



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 698 20 054 T3** 2010.05.06

(12) **Übersetzung der geänderten europäischen Patentschrift**

(97) **EP 0 988 378 B2**  
(21) Deutsches Aktenzeichen: **698 20 054.3**  
(86) PCT-Aktenzeichen: **PCT/GB98/00219**  
(96) Europäisches Aktenzeichen: **98 901 380.0**  
(87) PCT-Veröffentlichungs-Nr.: **WO 1998/032845**  
(86) PCT-Anmeldetag: **26.01.1998**  
(87) Veröffentlichungstag  
der PCT-Anmeldung: **30.07.1998**  
(97) Erstveröffentlichung durch das EPA: **29.03.2000**  
(97) Veröffentlichungstag  
der Patenterteilung beim EPA: **26.11.2003**  
(97) Veröffentlichungstag  
des geänderten Patents beim EPA: **18.11.2009**  
(47) Veröffentlichungstag im Patentblatt: **06.05.2010**

(51) Int Cl.<sup>8</sup>: **C12N 15/10** (2006.01)  
**C07K 16/00** (2006.01)  
**C12Q 1/68** (2006.01)

**Patentschrift wurde im Einspruchsverfahren geändert**

(30) Unionspriorität:  
**9701425**      **24.01.1997**      **GB**

(73) Patentinhaber:  
**Bioinvent International AB, Lund, SE**

(74) Vertreter:  
**Andrae Flach Haug, 83022 Rosenheim**

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

(72) Erfinder:  
**SÖDERLIND, Hans, Ulf, S-247 35 Södra Sandby,  
SE; BORREBAECK, Arne, Carl, S-245 62 Hjärup,  
SE**

(54) Bezeichnung: **METHODE ZUR IN VITRO MOLEKULAREN EVOLUTION EINER PROTEINFUNKTION**

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**

## Gebiet der Erfindung

**[0001]** Die vorliegende Erfindung betrifft ein Verfahren für die molekulare Evolution einer Proteinfunktion in vitro. Im Einzelnen, aber nicht ausschließlich, betrifft sie das Shuffling von Abschnitten einer Polynucleotidsequenz innerhalb einer codierenden Sequenz.

## Hintergrund der Erfindung

**[0002]** Die Funktion eines Proteins kann in vitro mittels verschiedener Verfahren modifiziert und verbessert werden, zu denen die ortsgerichtete Mutagenese (Moore et al. 1987), das kombinatorische Klonieren (Huse et al. 1989, Marks et al. 1992) und die Zufallsmutagenese in Kombination mit geeigneten Selektionssystemen (Barbas et al. 1992) gehören.

**[0003]** Das Verfahren der Zufallsmutagenese zusammen mit einer Selektion wurde in einigen Fällen zur Verbesserung der Proteinfunktion eingesetzt, und es existieren zwei unterschiedliche Strategien. Erstens die Randomisierung der gesamten Gensequenz in Kombination mit der Selektion einer Proteinvariante (Mutante) mit den gewünschten Eigenschaften, gefolgt von einer neuen Runde aus Zufallsmutagenese und Selektion. Dieses Verfahren kann dann wiederholt werden, bis eine Proteinvariante gefunden worden ist, die als optimal angesehen wird (Moore et al. 1996). Hier ist der herkömmliche Weg der Einführung von Mutationen eine fehlerhafte PCR (Leung et al. 1989) mit einer Mutationshäufigkeit von ungefähr 0,7%.

**[0004]** Zweitens können definierte Abschnitte des Gens mit degenerierten Primern mutagenisiert werden, was Mutationshäufigkeiten von bis zu 100% möglich macht (Griffiths et al. 1994, Yang et al. 1995). Je höher die eingesetzte Mutationshäufigkeit ist, desto begrenzter ist der Bereich des Gens, der den Mutationsvorgängen unterzogen werden kann.

**[0005]** Die Zufallsmutagenese wurde auf dem Gebiet des Antikörper-Engineering extensiv eingesetzt. In vivo gebildete Antikörpergene können in vitro kloniert werden (Larrick et al. 1989), und Zufallskombinationen der Gene, die für die variablen schweren und leichten Gene codieren, können einer Selektion unterzogen werden (Marks et al. 1992). Ausgewählte funktionelle Antikörperfragmente können weiter unter Einsatz einer Zufallsmutagenese und weiteren Selektionsrunden verbessert werden (Hoogenboom et al. 1992).

**[0006]** An die Strategie der Zufallsmutagenese schließt sich eine Selektion an. Varianten mit interessanten Charakteristika können selektiert werden, und die mutagenisierten DNA-Abschnitte verschiedener Varianten, von denen jede ein interessantes Charakteristikum zeigt, werden zu einer codierenden Sequenz kombiniert (Yang et al. 1995). Das ist ein sequenzieller Prozess aus mehreren Schritten, und potenzielle synergistische Effekte verschiedener Mutationen in unterschiedlichen Abschnitten können verloren gehen, da sie nicht einer Selektion in Kombination unterzogen werden. Somit schließen diese beiden Strategien nicht die gleichzeitige Mutagenese definierter Abschnitte und Selektion einer Kombination dieser Abschnitte ein. Ein weiterer Prozess beinhaltet die kombinatorische Paarung von Genen, die beispielsweise zur Verbesserung der Antikörperaffinität eingesetzt werden kann (Marks et al. 1992). Hier sind die drei CDR-Abschnitte in jedem variablen Gen fixiert, und diese Technologie ermöglicht kein Shuffling einzelner CDR-Abschnitte zwischen den Klonen.

**[0007]** Die Selektion funktioneller Proteine aus molekularen Bibliotheken wurde durch die Entwicklung der Phage-Display-Technologie revolutioniert (Parmley et al. 1987, McCafferty et al. 1990, Barbas et al. 1991). Hier ist der Phänotyp (das Protein) direkt mit seinem entsprechenden Genotyp (der DNA) verknüpft, und das ermöglicht eine direkte Klonierung des genetischen Materials, das dann weiteren Modifizierungen zur Verbesserung der Proteinfunktion unterzogen werden kann. Das Phage-Display wurde zur Klonierung funktioneller bindender Agenzien aus einer Vielzahl molekularer Bibliotheken mit einer Größe von bis zu  $10^{11}$  Transformanten verwendet (Griffiths et al. 1994). Somit kann das Phage-Display eingesetzt werden, um funktionelle bindende Agenzien aus molekularen Bibliotheken direkt zu klonieren, und es kann auch dazu verwendet werden, die ursprünglich selektierten Klone weiter zu verbessern.

**[0008]** Die Zufallskombination der DNA verschiedener mutierter Klone ist eine wirksamere Art und Weise, den Sequenzraum zu durchsuchen. Das Konzept des DNA-Shuffling (Stemmer, 1994) setzt eine Zufallsfragmentierung der DNA und den Zusammenbau der Fragmente zu einer funktionellen codierenden Sequenz ein. Bei diesem Verfahren ist es möglich, chemisch synthetisierte DNA-Sequenzen einzuführen und auf diese Weise die Variation gezielt in definierten Stellen des Gens durchzuführen, dessen DNA-Sequenz bekannt ist (Cramer

et al. 1995). Theoretisch ist es auch möglich, ein Shuffling von DNA zwischen beliebigen Klonen durchzuführen. Wenn jedoch das resultierende, durch Shuffling zusammengesetzte Gen bezüglich der Expression und der Aktivität funktionell sein soll, dann müssen die Klone, die einem Shuffling unterzogen werden sollen, verwandt oder sogar identisch sein, abgesehen von wenigen Zufallsmutationen. Das DNA-Shuffling zwischen genetisch unterschiedlichen Klonen führt im Allgemeinen zu nicht-funktionellen Genen.

#### Zusammenfassung der Erfindung

**[0009]** Die vorliegende Erfindung stellt ein Verfahren zur Generierung einer Polynucleotidbibliothek, wie es in den begleitenden Ansprüchen definiert ist, bereit. Ebenfalls beschrieben, aber nicht beansprucht, wird hier ein Verfahren zur Gewinnung einer Polynucleotidsequenz, die für ein Protein mit gewünschten Eigenschaften codiert, das den Schritt des Einbauens wenigstens einer Variante eines Nucleotidabschnittes (einer Motivvariante) in definierte Nucleotidabschnitte (Gerüstsequenz), die sich von einer Ausgangs-Polynucleotidsequenz ableiten, umfasst. Die neue zusammengesetzte Polynucleotidsequenz kann dann exprimiert werden, und das resultierende Protein kann zur Bestimmung seiner Eigenschaften einem Screening unterzogen werden.

**[0010]** Das vorliegende Verfahren zur Generierung einer Polynucleotidbibliothek ermöglicht es, Proteineigenschaften zu verändern, indem die für das Protein codierende Polynucleotidsequenz auf spezifische Weise modifiziert wird. Das kann entweder durch a) das Ersetzen einer bestimmten Region der Nucleotidsequenz durch eine andere Nucleotidsequenz oder durch b) das Mutieren der bestimmten Region, so dass die Nucleotidsequenz verändert wird, erreicht werden. Diese bestimmten Regionen (Motivvarianten) werden in Gerüst- oder Rahmenabschnitte (Gerüstsequenz) der ursprünglichen Polynucleotidsequenz (Ausgangs-Polynucleotidsequenz) inkorporiert, die, wenn sie neu zusammengefügt werden, eine Polynucleotidbibliothek generieren, die für ein Protein mit veränderten Eigenschaften codiert. Die Charakteristika des codierten Proteins werden als Ergebnis einer Veränderung der Aminosäuresequenz entsprechend den Veränderungen in der codierenden Polynucleotidsequenz verändert.

**[0011]** Statt eine Sequenz auf zufällige Weise zu modifizieren und sich dann auf ein extensives Screening des gewünschten codierten Proteins zu verlassen, haben es die Erfinder der vorliegenden Erfindung wünschenswert gefunden, ein Verfahren bereit zu stellen, das ausgewählte Abschnitte (Motivvarianten) eines Proteins modifiziert, während andere beibehalten werden.

**[0012]** Die Motivvarianten können Abschnitte einer Nucleotidsequenz sein, die für bestimmte Abschnitte eines Proteins codiert, z. B. funktionelle Abschnitte eines Proteins (beispielsweise Schleifen) oder CDR-Bereiche in einem Antikörper.

**[0013]** Bei der Gerüstsequenz kann es sich um Abschnitte einer Nucleotidsequenz handeln, die beibehalten werden sollte; z. B. können sie für weitere strukturelle Bereiche des Proteins codieren, beispielsweise für Framework-Bereiche eines Antikörpers.

**[0014]** Die Motivvarianten können Nucleotidabschnitte sein, die von der gleichen Polynucleotidsequenz abstammen wie die Gerüstsequenz, d. h. von der Ausgangs-Polynucleotidsequenz, die aber so mutiert wurden, dass die codierende Sequenz gegenüber derjenigen der Ausgangssequenz verändert wurde. Zum Beispiel kann die Ausgangs-Polynucleotidsequenz für einen Antikörper codieren. Die Nucleotidsequenzen, die für die CDR-Bereiche des Antikörpers (Motivvarianten) codieren, können aus der restlichen codierenden Sequenz des Ausgangs-Polynucleotids ausgewählt werden, mutiert werden und dann mit der Gerüstsequenz, die aus der verbleibenden codierenden Sequenz stammt, neu zusammengesetzt werden. Der exprimierte Antikörper unterscheidet sich vom Wildtyp-Antikörper, der von dem Ausgangs-Polynucleotid exprimiert wird, nur hinsichtlich der CDR-Bereiche.

**[0015]** Alternativ kann die Motivvariante von einer Polynucleotidsequenz stammen, die für ein Protein codiert, das in sequenzieller Verwandtschaft zu dem Protein steht, das durch die Ausgangs-Polynucleotidsequenz codiert wird. Zum Beispiel können die CDR-Bereiche des einen Antikörpers (Antikörper A) durch die CDR-Bereiche eines anderen Antikörpers (Antikörper B) ersetzt werden.

**[0016]** In allen Fällen kann das resultierende exprimierte Protein einem Screening bezüglich gewünschter Charakteristika unterzogen werden. Gewünschte Charakteristika können Veränderungen biologischer Eigenschaften des Proteins sein. Beispielsweise kann die Tertiärstruktur des Proteins verändert sein. Das kann seine Bindungseigenschaften beeinflussen, seine Fähigkeit, durch Zellen oder in Zellen hinein sekretiert zu werden, oder, im Falle von Enzymen, seine katalytischen Eigenschaften. Wenn das Protein ein Antikörper oder ein Teil

davon ist, kann es wünschenswert sein, seine Fähigkeit zur spezifischen Bindung an ein Antigen zu verändern oder seine Bindungseigenschaften gegenüber dem Ausgangs-Antikörper zu verbessern.

**[0017]** Beschrieben, aber nicht beansprucht, wird hier ein Verfahren zur Gewinnung eines Proteins mit gewünschten Eigenschaften durch das Inkorporieren von Varianten von Peptidabschnitten (Motivvarianten) in definierte Peptidbereiche (Gerüstsequenz), wobei das Verfahren die folgenden Schritte umfasst:

- a) Unterziehen einer Ausgangs-Polynucleotidsequenz, die für ein oder mehrere Proteinmotiv(e) codiert, einer Mutagenese zur Erzeugung einer Vielzahl unterschiedlich mutierter Derivate von ihr, oder Gewinnen eines Ausgangs-Polynucleotids, das für eine Vielzahl von Varianten von Peptidmotiven mit unbekannter Sequenz codiert,
- b) Bereitstellen einer Vielzahl von Oligonucleotidpaaren, wobei jedes Paar räumlich getrennt voneinander vorliegenden Abschnitten der Ausgangs-Polynucleotidsequenz entspricht, die eine dazwischen liegende Variante des Proteinmotivs begrenzen, und Verwenden eines jeden genannten Oligonucleotidpaares als Primer für die PCR zur Amplifizierung des dazwischen liegenden Motivs,
- c) Gewinnen einzelsträngiger Nucleotidsequenzen von den so isolierten amplifizierten Nucleotidsequenzen, und
- d) Zusammenfügen von Nucleotidsequenzen, die für ein Protein codieren, durch das Inkorporieren von Nucleotidsequenzen, die aus dem Schritt c) oben stammen, in eine Nucleotidsequenz, die für eine Gerüstsequenz codiert.

**[0018]** Das Verfahren kann ferner den Schritt des Expressierens des resultierenden Proteins, das durch die zusammengefügte Nucleotidsequenz codiert wird, und das Screenen bezüglich gewünschter Eigenschaften umfassen.

**[0019]** Vorzugsweise handelt sich bei der Ausgangs-Polynucleotidsequenz um DNA, von der DNA-Sequenzen stammen, die für die Motivvarianten und Gerüstsequenzen codieren.

**[0020]** Vorzugsweise sind die Oligonucleotidpaare einzelsträngige Oligonucleotidprimer. Einer aus dem genannten Paar kann mit einem Partner eines spezifischen Bindungspaares (member of a specific binding pair, MSBP) verknüpft sein. Der MSBP ist vorzugsweise Biotin, dessen spezifischer Bindungspartner z. B. Streptavidin sein könnte. Über den Einsatz des spezifischen Bindungspaares können die amplifizierten Nucleotidsequenzen isoliert werden.

**[0021]** Die Zufallsmutagenese kann mittels eines beliebigen herkömmlichen Verfahrens durchgeführt werden, aber ein geeignetes Verfahren ist eine fehlerhafte PCR.

**[0022]** Das in Frage kommende Protein könnte beispielsweise ein Antikörper oder ein Antikörper-Fragment sein, der bzw. das die gewünschten Eigenschaften zeigt. Beispiele für Antikörper-Fragmente, die imstande sind, ein Antigen oder einen anderen Bindungspartner zu binden, sind das Fab-Fragment, das aus den VL-, VH-, C1- und CH1-Domänen besteht; das Fd-Fragment, das aus den VH- und CH1-Domänen besteht; das Fv-Fragment, das aus den VL- und VH-Domänen eines einzelnen Arms eines Antikörpers besteht; das dAb-Fragment, das aus einer VH-Domäne besteht; isolierte CDR-Bereiche und F(ab')<sub>2</sub>-Fragmente, ein bivalentes Fragment, das zwei Fab-Fragmente enthält, die über eine Disulfidbrücke im Hinge-Bereich verknüpft sind. Einzelkettige Fv-Fragmente sind ebenfalls eingeschlossen.

**[0023]** Bei einem Ansatz könnten nach der Zufallsmutagenese der DNA, die für den Antikörper codiert, oder eines Teiles dieser DNA (z. B. desjenigen, der für die Fab-Bereiche oder die variablen Bereiche codiert) Oligonucleotidprimer synthetisiert werden, die Sequenzen entsprechen, die die CDRs begrenzen (die Motivvarianten), so dass die DNA, die für die CDRs codiert, amplifiziert wird, zusammen mit möglichen Mutationen, zu denen es in den CDRs gekommen sein könnte. Diese können in den Neuzusammenbau der für den Antikörper codierenden Sequenz inkorporiert werden, und zwar unter Verwendung der amplifizierten DNA-Sequenzen der CDR sowie der DNA-Sequenzen des nicht-mutierten Gerüst-Frameworks (FR), was zur Generierung einer Polynucleotidbibliothek führt, die für einen Antikörper codiert, der eine neuartige Kombination von CDRs aufweist und der potenziell veränderte Eigenschaften zeigt, bezüglich derer auf herkömmliche Weise selektiert oder ein Screening durchgeführt werden kann.

**[0024]** Bei einem anderen Ansatz könnten, statt die CDRs zu mutieren und sie wieder zu einem Antikörper zusammensetzen, der in enger Beziehung zum Ausgangsantikörper, aus dem die CDRs gewonnen wurden, steht, die CDRs von einem oder mehreren existierenden Antikörper(n) genommen werden, aber eine unbekannte Sequenz haben. Unter Verwendung von Oligonucleotidprimern, die Sequenzen repräsentieren, die die

verschiedenen CDRs begrenzen, können die einzelnen CDRs amplifiziert, isoliert und in ein vorher festgelegtes Gerüst eingebaut werden, wodurch eine Polynucleotidbibliothek generiert wird.

**[0025]** Natürlich könnten Kombinationen der obigen Ansätze eingesetzt werden, wobei CDRs aus einem oder mehreren Ausgangsantikörper(n) genommen werden und in einem Gerüst unter Generierung einer Polynucleotidbibliothek zusammengebaut werden, die für einen vollständig neuen, sekundären Antikörper codiert, und dann, nach dem Screenen für das Gewinnen eines sekundären Antikörpers mit den gewünschten Eigenschaften, könnte die DNA, die für ihn codiert, mutiert werden, die CDRs könnten amplifiziert und isoliert werden und dann mit nicht-mutierter DNA des sekundären Antikörpers, bei der es sich um Nicht-CDR-DNA (Gerüst-DNA) handelt, zusammengebaut werden, um Varianten des sekundären Antikörpers zu erzeugen, die in den CDRs mutiert sind, und die einem Screening bezüglich verbesserter Eigenschaften im Hinblick auf den ursprünglich ausgewählten sekundären Antikörper unterzogen werden können.

**[0026]** Die vorliegende Erfindung ermöglicht eine neue Art der Isolierung von DNA-Sequenzen aus genetisch verwandten Klonen, die funktionell unterschiedlich sind. Genetisch verwandte Klone sind solche, die zu einer bestimmten Strukturklasse gehören, z. B. den Immunglobulinen oder den Alpha/Beta-Fässern. Die Erfindung ermöglicht sowohl die Isolierung als auch die Zufallskombination funktioneller Sequenzen dieser verwandten Klone zu einer gegebenen DNA-Sequenz. Diese funktionellen Sequenzen können Schleifen sein, die eine Bindung oder Katalyse bewirken.

**[0027]** Das Konzept der Erfindung wird anhand von Antikörpermolekülen demonstriert, wobei CDR-Bereiche verschiedener Keimbahnsequenzen isoliert und auf zufällige Weise in einer definierten Framework-Sequenz kombiniert werden können. Die Erfindung erweitert die Komplexität der molekularen Bibliotheken, aus denen mittels der Technik des Phage-Display selektiert werden kann. Das Konzept der Erfindung wird auch anhand der Affinitätsmaturierung von Antikörperfragmenten über die Isolierung und Zufallskombination mutierter CDR-Bereiche demonstriert.

**[0028]** Es ist nicht möglich, das Konzept des DNA-Shuffling (Stemmer 1994) für die Isolierung spezifischer Sequenzen und deren Zufallskombination zu einer gegebenen Gen-Sequenz einzusetzen, da es nicht möglich ist, einzelne in vivo erzeugte DNA-Bereiche mittels des DNA-Shuffling zu amplifizieren. Die Kombination vollständiger Gen-Sequenzen ist möglich, aber die hier definierten Bereiche können keinem Shuffling unterzogen werden. Statt dessen wird die gesamte DNA einem Shuffling unterzogen. Somit können DNA-Sequenzen von genetisch verwandten Klonen, die funktionell unterschiedlich sind, z. B. Proteine, die zu Strukturklassen wie Immunglobulinen oder Alpha/Beta-Fässern gehören, nicht auf eine solche Weise einem Shuffling unterzogen werden, dass spezifische Bereiche konstant gehalten und andere Bereiche einem Shuffling unterzogen werden.

**[0029]** Das durch die vorliegende Erfindung bereit gestellte System bietet ein einfaches Verfahren zur Generierung einer Bibliothek von Polynucleotiden, die zufällige Kombinationen funktioneller Bereiche von Proteinen (z. B. Schleifen) mit einem definierten (spezifisch ausgewählten) Gerüst darstellen, d. h. für das Shuffling von Schleifen zu einer gegebenen Tertiärstruktur eines Proteins, um neue Proteinfunktionen zu finden. Die Technologie des DNA-Shuffling führt außerdem Mutationen mit einer Häufigkeit von 0,7% ein (Stemmer, 1994). Somit ermöglicht die bekannte Technologie des DNA-Shuffling (Stemmer 1994) nicht das Shuffling nicht-mutierter Bereiche, da der Prozess selbst Mutationen in Zufallspositionen einführt, einschließlich der Gerüstregionen.

**[0030]** Im Gegensatz dazu erlaubt die Erfindung die Mutagenese definierter DNA-Sequenzen zusammen mit dem Shuffling und dem Zusammenbau dieser DNA-Stücke zu einer codierenden Sequenz, und sie ermöglicht die Kombination der Mutagenese definierter Bereiche und der anschließenden Selektion dieser Bereiche.

**[0031]** Die Erfindung ermöglicht es, dass unterschiedliche Bereiche der DNA aus unterschiedlichen Sequenzen (Klonen) einem Shuffling unterzogen und zufällig kombiniert werden, wodurch eine Polynucleotidbibliothek generiert wird. Das erhöht die genetische Variation, aus der funktionelle Antikörperfragmente selektiert werden, und erhöht somit die Wahrscheinlichkeit, Proteine mit den gewünschten Eigenschaften zu selektieren. Man kann sehen, dass durch das Zufalls-Shuffling von lediglich einhundert CDRs in jeder Position der VH und VL eines Fragmentes  $10^{12}$  Kombinationen erhalten werden können, wodurch die Variabilität, die sich normalerweise im Immunsystem findet, erhöht wird.

**[0032]** Die Erfindung stellt die Amplifizierung definierter Bereiche aus z. B. einer cDNA-Bibliothek unter Verwendung von zwei Primern bereit, von denen der eine biotinyliert ist. Durch den Einsatz der MSBP-Gruppe, z. B. Biotin, kann einzelsträngige DNA isoliert und im Prozess des Genzusammenbaus eingesetzt werden. Die

Erfinder der vorliegenden Erfindung haben das anhand der Amplifizierung verschiedener cDNA-Bereiche aus einer Bibliothek für Antikörpergene und die zufällige Kombination dieser CDR-Bereiche in einem vorgegebenen Framework-Bereich demonstriert. Somit können definierte DNA-Bereiche (Framework-Bereiche) mit Zufallsbereichen einer DNA (CDR-Bereichen) durchsetzt werden, die aus einem In-vivo-System stammen oder chemisch synthetisiert werden können.

**[0033]** Die vorliegende Erfindung stellt auch Polynucleotidsequenzen und die Proteine, für die sie codieren, die mittels der oben beschriebenen Verfahren erzeugt werden, bereit. Es werden auch Vektoren bereit gestellt, die die Polynucleotidsequenzen inkorporiert enthalten, sowie Wirtszellen, die mittels der Vektoren transformiert sind.

**[0034]** Die vorliegende Erfindung stellt auch eine Polynucleotidbibliothek bereit, die Polynucleotide umfasst, die mittels des oben beschriebenen Verfahrens erzeugt wurden, und die für ein Phage-Display eingesetzt werden können.

**[0035]** Es wurde nun Aspekte und Ausführungsformen der vorliegenden Erfindung anhand von Beispielen unter Bezugnahme auf die begleitenden Figuren veranschaulicht. Weitere Aspekte und Ausführungsformen werden Fachleuten auf diesem Gebiet klar sein. Alle in diesem Text erwähnten Dokumente werden hier mit der entsprechenden Quellenangabe aufgenommen.

#### Kurze Beschreibung der Zeichnungen

**[0036]** [Fig. 1](#) zeigt das Shuffling spezifischer DNA-Sequenzen zwischen verschiedenen Klonen, basierend auf dem Zusammenbau von Gensequenzen aus einem Satz überlappender Oligonucleotide, anhand eines PCR-Protokolls aus nur einem Schritt.

**[0037]** [Fig. 2](#) zeigt unterschiedliche Geschwindigkeitskonstanten der Dissoziation für verschiedene Klone mit CDRs, die einem Shuffling unterzogen wurden. Eine niedrige Säule steht für eine niedrige Dissoziationsgeschwindigkeit, eine hohe Säule steht für eine hohe Dissoziationsgeschwindigkeit. Der Klon 36 ist das ursprüngliche, nicht-mutierte Antikörperfragment.

**[0038]** [Fig. 3](#) zeigt die Ergebnisse einer Analyse eines mittels HPLC auf einer Superose-S-200-FPLC-Säule (Pharmacia) in PBS-Puffer affinitätsgereinigten scFv-Antikörperfragments. Peak 1 ist die monomere Form des Antikörperfragments, Peak 2 ist eine geringe Menge einer Verunreinigung, und Peak 3 ist NaN<sub>3</sub> (Natriumazid), das als Konservierungsmittel verwendet wurde.

**[0039]** [Fig. 4](#) zeigt eine schematische Darstellung der Amplifizierung definierter DNA-Sequenzen und deren Shuffling in ein Master-Framework. Es werden nur die CDR-Bereiche amplifiziert. [Fig. 4A](#): Zusammenbau von Genen für die VH-Domäne. Bei der Matrize handelt es sich um scFv-B11, das mit fehleranfälliger PCR mutiert wurde. Eine einzelne CDR wird mittels zweier Primer amplifiziert, die an die jeweilige CDR angrenzen, und einer dieser Primer ist am 5'-Ende biotinyliert. Die jeweilige CDR wird amplifiziert, und es wird doppelsträngige DNA (dsDNA) erzeugt, wobei die Mutationen auf die CDR fokussiert sind, da die beiden Amplifizierungsprimer keine Mutationen enthalten. Diese DNA wird in zwei einzelsträngige DNA-Moleküle aufgetrennt. Das Molekül ohne Biotin wird im Zusammenbau des Gens eingesetzt. Die Primer 725, 729, 730, 728 und 727 werden in einer DNA-Syntheseapparatur synthetisiert, und die Primer H2, H3 und H5 enthalten mutierte CDR und werden wie oben amplifiziert. [Fig. 4B](#): Zusammenbau von Genen für die VL-Domäne. Die CDRs werden auf die gleiche Weise wie in A amplifiziert. Die Primer 759, 738, 745, 744 und 880 werden in einer DNA-Syntheseapparatur synthetisiert, und die Primer L2, L3 und L5 enthalten mutierte CDR und werden wie oben amplifiziert.

**[0040]** [Fig. 5](#) zeigt die Ausrichtung der Peptidsequenzen für die Klone 3, 11 und 31 bezüglich des ursprünglichen, nicht-mutierten Antikörperfragments (wt). Die CDR-Bereiche sind markiert. Mutationen in den Klonen 3, 11 und 31 sind unterstrichen.

**[0041]** [Fig. 6](#) zeigt die Prinzipien für die Isolierung der einzelsträngigen DNA für das Shuffling definierter DNA-Bereiche.

**[0042]** [Fig. 7](#) zeigt die Länge der schweren CDR3-Kette aus verschiedenen Klonen. Diese CDR-Abschnitte wurden aus verschiedenen Keimbahnsequenzen amplifiziert und auf zufällige Weise in einen definierten Framework-Bereich (aus der DP-47-Sequenz) kloniert.

**[0043]** Fig. 8 ist eine schematische Darstellung der Amplifizierung definierter DNA-Sequenzen und deren Shuffling in ein Master-Framework. Alle in diesem Genzusammenbau eingesetzten Oligonucleotide werden mittels PCR amplifiziert, aber nur die CDR-Bereiche enthalten eine genetische Variation. **Fig. 8A:** Zusammenbau von Genen für die VH-Domäne. Die Matrize für die Amplifizierung des Framework-Bereichs ist scFv-B11, während die CDRs von cDNA amplifiziert werden, die aus peripheren Blutlymphozyten, aus Tonsillen und aus der Milz präpariert wurde. Das jeweilige DNA-Fragment wird unter Verwendung von zwei Primern amplifiziert, die an den Enden der Fragmente, die amplifiziert werden sollen, lokalisiert sind, und einer dieser Primer ist am 5'-Ende biotinyliert. Das jeweilige DNA-Fragment wird amplifiziert, und es wird doppelsträngige DNA (dsDNA) erzeugt. Diese DNA wird in zwei einzelsträngige DNA-Moleküle aufgetrennt. Das Molekül ohne Biotin wird im Genzusammenbau eingesetzt, d. h. die Primer H1, H4, H6, und diese Primer enthalten keine Variationen. Die Primer HCDR1, HCDR2 und HCDR3 enthalten unterschiedliche CDRs und werden unter Verwendung von zwei Primern amplifiziert, die an die jeweiligen CDRs angrenzen, und einer dieser Primer ist am 5'-Ende biotinyliert. Die jeweilige CDR wird amplifiziert, und es wird doppelsträngige DNA (dsDNA) erzeugt, wobei die Variation auf die CDR fokussiert ist, da die beiden Amplifizierungsprimer keine Mutationen enthalten. Diese DNA wird in zwei einzelsträngige DNA-Moleküle aufgetrennt und in einem Genzusammenbau der VH-Domäne im Bibliotheksformat eingesetzt, d. h. die Variation in den CDRs stammt aus unterschiedlichen Keimbahnsequenzen. Die Primer BT25 und BT26 werden in einer DNA-Syntheseapparatur synthetisiert. **Fig. 8B:** Zusammenbau von Genen für die VL-Domäne. Es handelt sich im Prinzip um das gleiche Verfahren wie in A. Die Primer L1, L4 und L6 werden mittels PCR amplifiziert und erzeugt und enthalten keine Variation. LCDR1, LCDR2 und LCDR3 enthalten jeweils eine unterschiedliche CDR. Die Primer BT7 und BT12 werden in einer DNA-Syntheseapparatur synthetisiert.

**[0044]** Fig. 9 zeigt die Variation in einer Bibliothek, die gemäß der Fig. 8 konstruiert wurde. Der scFv-Bereich der Klone der Bibliothek und das ursprüngliche scFv-B11, die an FITC (Fluoresceinisothiocyanat) binden, wurden mittels PCR synthetisiert. Die gereinigten PCR-Produkte wurden mit BstNI geschnitten und auf einem 2,5%igen Agarosegel aufgetrennt. Die Klone 1–15 finden sich in den Spuren 2–16, die Klone 16–29 finden sich in den Spuren 18–31. Das ursprüngliche scFv-B11 findet sich in der Spur 32. Die Analyse ergab, dass 28 Klone gemäß ihrem Restriktionsmuster und ihrer Fragmentgröße 13 verschiedenen Gruppen zugeordnet werden konnten. Acht Klone (1, 2, 8, 10, 12, 16, 26, 27) kamen nur einmal vor, zwei Klone (17, 24) erschienen ähnlich, eine Gruppe von Klonen (18, 23, 29) hatte 3 ähnliche Mitglieder, 2 Gruppen (5, 15, 14, 19 und 3, 4, 6, 11) hatten 4 Mitglieder und eine Gruppe (7, 9, 13, 20, 21, 22, 25) hatte 7 ähnliche Mitglieder. Dieses Experiment unterschätzt die Variation in der Bibliothek, da BstNI nur einen Teil der Sequenzvariabilität erkennt. Außerdem erlaubte die Auflösung des Gels den Nachweis kleiner Größenunterschiede nicht, und sie trennte keine Fragmente von 100 bp auf.

**[0045]** Fig. 9B zeigt Klone, die in dem Experiment, das beispielhaft in der Fig. 9A gezeigt wurde, ähnliche Restriktionsmuster hatten, nach einem Schneiden sowohl mit BstNI als auch mit BamHI und einer Auftrennung auf 3%igen Agarosegelen. Zur Erleichterung des Vergleiches wurden die im Experiment A beschriebenen Gruppen ähnlicher Klone auf den Gelen zusammengruppiert. Die Klone 8 und 28 aus dem Experiment A wurden aus Platzgründen weg gelassen.

Gel I) Spur 1–8: Standard, Klon 5, 15, 14, 19, 2, 27 bzw. ursprüngliches scFv-B11

Gel II) Spur 1–8: Standard, Klon 16, 17, 24, 18, 23, 29 bzw. 26

Gel III) Spur 1–8: Standard, Klon 7, 9, 13, 20, 21, 22 bzw. 25

Gel VI) Spur 1–8: Standard, Klon 3, 4, 6, 11, 1, 10 bzw. 12

**[0046]** Unter diesen verbesserten experimentellen Bedingungen zeigten im Wesentlichen alle Klone unterschiedliche Restriktionsmuster/Fragmentgrößen. Alle Klone unterschieden sich vom ursprünglichen scFv-B11-Gen (Spur 8, Gel 1). Außerdem zeigte sich, dass die Gruppen von Klonen, die in der Fig. 9A ähnlich erschienen, unterschiedlich waren, wie aus der Analyse in der Fig. 9B hervor geht. Siehe Klon 5, 15, 14, 19 (Spuren 2–5 im Gel I), Klon 17, 24 (Spuren 3–4 im Gel II), Klon 18, 23, 29 (Spuren 5–7 im Gel II), Klone 7, 9, 13, 20, 21, 22, 25 (Spuren 2–8 im Gel III) und Klone 3, 4, 6, 11 (Spuren 2–5 im Gel IV).

**[0047]** Als Schlussfolgerung lässt sich festhalten, dass diese Experimente nahe legen, dass die Bibliothek eine hohe Variabilität zeigt.

Detaillierte Beschreibung und beispielhafte Darstellung der Erfindung

**[0048]** Ein Aspekt des Verfahrens des DNA-Shuffling kann anhand der folgenden Schritte in der Fig. 1 veranschaulicht werden.

- [0049]** A: Ein Gen, das für ein interessantes Protein codiert, wird in überlappende Oligonucleotide unterteilt.
- [0050]** B: Die Oligonucleotide werden mittels PCR zu einer Gensequenz der vollen Länge zusammengefügt.
- [0051]** C: Die Gensequenz wird einer Mutagenese unterzogen, z. B. durch fehleranfällige PCR.
- [0052]** D: Es werden Oligonucleotidpaare synthetisiert, wobei jedes Paar einen Bereich abdeckt, der durch eines der Oligonucleotide des Schrittes A oben definiert wird, außer einem Bereich, der in der Mitte des Oligonucleotids aus dem Schritt A lokalisiert ist. Dieser nicht abgedeckte Bereich ist die DNA-Sequenz, die nach der PCR-Amplifizierung einem Shuffling unterzogen werden kann. Diese beiden synthetisierten Oligonucleotide können somit als Amplifizierungsprimer für die Amplifizierung der nicht abgedeckten Region verwendet werden.
- [0053]** E: Einer dieser Amplifizierungsprimer ist biotinyliert, und das doppelsträngige PCR-Produkt kann dann mittels gut bekannter Streptavidin-Systeme isoliert werden.
- [0054]** F: Aus den so isolierten amplifizierten Oligonucleotiden kann eine einzelsträngige DNA-Sequenz erhalten werden, die DNA aus dem nicht abgedeckten oben erwähnten Bereich enthält, die dann als Oligonucleotid in einem neuen Zusammenbau der Gensequenz, wie er im Schritt A beschrieben wurde, eingesetzt werden kann.
- [0055]** G: Wenn DNA-Sequenzen aus unterschiedlichen Klonen und aus unterschiedlichen Abschnitten der mutierten Gensequenz amplifiziert und einzelsträngig gemacht werden, dann werden sie beim PCR-Prozess des Genzusammenbaus auf zufällig Weise kombiniert. Diese Zufallskombination ist die Basis für die molekulare Evolution in vitro.

#### Beispiele

**[0056]** Die Urheber der vorliegenden Erfindung haben das Konzept des Shufflings definierter DNA unter unterschiedlichen experimentellen Bedingungen demonstriert. Erstens wird das Shuffling von in vitro mutierten CDR-Bereichen in ein Antikörperfragment zum Zwecke einer Affinitätsmaturierung (Beispiel 1 und 2) beispielhaft dargestellt, und zweitens wird das Shuffling von in vivo gebildeten CDRs zur Erzeugung einer hochvariablen Antikörperbibliothek (Beispiel 3 und 4) beispielhaft dargestellt.

#### 1. Affinitätsmaturierung

**[0057]** Es wurde ein Modellsystem auf der Basis des scFv-B11-Antikörperfragments, das an FITC bindet, entwickelt. Das vollständige Gen, das für dieses scFv codiert, wurde aus einem Satz von 12 Oligonucleotiden zusammengesetzt ([Fig. 4A](#) und [Fig. 4B](#)), die die bekannte DNA-Sequenz des scFv-B11 repräsentierten, und es konnte die funktionelle Bindung des Genprodukts an FITC verifiziert werden. Diese Gensequenz wurde dann mittels fehleranfälliger PCR mutagenisiert, und die DNA, die für die CDR-Bereiche codierte, wurde wie oben unter Verwendung der Amplifizierungsprimer, von denen einer biotinyliert ist amplifiziert. (Die CDR-Bereiche sind diejenigen Teile des Antikörpermoleküls, die in die Bindung des Antigens, in diesem Falle des FITC, verwickelt sind).

**[0058]** Es wurden alle sechs CDR-Bereiche amplifiziert, und es wurde ein neues Gen unter Verwendung von sechs Oligonucleotiden zusammengebaut, die aus dem ersten Zusammenbau von 12 Oligonucleotiden (siehe oben) ausgewählt worden waren (diese waren nicht mutagenisiert), sowie von sechs aus der Amplifizierung der mutagenisierten CDR-Bereiche. Die Selektion funktioneller Antikörperfragmente, die das FITC banden, wurde mittels des Phage-Display durchgeführt. 50% der Klone banden an FITC, wobei sie andere Dissoziationsgeschwindigkeiten zeigten als das ursprüngliche scFv-B11, wie mittels des BIAcore-Biosensors bestimmt wurde ([Fig. 2](#)). Das zeigt, dass die Klone bezüglich der Art und Weise, wie sie das FITC erkannten, verändert waren.

**[0059]** Von den 16 Klonen, für die im BIAcore-System gefunden wurde, dass sie an FITC binden ([Fig. 2](#)), wurden die Klone, 3, 11, 27 und 31 für eine genauere Analyse ausgewählt, da diese Klone die größeren Veränderungen bezüglich der Dissoziationsgeschwindigkeit zeigten. Diese Klone wurden exprimiert und auf einer Säule mit FITC-konjugiertem BSA affinitätsgereinigt und mit einem Puffer mit niedrigem pH eluiert. Die gereinigten scFv-Antikörperfragmente wurden weiter mittels HPLC unter Verwendung einer Superdex-200-FPLC-Säule von Pharmacia, die die Fähigkeit zur Trennung der monomeren und dimeren Form der Antikörper hatte, gerei-



nigt und analysiert. Bei allen Klonen dominierte die monomere Form (ein typisches Größenprofil ist in der [Fig. 3](#) gezeigt). Diese wurde dann gereinigt und in einer detaillierten Affinitätsanalyse unter Verwendung eines BIA-core-Biosensors eingesetzt (Tabelle 1).

Tabelle 1

## Affinitätsbestimmung für ausgewählte Klone

Klon	$k_{\text{ASS}} \text{ (M}^{-1}\text{s}^{-1}\text{)}$	$k_{\text{DISS}} \text{ (s}^{-1}\text{)}$	$K_{\text{A}} \text{ (M}^{-1}\text{)}$
Nr. 3	$2,0 \times 10^5$	$4,3 \times 10^{-3}$	$4,8 \times 10^7$
Nr. 11	$2,6 \times 10^5$	$3,3 \times 10^{-3}$	$7,8 \times 10^7$
Nr. 27	$5,0 \times 10^5$	$16,0 \times 10^{-3}$	$3,1 \times 10^7$
Nr. 31	$1,2 \times 10^5$	$5,4 \times 10^{-3}$	$2,1 \times 10^7$
(ursprüngliches FITC-B11)	$2,7 \times 10^5$	$9,7 \times 10^{-3}$	$2,8 \times 10^7$

**[0060]** Der Klon Nr. 11 zeigte eine Affinität, die 2,8-fach höher war als diejenige des ursprünglichen scFv-B11-Antikörperfragments. Diese Steigerung basiert auf einer niedrigeren Dissoziationsgeschwindigkeit der. Ein Klon (Nr. 27) zeigte eine 2-fache Erhöhung der Assoziationsgeschwindigkeit. Die Gesamtaffinität dieses Klons war jedoch aufgrund einer höheren Dissoziationsgeschwindigkeit derjenigen des ursprünglichen FITC-B11-Klons ähnlich. Die Verteilung der unterschiedlichen Assoziations- und Dissoziationsgeschwindigkeiten innerhalb der Klone wurde als Quelle für ein CDR-Reshuffling für eine weitere Verbesserung der Affinitäten in Betracht gezogen.

**[0061]** Drei Klone wurden sequenziert. In der VH-Region (d. h. der Hälfte des scFv-B11, die drei CDR-Bereiche trug) wurden alle Mutationen wie erwartet in den CDR-Bereichen gefunden, da diese die einzigen Bereiche waren, die unter Verwendung der Amplifizierungsprimer mutagenisiert und amplifiziert worden waren. Interessanterweise waren alle CDR-Bereiche verschieden und trugen unterschiedliche Mutationen ([Fig. 5](#)). Allerdings wurde im Falle des CDR-Bereichs 2 in allen 3 Klonen die gleiche Mutation (eine Substitution eines Tyrosins durch ein Histidin) gefunden (der Rest der CDR-Bereiche unterschied sich zwischen den Klonen).

**[0062]** Außerdem wurde gefunden, dass die Mutationshäufigkeiten zwischen 2% und 4% lagen, wie anhand der Basenveränderungen in der 90 bp langen Sequenz, die aus den drei CDR-Regionen bestand, bestimmt wurde. Das liegt über der Mutationshäufigkeit der fehlerhaften PCR und zeigt, dass eine Kombination einzelner CDR-Bereiche aus unterschiedlichen Klonen vorliegt.

## 2. Affinitätsmaturierung-Reshuffling

**[0063]** Für die Durchführung eines zweiten Shuffling (Reshuffling) wurden Klone, die bezüglich ihrer Bindungsaffinität für FITC ausgewählt worden waren, in einer weiteren Runde einer CDR-Amplifizierung und Bibliothekskonstruktion eingesetzt. Theoretisch enthält die Bibliothek mutierte, einem Shuffling unterzogene CDR-Bereiche, die bezüglich einer verbesserten Bindung an FITC ausgewählt wurden. Auf diese Weise konnten neue Kombinationen von CDR-Bereichen, die bezüglich ihrer Bindung verbessert waren, konstruiert werden, und die Bibliothek konnte einer Selektion bezüglich bindender Agenzien mit verbesserten Affinitäten unterzogen werden.

**[0064]** Der Pool aller Klone, die aus der Selektionsprozedur erhalten wurden (wie im Beispiel 1 detailliert beschrieben wurde), wurde als Matrize für die CDR-Amplifizierungen verwendet. Es wurde unter Verwendung der in der Tabelle 2 aufgeführten Primer durchgeführt eine Amplifizierung für jede CDR.

Tabelle 2

Sequenzen der beim CDR-Shuffling eingesetzten Primer  
B = Biotin-markierter 5'-Primer

**Primer für die CDR-Amplifizierung**

764 5' B-GTC CCT GAG ACT CTC CTG TGC AGC CTC TGG ATT CAC CTT T 3'  
875 5' TCC CTG GAG CCT GGC GGA CCC A 3'  
876 5' CGC CAG GCT CCA GGG AAG GGG CTG GAG TGG GTC TCA 3'  
765 5' B-GGA ATT GTC TCT GGA GAT GGT GAA 3'  
799 5' GAG CCG AGG ACA CGG CCG TGT ATT ACT GTG CAA GA 3'  
766 5' B-GCG CTG CTC ACG GTG ACC AGG GTA CCT TGG CCC CA 3'  
767 5' B-AGC GTC TGG GAC CCC CGG GCA GAG GGT CAC CAT CTC TTG T 3'  
800 5' GGG CCG TTC CTG GGA GCT GCT GGT ACC A 3'  
801 5' GCT CCC AGG AAC GGC CCC CAA ACT CCT CAT CTA T 3'  
768 5' B-GAC TTG GAG CCA GAG AAT CGG TCA GGG ACC CC 3'  
802 5' CTC CGG TCC GAG GAT GAG GCT GAT TAT TAC TGT 3'  
769 5' B-CGT CAG CTT GGT TCC TCC GCC GAA 3'

**Framework VH**

727 5' CCG CCG GAT CCA CCT CCG CCT GAA CCG CCT CCA CCG CTG CTC  
ACG GTG ACC A 3'  
728 5' GAC CGA TGG ACC TTT GGT ACC GGC GCT GCT CAC GGT GAC CA 3'  
729 5' GAG GTG CAG CTG TTG GAG TCT GGG GGA GGC TTG GTA CAG CCT  
GGG GGG TCC CTG AGA CTC TCC TGT 3'  
730 5' GGC CGT GTC CTC GGC TCT CAG GCT GTT CAT TTG CAG ATA CAG  
CGT GTT CTT GGA ATT GTC TCT GGA GAT GGT 3'

**Framework VL**

738 5' CAG TCT GTG CTG ACT CAG CCA CCC TCA GCG TCT GGG ACC CCC  
G 3'  
744 5' ACT AGT TGG ACT AGC CAC AGT CCG TGG TTG ACC TAG GAC CGT  
CAG CTT GGT TCC TCC GC 3'  
745 5' CTC ATC CTC GGA CCG GAG CCC ACT GAT GGC CAG GGA GGC TGA  
GGT GCC AGA CTT GGA GCC AGA GAA TCG 3'  
1129 5' CAG GCG GAG GTG GAT CCG GCG GTG GCG GAT CGC AGT CTG TGC  
TGA CTC AGC CAC CCT CAG CGT CTG GGA CCC CCG 3'

**Amplifizierungsprimer VH/VL-Zusammenbau**

1125 5' ACT CGC GGC CCA ACC GGC CAT GGC CGA GGT GCA GCT GTT GGA  
G 3'  
1126 5' CAA CTT TCT TGT CGA CTT TAT CAT CAT CAT CTT TAT AAT CAC  
CTA GGA CCG TCA GCT TGG T 3'

**[0065]** Die Amplifizierung wurde unter Einsatz der folgenden Parameter durchgeführt: 100 ng Matrize ( $1,6 \times 10^8$  CFU, Bakterien 6 Stunden gewachsen), 60 pmol eines jeden Primers, 5 Einheiten PFU-Polymerase (Stratagene),  $1 \times$  PFU-Puffer, 500  $\mu$ M dNTPs, Reaktionsvolumen 100  $\mu$ l, Vorerhitzen bei 96°C für 10 Minuten, 96°C für 1 Minute, 68°C für 1 Minute, 72°C für 1 Minute und 25 Zyklen, 72°C für 10 Minuten.

**[0066]** Dieses Verfahren war im Wesentlichen das gleiche wie das der CDR-Amplifizierung im Beispiel 1. Die amplifizierten CDRs wurde für einen Einbau in die für VH und VL codierende Sequenz gemäß der [Fig. 1](#), [Fig. 4A](#) und [Fig. 4B](#) und der Tabelle 3 eingesetzt.

Tabelle 3

PCR-Parameter für den Zusammenbau der VH- und VL-Gensequenzen beim CDR- Shuffling

VL	VH	
Primer 759	Primer 725	30 pmol
Primer 738	Primer 729	0,6 pmol
Primer L2	Primer H2	0,6 pmol
Primer L3	Primer H3	0,6 pmol
Primer 745	Primer 730	0,6 pmol
Primer L5	Primer H5	0,6 pmol
Primer 744	Primer 728	0,6 pmol
Primer 880	Primer 727	30 pmol
Taq	Taq	10 Einheiten
dNTPs	dNTPs	200 µM
1 × Taq-Puffer	1 × Taq-Puffer	ad 100 µl

Vorerhitzen auf 95°C für 10 Minuten, 20 Zyklen, 95°C für 1 Minute, 68°C für 1 Minute, 72°C für 1 Minute, 72°C für 10 Minuten.

**[0067]** Die VH und VL wurden dann zu einer für scFv codierenden Sequenz gemäß Standardverfahren (Griffiths et al. 1994) zusammengebaut. Die resultierende Bibliothek wurde einem Panning-Schritt unterzogen, um bindende Agenzien mit verbesserten Affinitäten für FITC zu selektieren. Die Selektionsprozedur für die Reshuffling-Bibliothek war im Wesentlichen die gleiche wie diejenige der anfangs dem Shuffling unterzogenen Bibliothek. Die Gesamtzahl der nach der Selektion erhaltenen Klone lag bei 510. Sechs Klone (B) wurden aus diesem neuen Pool ausgewählt und getestet und mit sechs Klonen (A) aus dem ersten Pool, die aus der dem Shuffling unterzogenen Bibliothek stammten, verglichen (Tabelle 4).

Tabelle 4

Dissoziationsgeschwindigkeiten einzelner Klone, die aus der dem Shuffling unterzogenen Bibliothek (Klone A) und aus der dem Reshuffling unterzogenen Bibliothek (Klone B) ausgewählt waren

Klon	$K_{\text{DISS}} (\text{s}^{-1} \times 10^{-3})$
scFv-B11 (ursprüngliches)	12,9
1A	6,3
12A	5,7
13A	9,0
14A	9,7
16A	1,8
17A	7,9
22B	0,2
31B	0,3
32B	9,8
33B	6,8
34B	7,3
35B	8,7

**[0068]** Zwei Klone aus den Reshuffling-Experimenten (22B und 31B) wiesen wesentlich niedrigere Dissoziationsgeschwindigkeiten auf, was anzeigte, dass der Reshuffling-Prozess bindende Agenzien mit verbesserten Affinitäten lieferte.

### 3. Klonierung und Shuffling definierter DNA-Abschnitte

**[0069]** In unserem System ist es möglich, definierte Abschnitte aus einer cDNA-Bibliothek unter Verwendung von zwei Primern, von denen einer biotinyliert ist, zu amplifizieren. Mit Hilfe der Biotingruppe kann einzelsträngige DNA isoliert und im Prozess des Genzusammenbaus eingesetzt werden ([Fig. 6](#)). Wir haben das mit der Amplifizierung diverser CDR-Bereiche aus einer Bibliothek aus Antikörpergenen und der Zufallskombination dieser CDR-Bereiche in einem vorgegebenen Framework-Bereich gezeigt. Somit können definierte DNA-Abschnitte (Framework-Bereiche) mit zufälligen DNA-Abschnitten (CDR-Bereichen) durchsetzt werden, die einen In-vivo-Ursprung haben (Tabelle 5). Der CDR3-Bereich variiert bezüglich der Größe ([Fig. 7](#)). Alternativ könnten diese Bereiche chemisch synthetisiert werden.

Tabelle 5

Kombination von CDR-Bereichen aus unterschiedlichen Keimbahnsequenzen, wobei die CDR-Bereiche in das DP-47-Framework, das für die variable schwere Domäne codiert, transplantiert wurden. Für CDR1 und CDR2 ist der vorgeschlagene Keimbahnursprung angegeben. Für CDR3 ist die Zahl der Reste im CDR-Bereich angegeben. n. b. = nicht bestimmt

Klon	CDR1	CDR2	CDR3
1	DP-35	DP-42	12
2	DP-49	DP-53	13
3	n. b.	DP-51	11
4	DP-32	DP-47	10
5	DP-41	DP-47	8
6	DP-32	DP-77	9
7	DP-31	DP-47	7
8	DP-49	DP-35	5
9	DP-49	DP-35	n. b.
10	DP-48	DP-48	n. b.
11	DP-51	DP-47	10
12	DP-34	DP-31	n. b.
13	DP-85	DP-53	4
14	DP-31	DP-77	10
15	DP-34	DP-53	4

#### 4. Konstruktion der Bibliothek

**[0070]** Es wurde eine Genbibliothek konstruiert, die für scFv-Antikörperfragmente codierte. Die für diese Bibliothek eingesetzte Strategie basiert auf dem Einbau eines Sets von Oligonucleotiden in eine Sequenz, die für VH- und VL-Antikörperdomänen codiert ([Fig. 8A](#), [Fig. 8B](#)). Native, in vivo gebildete CDR-Bereiche können einem Shuffling unterzogen und in ein vorgegebenes Master-Framework eingebaut werden. In diesem Beispiel haben wir dieses Konzept weiter entwickelt und sowohl für VH als auch für VL codierende Gensequenzen mit nativen CDR-Bereichen in ein vorgegebenes Master-Framework eingebaut. Somit wurden alle sechs CDR-Positionen einem Shuffling unterzogen. Die Ausgangsmatrize für die CDR-Amplifizierung war cDNA aus B-Zellen des peripheren Blutes, aus der Milz, den Tonsillen und den Lymphknoten. Oligonucleotide, die für die Framework-Bereiche codieren, wurden auch unter Einsatz der Strategie mit zwei flankierenden Primern amplifiziert, von denen einer biotinyliert ist (Primer L1, H1, L4, H4, L6, H6). Die verwendeten Primer werden in der Tabelle 6 und in der [Fig. 8A](#) und [Fig. 8B](#) beschrieben.

Tabelle 6

Sequenzen der bei der Konstruktion der Bibliothek eingesetzten Primer  
B = Biotin-markierter 5'-Primer

#### Zusammenbau von VH und VL

BT1. 5' ACA GTC ATA ATG AAA TAG CTA TTG C 3'  
 BT2. 5' B-GC ACA GGA GAG TCT CA 3'  
 BT3. 5' B-CA CCA TCT CCA GAG ACA ATT CC 3'  
 BT4. 5' GGC CGT GTC CTC GGC TCT 3'  
 BT5. 5' B-TG GTC ACC GTG AGC AGC 3'  
 BT6. 5' CCG CCG GAT CCA CCT 3'  
 BT7. 5' CAG GCG GAG GTG GAT CCG GC 3'

BT8. 5' B-CG GGG GTC CCA GAC GCT 3'  
 BT9. 5' B-CG ATT CTC TGG CTC CAA GT 3'  
 BT10. 5' CTC ATC CTC GGA CCG GA 3'  
 BT11. 5' B-TC GGC GGA GGA ACC AAG CT 3'  
 BT12. 5' TGG CCT TGA TAT TCA CAA ACG AAT 3'

#### Amplifizierung von In-vivo-CDRs

BT13. 5' B-TC CCT GAG ACT CTC CTG TGC AGC CTC TGG ATT CAC CTT 3'  
 BT14. 5' TTC CCT GGA GCC TGG CGG ACC CA 3'  
 BT15. 5' B-GG AAT TGT CTC TGG AGA TGG TGA A 3'  
 BT16. 5' GTC CGC CAG GCT CCA 3'  
 BT17. 5' B-CG CTG CTC ACG GTG ACC AGT GTA CCT TGG CCC CA 3'  
 BT18. 5' AGA GCC GAG GAC ACG GCC GTG TAT TAC TGT 3'  
 BT19. 5' B-AG CGT CTG GGA CCC CCG GGC AGA GGG TCA CCA TCT CTT 3'  
 BT20. 5' GGG CCG TTC CTG GGA GCT GCT GAT ACC A 3'  
 BT21. 5' GCT CCC AGG AAC GGC CCC CAA ACT CCT CAT CTA T 3'  
 BT22. 5' B-GA CTT GGA GCC AGA GAA TCG GTC AGG GAC CCC 3'  
 BT23. 5' B-GT CAG CTT GGT TCC TCC GCC GAA 3'  
 BT24. 5' CTC CGG TCC GAG GAT GAG GCT GAT TAT TAC T 3'

#### Zusammenbau von VH und VL

BT25. 5' B-TA CCT ATT GCC TAC GGC AGC CGC TGG ATT GTT ATT ACT CGC GGC  
 CCA GCC GGC CAT GGC CGA 3'  
 BT26. 5' CCG CCG GAT CCA CCT CCG CCT GAA CCG CCT CCA CCG CTG CTC ACG  
 GTG ACC A 3'

#### Amplifizierungsprimer für 2. Zusammenbau

BT27. 5' B-TGG CCT TGA TAT TCA CAA ACG AAT 3'  
 BT28. 5' B-ACG GCA GCC GCT GGA TTG 3'

**[0071]** Die PCR-Parameter für die Amplifizierung der CDR und der Framework-Region waren im Wesentlichen die gleichen, wie sie im Beispiel 2 beschrieben wurden. Die PCR-Parameter für den Zusammenbau der für VH und VL codierenden Gene werden in der Tabelle 7 beschrieben.

Tabelle 7

PCR-Parameter für den Zusammenbau der VH- und VL-Gensequenzen für die Konstruktion der Bibliothek

VH	VL	
Primer BT25	Primer BT7	30 pmol
Primer H1	Primer L1	0,6 pmol
Primer HCDR1	Primer LCDR1	0,6 pmol
Primer HCDR2	Primer LCDR2	0,6 pmol
Primer H4	Primer L4	0,6 pmol
Primer HCDR5	Primer LCDR3	0,6 pmol
Primer H6	Primer L6	0,6 pmol
Primer BT26	Primer BT12	30 pmol
Taq	Taq	10 Einheiten
dNTPs	dNTPs	200 µM
1 × Taq-Puffer	1 × Taq-Puffer	ad 100 µl

Vorerhitzen auf 95°C für 10 Minuten, 20 Zyklen: 95°C für 1 Minute, 68°C für 1 Minute, 72°C für 1 Minute, 72°C für 10 Minuten.

**[0072]** Die zusammengebauten VH- und VL-Gensequenzen wurden mittels Standardprotokollen in eine für scFv codierende Sequenz eingebaut (Griffiths et al. 1994). Es wurde eine Bibliothek von  $1,1 \times 10^9$  Mitgliedern konstruiert, und von den 40 getesteten Klonen enthielten alle 40 ein Insert der richtigen Größe, wie mittels PCR und Agarose-Gelelektrophorese bestimmt wurde. Um die Variabilität in dieser Bibliothek zu testen, wurden PCR-amplifizierte und gereinigte Inserts einer Spaltung mit BsTN1 und BamH1 unterzogen. Die Klone zeigten unterschiedliche Restriktionsmuster, wie anhand einer Agarose-Gelelektrophorese und eines Vergleichs mit der Kontrolle scFv-B11 bestimmt wurde (**Fig. 9**).

**[0073]** Um die Häufigkeit der Klone abzuschätzen, die zur Expression von scFv-Antikörperfragmenten fähig waren, wurden Klone aus der Bibliothek, die die FLAG-Sequenz (Hopp et al. 1989) enthielten, sowie Kontrollbakterien mit und ohne FLAG-Sequenz in niedriger Dichte auf Luria-Broth-Platten ausplattiert, die 100 µg/ml Ampicillin, 25 µg/ml Tetracyclin und 1% Glucose enthielten. Man ließ die Platten über Nacht bei 37°C wachsen, und dann wurden sie mittels Standardverfahren auf Nitrocellulose-Filter übertragen (Sambrook et al. 1989). Zur Induktion der scFv-Gene in den Bakterien wurden die Filter 4 Stunden auf Platten inkubiert, die 0,5 mM Isopropyl-thio-β-D-galactosid (IPTG), aber keine Glucose enthielten. Die Bakterien wurden dann durch eine Behandlung mit Lysozym/Chloroform lysiert, die Filter wurden gewaschen und mit einem Anti-FLAG-M2-Antikörper (Kodak) inkubiert, gefolgt von einem sekundären Peroxidase-konjugierten Anti-Maus-Antikörper (P260 Dakopatts), und der Nachweis erfolgte mit DAB (3,3'-Diaminobenzidintetrahydrochlorid, Sigma)(Tabelle 8).

Tabelle 8

## Häufigkeit intakter Antikörpergene in der Bibliothek

Bibliotheks-Pool	Getestete Klone	FLAG-positive Klone	Prozent positive Klone
A	145	88	60
B	77	52	67
C	158	105	66
D	68	48	70
alle Bibliotheks-Pools	448	293	65,4
Positivkontrolle pFAB5cHis scFvB11	64	64	100
Negativkontrolle pFAB5cHis	30	0	0

**[0074]** Der Anti-FLAG-Antikörper weist eine FLAG-Sequenz nach, die stromabwärts des scFv-Gens in den Bibliothekskonstrukten sowie im Kontrollvektor pFAB5cHis scFvB11 enthalten ist, aber nicht im ursprünglichen Vektor pFAB5cHis. Klone, an die der Anti-FLAG-Antikörper bindet, enthalten somit einen intakten offenen Leserahmen des scFv-Gens.

## Literaturstellen

- Barbas, CF et al.: Proc Natl Acad Sci USA, 88: 7978–82 (1991)  
 Barbas, CF et al.: Proc Natl Acad Sci USA, 89: 4457–61 (1992)  
 Cramer, A et al.: Biotechniques, 18: 194–196 (1995)  
 Griffiths, AD et al.: EMBO J, 13: 3245–3260 (1994)  
 Hoogenboom, HR et al.: J Mol Biol, 227: 381–8 (1992)  
 Hopp, TP et al.: Biotechniques 7: 580–589 (1989)  
 Huse, WD et al.: Science, 246: 1275–81 (1989)  
 Larrick, JW et al.: Biochem Biophys Res Commun, 160: 1250–6 (1989)  
 Leung, DW et al.: Technique, 1: 11–15 (1989)  
 Marks, JD et al.: Biotechnology, 10: 779–83 (1992)  
 McCafferty, J et al.: Nature, 348: 552–4 (1990)  
 Moore, JC et al.: Nature Biotechnology, 14: 458–467 (1996)  
 Parmley, SF et al.: Gene, 73: 305–318 (1988)  
 Roberts, S et al.: Nature, 328: 731–4 (1987)  
 Sambrook, J et al.: Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory Press 1989  
 Stemmer, WP: Nature, 370: 389–391 (1994)  
 Yang, WP et al.: J Mol Biol, 254: 392–403 (1995)

## Patentansprüche

1. Verfahren zur Erzeugung einer Polynucleotid-Bibliothek, das umfasst die Schritte des Gewinnens eines Ausgangspolynucleotids, das für eine oder mehrere Variante(n) von Proteinmotiven codiert;

- Bereitstellen einer Vielzahl von Oligonucleotidpaaren, wobei jedes Paar räumlich getrennt voneinander vorliegenden Abschnitten der Ausgangspolynucleotidsequenz entspricht, die eine dazwischen liegende Variante eines Proteinmotivs begrenzen, und Verwenden eines jeden genannten Oligonucleotidpaares als Primer für die PCR zur Amplifizierung des dazwischen liegenden Motivs;
- Gewinnen einzelsträngiger Nucleotidsequenzen von den so isolierten amplifizierten Nucleotidsequenzen;
- Zusammenfügen von Polynucleotidsequenzen durch das Inkorporieren von Nucleotidsequenzen, die aus dem Schritt b) oben stammen, in Nucleotidsequenzen, die für eine Gerüstsequenz codieren; und
- Einfügen der genannten Polynucleotidsequenzen in geeignete Vektoren.

2. Verfahren gemäß Anspruch 1, das außerdem die Schritte der Expression der genannten Polynucleotidsequenzen zur Gewinnung einer entsprechenden Bibliothek von Polypeptidsequenzen umfasst.



3. Verfahren gemäß Anspruch 2, das außerdem das Screening der Bibliothek bezüglich eines Polypeptids mit den gewünschten Eigenschaften umfasst.

4. Verfahren, das, nach der Erzeugung einer Polynucleotid-Bibliothek gemäß Anspruch 1, das Selektieren und Exprimieren der genannten Polynucleotidsequenzen zur Gewinnung von Polypeptiden mit den gewünschten Eigenschaften umfasst.

5. Verfahren gemäß einem beliebigen der Ansprüche 1 bis 4, wobei die zusammengefügte Polynucleotidsequenz für einen Antikörper oder einen Teil von diesem codiert.

6. Verfahren gemäß Anspruch 5, wobei die zusammengefügte Polynucleotidsequenz für einen Teil eines Antikörpers codiert, und wobei dieser Teil ein Fab-, Fd-, Fv-, dAb-, F(ab')<sub>2</sub>- oder scFv-Fragment oder eine isolierte CDR ist.

7. Verfahren gemäß Anspruch 3 oder 4, wobei ein Polypeptid mit den gewünschten Eigenschaften ein Antikörper oder ein Teil von diesem ist.

8. Verfahren gemäß Anspruch 7, wobei ein Polypeptid mit den gewünschten Eigenschaften ein Teil eines Antikörpers ist, und wobei dieser Teil ein Fab-, Fd-, Fv-, dAb-, F(ab')<sub>2</sub>- oder scFv-Fragment oder eine isolierte CDR ist.

Es folgen 10 Blatt Zeichnungen

Fig.1.

Shuffling definierter DNA-Abschnitte

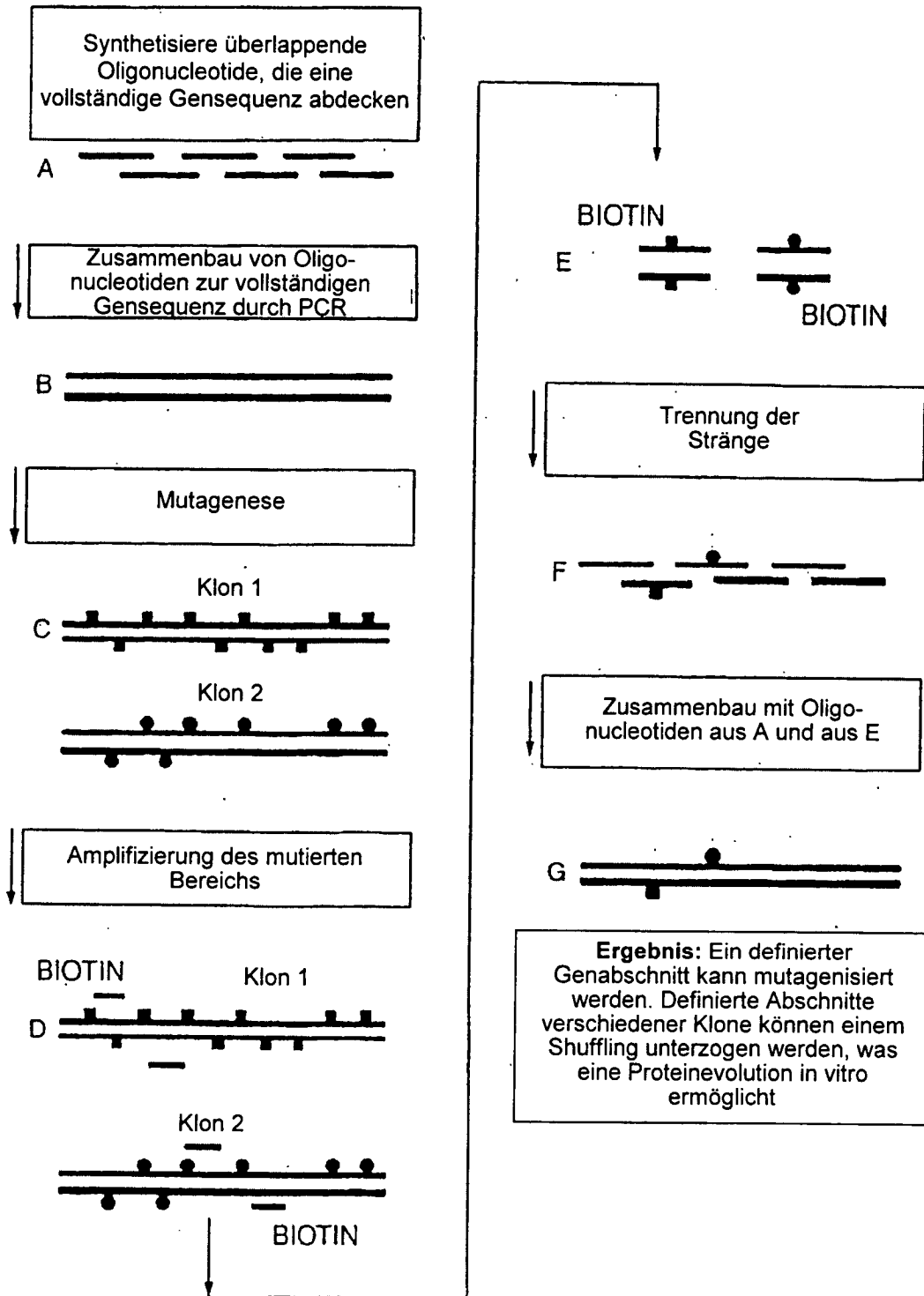


Fig.2.

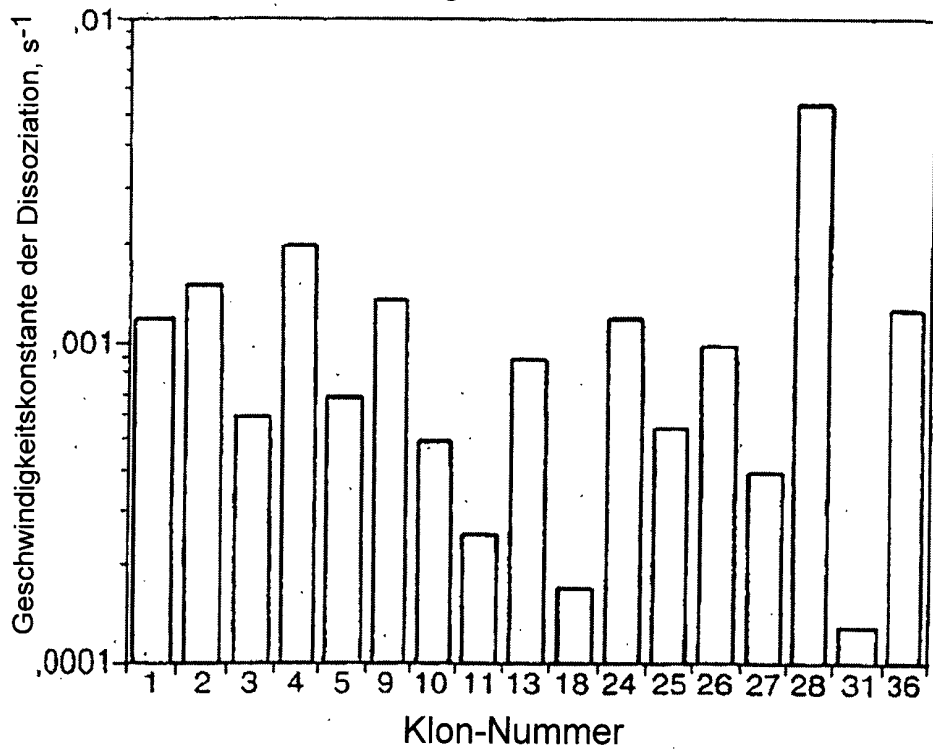


Fig.3.

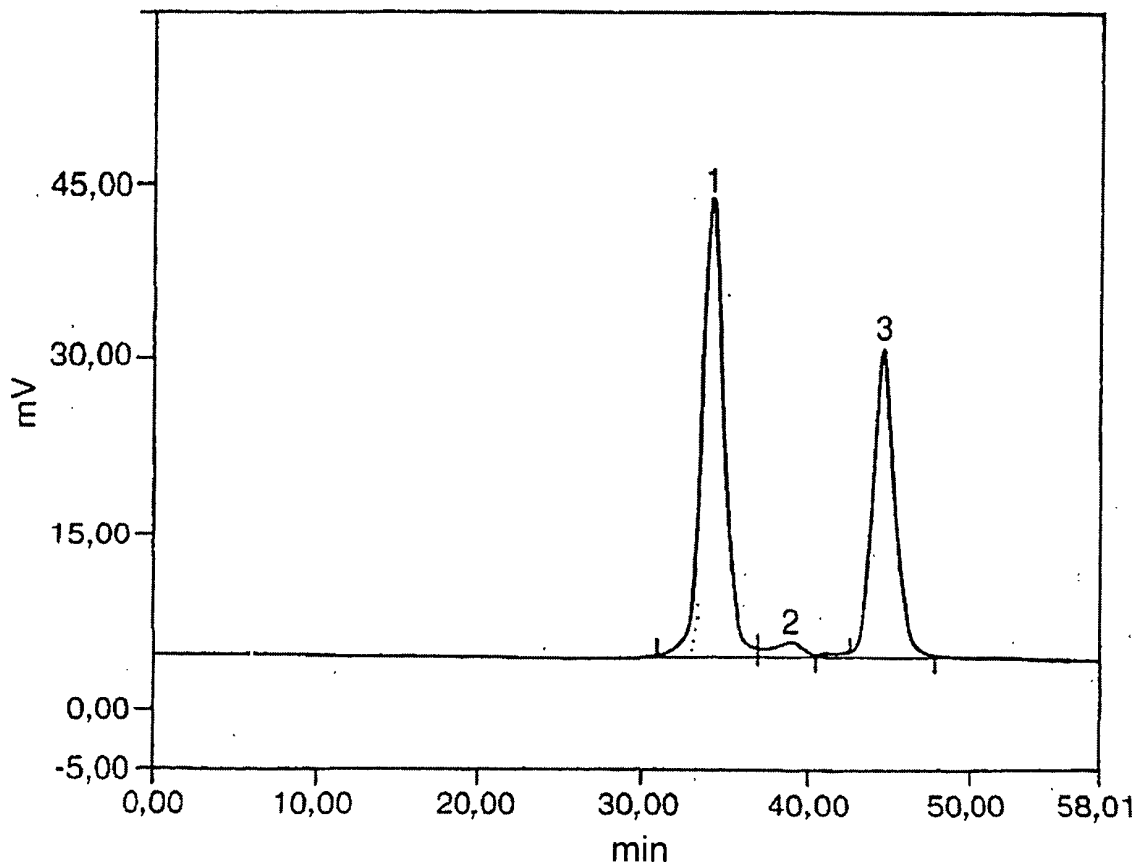
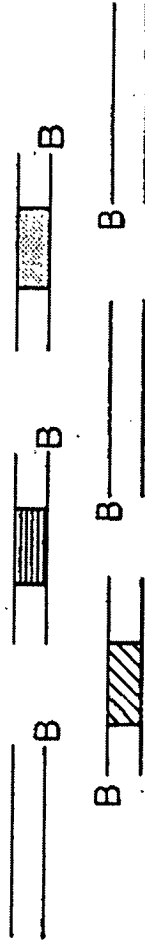
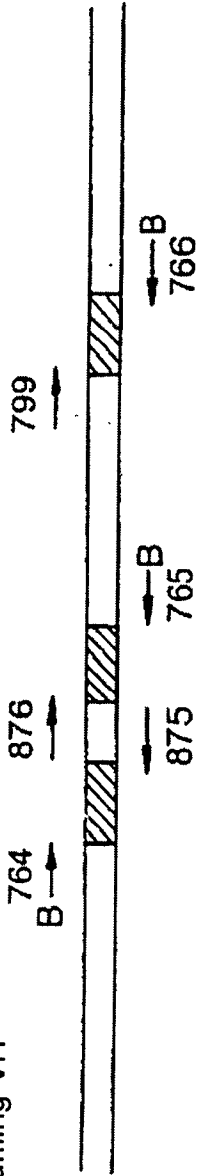


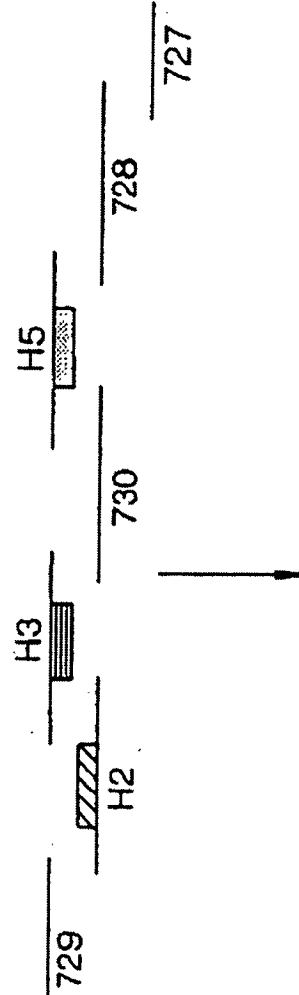
Fig.4A.

CDR-Shuffling VH



725

Zusammenbau 1



CDR-Shuffling VL

Fig. 4B.

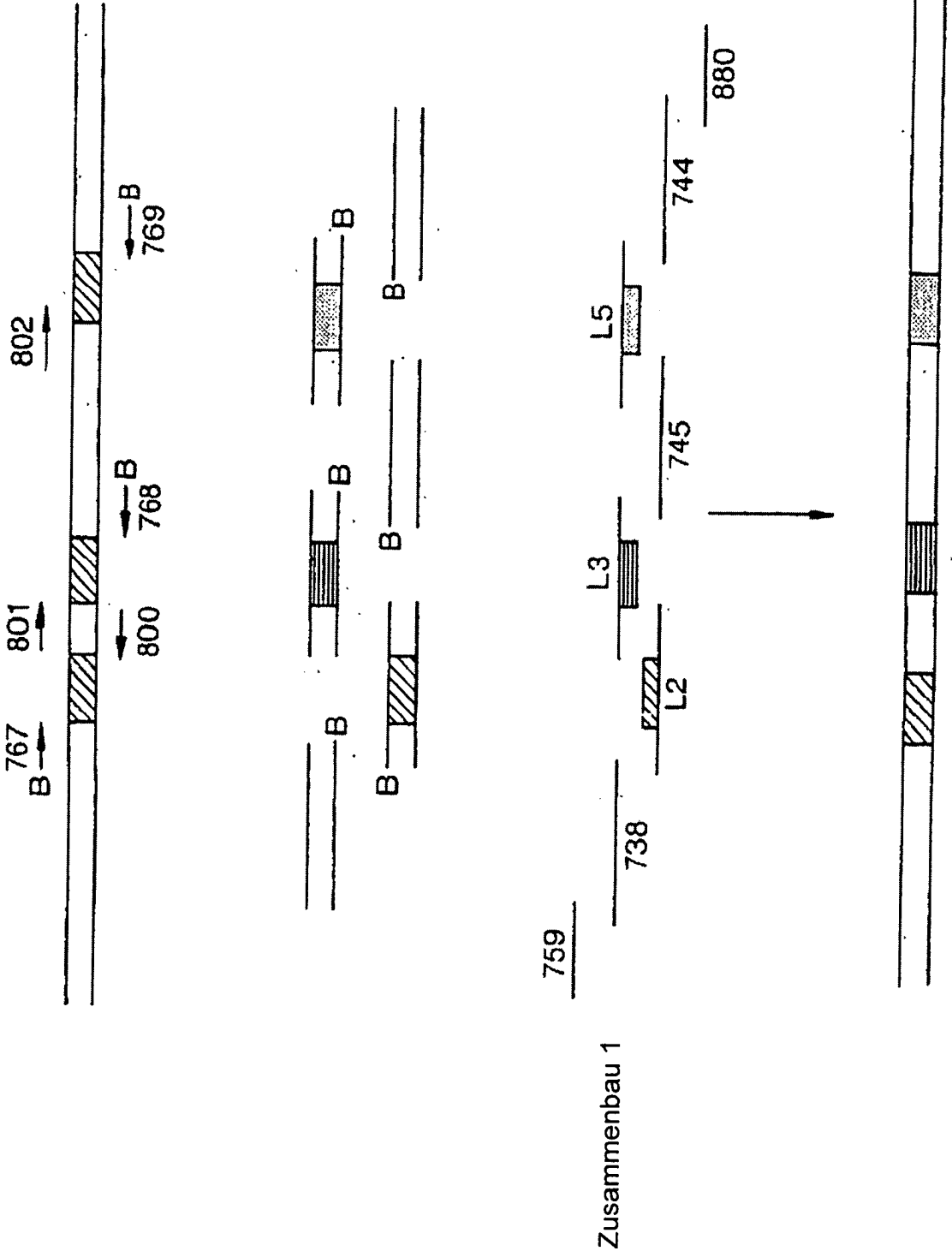


Fig.5.

			CDR1
KLON 31	LAAQPAMAEVQLLESGGGLVQPGGSLRLSCAASGFTF	SR	YAMSWVRQAPG
KLON 3	LAAQPAMAEVQLLESGGGLVQPGGSLRLSCAASGFTF	SG	YAMSWVRQAPG
KLON 11	LAAQPAMAEVQLLESGGGLVQPGGSLRLSCAASGFTF	SS	YAMSWVRQAPG
URSPRUNG	LAAQPAMAEVQLLESGGGLVQPGGSLRLSCAASGFTF	SS	YAMSWVRQAPG
		CDR2	
KLON 31	KGLEWVSAISGSGGTHYADSVKGRFTISRDN	SKNTLYLQMN	SLRAEDTA
KLON 3	KGLEWVSAISGSGGTHYADSVKGRFTISRDN	SKNTLYLQMN	SLRAEDTA
KLON 11	KGLEWVSAISGSGGTHYADSVKGRFTISRDN	SKNTLYLQMN	SLRAEDTA
URSPRUNG	KGLEWVSAISGSGGTYADSVKGRFTISRDN	SKNTLYLQMN	SLRAEDTA
		CDR3	
KLON 31	VYYCARI	GQF	WGQGLVTVSSGGSGGSQ
KLON 3	VYYCARI	GRF	WGQGLVTVSSGGSGGSQ
KLON 11	VYYCARI	GQF	WGQGLVTVSSGGSGGSQ
URSPRUNG	VYYCARI	GQF	WGQGLVTVSSGGSGGSQ

Fig.6.

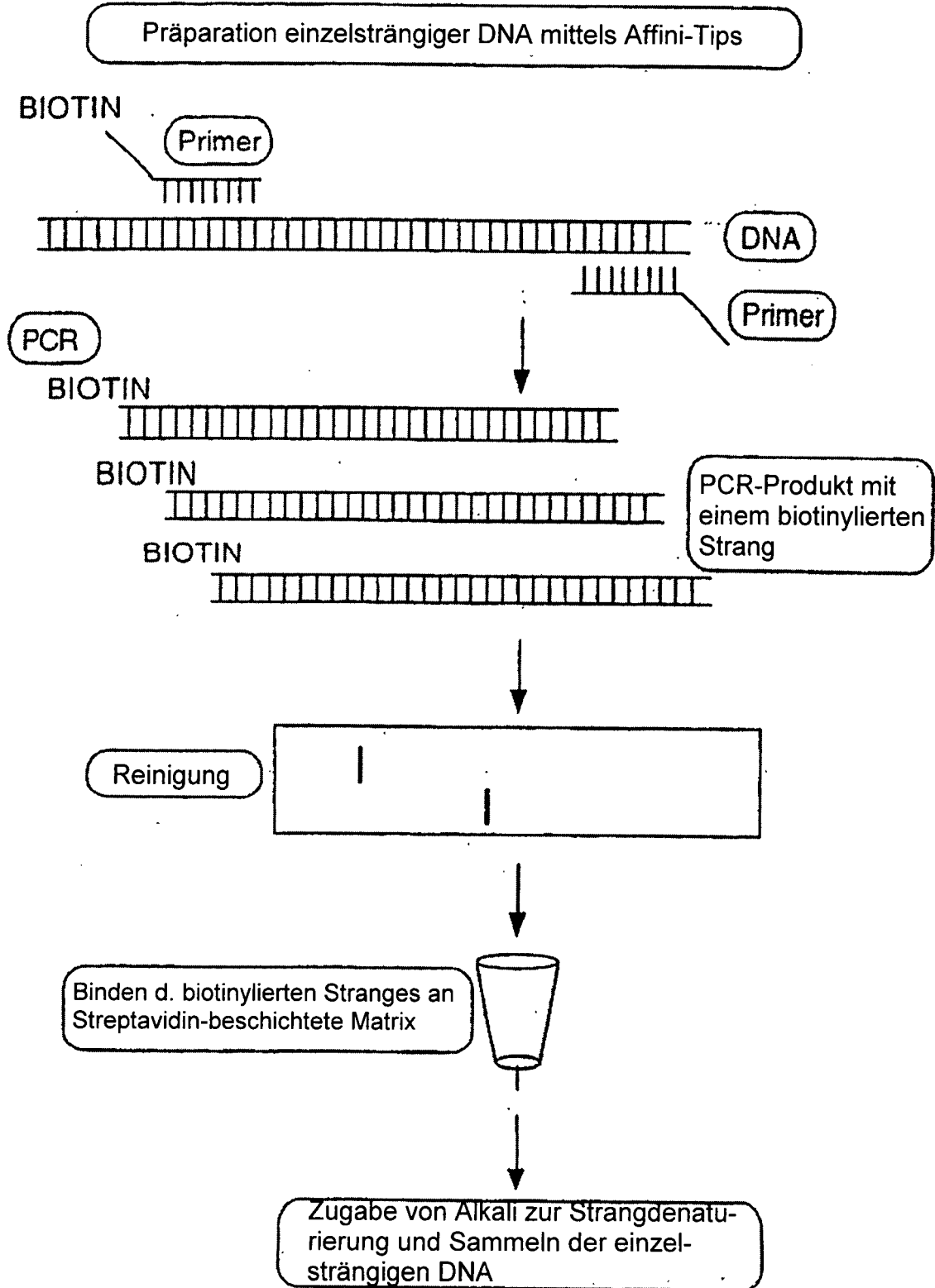
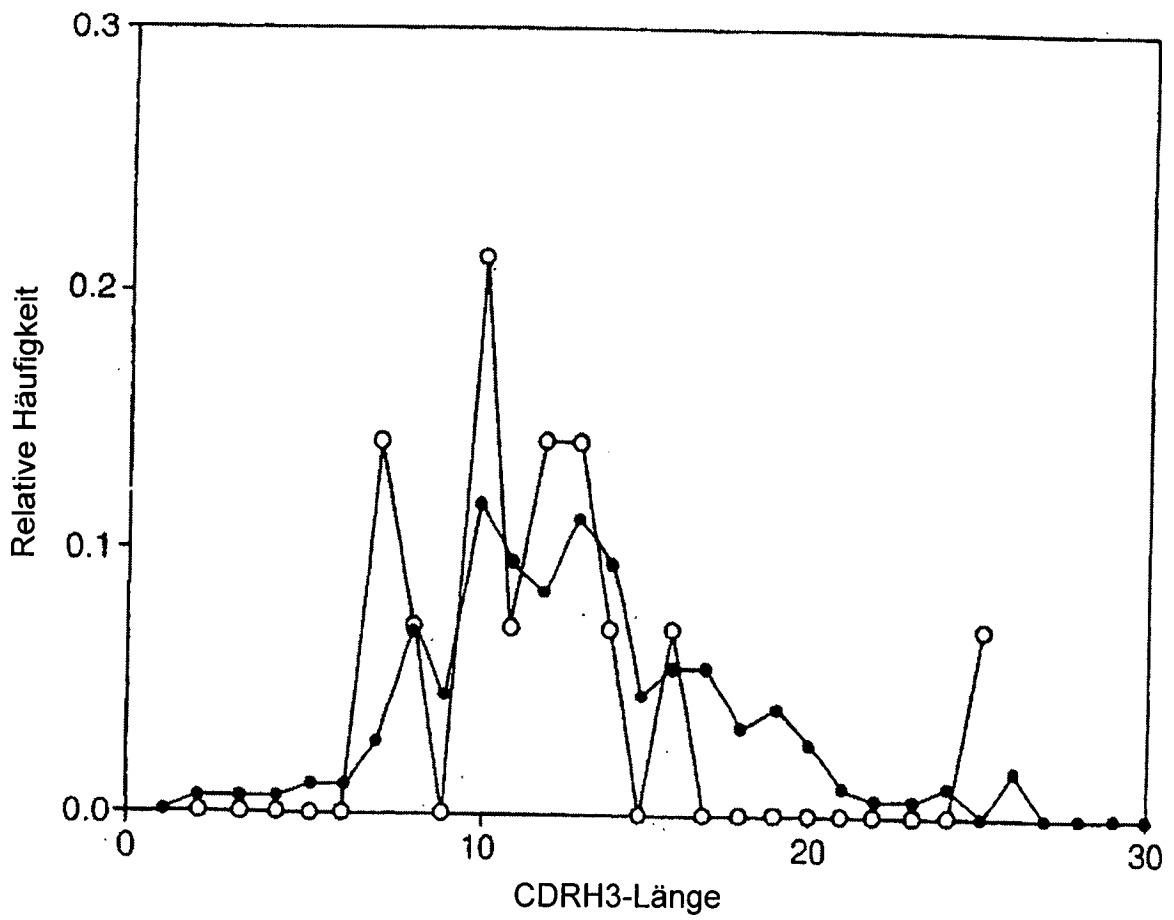


Fig.7.

—●— Länge nach Kabat  
—○— experimentell





Konstruktion der Bibliothek

Fig.8A.

Verschiedene Matrizen, z.B. scFv-B11 (für Framework-Amplifizierung), cDNA von Lymphozyten des peripheren Blutes, Tonsillen-cDNA, Milz-cDNA (für CDR-Amplifizierung)

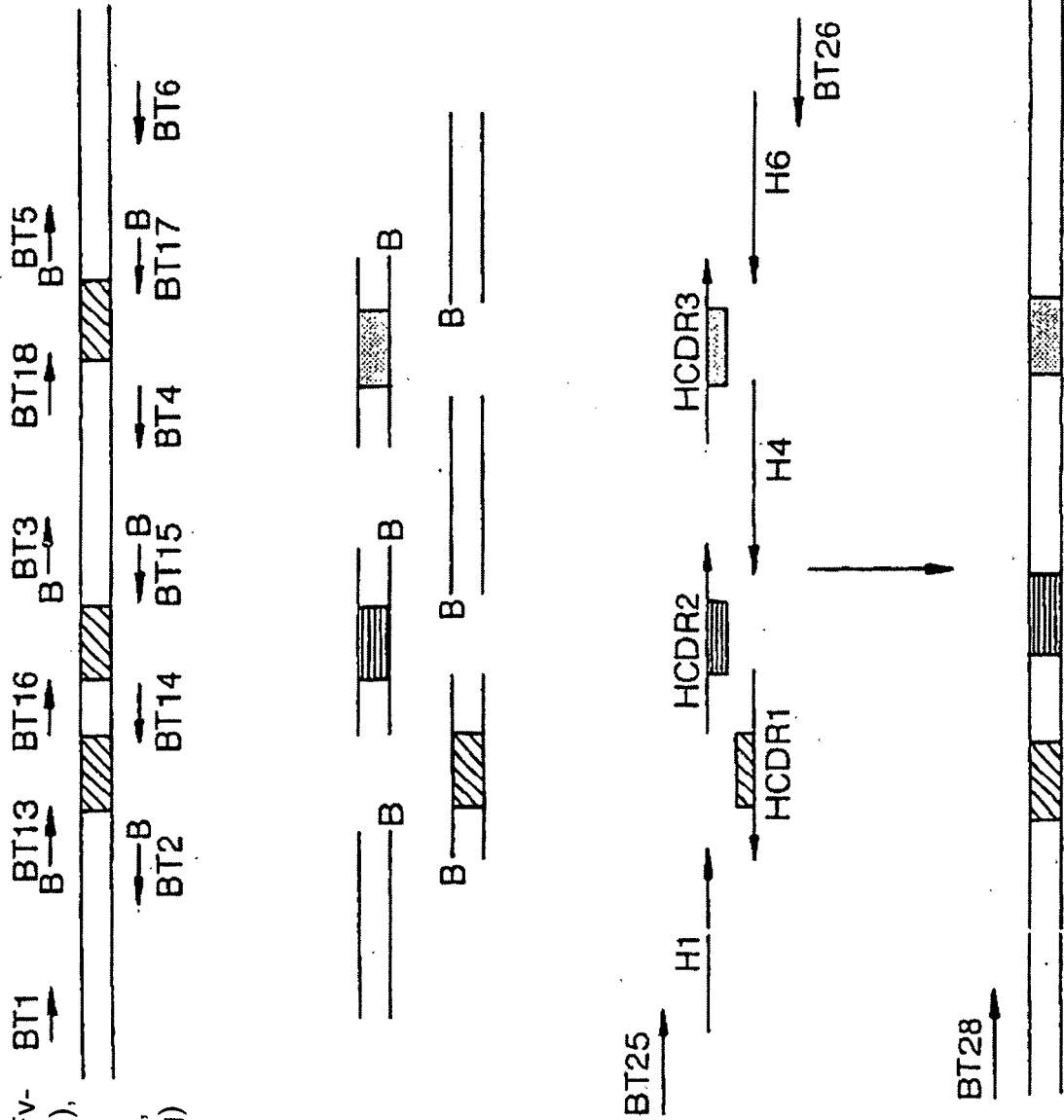


Fig. 8B.

Konstruktion der Bibliothek

Verschiedene Matrizen, z.B. scFv-B11 (für Framework-Amplifizierung), cDNA von Lymphozyten des peripheren Blutes, Tonsillen-cDNA, Milz-cDNA (für CDR-Amplifizierung)

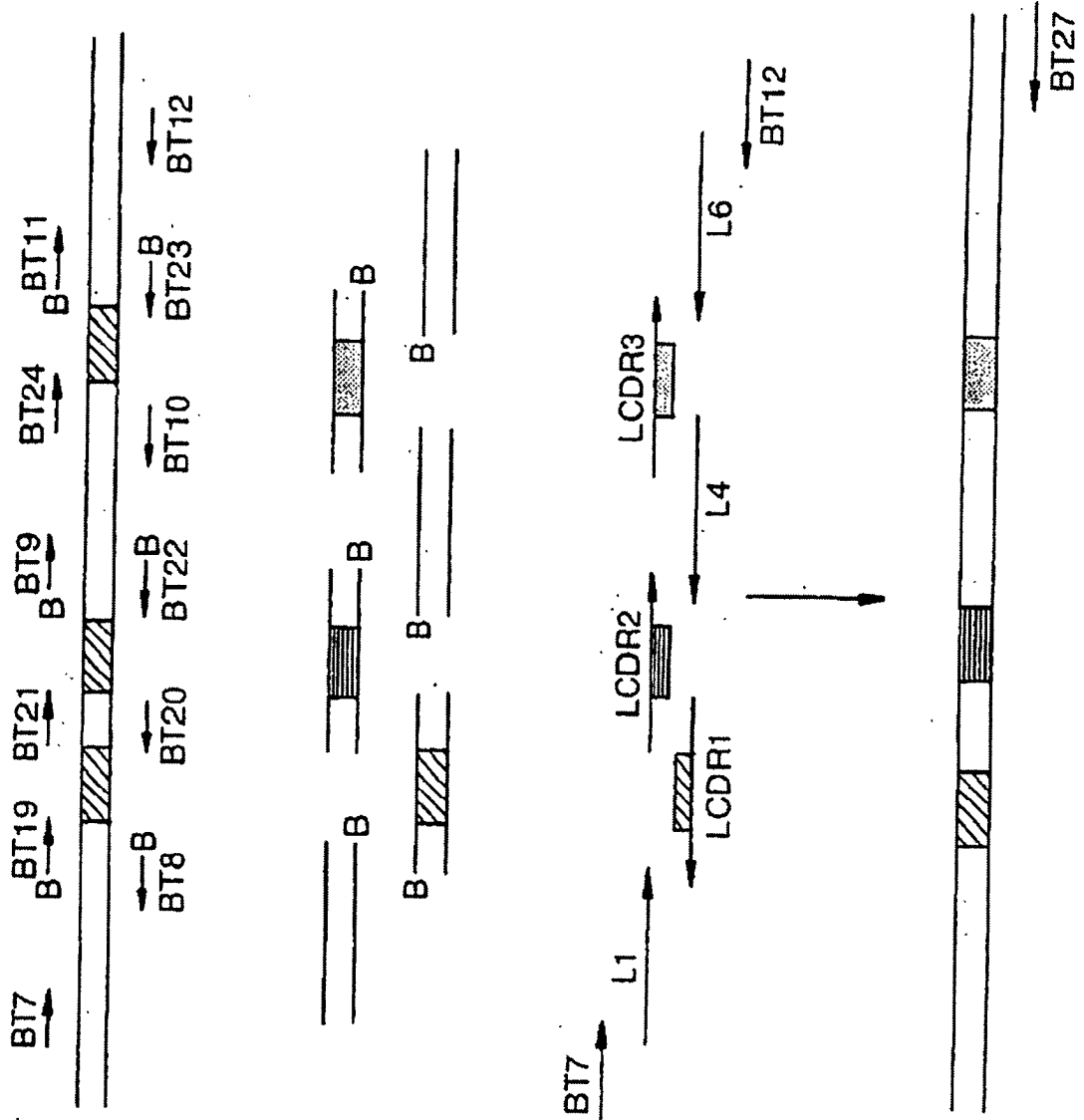


Fig.9A.

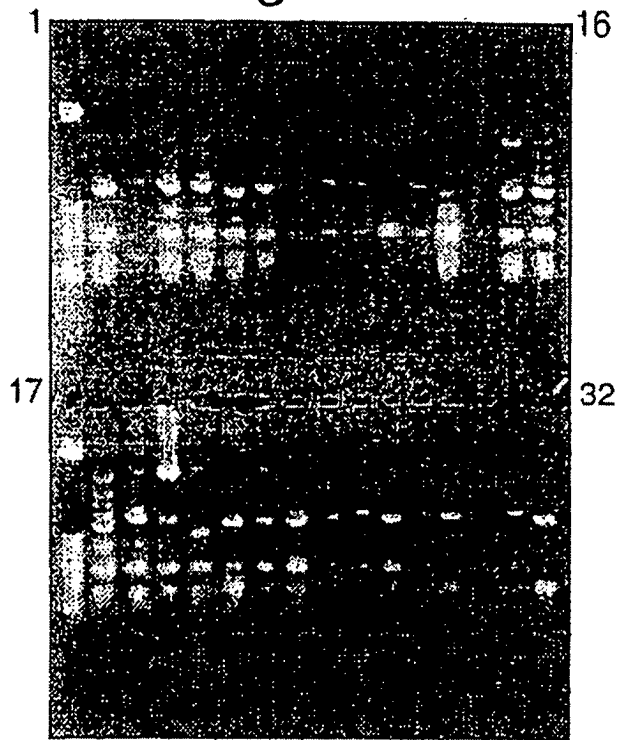


Fig.9B.

