(19) **Bundesrepublik Deutschland**
**Deutsches Patent- und Markenamt**

(10) **DE 694 33 811 T2** 2005.06.23

(12) # Übersetzung der europäischen Patentschrift

(97) **EP 0 679 196 B1**
(21) Deutsches Aktenzeichen: **694 33 811.7**
(86) PCT-Aktenzeichen: **PCT/US94/00193**
(96) Europäisches Aktenzeichen: **94 906 047.9**
(87) PCT-Veröffentlichungs-Nr.: **WO 94/016101**
(86) PCT-Anmeldetag: **06.01.1994**
(87) Veröffentlichungstag der PCT-Anmeldung: **21.07.1994**
(97) Erstveröffentlichung durch das EPA: **02.11.1995**
(97) Veröffentlichungstag der Patenterteilung beim EPA: **26.05.2004**
(47) Veröffentlichungstag im Patentblatt: **23.06.2005**

(51) Int Cl.[7]: **C12Q 1/68**

(30) Unionspriorität:
**1323** **07.01.1993** **US**

(84) Benannte Vertragsstaaten:
**AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LI, LU, MC, NL, PT, SE**

(73) Patentinhaber:
**Sequenom, Inc., San Diego, Calif., US**

(72) Erfinder:
**Köster, Hubert, Concord, US**

(74) Vertreter:
**Weickmann & Weickmann, 81679 München**

(54) Bezeichnung: **DNS - SEQUENZIERUNG DURCH MASSENSPEKTRONOMIE**

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

Hintergrund der Erfindung

**[0001]** Da die genetische Information durch die Sequenz der vier DNS-Bausteine Desoxyadenosin- (dpA), Desoxyguanosin- (dpG), Desoxycytidin- (dpC) und Desoxythymidin-5'-phosphat (dpT) repräsentiert ist, ist DNS-Sequenzierung eine der grundlegendsten Technologien in der Molekularbiologie und in den Lebenswissenschaften im Allgemeinen. Die Leichtigkeit und die Geschwindigkeit, mit der DNS-Sequenzen erhalten werden können, beeinflusst stark darauf bezogene Technologien wie z. B. die Entwicklung und Herstellung neuer therapeutischer Mittel und neuer und nützlicher Varietäten von Pflanzen und Mikroorganismen durch rekombinante DNS-Technologie. Insbesondere hilft die Enträtselung der DNS-Sequenz beim Verständnis menschlicher pathologischer Zustände, einschließlich genetischer Erkrankungen, Krebs und AIDS. In einigen Fällen können sehr feine Unterschiede wie z. B. eine Deletion, Hinzufügung oder Substitution eines Nukleotids schwere, in einigen Fällen sogar fatale Konsequenzen schaffen. Vor kurzem wurde DNS-Sequenzierung die Kerntechnologie des Humanen Genomsequenzierungsprojektes (z. B. J. E. Bishop und M. Waldholz, 1991, Genome: The Story of the Most Astonishing Scientific Adventure of Our Time – The Attempt to Map All the Genes in the Human Body, Simon & Schuster, New York). Die Kenntnis der vollständigen humanen genomischen DNS-Sequenz hilft sicherlich, menschliche Krankheiten zu verstehen, zu diagnostizieren, ihnen vorzubeugen und sie zu behandeln. Um in der Lage zu sein, die Bestimmung der etwa 3 Milliarden Basenpaare des menschlichem Genoms in einem vernünftigen Zeitrahmen und auf ökonomische Weise in Angriff zu nehmen, müssen schnelle, zuverlässige, empfindliche und billige Methoden entwickelt werden, die ebenso die Möglichkeit zur Automation eröffnen. Die vorliegende Erfindung stellt eine solche Technologie bereit.

**[0002]** Neuere Zusammenfassungen der heutigen Methoden zusammen mit zukünftigen Richtungen und Trends wurden von Barrell (The FASEB Journal 5, 40–45 (1991)) und Trainor (Anal. Chem. 62. 418–26, (1990)) verfasst.

**[0003]** Gegenwärtig wird DNS-Sequenzierung entweder durch das chemische Abbauverfahren von Maxam und Gilbert (Methods in Enzymology 65, 499–560 (1980)) oder durch das enzymatische Didesoxynukleotid-Terminationsverfahren von Sanger et al. (Proc., Natl. Acad. Sci. USA 74, 5463–67 (1977)) durchgeführt. Im chemischen Verfahren führen basenspezifische Modifikationen zu einer basenspezifischen Spaltung des radioaktiv- oder fluoreszenzmarkierten DNS-Fragments. Mit den vier getrennten basenspezifischen Spaltungsreaktionen werden vier Sätze verschachtelter Fragmente hergestellt, die nach der Länge durch Polyacrylamidgelelektrophorese (PAGE) getrennt werden. Nach Autoradiographie kann die Sequenz direkt gelesen werden, da jedes Band (Fragment) im Gel aus einem basenspezifischen Spaltungsereignis herrührt. Daher werden die Fragmentlängen der vier "Leitern" direkt in eine spezifische Position in der DNS-Sequenz übersetzt.

**[0004]** Im enzymatischen Kettenterminationsverfahren werden die vier basenspezifischen Sätze von DNS-Fragmenten gebildet, indem mit einem Primer/Template-System begonnen wird, das den Primer in die unbekannte DNS-Sequenzregion verlängert und dabei das Template kopiert und einen komplementären Strang in Anwesenheit von kettenterminierenden Reagenzien durch DNS-Polymerasen synthetisiert, wie z. B. durch das Klenow-Fragment von E. coli-DNS-Polymerase I, eine DNS-Polymerase aus Thermus aquaticus, Taq-DNS-Polymerase oder eine modifizierte T7-DNS-Polymerase, Sequenase (Tabor et al., Proc. Natl. Acad. Sci. USA 84, 4767–4771 (1987)). Hier wird das kettenterminierende Ereignis durch Einschließen von nur einem der kettenterminierenden Didesoxynukleosidtriphosphate ddATP, ddGTP, ddTTP bzw. ddCTP in einer begrenzten kleinen Konzentration in die vier getrennten Reaktionsmischungen zusätzlich zu den vier normalen Desoxynukleosidtriphosphaten dATP, dGTP, dTTP und dCTP erreicht. Die vier Sätze der resultierenden Fragmente bilden nach Elektrophorese vier basenspezifische Leitern, aus denen die DNS-Sequenz bestimmt werden kann.

**[0005]** Eine neuere Modifikation der Sequenzierungsstrategie nach Sanger involviert den Abbau von Phosphorthioat enthaltenden DNS-Fragmenten, die durch Verwendung von alpha-Thio-dNTP anstatt der normalerweise verwendeten ddNTPs während der durch eine DNS-Polymerase vermittelten Primerverlängerungsreaktion erhalten werden (Labeit et al., DNA 5, 173–177 (1986); Amersham, PCT-Anmeldung GB86/00349; Eckstein et al., Nucleic Acids Res, 16, 9947 (1988)). Hier werden die vier Sätze basenspezifischer Sequenzierungsleitern durch beschränkten Verdau mit Exonuklease III oder Schlangengiftphosphodiesterase, durch nachfolgende Trennung auf PAGE und Sichtbarmachung durch Radioisotopenmarkierung von entweder dem Primer oder von einem der dNTPs erhalten. In einer weiteren Modifikation wird die basenspezifische Spaltung durch Alkylieren des Schwefelatoms in der modifizierten Phosphodiesterbindung, gefolgt von einer Wärmebehandlung, erzielt (Max-Planck-Gesellschaft DE 39 30 312 A1). Beide Methoden können mit der Amplifizierung

der DNS durch die Polymerasekettenreaktion (polymerase chain reaction, PCR) kombiniert werden.

**[0006]** Im Voraus muss die DNS, die sequenziert werden soll, in sequenzierbare Stücke von gegenwärtig nicht mehr als 500 bis 1000 Nukleotiden fragmentiert werden. Ausgehend von einem Genom ist dies ein Mehrschrittverfahren, das Klonierungs- und Subklonierungsschritte involviert, die verschiedene und passende Klonierungsvektoren verwenden, wie z. B. YAC, Cosmide, Plasmide und M13-Vektoren (Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989)). Schließlich werden zum Sequenzieren nach Sanger die Fragmente von etwa 500 bis 1000 Basenpaaren in eine spezifische Restriktionsstelle der replikativen Form I (RF I) eines Derivats des M13-Bakteriophagen (Viera und Messing, Gene 19, 259 (1982)) integriert, und dann wird die doppelsträngige Form in die einzelsträngige zirkuläre Form transformiert, die als Template für das Sequenzierungsverfahren nach Sanger dient, die eine Bindungsstelle für einen durch chemische DNS-Synthese (Sinha, Biernat, McManus und Köster, Nucleic Acids Res. 12, 4539–57 (1984); US-Patent Nr. 4725677) erhaltenen universellen Primer stromaufwärts der Restriktionsstelle hat, in die das unbekannte DNS-Fragment inseriert ist. Unter spezifischen Bedingungen können unbekannte DNS-Sequenzen, die in doppelsträngige superspiralisierte Plasmid-DNS integriert sind, direkt durch das Sanger-Verfahren sequenziert werden (Chen und Seeburg, DNA 4, 165–170 (1985) und Lim et al., Gene Anal. Techn. 5, 32–39 (1988)), und mit der Polymerasekettenreaktion (PCR)(PCR Protocols: A Guide to Methods and Applications, Hrsg. Innis et al., Academic Press, San Diego (1990)) könnten Klonierungs- oder Subklonierungsschritte durch direktes Sequenzieren von chromosomaler DNS weggelassen werden, indem erst die DNS-Segmente durch PCR amplifiziert werden und dann das Sequenzierungsverfahren nach Sanger angewendet wird (Innis et al., Proc. Natl. Acad. Sci. USA 85, 9436–9440 (1988)). In diesem Falle muss jedoch die DNS-Sequenz in der interessierenden Region wenigstens soweit bekannt sein, dass sie an den Sequenzierungsprimer bindet.

**[0007]** Um fähig zu sein, die Sequenz von der PAGE zu lesen, müssen detektierbare Markierungen in entweder dem Primer (sehr oft am 5'-Ende) oder in einem der Desoxynukleosidtriphosphate dNTP verwendet werden. Die Verwendung von Radioisotopen wie z. B. $^{32}P$, $^{33}P$ oder $^{35}S$ ist nach wie vor die am häufigsten verwendete Technik. Nach PAGE werden die Gele mit einem Röntgenfilm exponiert, und die Silberkornexposition wird analysiert. Die Verwendung von Radioisotopenmarkierungen schafft mehrere Probleme. Die meisten Markierungen, die für autoradiographische Detektion von Sequenzierungsfragmenten verwendbar sind, haben relativ kurze Halbwertszeiten, was die brauchbare Zeit der Markierungen beschränken kann. Die Emission hochenergetischer beta-Strahlung, insbesondere von $^{32}P$, kann zum Abbau der Produkte durch Radiolyse führen; so dass die Probe sehr schnell nach der Markierung verwendet werden sollte. Des Weiteren kann hochenergetische Strahlung ebenso eine Zerstörung der Bandenschärfe durch Streuung verursachen. Einige dieser Probleme können durch Verwendung von weniger energetischen Isotopen reduziert werden, wie z. B. $^{33}P$ oder $^{35}S$ (siehe z. B. Ornstein et al., Biotechniques 3, 476 (1985)), hier müssen jedoch längere Expositionszeiten toleriert werden. Vor allem wirft die Verwendung von Radioisotopen signifikante Gesundheitsrisiken für den Experimentator auf, und in großen Sequenzierungsprojekten sind Dekontaminierung und Handhabung des radioaktiven Mülls andere schwerwiegende Probleme und Lasten.

**[0008]** Als Antwort auf die oben erwähnten Probleme, die sich auf die Verwendung von radioaktiven Markierungen beziehen, wurden nichtradioaktive Markierungstechniken erprobt und in den vergangenen Jahren in teilweise automatisierte DNS-Sequenzierungsverfahren integriert. All diese Verbesserungen verwenden die Sequenzierungsstrategie nach Sanger. Die Fluoreszenzmarkierung kann an den Primer (Smith et al., Nature 321, 674–679 (1986) und EPO-Patent Nr. EP 0233053, Du Pont De Nemours EPO-Anmeldung Nr. 0359225; Ansorge et al., J. Biochem. Biophys. Methods 13, 325–32 (1986)) oder an die kettenterminierenden Didesoxynukleosidtriphosphate (Prober et al. Science 238, 336–41 (1987); Applied Biosystems, PCT-Anmeldung WO 91/05060) angehängt werden. Basierend auf entweder der Markierung des Primers oder des ddNTPs wurden Systeme von Applied Biosystems (Smith et al., Science 235, G89 (1987); US-Patente Nr. 570973 und 689013), Du Pont De Nemours (Prober et al. Science 238, 336–341 (1987); US-Patente Nr. 881372 und 57566), Pharmacia-LKB (Ansorge et al. Nucleic Acids Res. 15, 4593–4602 (1987) und EMBL Patentanmeldung DE P3724442 und P3805808.1) und Hitachi (JP 1-90844 und DE 40 11 991 A1) entwickelt. Ein etwas ähnlicher Ansatz wurde von Brumbaugh et al. (Proc. Natl. Sci. USA 85, 5610–14 (1988) und US-Patent Nr. 4,729,947) entwickelt. Ein verbessertes Verfahren für das Du Pont-System ist beschrieben, welches zwei elektrophoretische Spuren mit zwei verschiedenen spezifischen Markierungen pro Spur verwendet (PCT-Anmeldung WO 92/02635). Ein anderer Ansatz verwendet fluoreszenzmarkiertes Avidin und biotinmarkierte Primer. Hier reagieren die Sequenzierungsleitern, die mit Biotin enden, während der Elektrophorese mit dem markierten Avidin, was zur Detektion individueller Sequenzierungsbanden führt (Brumbaugh et al., US-Patent Nr. 594676).

**[0009]** Vor kürzester Zeit wurden noch empfindlichere nichtradioaktive Markierungstechniken für DNS entwickelt, die Chemilumineszenz verwenden, die durch Enzyme triggerbar und amplifizierbar ist (Beck, O'Keefe,

Coull und Köster, Nucleic Acids Res. 17, 5115–5123 (1989) und Beck und Köster, Anal. Chem. 62, 2258–2270 (1990)). Diese Markierungsverfahren wurden mit Multiplex-DNS-Sequenzierung kombiniert (Church et al. Science 240, 185–188 (1988), um eine Strategie bereitzustellen, die auf Hochdurchsatz-DNS-Sequenzierung zielt (Köster et al., Nucleic Acids Res. Symposium Ser. Nr. 24, 318–321 (1991), University of Utah, PCT-Anmeldung Nr. WO 90/15883); diese Strategie leidet noch am Nachteil, dass sie sehr arbeitsaufwendig und schwer zu automatisieren ist.

**[0010]** Als Versuch, die DNS-Sequenzierung zu vereinfachen, wurden feste Träger eingeführt. In den meisten bisher publizierten Fällen wird der Templatestrang zur Sequenzierung (mit oder ohne PCR-Amplifizierung) auf einem festen Träger immobilisiert, wobei am häufigsten die starke Biotin-Avidin/Streptavidin-Wechselwirkung verwendet wird (Orion-Yhtymä Oy, US-Patent Nr. 277643; M. Uhlen et al. Nucleic Acids Res. 16, 3025–38 (1988); Cemu Bioteknik, PCT-Anmeldung Nr. WO 89/09282 und Medical Research Council, GB, PCT-Anmeldung Nr. WO 92/03575). Die Primerverlängerungsprodukte, die auf dem immobilisierten Templatestrang synthetisiert werden, werden von Enzymen, anderen Sequenzierungsreagenzien und Nebenprodukten durch einen Waschschritt gereinigt und dann unter denaturierenden Bedingungen durch Lösen der Wasserstoffbindungen zwischen den Watson-Crick-Basenpaaren freigesetzt und einer PAGE-Trennung unterzogen. In einem anderen Ansatz werden die Primerverlängerungsprodukte (nicht das Template) aus der DNS-Sequenzierungsreaktion an einen festen Träger über Biotin/Avidin gebunden (Du Pont De Nemours, PCT-Anmeldung WO 91/11533). Im Gegensatz zu den oben erwähnten Verfahren wird hier die Wechselwirkung zwischen Biotin und Avidin durch Verwenden denaturierender Bedingungen (Formamid/EDTA) überwunden, um die Primerverlängerungsprodukte der Sequenzierungsreaktion vom festen Träger zur PAGE-Trennung freizusetzen. Als feste Träger werden Kügelchen (z. B. magnetische Kügelchen (Dynabeads) und Sepharose-Kügelchen), Filter, Kapillaren, Plastikstäbchen (z. B. Polystyrrolstreifen) und Mikrotiterlöcher vorgeschlagen.

**[0011]** Alle soweit diskutierten Verfahren haben einen zentralen Schritt gemeinsam: Polyacrylamidgelelektrophorese (PAGE). In vielen Fällen stellt dies ein hauptsächliches Hindernis und eine hauptsächliche Beschränkung für jede dieser Verfahren dar. Die Herstellung eines homogenen Gels durch Polymerisierung, Laden der Proben, die Elektrophorese selbst, die Detektion des Sequenzierungsmusters (z. B. durch Autoradiographie), Entfernen des Gels und Saubermachen der Glasplatten, um ein anderes Gel herzustellen, sind sehr arbeitsaufwendige und zeitraubende Verfahren. Überdies ist das ganze Verfahren fehleranfällig, schwer zu automatisieren, und es wird ein ausgebildetes und geübtes Personal benötigt, um die Reproduzierbarkeit und Zuverlässigkeit zu verbessern. Im Falle einer radioaktiven Markierung kann die Autoradiographie selbst von Stunden bis Tagen dauern. Im Falle einer Fluoreszenzmarkierung kann wenigstens die Detektion der Sequenzierungsbanden automatisch durchgeführt werden, wenn Vorrichtungen zum Laserscannen verwendet werden, die in kommerziell verfügbare DNS-Sequenzierer integriert sind. Ein Problem, das mit der Fluoreszenzmarkierung verbunden ist, ist der Einfluss der vier verschiedenen basenspezifischen Fluoreszenz-Tags auf die Mobilität der Fragmente während der Elektrophorese und eine mögliche Überlappung bei der spektralen Bandbreite der vier spezifischen Farbstoffe, was die Unterscheidungskraft zwischen benachbarten Banden reduziert und dadurch die Wahrscheinlichkeit von Sequenzunklarheiten erhöht. Artefakte werden ebenso durch basenspezifische Wechselwirkungen mit der Polyacrylamidgelmatrix (Frank und Köster, Nucleic Acids Res. 6, 2069 (1979)) und durch die Bildung von Sekundärstrukturen produziert, was zu "Bandenkompressionen" führt und es daher nicht erlaubt, die Sequenz zu lesen. Dieses Problem wurde teilweise durch Verwendung von 7-Desazadesoxyguanosintriphosphaten (Barr et al., Biotechniques 4, 428 (1986)) überwunden. Jedoch werden die Gründe für einige Artefakte und auffallende Banden noch untersucht und bedürfen weiterer Verbesserung des gelelektrophoretischen Verfahrens.

**[0012]** Eine neuere Innovation bei der Elektrophorese ist die Kapillarzonenelektrophorese (capillary zone electrophoresis, CZE)(Jorgenson et al., J. Chromatography 352, 337 (1986); Gesteland et al., Nucleic Acids Res. 18, 1415–1419 (1990)), die im Vergleich mit der Schichtgelelektrophorese (PAGE) die Auflösung der Trennung signifikant erhöht, die Zeit für einen elektrophoretischen Lauf reduziert und die Analyse sehr kleiner Proben erlaubt. Hier entstehen jedoch andere Probleme aufgrund der Miniaturisierung des ganzen Systems, wie z. B. Wandeffekte und die Notwendigkeit eines hochempfindlichen Online-Detektionsverfahrens. Verglichen mit PAGE wird ein anderes Hindernis durch die Tatsache geschaffen, dass CZE nur ein "Einspur"-Verfahren ist, während PAGE Proben in vielen Spuren simultan elektrophoretisiert werden können.

**[0013]** Aufgrund der schweren Beschränkungen und Probleme in Bezug auf PAGE als ein integraler und zentraler Teil des Standard-DNS-Sequenzierungsprotokolls wurden mehrere Verfahren vorgeschlagen, DNS-Sequenzierung ohne einen elektrophoretischen Schritt zu machen. Ein Ansatz erfordert Hybridisierung oder Fragmentierungssequenzierung (Bains, Biotechnology 10, 757–58 (1992) und Mirzabekov et al., FEBS Letters 256, 118–122 (1989)), wobei die spezifische Hybridisierung bekannter kurzer Oligonukleotide (z. B. Octadesoxynu-

kleotide, die 65.536 verschiedene Sequenzen ergeben) an eine komplementäre DNS-Sequenz verwendet wird. Eine positive Hybridisierung enthüllt eine kurze Strecke der unbekannten Sequenz. Wiederholen dieses Verfahrens durch Ausführen von Hybridisierungen mit allen möglichen Octadesoxynukleotiden sollte theoretisch die Sequenz bestimmen. In einem vollständig anderen Ansatz wird eine schnelle Sequenzierung von DNS durch unilateralen Abbau eines einzelnen immobilisierten DNS-Fragments durch eine Exonuklease in einem sich bewegenden Flussstrom und durch Detektion der abgespaltenen Nukleotide durch ihren spezifischen Fluoreszenz-Tag durch Laseranregung gemacht (Jett et al., J. Biomolecular Structure & Dynamics 7, 301–309, (1989); US-Energieministerium, PCT-Anmeldung Nr. WO 89/03432). In einem anderen von Hyman (Anal. Biochem. 174, 423–436 (1988)) vorgeschlagenen System wird das Pyrophosphat zum Bestimmen der DNS-Sequenz verwendet, das erzeugt wird, wenn das korrekte Nukleotid an die wachsende Kette auf einem Primer-Template-System gebunden wird. Die verwendeten Enzyme und die DNS werden durch Festphasen (DEAE-Sepharose und Sepharose) entweder durch ionische Wechselwirkungen oder durch kovalente Bindung festgehalten. In einem kontinuierlichen Durchflusssystem wird die Pyrophosphatmenge durch Bioluminezenz (Luciferase) bestimmt. Ein Syntheseansatz zur DNS-Sequenzierung wird ebenso von Tsien et al. (PCT-Anmeldung Nr. WO 91/06678) verwendet. Hier werden die ankommenden dNTPs am 3'-Ende durch verschiedene Blockierungsgruppen geschützt, wie z. B. Acetyl- oder Phosphatgruppen, die vor dem nächsten Verlängerungsschritt entfernt werden, was dieses Verfahren sehr langsam im Vergleich mit Standardsequenzierungsverfahren macht. Die Template-DNS wird auf einem Polymerträger immobilisiert. Um den Einbau zu detektieren, wird zusätzlich eine fluoreszierende oder radioaktive Markierung in die modifizierten dNTPs eingebaut. Dieselbe Patentanmeldung beschreibt ebenso einen Apparat, der zur Automatisierung des Verfahrens konstruiert ist.

**[0014]** Massenspektrometrie stellt im Allgemeinen ein Mittel zum "Wiegen" individueller Moleküle bereit, indem man die Moleküle im Vakuum ionisiert und man sie durch Volatilisieren "fliegen" lässt. Unter dem Einfluss von Kombinationen von elektrischen und magnetischen Feldern folgen die Ionen Trajektorien abhängig von ihrer individuellen Masse (m) und Ladung (z). Im Bereich von Molekülen mit niedrigem Molekulargewicht ist Massenspektrometrie seit langem Teil des physikalisch-organischen Routinerepertoires zur Analyse und Kennzeichnung organischer Moleküle durch die Bestimmung der Masse des Ions des Ausgangsmoleküls. Des Weiteren wird durch Einrichten von Kollisionen dieses Ions des Ausgangsmoleküls mit anderen Partikeln (z. B. Argonatomen) das Ion des Moleküls durch die sogenannte kollisionsinduzierte Dissoziation (collision induced dissociation, CID) fragmentiert, wobei Sekundärionen gebildet werden. Das Fragmentierungsmuster/der Weg erlaubt sehr oft die Ableitung einer detaillierten strukturellen Information. Viele Anwendungen von massenspektrometrischen Verfahren sind im Stand der Technik bekannt, insbesondere in den Biowissenschaften, und können in Methods of Enzymology, Band 193: "Mass Spectrometry" (Herausgeber J. A. McCloskey), 1990, Academic Press, New York) zusammengefasst gefunden werden.

**[0015]** Aufgrund der anscheinenden analytischen Vorteile von Massenspektrometrie bei der Bereitstellung einer hohen Detektionsempfindlichkeit, der Genauigkeit der Massemessungen, der detaillierten Strukturinformationen durch CID in Verbindung mit einer MS/MS-Konfiguration und der Geschwindigkeit als auch einer Online-Datenübertragung auf einen Computer besteht erhebliches Interesse an der Verwendung von Massenspektrometrie für die Strukturanaylse von Nukleinsäuren. Neuere Übersichtsarbeiten, die dieses Feld zusammenfassen, schließen K. H. Schram, "Mass Spectrometry of Nucleic Acid Components, Biomedical Applications of Mass Spectrometry" 34, 203–287 (1990) und P. F. Crain, "Mass Spectrometry Techniques in Nucleic Acids Research, "Mass Spectrometry Reviews 9, 505–554 (1990) ein. Die größte Hürde beim Anwenden von Massenspektrometrie auf Nukleinsäuren ist die Schwierigkeit, diese sehr polaren Biopolymere zu volatilisieren. Daher ist "Sequenzierung" auf synthetische Oligonukleotide mit niedrigem Molekulargewicht beschränkt, indem die Masse des Ions des Ausgangsmoleküls bestimmt wird und durch dieses die schon bekannte Sequenz bestätigt wird oder alternativ die bekannte Sequenz durch die Erzeugung von Sekundärionen (Fragmentionen) über CID in einer MS/MS-Konfiguration bestätigt wird, die insbesondere zur Ionisation und Volatilisierung das Verfahren der schnellen Atombombardierung (fast atomic bombardment, FAB-Massenspektrometrie) oder Plasmadesorption (PD-Massenspektrometrie) verwendet. Als ein Beispiel wurde die Anwendung von FAB auf die Analyse von geschützten dimeren Blöcken zur chemischen Synthese von Oligodesoxynukleotiden beschrieben (Köster et al. Biomedical Environmental Mass Spectrometry 14, 111–116 (1987)).

**[0016]** Zwei noch neuere Ionisations-/Desorptionstechniken sind Elektrospray/Ionenspray (ES) und matrixassistierte Laser-Desorption/Ionisation (MALDI). ES-Massenspektrometrie wurden durch Fenn et al. eingeführt (J. Phys. Chem. 88, 4451–59 (1984); PCT-Anmeldung Nr. WO 90/14148), und gegenwärtige Anwendungen sind in neueren Übersichtsartikeln zusammengefasst (R. D. Smith et al., Anal. Chem. 62, 882–89 (1990) und B. Ardrey, Electrospray Mass Spectrometry, Spectroscopy Europe, 4, 10–18 (1992)). Die Molekulargewichte des Tetradecanukleotids d(CATGCCATGGCATG)(SEQ ID NO: 1)(Covey et al. "The Determination of Protein,

Oligonukleotide and Peptide Molecular Weights by Ionspray Mass Spectrometry, "Rapid Communications in Mass Spectrometry, 2, 249–256 (1988)), des 21mers d(AAATTGTGCACATCCTGCAGC)(SEQ ID NO: 2) und einer tRNS mit 76 Nukleotiden ohne Angabe von Einzelheiten (Methods in Enzymology, 193, "Mass Spectrometry" (Herausgeber McCloskey), Seite 425, 1990, Academic Press, New York) wurden publiziert. Als Massenanalysator wird am häufigsten ein Quadrupol verwendet. Die Bestimmung von Molekulargewichten in femtomolaren Probenmengen ist aufgrund der Anwesenheit von vielen Ionenpeaks sehr genau, die alle für die Massenberechnung verwendet werden könnten.

**[0017]** MALDI-Massenspektrometrie kann im Gegensatz besonders attraktiv sein, wenn eine Flugzeit-(time-of-flight, TOF)-Konfiguration als Massenanalysator verwendet wird. Die MALDI-TOF-Massenspektrometrie wurde durch Hillenkamp et al. eingeführt ("Matrix Assisted UV-Laser Desorption/Ionization: A New Approach to Mass Spectrometry of Large Biomolecules", Biological Mass Spectrometry (Herausgeber Burlingame und McCloskey), Elsevier Science Publishers, Amsterdam, Seiten 49–60, 1990). Da in den meisten Fällen keine mehrfachen molekularen Ionenpeaks mit dieser Technik produziert werden, sehen die Massenspektren im Vergleich zur ES-Massenspektrometrie im Prinzip einfacher aus. Obwohl DNS-Moleküle mit einem Molekulargewicht von bis zu 410.000 Dalton desorbiert und volatilisiert werden könnten (Williams et al. "Volatilization of High Molecular Weight DNA by Pulsed Laser Ablation of Frozen Aqueous Solutions", Science, 246, 1585–87 (1989)), wurde diese Technik bis jetzt nur verwendet, um die Molekulargewichte relativ kleiner Oligonukleotide mit bekannter Sequenz zu bestimmen, z. B. von Oligothymidylsäuren mit bis zu 18 Nukleotiden (Huth-Fehre et al., "Matrix-Assisted Laser Desorption Mass Spectrometry of Oligodeoxythymidylic Acids," Rapid Communications in Mass Spectrometry, 6, 209–13 (1992)) und von einer doppelsträngigen DNS mit 28 Basenpaaren (Williams et al., "Time-of-Flight Mass Spectrometry of Nucleic Acids by Laser Ablation and Ionization from a Frozen Aqueous Matrix," Rapid Communications in Mass Spectrometry, 4, 348–351 (1990)). In einer Veröffentlichung (Huth-Fehre et al., 1992, supra) wurde gezeigt, dass eine Mischung aller Oligothymidylsäuren mit von n = 12 bis n = 18 Nukleotiden aufgelöst werden konnte.

**[0018]** In US-Patent Nr. 5,064,754 werden RNS-Transkripte hergestellt, die mit DNS verlängert sind und die beide zu der zu sequenzierenden DNS komplementär sind, indem NTPs, dNTPs und als terminierende Nukleotide ddNTPs eingebaut werden, die an der 5'-Position der Zuckergruppe mit einer Kombination der Isotope $^{12}C$, $^{13}C$, $^{14}C$, $^{1}H$, $^{2}H$, $^{3}H$, $^{16}O$, $^{17}O$ und $^{18}O$ substituiert sind. Die erhaltenen Polynukleotide werden zu 3'-Nukleotiden abgebaut, an der N-glycosidischen Bindung gespalten, und die isotopenmarkierte 5'-Funktionalität wird durch Periodat-Oxidation entfernt, und die resultierende Formaldehydspezies wird durch Massenspektrometrie bestimmt. Eine spezifische Kombination von Isotopen dient dazu, basenspezifisch zwischen internen Nukleotiden, die aus dem Einbau von NTPs und dNTPs herrühren, und terminalen Nukleotiden, die durch Verbinden von ddNTPs an das Ende der Polynukleotidkette verursacht werden, zu unterscheiden. Eine Serie von RNS/DNS-Fragmenten wird hergestellt und in einer Ausführungsform durch Elektrophorese getrennt, und mit Hilfe des sogenannten Matrixanalyseverfahrens wird die Sequenz abgeleitet.

**[0019]** Im japanischen Patent Nr. 59-131909 wird ein Instrument beschrieben, das Nukleinsäurefragmente detektiert, die entweder durch Elektrophorese, Flüssigkeitschromatographie oder Hochgeschwindigkeitsgelfiltration getrennt werden. Eine massenspektrometrische Detektion wird durch Einbauen von Atomen in die Nukleinsäuren erzielt, die normalerweise nicht in DNS auftreten, wie z. B. S, Br, I oder Ag, Au, Pt, Os, Hg. Das Verfahren wird jedoch nicht auf Sequenzierung von DNS unter Verwendung des Verfahrens nach Sanger angewendet. Insbesondere schlägt es keine basenspezifische Korrelation derartiger Elemente mit einem individuellen ddNTP vor.

**[0020]** Die PCT-Anmeldung Nr. WO 89/12694 (Brennan et al., Proc. SPIE-Int. Soc. Opt. Eng. 1206, (New Technol. Cytom. Mol. Biol.), Seiten 60–77 (1990); und Brennan, US-Patent Nr. 5,003,059) verwendet die Methodologie nach Sanger zur DNS-Sequenzierung durch Verwendung einer Kombination von entweder den vier stabilen Isotopen $^{32}S$, $^{33}S$, $^{34}S$, $^{36}S$ oder $^{35}Cl$, $^{37}Cl$, $^{79}Br$, $^{81}B$, um die kettenterminierenden ddNTPs spezifisch zu markieren. Die Schwefelisotope können sich entweder an der Base oder an der Alphaposition der Triphosphatgruppe befinden, während die Halogenisotope sich entweder an der Base oder an der 3'-Position des Zuckerringes befinden. Die Sequenzierungsreaktionsmischungen werden durch eine elektrophoretische Technik wie z. B. CZE getrennt, in eine Verbrennungseinheit überführt, in der die Schwefelisotope der eingebauten ddNTPs bei etwa 900°C in einer Sauerstoffatmosphäre umgewandelt werden. Das $SO_2$, das mit Massen von 64, 65, 66 oder 68 erzeugt wird, wird online durch Massenspektrometrie bestimmt, wobei z. B. als Massenanalysator ein Quadrupol mit einem Einzelionen-Multiplier verwendet wird, um den Ionenstrom zu detektieren.

**[0021]** Ein ähnlicher Ansatz wird in US-Patent Nr. 5,002,868 (Jacobson et al., Proc. SPIE-Int. Soc. Opt. Eng. 1435, (Opt. Methods Ultrasensitive Detect. Anal. Tech. Appl.), 26–35 (1991)) vorgeschlagen, wobei Sequen-

zierung nach Sanger mit vier ddNTPs, die spezifisch an der Alphaposition der Triphosphatgruppe mit einem der vier stabilen Schwefelisotope wie oben beschrieben substituiert sind, und eine nachfolgende Trennung der vier Sätze verschachtelter Sequenzen durch Röhrengelelektrophorese verwendet werden. Der einzige Unterschied ist die Verwendung von Resonanzionisationsspektroskopie (RIS) in Verbindung mit einem magnetischen Sektormassenanalysator, wie in US-Patent Nr. 4,442,354 offenbart ist, um die Schwefelisotope zu detektieren, die den spezifischen Nukleotidterminatoren entsprechen und dadurch die Zuweisung der DNS-Sequenz erlauben.

**[0022]** Die EPO-Patentanmeldungen Nr. 0360676 A1 und 0360677 A1 beschreiben ebenfalls eine Sequenzierung nach Sanger, wobei stabile Isotopensubstitutionen in den ddNTPs, wie z. B. D, $^{13}C$, $^{15}N$, $^{17}O$, $^{18}O$, $^{32}S$, $^{33}S$, $^{34}S$, $^{36}S$, $^{19}F$, $^{35}Cl$, $^{37}Cl$, $^{79}Br$, $^{81}Br$, und $^{127}I$ oder funktionelle Gruppen wie z. B. $CF_3$ oder $Si(CH_3)_3$, gemäß der chemischen Funktionalität an der Base, dem Zucker oder der Alphaposition der Triphosphatgruppe verwendet werden. Die Sequenzierungsreaktionsmischungen nach Sanger werden durch Röhrengelelektrophorese getrennt. Das Effluat wird durch das Elektrospray/Thermospray-Verneblerverfahren in ein Aerosol verwandelt und dann durch ein heißes Plasma (7.000 bis 8.000°K) atomisiert und ionisiert und durch einen einfachen Massenanalysator analysiert. Es wird ein Instrument vorgeschlagen, das einen in die Lage versetzt, die Analyse der Sequenzierungsreaktionsmischung nach Sanger zu automatisieren, das aus einer Röhrenelektrophorese, einem Vernebler und einem Massenanalysator besteht.

**[0023]** Die Anwendung von Massenspektrometrie zum Durchführen von DNS-Sequenzierungen durch das Hybridisierungs-Fragmentverfahren (siehe oben) wurde kürzlich vorgeschlagen (Bains, "DNA Sequencing by Mass Spectrometry: Outline of a Potential Future Application", Chimicaoggi 9, 13–16 (1991)).

## Zusammenfassung der Erfindung

**[0024]** Die Erfindung beschreibt ein neues Verfahren zur DNS-Sequenzierung. Die Verbesserungen gegenüber den existierenden DNS-Sequenzierungstechnologien schließen hohe Geschwindigkeit, hohen Durchsatz, keine Notwendigkeit für Elektrophorese (und daher keine Gel-Leseartefakte aufgrund der vollständigen Abwesenheit eines elektrophoretischen Schritts) und keine teuren Reagenzien ein, die verschiedene Substitutionen mit stabilen Isotopen involvieren. Die Erfindung verwendet die Sequenzierungsstrategie nach Sanger und setzt die Sequenzinformation durch Analyse der verschachtelten Fragmente zusammen, die durch basenspezifische Kettentermination über ihre unterschiedlichen Molekularmassen erhalten werden, wobei Massenspektrometrie verwendet wird, z. B. MALDI- oder ES-Massenspektrometrie. Ein weiterer Anstieg des Durchsatzes kann erhalten werden, indem Massemodifikationen im Oligonukleotidprimer, den kettenterminierenden Nukleosidtriphosphaten und/oder den kettenverlängernden Nukleosidtriphosphaten eingeführt werden, als auch indem integrierte Tag-Sequenzen verwendet werden, die Multiplexen durch Hybridisierung von tagspezifischen Sonden mit massedifferenzierten Molekulargewichten ermöglichen.

## Kurze Beschreibung der Abbildungen

**[0025]** Abb. 1 ist eine Darstellung des Verfahrens zur Herstellung der Proben, die durch Massenspektrometrie analysiert werden sollen. Dieses Verfahren erfordert die Insertion eines DNS-Fragments unbekannter Sequenz in einen Klonierungsvektor, wie z. B. Derivate von M13, pUC oder Phagemide, Transformieren der doppelsträngigen Form in die einzelsträngige Form, Durchführen der vier Sequenzierungsreaktionen nach Sanger, vorübergehendes Binden der basenspezifisch terminierten verschachtelten Fragmentfamilie an einen festen Träger, Entfernen aller Nebenprodukte durch einen Waschschritt, Konditionieren der verschachtelten DNS- oder RNS-Fragmente durch z. B. ein Kationenaustausch- oder Modifizierungsreagenz und entweder direktes Überführen der immobilisierten verschachtelten Fragmente zur massenspektrometrischen Analyse oder Abspalten der gereinigten Fragmentfamilie vom Träger und Abziehen des Spaltungsreagenzes.

**[0026]** Abb. 2A zeigt die Sequenzierungsprodukte nach Sanger eines hypothetischen DNS-Fragments mit einer Länge von 50 Nukleotiden (SEQ ID NO: 3) mit einer annähernd gleich balancierten Basenzusammensetzung unter Verwendung von ddTTP als terminierendem Desoxynukleosidtriphosphat. Die molekularen Massen der verschiedenen kettenterminierten Fragmente sind angegeben.

**[0027]** Abb. 2B zeigt ein idealisiertes Massenspektrum von solch einer DNS-Fragmentmischung.

**[0028]** Die Abb. 3A und Abb. 3B zeigen in Analogie zu den Abb. 2A und Abb. 2B Daten für dieselbe Modellsequenz (SEQ ID NO: 3) mit ddATP als Kettenterminator.

**[0029]** Die Abb. 4A und Abb. 4B zeigen Daten analog zu den Abb. 2A und Abb. 2B, wenn ddGTP als Kettenterminator für dieselbe Modellsequenz (SEQ ID NO: 3) verwendet wird.

**[0030]** Die Abb. 5A und Abb. 5B veranschaulichen die Ergebnisse, die erhalten werden, wenn die Kettentermination mit ddCTP als Kettenterminator auf ähnliche Weise wie in den Abb. 2A und Abb. 2B für dieselbe Modellsequenz (SEQ ID NO: 3) durchgeführt wird.

**[0031]** Abb. 6 fasst die Ergebnisse der Abb. 2A bis Abb. 5B zusammen, wobei sie die Korrelation von Molekulargewichten der verschachtelten vier Fragmentfamilien mit der DNS-Sequenz zeigt (SEQ ID NO: 3).

**[0032]** Die Abb. 7A und Abb. 7B veranschaulichen die allgemeine Struktur von massemodifizierten Sequenzierungsnukleinsäureprimern oder Tag-Sequenzierungssonden für entweder eine DNS- oder RNS-Sequenzierung nach Sanger.

**[0033]** Die Abb. 8A und Abb. 8B zeigen die allgemeine Struktur für die massemodifizierten Triphosphate für entweder eine DNS- oder RNS-Sequenzierung nach Sanger. Allgemeine Formeln für die kettenverlängernden und die kettenterminierenden Nukleosidtriphosphate werden demonstriert.

**[0034]** Abb. 9 umreißt beispielhaft verschiedene Linkerchemien (X) mit entweder Polyethylenglycol oder endständig monoalkyliertem Polyethylenglycol (R).

**[0035]** Abb. 10 veranschaulicht ähnliche Linkerchemien, wie in Abb. 8A und Abb. 8B gezeigt, und beschreibt verschiedene massemodifizierende Gruppen (R).

**[0036]** Abb. 11 umreißt, wie multiplex-massenspektrometrische Sequenzierung unter Verwendung der massemodifizierten Nukleinsäureprimer (UP) arbeiten kann.

**[0037]** Abb. 12 zeigt das Verfahren zur multiplex-massenspektrometrischen Sequenzierung, das massemodifizierte kettenverlängernde und/oder terminierende Nukleosidtriphosphate verwendet.

**[0038]** Abb. 13 zeigt eine multiplex-massenspektrometrische Sequenzierung durch Einbeziehen der Hybridisierung von massemodifizierten Sonden, die für eine Tag-Sequenz spezifisch sind.

**[0039]** Abb. 14 zeigt ein MALDI-TOF-Spektrum einer Mischung von Oligothymidylsäuren $d(pT)_{12-18}$.

**[0040]** Abb. 15 zeigt eine Überlagerung von MALDI-TOF-Spektren des 50mers d(TAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATGACTAGCT) (SEQ ID NO: 3)(500 fmol) und $dT(pdT)_{99}$ (500 fmol).

**[0041]** Die Abb. 16A bis Abb. 16M zeigen die MALDI-TOF-Spektren aller 13 DNS-Sequenzen, die die verschachtelten dT-terminierten Fragmente der DNS-Sequenzierungssimulation nach Sanger aus Abb. 2 darstellen. Jeweils 500 fmol wie folgt: Abb. 16A ist ein 7mer; Abb. 16B ist ein 10mer; Abb. 16C ist ein 11mer; Abb. 16D ist ein 19mer; Abb. 16E ist ein 20mer; Abb. 16F ist ein 24mer; Abb. 16G ist ein 26mer; Abb. 16H ist ein 33mer; Abb. 16I ist ein 37mer; Abb. 16J ist ein 38mer; Abb. 16K ist ein 42mer; Abb. 16L ist ein 46mer und Abb. 16M ist ein 50mer.

**[0042]** Die Abb. 17A und Abb. 17B zeigen die Überlagerung der Spektren aus Abb. 16. Die beiden Tafeln zeigen zwei verschiedene Maßstäbe und die Spektren, die bei diesen Maßstäben analysiert wurden. Abb. 17A zeigt die Überlagerung der Spektren aus Abb. 16A bis Abb. 16F. Der Buchstabe über jedem Peak entspricht den originalen Spektren des Fragments in Abb. 16. Peak B entspricht z. B. Abb. 16B, Peak C entspricht Abb. 16C, etc.

**[0043]** Abb. 18 zeigt die überlagerten MALDI-TOF-Spektren der MALDI-MS-Analyse von massemodifizierten Oligonukleotiden, wie in Beispiel 21 beschrieben.

**[0044]** Abb. 19 veranschaulicht verschiedene Linkerchemien zwischen dem festen Träger (P) und dem Nukleinsäureprimer (NS) durch eine starke elektrostatische Wechselwirkung.

**[0045]** Die Abb. 20A und Abb. 20B veranschaulichen verschiedene Linkerchemien zwischen dem festen Träger (P) und dem Nukleinsäureprimer (NS) durch einen Ladungsübertragungskomplex eines Ladungsüber-

tragungsakzeptors (A) und eines Ladungsübertragungsdonors (D).

**[0046]** Abb. 21 veranschaulicht verschiedene Linkerchemien zwischen dem festen Träger (P) und dem Nukleinsäureprimer (NS) durch einen stabilen organischen Rest.

**[0047]** Abb. 22 veranschaulicht eine mögliche Linkerchemie zwischen einem festen Träger (P) und dem Nukleinsäureprimer (NS) durch Watson-Crick-Basenpaaren.

**[0048]** Abb. 23 veranschaulicht ein Verbinden des festen Trägers (P) und des Nukleinsäureprimers (NS) durch eine photolytisch spaltbare Bindung.

## Detaillierte Beschreibung der Erfindung

**[0049]** Die Erfindung beschreibt ein verbessertes Verfahren zur DNS-Sequenzierung. Insbesondere verwendet die Erfindung Massenspektrometrie, wie z. B. matrixassistierte Laserdesorptions/Ionisations- (MALDI) oder Elektrospray- (ES) Massenspektrometrie (MS), um die Sequenzierungsreaktionsmischungen nach Sanger zu analysieren.

**[0050]** Bei Sequenzierung nach Sanger werden vier Familien von kettenterminierten Fragmenten erhalten. Die Massedifferenz pro zugefügtem Nukleotid ist 289,19 für dpC, 313,21 für dpA, 329,21 für dpG bzw. 304,2 für dpT.

**[0051]** Erfindungsgemäß können die Molekulargewichte der vier spezifisch terminierten Fragmentfamilien simultan durch MS bestimmt werden, entweder durch Mischen der Produkte aller vier Reaktionsläufe in wenigstens zwei getrennten Reaktionsgefäßen (d. h. alle laufen getrennt, oder zwei zusammen oder drei zusammen) oder durch Laufenlassen einer Reaktion, die alle vier kettenterminierenden Nukleotide (d. h. eine Reaktionsmischung, umfassend dTTP, ddTTP, dATP, ddATP, dCTP, ddCTP, dGTP, ddGTP) in einem Reaktionsgefäß hat. Durch simultanes Analysieren aller vier basenspezifisch terminierten Reaktionsprodukte wurden alle Molekulargewichtswerte tatsächlich interpoliert. Ein Vergleich der Massedifferenz, die zwischen Fragmenten mit bekannten Massen für jedes kettenterminierende Nukleotid gemessen wird, ermöglicht die Zuordnung einer Sequenz, die ausgeführt werden soll. In einigen Fällen kann es wünschenswert sein, wie unten diskutiert, die Masse der kettenterminierenden Nukleotide zu modifizieren, um die Differenz im Molekulargewicht zwischen jedem Nukleotid zu erweitern. Es ist den Fachleuten klar, wann Massemodifizierung der kettenterminierenden Nukleotide wünschenswert ist und wann sie z. B. von der Auflösungsfähigkeit des bestimmten verwendeten Spektrometers abhängen kann. Es kann z. B. wünschenswert sein, vier kettenterminierende Nukleotide ddTTP, ddCTP[1], ddATP[2] und ddGTP[3] herzustellen, wobei ddCTP[1], ddATP[2] und ddGTP[3] massemodifiziert werden, um Molekulargewichte zu haben, die durch das bestimmte verwendete Spektrometer von einander aufgelöst werden können.

**[0052]** Die Begriffe kettenverlängernde Nukleotide und kettenterminierende Nukleotide sind im Stand der Technik gut bekannt. Für DNS schließen kettenverlängernde Nukleotide 2'-Desoxyribonukleotide ein, und kettenterminierende Nukleotide schließen 2',3'-Didesoxyribonukleotide ein. Für RNS schließen kettenverlängernde Nukleotide Ribonukleotide ein, und kettenterminierende Nukleotide schließen 3'-Desoxyribonukleotide ein. Der Begriff Nukleotid ist ebenso im Stand der Technik gut bekannt. Für die erfindungsgemäßen Zwecke schließen Nukleotide Nukleosidmono-, -di-, und -triphosphate ein. Nukleotide schließen ebenso modifizierte Nukleotide ein, wie z. B. Phosphorthioatnukleotide.

**[0053]** Da Massenspektrometrie ein serielles Verfahren ist, kann im Gegensatz zur gegenwärtig verwendeten Schichtgelelektrophorese, die ermöglicht, dass mehrere Proben parallel verarbeitet werden, eine weitere Verbesserung durch multiplexmassenspektrometrische DNS-Sequenzierung erzielt werden, die simultane Sequenzierung von mehr als einem DNS- oder RNS-Fragment erlaubt. Wie unten detaillierter beschrieben, kann der Bereich von etwa 300 Masseeinheiten zwischen einem hinzugefügten Nukleotid verwendet werden, indem entweder massemodifizierte Nukleinsäuresequenzierungsprimer oder kettenverlängernde und/oder terminierende Nukleosidtriphosphate verwendet werden, um das Molekulargewicht der basenspezifisch terminierten Fragmente einer bestimmten DNS- oder RNS-Spezies zu verschieben, die auf vorbestimmte Weise sequenziert werden soll. Zum ersten Mal können mehrere Sequenzierungsreaktionen parallel massenspektrometrisch analysiert werden. Multiplex-massenspektrometrische DNS-Sequenzierung kann durchgeführt werden, indem die Fragmentfamilien durch spezifische Oligonukleotide (Tag-Sonden) massemodifiziert werden, die an spezifischen Tag-Sequenzen innerhalb jeder der Fragmentfamilien hybridisieren. Die Tag-Sonden können kovalent an die individuelle und spezifische Tag-Sequenz vor der Massenspektrometrie gebunden werden.

**[0054]** In einer erfindungsgemäßen Ausführungsform werden die Molekulargewichtswerte von wenigstens zwei basenspezifisch terminierten Fragmenten durch Verwendung von Massenspektrometrie gleichzeitig bestimmt. Die Molekulargewichtswerte von vorzugsweise wenigstens fünf und stärker bevorzugt wenigstens zehn basenspezifisch terminierten Fragmenten werden durch Massenspektrometrie bestimmt. Ebenso sind Bestimmungen der Molekulargewichtswerte von wenigstens 20 basenspezifisch terminierten Fragmenten und wenigstens 30 basenspezifisch terminierten Fragmenten in die Erfindung eingeschlossen. Weiterhin können die verschachtelten basenspezifisch terminierten Fragmente eines spezifischen Satzes von allen Reaktanten und den Nebenprodukten gereinigt werden, aber werden nicht voneinander getrennt. Der gesamte Satz der verschachtelten basenspezifisch terminierten Fragmente wird gleichzeitig analysiert, und die Molekulargewichtswerte werden bestimmt. Wenigstens zwei basenspezifisch terminierte Fragmente werden gleichzeitig durch Massenspektrometrie analysiert, wenn die Fragmente in derselben Probe enthalten sind.

**[0055]** Im Allgemeinen beginnt das gesamte massenspektrometrische DNS-Sequenzierungsverfahren mit einer Bibliothek kleiner genomischer Fragmente, die nach einem ersten zufälligen oder spezifischen Schneiden der genomischen DNS in große Stücke erhalten wird, die dann in verschiedenen Subklonierungsschritten bezüglich der Größe reduziert und in Vektoren inseriert werden, wie Derivaten von M13 oder pUC (z. B. M13mp18 oder M13mp19)(siehe **Abb. 1**). In einem unterschiedlichen Ansatz werden die in Vektoren wie z. B. M13 inserierten Fragmente durch Subklonierung erhalten, die mit einer cDNS-Bibliothek startet. In noch einem anderen Ansatz werden die DNS-Fragmente, die sequenziert werden sollen, durch die Polymerasekettenreaktion erzeugt (z. B. Higuchi et al., "A General Method of in vitro Preparation and Mutagenesis of DNA Fragments: Study of Protein and DNA Interactions", Nucleic Acids Res., 16, 7351–67 (1988)). Wie im Stand der Technik bekannt, kann Sequenzierung nach Sanger mit einem Nukleinsäureprimer (UP), der an den Plusstrang bindet oder mit einem anderen Nukleinsäureprimer, der an den entgegengesetzte Minusstrang bindet, beginnen. Daher kann entweder die komplementäre Sequenz beider Stränge einer gegebenen unbekannten DNS-Sequenz erhalten werden (Verminderung der Unklarheit bei der Sequenzbestimmung vorausgesetzt), oder die Länge der Sequenzinformation, die aus einem Klon erhältlich ist, kann durch Erzeugen von Sequenzinformation von beiden Enden des unbekannten vektorinserierten DNS-Fragments verlängert werden.

**[0056]** Der Nukleinsäureprimer trägt vorzugsweise am 5'-Ende eine Bindungsfunktionalität L, die einen Spacer mit ausreichender Länge einschließen kann und die mit einer geeigneten Funktionalität L' auf einem festen Träger wechselwirken kann, um eine reversible Bindung zu bilden, z. B. eine photospaltbare Bindung. Da jede der vier Sequenzierungsfamilien nach Sanger mit einem Nukleinsäureprimer beginnt (L-UP, **Abb. 1**), kann diese Fragmentfamilie an den festen Träger gebunden werden, indem sie mit funktionellen Gruppen L' auf der Oberfläche eines festen Trägers reagiert und dann intensiv gewaschen wird, um alle Puffersalze, Triphosphate, Enzyme, Reaktionsnebenprodukte etc. zu entfernen. Außerdem kann es für die massenspektrometrische Analyse auf dieser Stufe wichtig sein, das Kation am Phosphatgerüst des DNS-Fragments auszutauschen, um Peakverbreiterung aufgrund einer Heterogenität der Kationen zu eleminieren, die pro Nukleotideinheit gebunden sind. Da die L-L'-Bindung nur vorübergehender Natur mit dem Ziel ist, verschachtelte DNS- oder RNS-Fragmente nach Sanger zu fangen, um sie für die massenspektrometrische Analyse sauber zu konditionieren, gibt es verschiedene Chemien, die diesem Zweck dienen können. Zusätzlich zu den angegebenen Beispielen, in denen die verschachtelten Fragmente kovalent an den festen Träger gebunden sind, gewaschen und vom Träger für die massenspektrometrische Analyse abgespalten werden, kann die vorübergehende Bindung so beschaffen sein, dass sie unter den Bedingungen der Massenspektrometrie gespalten wird, d. h. eine photospaltbare Bindung wie z. B. ein Ladungsübertragungskomplex oder ein stabiler organischer Rest. Außerdem kann die Bindung mit L' gebildet werden, das eine quartäre Ammoniumgruppe ist (einige Beispiele sind in **Abb. 19** gegeben). In diesem Fall trägt die Oberfläche des festen Trägers vorzugsweise negative Ladungen, die das negativ geladene Nukleinsäuregerüst abweisen und dadurch die Desorption erleichtern. Eine Desorption findet entweder durch die Wärme statt, die durch den Laserpuls geschaffen wird und/oder abhängig von L' durch spezifische Absorption von Laserenergie, die in Resonanz mit dem L'-Chromophor ist (siehe z. B. Beispiele, die in **Abb. 19** gegeben sind). Die Funktionalitäten L und L' können ebenso einen Ladungsübertragungskomplex bilden und dabei die vorübergehende L-L'-Bindung bilden. Verschiedene Beispiele für geeignete Funktionalitäten mit entweder Akzeptor- oder Donoreigenschaften sind ohne Beschränkung in den **Abb. 20A** und **Abb. 20B** beschrieben. Da in vielen Fällen das "Ladungsübertragungsband" durch UV/VIS-Spektrometrie bestimmt werden kann (siehe z. B. Organic Charge Transfer Complexes von R. Foster, Academic Press, 1969), kann die Laserenergie auf die korrespondierende Energie der Ladungsübertragungswellenlänge abgestimmt werden, und daher kann eine spezifische Desorption vom festen Träger initiiert werden. Die Fachleute werden erkennen, dass mehrere Kombinationen diesem Zweck dienen können und dass die Donorfunktionalität entweder auf dem festen Träger sein kann oder an die verschachtelten DNS-/RNS-Fragmente nach Sanger gekoppelt sein kann oder umgekehrt.

**[0057]** In noch einem anderen Ansatz kann die vorübergehende L-L'-Bindung gebildet werden, indem homolytisch relativ stabile Reste gebildet werden, wie in Abb. 21 beispielhaft dargestellt. In Beispiel 4 aus Abb. 21 wird eine Kombination der Ansätze gezeigt, die Ladungsübertragungskomplexe und stabile organische Reste verwenden. Hier werden die verschachtelten DNS-/RNS-Fragmente nach Sanger über die Bildung eines Ladungsübertragungskomplexes gefangen. Unter dem Einfluss des Laserpulses findet sowohl Desorption (wie oben diskutiert) als auch Ionisation an der Position des Restes statt. In den anderen Beispielen der Abb. 21 wird unter dem Einfluss des Laserpulses die L-L'-Bindung gespalten, und die verschachtelten DNS-/RNS-Fragmente nach Sanger werden an der gebildeten Position des Restes desorbiert und nachfolgend ionisiert. Die Fachleute erkennen, dass andere organische Reste ausgewählt werden können, und dass im Verhältnis zu den benötigten Dissoziationsenergien, um die Bindung zwischen ihnen homolytisch zu spalten, eine entsprechende Laserwellenlänge ausgewählt werden kann (siehe z. B. Reactive Molecules von C. Wentrup. John Wiley & Sons, 1984). In noch einem anderen Ansatz werden die verschachtelten DNS-/RNS-Fragmente nach Sanger über Watson-Crick-Basenpaarung mit einem an den festen Träger gebundenen Oligonukleotid gefangen, das entweder zu der Sequenz des Nukleinsäureprimers oder zu der Oligonukleotidsequenz des Tags komplementär ist (siehe Abb. 22). Die gebildete Duplex wird unter dem Einfluss des Laserpulses gespalten, und eine Desorption kann initiiert werden. Die an den festen Träger gebundene Basensequenz kann durch natürliche Oligoribo- oder Oligodesoxyribonukleotide als auch Analoga (z. B. ein thiomodifiziertes Phosphodiester- oder Phosphotriestergerüst) oder durch Verwenden von Oligonukleotidmimetika präsentiert werden, wie z. B. PNS-Analoga (siehe z. B. Nielsen et al., Science, 254, 1497 (1991)), die die Basensequenz weniger empfänglich für enzymatischen Abbau machen und daher die gesamte Stabilität der an den festen Träger gebundenen Fängerbasensequenz erhöhen. Mit geeigneten L-L'-Bindungen kann eine Spaltung direkt mit einem Laser erhalten werden, der auf die Energie abgestimmt ist, die zur Bindungsspaltung notwendig ist. Daher können die immobilisierten verschachtelten Fragmente nach Sanger direkt während der massenspektrometrischen Analyse abgelöst werden.

**[0058]** Um die massenspektrometrische Leistungsfähigkeit zu erhöhen, kann es notwendig sein, dass Phosphodiestergerüst vor der MS-Analyse zu modifizieren. Dies kann z. B. durch Verwenden von alpha-thiomodifizierten Nukleotiden zur Kettenverlängerung und -termination erreicht werden. Mit Alkylierungsmitteln wie z. B. Alkyliodiden, Iodazetamid, β-Iodethanol, 2,3-Epoxy-1-propanol (siehe Abb. 10) werden die Monothiophosphodiesterbindungen der verschachtelten Fragmente nach Sanger in Phosphotriesterbindungen überführt. Multiplexen durch Massemodifizierung wird in diesem Falle durch Massemodifizieren der Nukleinsäureprimer (UP) oder der Nukleosidtriphosphate an der Zucker- oder der Basengruppe erhalten. Fachleute können sich andere Modifizierungen der verschachtelten Fragmente nach Sanger vorstellen. In einer Ausführungsform der Erfindung ermöglicht die Linkerchemie die Abspaltung der so gereinigten verschachtelten DNS enzymatisch, chemisch oder physikalisch. Als Beispiel kann die L-L'-Chemie von der Art einer Disulfidbindung sein (zum Beispiel chemisch durch Mercaptoethanol oder Dithioerythritol spaltbar), eines Biotin/Streptavidinsystems, eines heterobifunktionellen Derivats einer Tritylethergruppe (Köster et al., "A Versatile Acid-Labile Linker for Modification of Synthetic Biomolecules", Tetrahedron Letters 31, 7095 (1990)), die unter milden sauren Bedingungen gespalten werden kann, einer Lävulinylgruppe, die unter nahezu neutralen Bedingungen mit einem Hydrazinium-/Acetatpuffer spaltbar ist, einer Arginin-Arginin- oder Lysin-Lysinbindung, die durch ein Endopeptidaseenzym wie Trypsin spaltbar ist, oder einer Pyrophosphatbindung, die durch eine Pyrophosphatase spaltbar ist, einer photospaltbaren Bindung, die z. B. physikalisch gespalten werden kann, oder kann ähnlicher Art (siehe z. B. Abb. 23) sein. Gegebenenfalls kann ein anderer Kationenaustausch vor der massenspektrometrischen Analyse durchgeführt werden. Falls eine enzymspaltbare Bindung verwendet wird, um die verschachtelten Fragmente zu immobilisieren, kann das zum Spalten der Bindung verwendete Enzym als ein interner Massestandard während der MS-Analyse dienen.

**[0059]** Das Reinigungsverfahren und/oder das Ionenaustauschverfahren kann durch eine Anzahl anderer Verfahren anstatt oder in Verbindung mit einer Immobilisierung auf einem festen Träger ausgeführt werden. Zum Beispiel können die basenspezifisch terminierten Produkte von den Reaktanten durch Dialyse, Filtration (einschließlich Ultrafiltration) und Chromatographie getrennt werden. Entsprechend können diese Techniken verwendet werden, um das Kation des Phosphatgerüstes mit einem Gegenion auszutauschen, das die Peakverbreiterung vermindert.

**[0060]** Die basenspezifisch terminierten Fragmentfamilien können durch Standardsequenzierung nach Sanger erzeugt werden, wobei das Große Klenowfragment aus E. coli-DNS-Polymerase I verwendet wird, durch Sequenase, Taq-DNS-Polymerase und andere für diesen Zweck geeignete DNS-Polymerasen, wobei verschachtelte DNS-Fragmente für die massenspektrometrische Analyse erzeugt werden. Es ist jedoch ein Teil der Erfindung, dass die basenspezifisch terminierten RNS-Transkripte der DNS-Fragmente, die sequenziert werden sollen, ebenso für die massenspektrometrische Sequenzbestimmung verwendet werden können. In

diesem Falle können verschiedene RNS-Polymerasen wie z. B. die SP6- oder die T7-RNS-Polymerase auf geeigneten Vektoren verwendet werden, die zum Beispiel den SP6- oder den T7-Promotor enthalten (z. B. Axelrod et al., "Transcription from Bacteriophage T7 and SP6 RNA Polymerase Promoters in the Presence of 3'-Deoxyribonukleoside 5'-triphosphate Chain Terminators", Biochemistry 24, 5716–23 (1985)). In diesem Falle werden die unbekannten DNS-Sequenzfragmente stromabwärts derartiger Promotoren inseriert. Eine Transkription kann ebenso durch einen Nukleinsäureprimer initiiert werden (Pitulle et al., "Initiator Oligonucleotides for the Combination of Chemical and Enzymatic RNA Synthesis", Gene 112, 101–105 (1992)), der in einer erfindungsgemäßen Ausführungsform geeignete Bindungsfunktionalitäten L trägt, die wie oben umrissen die Immobilisierung der verschachtelten RNS-Fragmente vor der massenspektrometrischen Analyse zur Reinigung und/oder zur geeigneten Modifizierung und/oder Konditionierung ermöglichen.

**[0061]** Für dieses Immobilisierungsverfahren der DNS-/RNS-Sequenzierungsprodukte zur massenspektrometrischen Analyse können verschiedene feste Träger verwendet werden, z. B. Kügelchen (Silicagel, kontrolliertes Porenglas, magnetische Kügelchen, Sephadex-/Sepharosekügelchen, Cellulosekügelchen etc.), Kapillaren, Glasfiberfilter, Glasoberflächen, Metalloberflächen oder Plastikmaterial. Beispiele verwendbarer Plastikmaterialien schließen Membranen im Filter- oder Mikrotiterplattenformat ein, wobei die letzteren die Automatisierung des Reinigungsverfahrens durch Verwenden von Mikrotiterplatten ermöglichen, die als eine erfindungsgemäße Ausführungsform eine permeable Membran auf dem Boden des Loches tragen, die mit L' funktionalisiert ist. Membranen können auf Polyethylen, Polypropylen, Polyamid, Polyvinylidendifluorid und ähnlichem basieren. Beispiele geeigneter Metalloberflächen schließen Stahl, Gold, Silber, Aluminium und Kupfer ein. Nach Reinigung, Kationenaustausch und/oder Modifizierung des Phosphodiestergerüstes der L-L'-gebundenen verschachtelten Fragmente nach Sanger können sie chemisch, enzymatisch oder physikalisch vom festen Träger abgespalten werden. Ebenso können die L-L'-gebundenen Fragmente vom Träger abgespalten werden, wenn sie einer massenspektrometrischen Analyse unterzogen werden, wobei geeignet gewählte L-L'-Bindungen und entsprechende Laserenergien/-intensitäten wie oben und in den **Abb. 19**–**Abb. 23** beschrieben verwendet werden.

**[0062]** Die hochgereinigten vier basenspezifisch terminierten DNS- oder RNS-Fragmentfamilien werden dann in Bezug auf ihre Fragmentlänge über Bestimmung ihres jeweiligen Molekulargewichtes durch MALDI- oder ES-Massenspektrometrie analysiert.

**[0063]** Für ES werden die in Wasser oder einem flüchtigen Puffer gelösten Proben entweder kontinuierlich oder diskontinuierlich in ein Ionisationsinterface bei atmosphärischem Druck (atmospheric pressure ionization interface, API) injiziert und dann durch einen Quadrupol massenanalysiert. Mit Hilfe eines Computerprogramms werden die Molekulargewichtspeaks für das bekannte Molekulargewicht des Nukleinsäureprimers (UP) gesucht, und es wird bestimmt, welches der vier kettenterminierenden Nukleotide an das UP hinzugefügt wurde. Dies stellt das erste Nukleotid der unbekannten Sequenz dar. Dann kann das zweite, dritte, n-te Verlängerungsprodukt auf eine ähnliche Weise identifiziert werden, und dadurch wird die Nukleinsäuresequenz zugewiesen. Die Erzeugung mehrfacher Ionenpeaks, die durch die Verwendung von ES-Massenspektrometrie erhalten werden können, kann die Genauigkeit der Massebestimmung erhöhen.

**[0064]** Bei der MALDI-Massenspektrometrie können verschiedene Massenanalysatoren verwendet werden, z. B. Magnetsektor-/Magnetdeflektorinstrumente im Einzel- oder Tripelquadrupolmodus (MS/MS). Fouriertransformation und Flugzeit- (time-of-flight, TOF) Konfigurationen sind im Stand der Technik für Massenspektrometrie bekannt. Die **Abb. 2A** bis **Abb. 6** sind als Beispiel für die Daten gegeben, die erhalten werden können, wenn ein hypothetisches DNS-Fragment mit einer Länge von 50 Nukleotiden (SEQ ID NO: 3) sequenziert wird, das ein Molekulargewicht von 15.344,02 Dalton hat. Die Molekulargewichte werden angegeben, die für ddT- (**Abb. 2A** und **Abb. 2B**), ddA- (**Abb. 3A** und **Abb. 3B**), ddG- (**Abb. 4A** und **Abb. 4B**) und ddC- (**Abb. 5A** und **Abb. 5B**) terminierte Produkte berechnet wurden (entsprechend Fragmenten aus SEQ ID NO: 3), und die idealisierten vier MALDI-TOF-Massenspektren werden gezeigt. Alle vier Spektren werden überlagert, und hieraus kann die DNS-Sequenz erzeugt werden. Dies wird in der zusammenfassenden **Abb. 6** gezeigt, die demonstriert, wie die Molekulargewichte mit der DNS-Sequenz korrelieren. MALDI-TOF-Spektren wurden für die ddT-terminierten Produkte erzeugt (**Abb. 16A**–**Abb. 16M**), die denjenigen in den **Abb. 2A**–**Abb. 2B** entsprechen, und diese Spektren wurden überlagert (**Abb. 17A** und **Abb. 17B**). Die Korrelation berechneter Molekulargewichte der ddT-Fragmente und ihrer experimentell verifizierten Gewichte ist in Tabelle 1 gezeigt. Entsprechend können, wenn alle vier kettenterminierten Reaktionen kombiniert und dann durch Massenspektrometrie analysiert werden, die Molekulargewichtsunterschiede zwischen zwei benachbarten Peaks verwendet werden, um die Sequenz zu bestimmen. Für das Desorptions-/Ionisationsverfahren können zahlreiche Matrix-/Laserkombinationen verwendet werden.

TABELLE I

Korrelation der berechneten und experimentell verifizierten Molekulargewichte der 13 DNS-Fragmente der ABBILDUNGEN 2 und 16A–16M

| Fragment (n-mer) | berechnete Masse | experimentelle Masse | Unterschied |
|---|---|---|---|
| 7mer | 2104,45 | 2119,9 | +15,4 |
| 10mer | 3011,04 | 3026,1 | +15,1 |
| 11mer | 3315,24 | 3330,1 | +14,9 |
| 19mer | 5771,82 | 5788,0 | +16,2 |
| 20mer | 6076,02 | 6093,8 | +17,8 |
| 24mer | 7311,82 | 7374,9 | +63,1 |
| 26mer | 7945,22 | 7960,9 | +15,7 |
| 33mer | 10112,63 | 10125,3 | +12,7 |
| 37mer | 11348,43 | 11361,4 | +13,0 |
| 38mer | 11652,62 | 11670,2 | +17,6 |
| 42mer | 12872,42 | 12888,3 | +15,9 |
| 46mer | 14108,22 | 14125,0 | +16,8 |
| 50mer | 15344,02 | 15362,6 | +18,6 |

**[0065]** Um den Durchsatz auf einen Level zu erhöhen, der für großvolumige genomische und cDNS-Sequenzierungsprojekte notwendig ist, ermöglicht die vorliegende Erfindung die Verwendung von Multiplex-Massenspektrometrie, um mehr als eine Sequenz simultan zu bestimmen. Dies kann durch mehrere, wenn auch unterschiedliche Methodologien erzielt werden, wobei das Basisprinzip die Massemodifikation des Nukleinsäureprimers (UP), der kettenverlängernden und/oder kettenterminierenden Nukleosidtriphosphate ist, oder die Verwendung von massedifferenzierten Tag-Sonden, die mit spezifischen Tag-Sequenzen hybridisieren können. Der Begriff "Nukleinsäureprimer" wie hier verwendet umfasst Primer für sowohl DNS- und RNS-Sequenzierung nach Sanger.

**[0066]** Als Beispiel zeigt **Abb. 7A** eine allgemeine Formel des Nukleinsäureprimers (UP) und der Tag-Sonden (TS). Die massemodifizierende Gruppe kann zum Beispiel an entweder das 5'-Ende des Oligonukleotids ($M^1$), an die Nukleobase (oder Basen)($M^2$, $M^7$), an das Phosphatgerüst ($M^3$) und an die 2'-Position des Nukleosids (der Nukleoside)($M^4$, $M^6$) oder/und an die terminale 3'-Position ($M^5$) gebunden werden. Die Primerlänge kann zwischen einer Länge von 1 und 50 Nukleotiden variieren. Um die DNS-Sequenzierung nach Sanger zu primen, ist der Primer vorzugsweise im Bereich einer Länge von etwa 15 bis 30 Nukleotiden. Um die Transkription in einer RNS-Polymerase vermittelten Sequenzierungsreaktion nach Sanger künstlich zu primen, ist die Länge des Primers vorzugsweise im Bereich von etwa 2 bis 6 Nukleotiden. Wenn eine Tag-Sonde (TS) mit einer integrierten Tag-Sequenz einer Familie kettenterminierter Fragmente hybridisiert werden muss, beträgt seine bevorzugte Länge etwa 20 Nukleotide.

**[0067]** Die Tabelle in **Abb. 7B** beschreibt einige Beispiele massemodifizierter Primer-/Tag-Sondenkonfigurationen für eine DNS- als auch für eine RNS-Sequenzierung nach Sanger. Diese Liste ist jedoch nicht als limitierend gemeint, da zahlreiche Kombinationen von massemodifizierenden Funktionen und Positionen innerhalb des Oligonukleotidmoleküls möglich sind und als Teil der Erfindung betrachtet werden. Die massemodifizierende Funktionalität kann zum Beispiel ein Halogen, ein Azido oder von der Art XR sein, wobei X eine verbindende Gruppe und R eine massemodifizierende Funktionalität ist. Die massemodifizierende Funktionalität kann daher verwendet werden, um definierte Masse-Inkremente in das Oligonukleotidmolekül einzuführen.

**[0068]** In einer anderen Ausführungsform sind die Nukleotide, die für die Kettenverlängerung und/oder -termination verwendet werden, massemodifiziert. Beispiele solcher modifizierten Nukleotide sind in den **Abb. 8A** und **Abb. 8B** gezeigt. Hier kann die massemodifizierende Gruppe M entweder an die Nukleobase $M^2$ (im Falle

des $c^7$-Desazanukleosids ebenso an C-7, $M^7$), an die Triphosphatgruppe am alpha-Phosphat $M^3$ oder an die 2'-Position des Zuckerrings des Nukleosidtriphosphates $M^4$ und $M^6$ gebunden sein. Außerdem kann die massemodifizierende Funktionalität hinzugefügt werden, um die Kettentermination zu beeinflussen, zum Beispiel durch seine Bindung an die 3'-Position des Zuckerrings im Nukleosidtriphosphat $M^5$. Die Liste in **Abb. 8B** stellt Beispiele möglicher Konfigurationen zur Erzeugung kettenterminierender Nukleosidtriphosphate zur RNS- oder DNS-Sequenzierung nach Sanger dar. Für Fachleute ist es jedoch klar, dass viele andere Kombinationen diesem erfindungsgemäßen Zweck ebenso gut dienen können. Auf dieselbe Weise erkennen Fachleute, dass kettenverlängernde Nukleosidtriphosphate ebenso auf dieselbe Art mit zahlreichen Variationen und Kombinationen in Funktionalitäts- und Bindungspositionen massemodifiziert werden können.

**[0069]** Ohne den Umfang der Erfindung zu beschränken, gibt **Abb. 9** eine detailliertere Beschreibung bestimmter Beispiele, wie die Massemodifikation M für X in XR eingebracht werden kann als auch wie Oligo-/Polyethylenderivate für R verwendet werden können. Das massemodifizierende Inkrement ist in diesem Falle 44, d. h. fünf verschiedene massemodifizierte Spezies können erzeugt werden, indem nur m von 0 bis 4 verändert wird und dadurch Masseeinheiten von 45 (m = 0), 89 (m = 1), 133 (m = 2), 177 (m = 3) und 221 (m = 4) zum Nukleinsäureprimer (UP), zur Tag-Sonde (TS) bzw. zum Nukleosidtriphosphat hinzugefügt werden. Die Oligo-/Polyethylenglycole können ebenso durch ein niederes Alkyl wie zum Beispiel Methyl, Ethyl, Propyl, Isopropyl, t-Butyl und ähnliches monoalkyliert sein. Eine Auswahl der bindenden Funktionalitäten X ist ebenso veranschaulicht. Andere Chemien können in den massemodifizierten Verbindungen verwendet werden, wie zum Beispiel diejenigen, die in Oligonucleotides and Analogues, A Practical Approach, Hrsg. F. Eckstein, IRL Press, Oxford 1991, kürzlich beschrieben wurden.

**[0070]** In noch einer anderen Ausführungsform können verschiedene massemodifizierende Funktionalitäten R, die von Oligo-/Polyethylenglycolen verschieden sind, ausgewählt und über geeignete Linkerchemien X gebunden werden. Ohne jede Beschränkung sind einige Beispiele in **Abb. 10** gegeben. Eine einfache Massemodifikation kann durch Substituieren von H durch Halogene wie F, Cl, Br und/oder I oder Pseudohalogene wie zum Beispiel SCN, NCS oder durch Verwendung unterschiedlicher Alkyl-, Aryl- oder Aralkylgruppen wie zum Beispiel Methyl, Ethyl, Propyl, Isopropyl, t-Butyl, Hexyl, Phenyl, substituiertes Phenyl, Benzyl oder funktioneller Gruppen wie zum Beispiel $CH_2F$, $CHF_2$, $CF_3$, $Si(CH_3)_3$, $Si(CH_3)_2(C_2H_5)$, $Si(CH_3)(C_2H_5)_2$, $Si(C_2H_5)_3$ erzielt werden. Noch eine andere Massemodifikation kann erhalten werden, indem Homo- und Heteropeptide durch X an das UP, die TS oder die Nukleosidtriphosphate gebunden werden. Ein Beispiel, das zur Erzeugung massemodifizierter Spezies mit einem Masse-Inkrement von 57 verwendbar ist, ist die Bindung von Oligoglycinen, d. h. es werden Massemodifikationen von 74 (r = 1, m = 0), 131 (r = 1, m = 2), 188 (r = 1, m = 3), 245 (r = 1, m = 4) erzielt. Einfache Oligoamide können ebenso verwendet werden, z. B. können Massemodifikationen von 74 (r = 1, m = 0), 88 (r = 2, m = 0), 102 (r = 3, m = 0), 116 (r = 4, m = 0) etc. erhalten werden. Für Fachleute ist es offensichtlich, dass es zahlreiche Möglichkeiten zusätzlich zu denjenigen in **Abb. 10** und in der oben erwähnten Referenz gibt (Oligonucleotides and Analogues, F. Eckstein, 1991), auf vorbestimmte Weise viele verschiedene massemodifizierende Funktionalitäten in UP, TS und Nukleosidtriphosphate einzuführen, die für eine DNS- und RNS-Sequenzierung nach Sanger akzeptabel sind.

**[0071]** Wie hier verwendet bezeichnet der hochgestellte Index "0 – i" i + 1 massedifferenzierte Nukleotide, Primer und Tags. In einigen Beispielen kann der hochgestellte Index 0 (z. B. $NTP^0$, $UP^0$) eine nichtmodifizierte Spezies eines bestimmten Reaktanten bezeichnen, und der hochgestellte Index i (z. B. $NTP^i$, $NTP^1$, $NTP^2$ etc.) kann die i-te massemodifizierte Spezies dieses Reaktanten bezeichnen. Wenn zum Beispiel mehr als eine Nukleinsäurespezies (z. B. DNS-Klone) gleichzeitig durch Multiplex-DNS-Sequenzierung sequenziert werden muss, dann können i + 1 verschiedene massemodifizierte Nukleinsäureprimer ($UP^0$, $UP^1$ ... $UP^i$) verwendet werden, um jeden Satz von basenspezifisch terminierten Fragmenten zu unterscheiden, wobei jede Spezies von massemodifizierten $UP^i$ durch Massenspektrometrie vom Rest unterschieden werden kann.

**[0072]** Drei verschiedene Basisverfahren zur multiplex-massenspektrometrischen DNS-Sequenzierung, die die beschriebenen massemodifizierten Reagenzien verwenden, sind unten beschrieben:

A) Multiplexen durch Verwenden von massemodifizierten Nukleinsäureprimern (UP) zur DNS- oder RNS-Sequenzierung nach Sanger (siehe zum Beispiel **Abb. 11**),

B) Multiplexen durch Verwenden von massemodifizierten Nukleosidtriphosphaten als Kettenverlängerer und/oder Kettenterminatoren zur DNS- oder RNS-Sequenzierung nach Sanger (siehe zum Beispiel **Abb. 12**), und

C) Multiplexen durch Verwenden von Tag-Sonden, die spezifisch mit Tag-Sequenzen hybridisieren, die in einen Teil der vier basenspezifisch terminierten DNS-RNS-Fragmentfamilien nach Sanger integriert sind. Eine Massemodifikation kann hier erzielt werden, wie für die **Abb. 7A**, **Abb. 7B**, **Abb. 9** und **Abb. 10** beschrieben ist, oder alternativ durch Konstruktion unterschiedlicher Oligonukleotidsequenzen, die dieselbe

oder eine unterschiedliche Länge zu nichtmodifizierten Nukleotiden haben, die auf vorbestimmte Weise passend differenzierte Molekulargewichte (siehe zum Beispiel Abb. 13) erzeugen.

**[0073]** Das Multiplexverfahren durch massemodifizierte Nukleinsäureprimer (UP) wird durch ein Beispiel in Abb. 11 für das gleichzeitige Massenanalysieren von vier verschiedenen DNS-Klonen dargestellt. Die erste Reaktionsmischung wird durch Standard-DNS-Sequenzierung nach Sanger erhalten, wobei ein unbekanntes DNS-Fragment 1 (Klon 1) in einen geeigneten Vektor (z. B. M13mp18) integriert ist, wobei ein nichtmodifizierter Nukleinsäureprimer $UP^0$ und eine Standardmischung der vier nichtmodifizierten Desoxynukleosidtriphosphate $dNTP^0$ mit 1/10 eines der vier Didesoxynukleosidtriphosphaten $ddNTP^0$ verwendet wird. Eine zweite Reaktionsmischung für das DNS-Fragment 2 (Klon 2) wird erhalten, indem ein massemodifizierter Nukleinsäureprimer $UP^1$ und, wie oben, die vier nichtmodifizierten Nukleosidtriphosphate $dNTP^0$, die in jeder getrennten Sanger-Reaktion 1/10 der kettenterminierenden nichtmodifizierten Didesoxynukleosidtriphosphate $ddNTP^0$ enthalten, verwendet werden. In den beiden anderen Experimenten haben die vier Sanger-Reaktionen die folgenden Zusammensetzungen: DNS-Fragment 3 (Klon 3), $UP^2$, $dNTP^0$, $ddNTP^0$ und DNS-Fragment 4 (Klon 4), $UP^3$, $dNTP^0$, $ddNTP^0$. Zur massenspektrometrischen DNS-Sequenzierung werden alle basenspezifisch terminierten Reaktionen der vier Klone gepoolt und massenanalysiert. Die verschiedenen Massepeaks, die zu den vier didesoxyterminierten (z. B. ddT-terminierten) Fragmentfamilien gehören, werden spezifisch verlängerten und ddT-terminierten Fragmenten durch Suchen (zum Beispiel durch ein Computerprogramm) für die bekannten molekularen Ionenpeaks von $UP^0$, $UP^1$, $UP^2$ und $UP^3$, verlängert durch eines der vier Didesoxynukleosidtriphosphate $UP^0$-$ddN^0$, $UP^1$-$ddN^0$, $UP^2$-$ddN^0$ und $UP^3$-$ddN^0$, zugewiesen. Auf diese Weise werden die ersten Nukleotide der vier unbekannten DNS-Sequenzen der Klone 1 bis 4 bestimmt. Das Verfahren wird wiederholt, wobei die molekularen Massen der vier spezifischen ersten Verlängerungsprodukte gespeichert werden, bis die vier Sequenzen zugewiesen sind. Unzweideutige Masse/Sequenzzuweisungen sind sogar in einem Szenario des schlimmsten Falles möglich, in dem die vier massemodifizierten Nukleinsäureprimer durch dasselbe Didesoxynukleosidtriphosphat verlängert werden, wobei die Verlängerungsprodukte dann zum Beispiel $UP^0$-ddT, $UP^1$-ddT, $UP^2$-ddT und $UP^3$-ddt sind, die sich durch das bekannte Masse-Inkrement unterscheiden, das die vier Nukleinsäureprimer voneinander unterscheddet. In einer anderen Ausführungsform wird eine analoge Technik verwendet, die unterschiedliche Vektoren verwendet, die zum Beispiel die SP6- und/oder die T7-Promotorsequenzen enthalten und eine Transkription mit den Nukleinsäureprimern $UP^0$, $UP^1$, $UP^2$ und $UP^3$ und jeweils einer RNS-Polymerase (z. B: SP6- oder T7-RNS-Polymerase) mit kettenverlängernden und -terminierenden nichtmodifizierten Nukleosidtriphosphaten $NTP^0$ und 3'-$dNTP^0$ leisten. Dabei wird die DNS-Sequenz durch RNS-Sequenzierung nach Sanger bestimmt.

**[0074]** Abb. 12 veranschaulicht das Verfahren zum Multiplexen durch massemodifizierte kettenverlängernde oder/und -terminierende Nukleosidtriphosphate, in dem drei unterschiedliche DNS-Fragmente (3 Klone) gleichzeitig massenanalysiert werden. Die erste DNS-Sequenzierungsreaktion nach Sanger (DNS-Fragment 1, Klon 1) ist die Standardmischung, die nichtmodifizierte Nukleinsäureprimer $UP^0$, $dNTP^0$ verwendet und in jeder der vier Reaktionen eines der vier $ddNTP^0$. Die zweite (DNS-Fragment 2, Klon 2) und die dritte (DNS-Fragment 3, Klon 3) haben den folgenden Inhalt: $UP^0$, $dNTP^0$, $ddNTP^1$ bzw. $UP^0$, $dNTP^0$, $ddNTP^2$. In einer Variante dieses Verfahrens kann die Amplifizierung des Masse-Inkrementes beim Massemodifizieren der verlängerten DNS-Fragmente durch die Verwendung von gleich massemodifizierten Desoxynukleosidtriphosphaten (d. h. $dNTP^1$, $dNTP^2$) zur Kettenverlängerung entweder allein oder in Verbindung mit dem homologen gleich massemodifizierten Didesoxynukleosidtriphosphat erzielt werden. Für die drei oben beschriebenen Klone kann der Inhalt der Reaktionsmischung wie folgt sein: entweder $UP^0$/$dNTP^0$/$ddNTP^0$, $UP^0$/$dNTP^1$/$ddNTP^0$ und $UP^0$/$dNTP^2$/$ddNTP^0$ oder $UP^0$/$dNTP^0$/$IddNTP^0$, $UP^0$/$dNTP^1$/$ddNTP^1$ und $UP^0$/$dNTP^2$/$ddNTP^2$. Wie oben beschrieben kann DNS-Sequenzierung durch eine RNS-Sequenzierung nach Sanger durchgeführt werden, die nichtmodifizierte Nukleinsäureprimer $UP^0$ und eine geeignete Mischung kettenverlängernder und -terminierender Nukleosidtriphosphate verwendet. Die Massemodifizierung kann wiederum entweder im kettenterminierenden Nukleosidtriphosphat allein oder in Verbindung mit massemodifizierten kettenverlängernden Nukleosidtriphosphaten erfolgen. Multiplexen wird durch Poolen der drei basenspezifisch terminierten Sequenzierungsreaktionen (z. B. die ddTTP-terminierten Produkte) und simultanem Analysieren der gepoolten Produkte durch Massenspektrometrie erzielt. Die ersten Verlängerungsprodukte der bekannten Nukleinsäureprimersequenz werden wiederum z. B. durch ein Computerprogramm zugewiesen. Masse/Sequenzzuweisungen sind sogar im schlimmsten Falle möglich, in dem die Nukleinsäureprimer durch dasselbe Nukleotid, z. B. ddT, in allen drei Klonen verlängert/terminiert wurden. Die folgenden auf diese Weise erhaltenen Konfigurationen können durch ihre unterschiedlichen Massemodifikationen gut differenziert werden: $UP^0$-$ddT^0$, $UP^0$-$ddT^1$, $UP^0$-$ddT^2$.

**[0075]** DNS-Sequenzierung durch Multiplex-Massenspektrometrie kann durch Klonen der DNS-Fragmente, die sequenziert werden sollen, in "Plex-Vektoren" erzielt werden, die vektorspezifische "Tag-Sequenzen" wie

beschrieben enthalten (Köster et al., "Oligonucleotide Synthesis and Multiplex DNA Sequencing Using Chemiluminescent Detection", Nucleic Acids Res. Symposium Ser. Nr. 24, 318–321 (1991)), dann durch Poolen von Klonen verschiedener Plex-Vektoren für die DNS-Präparation und für die vier getrennten Sequenzierungsreaktionen nach Sanger, wobei Standard-dNTP$^0$/ddNTP$^0$ und ein Nukleinsäureprimer UP$^0$ verwendet werden, durch Reinigen der vier Multiplexfragmentfamilien durch Binden an einen festen Träger durch die Linkergruppe L am 5'-Ende von UP, durch Auswaschen aller Nebenprodukte und Abspalten der gereinigten Multiplex-DNS-Fragmente vom Träger oder Verwenden der L-L'-gebundenen verschachtelten Sanger-Fragmente als solche zur massenspektrometrischen Analyse wie oben beschrieben, durch Durchführen von Demultiplexen durch Hybridisieren mit spezifischen "Tag-Sonden", wobei eine nach der anderen hybridisiert wird, und durch darauffolgendes Analysieren durch Massenspektrometrie (siehe zum Beispiel **Abb. 13**). Als Referenzpunkt laufen die vier basenspezifisch terminierten Multiplex-DNS-Fragmentfamilien durch das Massenspektrometer, und alle ddT$^0$-, ddA$^0$-, ddC$^0$- und ddG$^0$-terminierten molekularen Ionenpeaks werden jeweils detektiert und gespeichert. Die Zuweisung von zum Beispiel ddT$^0$-terminierten DNS-Fragmenten zu einer spezifischen Fragmentfamilie wird durch eine andere massenspektrometrische Analyse nach Hybridisieren einer spezifischen Tag-Sonde (TS) an die entsprechende Tag-Sequenz geleistet, die in der Sequenz dieser spezifischen Fragmentfamilie enthalten ist. Nur diejenigen molekularen Ionenpeaks, die im Stande sind, mit der spezifischen Tag-Sonde zu hybridisieren, werden zu einer höheren molekularen Masse durch dasselbe bekannte Masse-Inkrement (z. B. der Tag-Sonde) verschoben. Diese verschobenen Ionenpeaks aufgrund aller Hybridisierungen einer spezifischen Tag-Sonde gehören zu derselben Fragmentfamilie. Für eine gegebene Fragmentfamilie wird dies für die verbleibenden kettenterminierten Fragmentfamilien mit derselben Tag-Sonde wiederholt, um die vollständige DNS-Sequenz zuzuweisen. Dieses Verfahren wird i–1 mal entsprechend den i gemultiplexten Klonen wiederholt (der i-te Klon wird standardmäßig identifiziert).

**[0076]** Die Differenzierung der Tag-Sonden für die unterschiedlichen gemultiplexten Klone können gerade durch die DNS-Sequenz und ihre Fähigkeit zur Watson-Crick-Basenpaarung an die Tag-Sequenz erhalten werden. Es ist im Stand der Technik gut bekannt, wie stringente Bedingungen berechnet werden, um eine spezifische Hybridisierung einer gegebenen Tag-Sonde mit einer gegebenen Tag-Sequenz bereitzustellen (siehe zum Beispiel Molecular Cloning: A laboratory manual. 2. Aufl., Hrsg. Sambrook, Fritsch und Maniatis (Cold Spring Harbor Laboratory Press: NY, 1989, Kapitel 11). Außerdem kann eine Differenzierung erhalten werden, indem die Tag-Sequenz für jeden Plex-Vektor konstruiert wird, so dass er eine ausreichende Massedifferenz hat, damit er gerade durch Veränderung der Länge oder der Basenzusammensetzung oder durch Massemodifikationen gemäß **Abb. 7A**, **Abb. 7B**, **Abb. 9** und **Abb. 10** einzigartig wird. Um die Duplex zwischen der Tag-Sequenz und der Tag-Sonde während der massenspektrometrischen Analyse intakt zu halten, stellt eine andere erfindungsgemäße Ausführungsform eine kovalente Bindung bereit, die zum Beispiel durch photoreaktive Gruppen wie zum Beispiel Psoralen und Ellipticin und durch andere Verfahren, die dem Fachmann bekannt sind, vermittelt wird (siehe zum Beispiel Helene et al., Nature 344, 358 (1990) und Thuong et al. "Oligonucleotides Attached to Intercalators. Photoreaktive and Cleavage Agents" in F. Eckstein, Oligonucleotides and Analogues: A practical Approach, IRL Press, Oxford, 1991, 283–306).

**[0077]** Die DNS-Sequenz wird nochmals enträtselt, indem nach dem Ionenpeak mit dem niedrigsten Molekulargewicht gesucht wird, der dem bekannten UP$^0$-Tag-Sequenz/Tag-Sondenmolekulargewicht plus dem ersten Verlängerungsprodukt, z. B. ddT$^0$, entspricht, dann dem zweiten, dem dritten etc.

**[0078]** In einer Kombination des letzteren Ansatzes mit den vorher beschriebenen Multiplexverfahren kann eine weitere Zunahme beim Multiplexen erzielt werden, indem zusätzlich zu den Tag-Sonden/Tag-Sequenz-wechselwirkungen massemodifizierte Nukleinsäureprimer (**Abb. 7A** und **Abb. 7B**) und/oder massemodifizierte Desoxynukleoside dNTP$^{0\text{-}i}$ und/oder Didesoxynukleosidtriphosphate ddNTP$^{0\text{-}i}$ verwendet werden. Fachleute erkennen, dass der Tag-Sequenz-/Tag-Sonden-Multiplexansatz nicht auf die DNS-Sequenzierung nach Sanger beschränkt ist, die verschachtelte DNS-Fragmente mit DNS-Polymerasen erzeugt. Die DNS-Sequenz kann ebenso bestimmt werden, indem die unbekannte DNS-Sequenz von passenden einen Promotor enthaltenden Vektoren (siehe oben) mit verschiedenen RNS-Polymerasen und Mischungen von NTP$^{0\text{-}i}$/3'-dNTP$^{0\text{-}i}$ transkribiert wird und auf diese Weise verschachtelte RNS-Fragmente erzeugt werden.

**[0079]** In noch einer anderen erfindungsgemäßen Ausführungsform kann die massemodifizierende Funktionalität durch ein zwei- oder vielschrittiges Verfahren eingeführt werden. In diesem Fall werden in einem ersten Schritt die Nukleinsäureprimer, die kettenverlängernden oder -terminierenden Nukleosidtriphosphate und/oder die Tag-Sonden durch eine Vorläuferfunktionalität modifiziert, wie zum Beispiel Azido -$N_3$, oder mit einer funktionellen Gruppe modifiziert, in der das R in XR H ist (**Abb. 7A**, **Abb. 7B**, **Abb. 9**) und dadurch vorläufige. Funktionen bereitstellt, d. h. -OH, -$NH_2$, -NHR, -SH, -NCS, -OCO$(CH_2)_r$COOH (r = 1–20), -NHCO$(CH_2)_r$COOH (r = 1–20), -$OSO_2$OH, -OCO$(CH_2)_r$I (r = 1–20), -OP(O-Alkyl)N$(Alkyl)_2$, die aber nicht darauf beschränkt sind. Diese

weniger massigen Funktionalitäten führen zu besseren Substrateigenschaften für die enzymatische DNS- oder RNS-Synthesereaktionen des DNS-Sequenzierungsverfahrens. Die geeignete massemodifizierende Funktionalität wird dann nach der Erzeugung der verschachtelten basenspezifisch terminierten DNS- oder RNS-Fragmente vor der Massenspektrometrie eingeführt. Einige Beispiele von Verbindungen, die als massemodifizierende Funktionalitäten dienen können, sind in den **Abb. 9** und **Abb. 10** beschrieben, ohne den Umfang der Erfindung zu beschränken.

**[0080]** Kits zur Sequenzierung von Nukleinsäuren durch Massenspektrometrie schließen Kombinationen der oben beschriebenen Sequenzierungsreaktanten ein. Zum Beispiel umfassen Kits in einer Ausführungsform Reaktanten für eine multiplexmassenspektrometrische Sequenzierung mehrerer unterschiedlicher Nukleinsäurespezies. Der Kit kann einen festen Träger einschließen, der eine Linkerfunktionalität ($L^1$) zur Immobilisierung von basenspezifisch terminierten Produkten hat, wenigstens einen Nukleinsäureprimer, der eine Linkergruppe (L) zur reversiblen und vorübergehenden Verbindung des Primers an den festen Träger durch zum Beispiel eine photospaltbare Bindung, einen Satz von kettenverlängernden Nukleotiden (z. B. dATP, dCTP, dGTP und dTTP, oder ATP, CTP, GTP und UTP), einen Satz kettenterminierender Nukleotide (wie z. B. 2',3'-Dideoxynukleotide zur DNS-Synthese oder 3'-Desoxynukleotide zur RNS-Synthese) und eine geeignete Polymerase zum Synthetisieren komplementärer Nukleotide. Primer und/oder terminierende Nukleotide können derartig massemodifiziert werden, so dass die basenspezifisch terminierten Fragmente, die aus einer der Nukleinsäurespezies, die sequenziert werden sollen, erzeugt werden, durch Massenspektrometrie von allen anderen unterschieden werden kann. Alternativ zur Verwendung der massemodifizierten Synthesereaktanten kann ein Satz von Tag-Sonden (wie oben beschrieben) in den Kit eingeschlossen werden. Der Kit kann ebenso geeignete Puffer als auch eine Anleitung zur Durchführung der Multiplex-Massenspektrometrie zum gleichzeitigen Sequenzieren vieler Nukleinsäurespezies enthalten.

**[0081]** In einer anderen Ausführungsform kann der Kit zum Nukleinsäuresequenzieren einen festen Träger wie oben beschrieben, einen Primer zum Initiieren der Synthese der komplementären Nukleinsäurefragmente, einen Satz kettenverlängernder Nukleotide und eine geeignete Polymerase umfassen. Die massemodifizierten kettenterminierenden Nukleotide werden derartig ausgewählt, dass das Hinzufügen von einem der Kettenterminatoren zu einer wachsenden komplementären Nukleinsäure durch Massenspektrometrie unterschieden werden kann.

## BEISPIEL 1

### Immobilisieren von Primerverlängerungsprodukten einer DNS-Sequenzierungsreaktion nach Sanger für eine massenspektrometrische Analyse über Disulfidbindungen

**[0082]** Als ein fester Träger werden Sequelonmembranen (Millipore Corp., Bedford, MA) mit Phenylisothiocyanatgruppen als Ausgangsmaterial verwendet. Die Membranscheiben mit einem Durchmesser von 8 mm werden mit einer Lösung von N-Methylmorpholin/Wasser/2-Propanol (NMM-Lösung)(2/49/49 V/V/V) befeuchtet, die überschüssige Flüssigkeit wird mit Filterpapier entfernt, und die Membranscheiben werden auf ein Stück Plastikfilm oder Aluminiumfolie gebracht, das sich auf einem Heizblock befindet, der auf 55°C eingestellt ist. Eine Lösung von 1 mM 2-Mercaptoethylamin (Cysteamin) oder 2,2'-Dithio-bis(ethylamin)(Cystamin) oder S-(2-Thiopyridyl)-2-thio-ethylamin (10 µl, 10 nmol) in NMM wird pro Scheibe hinzugefügt und auf 55°C erwärmt. Nach 15 min werden 10 µl NMM-Lösung pro Scheibe hinzugefügt und für weitere 5 min erwärmt. Ein Überschuss an Isothiocyanatgruppen kann durch Behandlung mit 10 µl einer 10 mM Lösung von Glycin in NMM-Lösung entfernt werden. Für Cystamin werden die Scheiben mit 10 µl einer Lösung von 1 M wässerigem Dithiothreitol (DTT)/2-Propanol (1 : 1 V/V) für 15 min bei Raumtemperatur behandelt. Dann werden die Scheiben ausgiebig in einem Filtrationssammler mit 5 Aliquots von jeweils 1 ml der NMM-Lösung gewaschen, dann mit 5 Aliquots von 1 ml Acetonitril/Wasser (1/1 V/V) gewaschen und danach getrocknet. Wenn sie nicht sofort verwendet werden, werden die Scheiben mit freien Thiolgruppen in einer Lösung von 1 M wässerigem Dithiothreitol/2-Propanol (1 : 1 V/V) gelagert, und vor der Verwendung wird DTT durch drei Waschungen mit jeweils 1 ml der NMM-Lösung entfernt. Die Primeroligonukleotide mit einer 5'-SH-Funktionalität können durch verschiedene Verfahren hergestellt werden (z. B. B. C. F. Chu et al., Nucleic Acids Res. 14, 5591–5603 (1986), Sproat et al., Nucleic Acids Res. 15, 4837–48 (1987) und Oliqonucleotides and Analogues: A Practical Approach (Herausgeber F. Eckstein), IRL Press Oxford, 1991). Sequenzierungsreaktionen gemäß dem Protokoll nach Sanger werden auf Standardweise durchgeführt (z. B. H. Swerdlow et al., Nucleic Acids Res. 18, 1415–19 (1990)). In der Anwesenheit von etwa 7–10 mM DTT kann der freie 5'-Thiolprimer verwendet werden, in anderen Fällen kann die SH-Funktionalität während der Sequenzierungsreaktion nach Sanger z. B. durch eine Tritylgruppe geschützt werden, die vor dem Ankern an den Träger auf die folgende Weise entfernt wird. Die vier Sequenzierungsreaktionen (jeweils 150 µl in einem Eppendorfröhrchen) werden durch eine 10minütige Inku-

bation bei 70°C beendet, um die DNS-Polymerase (z. B. Klenow-Fragment, Sequenase) zu denaturieren, und die Reaktionsmischungen werden ethanolgefällt. Die Überstände werden entfernt, und die Pellets werden mit 25 µl einer wässerigen 1 M Silbernitratlösung gevortext, und nach einer Stunde bei Raumtemperatur werden 50 µl einer wässerigen 1 M DTT-Lösung hinzugefügt und durch Vortexen gemischt. Nach 15 min werden die Mischungen zentrifugiert, und die Pellets werden zweimal mit 100 µl Ethylacetat durch Vortexen und Zentrifugieren gewaschen, um überschüssiges DTT zu entfernen. Die Primerverlängerungsprodukte mit einer freien 5'-Thiolgruppe sind nun an den thiolierten Membranträger unter milden oxidierenden Bedingungen gekoppelt. Im Allgemeinen ist es ausreichend, die 5'-thiolierten Primerverlängerungsprodukte in 10 µl 10 mM entgastem Triethylammoniumacetatpuffer (TEAA) pH 7,2 gelöst zu den thiolierten Membranträgern hinzuzufügen. Koppeln wird durch Trocknen der Proben auf den Membranscheiben mit einem kalten Ventilator erzielt. Das Verfahren kann wiederholt werden, indem die Membran mit 10 µl 10 mM TEAA-Puffer pH 7,2 angefeuchtet wird und wie oben getrocknet wird. Wenn 2-Thiopyridyl-derivatisierte Verbindungen verwendet werden, kann das Ankern durch die Freisetzung von Pyridin-2-thion spektrophotometrisch bei 343 nm beobachtet werden.

**[0083]** In einer anderen Variante dieses Ansatzes wird der Oligonukleotidprimer mit einer Aminogruppe am 5'-Ende funktionalisiert, die durch Standardverfahren während einer automatisierten DNS-Synthese eingebaut wird. Nach der Primerverlängerung während des Sequenzierungsverfahrens nach Sanger wird die primäre Aminogruppe mit 3-(2-Pyridyldithio)propionsäure-N-hydroxysuccinimidester (SPDP) zur Reaktion gebracht und danach an die thiolierten Träger gekoppelt und über die Freisetzung von Pyridyl-2-thion, wie oben beschrieben, beobachtet. Nach Denaturierung der DNS-Polymerase und Ethanolfällung der Sequenzierungsprodukte werden die Überstände entfernt, und die Pellets werden in 10 µl 10 mM TEAA-Puffer pH 7,2 gelöst, und 10 µl einer 2 mM Lösung von SPDP in 10 mM TEAA werden hinzugefügt. Die Reaktionsmischung wird gevortext und für 30 min bei 25°C inkubiert. Überschüssiges SPDP wird dann durch drei Extraktionen (Vortexen, Zentrifugieren) mit jeweils 50 µl Ethanol entfernt, und die resultierenden Pellets werden in 10 µl 10 mM TEAA-Puffer pH 7,2 gelöst und an die thiolierten Träger gekoppelt (siehe oben).

**[0084]** Die Primerverlängerungsprodukte werden gereinigt, indem die Membranscheiben dreimal mit jeweils 100 µl NMM-Lösung und dreimal mit jeweils 100 µl 10 mM TEAA-Puffer pH 7,2 gewaschen werden. Die gereinigten Primerverlängerungsprodukte werden durch drei aufeinanderfolgende Behandlungen mit 10 µl 10 mM 2-Mercaptoethanol in 10 mM TEAA-Puffer pH 7,2 freigesetzt, lyophilisiert und mit entweder ES- oder MALDI-Massenspektrometrie analysiert.

**[0085]** Das Verfahren kann ebenso für massemodifizierte Nukleinsäureprimer $UP^{0-i}$ in einer analogen und geeigneten Weise verwendet werden, wobei die chemischen Eigenschaften der massemodifizierenden Funktionalitäten in Betracht gezogen werden.

## BEISPIEL 2

### Immobilisierung von Primerverlängerungsprodukten der DNS-Sequenzierungsreaktion nach Sanger zur massenspektrometrischen Analyse über die Lävulinylgruppe

**[0086]** 5-Aminolävulinsäure wird an der primären Aminogruppe mit der Fmoc-Gruppe geschützt, wobei 9-Fluorenylmethyl-N-succinimidylcarbonat verwendet wird, und wird dann in den N-Hydroxysuccinimidester (NHS-Ester) unter Verwendung von N-Hydroxysuccinimid und Dicyclohexylcarbodiimid unter Standardbedingungen überführt. Für die Sequenzierungsreaktionen nach Sanger werden Nukleinsäureprimer $UP^{0-i}$ verwendet, die mit einer primären Aminogruppe am 5'-Ende funktionalisiert sind, die durch Standardverfahren während einer automatisierten DNS-Synthese mit Aminolinker-Phosphoamiditen als letzter Syntheseschritt eingeführt werden. Eine Sequenzierung nach Sanger wird unter Standardbedingungen durchgeführt (siehe oben). Die vier Reaktionsmischungen (jeweils 150 µl in einem Eppendorfröhrchen) werden für 10 min auf 70°C erwärmt, um die DNS-Polymerase zu inaktivieren, ethanolgefällt, zentrifugiert und in 10 µl 10 mM TEAA-Puffer pH 7,2 resuspendiert. 10 µl einer 2 mM Lösung des Fmoc-5-Aminolävulinyl-NHS-Esters in 10 mM TEAA-Puffer werden hinzugefügt, gevortext und für 30 min bei 25°C inkubiert. Der Überschuss des Reagenzes wird durch Ethanolfällung und Zentrifugation entfernt. Die Fmoc-Gruppe wird durch Resuspendieren der Pellets in 10 µl einer Lösung von 20% Piperidin in N,N-Dimethylformamid/Wasser (1 : 1 V/V) abgespalten. Nach 15 min bei 25°C wird Piperidin durch drei Präzipitationen/Zentrifugationen mit jeweils 100 µl Ethanol gründlich entfernt, die Pellets werden in 10 µl einer Lösung von N-Methylmorpholin, 2-Propanol und Wasser (2/10/88 V/V/V) resuspendiert, und sie werden auf den festen Träger gekoppelt, der eine Isothiocyanatgruppe trägt. Im Falle der DITC-Sequelonmembran (Millipore Corp., Bedford, MA) werden die Membranen wie in BEISPIEL 1 hergestellt, und Koppeln wird auf einem Wärmeblock bei 55°C, wie oben beschrieben, erzielt. RNS-Verlängerungsprodukte werden auf analoge Weise immobilisiert. Das Verfahren kann auf andere feste Träger mit Isothiocyanatgrup-

pen auf ähnliche Weise angewendet werden.

**[0087]** Die immobilisierten Primerverlängerungsprodukte werden ausgiebig dreimal mit jeweils 100 µl NMM-Lösung und dreimal mit 100 µl 10 mM TEAA-Puffer pH 7,2 gewaschen. Die gereinigten Primerverlängerungsprodukte werden durch drei aufeinanderfolgende Behandlungen mit 10 µl 100 mM Hydraziniumacetatpuffer pH 6,5 freigesetzt, lyophilisiert und durch entweder ES- oder MALDI-Massenspektrometrie analysiert.

BEISPIEL 3

Immobilisierung von Primerverlängerungsprodukten einer Sequenzierungsreaktion nach Sanger zur massenspektrometrischen Analyse über eine trypsinempfindliche Bindung

**[0088]** Membranscheiben aus Sequelon-DITC mit einem Durchmesser von 8 mm (Millipore Corp, Bedford, MA) werden mit 10 µl NMM-Lösung angefeuchtet (N-Methylmorpholin/Propanol-2/Wasser, 2/49/49 V/V/V), und ein Linkerarm wird durch eine Reaktion mit 10 µl einer 10 mM Lösung von 1,6-Diaminohexan in NMM eingeführt. Das überschüssige Diamin wird durch drei Waschschritte mit 100 µl NMM-Lösung entfernt. Unter Verwendung von Standardpeptidsyntheseprotokollen werden zwei L-Lysinreste durch zwei aufeinanderfolgende Kondensationen mit N-Fmoc-N-tBoc-L-Lysinpentafluorphenylester verbunden, die endständige Fmoc-Gruppe wird mit Piperidin in NNM entfernt, und die freie α-Aminogruppe wird an 1,4-Phenylendiisothiocyanat (DITC) gekoppelt. Überschüssiges DITC wird durch drei Waschschritte mit 100 µl 2-Propanol entfernt, und die N-tBoc-Gruppen werden mit Trifuoressigsäure gemäß Standardpeptidsyntheseverfahren entfernt. Die Nukleinsäureprimerverlängerungsprodukte werden aus Oligonukleotiden hergestellt, die eine primäre Aminogruppe am 5'-Terminus tragen. Die vier DNS-Sequenzierungsreaktionsmischungen nach Sanger (jeweils 150 µl in Eppendorfröhrchen) werden für 10 min auf 70°C erwärmt, um die DNS-Polymerase zu inaktivieren, ethanolgefällt, und die Pellets werden in 10 µl einer Lösung aus N-Methylmorpholin, 2-Propanol und Wasser (2/10/88, V/V/V) resuspendiert. Diese Lösung wird auf die Lys-Lys-DITC-Membranscheiben übertragen, wobei die Scheiben auf einem Wärmeblock, der auf 55°C eingestellt ist, befestigt werden. Nach dem Trocknen werden 10 µl der NMM-Lösung hinzugefügt, und das Trockenverfahren wird wiederholt.

**[0089]** Die immobilisierten Primerverlängerungsprodukte werden dreimal ausgiebig mit jeweils 100 µl NMM-Lösung und dreimal mit 100 µl 10 mM TEAA-Puffer pH 7,2 gewaschen. Für die massenspektrometrische Analyse wird die Bindung zwischen den Primerverlängerungsprodukten und dem festen Träger durch Behandlung mit Trypsin unter Standardbedingungen gespalten, und die freigesetzten Produkte werden entweder durch ES- oder MALDI-Massenspektrometrie analysiert, wobei Trypsin als interner Massestandard dient.

BEISPIEL 4

Immobilisierung von Primerverlängerungsprodukten einer Sequenzierungsreaktion nach Sanger zur massenspektrometrischen Analyse über eine Pyrophosphatbindung

**[0090]** Die DITC-Sequelonmembran (Scheiben mit einem Durchmesser von 8 mm) werden wie in BEISPIEL 3 hergestellt, und 10 µl einer 10 mM Lösung von 3-Aminopyridinadenindinukleotid (APAD) (Sigma) in NMM-Lösung werden hinzugefügt. Das überschüssige APAD wird durch Waschen mit 10 µl NMM-Lösung entfernt, und die Scheiben werden mit 10 µl 10 mM Natriumperiodat in NMM-Lösung behandelt (15 min, 25°C). Überschüssiges Periodat wird entfernt, und die Primerverlängerungsprodukte der vier DNS-Sequenzierungsreaktionen nach Sanger (jeweils 150 µl in Eppendorfröhrchen), die Nukleinsäureprimer mit einer primären Aminogruppe am 5'-Ende verwenden, werden ethanolgefällt, in 10 µl einer Lösung aus N-Methylmorpholin/2-Propanol/Wasser (2/10/88, V/V/V) gelöst und an die 2',3'-Dialdehydgruppen des immobilisierten NAD-Analogons gekoppelt.

**[0091]** Die Primerverlängerungsprodukte werden ausgiebig mit der NMM-Lösung (dreimal mit jeweils 100 µl) und mit 10 mM TEAA-Puffer pH 7,2 (dreimal mit jeweils 100 µl) gewaschen, und die gereinigten Primerverlängerungsprodukte werden durch Behandlung mit entweder NADase oder Pyrophosphatase in 10 mM TEAA bei pH 7,2 bei 37°C für 15 min freigesetzt, lyophilisiert und durch entweder ES- oder MALDI-Massenspektrometrie analysiert, wobei die Enzyme als interner Massestandard dienen.

BEISPIEL 5

Synthese von Nukleinsäureprimern, die durch Glycinreste an der 5'-Position des Zuckerrestes des endständigen Nukleosids massemodifiziert sind

**[0092]** Oligonukleotide werden durch automatisierte Standard-DNS-Synthese unter Verwendung von β-Cyanoethylphosphoamiditen (H. Köster, Nucleic Acids Res. 12, 4539 (1984)) synthetisiert, und eine 5'-Aminogruppe wird am Ende der Festphasen-DNS-Synthese (z. B: Agrawal et al., Nucleic Acids Res. 14, 6227–45 (1986) oder Sproat et al., Nucleic Acids Res. 15, 6181–96 (1987)) eingeführt. Die Gesamtmenge der mit 0,25 µmol CPG-gebundenem Nukleosid gestarteten Oligonukleotidsynthese wird mit konzentriertem wässerigem Ammoniak entschützt, über OligoPAK™-Kartuschen (Millipore Corp., Bedford, MA) gereinigt und lyophilisiert. Dieses Material mit einer 5'-terminalen Aminogruppe wird in 100 µl absolutem N,N-Dimethylformamid (DMF) gelöst und mit 10 µmol N-Fmoc-Glycinpentafluorphenylester für 60 min bei 25°C kondensiert. Nach Ethanolfällung und Zentrifugation wird die Fmoc-Gruppe durch 10minütige Behandlung mit 100 µl einer Lösung aus 20% Piperidin in N,N-Dimethylformamid abgespalten. Überschüssiges Piperidin, DMF und die Spaltprodukte der Fmoc-Gruppe werden durch Ethanolfällung entfernt, und das Präzipitat wird aus 10 mM TEAA-Puffer pH 7,2 lyophilisiert. Dieses Material wird nun entweder als Primer für DNS-Sequenzierungsreaktionen nach Sanger, oder ein oder mehrere Glycinreste (oder andere geeignete geschützte aktive Aminosäureester) werden hinzugefügt, um eine Serie von massemodifizierten Primeroligonukleotiden zu schaffen, die zur DNS- oder RNS-Sequenzierung nach Sanger geeignet sind. Immobilisierung dieser massemodifizierten Nukleinsäureprimer $UP^{0-i}$ kann nach Primerverlängerung während des Sequenzierungsverfahrens wie z. B. in den BEISPIELEN 1 bis 4 beschrieben erzielt werden.

BEISPIEL 6

Synthese von Nukleinsäureprimern, die am C-5 der heterocyclischen Base eines Pyrimidinnukleosids mit Glycinresten massemodifiziert sind

**[0093]** Ausgangsmaterial war 5-(3-Aminopropynyl-1)-3',5'-di-p-tolyldesoxyuridin, das nach Literaturverfahren hergestellt und 3',5'-de-O-acetyliert wurde (Haralambidis et al., Nucleic Acids Res. 15, 4857–76 (1987)). 0,281 g (1,0 mmol) 5-(3-Aminopropynyl-1)-2'-desoxyuridin wurden mit 0,927 g (2,0 mmol) N-Fmoc-Glycinpentafluorphenylester in 5 ml absolutem N,N-Dimethylformamid für 60 min bei Raumtemperatur in Anwesenheit von 0,129 g (1 mmol, 174 µl) N,N-Diisopropylethylamin zur Reaktion gebracht. Lösungsmittel wurden durch Rotationsverdampfen entfernt, und das Produkt wurde durch Silicagel-Chromatographie gereinigt (Kieselgel 60, Merck, Säule: 2,5 × 50 cm, Elution mit Chloroform/Methanolgemischen). Die Ausbeute war 0,44 g (0,78 mmol, 78%). Um einen weiteren Glycinrest hinzuzufügen, wurde die Fmoc-Gruppe durch eine 20minütige Behandlung mit einer 20%igen Lösung aus Piperidin in DMF entfernt, im Vakuum abgezogen, und der verbleibende Feststoff wurde dreimal mit 20 ml Ethylacetat extrahiert. Nach Entfernung des verbleibenden Ethylacetats wird N-Fmoc-Glycinpentafluorphenylester wie oben beschrieben gekoppelt. 5-(3-(N-Fmoc-Glycyl)-amidopropynyl-1)-2'-desoxyuridin wird in das 5'-O-dimethoxytritylierte Nukleosid-3'-O-β-cyanoethyl-N,N-diisopropylphosphoamidit überführt und in einer automatisierten Oligonukleotidsynthese durch Standardverfahren (H. Köster et al., Nucleic Acids Res. 12, 2261 (1984)) eingebaut. Diese glycinmodifizierten thymidinanalogen Bausteine zur chemischen DNS-Synthese können verwendet werden, um eines oder mehrere der Thymidin/Uridinnukleotide in der Nukleinsäureprimersequenz zu ersetzen. Die Fmoc-Gruppe wird am Ende der Festphasensynthese mit einer 20minütigen Behandlung mit einer 20%igen Lösung von Piperidin in DMF bei Raumtemperatur entfernt. DMF wird durch einen Waschschritt mit Acetonitril entfernt, und das Oligonukleotid wird entschützt und auf Standardweise gereinigt.

BEISPIEL 7

Synthese eines Nukleinsäureprimers, der am C-5 der heterocyclischen Base eines Pyrimidinnukleosids mit β-Alaninresten massemodifiziert ist

**[0094]** Das Ausgangsmaterial entsprach dem aus Beispiel 6. 0,281 g (1,0 mmol) 5-(3-Aminopropynyl-1)-2'-desoxyuridin wurden mit N-Fmoc-β-Alaninpentafluorphenylester (0,955 g, 2,0 mmol) in 5 ml N,N-Dimethylformamid (DMF) für 60 min bei Raumtemperatur in Anwesenheit von 0,129 g (174 µl, 1,0 mmol) N,N-Diisopropylethylamin zur Reaktion gebracht. Lösungsmittel wurden entfernt, und das Produkt wurde durch Silicagel-Chromatographie wie in Beispiel 6 beschrieben gereinigt. Die Ausbeute war 0,425 g (0,74 mmol, 74%). Ein weiterer β-Alaninrest kann auf exakt dieselbe Weise nach Entfernen der Fmoc-Gruppe hinzugefügt werden. Die Herstellung des 5'-O-dimethoxytritylierten Nukleosid-3'-O-β-cyanoethyl-N,N-diisopropylphosphoami-

dits aus 5-(3-(N-Fmoc-β-Alanyl)-amidopropynyl-1)-2'-desoxyuridin und Einbau in automatisierter Oligonukleotidsynthese wird unter Standardbedingungen durchgeführt. Dieser Baustein kann beliebige der Thymidin/Uridinreste in den Nukleinsäureprimersequenzen ersetzen. Im Falle von nur einem eingebauten massemodifizierten Nukleotid hätten die Nukleinsäureprimermoleküle, die gemäß den BEISPIELEN 6 und 7 hergestellt würden, einen Masseunterschied von 14 Dalton.

## BEISPIEL 8

### Synthese eines Nukleinsäureprimers, der am C-5 der heterocyclischen Base eines Pyrimidinnukleosids mit Ethylenglycolmonomethylether massemodifiziert ist

**[0095]** Als nukleosidischer Bestandteil wurde in diesem Beispiel 5-(3-Aminopropynyl-1)-2'-desoxyuridin verwendet (siehe BEISPIELE 6 und 7). Die massemodifizierende Funktionalität wurde wie folgt erhalten: 7,61 g (100,0 mmol) frisch destillierter Ethylenglycolmonomethylether gelöst in 50 ml absolutem Pyridin wurden mit 10,01 g (100,0 mmol) umkristallisiertem Bernsteinsäureanhydrid in Anwesenheit von 1,22 g (10,0 mmol) 4-N,N-Dimethylaminopyridin über Nacht bei Raumtemperatur zur Reaktion gebracht. Die Reaktion wurde durch Hinzufügen von Wasser (5,0 ml) beendet, die Reaktionsmischung wurde im Vakuum abgezogen, zweimal zusammen mit trockenem Toluol (jeweils 20 ml) abgezogen, und der Rückstand wurde in 100 ml Dichlormethan wieder aufgelöst. Die Lösung wurde nacheinander zweimal mit 10%iger wässeriger Zitronensäure (2 × 20 ml) und einmal mit Wasser (20 ml) extrahiert, und die organische Phase wurde über wasserfreiem Natriumsulfat getrocknet. Die organische Phase wurde im Vakuum abgezogen, der Rückstand wurde in 50 ml Dichlormethan gelöst und in 500 ml Pentan ausgefällt, und das Präzipitat wurde im Vakuum getrocknet. Die Ausbeute war 13,12 g (74,0 mmol, 74 %). 8,86 g (50,0 mmol), des succinylierten Ethylenglycolmonomethylethers wurden in 100 ml Dioxan gelöst, das 5% trockenes Pyridin (5 ml) enthielt, und 6,96 g (50,0 mmol) 4-Nitrophenol und 10,32 g (50,0 mmol) Dicyclohexylcarbodiimid wurden hinzugefügt, und die Reaktion lief bei Raumtemperatur für 4 Stunden. Dicyclohexylharnstoff wurde durch Filtration entfernt, das Filtrat wurde im Vakuum abgezogen, und der Rückstand wurde wieder in 50 ml wasserfreiem DMF aufgelöst. 12,5 ml (etwa 12,5 mmol 4-Nitrophenylester) dieser Lösung wurden verwendet, um 2,81 g (10,0 mmol) 5-(3-Aminopropynyl-1)-2'-desoxyuridin zu lösen. Die Reaktion wurde in Anwesenheit von 1,01 g (10,0 mmol, 1,4 ml) Triethylamin bei Raumtemperatur über Nacht durchgeführt. Die Reaktionsmischung wurde im Vakuum abgezogen, zusammen mit Toluol abgezogen, in Dichlormethan wieder aufgelöst und mit Silicagel (Si60, Merck, Säule 4 × 50 cm) mit Dichlormethan/Methanolmischungen chromatographiert. Die Fraktionen, die die gewünschte Verbindung enthielten, wurden gesammelt, abgezogen, in 25 ml Dichlormethan wieder gelöst und in 250 ml Pentan gefällt. Das getrocknete Präzipitat von 5-(3-N-(O-Succinylethylenglycolmonomethylether)-amidopropynyl-1)-2'-desoxyuridin (Ausbeute 65%) wird O-dimethoxytrityliert und in das Nukleosid-3'-O-β-cyanoethyl-N,N-diisopropylphosphoamidit überführt und als Baustein in automatisierten Oligonukleotidsynthesen gemäß Standardverfahren eingebaut. Das massemodifizierte Nukleotid kann einen oder mehrere der Thymidin/Uridinreste in der Nukleinsäureprimersequenz ersetzen. Entschützen und Reinigen der Primeroligonukleotide folgt ebenso Standardverfahren.

## BEISPIEL 9

### Synthese eines Nukleinsäureprimers, der am C-5 der heterocyclischen Base eines Pyrimidinnukleosids mit Ethylenglycolmonomethylether massemodifiziert ist

**[0096]** Das nukleosidische Ausgangsmaterial 5-(3-Aminopropynyl-1)-2'-desoxyuridin entsprach dem der vorhergehenden Beispiele. Die massemodifizierende Funktionalität wurde auf ähnliche Weise wie in BEISPIEL 8 erhalten. 12,02 g (100,0 mmol) frisch destillierter Diethylenglycolmonomethylether gelöst in 50 ml absolutem Pyridin wurde mit 10,01 g (100,0 mmol) umkristallisiertem Bernsteinsäureanhydrid in Anwesenheit von 1,22 g (10,0 mmol) 4-N,N-Dimethylaminopyridin (DMAP) über Nacht bei Raumtemperatur zur Reaktion gebracht. Die Aufarbeitung war wie in BEISPIEL 8 beschrieben. Die Ausbeute war 18,35 g (82,3 mmol, 82,3%). 11,06 g (50,0 mmol) des succinylierten Diethylenglycolmonomethylethers wurden in den 4-Nitrophenylester überführt, und darauf wurden 12,5 mmol mit 2,81 g (10,0 mmol) 5-(3-Aminopropynyl-1)-2'-desoxyuridin wie in BEISPIEL 8 beschrieben zur Reaktion gebracht. Die Ausbeute nach Silicagel-Säulenchromatographie und Fällung in Pentan war 3,34 g (6,9 mmol, 69%). Nach Dimethoxytritylierung und Überführung in das Nukleosid-β-cyanoethylphosphoamidit werden die massemodifizierten Bausteine in automatisierten chemischen DNS-Synthesen gemäß Standardverfahren eingebaut. Innerhalb der Sequenz des Nukleinsäureprimers $UP^{0-i}$ können ein oder mehrere der Thymidin/Uridinreste durch dieses massemodifizierte Nukleotid ersetzt werden. Im Falle von nur einem eingebauten massemodifizierten Nukleotid hätten die Nukleinsäureprimermoleküle aus den BEISPIELEN 8 und 9 einen Masseunterschied von 44,05 Dalton.

BEISPIEL 10

Synthese eines Nukleinsäureprimers, der am C-8 der heterocyclischen Base eines Desoxyadenosins mit Glycin massemodifiziert ist

**[0097]** Das Ausgangsmaterial war $N^6$-Benzoyl-8-bromo-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyadenosin, das gemäß der Literatur hergestellt wurde (Singh et al., Nucleic Acids Res. 18, 3339–45, (1990)). 632,5 mg (1,0 mmol) dieses 8-Bromodesoxyadenosinderivats wurden in 5 ml absolutem Ethanol suspendiert und mit 251,2 mg (2,0 mmol) Glycinmethylester (Hydrochlorid) in Anwesenheit von 241,4 mg (2,1 mmol, 366 µl) N,N-Diisopropylethylamin zur Reaktion gebracht und refluxiert, bis das nukleosidische Ausgangsmaterial verschwunden war (4–6 Stunden), was durch Dünnschichtchromatographie (thin layer chromatography, TLC) überprüft wurde. Das Lösungsmittel wurde abgezogen, und der Rückstand wurde durch Silicagel-Chromatographie (Säule 2,5 × 50 cm) gereinigt, wobei Lösungsmittelmischungen aus Chloroform/Methanol verwendet wurden, die 0,1 Pyridin enthielten. Die Produktfraktionen wurden vereinigt, das Lösungsmittel wurde abgezogen, die Fraktionen wurden in 5 ml Dichlormethan gelöst und in 100 ml Pentan gefällt. Die Ausbeute war 487 mg (0,76 mmol, 76%). Die Überführung in das entsprechende Nukleosid-β-cyanoethylphosphoamidit und Einbau in einer automatisierten chemischen DNS-Synthese wird unter Standardbedingungen durchgeführt. Während der endgültigen Entschützung mit konzentriertem wässerigem Ammoniak werden die Methylgruppen von den Glycingruppen entfernt. Der massemodifizierte Baustein kann einen oder mehrere Desoxyadenosin/Adenosinreste in der Nukleinsäureprimersequenz ersetzen.

BEISPIEL 11

Synthese eines Nukleinsäureprimers, der am C-8 der heterocyclischen Base eines Desoxyadenosins mit Glycylglycin massemodifiziert ist

**[0098]** Dieses Derivat wurde analog zu dem Glycinderivat aus BEISPIEL 10 hergestellt. 632,5 mg (1,0 mmol) $N^6$-Benzoyl-8-bromo-5'-O-(4,4'-dimethoxytrityl)-2'-desoxyadenosin wurden in 5 ml absolutem Ethanol suspendiert und mit 324,3 mg (2,0 mmol) Glycylglycinmethylester in Anwesenheit von 241,4 mg (2,1 mmol, 366 µl) N,N-Diisopropylethylamin zur Reaktion gebracht. Die Mischung wurde refluxiert, und die Vollständigkeit der Reaktion wurde durch TLC überprüft. Aufarbeitung und Reinigung waren zu der ähnlich, die in BEISPIEL 10 beschrieben wurde. Die Ausbeute nach Silicagel-Säulenchromatographie und Fällung in Pentan war 464 mg (0,65 mmol, 65%). Die Überführung in das Nukleosid-β-cyanoethylphosphoamidit und in synthetische Oligonukleotide wurde gemäß Standardverfahren durchgeführt. Falls nur einer der Desoxyadenosin-/Adenosinreste im Nukleinsäureprimer durch dieses massemodifizierte Nukleotid ersetzt ist, beträgt der Masseunterschied zwischen den Nukleinsäureprimern der BEISPIELE 10 und 11 57,03 Dalton.

BEISPIEL 12

Synthese eines Nukleinsäureprimers, der am C-2' des Zuckerrestes des 2'-Amino-2'-desoxythymidins mit Ethylenglycolmonomethyletherresten massemodifiziert ist

**[0099]** Ausgangsmaterial war 5'-O-(4,4-Dimethoxytrityl)-2'-amino-2'-desoxythymidin, das gemäß veröffentlichten Verfahren synthetisiert wurde (z. B. Verheyden et al., J. Org. Chem. 36, 250–254 (1971), Sasaki et al., J. Org. Chem. 41, 3138–3143 (1976) Imazawa et al., J. Org. Chem. 44, 2039–2041 (1979), Hobbs et al., J. Org. Chem. 42, 714–719 (1976), Ikehara et al., Chem. Pharm. Bull. Japan 26, 240–244 (1978), siehe ebenso PCT-Anmeldung WO 88/00201). 5'-O-(4,4-Dimethoxytrityl)-2'-amino-2'-desoxythymidin (559,62 mg, 1,0 mmol) wurde mit 2,0 mmol des 4-Nitrophenylesters eines succinylierten Ethylenglycolmonomethylethers (siehe BEISPIEL 8) in 10 ml trockenem DMF in Anwesenheit von 1,0 mmol (140 µl) Triethylamin für 18 Stunden bei Raumtemperatur zur Reaktion gebracht. Die Reaktionsmischung wurde im Vakuum abgezogen, zusammen mit Toluol abgezogen, wieder in Dichlormethan gelöst und mit Silicagel-Chromatographie (Si60, Merck, Säule 2,5 × 50 cm, Elutionsmittel: Chloroform/Methanolmischungen, die 0,1% Triethylamin enthalten) gereinigt. Die Fraktionen, die das Produkt enthielten, wurden vereinigt, abgezogen und in Pentan gefällt. Die Ausbeute war 524 mg (0,73 mmol, 73%). Die Überführung in das Nukleosid-β-cyanoethyl-N,N-diisopropylphosphoamidit und Einbau in einem automatisierten chemischen DNS-Syntheseprotokoll wird mit Standardverfahren durchgeführt. Das massemodifizierte Desoxythymidinderivat kann einen oder mehrere Thymidinreste im Nukleinsäureprimer ersetzen.

**[0100]** Auf analoge Weise werden die entsprechenden massemodifizierten Oligonukleotide unter Verwendung von 4-Nitrophenylester von succinyliertem Diethylenglycolmonomethylether (siehe BEISPIEL 9) und Tri-

ethylenglycolmonomethylether hergestellt. Im Falle von nur einem in die Sequenz eingebauten massemodifizierten Nukleosid beträgt der Masseunterschied zwischen den Ethylen-, Diethylen- und Triethylenglycolderivaten 44,05, 88,1 bzw. 132,15 Dalton.

BEISPIEL 13

Synthese eines Nukleinsäureprimers, der an der Internukleotidbindung über Alkylierung von Phosphorthioatgruppen massemodifiziert ist

**[0101]** Phosphorthioat enthaltende Oligonukleotide wurden gemäß Standardverfahren hergestellt (siehe z. B. Gait et al., Nucleic Acids Res. 19, 1183 (1991)). Eine, mehrere oder alle Internukleotidbindungen können auf diese Weise modifiziert werden. Die (–)-M13-Nukleinsäureprimersequenz (17mer) 5'-dGTAAAACGACGGCCAGT (SEQ ID NO: 6) wurde im 0,25-µmol-Maßstab auf einem DNS-Synthesizer synthetisiert, wobei eine Phosphorthioatgruppe nach dem letzten Synthesezyklus eingeführt wurde (G-zu-T-Kopplung). Sulfurisierung, Entschützung und Reinigung folgten Standardprotokollen. Die Ausbeute war 31,4 mmol (12,6 Gesamtausbeute), was 31,4 nmol Phosphorthioatgruppen entsprach. Alkylierung wurde durchgeführt, indem der Rückstand in 31,4 µl TE-Puffer (0,01 M Tris pH 8,0, 0,001 M EDTA) gelöst wurde und indem 16 µl einer Lösung von 20 mM Lösung von 2-Iodethanol (320 nmol, d. h. 10facher Überschuss in Bezug auf Phosphorthioatdiester) in N,N-Dimethylformamid (DMF) hinzugefügt wurden. Das alkylierte Oligonukleotid wurde durch Standardumkehrphasen-HPLC (RP-18 Ultraphere, Beckman, Säule: 4,5 × 250 mm, 100 mM Triethylammoniumacetat, pH 7,0 und ein Gradient von 5 bis 40% Acetonitril) gereinigt.

**[0102]** In einer Variation dieses Verfahrens wird der Nukleinsäureprimer, der ein oder mehrere Phosphorthioatphosphodiesterbindungen enthält, in den Sequenzierungsreaktionen nach Sanger verwendet. Die Primerverlängerungsprodukte der vier Sequenzierungsreaktionen werden wie exemplarisch in den BEISPIELEN 1–4 dargestellt gereinigt, vom festen Träger abgespalten, lyophilisiert und jeweils in 4 µl TE-Puffer pH 8,0 gelöst und alkyliert, indem 2 µl einer 20 mM Lösung von 2-Iodethanol in DMF hinzugefügt werden. Dann wird es durch ES- und/oder MALDI-Massenspektrometrie analysiert.

**[0103]** Auf analoge Weise können durch Verwendung anstatt von 2-Iodethanol von z. B. 3-Iodpropanol, 4-Iodbutanol massemodifizierte Nukleinsäureprimer mit einer Massedifferenz von 14,03, 28,06 bzw. 42,03 Dalton im Vergleich zum nichtmodifizierten Phosphorthioat-Phosphodiester enthaltenden Oligonukleotid erhalten werden.

BEISPIEL 14

Synthese von 2'-Amino-2'-desoxyuridin-5'-triphosphat und 3'-Amino-2',3'-didesoxythymidin-5'-triphosphat, die an der 2'- oder 3'-Aminofunktion mit Glycin- oder β-Alaninresten massemodifiziert sind

**[0104]** Ausgangsmaterial war 2'-Azido-2'-desoxyuridin, das gemäß der Literatur hergestellt wurde (Verheyden et al., J. Org. Chem. 36, 250 (1971)), das in einer Eintopfreaktion unter Standardreaktionsbedingungen am 5'-OH mit 4,4-Dimethoxytritylchlorid in Pyridin 4,4-dimethoxytrityliert und am 3'-OH mit Essigsäureanhydrid acetyliert wurde. Mit 191 mg (0,71 mmol) 2'-Azido-2'-desoxyuridin als Ausgangsmaterial wurden 396 mg (0,65 mmol, 90,8%) 5'-O-(4,4-Dimethoxytrityl)-3'-O-acetyl-2'-azido-2'-desoxyuridin nach Reinigung mit Silicagel-Chromatographie erhalten. Eine Reduktion der Azidogruppe wurde unter Verwendung von veröffentlichten Bedingungen durchgeführt (Barta et al., Tetrahedron 46, 587–594, (1990)). Die Ausbeute von 5'-O-(4,4-Dimethoxytrityl)-3'-O-acetyl-2'-amino-2'-desoxyuridin nach Silicagel-Chromatographie war 288 mg (0,49 mmol, 76%). Dieses geschützte 2'-Amino-2'-desoxyuridinderivat (588 mg, 1,0 mmol) wurde mit 2 Äquivalenten (927 mg, 2,0 mmol) N-Fmoc-Glycinpentafluorphenylester in 10 ml trockenem DMF über Nacht bei Raumtemperatur in Anwesenheit von 1,0 mmol (174 µl) N,N-Diisopropylethylamin zur Reaktion gebracht. Lösungsmittel wurden durch Abziehen im Vakuum entfernt, und der Rückstand wurde durch Silicagel-Chromatographie gereinigt. Die Ausbeute war 711 mg (0,71 mmol, 82%). Detritylierung wurde durch eine einstündige Behandlung mit 80%iger wässeriger Essigsäure bei Raumtemperatur erzielt. Der Rückstand wurde bis zur Trockene abgezogen, zweimal zusammen mit Toluol abgezogen, in 1 ml trockenem Acetonitril suspendiert und mit $POCl_3$ gemäß der Literatur 5'-phosphoryliert (Yoshikawa et al., Bull. Chem. Soc. Japan 42, 3505 (1969) und Sowa et al., Bull. Chem. Soc. Japan 48, 2048 (1975)) und direkt in einer Eintopfreaktion in das 5'-Triphosphat unter Verwendung von 3 ml einer 0,5 M-Lösung (1,5 mmol) von tetra(tri-n-Butylammonium)pyrophosphat in DMF gemäß der Literatur (z. B. Seela et al., Helvetica Chimica Acta 74, 1048 (1991)) transformiert. Die Fmoc- und die 3'-O-Acetylgruppen wurden durch eine einstündige Behandlung mit konzentriertem wässerigem Ammoniak bei Raumtemperatur entfernt, und die Reaktionsmischung wurde abgezogen und lyophilisiert. Ebenso folgte die Reini-

gung Standardverfahren unter Verwendung von Anionenaustauschchromatographie auf DEAE-Sephadex mit einem linearen Gradienten von Triethylammoniumbicarbonat (0,1 M–1,0 M). Triphosphat enthaltende Fraktionen (überprüft durch Dünnschichtchromatographie auf Polyethylenimincelluloseplatten) wurden gesammelt, abgezogen und lyophilisiert. Die Ausbeute (durch UV-Absorption der Uracilgruppe) war 68% (0,48 mmol).

**[0105]** Ein glycylglycinmodifiziertes 2'-Amino-2'-desoxyuridin-5'-triphosphat wurde durch Entfernen der Fmoc-Gruppe aus 5'-O-(4,4-Dimethoxytrityl)-3'-O-acetyl-2'-N-(N-9-fluorenylmethyloxycarbonyl-glycyl)-2'-amino-2'-desoxyuridin durch eine einstündige Behandlung mit einer 20%igen Lösung von Piperidin in DMF bei Raumtemperatur, Abziehen von Lösungsmitteln, zweifaches Abziehen zusammen mit Toluol und nachfolgende Kondensation mit N-Fmoc-Glycinpentafluorphenylester erhalten. Ausgehend von 1,0 mmol des 2'-N-Glycyl-2'-amino-2'-desoxyuridinderivates und folgend dem Verfahren wie oben beschrieben, wurden 0,72 mmol (72%) des entsprechenden 2'-(N-Glycylglycyl)-2'-amino-2'-desoxyuridin-5'-triphosphats erhalten.

**[0106]** Ausgehend von 5'-O-(4,4-Dimethoxytrityl)-3'-O-acetyl-2'-amino-2'-desoxyuridin und Koppeln mit N-Fmoc-β-Alaninpentafluorphenylester kann das entsprechende 2'-(N-β-Alanyl)-2'-amino-2'-desoxyuridin-5'-triphosphat synthetisiert werden. Diese modifizierten Nukleosidtriphosphate werden während des DNS-Sequenzierungsverfahrens nach Sanger in die Primerverlängerungsprodukte eingebaut. Die Massedifferenz zwischen den glycin-, β-alanin- und glycylglycinmassemodifizierten Nukleosiden beträgt pro eingebautem Nukleotid 58,06, 72,09 bzw. 115,1 Dalton.

**[0107]** Wenn man von 5'-O-(4,4-Dimethoxytrityl)-3'-amino-2',3'-didesoxythymidin (erhalten durch veröffentlichte Verfahren, siehe BEISPIEL 12) ausgeht, können die entsprechenden 3'-(N-Glycyl)-3'-amino-/, 3'-(N-Glycylglycyl)-3'-amino-/ und 3'-(N-β-Alanyl)-3'-amino-2',3'-didesoxythymidin-5'-triphosphate erhalten werden. Die massemodifizierten Nukleosidtriphosphate dienen als eine terminierende Nukleotideinheit in den DNS-Sequenzierungsreaktionen nach Sanger und stellen eine Massedifferenz pro terminiertem Fragment von 58,06, 72,09 bzw. 115,1 Dalton bereit, wenn sie im Multiplexsequenzierungsmodus verwendet werden. Die massedifferenzierten Fragmente können dann durch ES- und/oder MALDI-Massenspektrometrie analysiert werden.

## BEISPIEL 15

### Synthese von Desoxyuridin-5'-triphosphat, das am C-5 der heterocyclischen Base mit Glycin-, Glycylglycin oder β-Alaninresten massemodifiziert ist

**[0108]** 0,281 g (1,0 mmol) 5-(3-Aminopropynyl-1)-2'-desoxyuridin (siehe BEISPIEL 6) wurden entweder mit 0,927 g (2,0 mmol) N-Fmoc-Glycinpentafluorphenylester oder 0,955 g (2,0 mmol) N-Fmoc-β-Alaninpentafluorphenylester in 5 ml trockenem DMF über Nacht bei Raumtemperatur in Anwesenheit von 0,129 g N,N-Diisopropylethylamin (174 µl, 1,0 mmol) zur Reaktion gebracht. Lösungsmittel wurden durch Abziehen im Vakuum entfernt, und die Kondensationsprodukte wurden durch Flashchromatographie auf Silicagel gereinigt (Still et al., J. Org. Chem. 43, 2923–2925 (1978)). Die Ausbeute war 476 mg (0,85 mmol, 85%) für das Glycin- und 436 mg (0,76 mmol, 76%) für das β-Alaninderivat. Für die Synthese des Glycylglycinderivates wurde die Fmoc-Gruppe von 1,0 mmol Fmoc-Glycindesoxyuridinderivat durch eine einstündige Behandlung mit 20%igem Piperidin in DMF bei Raumtemperatur entfernt. Lösungsmittel wurden im Vakuum abgezogen, der Rückstand wurde zweimal zusammen mit Toluol abgezogen und mit 0,927 g (2,0 mmol) N-Fmoc-Glycinpentafluorphenylester kondensiert und wie oben beschrieben gereinigt. Die Ausbeute war 455 mg (0,72 mmol, 72%). Die Glycyl-, Glycylglycyl und β-Alanyl-2'-desoxyuridinderivate, die durch die Fmoc-Gruppe N-geschützt waren, wurden in die 3'-O-Acetylderivate durch Tritylierung mit 4,4-Dimethoxytritylchlorid in Pyridin und Acetylierung mit Essigsäureanhydrid in Pyridin in einer Eintopfreaktion überführt und danach durch eine einstündige Behandlung mit 80%iger wässeriger Essigsäure gemäß Standardverfahren detrityliert. Lösungsmittel wurden entfernt, die Rückstände wurden in 100 ml Chloroform gelöst und zweimal mit 50 ml 10%igem Natriumbicarbonat und einmal mit 50 ml Wasser extrahiert, mit Natriumsulfat getrocknet, das Lösungsmittel wurde abgezogen, und der Rückstand wurde mit Flashchromatographie über Silicagel gereinigt. Die Ausbeuten waren 361 mg (0,60 mmol, 71%) für das Glycyl-, 351 mg (0,57 mmol, 75%) für das β-Alanyl- bzw. 323 mg (0,49 mmol, 68%) für das Glycylglycyl-3-O'-acetyl-2'-desoxyuridinderivat. Phosphorylierung am 5'-OH mit $POCl_3$, Überführung in das 5'-Triphosphat durch eine in situ-Reaktion mit tetra(tri-n-Butylammonium)pyrophosphat in DMF, 3'-De-O-acetylierung, Spaltung der Fmoc-Gruppe, die endgültige Reinigung durch Anionenaustauschchromatographie auf DEAE-Sephadex wurde wie in BEISPIEL 14 beschrieben durchgeführt. Die Ausbeuten entsprechend der UV-Absorption der Uracilgruppe waren 0,41 mmol 5-(3-(N-Glycyl)-amidopropynyl-1)-2'-desoxyuridin-5'-triphosphat (84%), 0,43 mmol 5-(3-(N-β-Alanyl)-amidopropynyl-1)-2'-desoxyuridin-5'-triphosphat (75%) und 0,38 mmol 5-(3-(N-Glycylglycyl)-amidopropynyl-1)-2'-desoxyuridin-5'-triphosphat (78%).

**[0109]** Diese massemodifizierten Nukleosidtriphosphate wurden während der DNS-Primerverlängerungsreaktionen nach Sanger eingebaut.

**[0110]** Bei Verwendung von 5-(3-Aminopropynyl-1)-2',3'-desoxyuridin als Ausgangsmaterial und Ausführen einer analogen Reaktionssequenz wurden die entsprechenden Glycyl-, Glycylglycyl- und β-Alanyl-2',3'-didesoxyuridin-5'-triphosphate in Ausbeuten von 69, 63 bzw. 71% erhalten. Diese massemodifizierten Nukleosidtriphosphate dienen als kettenterminierende Nukleotide während der DNS-Sequenzierungsreaktionen nach Sanger. Die massemodifizierten Sequenzierungsleitern werden entweder durch ES- oder MALDI-Massenspektrometrie analysiert.

## BEISPIEL 16

### Synthese von 8-Glycyl- und 8-Glycylglycyl-2'-desoxyadenosin-5'-triphosphat

**[0111]** 727 mg (1,0 mmol) von $N^6$-(4-tert-Butylphenoxyacetyl)-8-glycyl-5'-(4,4-dimethoxytrityl)-2'-desoxyadenosin oder 800 mg (1,0 mmol) $N^6$-(4-tert-Butylphenoxyacetyl)-8-glycylglycyl-5'-(4,4-dimethoxytrityl)-2'-desoxyadenosin, die gemäß den BEISPIELEN 10 und 11 und der Literatur hergestellt wurden (Köster et al., Tetrahedron 37, 362 (1981), wurden mit Essigsäureanhydrid in Pyridin am 3'-OH acetyliert, an der 5'-Position mit 80% Essigsäure in einer Eintopfreaktion detrityliert und über Phosphorylierung mit $POCl_3$ und in situ-Reaktion mit tetra(tri-n-butylammonium)pyrophosphat wie in BEISPIEL 14 beschrieben in die Triphosphate überführt. Entschützung der $N^6$-tert-Butylphenoxyacetyl-, der 3'-O-Acetyl- und der O-Methylgruppe an den Glycinresten wurde mit konzentriertem wässerigem Ammoniak für neunzig Minuten bei Raumtemperatur erzielt. Das Ammoniak wurde durch Lyophilisierung entfernt, und der Rückstand wurde mit Dichlormethan gewaschen, das Lösungsmittel wurde durch Abziehen im Vakuum entfernt, und das verbleibende feste Material wurde durch Anionenaustauschchromatographie auf DEAE-Sephadex unter Verwendung eines linearen Gradienten von Triethylammoniumbicarbonat von 0,1 bis 1,0 M gereinigt. Die Nukleosidtriphosphat enthaltenden Fraktionen (überprüft durch TLC auf Polyethylenimincelluloseplatten) wurden vereinigt und lyophilisiert. Die Ausbeute des 8-Glycyl-2'-desoxyadenosin-5'-triphosphats (bestimmt durch UV-Absorption der Adeningruppe) war 57% (0,57 mmol). Die Ausbeute des 8-Glycylglycyl-2'-desoxyadenosin-5'-triphosphats war 51% (0,51 mmol).

**[0112]** Diese massemodifizierten Nukleosidtriphosphate wurden während der Primerverlängerung in den DNS-Sequenzierungsreaktionen nach Sanger eingebaut.

**[0113]** Bei Verwendung der entsprechenden $N^6$-(4-tert-Butylphenoxyacetyl)-8-glycyl- oder -glycylglycyl-5'-O-(4,4-dimethoxytrityl)-2',3'-didesoxyadenosinderivate als Ausgangsmaterialien, die gemäß Standardverfahren hergestellt wurden (siehe z. B. für die Einführung der 2',3'-Funktion: Seela et al., Helvetica Chimica Acta 74, 1048–1058 (1991)) und Verwendung einer analogen Reaktionssequenz wie oben beschrieben wurden die kettenterminierenden massemodifizierten Nukleosidtriphosphate 8-Glycyl- und 8-Glycylglycyl-2',3'-didesoxyadenosin-5'-triphosphate in Ausbeuten von 53 bzw. 47% erhalten. Die massemodifizierten Sequenzierungsfragmentleitern werden entweder durch ES- oder MALDI-Massenspektrometrie analysiert.

## BEISPIEL 17

### Massemodifikation von DNS-Sequenzierungsfragmentleitern nach Sanger durch Einbau von kettenverlängernden 2'-Desoxy- und kettenterminierenden 2',3'-Didesoxythymidin-5'-(alpha-S-)-triphosphaten und nachfolgende Alkylierung mit 2-Iodethanol und 3-Iodpropanol

**[0114]** 2',3'-Didesoxythymidin-5'-(alpha-S)-triphosphat wurde gemäß veröffentlichten Verfahren hergestellt (z. B. für die alpha-S-Triphosphatgruppe: Eckstein et al., Biochemistry 15, 1685 (1976) und Accounts Chem. Res. 12, 204, (1978) und für die 2',3'-Didesoxygruppe: Seela et al., Helvetica Chimica Acta, 74, 1048–1058 (1991)). DNS-Sequenzierungsreaktionen nach Sanger, die 2'-Desoxythymidin-5'-(alpha-S)-triphosphat verwenden, werden gemäß Standardprotokollen durchgeführt (z. B. Eckstein, Ann. Rev. Biochem. 54, 367 (1985)). Bei Verwendung von 2',3'-Didesoxythymidin-5'-(alpha-S)-triphosphaten wird dieses anstatt des nichtmodifizierten 2',3'-Didesoxythymidin-5'-triphosphates bei Standard-DNS-Sequenzierungen nach Sanger verwendet (siehe z. B. Swerdlow et al., Nucleic Acids Res. 18, 1415–1419 (1990)). Das Template (2 pmol) und der M13-Nukleinsäuresequenzierungsprimer (4 pmol), die gemäß BEISPIEL 1 modifiziert sind, werden durch Erwärmen für 5 min auf 65°C in 100 µl 10 mM Tris-HCl pH 7,5, 10 mM $MgCl_2$, 50 mM NaCl, 7 mM Dithiothreitol (DTT) annealt und langsam für eine Stunde auf 37°C gebracht. Die Sequenzierungsreaktionsmischung enthält, wie für die T-spezifische Terminationsreaktion exemplarisch dargestellt, in einem Endvolumen von 150 µl 200 µM (Endkonzentration) von jeweils dATP, dCTP, dTTP, 300 µM c7-Desaza-dGTP, 5 µM 2',3'-Didesoxythymi-

din-5'-(alpha-S)-triphosphat und 40 Einheiten Sequenase (United States Biochemicals). Die Polymerisation wird für 10 min bei 37°C durchgeführt, die Reaktionsmischung wird auf 70°C erwärmt, um die Sequenase zu inaktivieren, und ethanolgefällt und an thiolierte Sequelon-Membranscheiben (8 mm Durchmesser) gekoppelt, wie in BEISPIEL 1 beschrieben. Eine Alkylierung wird durchgeführt, indem die Scheiben mit 10 µl 10 mM Lösung von entweder 2-Iodethanol oder 3-Iodpropanol in NMM (N-Methylmorpholin/Wasser/2-Propanol, 2/49/49, V/V/V)(dreimal) behandelt werden, mit 10 µl NMM gewaschen werden (dreimal), und die alkylierten T-terminierten Primerverlängerungsprodukte vom Träger durch Behandeln mit DTT wie in BEISPIEL 1 beschrieben abgespalten werden. Eine Analyse von massemodifizierten Fragmentfamilien wird entweder durch ES- oder MALDI-Massenspektrometrie durchgeführt.

## BEISPIEL 18

### Analyse einer Mischung von Oligothymidylsäuren

**[0115]** Oligothymidylsäure, oligo-p$(dT)_{12-18}$ ist kommerziell erhältlich (United States Biochemicals, Cleveland, OH). Im Allgemeinen wurde eine 0,5 M Matrixlösung in Ethanol hergestellt. Verschiedene Matrizen wurden für dieses Beispiel und die Beispiele 19–21 verwendet, wie z. B. 3,5-Dihydrobenzoesäure, Sinapinsäure, 3-Hydroxypicolinsäure, 2,4,6-Trihydroxyacetophenon. Oligonukleotide wurden nach Reinigung durch HPLC lyophilisiert und in ultrareines Wasser aufgenommen (MilliQ, Millipore), wobei solche Mengen verwendet wurden, dass eine Konzentration von 10 pmol/µl als Stammlösung erhalten wurde. Ein Aliquot (1 µl) dieser Konzentration oder eine Verdünnung in ultrareinem Wasser wurde mit 1 µl der Matrixlösung auf einer flachen Metalloberfläche gemischt, die als Sondenspitze diente, und mit einem Ventilator mit kalter Luft getrocknet. In einigen Experimenten wurden Kationenaustauschkügelchen in Säureform zu der Mischung aus der Matrix- und Probelösung hinzugefügt.

**[0116]** MALDI-TOF-Spektren wurden in diesem Beispiel und den Beispielen 19–21 auf verschiedenen kommerziellen Instrumenten erhalten, wie z. B. Vision 2000 (Finnigan-MAT), VG TofSpec (Fisons Instruments), LaserTec Research (Vestec). Die Bedingungen in diesem Beispiel waren der lineare Negativionenmodus mit einer Beschleunigungsspannung von 25 kV. Das erzeugte MALDI-TOF-Spektrum wird in **Abb. 14** gezeigt. Eine Massekalibrierung wurde extern durchgeführt und allgemein durch Verwendung definierter Peptide im geeigneten Massebereich erzielt, wie z. B. durch Insulin, Gramicidin S, Trypsinogen, Rinderserumalbumin und Cytochrom C. Alle Spektren wurden erzeugt, indem ein Stickstofflaser mit 5-nsec-Pulsen bei einer Wellenlänge von 337 nm verwendet wurde. Die Laserenergie variierte zwischen $10^6$ und $10^7$ W/cm$^2$. Um das Signal-Rauschverhältnis allgemein zu verbessern, wurden die Intensitäten von 10 bis 30 Laserschüssen akkumuliert.

## BEISPIEL 19

### Massenspektrometrische Analyse eines 50mers und eines 99mers

**[0117]** Zwei große Oligonukleotide wurden durch Massenspektrometrie analysiert. Das 50mer d(TAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATGACTAGCT) (SEQ ID NO: 3) und dT$(pdT)_{99}$ wurden verwendet. Die Oligodesoxynukleotide wurden unter Verwendung von β-Cyanoethylphosphoamiditen synthetisiert und unter Verwendung von veröffentlichten Verfahren gereinigt (z. B. N. D. Sinha, J. Biernat, J. McManus und H. Köster, Nucleic Acids Res. 12, 4539 (1984)), wobei kommerziell erhältliche DNS-Synthesizer von entweder Millipore (Bedford, MA) oder Applied Biosystems (Foster City, CA) und eine HPLC-Ausrüstung und RP18-Umkehrphasensäulen von Waters (Milford, MA) verwendet wurden. Die Proben für die massenspektrometrische Analyse wurden wie in Beispiel 18 beschrieben hergestellt. Die Bedingungen, die für die MALDI-MS-Analyse für jedes Oligonukleotid verwendet wurden, waren 500 fmol von jedem Oligonukleotid, Reflektron-Positivionenmodus mit einer Beschleunigung von 5 kV und Nachbeschleunigung von 20 kV. Die erzeugten MALDI-TOF-Spektren wurden überlagert und sind in **Abb. 15** gezeigt.

## BEISPIEL 20

### Simulation der DNS-Sequenzierungsergebnisse aus **Abb. 2**

**[0118]** Die 13 DNS-Sequenzen, die die verschachtelten dT-terminierten Fragmente der DNS-Sequenzierung nach Sanger für das 50mer repräsentieren, das in Beispiel 19 beschrieben ist (SEQ ID NO: 3), wurden wie in Beispiel 19 beschrieben synthetisiert. Die Proben wurden wie in Beispiel 18 beschrieben behandelt, und 500 fmol von jedem Fragment wurden durch MALDI-MS analysiert, wie in Beispiel 18 beschrieben. Die resultierenden MALDI-TOF-Spektren sind in den **Abb. 16A**–**Abb. 16M** gezeigt. Die Bedingungen waren der Reflek-

tron-Positivionenmodus mit einer Beschleunigung von 5 kV und Nachbeschleunigung von 20 kV. Die berechneten molekularen Massen und die experimentellen molekularen Massen sind in Tabelle 1 gezeigt.

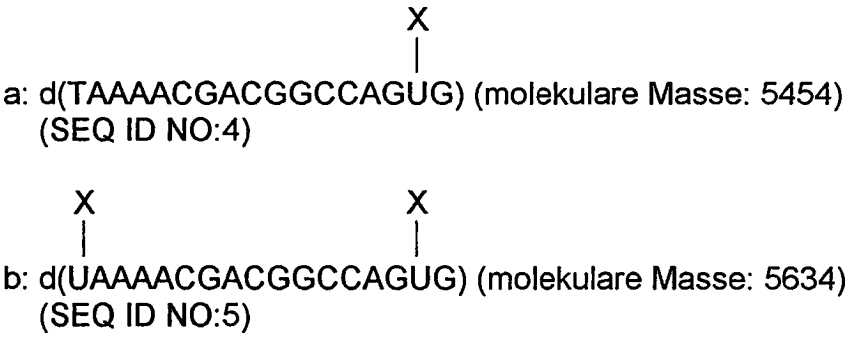
**[0119]** Die MALDI-TOF-Spektren wurden überlagert (**Abb. 17A** und **Abb. 17B**), um zu demonstrieren, dass die individuellen Peaks sogar zwischen dem 10mer und 11mer (obere Tafel) und dem 37mer und 38mer (untere Tafel) auflösbar sind. Die zwei Tafeln zeigen zwei unterschiedliche Maßstäbe und die in diesem Maßstab analysierten Spektrum.

BEISPIEL 21

MALDI-MS-Analyse eines massemodifizierten Oligonukleotids

**[0120]** Ein 17mer wurde am C-5 von einer oder zwei Desoxyuridingruppen massemodifiziert. 5-[13-(2-Methoxyethoxyl)-tridecyn-1-yl)-5'-O-(4,4-dimethoxytrityl)-2'-desoxyuridin-3'-β-cyanoethyl-N,N-diisopropylphosphoamidit wurde verwendet, um die modifizierten 17mere unter Verwendung der Verfahren wie in Beispiel 19 beschrieben zu synthetisieren.

**[0121]** Die modifizierten 17mere waren




wobei X = -C≡C-$(CH_2)_{11}$-OH
(nichtmodifiziertes 17mer: molekulare Masse: 5273)

**[0122]** Die Proben wurden wie in Beispiel 18 beschrieben behandelt, und 500 fmol von jedem modifizierten 17mer wurden durch MALDI-MS analysiert, wie in Beispiel 18 beschrieben. Die Bedingungen waren der Reflektron-Positivionenmodus mit einer Beschleunigung von 5 kV und Nachbeschleunigung von 20 kV. Die erzeugten MALDI-TOF-Spektren wurden überlagert und sind in **Abb. 18** gezeigt.

**[0123]** Alle der oben zitierten Referenzen und Veröffentlichungen werden hiermit durch die Referenz aufgenommen.

SEQUENZPROTOKOLL

(1) ALLGEMEINE ANGABEN:

(i) AMELDER:
(A) NAME: KOSTER, HUBERT
(B) STRASSE: 1640 MONUMENT STREET
(C) ORT: CONCORD
(D) LAND: MASSACHUSETTS
(E) STAAT: USA
(F) POSTLEITZAHL: 01742
(G) TELEFON: (508) 369-9790

(ii) BEZEICHNUNG DER ERFINDUNG: DNS-SEQUENZIERUNG DURCH MASSENSPEKTROMETRIE

(iii) ANZAHL DER SEQUENZEN: 6

(v) COMPUTER-LESBARE FORM:
(A) DATENTRAEGER: Diskette
(B) COMPUTER: IBM PC kompatibel
(C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: ASCII (Text)

(vi) DATEN DER VORLIEGENDEN ANMELDUNG:
(A) ANMELDENUMMER:
(B) ANMELDEDATUM: 06-JAN-1994
(C) KLASSIFIZIERUNG:

(vii) DATEN DER FRÜHEREN ANMELDUNG:
(A) ANMELDENUMMER: US 08/001,323
(B) ANMELDEDATUM: 07-JAN-1993
(C) KLASSIFIZIERUNG: 1807

(viii) ANWALTS-/AGENTENANGABEN:
(A) NAME: DeConti, Giulio A.
(B) REGISTRIERNUMMER: 31,503
(C) AKTENZEICHEN: HKI-003CP

(ix) TELEKOMMUNIKATIONSANGABEN:
(A) TELEFON: (617) 227-7400
(B) TELEFAX: (617) 227-5941

(2) ANGABEN ZU SEQ ID NO: 1:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 14 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: andere Nukleinsäure

(iii) HYPOTHETISCH: JA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

CATGCCATGG CATG 14

(2) ANGABEN ZU SEQ ID NO: 2:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 21 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: andere Nukleinsäure

(iii) HYPOTHETISCH: JA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

AAATTGTGCA CATCCTGCAG C 21

(2) ANGABEN ZU SEQ ID NO: 3:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 50 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: andere Nukleinsäure

(iii) HYPOTHETISCH: JA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

TAACGGTCAT TACGGCCATT GACTGTAGGA CCTGCATTAC ATGACTAGCT 50

(2) ANGABEN ZU SEQ ID NO: 4:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 17 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzelstrang

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: andere Nukleinsäure

(iii) HYPOTHETISCH: JA

(ix) MERKMAL
(A) NAME/SCHLÜSSEL: misc_feature
(B) LAGE: 16
(C) SONSTIGE ANGABEN: /Anmerkung= "N ist das massemodifizierte Desoxyuridin-Nukleotid"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

TAAAACGACG GCCAGNG 17

(2) ANGABEN ZU SEQ ID NO: 5:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 17 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: andere Nukleinsäure

(iii) HYPOTHETISCH: JA

(ix) MERKMAL
(A) NAME/SCHLÜSSEL: misc_feature
(B) LAGE: 1,16
(C) SONSTIGE ANGABEN: /Anmerkung= "N ist das massemodifizierte Desoxyuridin-Nukleotid"

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

NAAAACGACG GCCAGNG 17

(2) ANGABEN ZU SEQ ID NO: 6:

(i) SEQUENZKENNZEICHEN:
(A) LÄNGE: 17 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzelstrang
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: andere Nukleinsäure

(iii) HYPOTHETISCH: JA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

GTAAAACGAC GGCCAGT 17

**Patentansprüche**

1. Verfahren zum Sequenzieren einer Nukleinsäure, umfassend die Schritte:
a) Synthetisieren komplementärer Nukleinsäuren, die komplementär zu der Nukleinsäure sind, die sequenziert werden soll, ausgehend von einem Nukleinsäureprimer und in der Anwesenheit von kettenterminierenden und kettenverlängernden Nukleotiden, um vier Sätze von basenspezifisch terminierten komplementären Nukleinsäurefragmenten herzustellen,
b) gleichzeitiges Bestimmen des Molekulargewichtswertes jedes basenspezifisch terminierten Fragmentes durch matrixassistierte Laser-Desorptions-/Ionisationsmassenspektrometrie (MALDI-MS), und
c) Bestimmen der Nukleotidsequenz durch Anordnen der basenspezifisch terminierten Fragmente nach dem Molekulargewicht.

2. Verfahren zum Sequenzieren einer Nukleinsäure, umfassend die Schritte:
a) Synthetisieren komplementärer Nukleinsäuren, die komplementär zu der Nukleinsäure sind, die sequenziert werden soll, ausgehend von einem Nukleinsäureprimer und in der Anwesenheit von kettenterminierenden und kettenverlängernden Nukleotiden, um vier Sätze von basenspezifisch terminierten komplementären Nukleinsäurefragmenten herzustellen,
b) gleichzeitiges Bestimmen des Molekulargewichtswertes jedes basenspezifisch terminierten Fragmentes durch Massenspektrometrie, und
c) Bestimmen der Nukleotidsequenz durch Anordnen der basenspezifisch terminierten Fragmente nach dem Molekulargewicht,
wobei im Verfahren die basenspezifisch terminierten Nukleinsäurefragmente konditioniert werden.

3. Verfahren gemäß Anspruch 1 oder Anspruch 2, wobei die basenspezifisch terminierten Fragmente vor dem Schritt des Bestimmens der Molekulargewichtswerte durch Massenspektrometrie gereinigt werden.

4. Verfahren gemäß Anspruch 3, wobei die basenspezifisch terminierten Fragmente gereinigt werden, umfassend die Schritte
a) Immobilisieren der komplementären Nukleinsäuren auf einem festen Träger, und
b) Auswaschen aller verbliebenen Reaktanten und Nebenprodukte.

5. Verfahren gemäß Anspruch 4, weiterhin umfassend den Schritt des Entfernens der komplementären Nukleinsäuren vom festen Träger.

6. Verfahren gemäß Anspruch 1 oder Anspruch 2, wobei ein Gegenion des Phosphatgerüstes der komplementären Nukleinsäuren entfernt oder gegen ein zweites Gegenion ausgetauscht wird, wobei des zweite Gegenion den Schritt des Bestimmens der Molekulargewichtswerte durch Massenspektroskopie ermöglicht.

7. Verfahren gemäß Anspruch 1 oder Anspruch 2, wobei der Schritt des Bestimmens der Nukleotidsequenz weiterhin Interpolieren der Molekulargewichtswerte umfasst, die für jedes der basenspezifisch terminierten Fragmente bestimmt werden.

8. Verfahren gemäß Anspruch 2, wobei der Molekulargewichtswert von jedem Fragment durch matrixassistierte Laser-Desorptions-/Ionisationsmassenspektrometrie (MALDI-MS) bestimmt wird.

9. Verfahren gemäß Anspruch 2, wobei der Molekulargewichtswert von jedem Fragment durch Elektrospray-Massenspektrometrie (ES-MS) bestimmt wird.

10. Verfahren gemäß Anspruch 1 oder Anspruch 2, wobei die komplementäre Nukleinsäure unter Verwendung eines Nukleinsäureprimers, wenigstens eines Desoxynukleotids ausgewählt aus der Gruppe bestehend aus Desoxyadenosintriphosphat dATP, Desoxythymidintriphosphat dTTP, Desoxyguanosintriphosphat dGTP, Desoxycytidintriphosphat dCTP, Desoxyinosintriphosphat dITP, einem 7-Desazadesoxynukleosidtriphosphat $c^7$dGTP, einem 7-Desazadesoxynukleosidtriphosphat $c^7$dATP und einem 7-Desazadesoxynukleosidtriphosphat $c^7$dITP; wenigstens eines kettenabbrechenden Didesoxynukleotids ausgewählt aus der Gruppe bestehend aus Didesoxyadenosintriphosphat ddATP, Didesoxythymidintriphosphat ddTTP, Didesoxyguanosintriphosphat ddGTP und Didesoxycytidintriphosphat ddCTP, und einer DNA-Polymerase synthetisiert wird.

11. Verfahren gemäß Anspruch 1 oder Anspruch 2, wobei die komplementäre Nukleinsäure unter Verwendung eines Nukleinsäureprimers, wenigstens eines Nukleotids ausgewählt aus der Gruppe bestehend aus Adenosintriphosphat ATP, Uridintriphosphat UTP, Guanosintriphosphat GTP, Cytidintriphosphat CTP, Inosintri-

phosphat ITP, einem 7-Desazanukleosidtriphosphat $c^7$ATP, einem 7-Desazanukleosidtriphosphat $c^7$GTP und einem 7-Desazanukleosidtriphosphat $c^7$ITP; wenigstens eines kettenabbrechenden 3'-Desoxynukleotids ausgewählt aus der Gruppe bestehend aus Desoxyadenosintriphosphat 3'-dATP, Desoxyuridintriphosphat 3'-dUTP, Desoxyguanosintriphosphat 3'-dGTP und Desoxycytidintriphosphat 3'-dCTP, und einer RNA-Polymerase synthetisiert wird.

12. Verfahren gemäß Anspruch 1 oder Anspruch 2, wobei der Nukleinsäureprimer weiterhin eine Linkergruppe (L) zum reversiblen Immobilisieren des Primers auf einem festen Träger einschließt.

13. Verfahren gemäß Anspruch 12, wobei Sätze basenspezifisch terminierter Fragmente durch die Linkergruppe (L) an eine Funktionalität (L') an den Träger gekoppelt werden, wobei eine zeitweilige und spaltbare Verbindung der komplementären Nukleinsäure an den Träger geschaffen wird.

14. Verfahren gemäß Anspruch 13, wobei die zeitweilige und spaltbare Verbindung enzymatisch, chemisch oder physikalisch gespalten werden kann.

15. Verfahren gemäß Anspruch 14, wobei die zeitweilige und spaltbare Verbindung aus der Gruppe bestehend aus einer photospaltbaren Bindung, einer Bindung basierend auf einer starken elektrostatischen Wechselwirkung, einer Trityletherbindung, einer β-Benzoylpropionylgruppe, einer Levulinylgruppe, einer Disulfidbindung, einer Arginin/Argininbindung, einer Lysin/Lysinbindung, einer Pyrophosphatbindung und einer Bindung gebildet durch Watson-Crick-Basenpaarung ausgewählt wird.

16. Verfahren gemäß Anspruch 13, wobei die trägergebundenen basenspezifisch terminierten Fragmente vollständig gewaschen werden, um alle verbliebenen Reaktanten und Nebenprodukte der Sequenzierungsreaktion zu entfernen.

17. Verfahren gemäß Anspruch 16, wobei die basenspezifisch terminierten Fragmente vom festen Träger vor der Massenspektrometrie abgespalten werden.

18. Verfahren gemäß Anspruch 16, wobei die basenspezifisch terminierten Fragmente vom festen Träger während der Massenspektrometrie abgespalten werden.

19. Verfahren gemäß Anspruch 1 oder Anspruch 2, wobei die kettenverlängernden Nukleotide und/oder die kettenterminierenden Nukleotide und/oder der Primer modifizierte Nukleotide einschließen.

20. Verfahren gemäß Anspruch 19, wobei die modifizierten Nukleotide Phosphorthioatnukleotide einschließen.

21. Verfahren gemäß Anspruch 20, wobei die Phosphorthioatnukleotide alkylierte Phosphorthioatnukleotide umfassen.

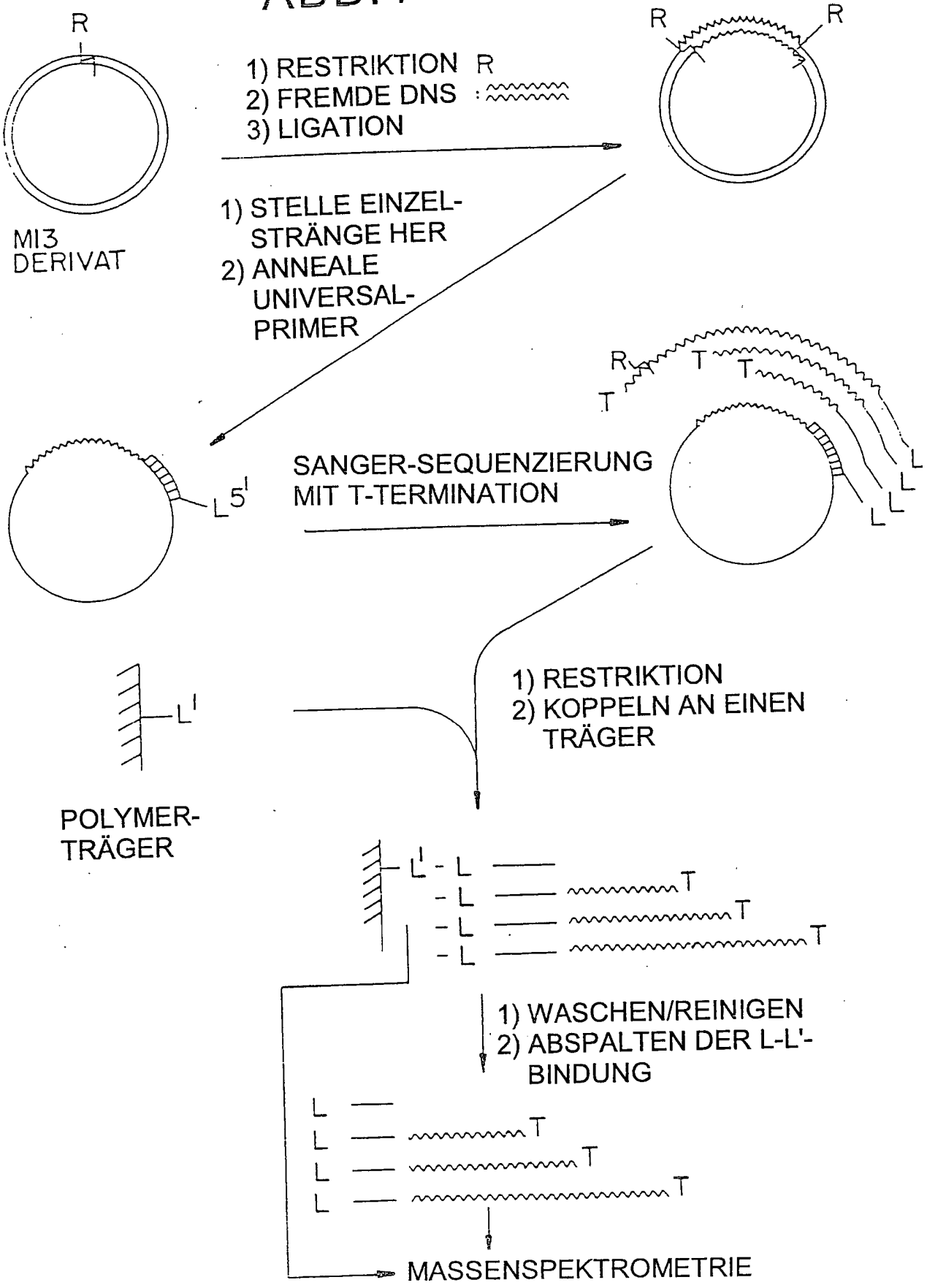
22. Verfahren gemäß Anspruch 4, wobei das Immobilisieren durch eine durch eine Pyrophosphatase spaltbare Bindung bewirkt wird.

23. Verfahren gemäß Anspruch 13, wobei die Verbindung durch eine durch Pyrophosphatase spaltbare Bindung bewirkt wird.

Es folgen 36 Blatt Zeichnungen

**Anhängende Zeichnungen**

ABB. 1



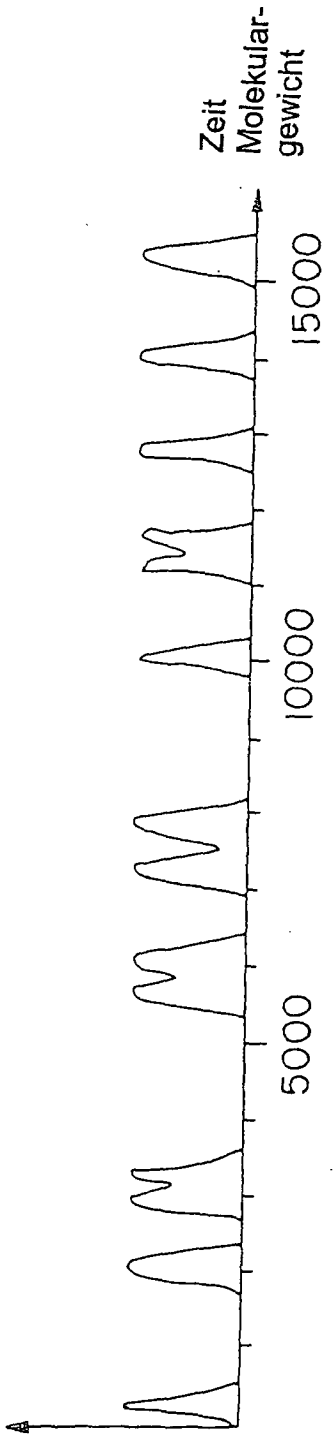
# ABB. 2A

| | |
|---|---|
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATGACTAGCT | |
| 5'-dT | 226.23 |
| 5'-dTAACGGT | 2104.45 |
| 5'-dTAACGGTCAT | 3011.04 |
| 5'-dTAACGGTCATT | 3315.24 |
| 5'-dTAACGGTCATTACGGCCAT | 5771.82 |
| 5'-dTAACGGTCATTACGGCCATT | 6076.02 |
| 5'-dTAACGGTCATTACGGCCATTGACT | 7311.82 |
| 5'-dTAACGGTCATTACGGCCATTGACTGT | 7945.22 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCT | 10112.63 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCAT | 11348.43 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATT | 11652.62 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACAT | 12872.42 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATGACT | 14108.22 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATGACTAGCT | 15344.02 |



DE 694 33 811 T2 2005.06.23

ABB. 2B




| | | verwendete Molekulargewichtsinkremente: | |
|---|---|---|---|
| 1 | 5'-dTp | | 305.20412 |
| 1 | 3'-ddT | | 225.22422 |
| 12 | dTp | | 304.19618 |
| 13 | dAp | | 313.20954 |
| 12 | dCp | | 289.18454 |
| 11 | dGp | | 329.20894 |
| 50 | Nukleotide | | |

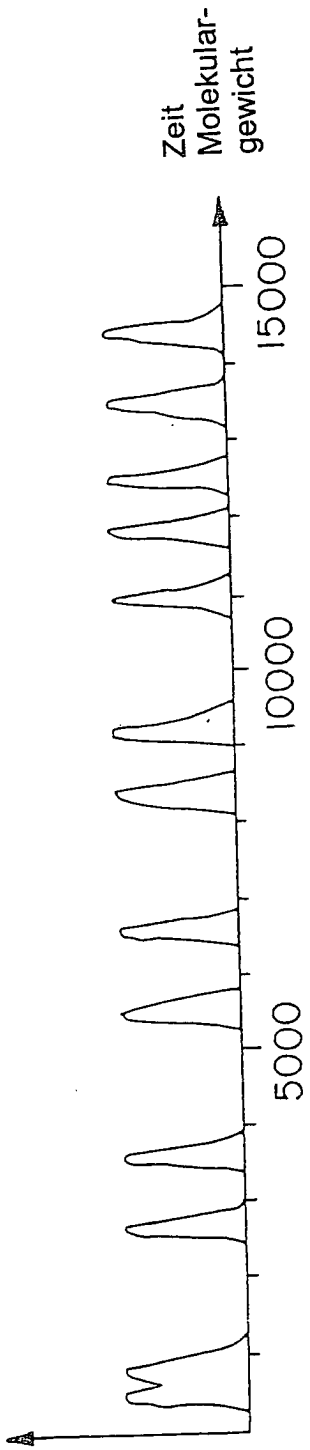
# ABB. 3A

| | |
|---|---|
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATGACTAGCT | |
| 5'-dTA | 530.43 |
| 5'-dTAA | 843.64 |
| 5'-dTAACGGTCA | 2697.83 |
| 5'-dTAACGGTCATTA | 3619.43 |
| 5'-dTAACGGTCATTACGGCCA | 5458.61 |
| 5'-dTAACGGTCATTACGGCCATTGA | 6709.42 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTA | 8249.42 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGA | 9221.05 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCA | 11035.22 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTA | 11956.82 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACA | 12559.21 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATGA | 13505.83 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATGACTA | 14412.42 |



DE 694 33 811 T2 2005.06.23

ABB. 3B




verwendete Molekulargewichtsinkremente:

| | | |
|---|---|---|
| 1 | 5'-dTp | 305.20412 |
| 1 | 3'-ddA | 225.22422 |
| 13 | dTp | 304.19618 |
| 12 | dAp | 313.20954 |
| 12 | dCp | 289.18454 |
| 11 | dGp | 329.20894 |
| 50 | Nukleotide | |

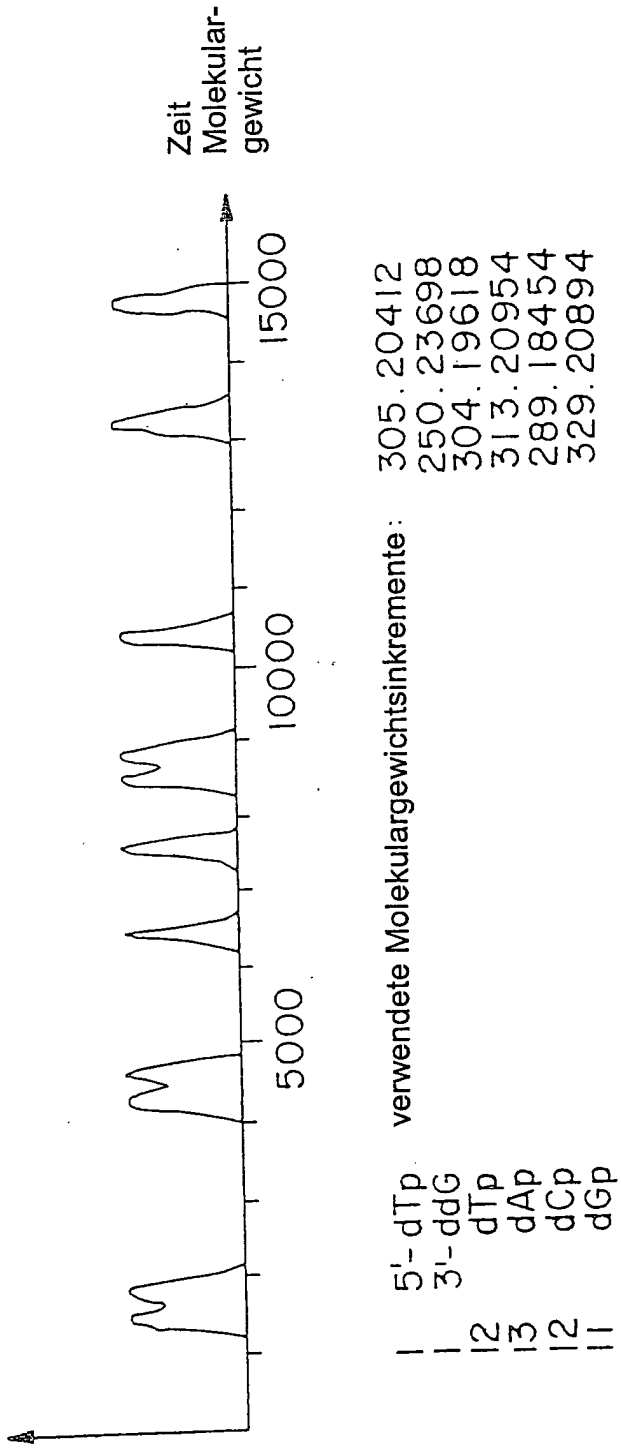
# ABB. 4A

| | |
|---|---|
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATGACTAGCT | |
| 5'-dTAACG | 1471.05 |
| 5'-dTAACGG | 1800.25 |
| 5'-dTAACGGTCATTACG | 4246.84 |
| 5'-dTAACGGTCATTACGG | 4576.05 |
| 5'-dTAACGGTCATTACGGCCATTG | 6405.23 |
| 5'-dTAACGGTCATTACGGCCATTGACTG | 7641.03 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAG | 8587.64 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGG | 8916.85 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTG | 10441.84 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATG | 13201.63 |
| 5'-dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATGACTAG | 14750.64 |



DE 694 33 811 T2 2005.06.23

ABB. 4B


Zeit
Molekular-
gewicht
15000
10000
5000
Ionen-
strom
verwendete Molekulargewichtsinkremente:
1 5'- dTp 305.20412
1 3'- ddG 250.23698
12 dTp 304.19618
13 dAp 313.20954
12 dCp 289.18454
11 dGp 329.20894
50 Nukleotide

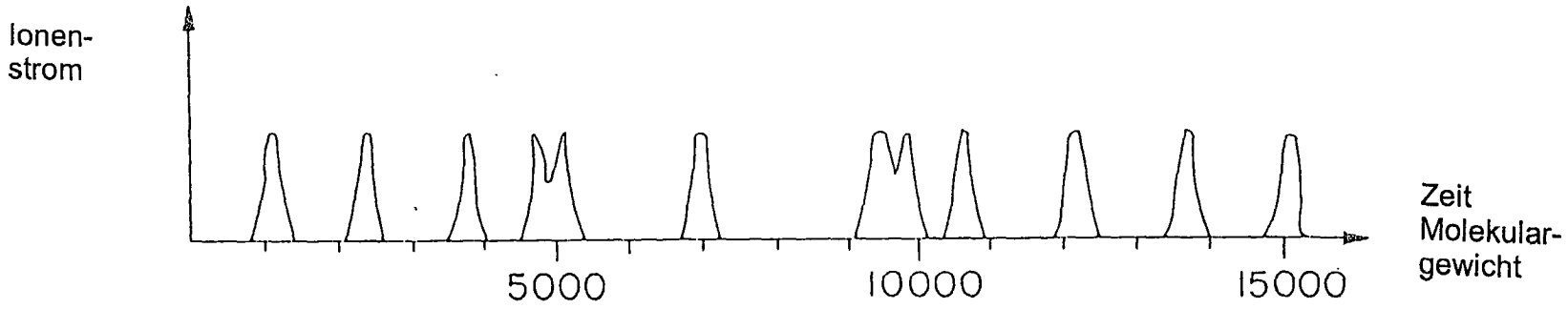
# ABB.5A

| | |
|---|---|
| 5'- dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATGACTAGCT | |
| 5'- dTAAC | 1141.84 |
| 5'- dTAACGGTC | 2393.63 |
| 5'- dTAACGGTCATTAC | 3917.63 |
| 5'- dTAACGGTCATTACGGC | 4865.23 |
| 5'- dTAACGGTCATTACGGCC | 5154.42 |
| 5'- dTAACGGTCATTACGGCCATTGAC | 7007.62 |
| 5'- dTAACGGTCATTACGGCCATTGACTGTAGGAC | 9519.25 |
| 5'- dTAACGGTCATTACGGCCATTGACTGTAGGACC | 9808.43 |
| 5'- dTAACGGTCATTACGGCCATTGACTGTAGGACCTGC | 10731.02 |
| 5'- dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTAC | 12255.02 |
| 5'- dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATGAC | 13804.02 |
| 5'- dTAACGGTCATTACGGCCATTGACTGTAGGACCTGCATTACATGACTAGC | 15039.82 |



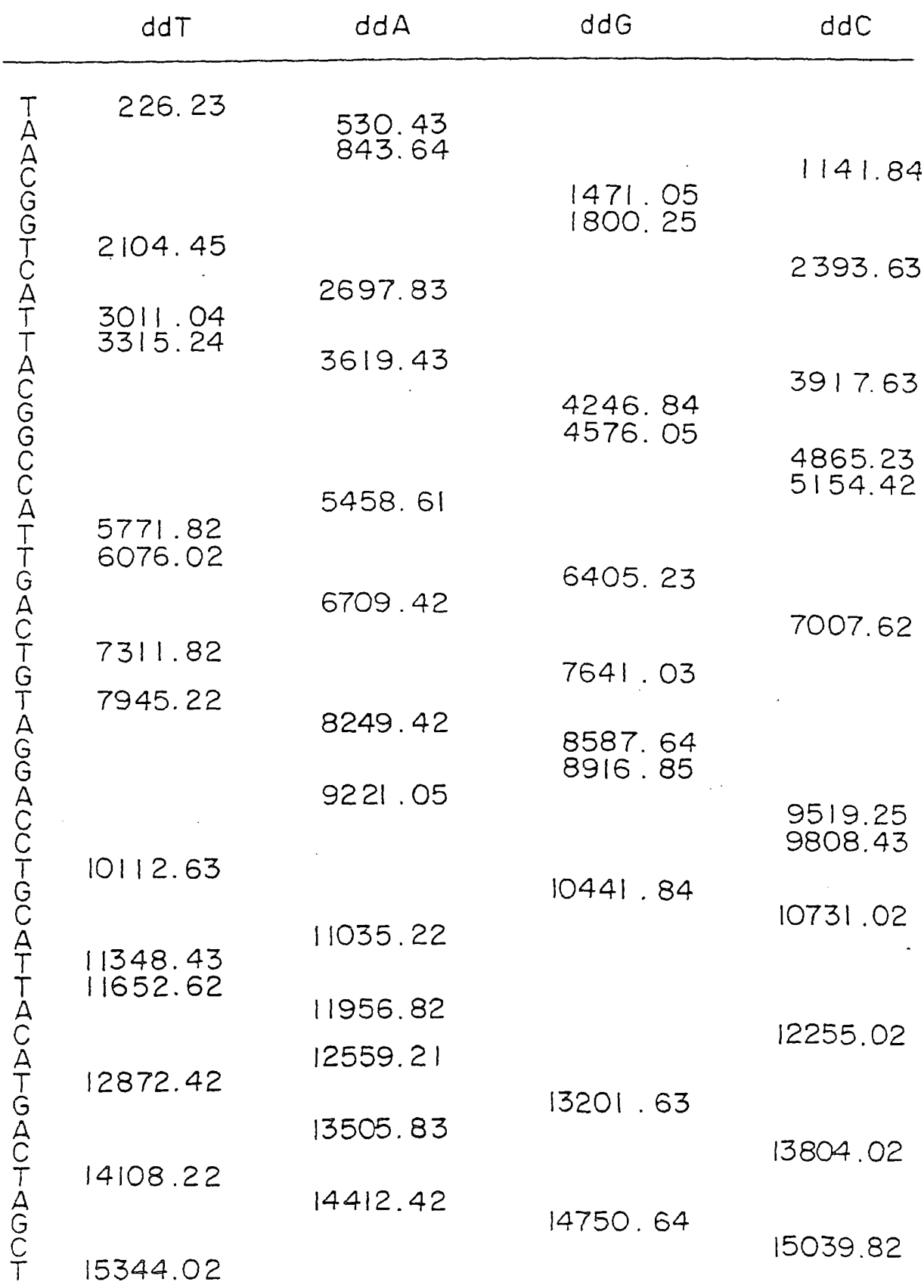
DE 694 33 811 T2 2005.06.23

# ABB. 5B

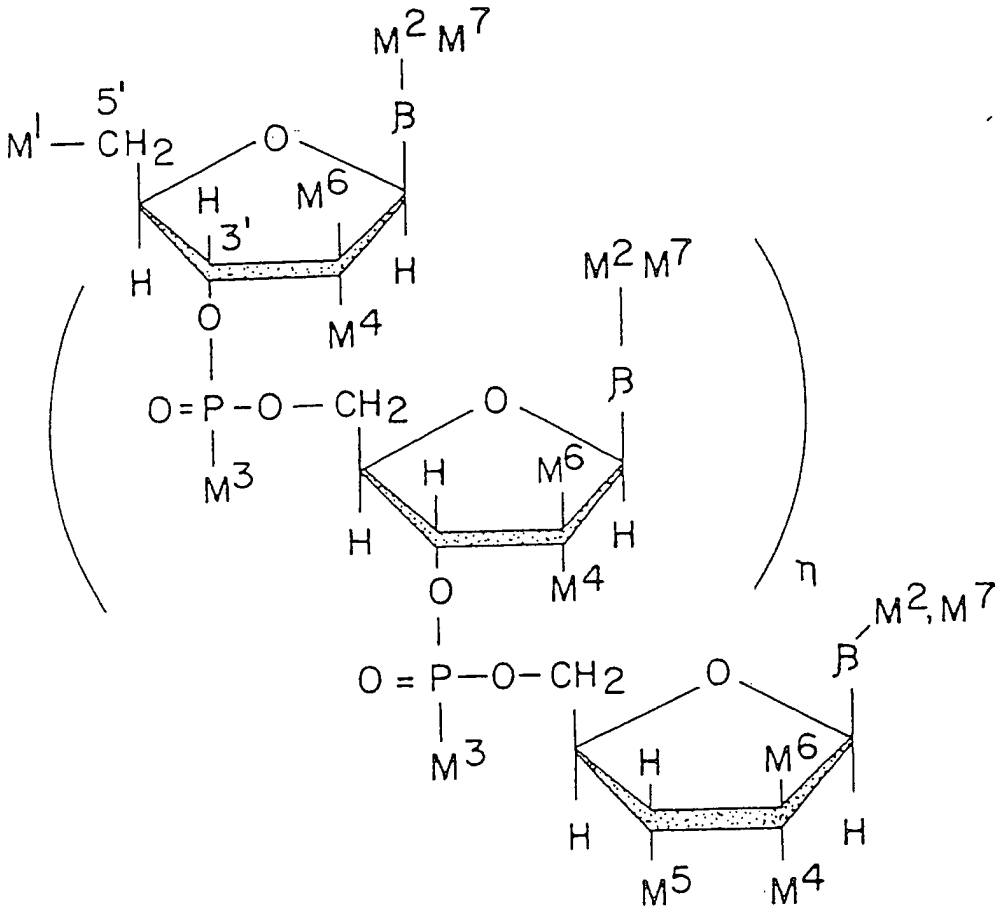



| | | |
|---|---|---|
| 1 | 5'-dTp | verwendete Molekulargewichtsinkremente : 305.20412 |
| 1 | 3'-ddC | 210.21258 |
| 12 | dTp | 304.19618 |
| 13 | dAp | 313.20954 |
| 12 | dCp | 289.18454 |
| 11 | dGp | 329.20894 |
| 50 | Nukleotide | |

# ABB. 6


ddT
ddA
ddG
ddC
T
A
A
C
G
G
T
C
A
T
T
A
C
G
G
C
C
A
T
T
G
A
C
T
G
T
A
G
G
A
C
C
T
G
C
A
T
T
A
C
A
T
G
A
C
T
A
G
C
T
226.23
530.43
843.64
1141.84
1471.05
1800.25
2104.45
2393.63
2697.83
3011.04
3315.24
3619.43
3917.63
4246.84
4576.05
4865.23
5154.42
5458.61
5771.82
6076.02
6405.23
6709.42
7007.62
7311.82
7641.03
7945.22
8249.42
8587.64
8916.85
9221.05
9519.25
9808.43
10112.63
10441.84
10731.02
11035.22
11348.43
11652.62
11956.82
12255.02
12559.21
12872.42
13201.63
13505.83
13804.02
14108.22
14412.42
14750.64
15039.82
15344.02

# ABB. 7A




η = 1 – 50

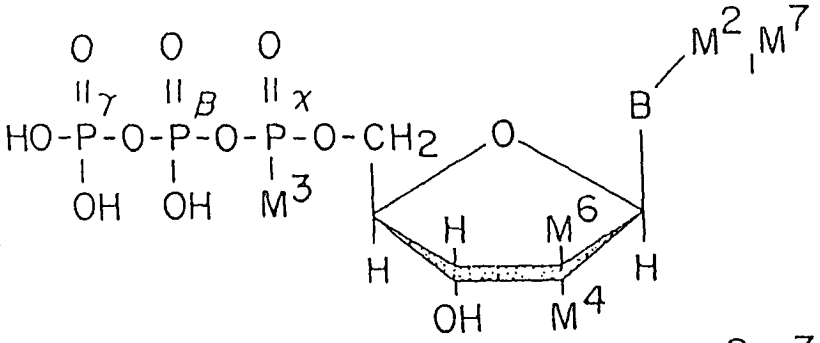
M = H, OH, XR, Halogen, $N_3$

# ABB.7B

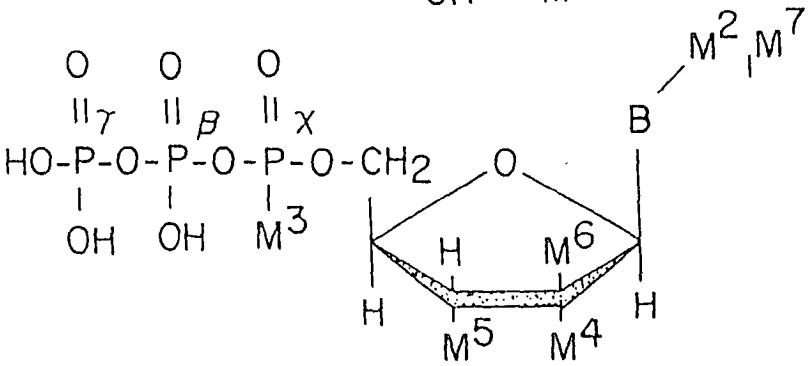
| | $M^1$ | $M^2$ | $M^3$ | $M^5$ |
|---|---|---|---|---|
| Typ Ia (basen-modifizierte DNS | OH | XR/Hal | OH | H |
| Typ Ib (basen-modifizierte RNS) | OH | XR/Hal | OH | OH |
| Typ IIa (5'-modifizierte DNS) | XR/Hal | H | OH | H |
| Typ IIb (5'-modifizierte RNS) | XR/Hal | H | OH | OH |
| Typ III (3'-modifiziert) | OH | H | OH | XR/Hal |
| Typ IVa (P-modifizierte DNS) | OH | H | XR | H |
| Typ IVb (P-modifizierte RNS) | OH | H | XR | OH |

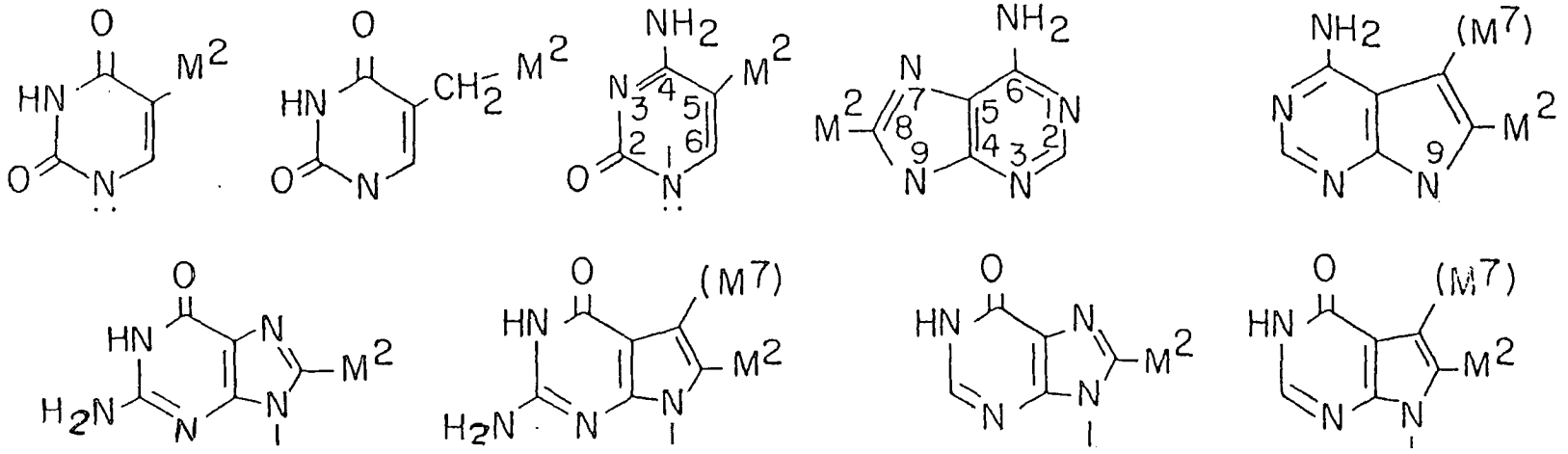
# ABB. 8A

Nukleosidtriphosphatelongatoren:



Nukleosidtriphosphatterminatoren:





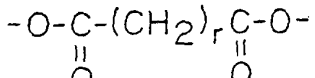
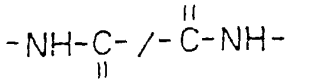
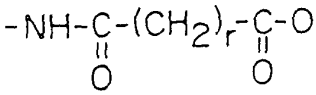
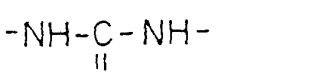
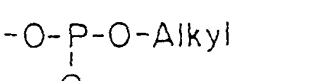
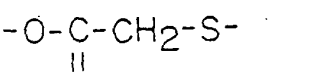
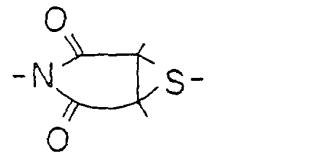


DE 694 33 811 T2 2005.06.23

ABB.8B

| | $M^2$ | $M^3$ | $M^4$ | $M^5$ |
|---|---|---|---|---|
| Typ A (DNS-Termination) | XR | OH | H | H |
| Typ B (DNS-Termination) | H | OH | H | XR |
| Typ C (DNS-Termination) | H | XR | H | H |
| Typ D (RNS-Termination) | XR | OH | OH | H |
| Typ E (RNS-Termination) | H | OH | OH | XR |
| Typ F (RNS-Termination) | H | XR | OH | H |

# ABB. 9

| X | R |
|---|---|
| -O- | $-(CH_2CH_2O)_m-CH_2CH_2-OH$ |
| | or $-(CH_2CH_2O)_m-CH_2CH_2O-Alkyl$ |
|  | $-(CH_2CH_2O)_m-CH_2CH_2-OH$ |
| | or $-(CH_2CH_2O)_m-CH_2CH_2-O-Alkyl$ |
|  | $-(CH_2CH_2O)_m-CH_2CH_2-OH$ |
| | or $-(CH_2CH_2O)_m-CH_2CH_2-O-Alkyl$ |
|  | $-(CH_2CH_2O)_m-CH_2CH_2-OH$ |
| | or $-(CH_2CH_2O)_m-CH_2CH_2-O-Alkyl$ |
|  | $-(CH_2CH_2O)_m-CH_2CH_2-OH$ |
| | or $-(CH_2CH_2O)_m-CH_2CH_2-O-Alkyl$ |
|  | $-(CH_2CH_2O)_m-CH_2CH_2-OH$ |
| | or $-(CH_2CH_2O)_m-CH_2CH_2-O-Alkyl$ |
| $-O-SO_2-O-$ | $-(CH_2CH_2O)_m-CH_2CH_2-OH$ |
| | or $-(CH_2CH_2O)_m-CH_2CH_2-O-Alkyl$ |
|  | $-(CH_2CH_2O)_m-CH_2CH_2-OH$ |
| | or $-(CH_2CH_2O)_m-CH_2CH_2-O-Alkyl$ |
|  | $-(CH_2CH_2O)_m-CH_2CH_2-OH$ |
| | or $-(CH_2CH_2O)_m-CH_2CH_2-O-Alkyl$ |
| -S- | $-(CH_2CH_2O)_m-CH_2CH_2-OH$ |
| | or $-(CH_2CH_2O)_m-CH_2CH_2-O-Alkyl$ |
| -NH- | $-(CH_2CH_2O)_m-CH_2CH_2-OH$ |
| | or $-(CH_2CH_2O)_m-CH_2CH_2-O-Alkyl$ |

m= 0, 1-200
r = 1-20

# ABB. 10

-H

Alkyl: $-(CH_2)_r-CH_3$ e.g. $-CH_3$, $-C_2H_5$,
und verzweigt, z. B. $-CH(CH_3)_2$

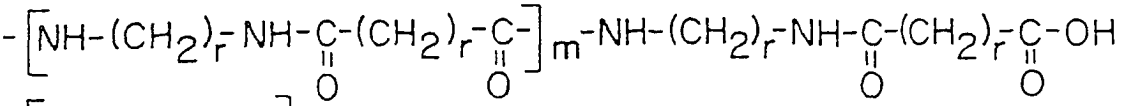
$ICH_2(CH_2)_r-O-H$

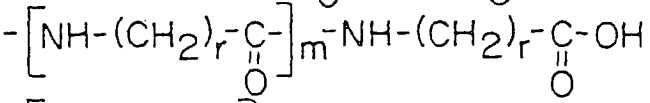
2,3-Epoxy-1-propanol

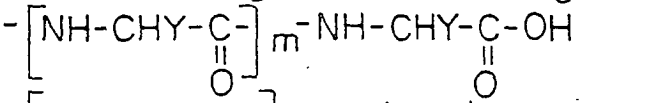
$-(CH_2)_m-CH_2-O-H$

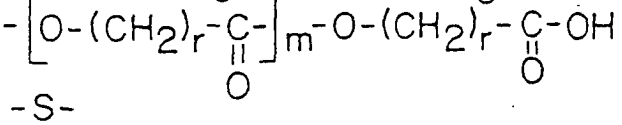
$-(CH_2)_m-CH_2-O-Alkyl$

$-(CH_2CH_2NH)_m-CH_2CH_2-NH_2$









-S-

$-Si(Alkyl)_3$

-Halogen

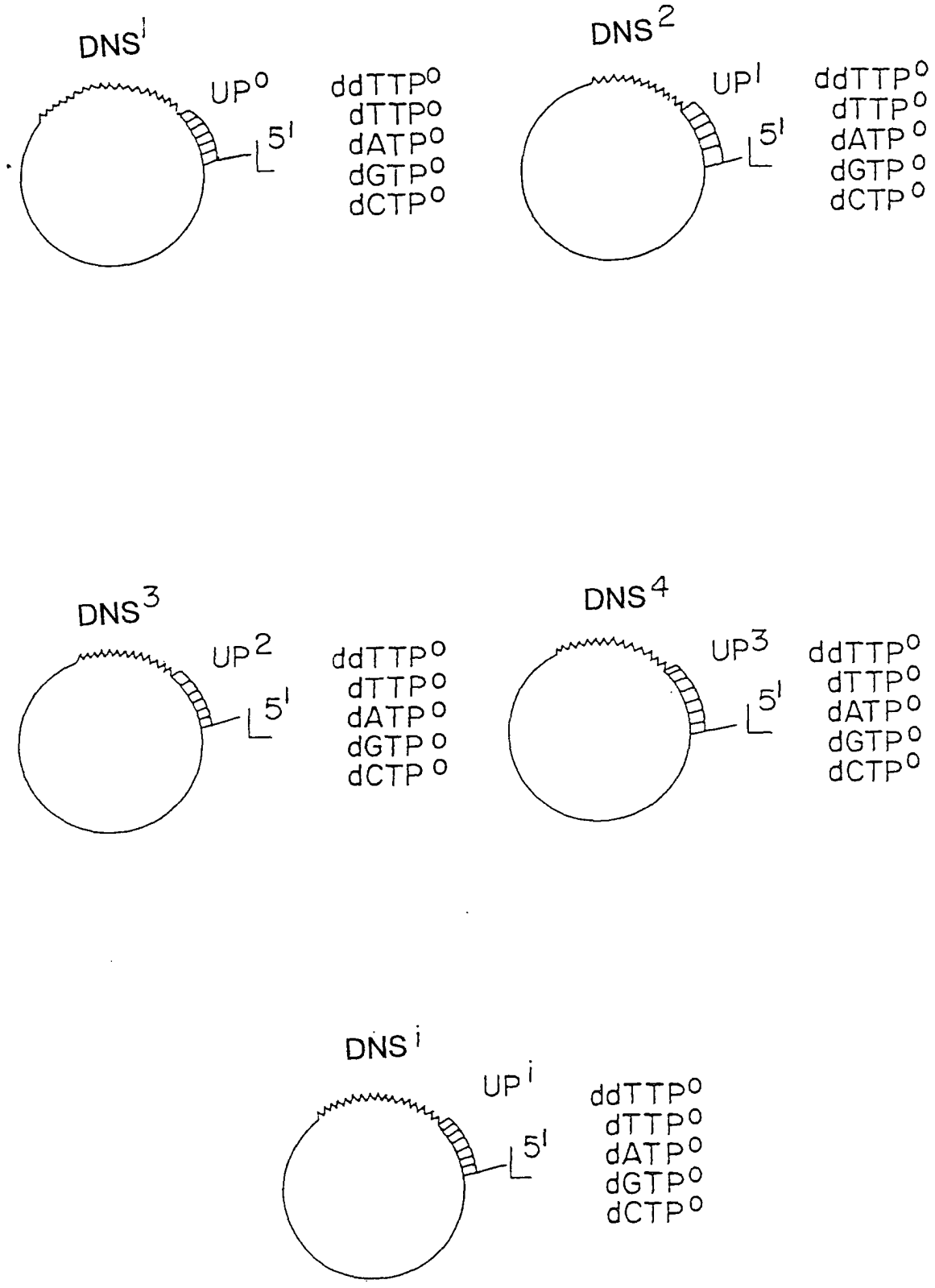
$-N_3$

$-CH_2F$, $-CHF_2$, $-CF_3$

m = 0, 1 - 200
r = 1 - 20

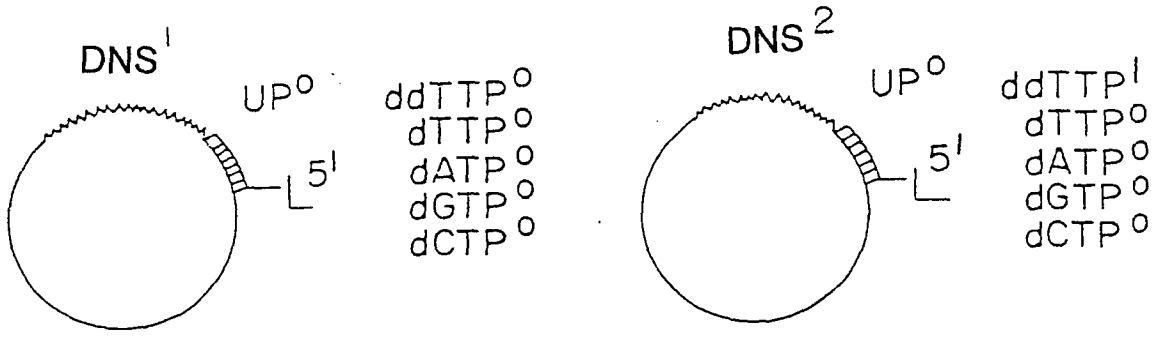
# ABB. 11

z. B. zur dT-Termination:



# ABB. 12

z. B. zur dT-Termination:

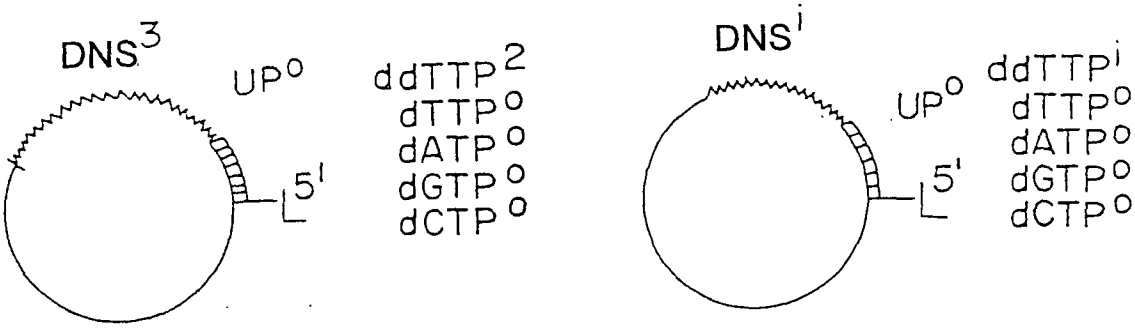



zur Signalamplifikation Elongation mit:

dTTP o
dATP o
dGTP o
dCTP o

dTTP 1
dATP o
dGTP o
dCTP o

z. B. zur dT- Termination:



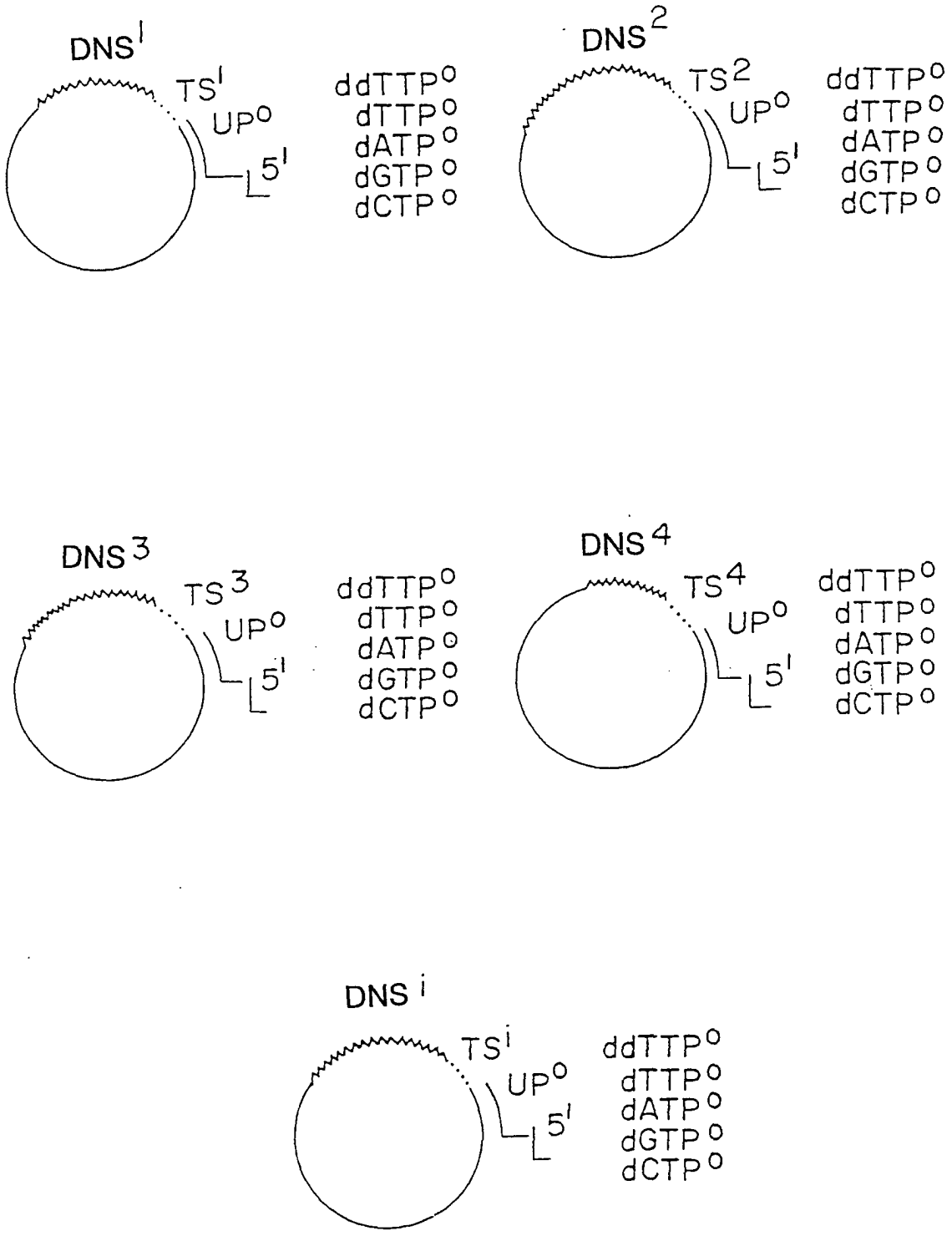

zur Signalamplifikation Elongation mit:

dTTP 2
dATP o
dGTP o
dCTP o

dTTP i
dATP o
dGTP o
dCTP o

# ABB. 13

z. B. zur dT- Termination:




TS= TAG-SEQUENZ

ABB. 14

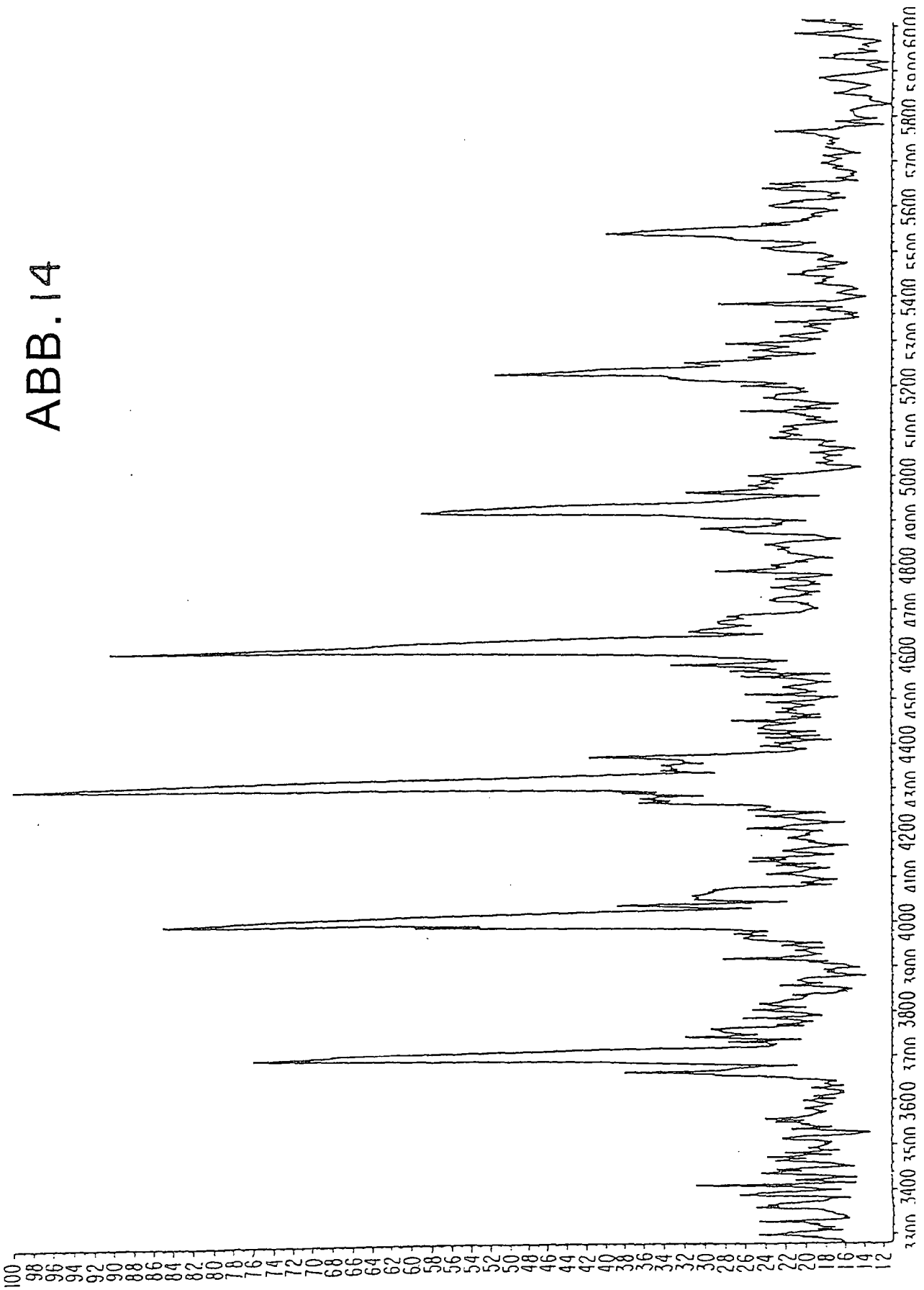

100
98
96
94
92
90
88
86
84
82
80
78
76
74
72
70
68
66
64
62
60
58
56
54
52
50
48
46
44
42
40
38
36
34
32
30
28
26
24
22
20
18
16
14
12
3400
3600
3800
4000
4200
4400
4600
4800
5000
5200
5400
5600
5800

ABB. 15

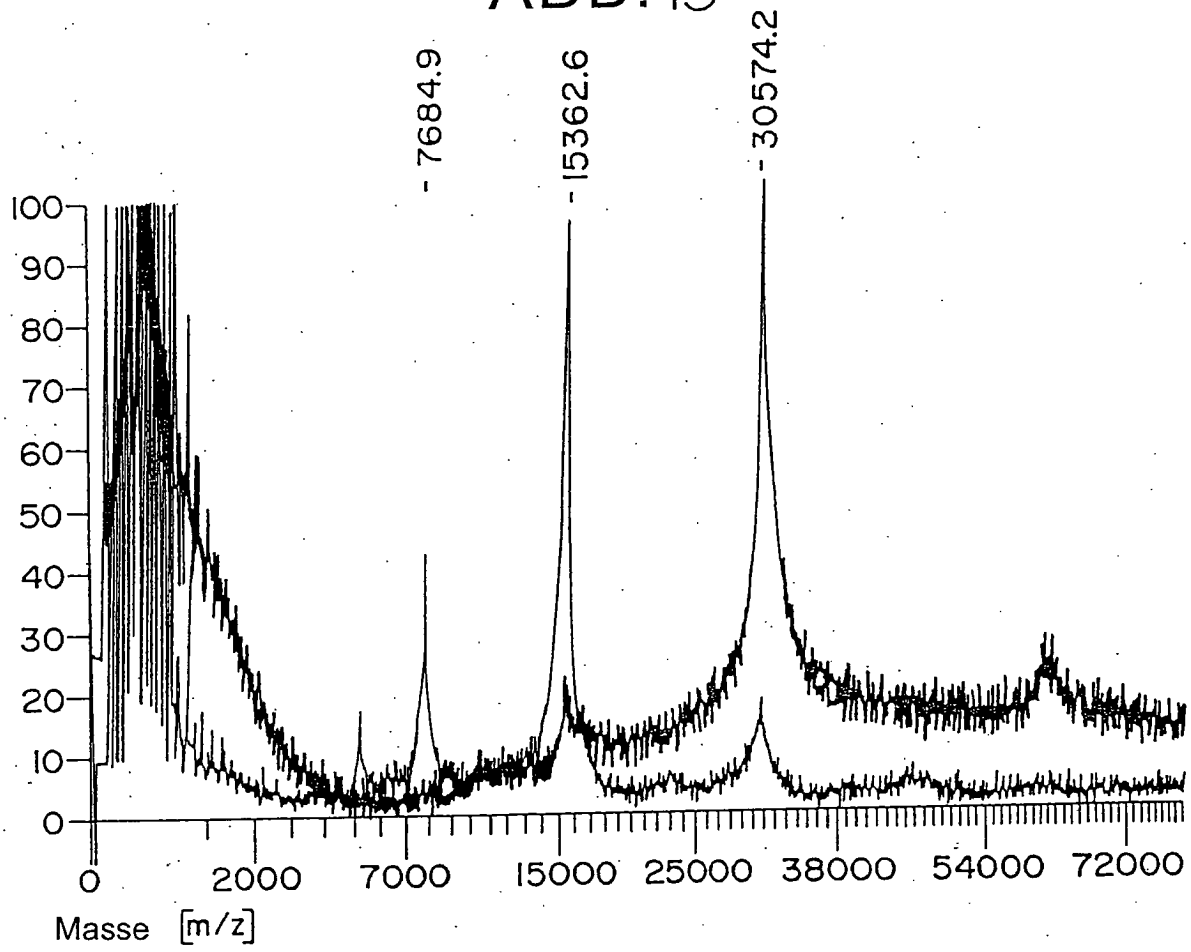

- 7684.9
- 15362.6
- 30574.2
100
90
80
70
60
50
40
30
20
10
0
0
2000
7000
15000
25000
38000
54000
72000
Masse [m/z]

ABB. 16A

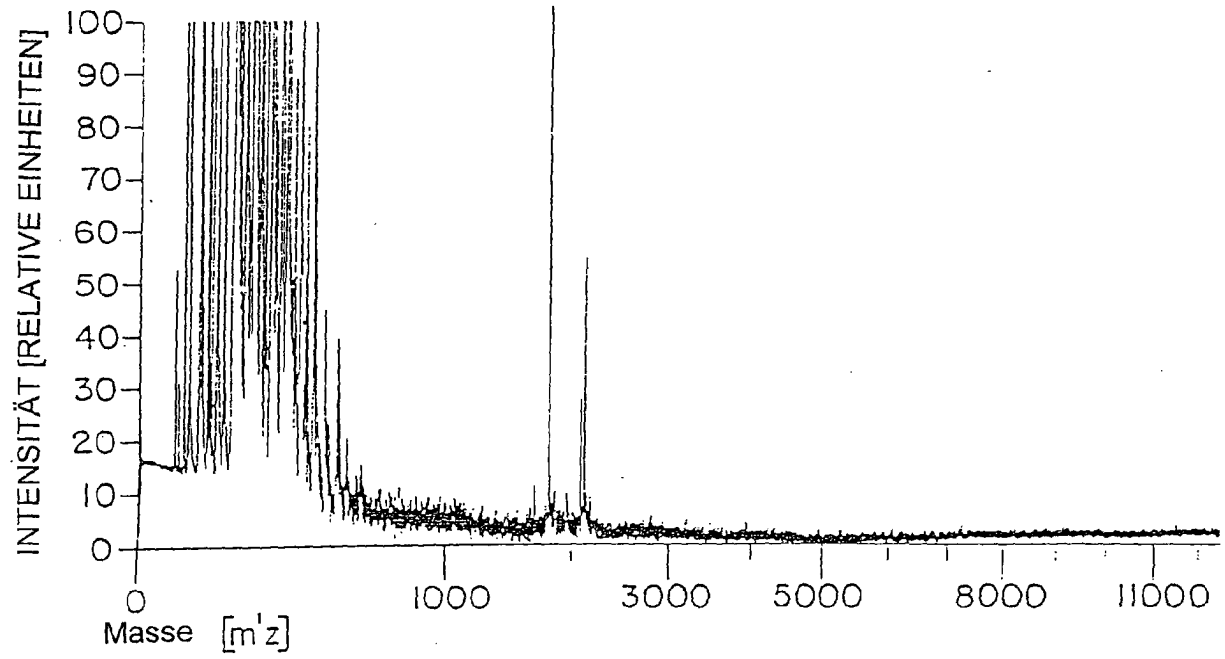



ABB. 16B

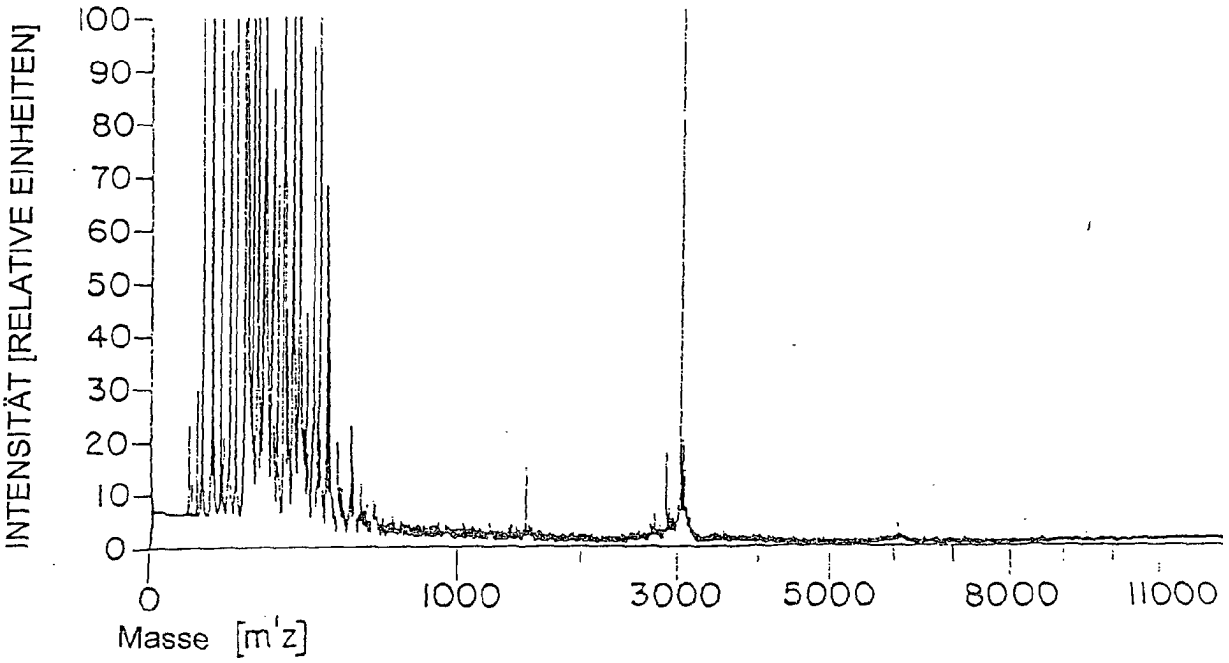


ABB. 16C

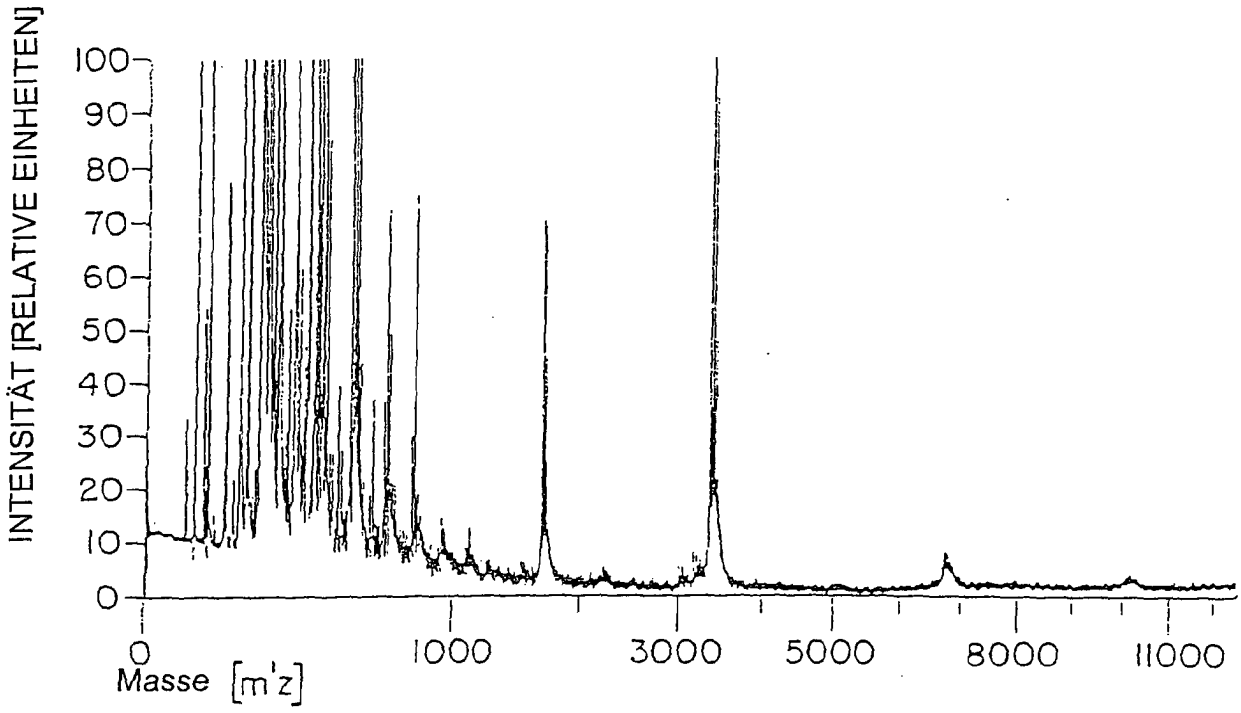



ABB. 16D

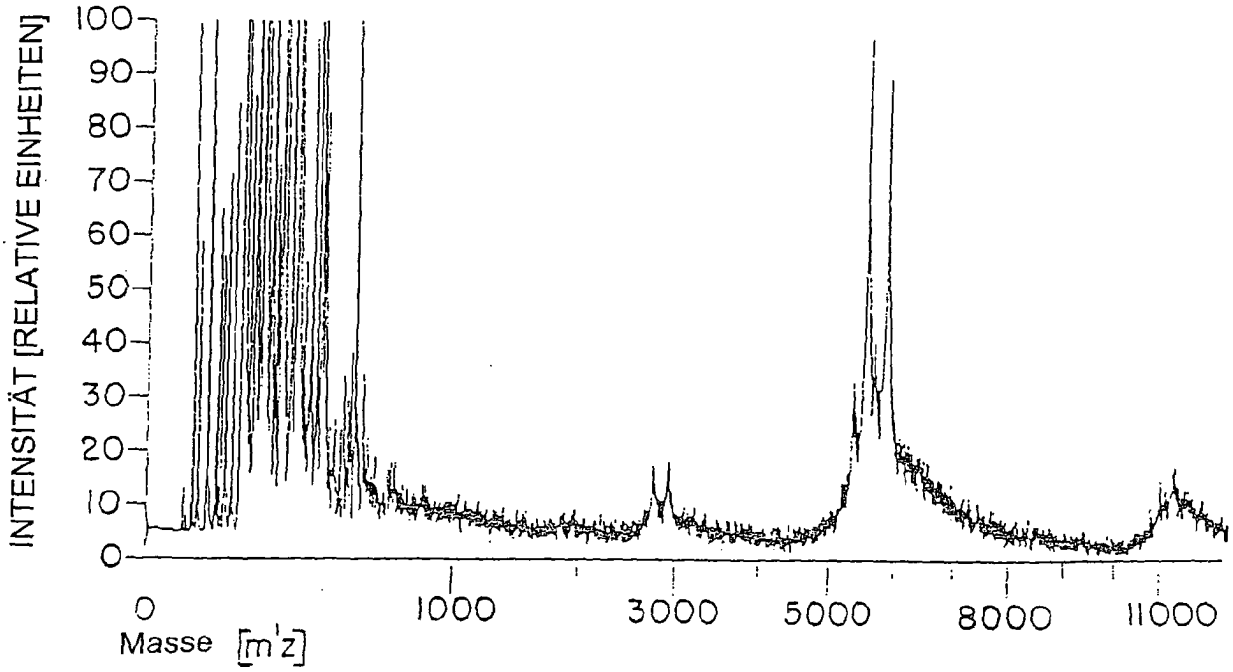


ABB. 16E

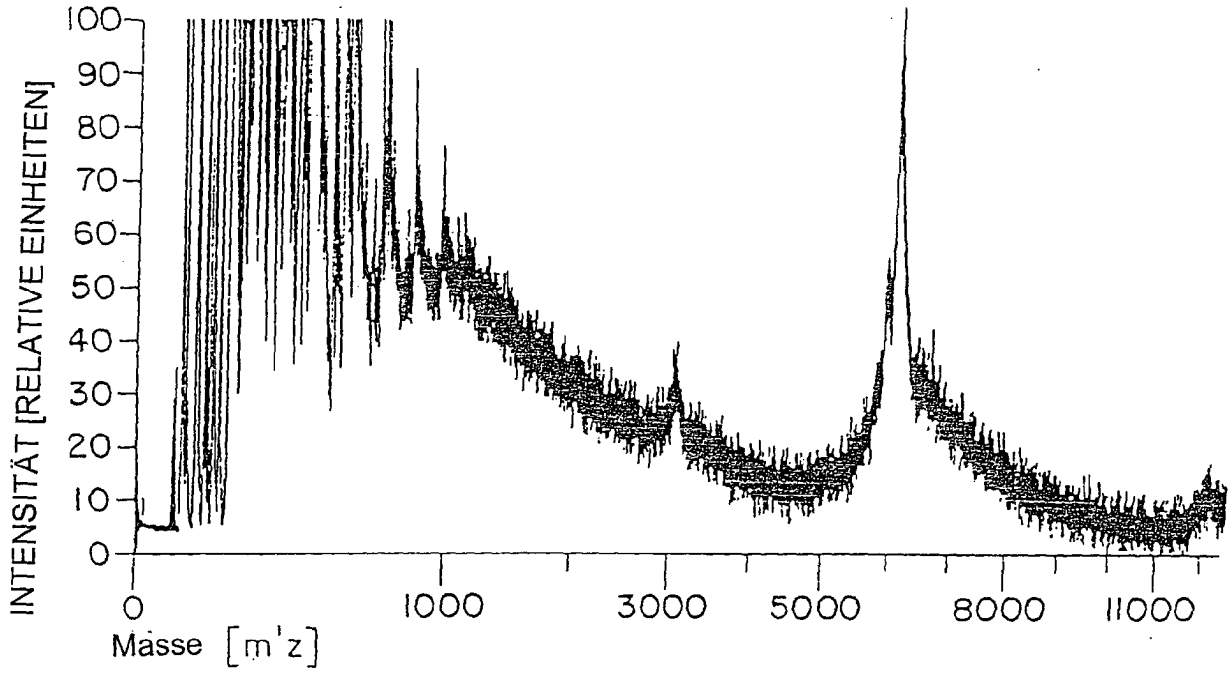



ABB. 16F

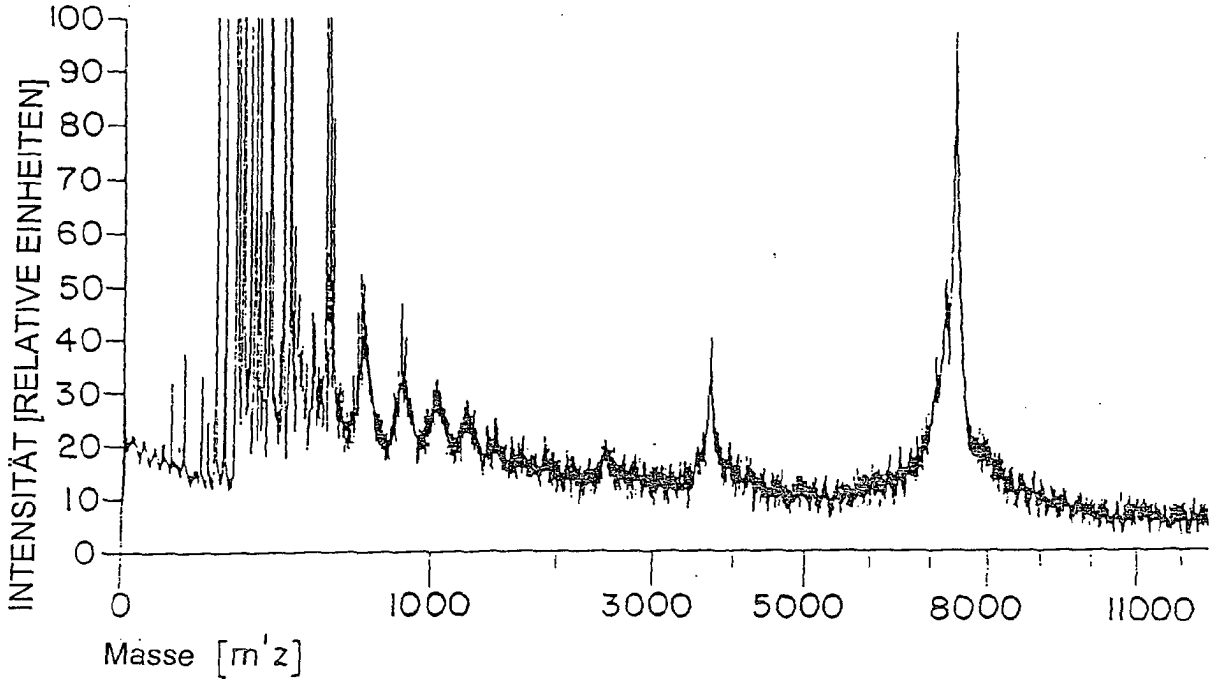


ABB. 16G

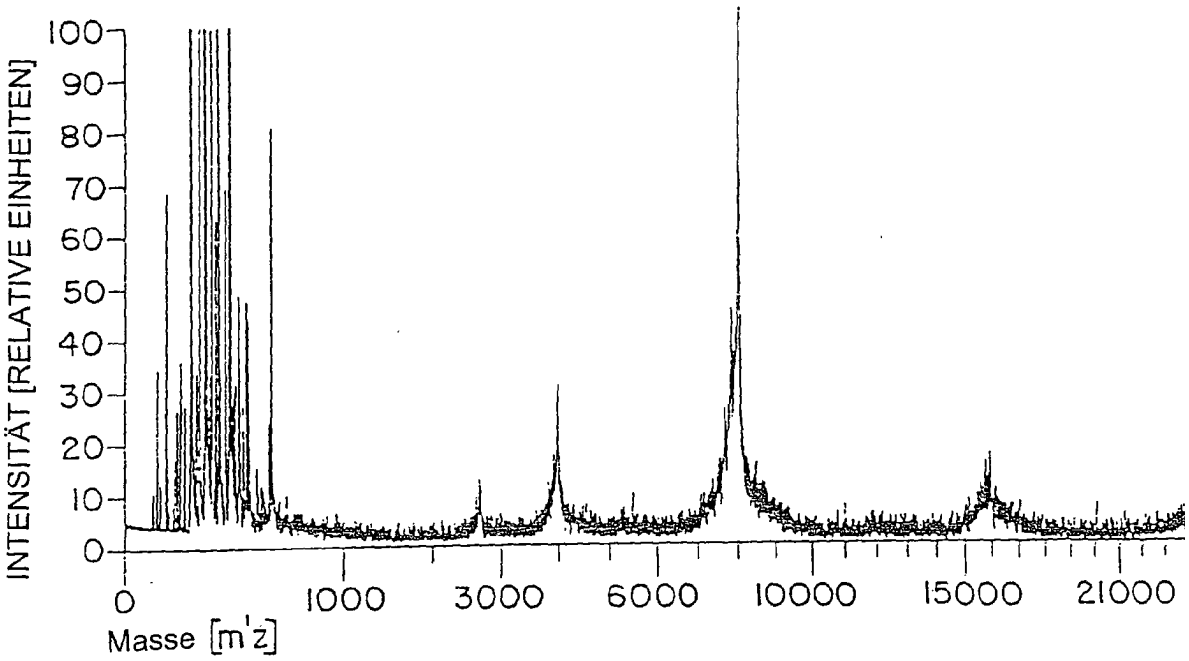



ABB. 16H

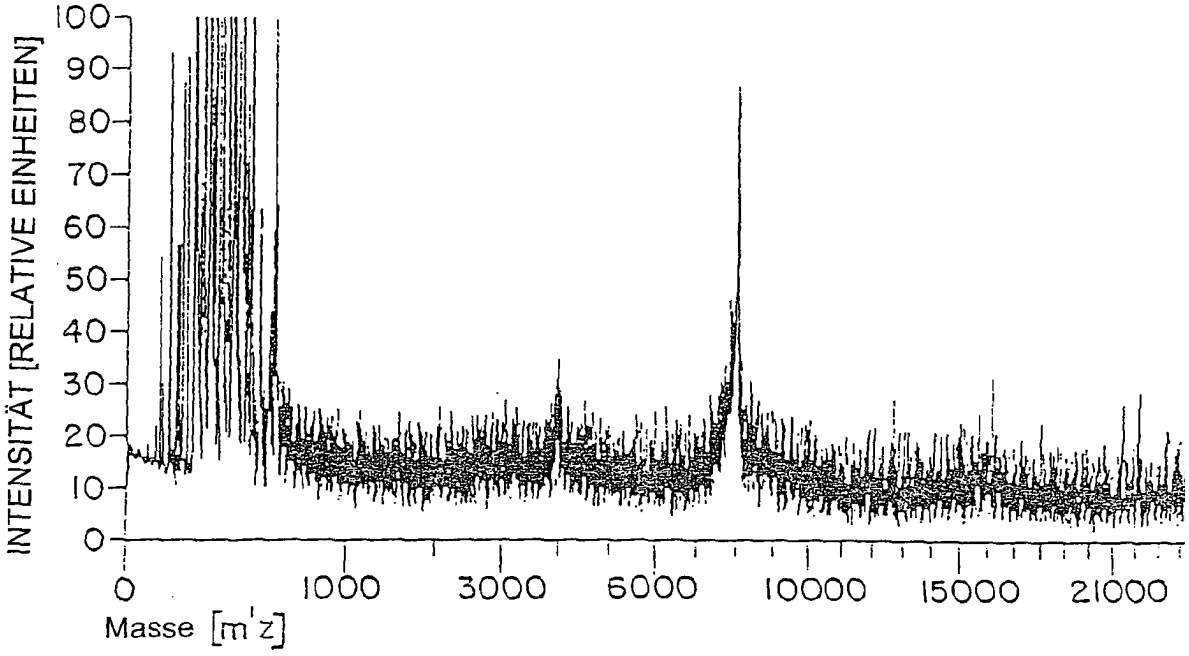


FIG. 16I

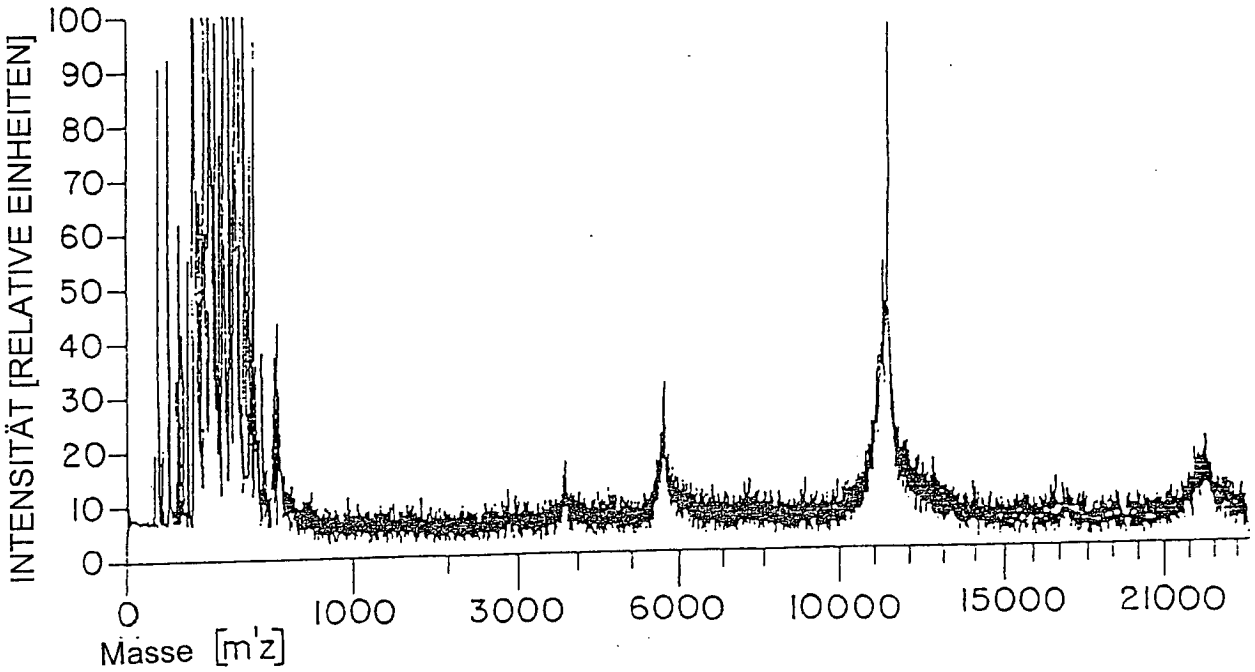



ABB. 16J

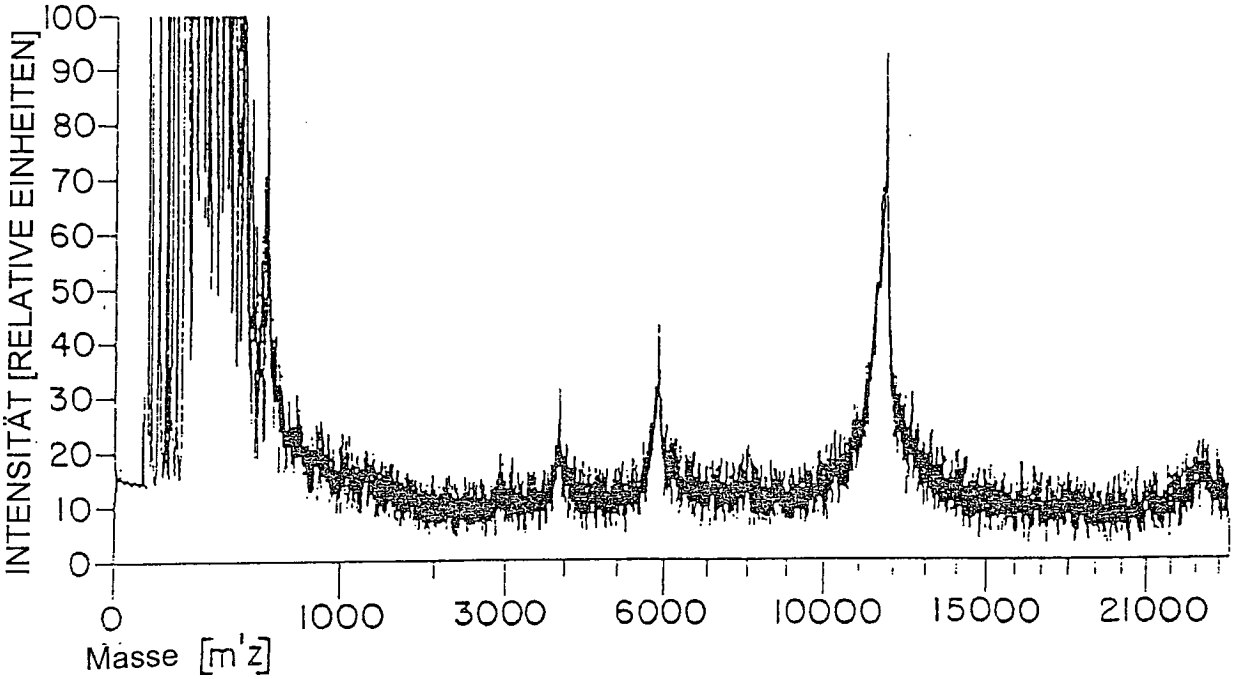


ABB. 16K

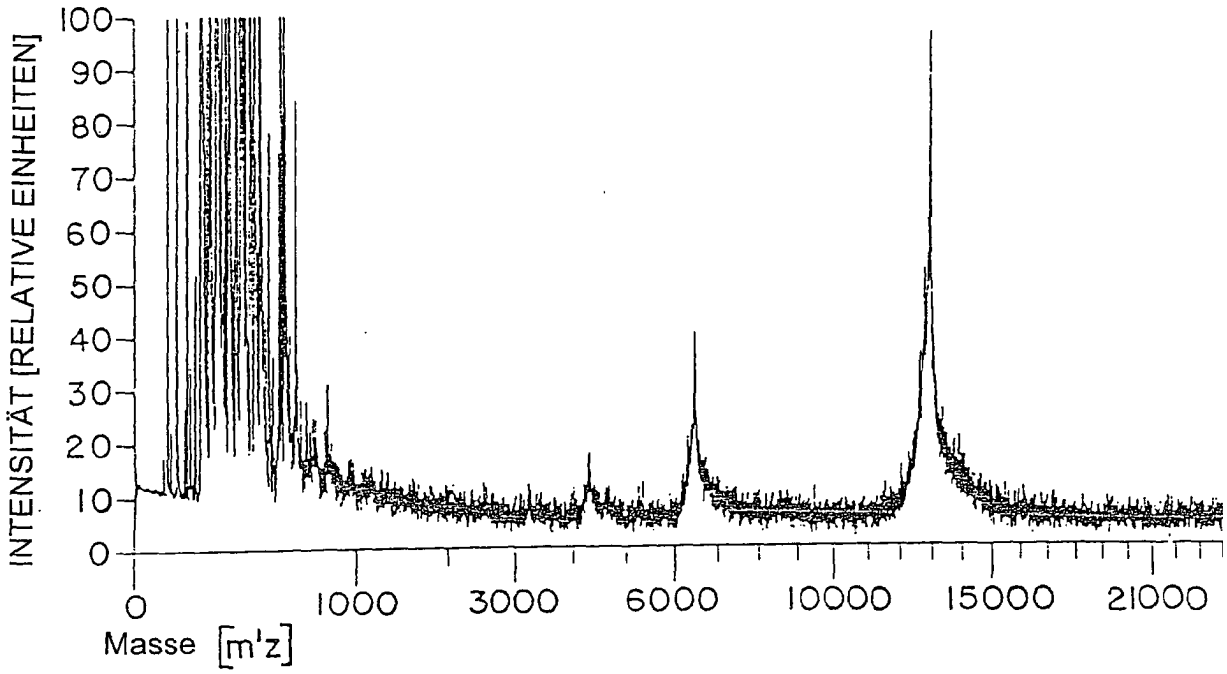



ABB. 16L

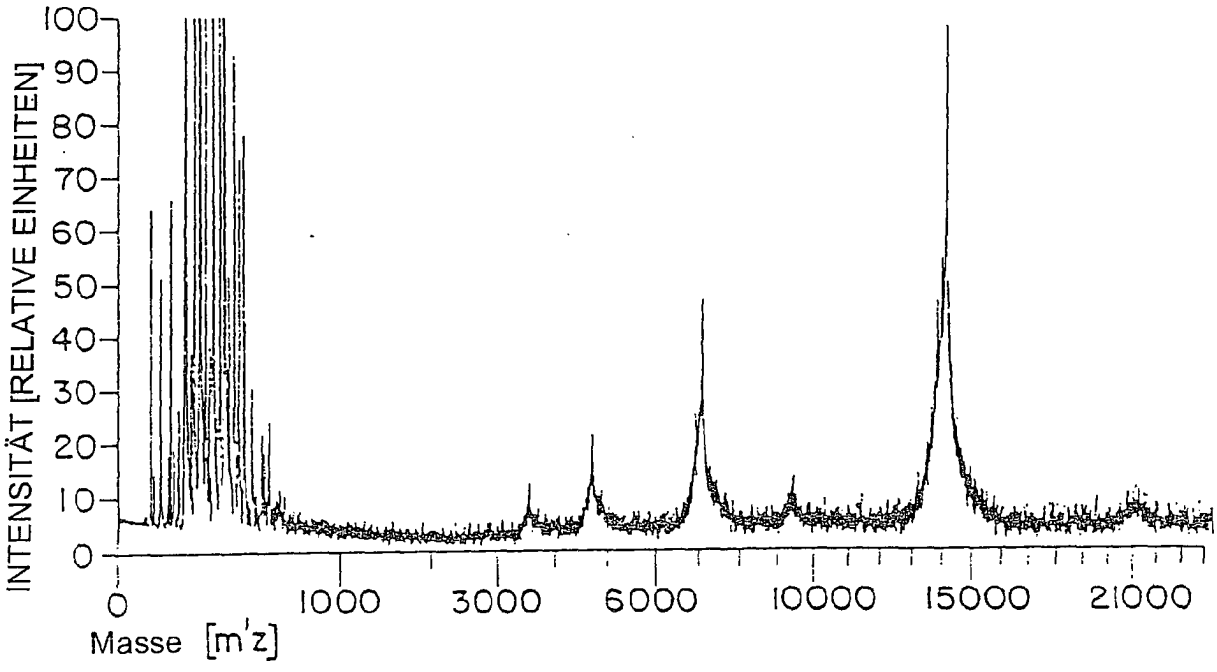


ABB. 16M

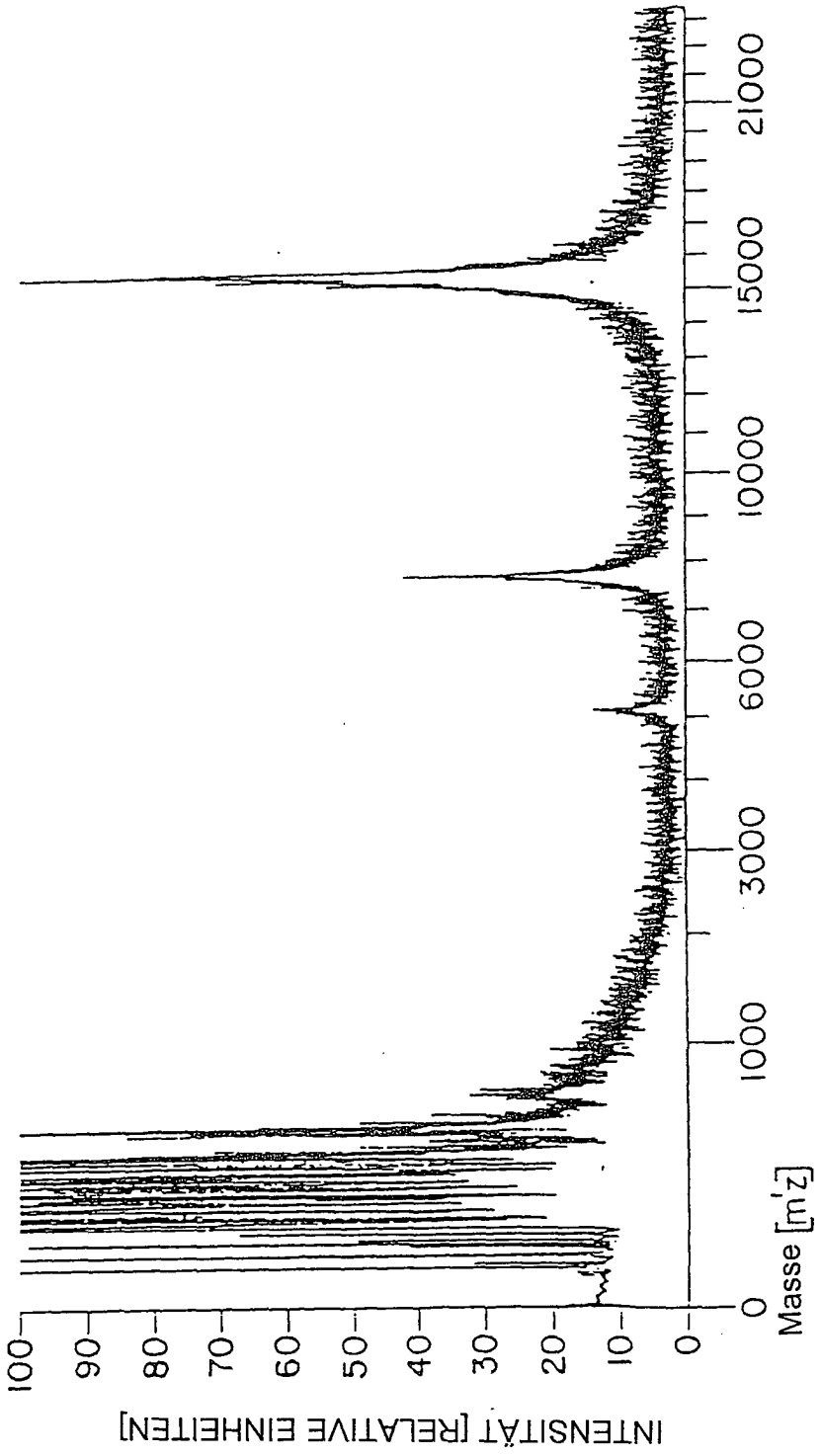

INTENSITÄT [RELATIVE EINHEITEN]
100
90
80
70
60
50
40
30
20
10
0
0
1000
3000
6000
10000
15000
21000
Masse [m/z]

ABB. 17A

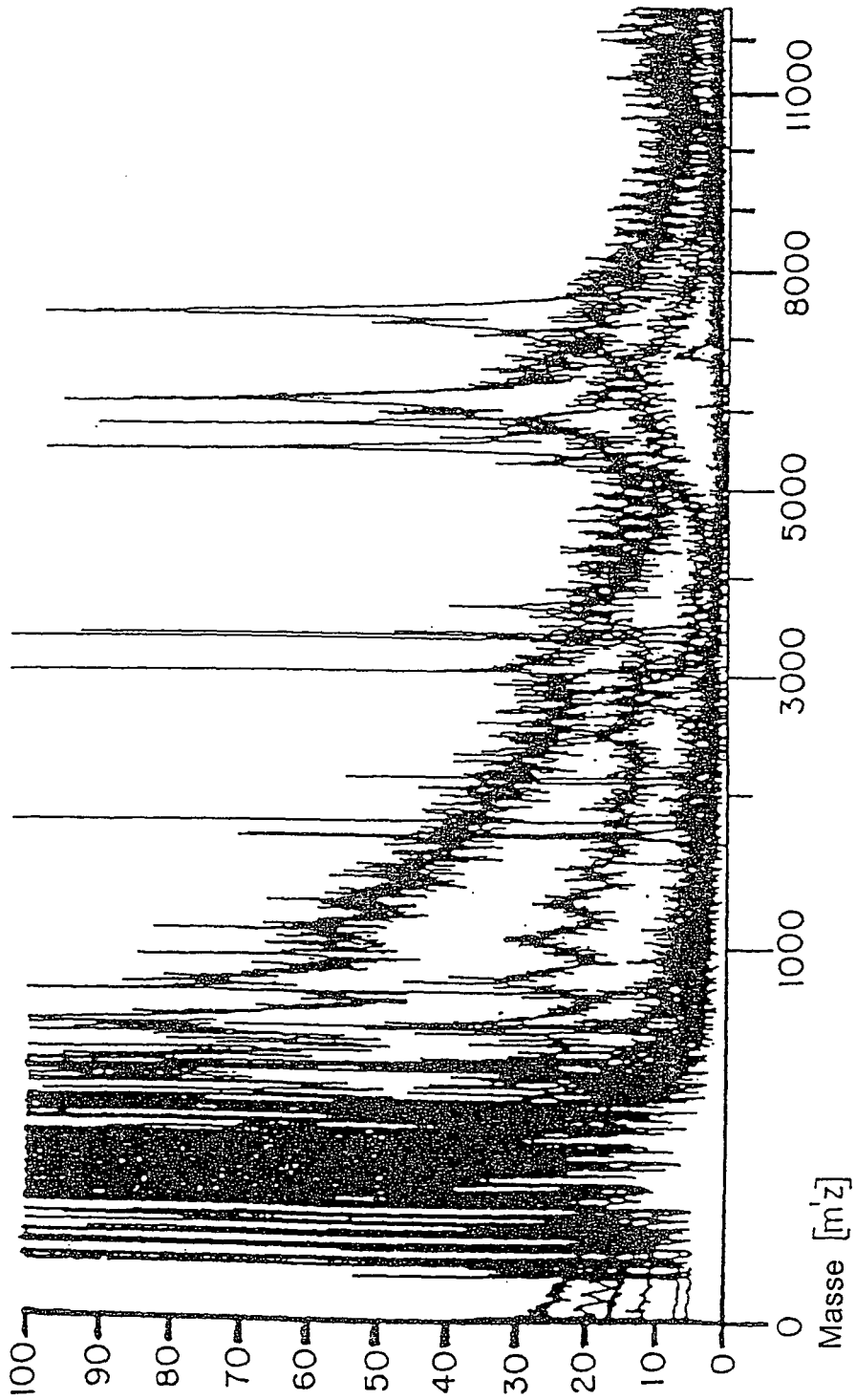

100
90
80
70
60
50
40
30
20
10
0
0
1000
3000
5000
8000
11000
Masse [m'z]

ABB. 17B

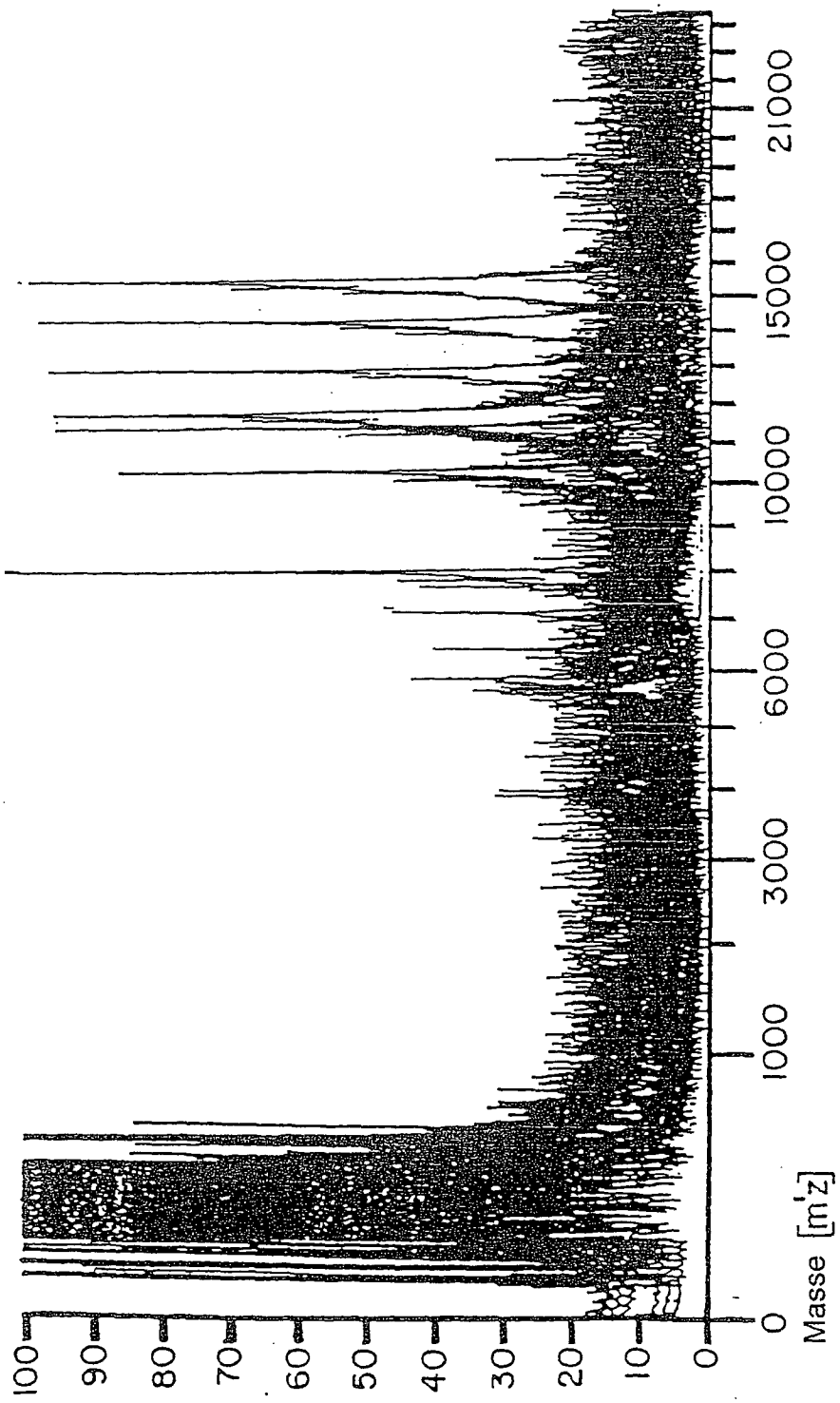
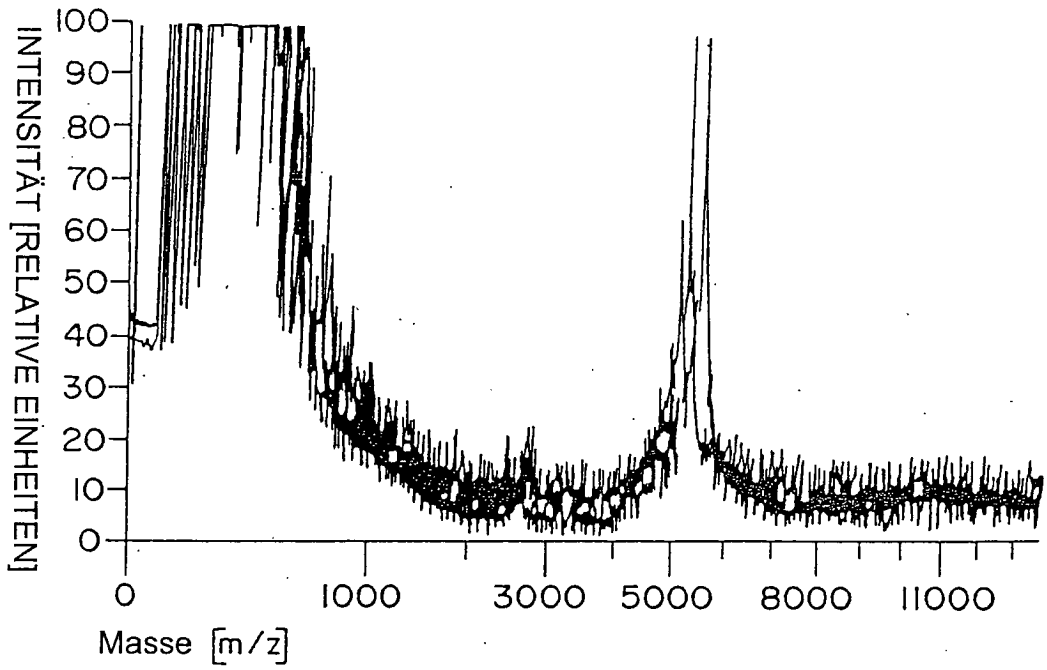
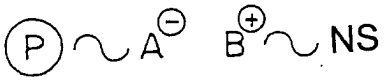

100
90
80
70
60
50
40
30
20
10
0
0
1000
3000
6000
10000
15000
21000
Masse [m'z]

ABB. 18


INTENSITÄT [RELATIVE EINHEITEN]
100
90
80
70
60
50
40
30
20
10
0
0
1000
3000
5000
8000
11000
Masse [m/z]

# ABB. 19

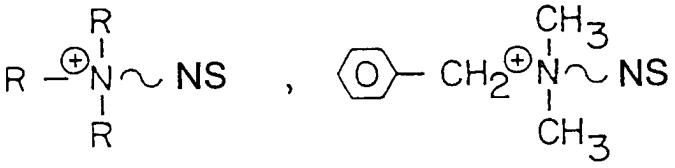


~ : Spacer

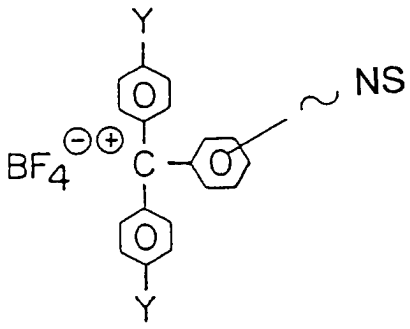
NS: Nukleinsäure (verschachtelte Sanger-Fragmente)

 : Polymerträger

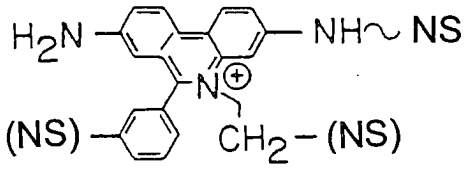
$A^{\ominus}$ : $-CO_2^{\ominus}$, $-SO_3^{\ominus}$, $PO_4^{2-}$, $SiO_4^{3-}$, $PO_3^{2-}$, $SiO_3^{3-}$

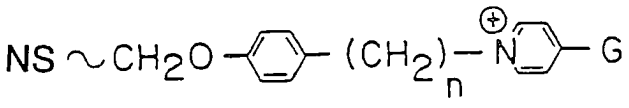
$B^{\oplus}$ : 

R = Alkyl, Aralkyl



Y = H, $OCH_3$, $NR_2$





n = 1-10

G = $NO_2$, CN, F, CL, Br

# ABB. 20A

P∿A  D∿NS  ∿ : Spacer

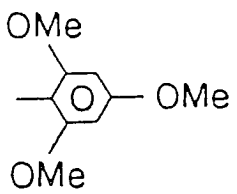
P– : Polymerträger

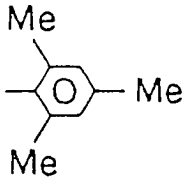
NS : Nukleinsäure (verschachtelte Sanger-Fragmente)

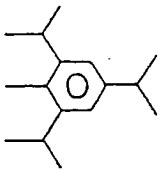
A : Ladungsübertragungsakzeptor

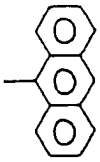
D : Ladungsübertragungsdonator

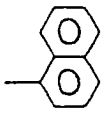
D:

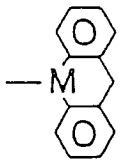


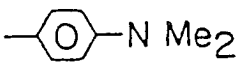



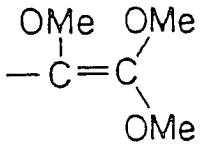



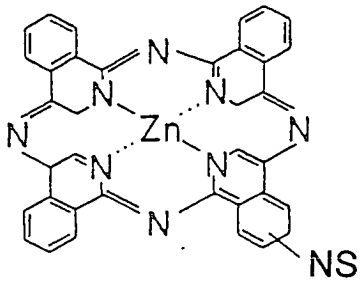




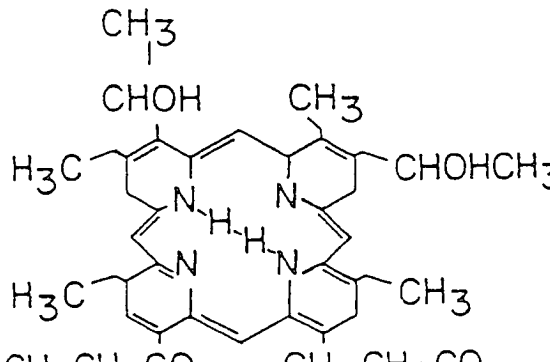


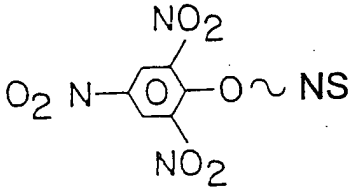

Zink-Phthalocyanin

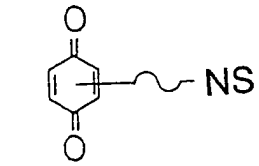


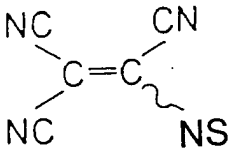

Hämatoporphymin

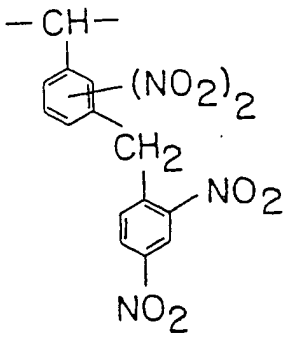
# ABB. 20B

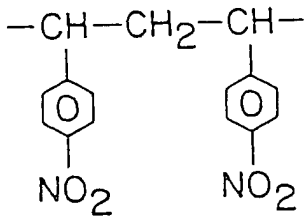
A:

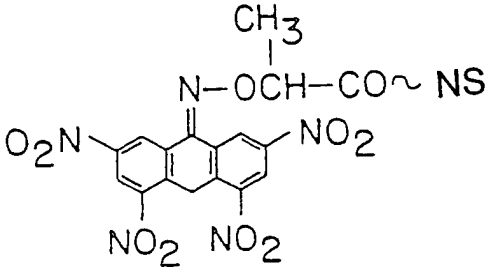




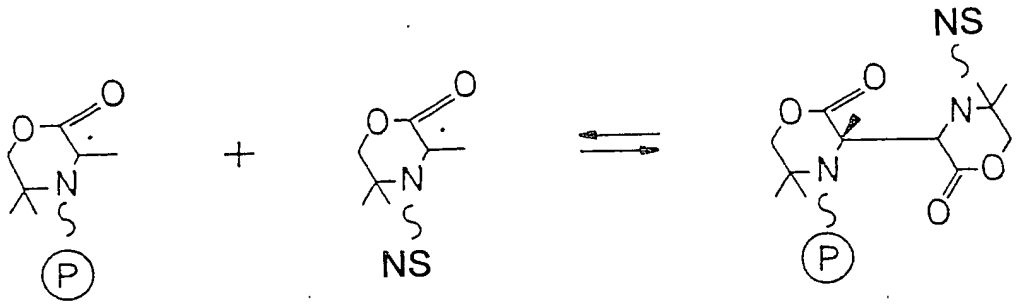




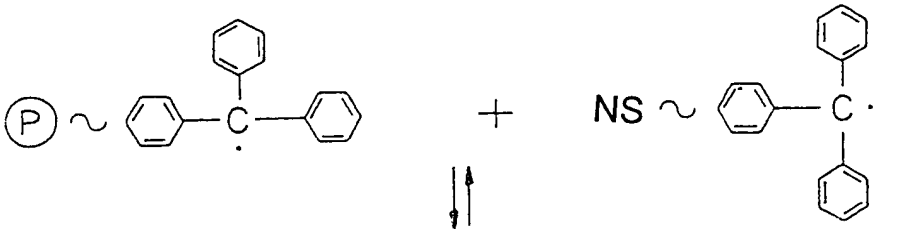




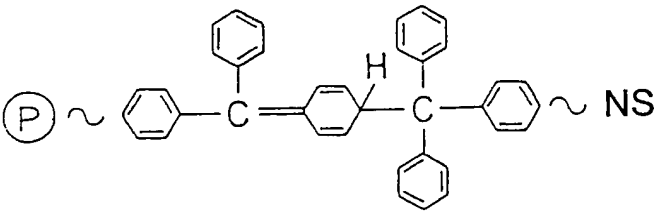
# ABB. 21

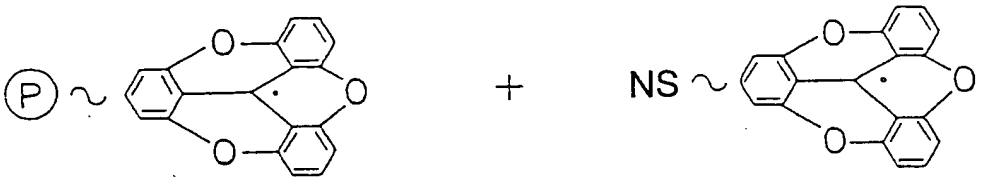


(1)

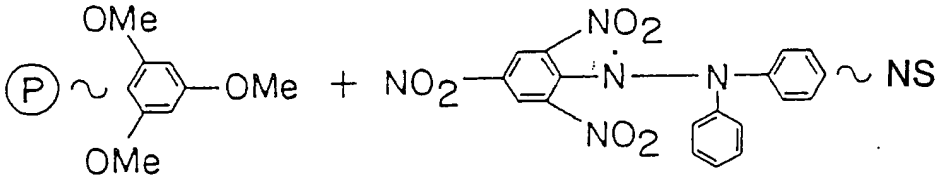


(2)

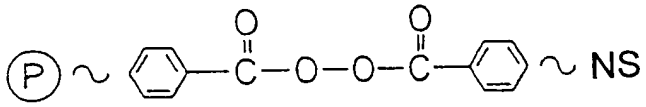




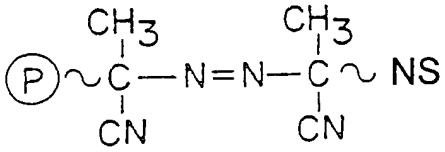
(3)



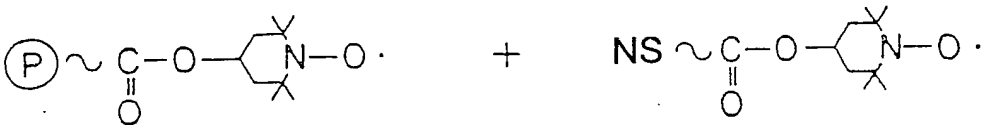
(4)



(5)

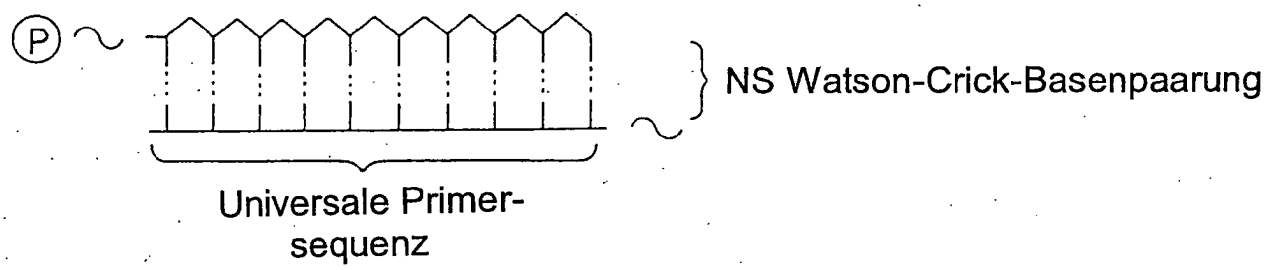


(6)



(7)

# ABB. 22




NS : Nukleinsäure (verschachtelte Sanger-Fragmente)

Ⓟ— : Polymerträger

: Oligoribo-/-Desoxyribo-
nukleotid oder
Oligonukleotidmimetikum
" analogon

# ABB. 23

Ⓟ ~ h·ν ~ NS

Ⓟ : POLYMERTRÄGER

~ : SPACER

h·ν : PHOTOSPALTBARE BINDUNG

NS : NUKLEINSÄURE
(VERSCHACHTELTE SANGER-FRAGMENTE)