleotidanalogen in einer Polymerasereaktion in einen wachsenden DNA-Strang integriert worden ist, welches die folgenden Schritte umfasst:

- i. Befestigen des 5'-Endes der Nukleinsäure an eine feste Oberfläche;
- ii. Befestigen eines Primers an die an der festen Oberfläche befestigte Nukleinsäure;
- iii. Hinzufügen einer Polymerase und eines oder mehrerer Nukleotidanaloga an die Nukleinsäure, um somit ein Nukleotidanalogen in den wachsenden DNA-Strang zu integrieren, wobei das integrierte Nukleotidanalogen die Polymerasereaktion terminiert und wobei jedes unterschiedliche Nukleotidanalogen umfasst: (a) eine Base ausgewählt aus der Gruppe bestehend aus Adenin, Guanin, Cytosin, Thymin und Uracil und deren Analoga; (b) eine eindeutige Markierung, die mittels eines spaltbaren Linkers an der Base oder an einem Analogon der Base befestigt ist; (c) eine Desoxyribose; und (d) eine spaltbare chemische Gruppe, um

eine -OH-Gruppe an einer 3'-Position der Desoxyribose zu schützen, wobei die spaltbare chemische Gruppe $-\text{CH}_2\text{OCH}_3$ oder $-\text{CH}_2\text{CH}=\text{CH}_2$ ist;

- iv. Waschen der festen Oberfläche, um nicht integrierte Nukleotidanaloga zu entfernen;
- v. Bestimmen der Identität der an dem Nukleotidanalogen befestigten eindeutigen Markierung, die in den wachsenden DNA-Strang integriert wurde, um somit das integrierte Nukleotidanalogen zu identifizieren;
- vi. Hinzufügen einer oder mehrerer chemischer Verbindungen, um jegliche nicht reaktierten -OH-Gruppen auf dem an der Nukleinsäure befestigten Primer oder auf einem Primerextensionsstrang, der durch das Hinzufügen eines oder mehrerer Nukleotide oder Nukleotidanaloga an den Primer gebildet wurde, permanent zu schützen;
- vii. Abspalten des spaltbaren Linkers zwischen dem Nukleotidanalogen, das in den wachsenden DNA-Strang integriert wurde, und der eindeutigen Markierung;
- viii. Abspalten der spaltbaren chemischen Gruppe, die die -OH-Gruppe an der 3'-Position der Desoxyribose schützt, um die -OH-Gruppe zu entschützen und Waschen der festen Oberfläche, um abgespaltene Verbindungen zu entfernen; und
- ix. Wiederholen der Schritte (iii) bis einschließlich (viii), um somit für jede Wiederholung die Identität des neu in den wachsenden DNA-Strang integrierten Nukleotidanalogs zu bestimmen;

wobei in dem Fall, dass es sich bei der eindeutigen Markierung um einen Farbstoff handelt, die Reihenfolge der Schritte (v) bis einschließlich (vii) folgendermaßen lautet: (v), (vi) und (vii); und

wobei in dem Fall, dass es sich bei der eindeutigen Markierung um einen Masse-Tag handelt, die Reihenfolge der Schritte (v) bis einschließlich (vii) folgendermaßen lautet: (vi), (vii) und (v).

[0045] In einer Ausführungsform eines jeglichen der hierin beschriebenen Nukleotidanaloga handelt es sich bei der Nukleotidbase um Adenin. In einer Ausführungsform handelt es sich bei der Nukleotidbase um Guanin. In einer Ausführungsform handelt es sich bei der Nukleotidbase um Cytosin. In einer Ausführungsform handelt es sich bei der Nukleotidbase um Thymin. In einer Ausführungsform handelt es sich bei der Nukleotidbase um Uracil. In einer Ausführungsform handelt es sich bei der Nukleotidbase um ein Analogon von Adenin. In einer Ausführungsform handelt es sich bei der Nukleotidbase um ein Analogon von Guanin. In einer Ausführungsform handelt es sich bei der Nukleotidbase um ein Analogon von Cytosin. In einer Ausführungsform handelt es sich bei der Nukleotidbase um ein Analogon von Thymin. In einer Ausführungsform handelt es sich bei der Nukleotidbase um ein Analogon von Uracil.

[0046] In unterschiedlichen Ausführungsformen jeglicher der hierin beschriebenen Erfindungen besteht die feste Oberfläche aus Glas, aus Silizium oder aus Gold. In unterschiedlichen Ausführungsformen handelt es sich bei der festen Oberfläche um ein magnetisches Kügelchen, einen Chip, einen Kanal in einem Chip oder einen porösen Kanal in einem Chip. In einer Ausführungsform handelt es sich bei der festen Oberfläche um Glas. In einer Ausführungsform handelt es sich bei der festen Oberfläche um Silizium. In einer Ausführungsform handelt es sich bei der festen Oberfläche um Gold. In einer Ausführungsform handelt es sich bei der festen Oberfläche um ein magnetisches Kügelchen. In einer Ausführungsform handelt es sich bei der festen Oberfläche um einen Chip. In einer Ausführungsform handelt es sich bei der festen Oberfläche um einen Kanal in einem Chip. In einer Ausführungsform handelt es sich bei der festen Oberfläche um einen porösen Kanal in einem Chip. Es können auch andere Materialien verwendet werden, vorausgesetzt das Material hat keine Auswirkungen auf die Schritte des Verfahrens.

[0047] In einer Ausführungsform umfasst der Schritt zur Befestigung der Nukleinsäure an die feste Oberfläche:

- i. Beschichten der festen Oberfläche mit einem Phosphinrest;
- ii. Befestigen einer Azidogruppe an das 5'-Ende der Nukleinsäure, und
- iii. Immobilisieren des 5'-Endes der Nukleinsäure an die feste Oberfläche mittels Interaktion zwischen dem Phosphinrest auf der festen Oberfläche und der Azidogruppe an dem 5'-Ende der Nukleinsäure.

[0048] In einer Ausführungsform umfasst der Schritt des Beschichtens der festen Oberfläche mit dem Phosphinrest:

- i. Beschichten der Oberfläche mit einem primären Amin, und
- ii. kovalente Kopplung eines N-Hydroxysuccinimidylesters von Triarylphosphin mit dem primären Amin.

[0049] In einer Ausführungsform handelt es sich bei der Nukleinsäure, die an der festen Oberfläche befestigt wird, um eine einzelsträngige Desoxyribonukleinsäure (DNA). In einer anderen Ausführungsform handelt es sich bei der Nukleinsäure, die in Schritt (i) an der festen Oberfläche befestigt wird, um eine doppelsträngige DNA, wobei nur ein Strang direkt an der festen Oberfläche befestigt wird und wobei der Strang, der nicht direkt

an der festen Oberfläche befestigt wird, vor der Durchführung von Schritt (ii) mittels Denaturierung entfernt wird. In einer Ausführungsform handelt es sich bei der Nukleinsäure, die an der festen Oberfläche befestigt wird, um eine Ribonukleinsäure (RNA) und bei der Polymerase in Schritt (iii) um eine reverse Transkriptase.

[0050] In einer Ausführungsform wird der Primer in Schritt (ii) an einem 3'-Ende der Nukleinsäure befestigt und der befestigte Primer umfasst eine stabile Schleife und eine -OH-Gruppe an einer 3'-Position einer Desoxyribose, welche in der Polymerasereaktion zum Selbstpriming in der Lage ist. In einer Ausführungsform umfasst der Schritt der Befestigung des Primers an der Nukleinsäure das Hybridisieren des Primers an die Nukleinsäure oder das Ligieren des Primers an die Nukleinsäure. In einer Ausführungsform wird der Primer mittels einer Ligationsreaktion an der Nukleinsäure befestigt, welche das 3'-Ende der Nukleinsäure mit dem 5'-Ende des Primers verknüpft.

[0051] In einer Ausführungsform werden eines oder mehrere von vier unterschiedlichen Nukleotidanaloga in Schritt (iii) hinzugefügt, wobei jedes unterschiedliche Nukleotidalogon eine unterschiedliche Base umfasst ausgewählt aus der Gruppe bestehend aus Thymin oder Uracil oder einem Analogon von Thymin oder Uracil, Adenin oder einem Analogon von Adenin, Cytosin oder einem Analogon von Cytosin sowie Guanin oder einem Analogon von Guanin, und wobei jedes der vier unterschiedlichen Nukleotidanaloga eine eindeutige Markierung umfasst.

[0052] Bei der spaltbaren chemischen Gruppe, die die -OH-Gruppe an der 3'-Position der Desoxyribose in dem Nukleotidalogon schützt, handelt es sich um $-CH_2OCH_3$ oder um $-CH_2CH=CH_2$. Jede beliebige dieser chemischen Gruppen 1) ist stabil während der Polymerasereaktion, 2) hat keine Auswirkungen auf die Erkennung des Nukleotidalogons als ein Substrat durch Polymerase und 3) ist spaltbar.

[0053] In einer Ausführungsform handelt es sich bei der eindeutigen Markierung, die an dem Nukleotidalogon befestigt ist, um einen fluoreszierenden Rest oder einen fluoreszierenden Halbleiterkristall. In weiteren Ausführungsformen wird der fluoreszierende Rest ausgewählt aus der Gruppe bestehend aus 5-Carboxyfluorescein, 6-Carboxyrhodamin-6G, N,N,N',N'-Tetramethyl-6-carboxyrhodamin und 6-Carboxy-X-rhodamin. In einer Ausführungsform handelt es sich bei dem fluoreszierenden Rest um 5-Carboxyfluorescein. In einer Ausführungsform handelt es sich bei dem fluoreszierenden Rest um 6-Carboxyrhodamin-6G, N,N,N',N'-Tetramethyl-6-carboxyrhodamin. In einer Ausführungsform handelt es sich bei dem fluoreszierenden Rest um 6-Carboxy-X-rhodamin.

[0054] In einer Ausführungsform handelt es sich bei der eindeutigen Markierung, die an dem Nukleotidalogon befestigt ist, um einen Fluoreszenzenergietransfer-Tag, welcher einen Energietransferdonor und einen Energietransferakzeptor umfasst. In weiteren Ausführungsformen handelt es sich bei dem Energietransferdonor um 5-Carboxyfluorescein oder Cyanin, wobei der Energietransferakzeptor ausgewählt ist aus der Gruppe bestehend aus Dichlorcarboxyfluorescein, Dichlor-6-carboxyrhodamin-6G, Dichlor-N,N,N',N'-tetramethyl-6-carboxyrhodamin und Dichlor-6-carboxy-X-rhodamin. In einer Ausführungsform handelt es sich bei dem Energietransferakzeptor um Dichlorcarboxyfluorescein. In einer Ausführungsform handelt es sich bei dem Energietransferakzeptor um Dichlor-6-carboxyrhodamin-6G. In einer Ausführungsform handelt es sich bei dem Energietransferakzeptor um Dichlor-N,N,N',N'-tetramethyl-6-carboxyrhodamin. In einer Ausführungsform handelt es sich bei dem Energietransferakzeptor um Dichlor-6-carboxy-X-rhodamin.

[0055] In einer Ausführungsform handelt es sich bei der eindeutigen Markierung, die an dem Nukleotidalogon befestigt ist, um einen Masse-Tag, der mittels eines Massenspektrometers detektiert und differenziert werden kann. In weiteren Ausführungsformen ist der Masse-Tag ausgewählt aus der Gruppe bestehend aus einer 2-Nitro- α -methyl-benzylgruppe, einer 2-Nitro- α -methyl-3-fluorbenzylgruppe, einer 2-Nitro- α -methyl-3,4-difluorbenzylgruppe und einer 2-Nitro- α -methyl-3,4-dimethoxybenzylgruppe. In einer Ausführungsform handelt es sich bei dem Masse-Tag um eine 2-Nitro- α -methyl-benzylgruppe. In einer Ausführungsform handelt es sich bei dem Masse-Tag um eine 2-Nitro- α -methyl-3-fluorbenzylgruppe. In einer Ausführungsform handelt es sich bei dem Masse-Tag um eine 2-Nitro- α -methyl-3,4-difluorbenzylgruppe. In einer Ausführungsform handelt es sich bei dem Masse-Tag um eine 2-Nitro- α -methyl-3,4-dimethoxybenzylgruppe. In einer Ausführungsform wird der Masse-Tag mittels der Verwendung eines parallelen Massenspektrometrysystems detektiert, welches eine Vielzahl von Massenspektrometern für atmosphärischen Druck und chemische Ionisation zur parallelen Analyse von einer Vielzahl von Masse-Tags umfassenden Proben umfasst.

[0056] In einer Ausführungsform ist die eindeutige Markierung durch einen spaltbaren Linker an der 5'-Position von Cytosin oder Thymin oder an einer 7-Position von Deazaadenin oder Deazaguanin befestigt. Die eindeutige Markierung könnte ebenfalls durch einen spaltbaren Linker an einer anderen Position in dem Nukleo-

tidanalogon befestigt sein, vorausgesetzt, dass die Befestigung der Markierung während der Polymerasereaktion stabil ist und das Nukleotidanalogon durch Polymerase als ein Substrat erkannt werden kann. Die spaltbare Markierung könnte z. B. an der Desoxyribose befestigt werden.

[0057] In einer Ausführungsform wird der Linker zwischen der eindeutigen Markierung und dem Nukleotidanalogon durch ein Mittel abgespalten, welches ausgewählt ist aus der Gruppe bestehend aus einem oder mehreren physikalischen Mitteln, chemischen Mitteln, physiko-chemischen Mitteln, Hitze und Licht. In einer Ausführungsform wird der Linker durch ein physikalisches Mittel abgespalten. In einer Ausführungsform wird der Linker durch ein physiko-chemisches Mittel abgespalten. In einer Ausführungsform wird der Linker durch Hitze abgespalten. In einer Ausführungsform wird der Linker durch Licht abgespalten. In einer Ausführungsform wird der Linker durch ultraviolettes Licht abgespalten. In einer weiteren Ausführungsform handelt es sich bei dem spaltbaren Linker um einen photospaltbaren Linker, der einen 2-Nitrobenzylrest umfasst.

[0058] In einer Ausführungsform wird die spaltbare chemische Gruppe, welche dazu verwendet wird, die -OH-Gruppe an der 3'-Position der Desoxyribose zu schützen, durch ein Mittel abgespalten, welches ausgewählt ist aus der Gruppe bestehend aus einem oder mehreren physikalischen Mitteln, chemischen Mitteln, physiko-chemischen Mitteln, Hitze und Licht. In einer Ausführungsform wird der Linker durch ein physiko-chemisches Mittel abgespalten. In einer Ausführungsform wird der Linker durch Hitze abgespalten. In einer Ausführungsform wird der Linker durch Licht abgespalten. In einer Ausführungsform wird der Linker durch ultraviolettes Licht abgespalten.

[0059] In einer Ausführungsform handelt es sich bei den chemischen Verbindungen, die in Schritt (vi) hinzugefügt werden, um jegliche nicht reagierte -OH-Gruppe auf dem an der Nukleinsäure befestigten Primer oder auf dem Primerextensionsstrang permanent zu schützen, um eine Polymerase oder eines oder mehrere unterschiedliche Didesoxynukleotide oder Analoga von Didesoxynukleotiden. In weiteren Ausführungsformen sind die unterschiedlichen Didesoxynukleotide ausgewählt aus der Gruppe bestehend aus 2',3'-Didesoxyadenosin-5'-triphosphat, 2',3'-Didesoxyguanosin-5'-triphosphat, 2',3'-Didesoxycytidin-5'-triphosphat, 2',3'-Didesoxythymidin-5'-triphosphat, 2',3'-Didesoxyuridin-5'-triphosphat sowie deren Analoga. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um 2',3'-Didesoxyguanosin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um 2',3'-Didesoxycytidin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um 2',3'-Didesoxythymidin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um 2',3'-Didesoxyuridin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um ein Analogon von 2',3'-Didesoxyadenosin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um ein Analogon von 2',3'-Didesoxyguanosin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um ein Analogon von 2',3'-Didesoxycytidin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um ein Analogon von 2',3'-Didesoxythymidin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um ein Analogon von 2',3'-Didesoxyuridin-5'-triphosphat.

[0060] In einer Ausführungsform werden in Schritt (vi) eine Polymerase und eines oder mehrere von vier unterschiedlichen Didesoxynukleotiden hinzugefügt, wobei jedes unterschiedliche Didesoxynukleotid ausgewählt ist aus der Gruppe bestehend aus 2',3'-Didesoxyadenosin-5'-triphosphat oder einem Analogon von 2',3'-Didesoxyadenosin-5'-triphosphat; 2',3'-Didesoxyguanosin-5'-triphosphat oder einem Analogon von 2',3'-Didesoxyguanosin-5'-triphosphat; 2',3'-Didesoxycytidin-5'-triphosphat oder einem Analogon von 2',3'-Didesoxycytidin-5'-triphosphat; 2',3'-Didesoxythymidin-5'-triphosphat oder 2',3'-Didesoxyuridin-5'-triphosphat oder einem Analogon von 2',3'-Didesoxythymidin-5'-triphosphat oder 2',3'-Didesoxyuridin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um 2',3'-Didesoxyadenosin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um ein Analogon von 2',3'-Didesoxyadenosin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um 2',3'-Didesoxyguanosin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um ein Analogon von 2',3'-Didesoxyguanosin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um 2',3'-Didesoxycytidin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um ein Analogon von 2',3'-Didesoxycytidin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um 2',3'-Didesoxythymidin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um 2',3'-Didesoxyuridin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um ein Analogon von 2',3'-Didesoxythymidin-5'-triphosphat. In einer Ausführungsform handelt es sich bei dem Didesoxynukleotid um ein Analogon von 2',3'-Didesoxyuridin-5'-triphosphat.

[0061] Es könnte ebenfalls eine andere Art von chemischer Verbindung, welche spezifisch mit der -OH-Gruppe reagiert, verwendet werden, um eine jegliche nicht reagierte -OH-Gruppe auf dem an der Nukleinsäure be-

festigten Primer oder auf einem Extensionsstrang, welcher durch Hinzufügen eines oder mehrerer Nukleotide oder Nukleotidanaloga zu dem Primer gebildet wurde, permanent zu schützen.

[0062] Die vorliegende Erfindung stellt ein Verfahren zur simultanen Sequenzierung einer Vielzahl unterschiedlicher Nucleinsäuren bereit, welches die gleichzeitige Anwendung jedes beliebigen der hierin offenbarten Verfahren zur Sequenzierung einer Nucleinsäure auf die Vielzahl der unterschiedlichen Nucleinsäuren umfasst. In anderen Ausführungsformen kann das Verfahren dazu verwendet werden, eine bis mehr als 100.000 unterschiedliche Nucleinsäuren gleichzeitig zu sequenzieren.

[0063] Die vorliegende Erfindung stellt die Verwendung eines jeden der hierin offenbarten Verfahren zur Detektion von Single Nucleotide Polymorphismen, zur genetischen Mutationsanalyse, zur seriellen Analyse der Genexpression, zur Genexpressionsanalyse, zur forensischen Identifikation, für Studien zur genetischen Krankheitsassoziation, zur DNA-Sequenzierung, zur genomischen Sequenzierung, zur Translationsanalyse oder zur Transkriptionsanalyse bereit.

[0064] Die vorliegende Erfindung stellt ein Verfahren zur Befestigung einer Nucleinsäure an eine feste Oberfläche bereit, welches umfasst:

- i. Beschichten der festen Oberfläche mit einem Phosphinrest,
- ii. Befestigen einer Azidogruppe an einem 5'-Ende der Nucleinsäure, und
- iii. Immobilisieren des 5'-Endes der Nucleinsäure an der festen Oberfläche mittels Interaktion zwischen dem Phosphinrest auf der festen Oberfläche und der Azidogruppe an dem 5'-Ende der Nucleinsäure.

[0065] In einer Ausführungsform umfasst der Schritt des Beschichtens der festen Oberfläche mit dem Phosphinrest:

- i. Beschichten der Oberfläche mit einem primären Amin, und
- ii. kovalente Kopplung eines N-Hydroxysuccinimidylesters von Triarylphosphin mit dem primären Amin.

[0066] In verschiedenen Ausführungsformen besteht die feste Oberfläche aus Glas, aus Silizium oder aus Gold. In verschiedenen Ausführungsformen handelt es sich bei der festen Oberfläche um ein magnetisches Kügelchen, um einen Chip, um einen Kanal in einem Chip oder um einen porösen Kanal in einem Chip.

[0067] In verschiedenen Ausführungsformen handelt es sich bei der Nucleinsäure, die an der festen Oberfläche befestigt ist, um eine einzel- oder doppelsträngige DNA oder eine RNA. In einer Ausführungsform handelt es sich bei der Nucleinsäure um eine doppelsträngige DNA und nur ein Strang ist an der festen Oberfläche befestigt. In einer weiteren Ausführungsform wird der Strang der doppelsträngigen DNA, der nicht an der festen Oberfläche befestigt ist, mittels Denaturierung entfernt.

[0068] Die vorliegende Erfindung stellt die Verwendung eines jeglichen der hierin offenbarten Verfahren zur Befestigung einer Nucleinsäure an einer Oberfläche zur Genexpressionsanalyse, zur auf Mikroarrays basierenden Genexpressionsanalyse oder Mutationsdetektion, zur Translationsanalyse, zur Transkriptionsanalyse oder für andere genetische Anwendungen bereit.

[0069] Die vorliegende Erfindung stellt ein Nucleotidanalogen bereit, welches umfasst:

- a) eine Base ausgewählt aus der Gruppe bestehend aus Adenin oder einem Analogon von Adenin, Cytosin oder einem Analogon von Cytosin, Guanin oder einem Analogon von Guanin, Thymin oder einem Analogon von Thymin sowie Uracil oder einem Analogon von Uracil;
- b) eine eindeutige Markierung, welche mittels eines spaltbaren Linkers an der Base oder an einem Analogon der Base befestigt ist;
- c) eine Desoxyribose; und
- d) eine spaltbare chemische Gruppe, um eine -OH-Gruppe an einer 3'-Position der Desoxyribose zu schützen, wobei es sich bei der spaltbaren chemischen Gruppe um $-CH_2OCH_3$ oder um $-CH_2CH=CH_2$ handelt.

[0070] In einer Ausführungsform des Nucleotidanalogs handelt es sich bei der spaltbaren chemischen Gruppe, welche die -OH-Gruppe an der 3'-Position der Desoxyribose schützt, um $-CH_2CH=CH_2$.

[0071] In einer Ausführungsform handelt es sich bei der eindeutigen Markierung um einen fluoreszierenden Rest oder einen fluoreszierenden Halbleiterkristall. In weiteren Ausführungsformen ist der fluoreszierende Rest ausgewählt aus der Gruppe bestehend aus 5-Carboxyfluorescein, 6-Carboxyrhodamin-6G, N,N,N',N'-Tetramethyl-6-carboxyrhodamin und 6-Carboxy-X-rhodamin.

[0072] In einer Ausführungsform handelt es sich bei der eindeutigen Markierung um einen Fluoreszenzenergietransfer-Tag, welcher einen Energietransferdonor und einen Energietransferakzeptor umfasst. In weiteren Ausführungsformen handelt es sich bei dem Energietransferdonor um 5-Carboxyfluorescein oder Cyanin, wobei der Energietransferakzeptor ausgewählt ist aus der Gruppe bestehend aus Dichlorcarboxyfluorescein, Dichlor-6-carboxyrhodamin-6G, Dichlor-N,N,N',N'-tetramethyl-6-Carboxyrhodamin und Dichlor-6-carboxy-X-rhodamin.

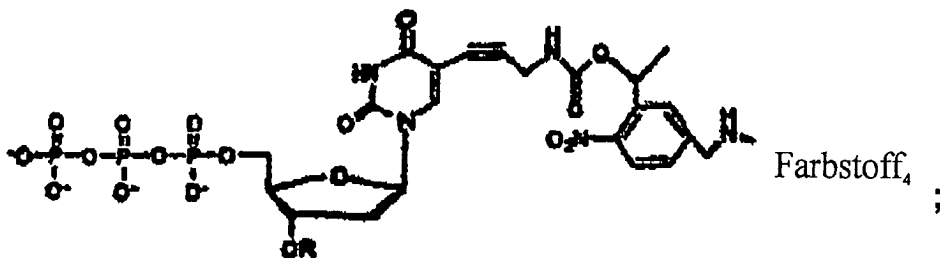
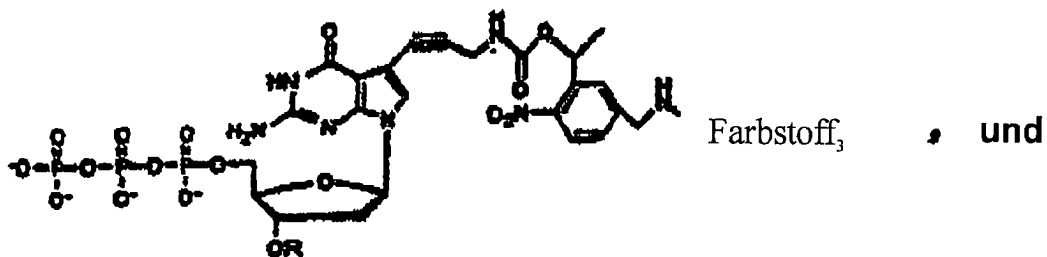
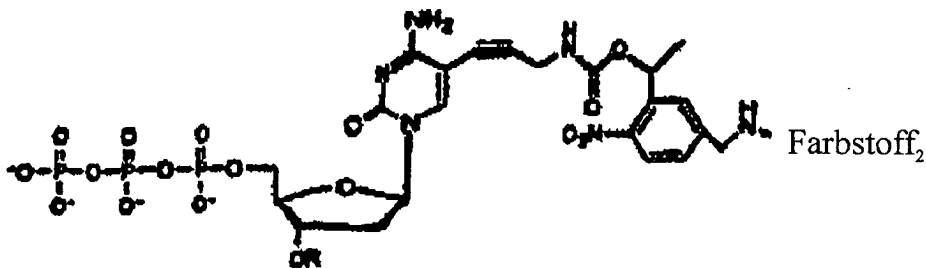
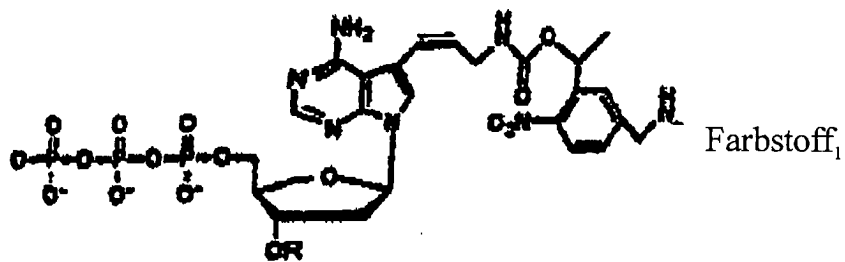
[0073] In einer Ausführungsform handelt es sich bei der eindeutigen Markierung um einen Masse-Tag, der mittels eines Massenspektrometers detektiert und differenziert werden kann. In weiteren Ausführungsformen ist der Masse-Tag ausgewählt aus der Gruppe bestehend aus einer 2-Nitro- α -methyl-benzylgruppe, einer 2-Nitro- α -methyl-3-Fluorbenzylgruppe, einer 2-Nitro- α -methyl-3,4-Difluorbenzylgruppe und einer 2-Nitro- α -methyl-3,4-Dimethoxybenzylgruppe.

[0074] In einer Ausführungsform ist die eindeutige Markierung durch einen spaltbaren Linker an einer 5-Position von Cytosin oder Thymin oder an einer 7-Position von Deazaadenin oder Deazaguanin befestigt. Die eindeutige Markierung könnte ebenfalls durch einen spaltbaren Linker an einer anderen Position in dem Nukleotidanalogen befestigt sein, vorausgesetzt dass die Befestigung der Markierung während der Polymerasereaktion stabil ist und dass das Nukleotidanalogen durch Polymerase als ein Substrat erkannt werden kann. Der spaltbare Linker könnte z. B. an der Desoxyribose befestigt sein.

[0075] In einer Ausführungsform ist der Linker zwischen der eindeutigen Markierung und dem Nukleotidanalogen spaltbar durch ein Mittel ausgewählt aus der Gruppe bestehend aus einem oder mehreren physikalischen Mitteln, chemischen Mitteln, physiko-chemischen Mitteln, Hitze und Licht. In einer weiteren Ausführungsform handelt es sich bei dem spaltbaren Linker um einen photospaltbaren Linker, welcher einen 2-Nitrobenzylrest umfasst.

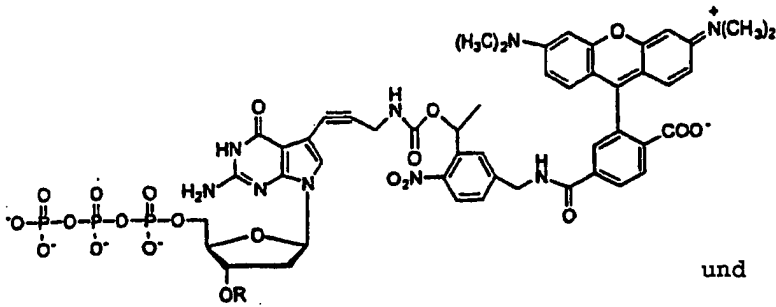
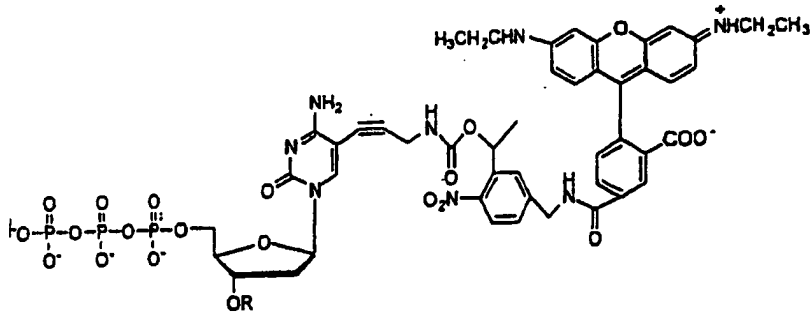
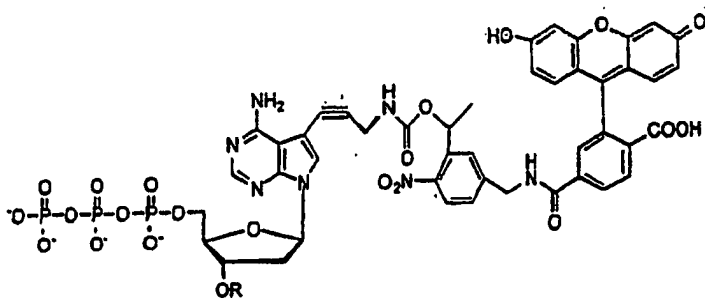
[0076] In einer Ausführungsform ist die spaltbare chemische Gruppe, welche dazu verwendet wird, die -OH-Gruppe an der 3'-Position der Desoxyribose zu schützen, spaltbar durch ein Mittel ausgewählt aus der Gruppe bestehend aus einem oder mehreren physikalischen Mitteln, chemischen Mitteln, physiko-chemischen Mitteln, Hitze und Licht.

[0077] In verschiedenen Ausführungsformen ist das Nukleotidanalogen ausgewählt aus der Gruppe bestehend aus:

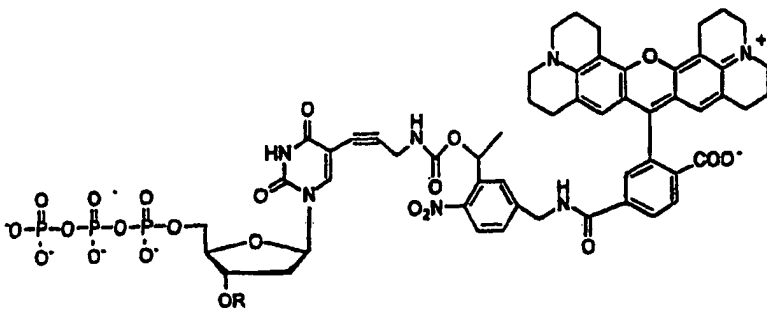


wobei Farbstoff₁, Farbstoff₂, Farbstoff₃ und Farbstoff₄ vier unterschiedliche eindeutige Markierungen sind; und wobei R -CH₂OCH₃ oder -CH₂CH=CH₂ ist.

[0078] In verschiedenen Ausführungsformen ist das Nucleotidanalogen ausgewählt aus der Gruppe bestehend aus:

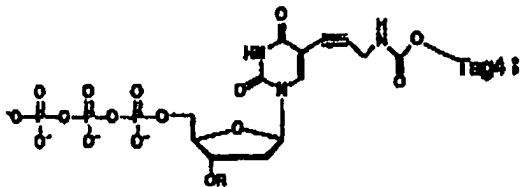
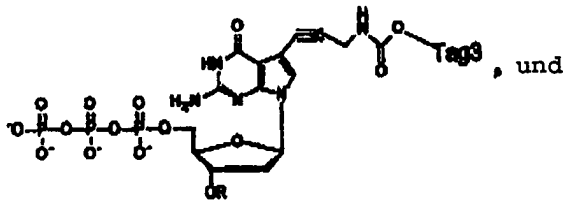
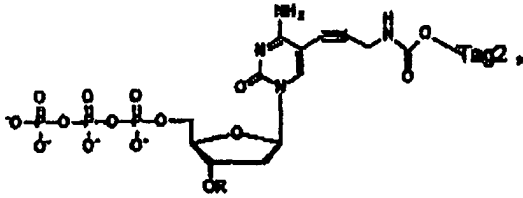
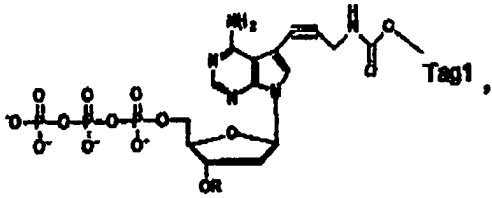


und



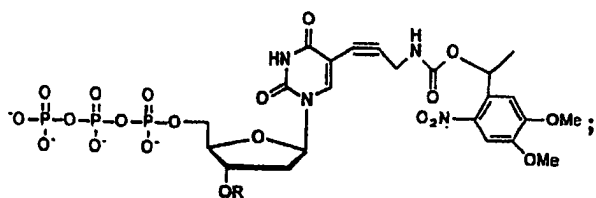
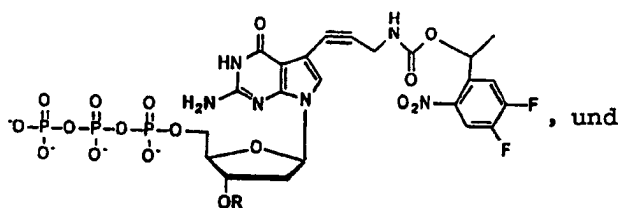
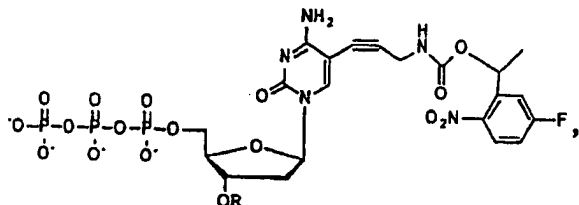
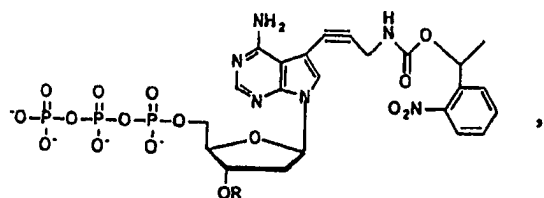
wobei R -CH₂OCH₃ oder -CH₂CH=CH₂ ist.

[0079] In verschiedenen Ausführungsformen ist das Nucleotidanalogen ausgewählt aus der Gruppe bestehend aus:



wobei Tag₁, Tag₂, Tag₃ und Tag₄ vier unterschiedliche Masse-Tag-Markierungen sind; und wobei R -CH₂OCH₃ oder -CH₂CH=CH₂ ist.

[0080] In verschiedenen Ausführungsformen ist das Nucleotidanalogen ausgewählt aus der Gruppe bestehend aus:



wobei R -CH₂OCH₃ oder -CH₂CH=CH₂ ist.

[0081] Die vorliegende Erfindung stellt die Verwendung eines jeglichen der hierin offenbarten Nukleotidanaloga zur Detektion von Single Nucleotide Polymorphismen, zur genetischen Mutationsanalyse, zur seriellen Analyse der Genexpression, zur Genexpressionsanalyse, zur forensischen Identifikation, für Studien zur genetischen Krankheitsassoziation, zur DNA-Sequenzierung, zur genomischen Sequenzierung, zur Translationsanalyse oder zur Transkriptionsanalyse bereit.

Experimentelle Details

1. Der Ansatz der Sequenzierung durch Synthese

[0082] Die Sequenzierung von DNA durch Synthese beinhaltet die Detektion der Identität eines jeden Nukleotids, während es in der Polymerasereaktion in den wachsenden DNA-Strang integriert wird. Die grundlegenden Anforderungen für das Funktionieren eines solchen Systems sind: (1) die Verfügbarkeit von 4 Nukleotidanaloga (aA, aC, aG, aT), von denen jedes mit einer eindeutigen Markierung versehen ist und einen die 3'-OH-Gruppe schützenden chemischen Rest enthält; (2) die 4 Nukleotidanaloga (aA, aC, aG, aT) müssen in der Polymerasereaktion effizient und zuverlässig von DNA-Polymerase als Terminatoren integriert werden; (3) der Tag und die die 3'-OH-Gruppe schützende Gruppe müssen mit hohem Ertrag entfernt werden, um die Integration und Detektion des nächsten Nukleotids zu gestatten; und (4) der wachsende DNA-Strang sollte das Waschen, die Detektion und die Abspaltungsprozesse überstehen, um an das DNA-Template annealed zu bleiben.

[0083] Der hierin offenbarte Ansatz der Sequenzierung durch Synthese ist in **Fig. 2A-2B** dargestellt. In **Fig. 2A** ist ein Beispiel gezeigt, in dem die eindeutigen Markierungen fluoreszierende Farbstoffe sind und die Oberfläche ein Chip ist; in **Fig. 2B** handelt es sich bei den eindeutigen Markierungen um Masse-Tags und bei der Oberfläche um in einen Chip eingeztete Kanäle. Der Synthese-Ansatz verwendet eine feste Oberfläche beispielsweise wie einen Glaschip mit einem immobilisierten DNA-Template, welches zum Selbstpriming in der Lage ist, um die Polymerasereaktion zu initiieren, sowie vier Nukleotidanaloga (3'-RO-A⁻LABEL1, 3'-RO-C⁻LABEL2,

$3\text{'-RO-G}_{\text{-LABEL3}}$, $3\text{'-RO-T}_{\text{-LABEL4}}$), von denen jedes mit einer eindeutigen Markierung versehen ist, z. B. einem fluoreszierenden Farbstoff oder einem Masse-Tag, an einer bestimmten Position auf der Purin- oder der Pyrimidinbase, sowie eine kleine spaltbare chemische Gruppe (R), um die 3'-OH-Gruppe zu schützen. Bei der Zugabe der vier Nucleotidanaloga und der DNA-Polymerase wird auf jedem Spot auf der Oberfläche nur ein Nucleotid analogon, das zu dem nächsten Nucleotid auf dem Template komplementär ist, von der Polymerase integriert (Schritt 1 in [Fig. 2A](#) und [Fig. 2B](#)).

[0084] Wie in [Fig. 2A](#) gezeigt ist, worin es sich bei den eindeutigen Markierungen um Farbstoffe handelt, wird, nachdem die überschüssigen Reagenzien entfernt und jegliche nicht integrierte Nucleotidanaloga auf dem Chip gewaschen wurden, ein Detektor verwendet, um die eindeutige Markierung zu detektieren. Es wird z. B. ein Vierfarben-Fluoreszenz-Imager verwendet, um die Oberfläche des Chips abzubilden, und die eindeutige Fluoreszenzemission eines spezifischen Farbstoffs auf den Nucleotidanaloga auf jedem Spot des Chips offenbart die Identität des integrierten Nucleotids (Schritt 2 in [Fig. 2A](#)). Nach dem Imaging wird die kleine Menge an nicht reagierter 3'-OH-Gruppe auf dem selbstgeprimten Templaterest durch überschüssige Didesoxynucleosidtriphosphate (ddNTPs) (ddATP, ddGTP, ddTTP und ddCTP) und DNA-Polymerase geschützt, um eine Interferenz mit dem nächsten Durchgang der Synthese zu vermeiden (Schritt 3 in [Fig. 2A](#)) – ein Konzept, das dem schützenden Schritt bei der automatischen Festphasen DNA-Synthese ähnlich ist (Caruthers, 1985). Die ddNTPs, welchen eine 3'-Hydroxylgruppe fehlt, werden aufgrund ihrer im Vergleich zu den mit Farbstoff markierten Nucleotiden geringen Größe und der ausgezeichneten Effizienz, mit der sie von der DNA-Polymerase integriert werden, ausgewählt, um das nicht reagierte 3'-OH des Nucleotids zu schützen. Der Farbstoffrest wird dann durch Licht abgespalten (~350 nm) und die das 3'-OH schützende R-Gruppe wird chemisch entfernt, um mit hohem Ertrag freie 3'-OH-Gruppe zu erzeugen (Schritt 4 in [Fig. 2A](#)). Es wird ein Waschschrift durchgeführt, um die abgespaltenen Farbstoffe und die R-Gruppe wegzuwaschen. Der selbstgeprimte DNA-Rest auf dem Chip ist in dieser Phase bereit für den nächsten Reaktionszyklus zur Identifikation der nächsten Nucleotidsequenz der Template-DNA (Schritt 5 in [Fig. 2A](#)).

[0085] Es ist nunmehr eine routinemäßige Vorgehensweise, einzelsträngige DNA in hoher Dichte (> 10.000 Spots pro Chip) auf einem 4 cm × 1 cm Glaschip zu immobilisieren (Skena et al., 1995). Folglich können in dem hierin offenbarten DNA-Sequenzierungssystem mehr als 10.000 Basen nach jedem Zyklus identifiziert werden, und nach 100 Zyklen sind eine Million Basenpaare aus einem Sequenzierungschip erzeugt worden.

[0086] Geeignete DNA-Polymerasen schließen ein Thermo Sequenase, Taq FS DNA-Polymerase, T7 DNA-Polymerase sowie Vent (exo-) DNA-Polymerase. Die Fluoreszenzemission von jedem spezifischen Farbstoff kann unter Verwendung eines Fluorimeters detektiert werden, das mit einer Vorrichtung zur Fluoreszenzdetektion von einem Glasobjektträger ausgestattet ist. Zur Bewertung im großen Maßstab kann ein Vielfarben-Scanningsystem verwendet werden, welches dazu in der Lage ist, eine Vielzahl von verschiedenen fluoreszierenden Farbstoffen (500 nm–700 nm) auf einem Glasobjektträger zu detektieren (GSI Lumonics ScanArray 5000 Standard Biochip Scanning System).

[0087] Ein Beispiel für den Ansatz der Sequenzierung durch Synthese unter Verwendung von Masse-Tags ist in [Fig. 2B](#) gezeigt. Der Ansatz verwendet eine feste Oberfläche, so wie poröse Kanäle aus Silica-Glas in einem Chip, mit immobilisiertem DNA-Template, welches zur Initiation der Polymerasereaktion zum Selbstpriming in der Lage ist, sowie vier Nucleotidanaloga ($3\text{'-RO-A}_{\text{-Tag1}}$, $3\text{'-RO-C}_{\text{-Tag2}}$, $3\text{'-RO-G}_{\text{-Tag3}}$, $3\text{'-RO-T}_{\text{-Tag4}}$), von denen jedes an der spezifischen Position der Base mit einem eindeutigen photospaltbaren Masse-Tag markiert ist und eine kleine spaltbare chemische Gruppe (R), um die 3'-OH-Gruppe zu schützen. Bei der Zugabe der vier Nucleotidanaloga und der DNA-Polymerase wird in jedem Kanal des Glaschips nur ein Nucleotid analogon, das zu dem nächsten Nucleotid auf dem Template komplementär ist, von der Polymerase integriert (Schritt 1 in [Fig. 2B](#)). Nachdem die überschüssigen Reagenzien entfernt und jegliche nicht integrierten Nucleotidanaloga auf dem Chip gewaschen wurden, wird die kleine Menge an nicht reagierter 3'-OH-Gruppe auf dem selbstgeprimten Templaterest durch überschüssige ddNTPs (ddATP, ddGTP, ddTTP und ddCTP) und DNA-Polymerase geschützt, um eine Interferenz mit dem nächsten Durchgang der Synthese zu vermeiden (Schritt 2 in [Fig. 2B](#)). Die ddNTPs werden aufgrund ihrer im Vergleich zu den markierten Nucleotiden geringen Größe und der ausgezeichneten Effizienz, mit der sie von der DNA-Polymerase integriert werden, ausgewählt, um das nicht reagierte 3'-OH des Nucleotids zu schützen. Die Masse-Tags werden durch Bestrahlung mit Licht (~350 nm) abgespalten (Schritt 3 in [Fig. 2B](#)) und dann mit einem Massenspektrometer detektiert. Die eindeutige Masse eines jeden Tags ergibt die Identität des Nucleotids in jedem Kanal (Schritt 4 in [Fig. 2B](#)). Die R-schützende Gruppe wird dann chemisch entfernt und gewaschen, um mit hohem Ertrag freie 3'-OH-Gruppen zu erzeugen (Schritt 5 in [Fig. 2B](#)). Der selbstgeprimte DNA-Rest auf dem Chip ist in dieser Phase bereit für den nächsten Reaktionszyklus, um die nächste Nucleotidsequenz der Template-DNA zu identifizieren (Schritt 6 in [Fig. 2B](#)).

[0088] Seit der Entwicklung neuer Ionisationstechniken, so wie matrixgestützte Laser-Desorptions-Ionisation (MALDI) und Elektrospray-Ionisation (ESI), wurde die Massenspektrometrie zu einem unverzichtbaren Werkzeug in vielen Gebieten der biomedizinischen Forschung. Wenngleich diese Ionisationsverfahren zur Analyse bioorganischer Moleküle, so wie Peptide und Proteine, geeignet sind, so sind doch Verbesserungen sowohl bei der Detektion als auch bei der Herstellung der Proben für den Einsatz der Massenspektrometrie in DNA-Sequenzierungsanwendungen erforderlich. Da der hierin offenbarte Ansatz kleine und stabile Masse-Tags verwendet, besteht kein Bedarf, große DNA-Sequenzierungsfragmente direkt zu detektieren, und es ist nicht notwendig, MALDI- oder ESI-Verfahren zur Detektion zu verwenden. Bei der Atmospheric Pressure Chemical Ionization (APCI) handelt es sich um ein Ionisationsverfahren, das sich bei atmosphärischem Druck einer Gasphasen-Ion-Molekül-Reaktion bedient (Dizidic et al., 1975). Bei diesem Verfahren werden Proben entweder mittels chromatographischer oder Flussinjektion in einen pneumatischen Vernebler eingebracht, wo sie durch einen Hochgeschwindigkeitsstrahl von Stickstoffgas in kleine Tröpfchen umgewandelt werden. Gelangen das erhitzte Gas und die Lösung in den Reaktionsbereich, so wird die überschüssige Menge an Lösungsmittel mittels Corona-Entladung ionisiert. Diese ionisierte mobile Phase fungiert gegenüber den Proben als das ionisierende Agens und ergibt pseudomolekulare $(M+H)^+$ und $(M-H)^-$ -Ionen. Aufgrund des Corona-Entladungs-Ionisationsverfahrens kann eine hohe Ionisationseffizienz erreicht werden, wobei stabile Ionisationsbedingungen mit einer Detektionssensitivität unterhalb des Femtomolbereichs für kleine und stabile organische Verbindungen aufrecht erhalten werden. Aufgrund der eingeschränkten Detektion von großen Molekülen wurde APCI jedoch zur Analyse von Peptiden und Nukleinsäuren durch ESI und MALDI ersetzt. Da es sich in dem offenbarten Ansatz bei den zu detektierenden Masse-Tags um relativ kleine und sehr stabile organische Moleküle handelt, ist die durch die Verwendung von ESI und MALDI gewonnene Fähigkeit, große biologische Moleküle zu detektieren, nicht vonnöten. APCI verfügt gegenüber ESI und MALDI über einige Vorteile, da es keine mühsame Probenvorbereitung erfordert, so wie Entsalzen oder Vermischen mit Matrix, um Kristalle auf einer Targetplatte zu präparieren. Bei ESI beeinträchtigen die Beschaffenheit der Proben und die Bedingungen zur Vorbereitung von Proben (d.h. die Anwesenheit von Puffer oder anorganischen Salzen) die Ionisationseffizienz. MALDI erfordert die Zugabe von Matrix vor der Einführung der Proben in das Massenspektrometer, und seine Geschwindigkeit wird oftmals dadurch eingeschränkt, dass erst nach einem idealen Bestrahlungsspot gesucht werden muss, um interpretierbare Massenspektren zu erhalten. Diese Einschränkungen werden durch APCI überwunden, da die Lösung mit den Masse-Tags direkt ohne zusätzliche Probenaufreinigung oder Vorbereitung in das Massenspektrometer injiziert werden kann. Da die mit Masse-Tags versehenen Proben volatil sind und kleine Massenzahlen aufweisen, können diese Verbindungen leicht und mit einer hohen Sensitivität mittels APCI-Ionisation detektiert werden. Dieses System kann für den Hochdurchsatzbetrieb aufgerüstet werden.

[0089] Jede Komponente des Systems der Sequenzierung durch Synthese wird nachfolgend detaillierter beschrieben.

2. Konstruktion einer Oberfläche enthaltend immobilisierte selbstgeprimte DNA-Reste

[0090] Das auf einer Oberfläche immobilisierte einzelsträngige DNA-Template wird gemäß dem in [Fig. 3](#) gezeigten Schema präpariert. Bei der Oberfläche kann es sich z. B. um einen Glaschip, so wie einen 4 cm × 1 cm Glaschip, oder um Kanäle in einem Glaschip handeln. Die Oberfläche wird zunächst mit 0,5 M NaOH behandelt, mit Wasser gewaschen und dann unter Ausbildung einer primären Aminoberfläche in einer hohen Dichte mit 3-Aminopropyltrimethoxysilan in wässrigem Ethanol beschichtet (Woolley et al., 1994). N-Hydroxysuccinimidyl (NHS)-Ester von Triarylphosphin (1) wird kovalent mit der primären Aminogruppe gekoppelt, wobei die Aminoberfläche in eine neue Triarylphosphinoberfläche umgewandelt wird, welche spezifisch mit DNA enthaltend eine Azidogruppe (2) reagiert, wobei ein Chip mit immobilisierter DNA gebildet wird. Da sich die Azidogruppe nur am 5'-Ende der DNA befindet und die Kopplungsreaktion durch die eindeutige Reaktion des Triarylphosphinrestes mit der Azidogruppe in wässriger Lösung stattfindet (Saxon und Bertozzi, 2000), wird eine solche DNA-Oberfläche optimale Hybridisierungsbedingungen bereitstellen.

[0091] Der NHS-Ester von Triarylphosphin (1) wird gemäß dem in [Fig. 4](#) gezeigten Schema präpariert. 3-Di-phenylphosphino-4-methoxycarbonyl-Benzoesäure (3) wird gemäß dem von Bertozzi et al. beschriebenen Verfahren (Saxon und Bertozzi, 2000) präpariert. Die Behandlung von (3) mit N-Hydroxysuccinimid bildet den entsprechenden NHS-Ester (4). Die Kopplung von (4) mit einem Aminocarboxylsäurerest ergibt eine Verbindung (5) mit einem langen Linker ($n = 1$ bis 10) zur optimierten Kopplung mit DNA auf der Oberfläche. Die Behandlung von (5) mit N-Hydroxysuccinimid erzeugt den NHS-Ester (1), der nun bereit ist für die Kopplung mit der mit primärem Amin beschichteten Oberfläche ([Fig. 3](#)).

[0092] Die mit Azido markierte DNA (2) wird gemäß dem in [Fig. 5](#) gezeigten Schema synthetisiert. Die Behandlung von Ethylester von 5-Bromvaleriansäure mit Natriumazid und nachfolgende Hydrolyse ergibt 5-Azi-

dovaleriansäure (Khoukhi et al., 1987), welche anschließend zur Kopplung mit einem durch den Aminolinker modifizierten Oligonukleotidprimer in einen NHS-Ester umgewandelt wird. Die Verwendung des azidomarkierten Primers zur Durchführung der Polymerase-Kettenreaktion (PCR) erzeugt ein azidomarkiertes DNA-Templat (2) zur Kopplung mit der durch das Triarylphosphin modifizierten Oberfläche ([Fig. 3](#)).

[0093] Der selbstgeprimte DNA-Templaterest auf dem Sequenzierungschip wird unter Verwendung enzymatischer Ligation wie in [Fig. 6 \(A & B\)](#) gezeigt konstruiert. Es wird ein Oligonukleotidprimer (B), mit einer 5'-phosphorylierten, 3'-OH-Gruppen-geschützten Schleife mittels eines Festphasen-DNA-Synthesizers synthetisiert. Der Primer (B) wird synthetisiert unter Verwendung eines modifizierten C-Phosphoramidits, dessen 3'-OH am 3'-Ende des Oligonukleotids entweder durch eine MOM (-CH₂OCH₃)- oder eine Allyl (-CH₂CH=CH₂)-Gruppe (in [Fig. 6](#) mit "R" gekennzeichnet) geschützt ist, um die Selbstligation des Primers in der Ligationsreaktion zu verhindern. Der mit einer Schleife versehene Primer kann daher nur an das 3'-Ende der DNA-Templates ligieren, die unter Verwendung von T4 RNA-Ligase (Zhang et al., 1996) auf dem Sequenzierungschip immobilisiert wurden, um den selbstgeprimten DNA-Templaterest (A) zu bilden. Der mit einer Schleife versehene Primer (B) ist so konstruiert, dass er eine äußerst stabile Schleife enthält (Antao et al., 1991) sowie einen Stamm mit der Sequenz von M13 reversem DNA Sequenzierungsprimer zum effizienten Priming in der Polymerasereaktion, wenn der Primer erst einmal an die immobilisierte DNA auf dem Sequenzierungschip ligiert und die 3'-OH-Schutzgruppe chemisch abgespalten wurde (Ireland et al., 1986; Kamal et al., 1999).

3. Sequenzierung mittels Synthesebewertung unter Verwendung der Nukleotidanaloga ${}^{3\text{'-HO}}\text{-A-}_{\text{Farbstoff1}}$, ${}^{3\text{'-HO}}\text{-C-}_{\text{Farbstoff2}}$, ${}^{3\text{'-HO}}\text{-G-}_{\text{Farbstoff3}}$, ${}^{3\text{'-HO}}\text{-T-}_{\text{Farbstoff4}}$

[0094] Es wurde ein Schema zur Bewertung der Photospaltungseffizienz unter Verwendung unterschiedlicher Farbstoffe und Testens des Ansatzes der Sequenzierung durch Synthese entwickelt. Es werden vier Nukleotidanaloga ${}^{3\text{'-HO}}\text{-A-}_{\text{Farbstoff1}}$, ${}^{3\text{'-HO}}\text{-C-}_{\text{Farbstoff2}}$, ${}^{3\text{'-HO}}\text{-G-}_{\text{Farbstoff3}}$, ${}^{3\text{'-HO}}\text{-T-}_{\text{Farbstoff4}}$, von denen jedes durch einen photospaltbaren Linker mit einem eindeutigen fluoreszierenden Farbstoff markiert ist, synthetisiert und in dem Ansatz der Sequenzierung durch Synthese verwendet. Beispiele für Farbstoffe schließen ein, sind jedoch nicht beschränkt auf: Farbstoff1 = FAM, 5-Carboxyfluorescein; Farbstoff2 = R6G, 6-Carboxyrhodamin-6G; Farbstoff3 = TAM, N,N,N',N'-Tetramethyl-6-carboxyrhodamin; und Farbstoff4 = ROX, 6-Carboxy-X-rhodamin. Die Strukturen der vier Nukleotidanaloga sind in [Fig. 7](#) gezeigt (R=h).

[0095] Der photospaltbare 2-Nitrobenzylrest wurde verwendet, um Biotin mit DNA und Protein zur effizienten Entfernung mittels UV-Licht (~350 nm) zu verknüpfen (Olejnik et al., 1995, 1999). In dem hierin offenbarten Ansatz, wird die 2-Nitrobenzylgruppe dazu verwendet, den fluoreszierenden Farbstoff und das Nukleotid miteinander zu verbinden, um die mit Farbstoff markierten Nukleotide zu bilden, wie sie in [Fig. 7](#) gezeigt sind.

[0096] In [Fig. 8](#) ist die Synthese von ${}^{3\text{'-HO}}\text{-G-}_{\text{Farbstoff3}}$ (Farbstoff3 = Tam) als ein repräsentatives Beispiel gezeigt. Es wird 7-Deaza-alkynylamino-dGTP unter Verwendung wohl etablierter Verfahren (Prober et al., 1987; Lee et al., 1992, und Hobbs et al., 1991) präpariert. Linker-Tam wird synthetisiert, indem man den Photospaltbaren Linker (Rollaf, 1982) mit NHS-Tam koppelt. 7-Deaza-alkynylamino-dGTP wird dann mit dem Linker-Tam gekoppelt, um ${}^{3\text{'-HO}}\text{-G-}_{\text{TAM3}}$ zu produzieren. Die Nukleotidanaloga mit einem freien 3'-OH (d.h. R=H) sind gute Substrate für die Polymerase. Es wird ein immobilisiertes DNA-Template synthetisiert ([Fig. 9](#)), das einen Abschnitt der Nukleotidsequenz ACGTACGACGT (SEQ ID NR: 1) enthält, die nach der Primingstelle keine wiederholten Sequenzen mehr aufweist. Es werden ${}^{3\text{'-HO}}\text{-A-}_{\text{Farbstoff1}}$ und DNA-Polymerase zu dem selbstgeprimten DNA-Rest hinzugefügt und dieser an der 3'-Stelle der DNA integriert. Es werden dann die Schritte aus [Fig. 2A](#) befolgt (der chemische Abspaltungsschritt ist hier nicht erforderlich, da das 3'-OH frei ist), um das Fluoreszenzsignal von Farbstoff-1 bei 520 nm zu detektieren. Als nächstes wird ${}^{3\text{'-HO}}\text{-C-}_{\text{Farbstoff2}}$ zugegeben, um das Fluoreszenzsignal von Farbstoff-2 bei 550 nm abzubilden. Als nächstes wird ${}^{3\text{'-HO}}\text{-G-}_{\text{Farbstoff3}}$ zugegeben, um das Fluoreszenzsignal von Farbstoff-3 bei 580 nm abzubilden, und schließlich wird ${}^{3\text{'-HO}}\text{-T-}_{\text{Farbstoff4}}$ zugegeben, um das Fluoreszenzsignal von Farbstoff-4 bei 610 nm abzubilden.

Ergebnisse bezüglich der Effizienz der photochemischen Abspaltung

[0097] Die erwarteten Photolyseprodukte von DNA, welche an ihrem 3'-Ende einen photospaltbaren fluoreszierenden Farbstoff enthält, sind in [Fig. 10](#) gezeigt. Der 2-Nitrobenzylrest wurde in einer Vielzahl von Studien erfolgreich als eine photospaltbare schützende Gruppe verwendet (Pillai, 1980). Die Effizienz des Photospaltungsschritts hängt von mehreren Faktoren ab, einschließlich der Effizienz der Lichtabsorption durch den 2-Nitrobenzylrest, die Effizienz des primären photochemischen Schritts und die Effizienz der sekundären thermischen Prozesse, welche zu dem abschließenden Abspaltungsprozess führen (Turro, 1991). Burgess et al. (1997) berichteten von der erfolgreichen Photospaltung eines durch einen 2-Nitrobenzyl-Linker auf einem Nu-

kleotidrest befestigten fluoreszierenden Farbstoffs, was zeigt, dass der fluoreszierende Farbstoff den Photospaltungsprozess nicht quencht. Es wurde ebenfalls eine auf dem 2-Nitrobenzyl-Chromophor basierende photolabile Schutzgruppe für biologische Markierungsanwendungen entwickelt, die Photospaltung beinhalten (Olejnik et al., 1999). Das hierin offenbarte Protokoll wird dazu verwendet, den in [Fig. 10](#) gezeigten Photospaltungsprozess zu optimieren. Die Absorptionsspektren von 2-Nitrobenzyl-Verbindungen werden untersucht und quantitativ mit den Absorptionsspektren der fluoreszierenden Farbstoffe verglichen. Da ein 1:1-Verhältnis der Anzahl von 2-Nitrobenzylresten und Farbstoffmolekülen vorliegen wird, wird das Verhältnis der Extinktionskoeffizienten dieser beiden Spezies die Konkurrenz um die Lichtabsorption bei spezifischen Wellenlängen reflektieren. Auf der Basis dieser Informationen können die Wellenlängen bestimmt werden, bei denen es bei der Absorption der 2-Nitrobenzylreste zu der stärksten Konkurrenz kam, ähnlich wie in dem von Olejnik et al. (1995) berichteten Ansatz.

[0098] Es kann ein Photolyseaufbau, der einen hohen Durchsatz an monochromatischem Licht aus einer 1.000 Watt-Hochdruck-Xenonlampe (LX1000UV, ILC) gestattet, in Verbindung mit einem Monochromator (Kratos, Schoeffel Instruments) verwendet werden. Dieses Instrument ermöglicht die Bewertung der Photospaltung von Modellsystemen als eine Funktion der Intensität und der Anregungswellenlänge des absorbierten Lichts. Zur Bestimmung des Ausmaßes der Photospaltung wird eine standardanalytische Analyse verwendet. Auf der Basis dieser Informationen kann die Effizienz der Photospaltung als eine Funktion der Wellenlänge bestimmt werden. Die Wellenlänge, bei der die Photospaltung am wirkungsvollsten ist, kann zur Verwendung in den Sequenzierungssystemen ausgewählt werden.

[0099] Es wurden unter Verwendung eines Modellsystems, wie es in [Fig. 11](#) gezeigt ist, Photospaltungsergebnisse erhalten. Die Kopplung von PC-LC-Biotin-NHS-Ester (Pierce, Rockford, IL, USA) mit 5-(Aminoacetamido)-Fluorescein (5-Amino-FAM) (Molecular Probes, Eugene, OR, USA) in Dimethylsulfonyloxid (DM-SO)/NaHCO₃ (pH = 8,2) über Nacht bei Raumtemperatur ergibt PC-LC-Biotin-FAM, welches sich aus einem Biotin an einem Ende, einer photospaltbaren 2-Nitrobenzylgruppe in der Mitte und einem Farbstoff-Tag (FAM) am anderen Ende zusammensetzt. Dieser photospaltbare Rest ahmt die in [Fig. 10](#) gezeigten konstruierten photospaltbaren Reste genau nach. Daher belegt die erfolgreiche Photolyse des PC-LC-Biotin-FAM-Restes das Prinzip der Hochleistungsphotolyse, so wie in dem DNA-Sequenzierungssystem verwendet. Zur Untersuchung der Photolyse wird PC-LC-Biotin-FAM zunächst auf einem mit Streptavidin beschichteten Mikroskopobjektträger aus Glas (XENOPORE, Hawthorne, NJ, USA) immobilisiert. Nach dem Abwaschen des nicht immobilisierten PC-LC-Biotin-FAM wurde das Fluoreszenzemissionsspektrum des immobilisierten PC-LC-Biotin-FAM wie in [Fig. 12](#) (Spektrum a) gezeigt aufgenommen. Die starke Fluoreszenzemission zeigt an, dass PC-LC-Biotin-FAM erfolgreich an der mit Streptavidin beschichteten Objektträgeroberfläche immobilisiert wurde. Es wurde dann die Photospaltbarkeit des 2-Nitrobenzyl-Linkers mittels Bestrahlung bei 350 nm getestet. Nach 10 Minuten Photolyse ($\lambda_{\text{irr}} = 350 \text{ nm}$; $\sim 0,5 \text{ mW/cm}^2$) und vor jeglichem Waschen wurde das Fluoreszenzemissionsspektrum des gleichen Spots auf dem Objektträger aufgenommen, welches keine Abnahme der Intensität zeigte ([Fig. 12](#), Spektrum b), was darauf hindeutete, dass der Farbstoff (FAM) während des Photolysevorgangs bei 350 nm nicht gebleicht wurde. Nach dem Waschen des Glasobjektträgers mit HPLC-Wasser im Anschluss an die Photolyse zeigte das Fluoreszenzemissionsspektrum des gleichen Spots auf dem Objektträger eine signifikante Abnahme der Intensität ([Fig. 12](#), Spektrum c), was darauf hindeutet, dass ein Großteil des Fluoreszenzfarbstoffs (FAM) von dem immobilisierten Biotinrest abgespalten und durch den Waschvorgang entfernt worden war. Dieses Experiment zeigt, dass die Abspaltung des Fluoreszenzfarbstoffs mit hoher Effizienz durch die Verwendung des photospaltbaren 2-Nitrobenzyl-Linkers erreicht werden kann.

4. Bewertung der Sequenzierung durch Synthese unter Verwendung der Nukleotidanaloga $3'\text{-RO-A-Farbstoff1}$,



[0100] Wurden die Schritte und Bedingungen in Abschnitt 3 erst einmal optimiert, so kann zur weiteren Untersuchung des Systems mit der Synthese der Nukleotidanaloga $3'\text{-RO-A-Farbstoff1}$, $3'\text{-RO-C-Farbstoff2}$, $3'\text{-RO-G-Farbstoff3}$, $3'\text{-RO-T-Farbstoff4}$ fortgefahren werden. Das 3'-OH liegt hier geschützt in allen vier Nukleotidanaloga vor, welche dann mit der DNA-Polymerase zusammengemischt und zur Bewertung des Sequenzierungssystems unter Verwendung des Schemas in [Fig. 9](#) verwendet werden können. Die MOM-(-CH₂OCH₃) oder Allyl-(-CH₂CH=CH₂) Gruppe wird unter Verwendung wohl etablierter Syntheseverfahren verwendet, um die 3'-OH-Gruppe zu schützen ([Fig. 13](#)) (Fuji et al., 1975; Metzker et al., 1994). Diese Gruppen können, wie in [Fig. 14](#) gezeigt, mit hoher Ausbeute chemisch entfernt werden (Ireland et al., 1986; Kamal et al., 1999). Die chemische Abspaltung der MOM- und Allyl-Gruppen ist relativ schonend und spezifisch, um den DNA-Template-Rest nicht abzubauen. So dauert z. B. die Abspaltung der Allylgruppe 3 Minuten bei einer Ausbeute von mehr als 93% (Kamal et al., 1999), während von der MOM-Gruppe berichtet wird, sie würde zu annähernd 100% abgespalten werden (Ireland et al., 1986).

5. Die Verwendung von energietransfergekoppelten Farbstoffen zur Optimierung des Systems der Sequenzierung durch Synthese

[0101] Die spektralen Eigenschaften der fluoreszierenden Tags können durch die Verwendung von energietransfergekoppelten (ET) Farbstoffen optimiert werden. Es wurde gezeigt, dass der ET-Primer und die ET-Didesoxynukleotide ein ausgezeichnetes Set von Reagenzien für die Vierfarben-DNA-Sequenzierung sind, das die Verwendung eines Lasers zur Anregung multipler Sets von fluoreszierenden Tags ermöglicht (Ju et al., 1995). Es wurde gezeigt, dass DNA-Polymerase (Thermo Sequenase und Taq FS) dazu in der Lage ist, die mit ET-Farbstoff markierten Didesoxynukleotide effizient zu integrieren (Rosenblum et al., 1997). Diese mit ET-Farbstoff markierten Sequenzierungsreagenzien werden inzwischen weitgehend in Großprojekten der DNA-Sequenzierung verwendet, so wie im Human Genom Projekt. Eine Bibliothek von mit ET-Farbstoff markierten Nukleotidanaloga kann, wie in **Fig. 15** gezeigt, zur Optimierung des DNA-Sequenzierungssystems synthetisiert werden. Das Set von ET-Farbstoffen (FAM-Cl₂FAM, FAM-Cl₂R6G, FAM-Cl₂TAM, FAM-Cl₂ROX) unter Verwendung von FAM als einem Donor und Dichloro (FAM, R6G, TAM, ROX) als Akzeptoren wurde in der Literatur beschrieben (Lee et al., 1997) und stellt ein Set von im Handel erhältlichen DNA-Sequenzierungsreagenzien dar. Es wurde bewiesen, dass diese ET-Farbstoff-Sets eine verstärkte Fluoreszenzintensität produzieren und dass die an der 5'-Position von T und C und der 7'-Position von G und A mit diesen ET-Farbstoffen markierten Nukleotide exzellente Substrate von DNA-Polymerase sind. Alternativ dazu kann ein ET-Farbstoff-Set unter Verwendung von Cyanin (Cy2) als einem Donor und Cl₂FAM, Cl₂R6G, Cl₂TAM oder Cl₂ROX als Energieakzeptoren konstruiert werden. Da Cy2 im Vergleich zu den Rhodamin- und Fluoresceinderivaten über eine höhere molare Absorbanz verfügt, produziert ein ET-System unter Verwendung von Cy2 als einem Donor viel stärkere Fluoreszenzsignale als das System, das FAM als einen Donor verwendet (Hung et al., 1996). **Fig. 16** zeigt ein Syntheschema für ein mit ET-Farbstoff markiertes Nukleotidanalogen mit Cy2 als einem Donor und Cl₂FAM als einem Akzeptor unter Verwendung einer ähnlichen Kopplungschemie wie die für die Synthese eines Energietransfersystems unter Verwendung von FAM als Donor (Lee et al., 1997). Die Kopplung von Cl₂FAM (I) mit dem Spacer 4-Aminomethylbenzoesäure (II) ergibt III, welches dann in NHS-Ester IV umgewandelt wird. Die Kopplung von IV mit Amino-Cy2 und die anschließende Umwandlung der resultierenden Verbindung in einen NHS-Ester ergibt V, welches sich anschließend an ein Amino-Photolinker-Nukleotid VI koppelt und so das mit ET-Farbstoff markierte Nukleotid VII ergibt.

6. Bewertung der Sequenzierung durch Synthese unter Verwendung der Nukleotidanaloga $3\text{-HO}^-\text{A-Farbstoff1}$, $3\text{-HO}^-\text{C-Farbstoff2}$, $3\text{-HO}^-\text{G-Farbstoff3}$, $3\text{-HO}^-\text{T-Farbstoff4}$

[0102] Die Vorläufer von vier Beispielen von Masse-Tags sind in **Fig. 17** gezeigt. Bei den Vorläufern handelt es sich um: (a) Acetophenon; (b) 3-Fluoracetophenon; (c) 3,4-Difluoracetophenon; und (d) 3,4-Dimethoxyacetophenon. Bei der Nitrierung und Reduktion werden aus den vier Vorläufern vier photoaktive Tags hergestellt und dazu verwendet, für die Identität eines jeden der vier Nukleotide (A, C, G, T) zu kodieren. Für die vier Masse-Tag-Vorläufer (a, b, c, d) werden saubere APCI-Massenspektren erhalten, wie in **Fig. 18** gezeigt. Der Spitzenwert mit einem Verhältnis von Masse zu Ladung (m/z) von 121 ist a, von 139 ist b, von 157 ist c und von 181 ist d. Dieses Ergebnis zeigt, dass diese vier Masse-Tags extrem stabil sind und in einem APCI-Massenspektrometer ohne Nebensignaleffekte zwischen den Masse-Tags sehr hohe Auflösungsdaten liefern. In den nachfolgend gezeigten Beispielen überträgt sich jedes der eindeutigen m/z-Verhältnisse eines jeden Masse-Tags auf die Identität des Nukleotids [Tag-1 (m/z, 150) = A; Tag-2 (m/z, 168) = C; Tag-3 (m/z, 186) = G; Tag-4 (m/z, 210) = T].

[0103] Es können unterschiedliche Kombinationen von Masse-Tags und Nukleotiden verwendet werden, wie durch das allgemeine Schema angedeutet: $3\text{-HO}^-\text{A-Tag1}$, $3\text{-HO}^-\text{C-Tag2}$, $3\text{-HO}^-\text{G-Tag3}$, $3\text{-HO}^-\text{T-Tag4}$, wobei Tag1, Tag2, Tag3 und Tag4 vier unterschiedliche eindeutige spaltbare Masse-Tags sind. In **Fig. 19** sind vier spezifische Beispiele von Nukleotidanaloga gezeigt. In **Fig. 19** ist "R" H, wenn die 3'-OH-Gruppe nicht geschützt ist. Wie zuvor besprochen, wurde der photospaltbare 2-Nitrobenzylrest dazu verwendet, Biotin zur effizienten Entfernung mittels Bestrahlung durch UV-Licht (~350 nm) an DNA und Protein zu knüpfen (Olejnik et al., 1995, 1999). Es werden vier verschiedene 2-Nitrobenzylgruppen mit unterschiedlichem Molekulargewicht als Masse-Tags verwendet, um die wie in **Fig. 19** gezeigten mit Masse-Tags markierten Nukleotide zu bilden: 2-Nitro- α -methyl-Benzyl (Tag-1) kodiert für A; 2-Nitro- α -Methyl-3-Fluorbenzyl (Tag-2) kodiert für C; 2-Nitro- α -methyl-3,4-Difluorbenzyl (Tag-3) kodiert für G; 2-Nitro- α -Methyl-3,4-Dimethoxybenzyl (Tag-4) kodiert für T.

[0104] Als ein repräsentatives Beispiel ist die Synthese des NHS-Esters eines Masse-Tags (Tag-3) in **Fig. 20** gezeigt. Es wird ein ähnliches Schema verwendet, um die anderen Masse-Tags zu konstruieren. Die Synthese von $3\text{-HO}^-\text{G-Tag3}$ unter Verwendung wohl etablierter Verfahren (Prober et al, 1987; Lee et al., 1992 und Hobbs et

al., 1991) ist in [Fig. 21](#) gezeigt. Es wird zunächst 7-Propargylamino-dGTP präpariert, indem man 7-I-dGTP mit N-Trifluoracetylpropargylamin reagiert, welches dann mit dem NHS-Tag-3 gekoppelt wird, um $3'\text{-HO-G-Tag}_3$ zu produzieren. Bei den Nukleotidanaloga mit einem freien 3'-OH handelt es sich um gute Substrate für die Polymerase.

[0105] Der Ansatz der Sequenzierung durch Synthese kann mit Masse-Tags unter Verwendung eines Schemas getestet werden, das dem in [Fig. 9](#) gezeigten Schema für Farbstoffe ähnlich ist. Es wird ein DNA-Templat, welches einen Abschnitt der Nukleotidsequenz enthält, die nach der Primingstelle keine sich wiederholenden Sequenzen aufweist, synthetisiert und in einem Glaskanal immobilisiert. Es werden $3'\text{-HO-A-Tag}_1$ und DNA-Polymerase zu dem selbstgeprimten DNA-Rest hinzugegeben, um den Einbau des Nukleotids in die 3'-Stelle der DNA zu gestatten. Es wird dann den Schritten aus [Fig. 2B](#) gefolgt (die chemische Abspaltung ist hier nicht erforderlich, da das 3'-OH frei ist), um den Masse-Tag aus Tag-1 ($m/z = 150$) zu detektieren. Als nächstes wird $3'\text{-HO-C-Tag}_2$ hinzugegeben und nach der Abspaltung von Tag-2 ($m/z = 168$) wird das resultierende Massenspektrum gemessen. Als nächstes werden wiederum $3'\text{-HO-G-Tag}_3$ und $3'\text{-HO-T-Tag}_4$ hinzugegeben und es werden die Massenspektren der Abspaltungsprodukte Tag-3 ($m/z = 186$) und Tag-4 ($m/z = 210$) gemessen. Beispiele für die erwarteten Photospaltungsprodukte sind in [Fig. 22](#) gezeigt. Der Photospaltungsmechanismus ist der gleiche, wie zuvor für den Fall beschrieben, in dem es sich bei den eindeutigen Markierungen um Farbstoffe handelt. Die Lichtabsorption (300–360 nm) durch den aromatischen 2-Nitrobenzylrest bewirkt die Reduktion der 2-Nitrogruppe zu einer Nitrosogruppe und eine Insertion von Sauerstoff in die Kohlenwasserstoffverbindung, die sich an der 2-Position befindet, gefolgt von der Abspaltung und Decarboxylierung (Pillai, 1980).

[0106] Die Synthese der Nukleotidanaloga $3'\text{-RO-A-Tag}_1$, $3'\text{-RO-C-Tag}_2$, $3'\text{-RO-G-Tag}_3$, $3'\text{-RO-T-Tag}_4$ kann zur weiteren Untersuchung des Systems fortgeführt werden wie zuvor für den Fall besprochen, in dem es sich bei den eindeutigen Markierungen um Farbstoffe handelt. Hier ist das 3'-OH geschützt in allen vier Nukleotidanaloga, welche dann mit DNA-Polymerase vermischt und zur Bewertung des Sequenzierungssystems unter Verwendung eines Schemas, das dem Schema aus [Fig. 9](#) ähnlich ist, verwendet werden können. Die MOM-(-CH₂OCH₃) oder die Allyl-(-CH₂CH=CH₂)-Gruppe wird verwendet, um die 3'-OH-Gruppe unter Verwendung wohl etablierter Syntheseverfahren zu schützen ([Fig. 13](#)) (Fuji et al, 1975; Metzker et al., 1994). Diese Gruppen können mit hoher Ausbeute chemisch entfernt werden, wie in [Fig. 14](#) gezeigt (Ireland et al., 1986; Kamal et al., 1999). Die chemische Abspaltung der MOM- und Allylgruppen ist relativ schonend und spezifisch, um den DNA-Templaterest nicht abzubauen.

7. Bewertung des kompletten Systems der Sequenzierung durch Synthese mittels der Sequenzierung von p53-Genen

[0107] Das Tumorsuppressorgen p53 kann als Modellsystem zur Bewertung des DNA-Sequenzierungssystems verwendet werden. Das p53-Gen ist eines der am häufigsten mutierten Gene bei Krebs im Menschen (O'Connor et al., 1997). Es wird zunächst ein Basenpaar-DNA-Template (nachfolgend gezeigt) synthetisiert, das am 5'-Ende eine Azidogruppe sowie einen Abschnitt der Sequenzen aus Exon 7 und Exon 8 des p53-Gens enthält:

```
5'-N3-TTCCTGCATGGGCGGCAIGAACCCGAGGCCCATCCTCACCATCATCAC
ACTGGAAGACTCCAGTGGTAATCTACTGGGACGGAACAGCTTTGAGGTGCA
TT-3' (SEQ ID NR: 2).
```

[0108] Dieses Template wird ausgewählt, um die Verwendung des Sequenzierungssystems zur Detektion von geclusterten Hot-Spot-Single-Basenmutationen zu erforschen. Die potentiell mutierten Basen sind in dem synthetischen Template unterstrichen (A, G, C und I). Das synthetische Template wird auf einem Sequenzierungschip oder in Glaskanälen immobilisiert, dann wird der Schleifenprimer wie in [Fig. 6](#) beschrieben an das immobilisierte Template ligiert und nachfolgend werden die Schritte aus [Fig. 2](#) befolgt, um die Sequenzierung zu bewerten. Es können mittels PCR erzeugte DNA-Templates zur weiteren Validierung des DNA-Sequenzierungssystems verwendet werden. Die Sequenzierungstemplates können mittels PCR unter Verwendung von flankierenden Primern (eines der Paare ist am 5'-Ende mit einer Azidogruppe markiert) in der Intronregion, die an jeder p53-Exon-Grenze gelegen ist, aus einem Pool genomischer DNA (Boehringer, Indianapolis, IN, USA) wie von Fu et al., (1998) beschrieben erzeugt und dann zur Bewertung der Sequenzierung auf dem DNA-Chip immobilisiert werden

Referenzen

Antao VP, Lai SY, Tinoco I Jr. (1991) A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucleic Acids Res.* 19:5901-5905.

- Axelrod VD, Vartikyan RM, Aivazashvili VA, Beabealashvili RS. (1978) Specific termination of RNA polymerase synthesis as a method of RNA and DNA sequencing. *Nucleic Acids Res.* 5(10): 3549-3563.
- Badman ER and Cooks RG. (2000) Cylindrical Ion Trap Array with Mass Selection by Variation in Trap Dimensions. *Anal. Chem.* 72 (20): 5079-5086.
- Badman ER and Cooks RG. (2000) A Parallel Miniature Cylindrical Ion Trap Array. *Anal. Chem.* 72 (14): 3291-3297.
- Bowling JM, Bruner KL, Cmarik JL, Tibbetts C. (1991) Neighboring nucleotide interactions during DNA sequencing gel electrophoresis. *Nucleic Acids Res.* 19: 3089-3097.
- Burgess K, Jacutin SE, Lim D, Shitangkoon A. (1997) An approach to photolabile, fluorescent protecting groups. *J. Org. Chem.* 62(15):5165-5168.
- Canard B, Cardona B, Sarfati RS. (1995) Catalytic editing properties of DNA polymerases. *Proc. Natl. Acad. Sci. USA* 92: 10859-10863.
- Caruthers MH. (1985) Gene synthesis machines: DNA chemistry and its uses. *Science* 230: 281-285.
- Chee M, Yang R, Hubbell E, Berno. A, Huang. XC., Stern D, Winkler, J, Lockhart DJ, Morris M S, Fodor, SP. (1996) Accessing genetic information with high-density DNA arrays. *Science.* 274: 610-614.
- Cheeseman PC. Method For Sequencing Polynucleotides, United States Patent No. 5,302,509, issued April 12, 1994.
- Dizidic I, Carrol, DI, Stillwell, RN, and Horning, MG.(1975) Atmospheric pressure ionization (API) mass spectrometry: formation of phenoxide ions from chlorinated aromatic compounds. *Anal. Chem.*, 47: 1308-1312.
- Fu DJ, Tang K, Braun A, Reuter D, Darnhofer-Demar B, Little DP, O'Donnell MJ, Cantor CR, Koster H. (1998) Sequencing exons 5 to 8 of the p53 gene by MALDI-TOF mass spectrometry. *Nat Biotechnol.* 16: 381-384.
- Fuji K, Nakano S, Fujita E. (1975) An improved method for methoxymethylation of alcohols under mild acidic conditions. *Synthesis* 276-277.
- Hobbs FW Jr, Coccuzza AJ. Alkynylamino-Nucleotides. United States Patent No. 5,047,519, issued September 10, 1991.
- Hung SC; Ju J; Mathies RA; Glazer AN. (1996) Cyanine dyes with high absorption cross section as donor chromophores in energy transfer primers. *Anal Biochem.* 243(1): 15-27.
- Hyman ED, (1988) A new method of sequencing DNA. *Analytical Biochemistry* 174: 423-436.
- Ireland RE, Varney MD (1986) Approach to the total synthesis of chlorothricolide-synthesis of (+/-)-19.2-O-dihydro-24-O-methylchlorothricolide, methyl-ester, ethyl carbonate. *J. Org. Chem.* 51: 635-648.
- Ju J, Glazer AN, Mathies RA. (1996) Cassette labeling for facile construction of energy transfer fluorescent primers. *Nucleic Acids Res.* 24: 1144-1148.
- Ju J, Ruan C, Fuller CW, Glazer AN Mathies RA. (1995) Energy transfer fluorescent dye-labeled primers for DNA sequencing and analysis. *Proc. Natl. Acad. Sci. USA* 92: 4347-4351.
- Kamal A, Laxman B, Rao NV. (1999) A mild and rapid regeneration of alcohols from their allylic ethers by chlorotrimethylsilane/sodium iodide. *Tetrahedron letters* 40: 371-372.
- Kheterpal I, Scherer J, Clark SM, Radhakrishnan A, Ju J, Ginther CL, Sensabaugh GF, Mathies RA. (1996) DNA Sequencing Using a Four-Color Confocal Fluorescence Capillary Array Scanner. *Electrophoresis.* 17: 1852-1859.
- Khokhi N, Vaultier M, Carrie R. (1987) Synthesis and reactivity of methyl-azido butyrates and ethyl-azido valerates and of the corresponding acid chlorides as useful reagents for the aminoalkylation. *Tetrahedron* 43:1811-1822.
- Lee LG, Connell CR, Woo SL, Cheng RD, Mcardle BF, Fuller CW, Halloran ND, Wilson RK. (1992) DNA sequencing with dye-labeled terminators and T7 DNA-polymerase-effect of dyes and dNTPs on incorporation of dye-terminators and probability analysis of termination fragments. *Nucleic Acids Res.* 20: 2471-2483.
- Lee LG, Spurgeon SL, Heiner CR, Benson SC, Rosenblum BB, Menchen SM, Graham RJ, Constantinescu A, Upadhyya KG, Cassel JM, (1997) New energy transfer dyes for DNA sequencing. *Nucleic Acids Res.* 25: 2816-2822.
- Liu H. H., Felton C., Xue Q. F., Zhang B., Jedrzejewski P., Karger B. L. and Foret F. (2000) Development of multichannel Devices with an Array of Electrospray tips for high-throughput mass spectrometry. *Anal. Chem.* 72: 3303-3310.
- Metzker ML, Raghavachari R, Richards S, Jacutin SE, Civitello A, Burgess K, Gibbs RA. (1994) Termination of DNA synthesis by novel 3'-modified-deoxyribonucleoside 5'-triphosphates. *Nucleic Acids Res.* 22: 4259-4267.
- O'Connor PM, Jackman J, Bae I, Myers TG, Fan S, Mutoh M, Scudiero DA, Monks A, Sausville EA, Weinstein JN, Friend S, Fornace AJ Jr, Kohn KW. (1997) Characterization of the p53 tumor suppressor pathway in cell lines of the National Cancer Institute anticancer drug screen and correlations with the growth-inhibitory potency of 123 anticancer agents. *Cancer Res.* 57: 4285-4300.
- Olejnik J, Ludemann HC, Krzymanska-Olejnik E, Berkenkamp S, Hillenkamp F, Rothschild KJ. (1999) Photocleavable peptide-DNA conjugates: synthesis and applications to DNA analysis using MALDI-MS. *Nucleic Acids Res.* 27:4626-4631.

- Olejnik J, Sonar S, Krzymanska-Olejnik E, Rothschild KJ. (1995) Photocleavable biotin derivatives: a versatile approach for the isolation of biomolecules. *Proc. Natl. Acad. Sci. USA.* 92: 7590-7594.
- Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J. (1994) Structures of ternary complexes of rat DNA polymerase β , a DNA template-primer, and ddCTP. *Science* 264: 1891-1903.
- Pennisi E. (2000) DOE Team Sequences Three Chromosomes. *Science* 288: 417-419.
- Pillai VNR. (1980) Photoremovable Protecting Groups in Organic Synthesis. *Synthesis* 1-62.
- Prober SM, Trainor GL, Dam RJ, Hobbs FW, Robertson CW, Zagursky RJ, Cocuzza AJ, Jensen MA, Baumeister K. (1987) A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* 238: 336-341.
- Rollaf F. (1982) Sodium-borohydride reactions under phase-transfer conditions-reduction of azides to amines. *J. Org. Chem.* 47: 4327-4329.
- Ronaghi M, Uhlen M, Nyren P. (1998) A sequencing Method based on real-time pyrophosphate. *Science* 281: 364-365.
- Rosenblum BB, Lee LG, Spurgeon SL, Khan SH, Menchen SM, Heiner CR, Chef SM. (1997) New dye-labeled terminators for improved DNA sequencing patterns. *Nucleic Acids Res.* 25: 4500-4504.
- Roses A. (2000) Pharmacogenetics and the practice of medicine. *Nature.* 405: 857-865.
- Salas-Solano O, Carrilho E, Kotler L, Miller AW, Goetzinger W, Sosic Z, Karger BL, (1998) Routine DNA sequencing of 1000 bases in less than-one hour by capillary electrophoresis with replaceable linear polyacrylamide solutions. *Anal. Chem.* 70: 3996-4003.
- Saxon E and Bertozzi CR (2000) Cell surface engineering by a modified Staudinger reaction. *Science* 287: 2007-2010.
- Schena M, Shalon D, Davis, R. Brown P. O. (1995) Quantitative monitoring of gene expression patterns with a cDNA microarray. *Science* 270: 467-470.
- Simpson PC, Adam DR, Woolley T, Thorsen T, Johnston R, Sensabaugh GF, and Mathies RA. (1998) High-throughput genetic analysis using microfabricated 96-sample capillary array electrophoresis microplates. *Proc. Natl. Acad. Sci. U.S.A.* 95: 2256-2261.
- Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd C, Connell CR, Heiner C, Kent SBH, Hood LE. (1986) Fluorescence detection in automated DNA sequencing analysis. *Nature* 321: 674-679.
- Tabor S, Richardson C. C. (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* 84:4767-4771.
- Tabor S. & Richardson, CC. (1995) A single residue in DNA polymerases of the Escherichia coli DNA polymerase I family is critical for distinguishing between deoxy-and dideoxyribonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* 92: 6339-6343.
- Turro NJ. (1991) *Modern Molecular Photochemistry*; University Science Books, Mill Valley, CA.
- Velculescu VE, Zhang, I, Vogelstein, B. and Kinzler KW (1995) Serial Analysis of Gene Expression. *Science* 270: 484-487.
- Welch MB, Burgess K, (1999) Synthesis of fluorescent, photolabile 3'-O-protected nucleoside triphosphates for the base addition sequencing scheme. *Nucleosides and Nucleotides* 18: 197-201.
- Woolley AT, Mathies RA. (1994) Ultra-high-speed DNA fragment separations using microfabricated capillary array electrophoresis chips. *Proc. Natl. Acad. Sci. USA.* 91: 11348-11352.
- Woolley AT, Sensabaugh GF and Mathies RA. (1997) High-Speed DNA Genotyping Using Microfabricated Capillary Array Electrophoresis Chips, *Anal. Chem.* 69 (11); 2181-2186.
- Yamakawa H, Ohara O. (1997) A DNA cycle sequencing reaction that minimizes compressions on automated fluorescent sequencers. *Nucleic. Acids. Res.* 25: 1311-1312.
- Zhang XH, Chiang VL, (1996) Single-stranded DNA ligation by T4 RNA ligase for PCR cloning of 5'-noncoding fragments and coding sequence of a specific gene. *Nucleic Acids Res.* 24: 990-991.
- Zhang B., Liu H. Karger BL. Foret F. (1999) Microfabricated devices for capillary electrophoresis electrospray mass spectrometry. *Anal. Chem.* 71: 3258-3264.
- Zhu Z, Chao J, Vu H, Waggoner AS. (1994) Directly labeled DNA probes using fluorescent nucleotides with different length linkers. *Nucleic Acids Res.* 22: 3418-3422.

SEQUENZPROTOKOLL

<110> The Trustees Of Columbia University In The City Of

<120> Massives paralleles Verfahren zur Dekodierung von DNA und RNA

<130> 0575/62239-B-PCT/JPW

<140> Noch nicht bekannt

<141> 2001-10-05

<150> 09/684,670

<151> 2000-10-06

<160> 2

<170> PatentIn Ver. 2.1

<210> 1

<211> 11

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: Template

<400> 1

acgtacgacg t 11

<210> 2

<211> 101

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: Template

<400> 2

ttcctgcatg ggcggcatga aeccgaggcc caccctcacc atcatcacac tggaagactc 60
cagtggaat ctactgggac ggaacagctt tgaggtgcat t 101

Patentansprüche

1. Verfahren zum Sequenzieren einer Nukleinsäure durch sequentielles Bestimmen der Identität von Nukleotidanaloga nachdem die Nukleotidanaloga in einen wachsenden Strang von DNA eingebaut sind, in einer Polymerasereaktion, die die folgenden Schritte umfasst:

- (i) Anheften eines 5' Endes der Nukleinsäure an eine feste Oberfläche;
- (ii) Anheften eines Primers an die an die feste Oberfläche angeheftete Nukleinsäure;
- (iii) Zugeben einer Polymerase und eines oder mehrerer verschiedener Nukleotidanaloga zu der Nukleinsäure um dadurch ein Nukleotidalogon in den wachsenden Strang von DNA einzubauen, wobei das eingebaute Nukleotidalogon die Polymerasereaktion beendet, und wobei jedes verschiedene Nukleotidalogon umfasst (a) eine Base ausgewählt aus der Gruppe bestehend aus Adenin, Guanin, Cytosin, Thymin, und Uracil, und ihrer Analoga; (b) eine eindeutige Markierung anheftet über einen spaltbaren Linker an die Base oder an ein Analogon der Base; (c) eine Desoxyribose; und (d) eine spaltbare chemische Gruppe um eine -OH-Gruppe an einer 3'-Position der Desoxyribose zu schützen, wobei die spaltbare chemische Gruppe $-\text{CH}_2\text{OCH}_3$ oder $-\text{CH}_2\text{CH}=\text{CH}_2$ ist;

- (iv) Waschen der festen Oberfläche, um nicht-eingebaute Nukleotidanaloga zu entfernen;
 - (v) Bestimmen der Identität der eindeutigen Markierung, angehängt an das Nukleotidanalogen, das in den wachsenden Strang von DNA eingebaut worden ist, um dadurch das eingebaute Nukleotidanalogen zu identifizieren;
 - (vi) Zugeben einer oder mehrere chemischer Verbindungen, um dauerhaft jede nicht-reagierte -OH-Gruppe auf dem Primer, angehängt an die Nukleinsäure oder auf einem Primer-Extension-Strang, gebildet durch Zugabe eines oder mehrerer Nukleotide oder Nukleotidanaloga zu dem Primer zu schützen;
 - (vii) Spalten des spaltbaren Linkers zwischen dem Nukleotidanalogen, das in den wachsenden Strang von DNA eingebaut worden ist, und der eindeutigen Markierung;
 - (viii) Spalten der spaltbaren chemischen Gruppe, die die -OH-Gruppe an der 3'-Position der Desoxyribose schützt, um die -OH-Gruppe zu entschützen, und Waschen der festen Oberfläche um gespaltene Verbindungen zu entfernen; und
 - (ix) Wiederholen der Schritte (iii) bis (viii), um für jede Wiederholung, die Identität des neu eingebauten Nukleotidanalogs in den wachsenden Strang von DNA zu bestimmen;
- wobei in dem Fall, dass es sich bei der eindeutigen Markierung um einen Farbstoff handelt, die Reihenfolge der Schritte (v) bis einschließlich (vii) folgendermaßen lautet: (v), (vi) und (vii); und
- wobei in dem Fall, dass es sich bei der eindeutigen Markierung um einen Masse-Tag handelt, die Reihenfolge der Schritte (v) bis einschließlich (vii) folgendermaßen lautet: (vi), (vii) und (v).

2. Verfahren nach Anspruch 1, wobei die feste Oberfläche Glas, Silikon oder Gold ist.

3. Verfahren nach Anspruch 1, wobei die feste Oberfläche ein Magnetkügelchen, ein Chip, ein Kanal in einem Chip oder ein poröser Kanal in einem Chip ist.

4. Verfahren nach Anspruch 1, wobei der Schritt des Anheftens der Nukleinsäure an die feste Oberfläche umfasst:

- (i) Beschichten der festen Oberfläche mit einem Phosphinrest,
- (ii) Anheften einer Azidogruppe an das 5'-Ende der Nukleinsäure, und
- (iii) Immobilisieren des 5'-Endes der Nukleinsäure auf der festen Oberfläche durch Wechselwirkung zwischen dem Phosphinrest auf der festen Oberfläche und der Azidogruppe am 5'-Ende der Nukleinsäure.

5. Verfahren nach Anspruch 4, wobei der Schritt des Beschichtens der festen Oberfläche mit dem Phosphinrest umfasst:

- (i) Beschichten der Oberfläche mit einem primären Amin, und
- (ii) Kovalentes Binden eines N-Hydroxysuccinimidylesters von Triarylphosphin mit dem primären Amin.

6. Verfahren nach Anspruch 1, wobei die Nukleinsäure, die an die feste Oberfläche angeheftet wird, eine einzelsträngige DNA ist.

7. Verfahren nach Anspruch 1, wobei die Nukleinsäure, die in Schritt (i) an die feste Oberfläche angeheftet wird, eine doppelsträngige DNA ist, wobei nur ein Strang direkt an die feste Oberfläche angeheftet wird, und wobei der Strang, der nicht direkt an die feste Oberfläche angeheftet wird, vor dem Übergang zu Schritt (ii) durch Denaturieren entfernt wird.

8. Verfahren nach Anspruch 1, wobei die Nukleinsäure, die an die feste Oberfläche angeheftet wird, eine RNA ist, und die Polymerase in Schritt (iii) reverse Transkriptase ist.

9. Verfahren nach Anspruch 1, wobei der Primer an ein 3'-Ende der Nukleinsäure in Schritt (ii) angeheftet wird, und wobei der angeheftete Primer eine stabile Schleife und eine -OH-Gruppe an einer 3'-Position einer Desoxyribose umfasst, fähig zum Selbstprimen in der Polymerasereaktion.

10. Verfahren nach Anspruch 1, wobei der Schritt des Anheftens des Primers an die Nukleinsäure Hybridisieren des Primers an die Nukleinsäure oder Ligieren des Primers an die Nukleinsäure umfasst.

11. Verfahren nach Anspruch 1, wobei eines oder mehrere von vier verschiedenen Nukleotidanaloga in Schritt (iii) zugegeben wird, wobei jedes verschiedene Nukleotidanalogen eine verschiedene Base umfasst, ausgewählt aus der Gruppe bestehend aus Thymin oder Uracil oder einem Analogon von Thymin oder Uracil, Adenin oder einem Analogon von Adenin, Cytosin oder einem Analogon von Cytosin, und Guanin oder einem Analogon von Guanin, und wobei jedes der vier verschiedenen Nukleotidanaloga eine eindeutige Markierung umfasst.

12. Verfahren nach Anspruch 1, wobei die spaltbare chemische Gruppe, die die -OH-Gruppe an der 3'-Position der Desoxyribose in dem Nukleotidanalogen schützt $-\text{CH}_2\text{CH}=\text{CH}_2$ ist.
13. Verfahren nach Anspruch 1, wobei die eindeutige Markierung, die an das Nukleotidanalogen angeheftet ist, ein fluoreszierender Rest oder ein fluoreszierender Halbleiterkristall ist.
14. Verfahren nach Anspruch 13, wobei der fluoreszierende Rest ausgewählt ist aus der Gruppe bestehend aus 5-Carboxyfluorescein, 6-Carboxyrhodamin-6G, N,N,N',N'-Tetramethyl-6-carboxyrhodamin, und 6-Carboxy-X-rhodamin.
15. Verfahren nach Anspruch 1, wobei die eindeutige Markierung, die an das Nukleotidanalogen angeheftet ist, ein Fluoreszenzenergietransfer-Tag ist, die einen Energietransferdonor und einen Energietransferakzeptor umfasst.
16. Verfahren nach Anspruch 15, wobei der Energietransferdonor 5-Carboxyfluorescein oder Cyanin ist, und wobei der Energietransferakzeptor ausgewählt ist aus der Gruppe bestehend aus Dichlorcarboxyfluorescein, Dichlor-6-carboxyrhodamin-6G, Dichlor-N,N,N',N'-tetramethyl-6-carboxyrhodamin, und Dichlor-6-carboxy-X-rhodamin.
17. Verfahren nach Anspruch 1, wobei die eindeutige Markierung, die an das Nukleotidanalogen angeheftet ist, ein Masse-Tag ist, die über ein Massenspektrometer detektiert und unterschieden werden kann.
18. Verfahren nach Anspruch 17, wobei das Masse-Tag ausgewählt ist aus der Gruppe bestehend aus einer 2-Nitro- α -methyl-benzyl-Gruppe, einer 2-Nitro- α -methyl-3-fluorbenzyl-Gruppe, einer 2-Nitro- α -methyl-3,4-difluorbenzyl-Gruppe, und einer 2-Nitro- α -methyl-3,4-dimethoxybenzyl-Gruppe.
19. Verfahren nach Anspruch 1, wobei die eindeutige Markierung über einen spaltbaren Linker an eine 5-Position von Cytosin oder Thymin oder an eine 7-Position von Deazaadenin oder Deazaguanin angeheftet ist.
20. Verfahren nach Anspruch 1, wobei der spaltbare Linker zwischen der eindeutigen Markierung und dem Nukleotidanalogen gespalten wird durch ein Mittel ausgewählt aus der Gruppe bestehend aus einem oder mehreren physikalischen Mitteln, einem chemischen Mittel, einem physikalisch-chemischen Mittel, Hitze und Licht.
21. Verfahren nach Anspruch 20, wobei der spaltbare Linker ein photospaltbarer Linker ist, der einen 2-Nitrobenzylrest umfasst.
22. Verfahren nach Anspruch 1, wobei die spaltbare chemische Gruppe, verwendet um die -OH-Gruppe an der 3'-Position der Desoxyribose zu schützen, gespalten wird durch ein Mittel ausgewählt aus der Gruppe bestehend aus einem oder mehreren physikalischen Mitteln, einem chemischen Mittel, eines physikalisch-chemischen Mittel, Hitze und Licht.
23. Verfahren nach Anspruch 1, wobei die in Schritt (vi) zugegebenen chemischen Verbindungen zum permanenten Schützen jeder nicht-reagierten -OH-Gruppe auf dem Primer, angeheftet an die Nukleinsäure oder auf dem Primer-Extensionstrang, eine Polymerase und ein oder mehrere verschiedene Didesoxynukleotide oder Analoga von Didesoxynukleotiden sind.
24. Verfahren nach Anspruch 23, wobei die verschiedenen Didesoxynukleotide ausgewählt sind aus der Gruppe bestehend aus 2',3'-Didesoxyadenosin-5'-triphosphat, 2',3'-Didesoxyguanosin-5'-triphosphat, 2',3'-Didesoxycytidin-5'-triphosphat, 2',3'-Didesoxythymidin-5'-triphosphat, 2',3'-Didesoxyuridin-5'-triphosphat, und ihre Analoga.
25. Verfahren nach Anspruch 1, wobei eine Polymerase und ein oder mehrere von vier verschiedenen Didesoxynukleotiden in Schritt (vi) zugegeben werden, und wobei jedes verschiedene Didesoxynukleotid ausgewählt ist aus der Gruppe bestehend aus 2',3'-Didesoxyadenosin-5'-triphosphat oder einem Analogon von 2',3'-Didesoxyadenosin-5'-triphosphat; 2',3'-Didesoxyguanosin-5'-triphosphat oder einem Analogon von 2',3'-Didesoxyguanosin-5'-triphosphat; 2',3'-Didesoxycytidin-5'-triphosphat oder einem Analogon von 2',3'-Didesoxycytidin-5'-triphosphat; und 2',3'-Didesoxythymidin-5'-triphosphat oder 2',3'-Didesoxyuridin-5'-triphosphat oder einem Analogon von 2',3'-Didesoxythymidin-5'-triphosphat oder einem Analogon von 2',3'-Didesoxyuridin-5'-triphosphat.

26. Verfahren nach Anspruch 17, wobei das Masse-Tag detektiert wird unter Verwendung eines parallelen Massenspektrometrysystems, das eine Vielzahl von Atmospheric Pressure Chemical Ionization-Massenspektrometern umfasst zur parallelen Analyse einer Vielzahl von Proben, die Masse-Tags umfassen.

27. Ein Verfahren zum gleichzeitigen Sequenzieren einer Vielzahl von verschiedenen Nukleinsäuren, welches das gleichzeitige Anwenden des Verfahrens nach Anspruch 1 auf die Vielzahl von verschiedenen Nukleinsäuren umfasst.

28. Verwendung des Verfahrens nach Anspruch 1 oder 27 zum Nachweis von Einzelnukleotidpolymorphismen, genetische Mutationsanalyse, serielle Analyse von Genexpression, Genexpressionsanalyse, Identifikation in der Forensik, Assoziationsstudien genetischer Krankheiten, DNA-Sequenzierung, genomische Sequenzierung, Translationsanalyse oder Transkriptionsanalyse.

29. Ein Verfahren zum Anheften einer Nukleinsäure an eine feste Oberfläche, welches umfasst:
 (i) Beschichten der festen Oberfläche mit einem Phosphinrest,
 (ii) Anheften einer Azidogruppe an ein 5'-Ende der Nukleinsäure, und
 (iii) Immobilisieren des 5'-Endes der Nukleinsäure auf der festen Oberfläche durch Wechselwirkung zwischen dem Phosphinrest auf der festen Oberfläche und der Azidogruppe am 5'-Ende der Nukleinsäure.

30. Verfahren nach Anspruch 29, wobei der Schritt des Beschichtens der festen Oberfläche mit dem Phosphinrest umfasst:

- (i) Beschichten der Oberfläche mit einem primären Amin, und
- (ii) Kovalentes Binden eines N-Hydroxysuccinimidylesters von Triarylphosphin mit dem primären Amin.

31. Verfahren nach Anspruch 29, wobei die feste Oberfläche Glas, Silikon oder Gold ist.

32. Verfahren nach Anspruch 29, wobei die feste Oberfläche ein Magnetkugeln, ein Chip, ein Kanal in einem Chip oder ein poröser Kanal in einem Chip ist.

33. Verfahren nach Anspruch 29, wobei die Nukleinsäure, die an die feste Oberfläche angeheftet wird, eine einzelsträngige DNA, eine doppelsträngige DNA oder eine RNA ist.

34. Verfahren nach Anspruch 33, wobei die Nukleinsäure eine doppelsträngige DNA ist und nur ein Strang an die feste Oberfläche angeheftet ist.

35. Verfahren nach Anspruch 34, wobei der Strang der doppelsträngigen DNA, der nicht an die feste Oberfläche angeheftet ist, durch Denaturieren entfernt wird.

36. Verwendung des Verfahrens nach Anspruch 29 zum Anheften einer Nukleinsäure an eine feste Oberfläche für Genexpressionsanalyse, Microarray-basierte Genexpressionsanalyse, Mutationsnachweis, Translationsanalyse, oder Transkriptionsanalyse dieser Nukleinsäure.

37. Ein Nukleotidanalogen welches umfasst:

- (a) eine Base ausgewählt aus der Gruppe bestehend aus Adenin oder einem Analogon von Adenin, Cytosin oder einem Analogon von Cytosin, Guanin oder einem Analogon von Guanin, Thymin oder einem Analogon von Thymin, und Uracil oder einem Analogon von Uracil;
- (b) eine eindeutige Markierung angeheftet über einen spaltbaren Linker an die Base oder an ein Analogon der Base;
- (c) eine Desoxyribose; und
- (d) eine spaltbare chemische Gruppe, um eine -OH-Gruppe an einer 3'-Position der Desoxyribose zu schützen, wobei die spaltbare chemische Gruppe $-\text{CH}_2\text{OCH}_3$ oder $-\text{CH}_2\text{CH}=\text{CH}_2$ ist.

38. Nukleotidanalogen nach Anspruch 37, wobei die spaltbare chemische Gruppe, welche die -OH-Gruppe an der 3'-Position der Desoxyribose in dem Nukleotidanalogen schützt $-\text{CH}_2\text{CH}=\text{CH}_2$ ist.

39. Nukleotidanalogen nach Anspruch 37, wobei die eindeutige Markierung ein fluoreszierender Rest oder ein fluoreszierender Halbleiterkristall ist.

40. Nukleotidanalogen nach Anspruch 39, wobei der fluoreszierende Rest ausgewählt ist aus der Gruppe bestehend aus 5-Carboxyfluorescein, 6-Carboxyrhodamin-6G, N,N,N',N'-Tetramethyl-6-carboxyrhodamin,

und 6-Carboxy-X-rhodamin.

41. Nukleotidanalogen nach Anspruch 37, wobei die eindeutige Markierung ein Fluoreszenzenergietransfer-Tag ist, die einen Energietransferdonor und einen Energietransferakzeptor umfasst.

42. Nukleotidanalogen nach Anspruch 41, wobei der Energietransferdonor 5-Carboxyfluorescein oder Cyanin ist, und wobei der Energietransferakzeptor ausgewählt ist aus der Gruppe bestehend aus Dichlorcarboxyfluorescein, Dichlor-6-carboxyrhodamin-6G, Dichlor-N,N,N',N'-tetramethyl-6-carboxyrhodamin, und Dichlor-6-carboxy-X-rhodamin.

43. Nukleotidanalogen nach Anspruch 37, wobei die eindeutige Markierung ein Masse-Tag ist, das durch ein Massenspektrometer detektiert und unterschieden werden kann.

44. Nukleotidanalogen nach Anspruch 43, wobei das Masse-Tag ausgewählt ist aus der Gruppe bestehend aus einer 2-Nitro- α -methyl-benzyl-Gruppe, einer 2-Nitro- α -methyl-3-fluorbenzyl-Gruppe, einer 2-Nitro- α -methyl-3,4-difluorbenzyl-Gruppe, und einer 2-Nitro- α -methyl-3,4-dimethoxybenzyl-Gruppe.

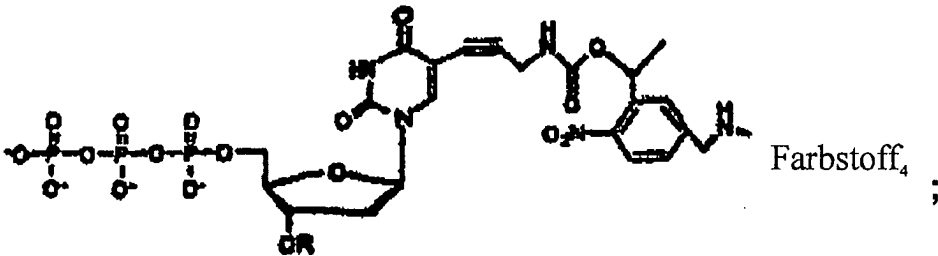
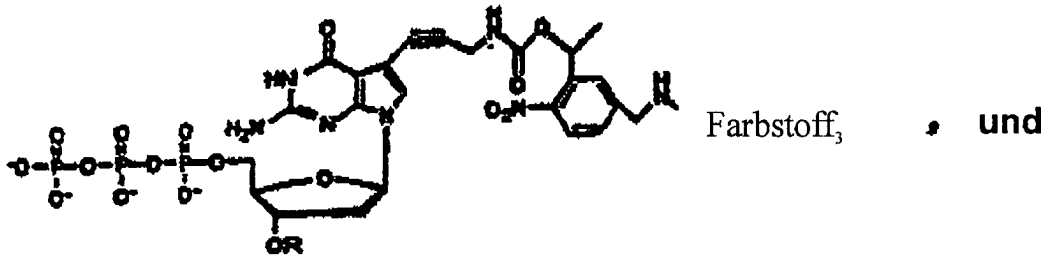
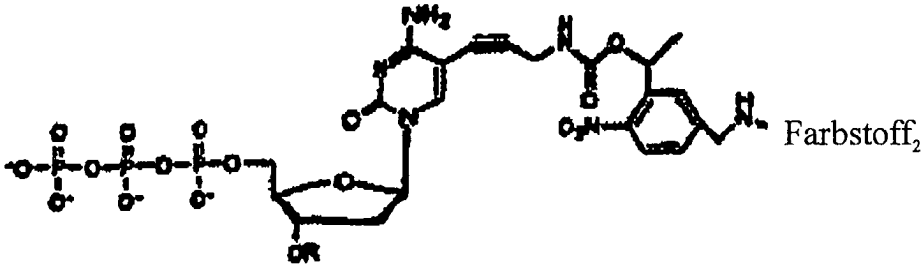
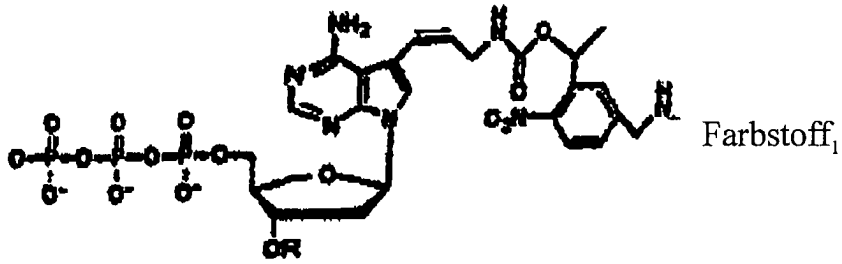
45. Nukleotidanalogen nach Anspruch 37, wobei die eindeutige Markierung über einen spaltbaren Linker an eine 5-Position von Cytosin oder Thymin oder an eine 7-Position von Deazaadenin oder Deazaguanin angeheftet ist.

46. Nukleotidanalogen nach Anspruch 37, wobei der Linker zwischen der eindeutigen Markierung und dem Nukleotidanalogen spaltbar ist durch ein Mittel ausgewählt aus der Gruppe bestehend aus einem oder mehreren physikalischen Mitteln, einem chemischen Mittel, einem physikalisch-chemischen Mittel, Hitze und Licht.

47. Nukleotidanalogen nach Anspruch 46, wobei der spaltbare Linker ein photospaltbarer Linker ist, der einen 2-Nitrobenzylrest umfasst.

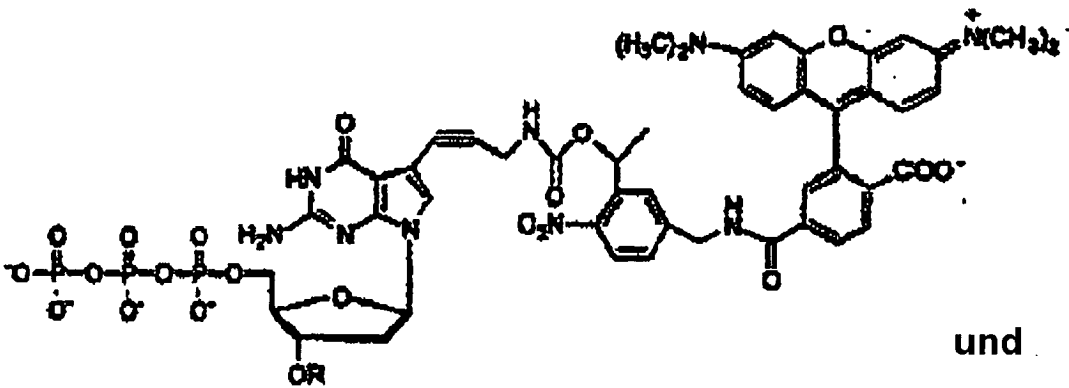
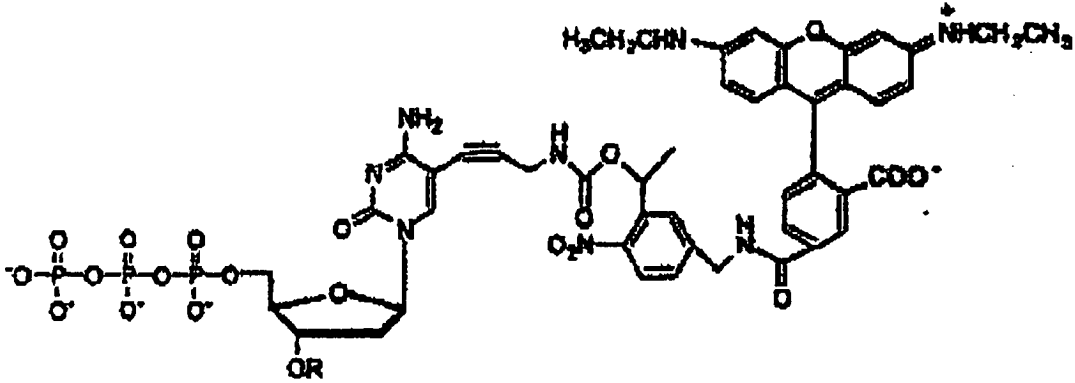
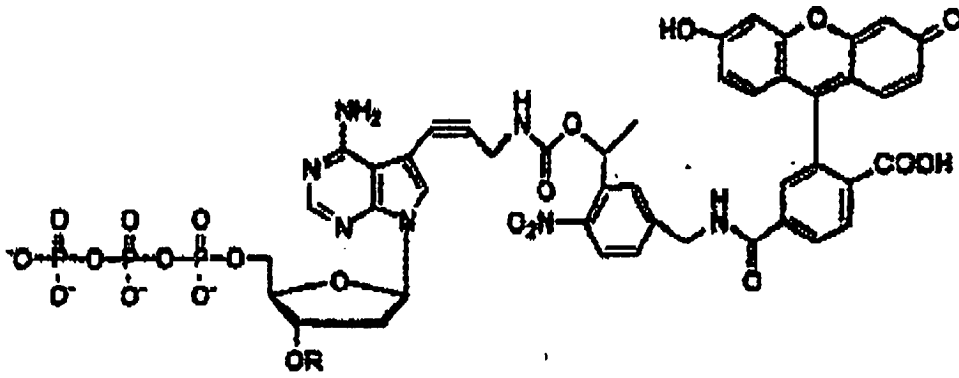
48. Nukleotidanalogen nach Anspruch 37, wobei die spaltbare chemische Gruppe, verwendet um die -OH-Gruppe an der 3'-Position der Desoxyribose zu schützen, spaltbar ist durch ein Mittel ausgewählt aus der Gruppe bestehend aus einem oder mehreren physikalischen Mitteln, einem chemischen Mittel, einem physikalisch-chemischen Mittel, Hitze und Licht.

49. Nukleotidanalogen nach Anspruch 37, wobei das Nukleotidanalogen ausgewählt ist aus der Gruppe bestehend aus:

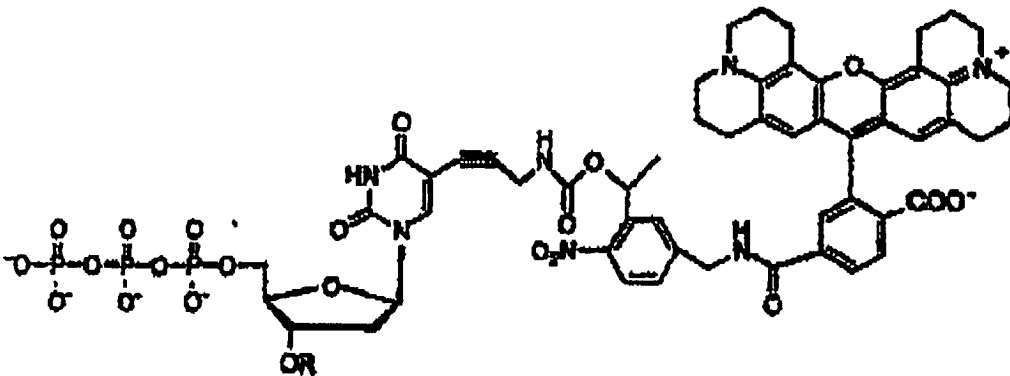


wobei Farbstoff₁, Farbstoff₂, Farbstoff₃ und Farbstoff₄ vier verschiedene Farbstoffmarkierungen sind; und wobei R -CH₂OCH₃ oder -CH₂CH=CH₂ ist.

50. Nukleotidanalogen nach Anspruch 49, wobei das Nukleotidanalogen ausgewählt ist aus der Gruppe bestehend aus:

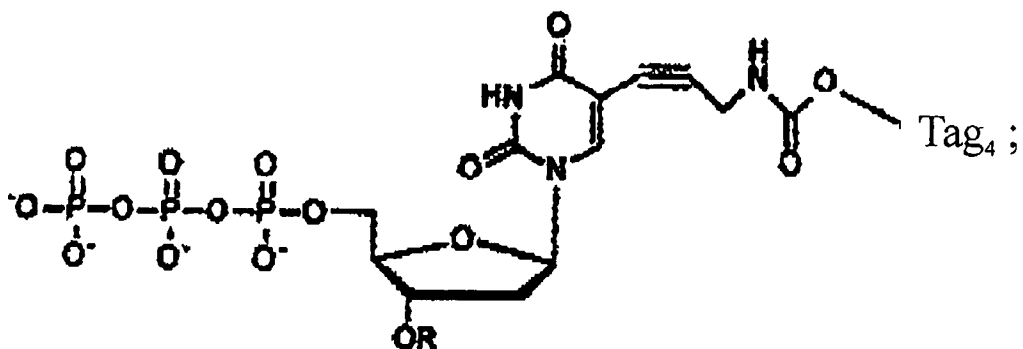
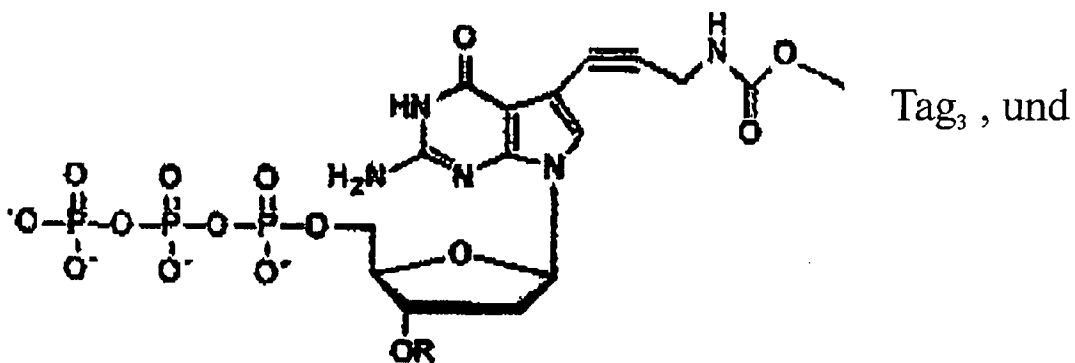
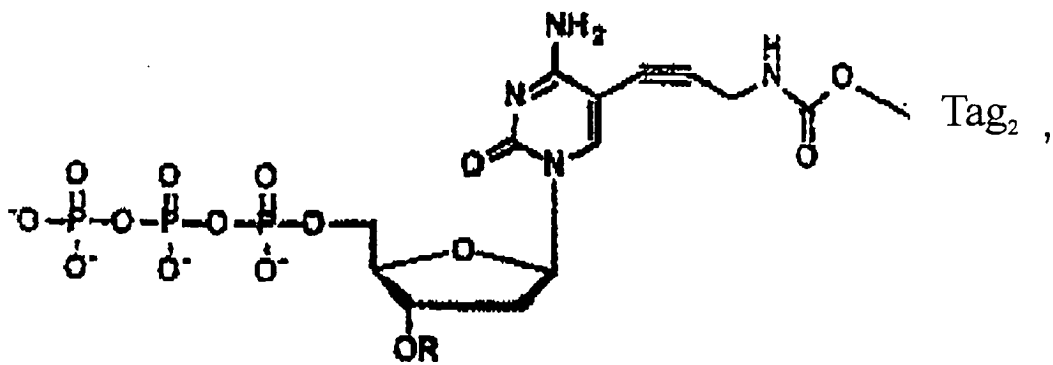
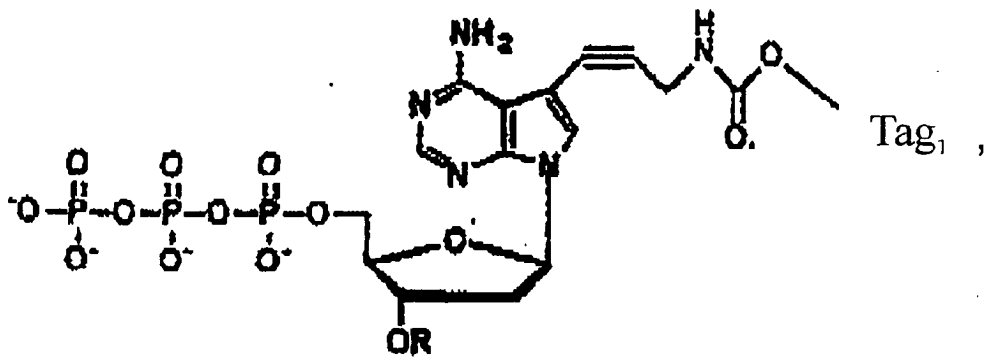


und



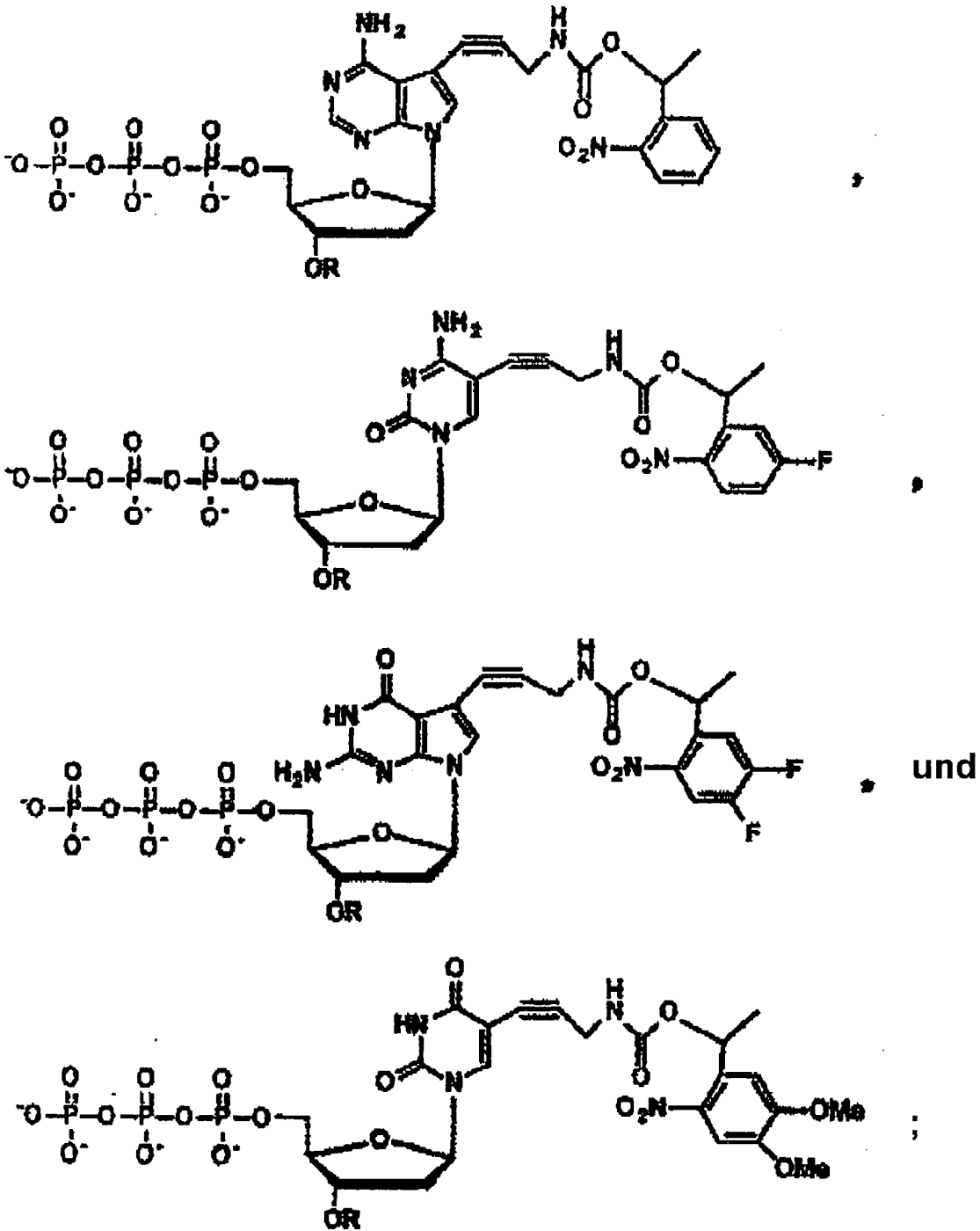
wobei R -CH₂OCH₃ oder -CH₂CH=CH₂ ist.

51. Nukleotidanalogen nach Anspruch 37, wobei das Nukleotidanalogen ausgewählt ist aus der Gruppe bestehend aus:



wobei Tag₁, Tag₂, Tag₃, und Tag₄ vier verschiedene Masse-Tag-Markierungen sind; und wobei R -CH₂OCH₃ oder -CH₂CH=CH₂ ist.

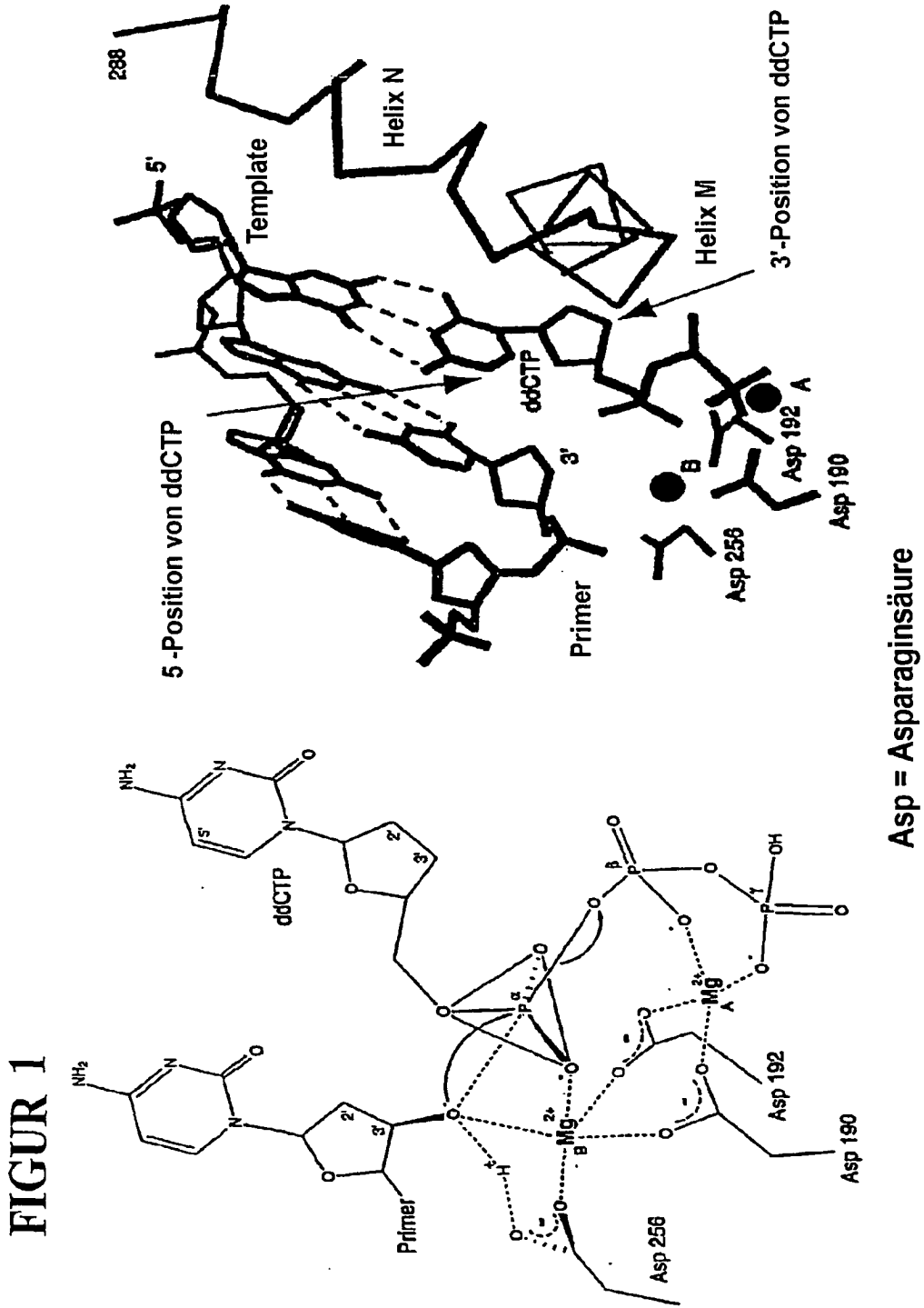
52. Nukleotidanalogen nach Anspruch 51, wobei das Nukleotidanalogen ausgewählt ist aus der Gruppe bestehend aus:

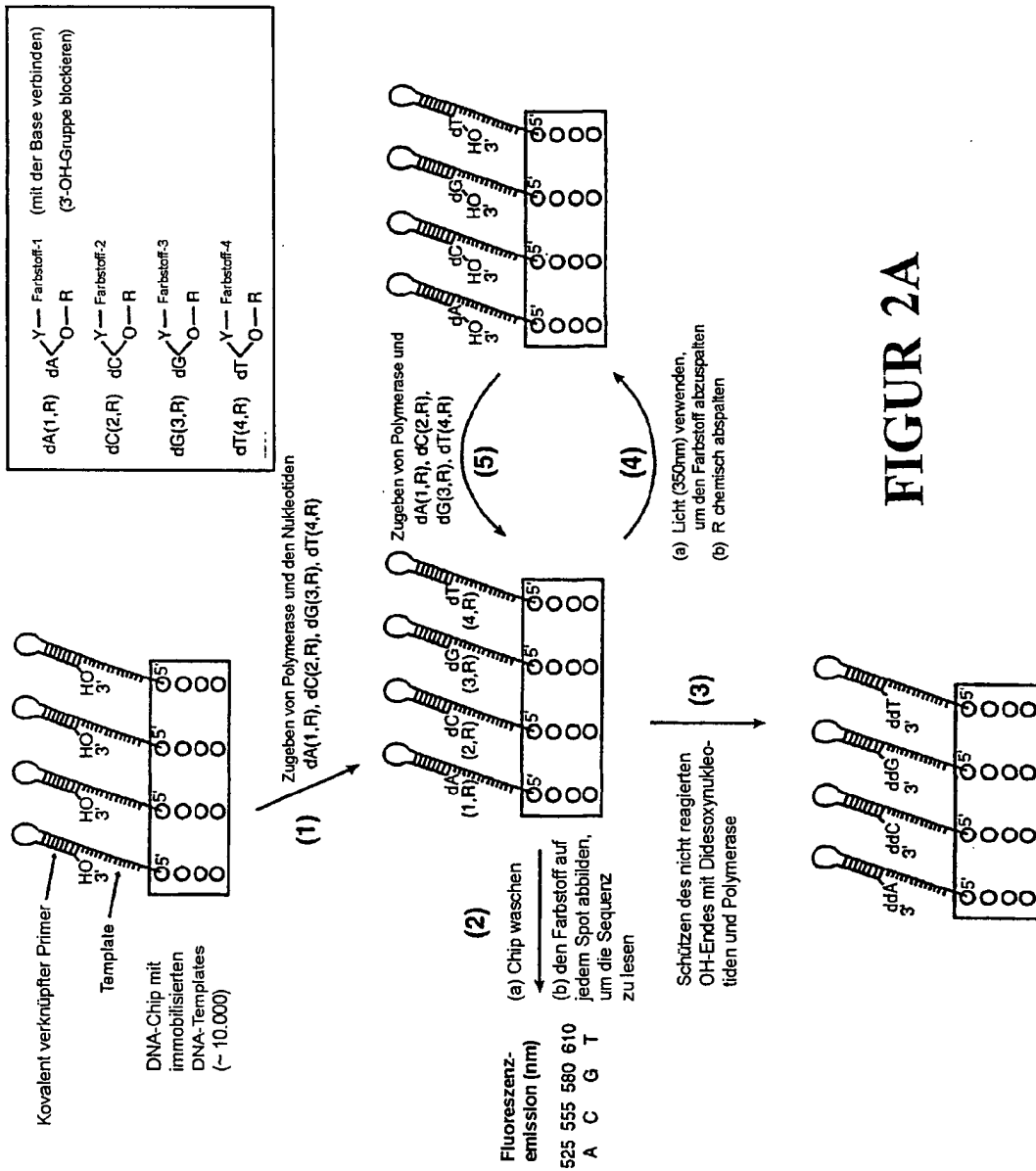


wobei R $-\text{CH}_2\text{OCH}_3$ oder $-\text{CH}_2\text{CH}=\text{CH}_2$ ist.

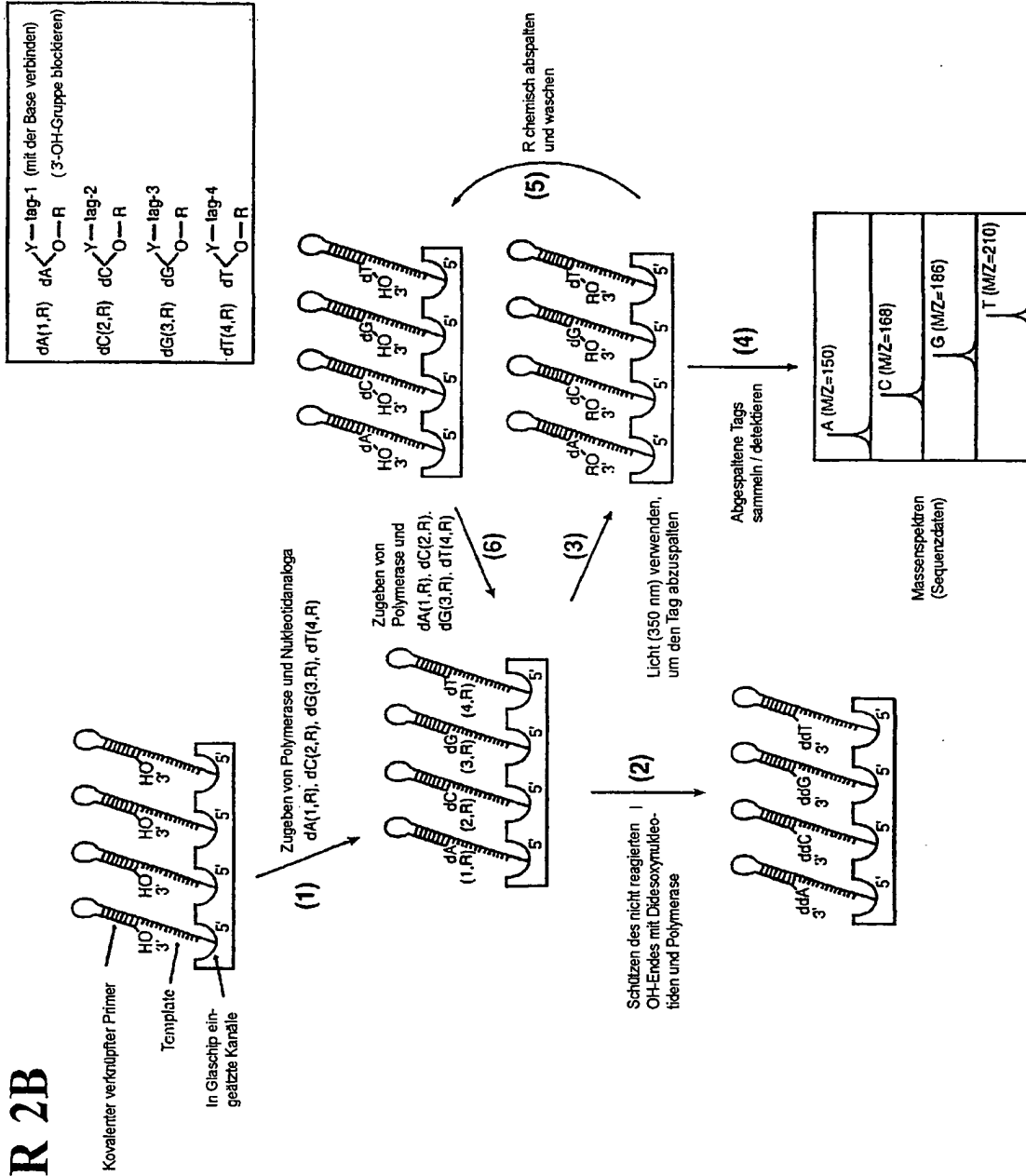
53. Verwendung des Nucleotidanalogs nach Anspruch 36 zum Nachweis von Einzelnucleotidpolymorphismen, zur genetischen Mutationsanalyse, serielle Analyse von Genexpression, Genexpressionsanalyse, Identifizierung in der Forensik, Assoziationsstudien genetischer Krankheiten, DNA-Sequenzierung, genomische Sequenzierung, Translationsanalyse oder Transkriptionsanalyse.

Es folgen 26 Blatt Zeichnungen

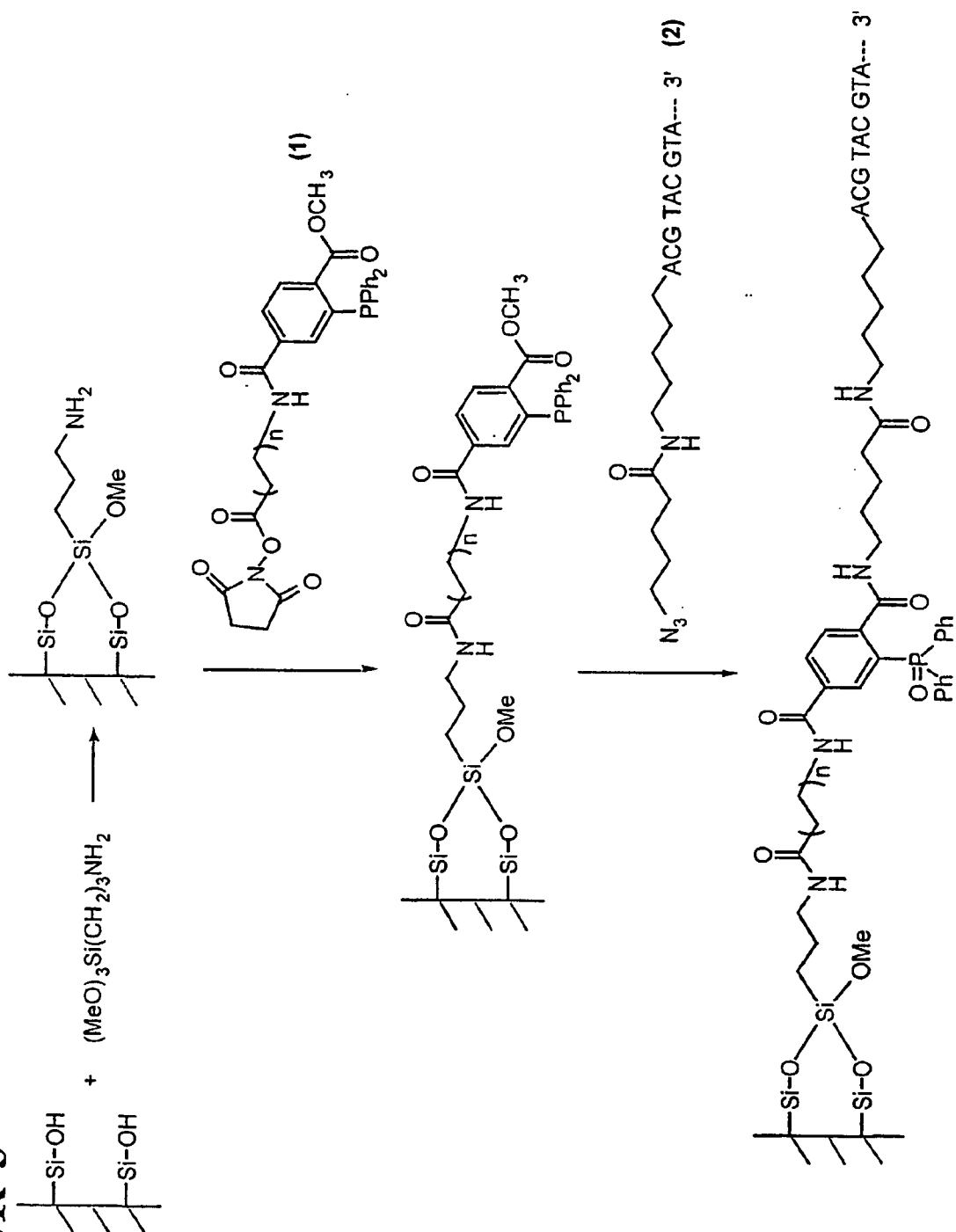




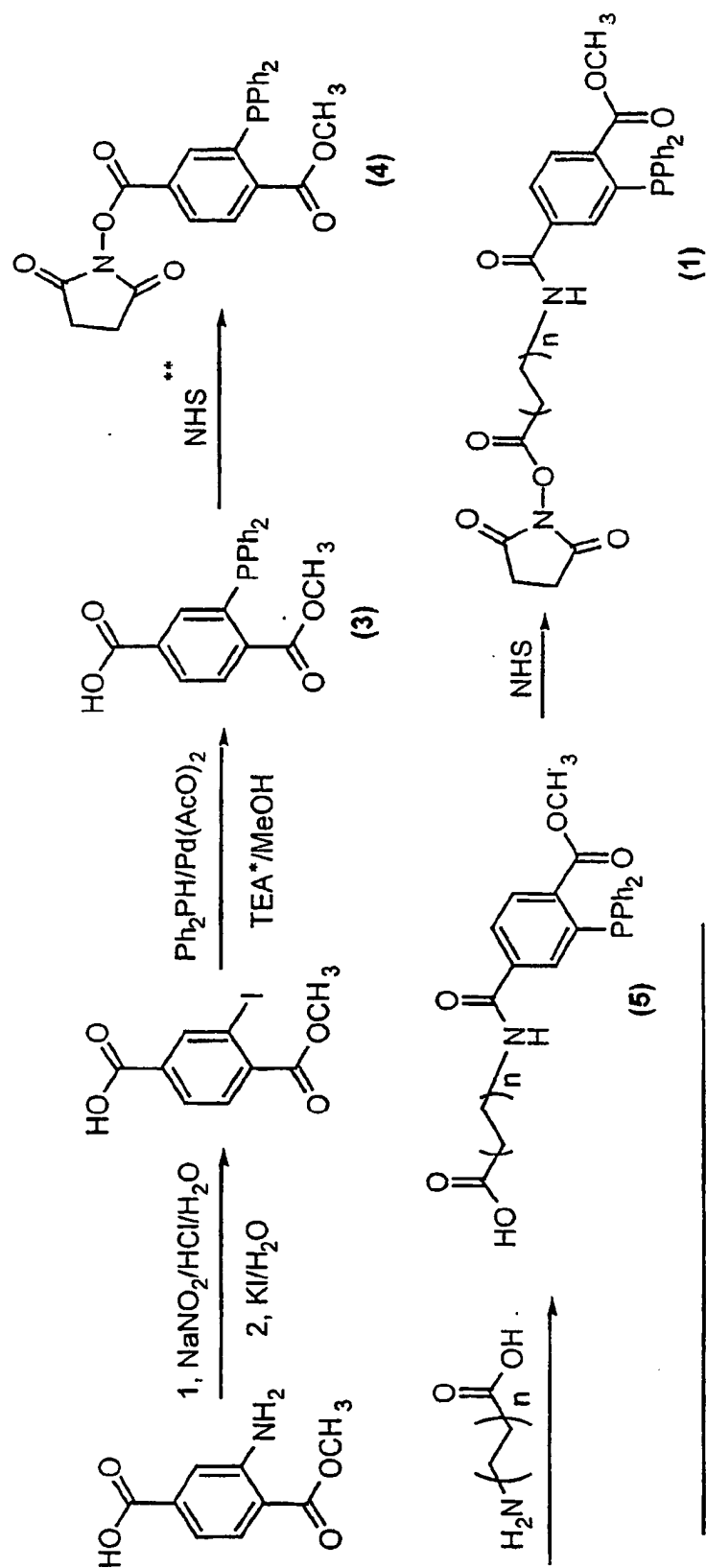
FIGUR 2B



FIGUR 3

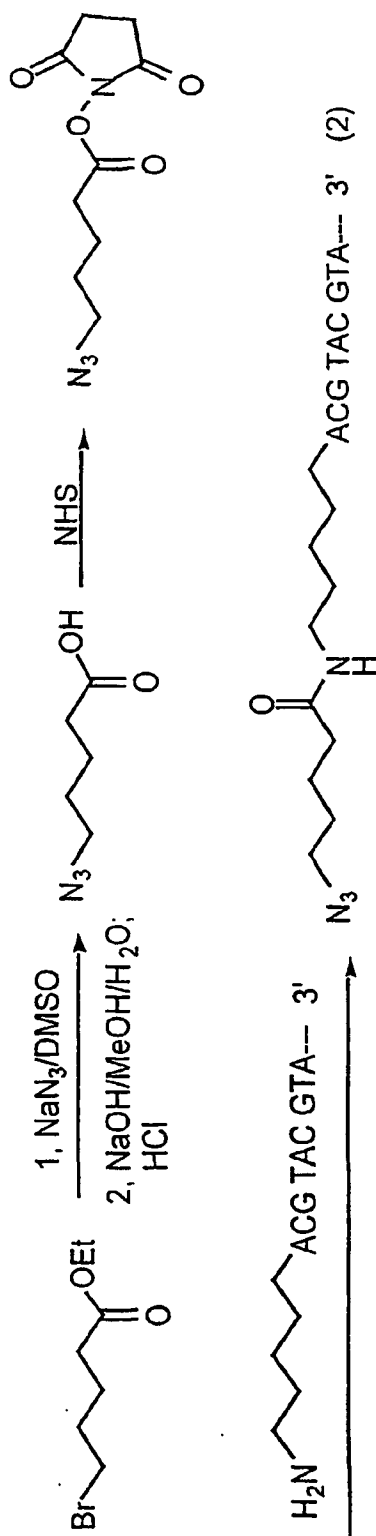


FIGUR 4

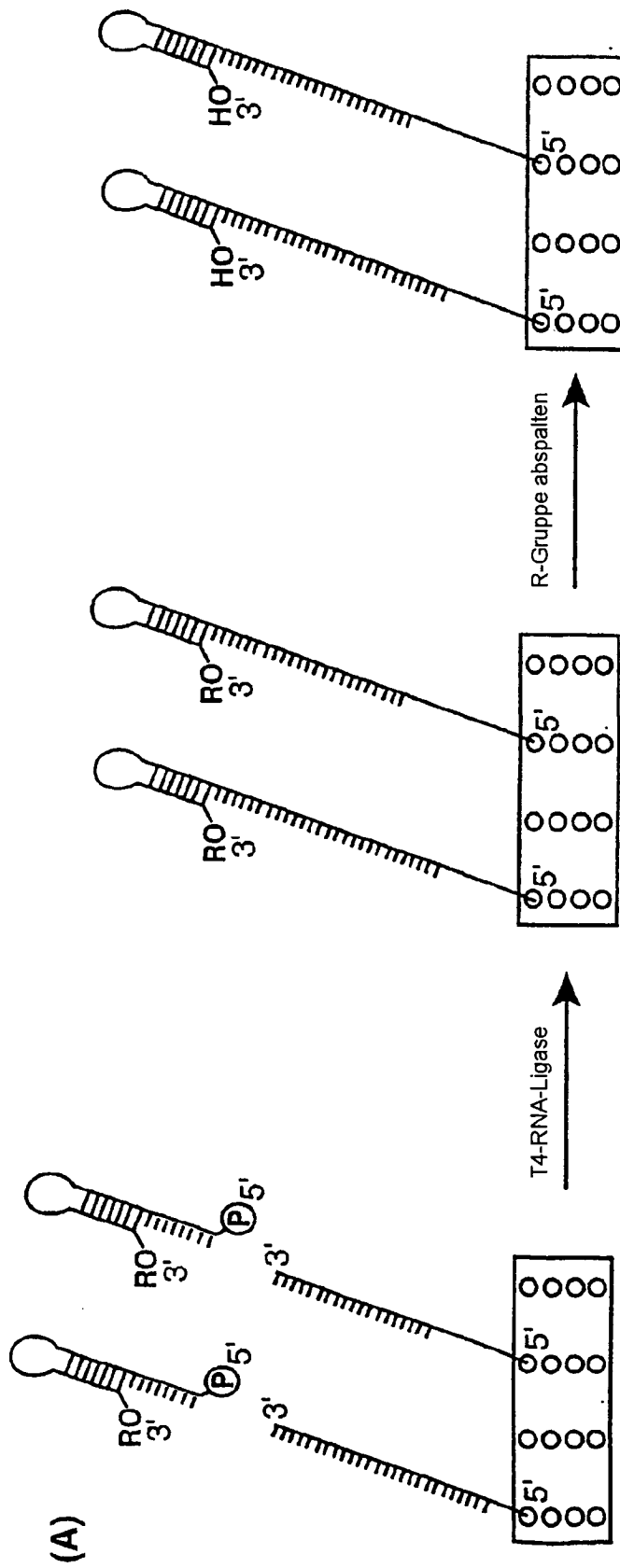


TEA = Triethylamin, NHS = N-Hydroxysuccinimid

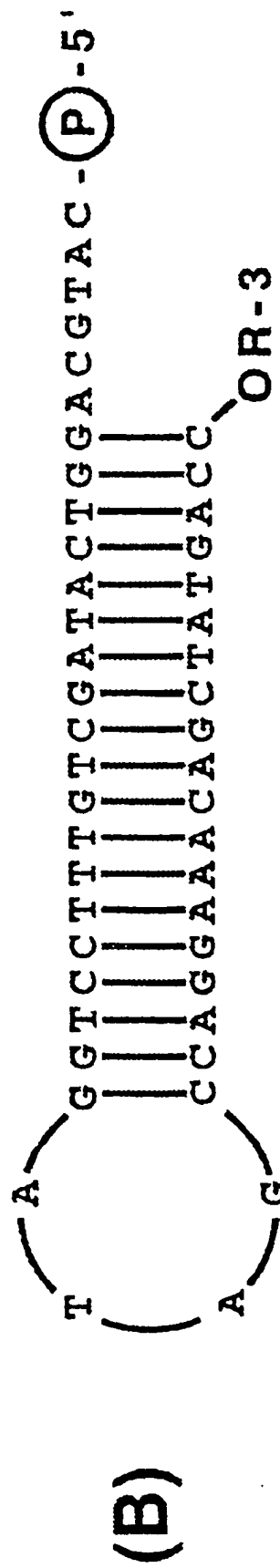
FIGUR 5



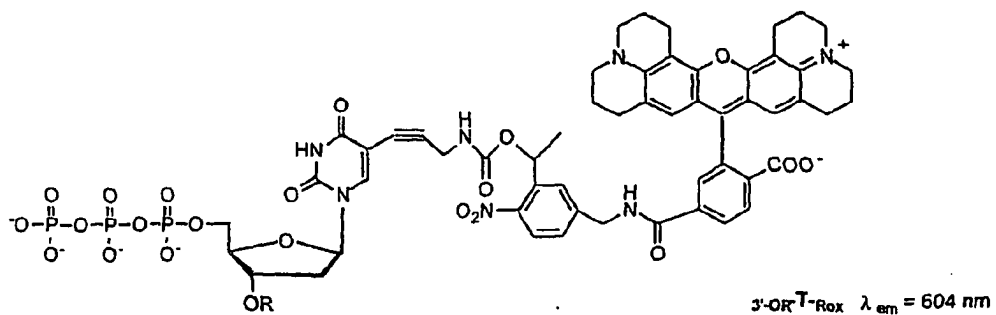
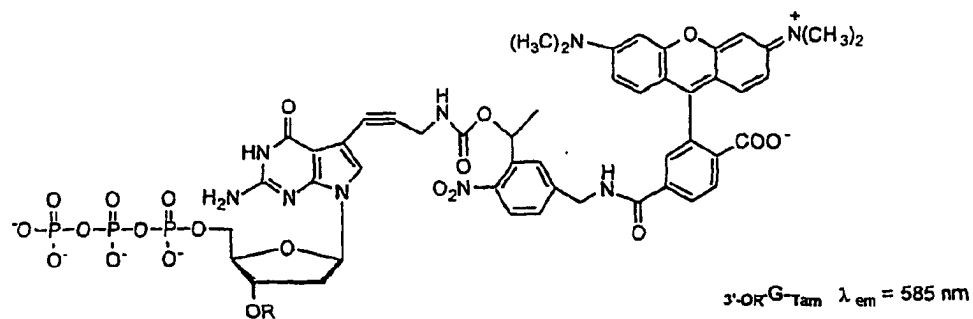
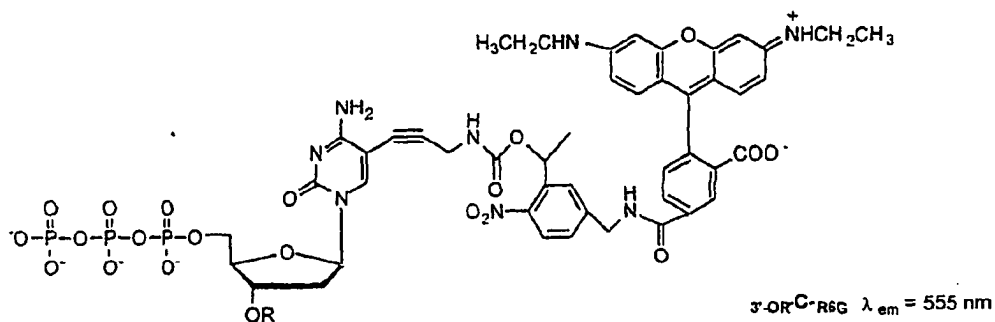
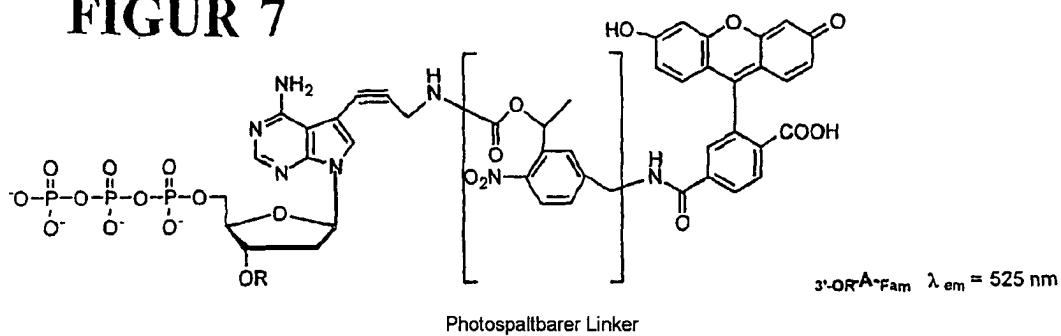
FIGUR 6A



FIGUR 6B

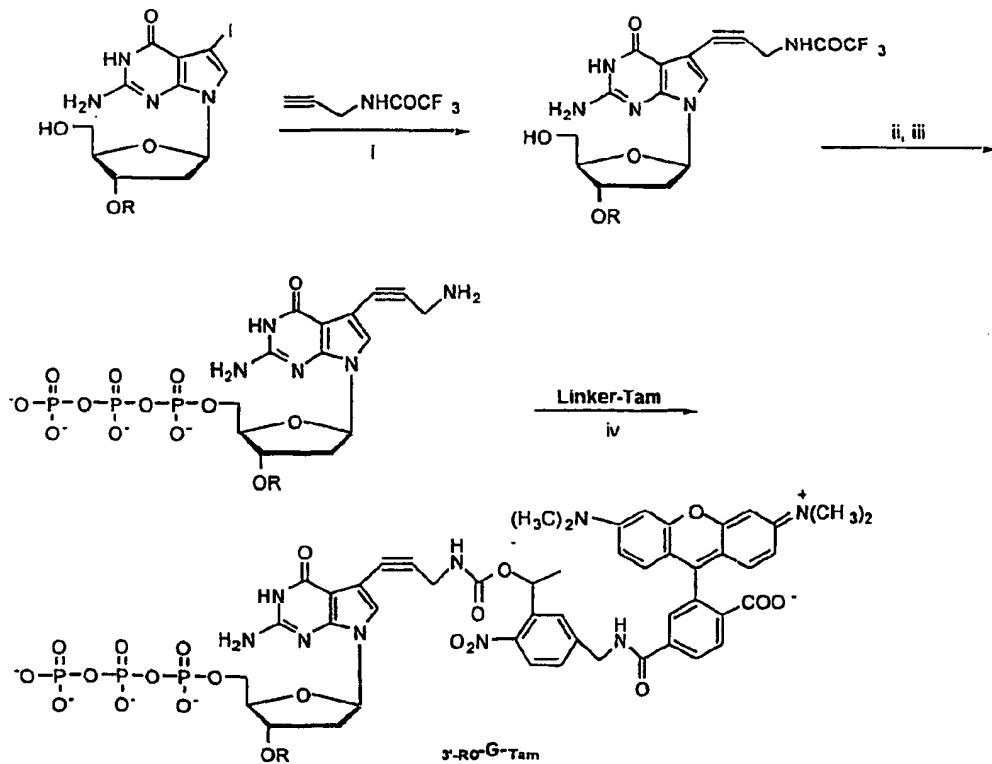
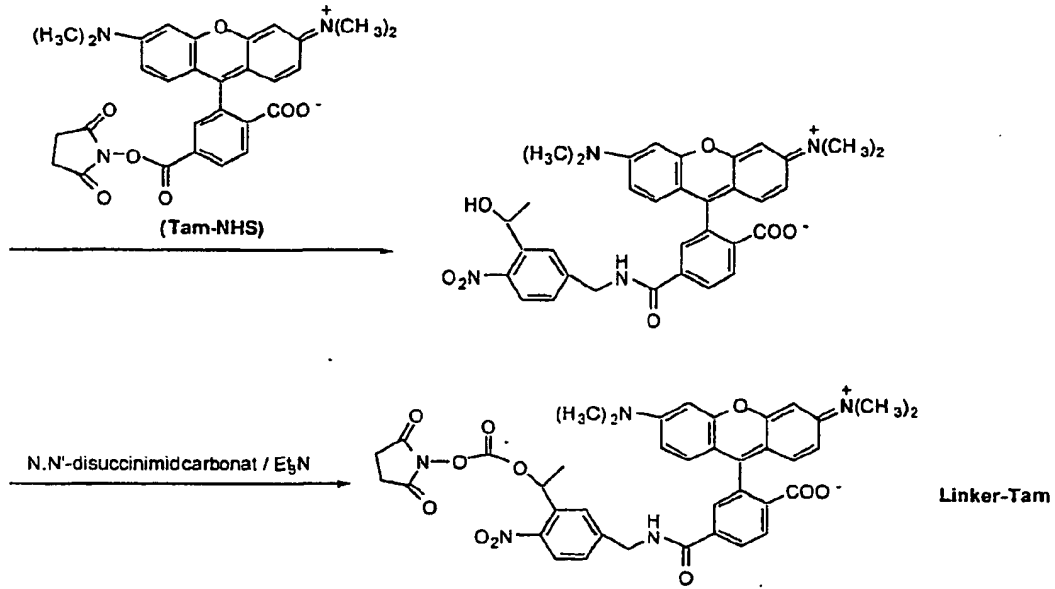
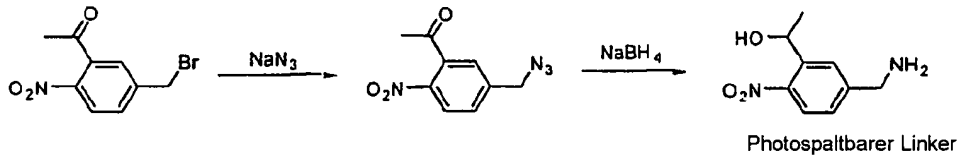


FIGUR 7

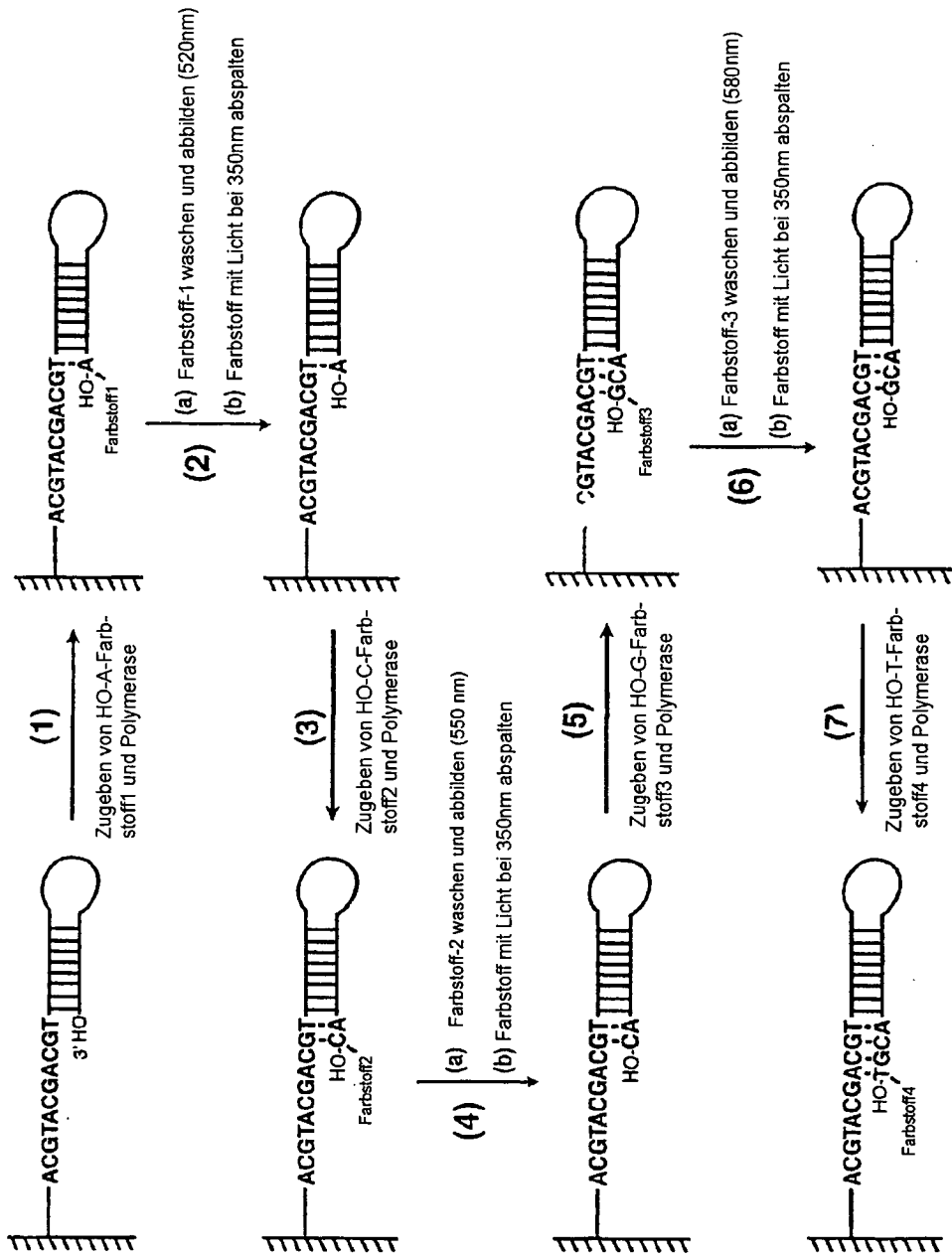


R = H, CH₂OCH₃ (MOM) oder CH₂-CH = CH₂ (Allyl)

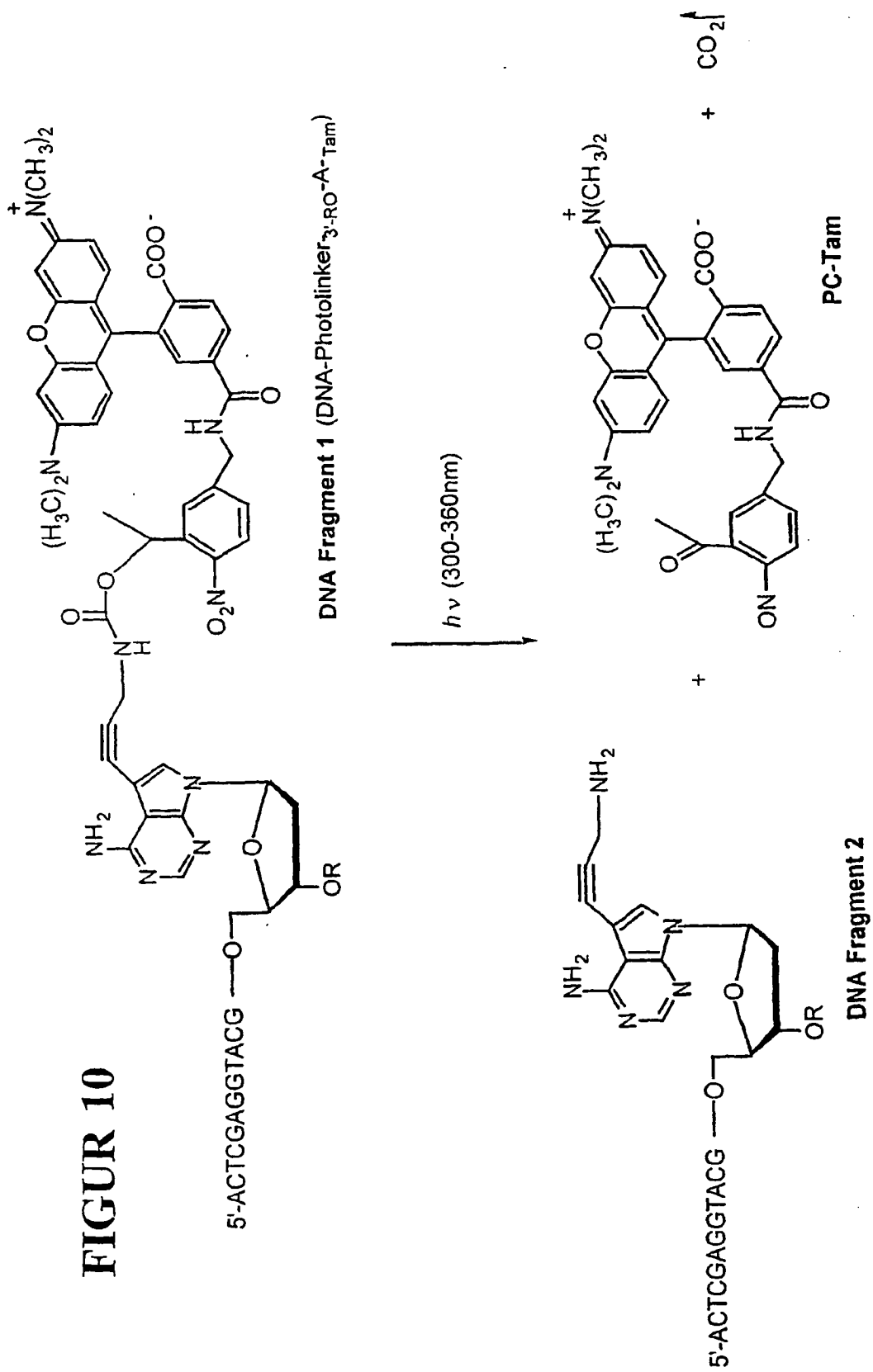
FIGUR 8

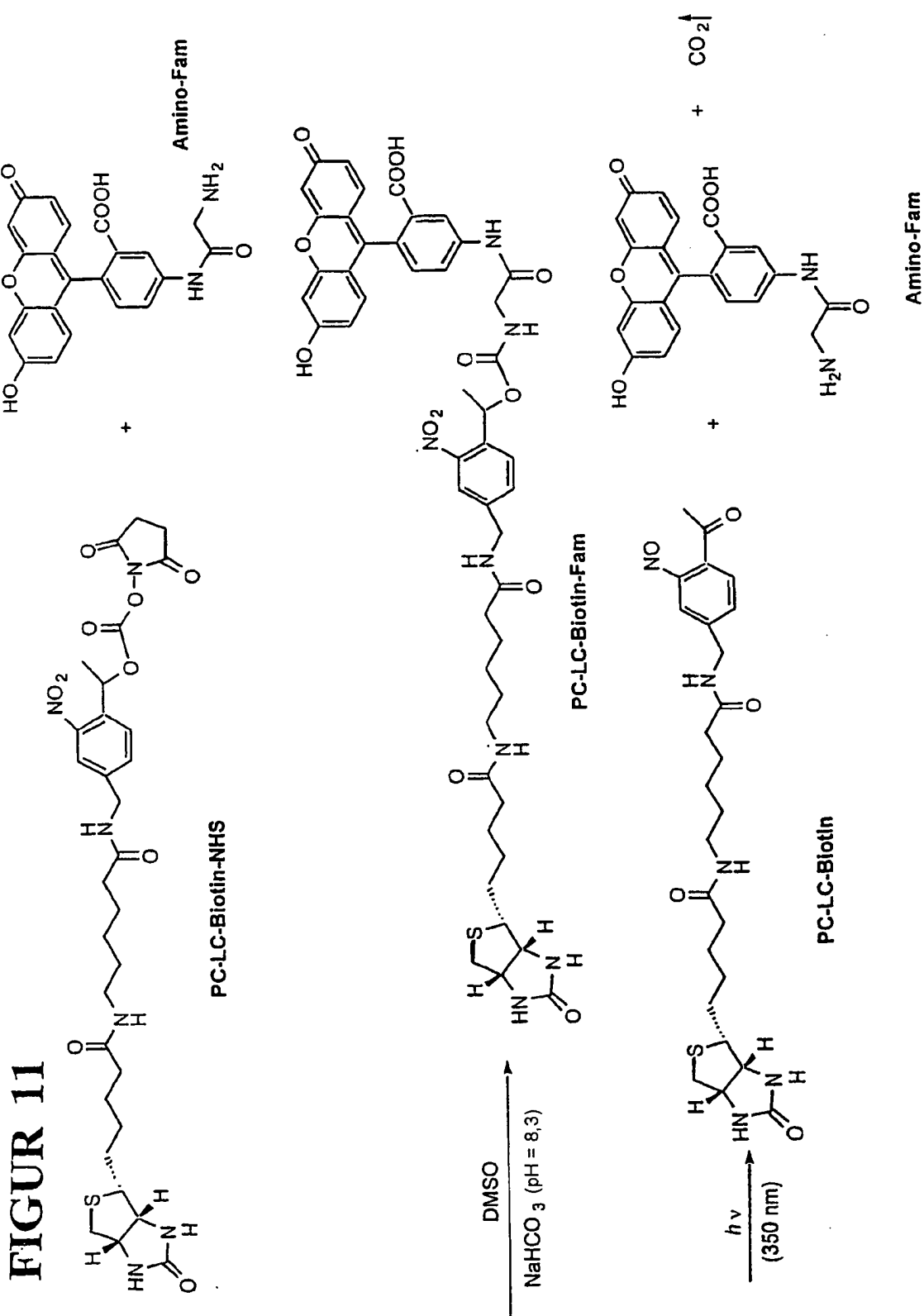


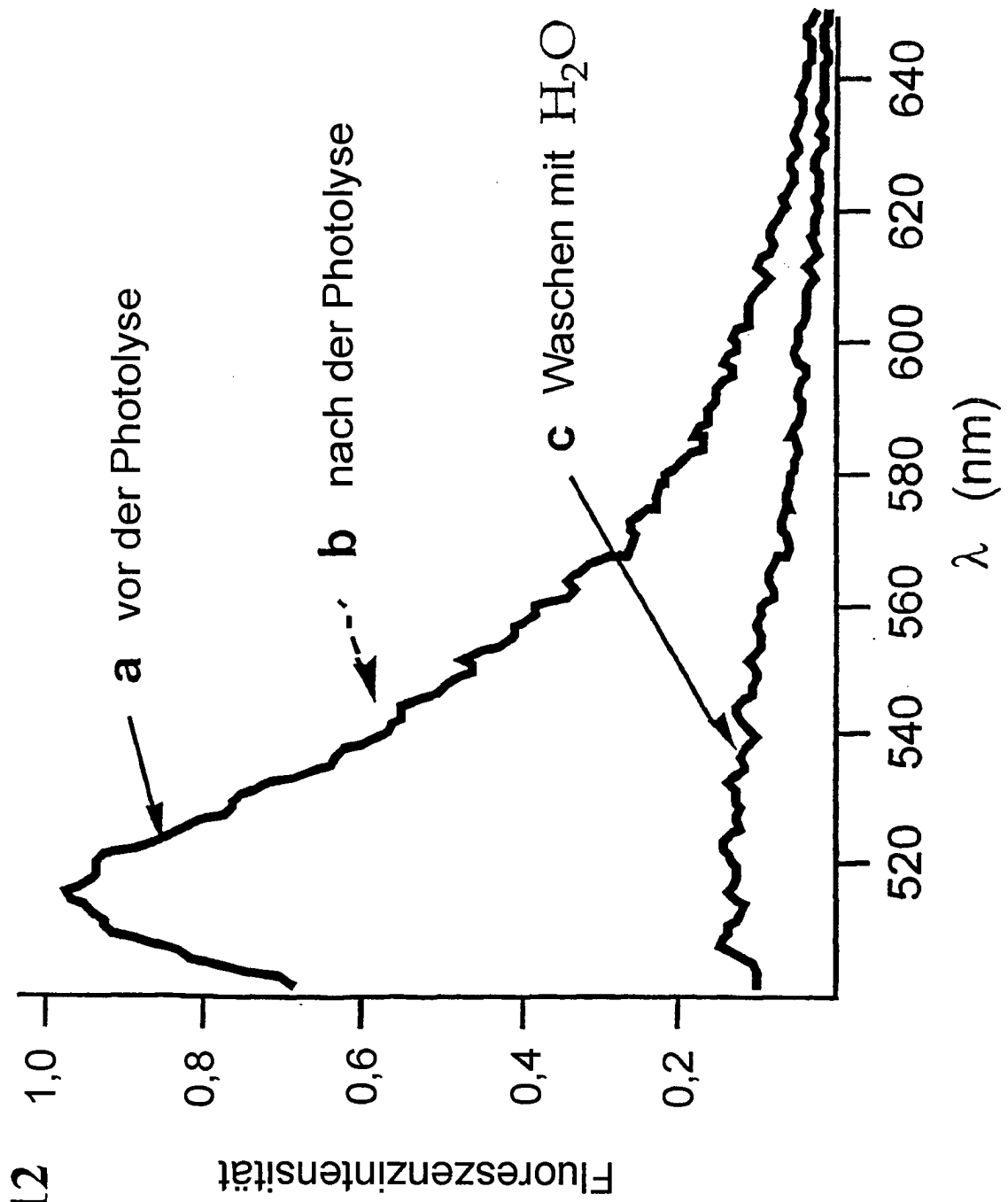
FIGUR 9



FIGUR 10

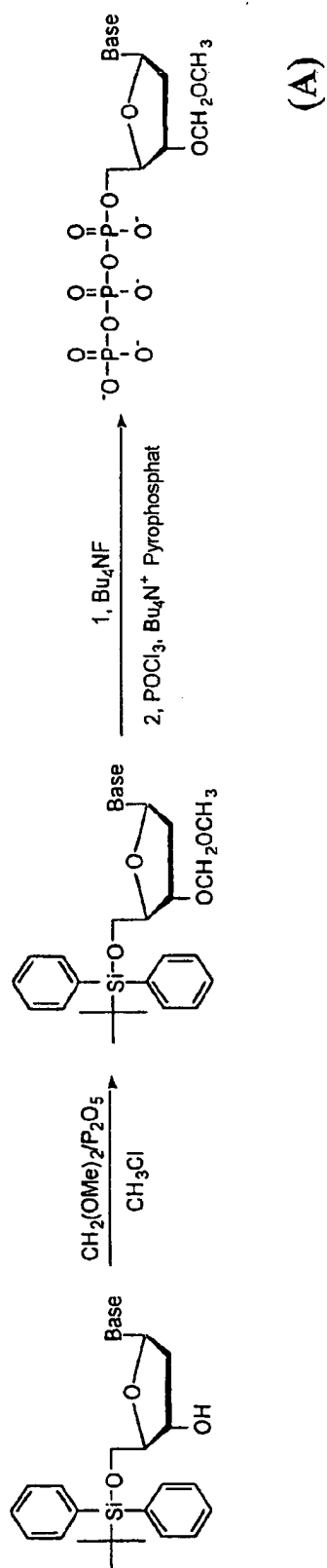




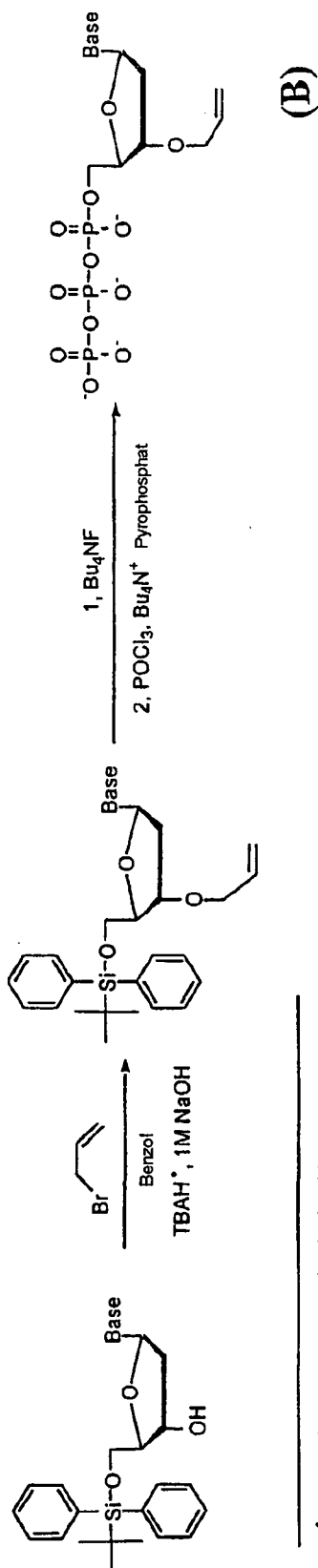


FIGUR 12

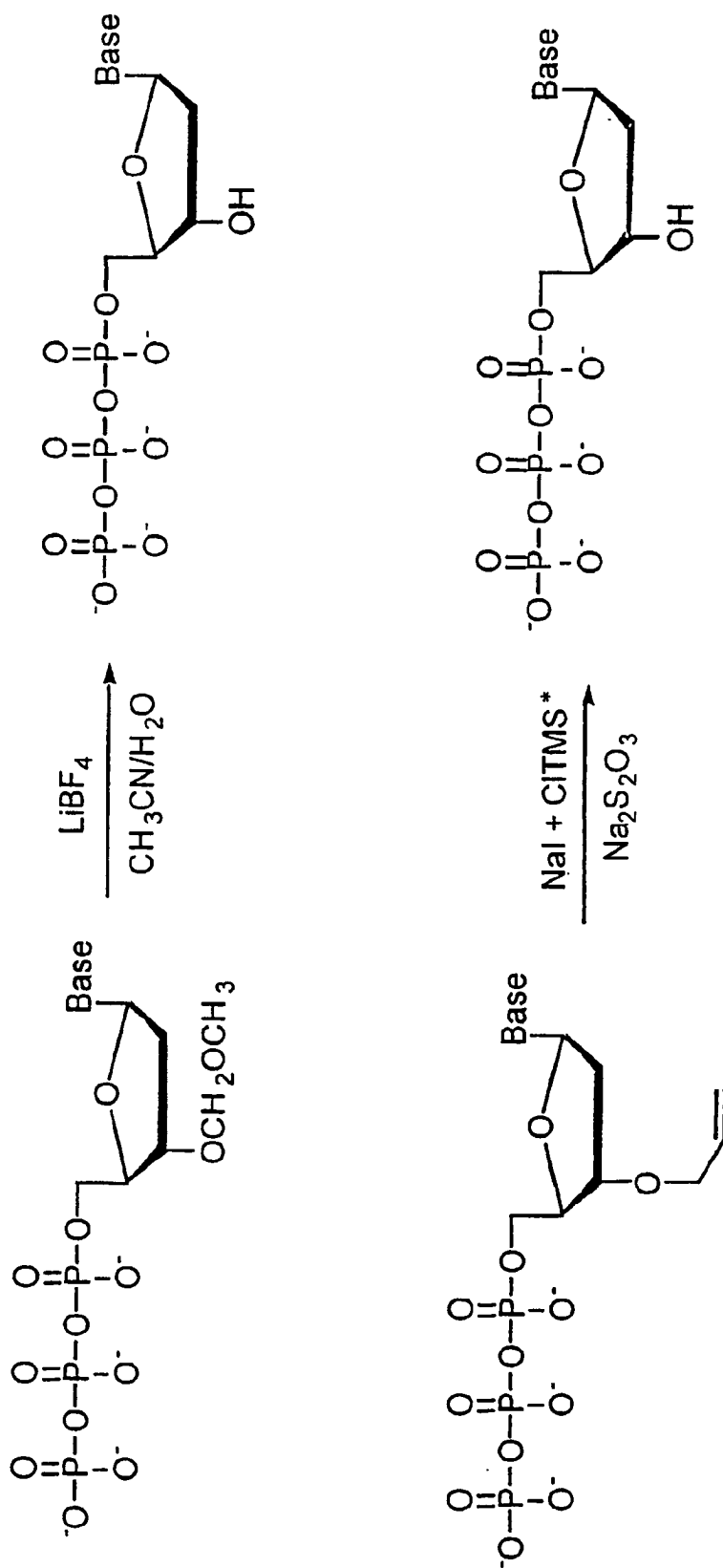
FIGUR 13A



FIGUR 13B

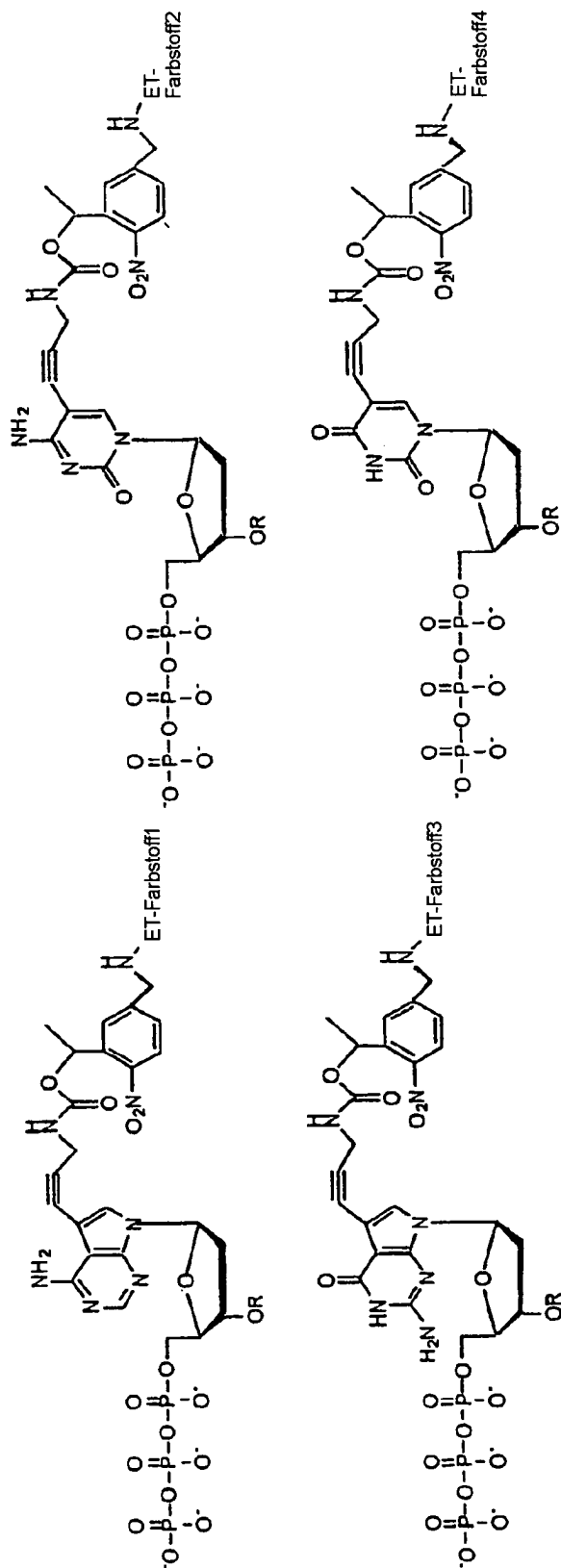


FIGUR 14



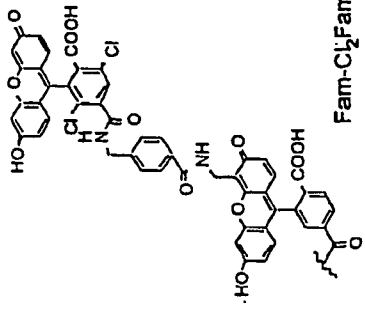
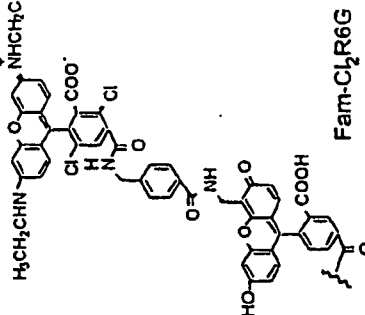
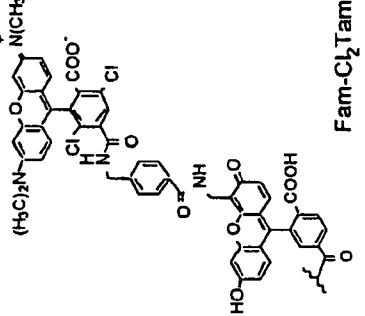
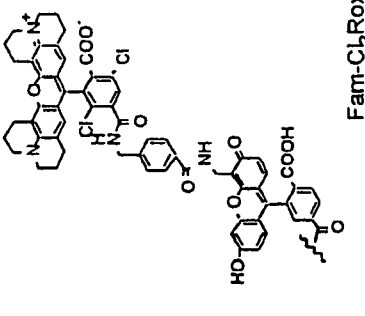
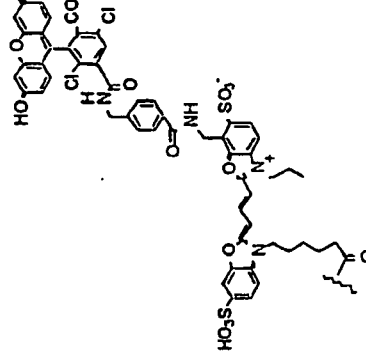
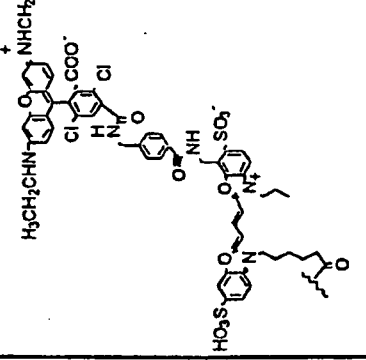
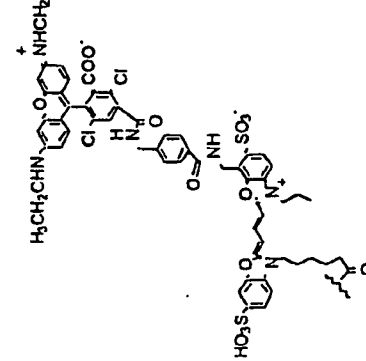
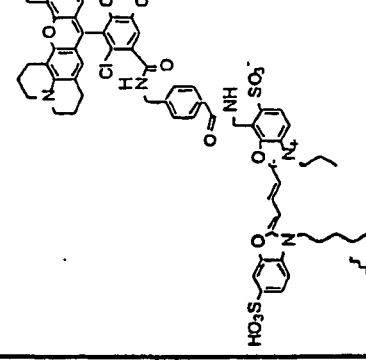
*CITMS = Chlortrimethylsilan

FIGUR 15A

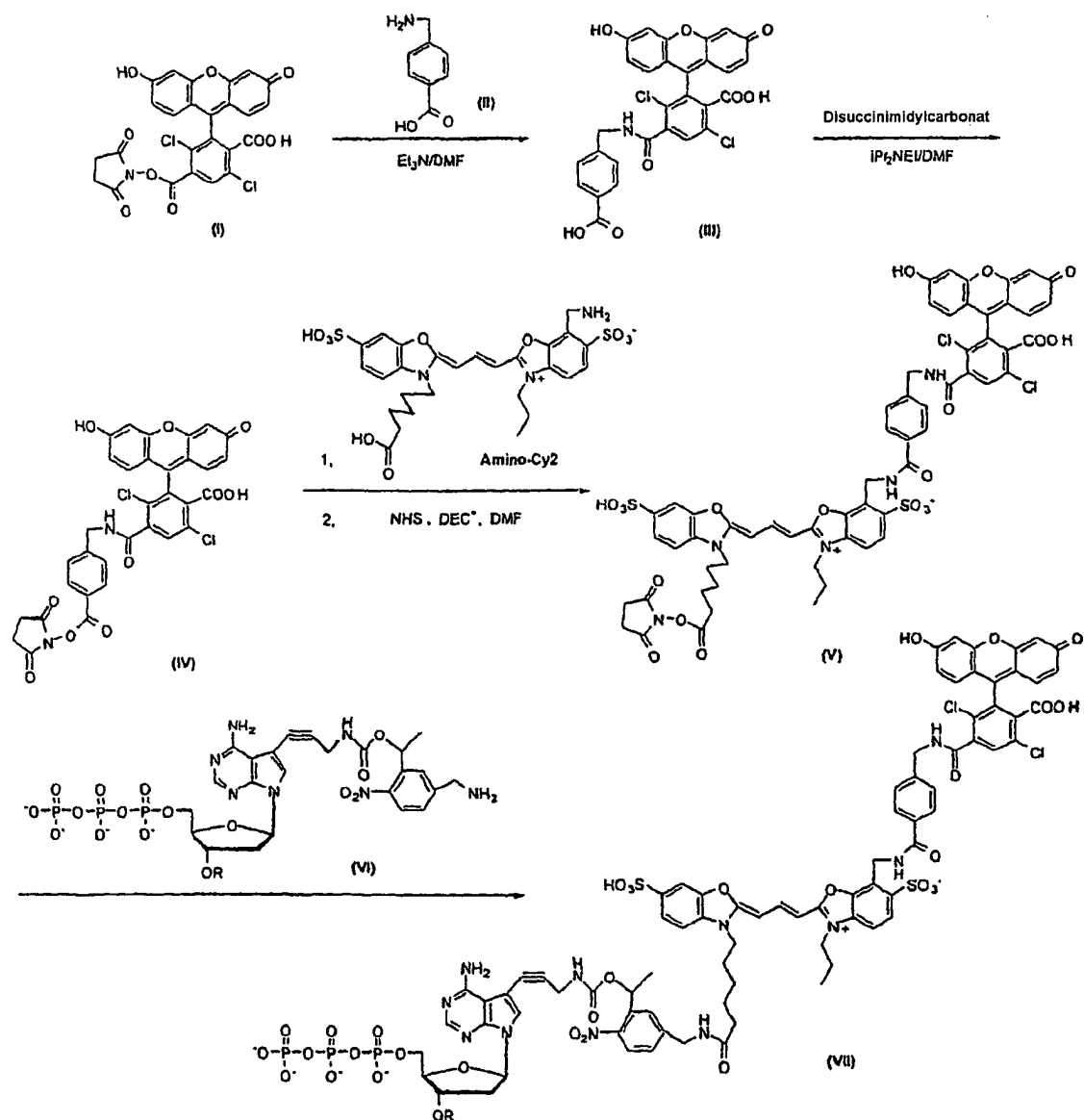


R = H, CH₂OCH₃ (MOM) oder CH₂-CH=CH₂ (Allyl)

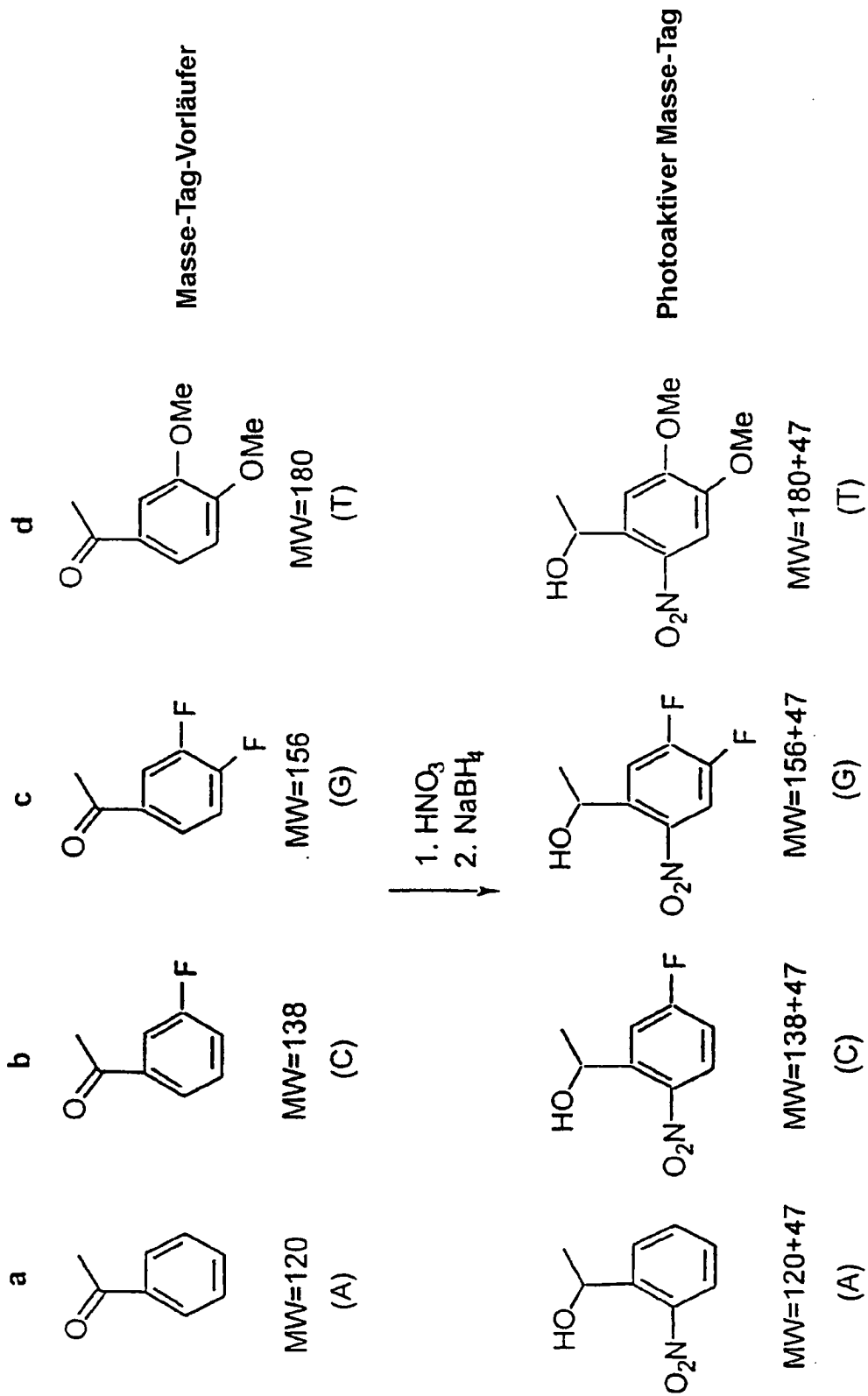
FIGUR 15B

ET Farbstoff ₁ $\lambda_{em} = 530 \text{ nm}$	ET Farbstoff ₂ $\lambda_{em} = 560 \text{ nm}$	ET Farbstoff ₃ $\lambda_{em} = 590 \text{ nm}$	ET Farbstoff ₄ $\lambda_{em} = 620 \text{ nm}$
 <p>Fam-C₁₂Fam</p>	 <p>Fam-C₁₂R6G</p>	 <p>Fam-C₁₂Tam</p>	 <p>Fam-C₁₂Rox</p>
 <p>Cy2-C₁₂Fam</p>	 <p>Cy2-C₁₂R6G</p>	 <p>Cy2-C₁₂Tam</p>	 <p>Cy2-C₁₂Rox</p>

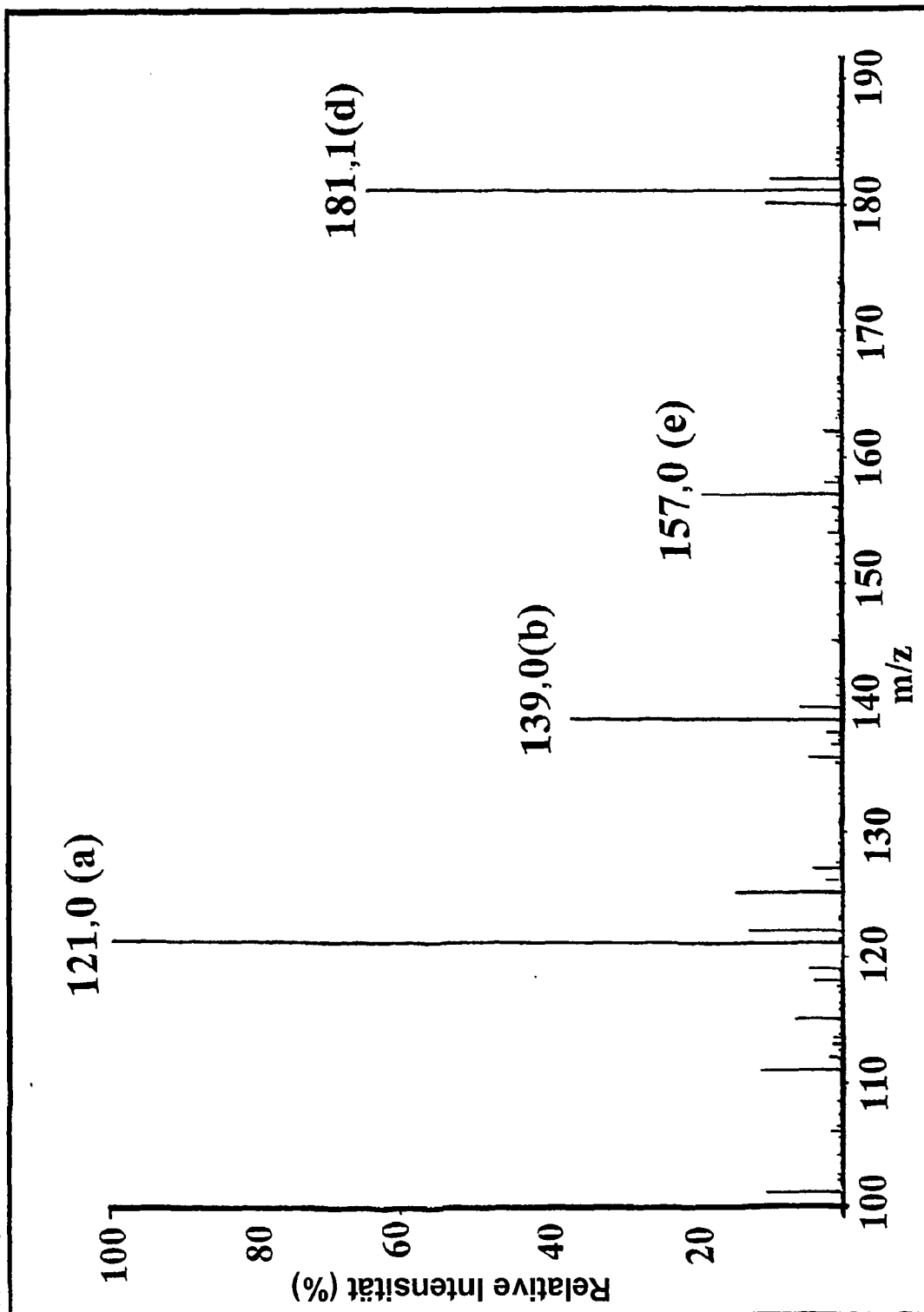
FIGUR 16



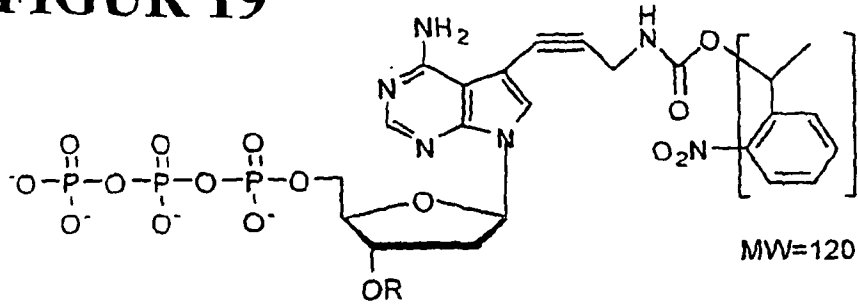
FIGUR 17



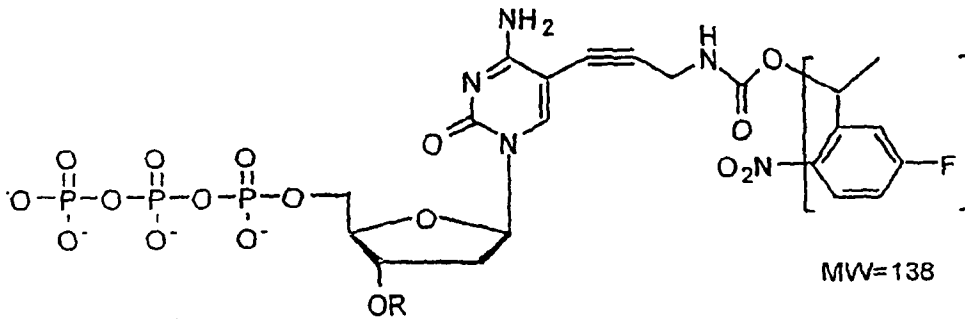
FIGUR 18



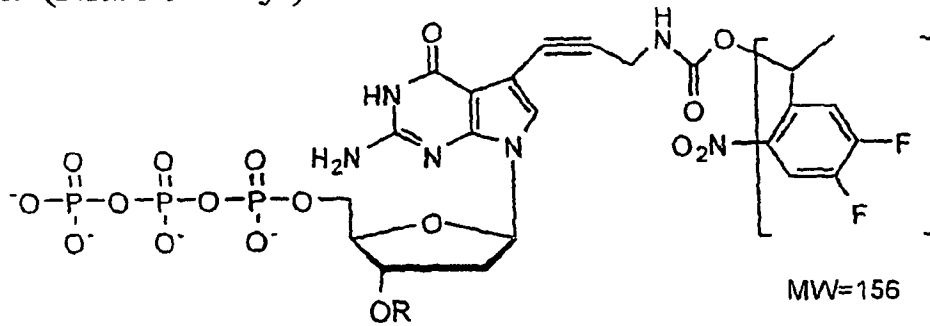
FIGUR 19



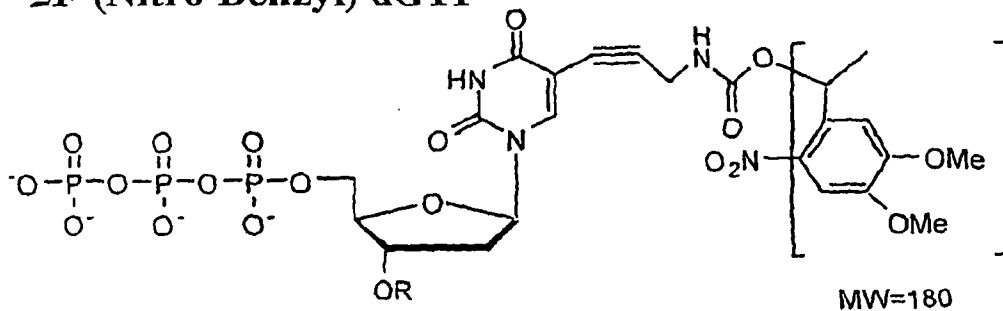
(Nitro-Benzyl)-dATP



F-(Nitro-Benzyl)-dCTP



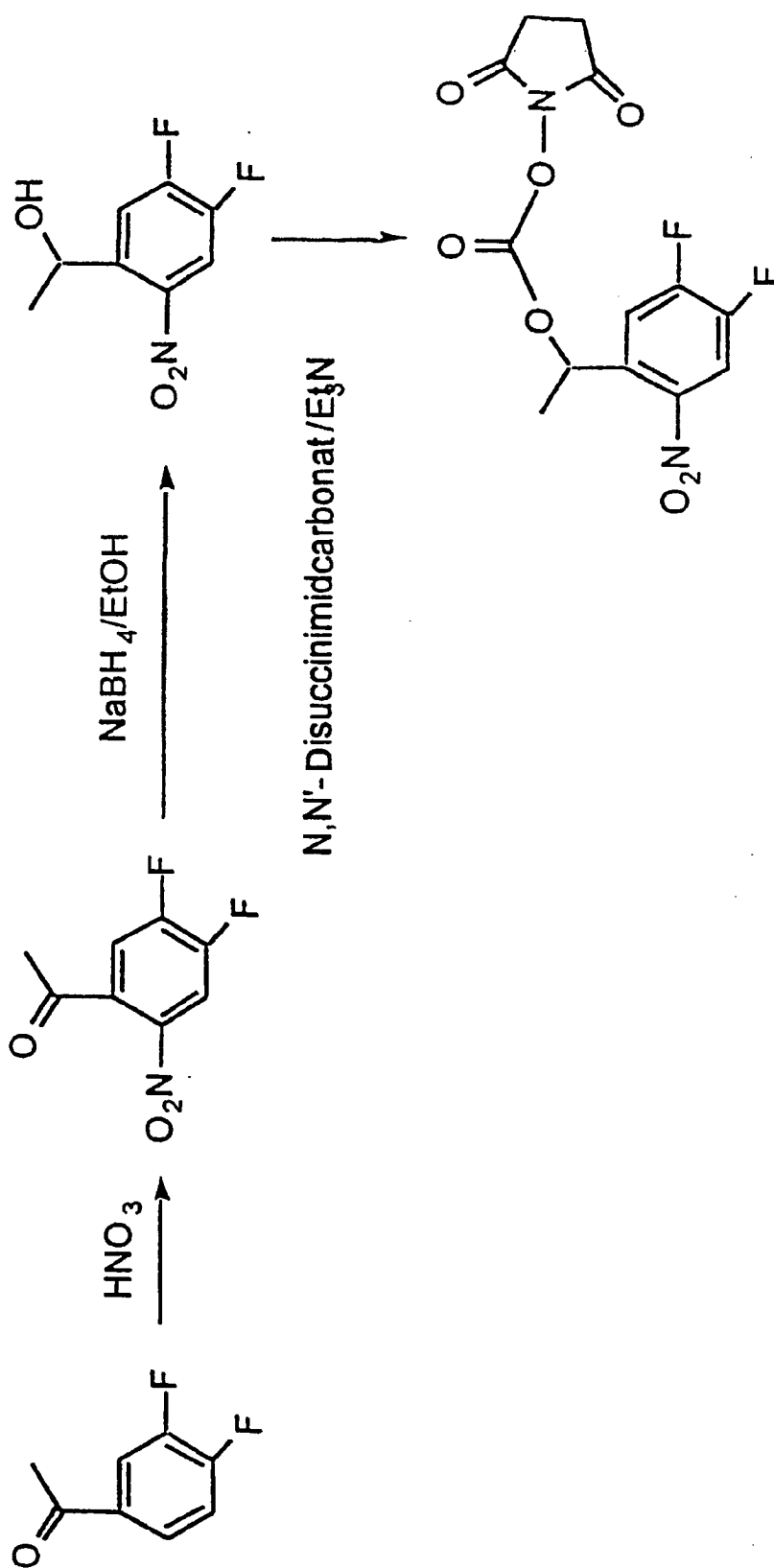
2F-(Nitro-Benzyl)-dGTP



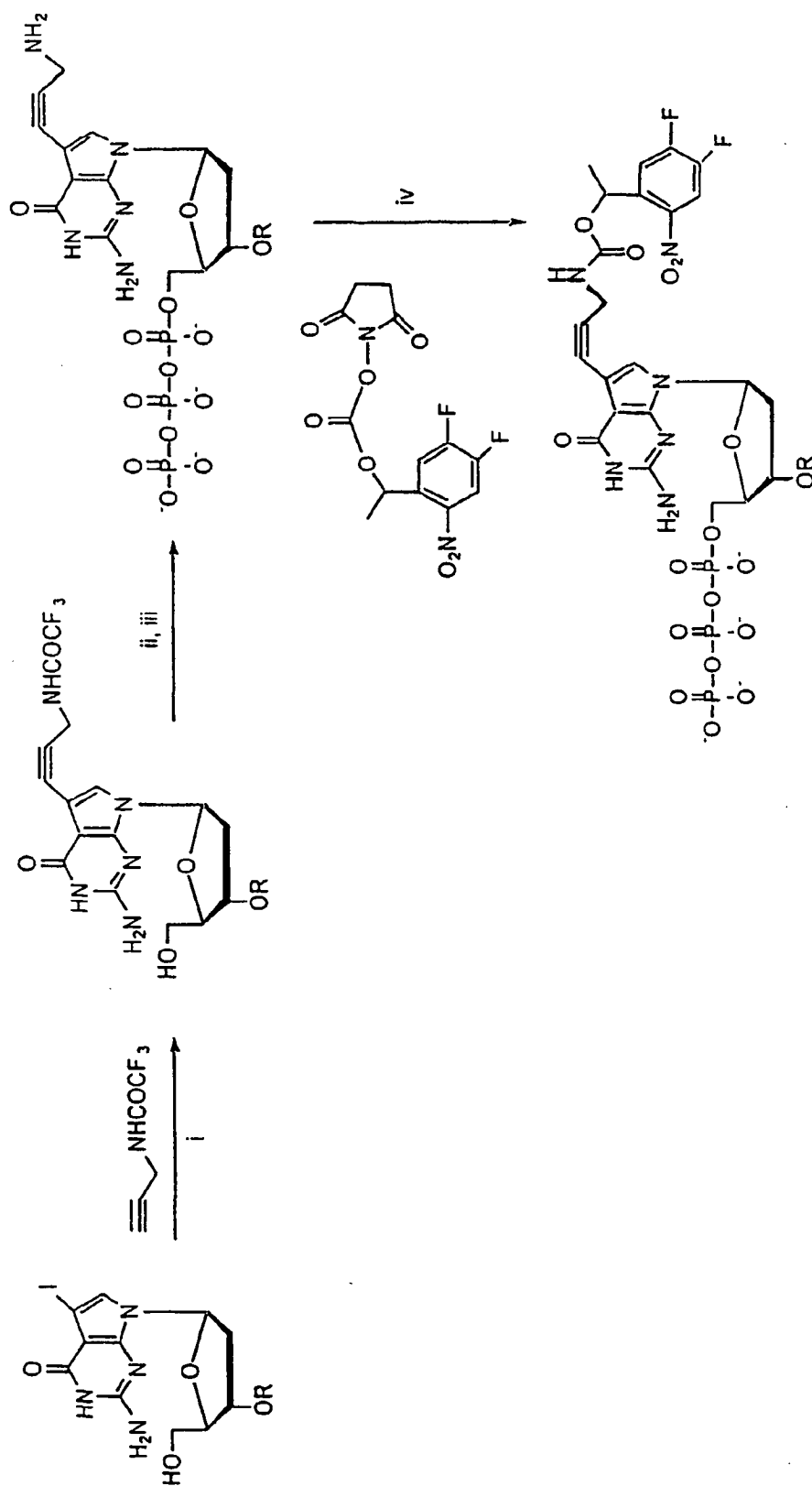
2(Meo)-(Nitro-Benzyl)-dTTP

R = H, MOM oder Allyl

FIGUR 20



FIGUR 21



FIGUR 22

