



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 602 14 682 T2** 2007.09.13

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 366 074 B1**

(51) Int Cl.<sup>8</sup>: **C07K 14/535** (2006.01)

(21) Deutsches Aktenzeichen: **602 14 682.8**

(86) PCT-Aktenzeichen: **PCT/EP02/02148**

(96) Europäisches Aktenzeichen: **02 732 448.2**

(87) PCT-Veröffentlichungs-Nr.: **WO 2002/070548**

(86) PCT-Anmeldetag: **28.02.2002**

(87) Veröffentlichungstag  
der PCT-Anmeldung: **12.09.2002**

(97) Erstveröffentlichung durch das EPA: **03.12.2003**

(97) Veröffentlichungstag  
der Patenterteilung beim EPA: **13.09.2006**

(47) Veröffentlichungstag im Patentblatt: **13.09.2007**

(30) Unionspriorität:  
**01105775**      **08.03.2001**      **EP**

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

(73) Patentinhaber:  
**Merck Patent GmbH, 64293 Darmstadt, DE**

(72) Erfinder:  
**CARR, J., Francis, Aberdeenshire AB23 8XU, GB;  
CARTER, Graham, Aberdeenshire AB21 7XB, GB**

(54) Bezeichnung: **MODIFIZIERTER HUMANER GRANULOZYTEN MAKROPHAGEN KOLONIEN STIMULIERENDER  
FAKTOR (GM-CSF) MIT REDUZIERTER IMMUNOGENITÄT**

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**

## GEBIET DER ERFINDUNG

**[0001]** Die vorliegende Erfindung betrifft GM-CSF-Polypeptide zur Verabreichung insbesondere an Menschen und insbesondere für therapeutische Verwendung. Die Polypeptide sind modifizierte Polypeptide, wobei die Modifizierung zu einer verringerten Neigung des Polypeptids führt, bei Verabreichung an das menschliche Subjekt eine Immunantwort hervorzurufen. Die Erfindung betrifft insbesondere die Modifizierung des humanen Granulozyten-Makrophagen-Kolonie-stimulierenden Faktors (GM-CSF), um zu GM-CSF-Proteinvarianten zu führen, die im Wesentlichen nicht immunogen oder weniger immunogen sind als jegliches nicht-modifiziertes Gegenstück, bei Verwendung in vivo. Weiterhin werden T-Zell-Epitop-Peptide offenbart, die von genanntem, nicht-modifiziertem Protein durch Mittel abstammen, mit denen es möglich ist, modifizierte GM-CSF-Varianten mit verringerter Immunisierungskraft zu erzeugen.

## STAND DER TECHNIK

**[0002]** Es gibt viele Fälle, wodurch die Effizienz eines therapeutischen Proteins durch eine unerwünschte Immunreaktion auf das therapeutische Protein eingeschränkt ist. Einige monoklonale Antikörper aus Mäusen waren als Therapien aussichtsreich für eine Anzahl an humanen Krankheitsbildern, aber versagten in bestimmten Fällen aufgrund der Induktion von deutlichen Graden an humaner antimuriner Antikörper-(HAMA)-Antwort [Schroff, R. W. et al (1985) *Cancer Res.* 45: 879–885; Shawler, D.L. et al (1985) *J. Immunol.* 135: 1530–1535]. Für monoklonale Antikörper wurden eine Anzahl an Techniken in dem Versuch entwickelt, die HAMA-Antwort zu verringern [WO 89/09622; EP 0239400; EP 0438310; WO 91/06667]. Diese rekombinanten DNA-Ansätze haben im Allgemeinen die genetische Information der Maus im endgültigen Antikörperkonstrukt verringert, während sie die humane genetische Information im endgültigen Konstrukt erhöhten. Dennoch haben die resultierenden „humanisierten“ Antikörper in einigen Fällen noch immer eine Immunantwort in Patienten hervorgehoben [Issacs J.D. (1990) *Sem. Immunol.* 2: 449, 456; Rebello, P.R. et al (1999) *Transplantation* 68: 1417–1420].

**[0003]** Antikörper sind nicht die einzige Klasse an Polypeptidmolekül, die als ein therapeutisches Mittel verabreicht wird, gegen das eine Immunantwort stattfinden kann. Sogar Proteine humanen Ursprungs und mit den gleichen Aminosäuresequenzen wie sie innerhalb von Menschen vorkommen, können noch immer eine Immunantwort in Menschen induzieren. Bekannte Beispiele umfassen die therapeutische Verwendung von Granulozyten-Makrophagen-Kolonie-stimulierendem Faktor [Wadhwa, M. et al (1999) *Clin. Cancer Res.* 5: 1353–1361] und Interferon alpha 2 [Russo, D. et al (1996) *Br. J. Haem.* 94: 300–305; Stein, R. et al (1988) *New Engl. J. Med.* 318: 1409–1413], unter anderen.

**[0004]** Ein Hauptfaktor bei der Induzierung einer Immunantwort ist die Gegenwart von Peptiden innerhalb des Proteins, die die Aktivität von T-Zellen über Präsentation auf MHC-Klasse-II-Molekülen, so genannter „T-Zell-Epitope“, stimulieren können. Derartige mögliche T-Zell-Epitope werden üblicherweise als jegliche Aminosäurerestsequenz mit einer Fähigkeit zur Bindung an MHC-Klasse-II-Molekülen definiert. Derartige T-Zell-Epitope können gemessen werden, um MHC-Bindung nachzuweisen. Implizit bedeutet ein „T-Zell-Epitop“ ein Epitop, das bei Bindung an MHC-Moleküle durch einen T-Zell-Rezeptor (TCR) erkannt werden kann, und das, zumindest im Prinzip, die Aktivierung dieser T-Zellen verursachen kann, indem ein TCR veranlasst wird, eine T-Zell-Antwort zu fördern. Es ist jedoch im Allgemeinen bekannt, dass bestimmte Peptide, die nachweislich an MHC-Klasse-II-Moleküle binden, in einer Proteinsequenz bewahrt werden können, da derartige Peptide als „selbst innerhalb des Organismus, in den das endgültige Protein verabreicht wird“, erkannt werden.

**[0005]** Es ist bekannt, dass bestimmte dieser T-Zell-Epitop-Peptide während des Abbaus von Peptiden, Polypeptiden oder Proteinen innerhalb von Zellen freigesetzt und anschließend durch Moleküle des Haupthistokompatibilitätskomplexes (MHC) präsentiert werden können, um die Aktivierung der T-Zellen auszulösen. Für durch MHC-Klasse-II-präsentierte Peptide kann eine derartige Aktivierung von T-Zellen, zum Beispiel, eine Antikörperantwort durch direkte Stimulierung der B-Zellen hervorrufen, um derartige Antikörper zu produzieren.

**[0006]** MHC-Klasse-II-Moleküle sind eine Gruppe von hoch polymorphen Proteinen, die eine zentrale Rolle bei der Helfer-T-Zellselektion und -aktivierung spielen. Die humane Leukozyten-Antigengruppe DR (HLA-DR) sind der vorherrschende Isotyp dieser Proteingruppe und sind der Hauptschwerpunkt der vorliegenden Erfindung. Die Isotypen HLA-DQ und HLA-DP führen jedoch gleichartige Funktionen aus, so dass die vorliegende Erfindung ebenso auf diese anwendbar ist. Die MHC-Klasse-II-DR-Moleküle bestehen aus einer alpha- und einer beta-Kette, die an ihren C-Termini durch die Zellmembran einfügen. Jedes Hetero-Dimer besitzt eine Li-

gandenbindungsdomäne, die an Peptide bindet, deren Länge zwischen 9 und 20 Aminosäuren schwankt, obwohl die Bindungsfurche maximal 11 Aminosäuren aufnehmen kann. Die Ligandenbindungsdomäne besteht aus Aminosäuren 1 bis 85 der Alphakette und Aminosäuren 1 bis 94 der Betakette. Bei DQ-Molekülen wurde kürzlich eine homologe Struktur nachgewiesen und es wird auch erwartet, dass die DP-Familie-Proteine sehr ähnlich sind. In Menschen sind ungefähr 70 verschiedene Allotypen des DR-Isotypen bekannt, von DQ sind es 30 verschiedene Allotypen und für DP sind 47 verschiedene Allotypen bekannt. Jedes Individuum trägt zwei bis vier DR-Allele, zwei DQ- und zwei DP-Allele. Die Struktur einer Anzahl an DR-Molekülen wurde gelöst und derartige Strukturen weisen auf eine offenendige Peptid-Bindungsfurche mit einer Anzahl an hydrophoben Taschen hin, die hydrophobe Reste (Taschenreste) des Peptids aufnehmen [Brown et al Nature (1993) 364: 33; Stern et al (1994) Nature 368: 215]. Polymorphismus, der die verschiedenen Allotypen des Klasse-II-Moleküls identifiziert, trägt zu einer breiten Diversität an verschiedenen Bindungsoberflächen für Peptide innerhalb einer Peptid-Bindungsfurche bei und stellt auf dem Populationsniveau bezogen auf die Fähigkeit, fremde Proteine zu erkennen und eine Immunantwort auf pathogene Organismen in Gang zu setzen, maximale Flexibilität sicher. Es gibt eine beträchtliche Menge an Polymorphismus innerhalb der Ligandenbindungsdomäne mit unterschiedlichen „Familien“ innerhalb verschiedener geographischer Populationen und ethnischer Gruppen. Dieser Polymorphismus wirkt sich auf die Bindungseigenschaften der Peptidbindungsdomäne aus, so dass verschiedene „Familien“ an DR-Molekülen Spezifitäten für Peptide mit unterschiedlichen Sequenzeigenschaften haben, auch wenn es einige Überlappung geben mag. Diese Spezifität bestimmt Erkennung von Th-Zell-Epitopen (Klasse-II-T-Zell-Antwort), die letztendlich für das Antreiben der Antikörperantwort zu  $\beta$ -Zell-Epitopen verantwortlich sind, die auf dem gleichen Protein anwesend sind, von dem das Th-Zell-Epitop abstammt. So wird die Immunantwort auf ein Protein in einem Individuum stark durch die T-Zell-Epitoperkennung beeinflusst, die eine Funktion der Peptidbindungsspezifität des HLA-DR-Allotyps dieses Individuums ist. Um T-Zell-Epitope in einem Protein oder Peptid im Kontext einer globalen Population zu identifizieren, ist es daher wünschenswert, die Bindungseigenschaften eines Satzes von HLA-DR-Allotypen zu betrachten, der so vielfältig wie möglich ist, um so einen Prozentsatz der Weltbevölkerung abzudecken, der so hoch wie möglich ist.

**[0007]** Eine Immunantwort auf ein therapeutisches Protein wie das Protein, das Gegenstand dieser Erfindung ist, erfolgt über den MHC-Klasse-II-Peptidpräsentationsweg. Hier werden exogene Proteine umschlossen und prozessiert für Präsentation in Verbindung mit MHC-Klasse-II-Molekülen vom DR-, DQ- oder DP-Typ. MHC-Klasse-II-Moleküle werden durch professionelle Antigen-präsentierende Zellen (APCs) wie Makrophagen und dendritische Zellen unter anderen exprimiert. Bindung eines MHC-Klasse-II-Peptidkomplexes durch einen verwandten T-Zellrezeptor auf der Oberfläche der T-Zelle, zusammen mit einer Querverbindung von bestimmten anderen Korezeptoren wie dem CD4-Molekül kann einen aktivierten Zustand innerhalb der T-Zelle hervorrufen.

**[0008]** Aktivierung führt zur Freisetzung von Cytokinen, die weiterhin andere Lymphozyten wie B-Zellen aktivieren, um Antikörper zu produzieren oder T-Killer-Zellen als eine volle zelluläre Immunantwort zu aktivieren. Die Fähigkeit eines Peptids an ein gegebenes MHC-Klasse-II-Molekül zur Präsentation auf der Oberfläche eines APCs zu binden, ist von einer Anzahl an Faktoren abhängig, insbesondere seiner Primärsequenz. Dies wird sowohl seine Neigung zur proteolytischen Spaltung wie auch seine Affinität zur Bindung innerhalb des Peptidbindungsspalts des MHC-Klasse-II-Moleküls beeinflussen. Der MHC-Klasse-II/Peptid-Komplex auf der APC-Oberfläche präsentiert einem bestimmten T-Zellrezeptor (TCR), der Determinanten erkennen kann, eine Bindungsseite, die sowohl von den exponierten Resten des Peptids wie auch von dem MHC-Klasse-II-Molekül bereitgestellt werden.

**[0009]** Im Stand der Technik gibt es Prozeduren zur Identifizierung von synthetischen Peptiden, die an MHC-Klasse-II-Moleküle binden können (z. B. WO 98/52976 und WO 00/34317). Derartige Peptide dienen vielleicht nicht in allen Situationen als T-Zell-Epitope, insbesondere in vivo, aufgrund der Verarbeitungswege oder anderer Phänomene. T-Zell-Epitopidentifizierung ist der erste Schritt zu einer Epitopentfernung. Die Identifizierung und Entfernung von möglichen T-Zell-Epitopen aus Proteinen wurde bereits früher offenbart. Nach dem Stand der Technik wurden Verfahren bereitgestellt, die den Nachweis von T-Zell-Epitopen üblicherweise durch Rechenmittel ermöglichen, indem nach erkannten Sequenzmotiven in experimentell bestimmten T-Zell-Epitopen gescannt wird, oder alternativ unter Verwendung von Rechentechniken zur Vorhersage von MHC-Klasse-II-bindenden Peptiden und insbesondere DR-bindenden Peptiden. WO 98/52976 und WO 00/34317 lehren rechnerische Threading-Ansätze zur Identifizierung von Polypeptidsequenzen mit dem Potenzial, ein Subset von humanen MHC-Klasse-II-DR-Allotypen zu binden. In diesen Lehren werden vorhergesagte T-Zell-Epitope durch die Verwendung von wohl überlegter Aminosäuresubstitution innerhalb der Primärsequenz des therapeutischen Antikörpers oder nicht-Antikörper-Proteins sowohl von nicht humaner wie auch von humaner Abstammung entfernt.

**[0010]** Andere Techniken, die lösliche Komplexe von rekombinanten MHC-Molekülen zusammen mit synthetischen Peptiden nutzen und an T-Zell-Klone aus peripheren Blutproben von menschlichen oder experimentellen tierischen Subjekten binden können, wurden im Stand der Technik verwendet [Kern, F. et al (1998) Nature Medicine 4: 975–978; Kwok, W.W. et al (2001) TRENDS in Immunology 22: 583–588] und können ebenfalls in einer Epitopidentifizierungsstrategie verwendet werden.

**[0011]** Wie oben dargestellt und als Konsequenz daraus, wäre es wünschenswert, T-Zell-Epitope aus einem gegebenen, im Prinzip therapeutisch wertvollen, aber ursprünglich immunogenen Peptid, Polypeptid oder Protein zu identifizieren und zu entfernen oder zumindest zu verringern.

**[0012]** Eines dieser therapeutisch wertvollen Moleküle ist der humane Granulozyten-Makrophagen-Kolonie-stimulierende Faktor (GM-CSF). GM-CSF ist ein saures Glykoprotein, das ursprünglich als die Produktion von Granulozyten und Monozyten aus ihren Knochenmarksvorläufermolekülen stimulierend definiert wurde. Das Protein enthält 127 Aminosäurereste und teilt eine deutliche Sequenzhomologie mit GM-CSF-Proteinen von anderen Säugetierarten. Die Verfügbarkeit von rekombinantem GM-CSF hat gezeigt, dass dieser wichtige hämopoetische Wachstumsfaktor auch die Bildung von erythrozytären und Megakaryozyt-Vorläufermolekülen stimulieren kann und zusätzlich reife Neutrophile, Monozyten, eosinophile und Mastzellen stimuliert, so dass diese funktionell aktiviert werden [Kaushanski, K. et al (1986) Proc. Natl. Acad. Sci. U.S.A. 83: 3101; Emerson, S.G. et al (1985) Eur. J. Biochem. 165: 7; Weisbart, R.H. et al (1985) Nature 314: 361; Grabstein, K.H. et al (1985) Science 232: 506].

**[0013]** Das primäre Translationsprodukt des humanen GM-CSF-Gens ist ein Protein mit 144 Aminosäureresten mit einer reifen (prozessierten) Form mit 127 Resten. Es besteht ein hoher Grad an Homologie zwischen GM-CSF-Proteinen aus anderen Säugetierarten und vollständige Konservierung der Disulfidstruktur, obwohl GM-CSF von anderen Spezies wie der Maus, keine humanen Zellen stimulieren können.

**[0014]** Andere haben GM-CSF-Moleküle und Analoga bereitgestellt [US 5 391 485; US 5 032 676; US 5 942 253; US 6 120 807], aber keine dieser Lehren erkennt die Wichtigkeit der T-Zell-Epitope für die immunogenen Eigenschaften des Proteins oder wurde so entworfen, um die genannten Eigenschaften auf einem spezifischen und kontrollierten Weg gemäß dem Schema der vorliegenden Erfindung direkt zu beeinflussen.

**[0015]** Die primäre 127-Aminosäuresequenz des humanen GM-CSF ist wie folgt:  
 APARSPSPSTQPWEHVNAIQEARRLLNLSRDTAAEMNETVEVISEMFDLQEPTCLQTRLELYKQG  
 LRGSLLKLGPLTMMASHYKQHCPTPETSCATQTITFESFKENLKDFFLLVIPFDCWEPVQE

**[0016]** Es besteht jedoch ein fortwährender Bedarf an GM-CSF-Analoga mit verbesserten Eigenschaften. Gewünschte Verbesserungen umfassen alternative Schemata und Modalitäten für die Expression und Reinigung des genannten Therapeutikums, aber auch und insbesondere Verbesserungen der biologischen Eigenschaften des Proteins.

**[0017]** Es besteht ein besonderer Bedarf an Verbesserung der in-vivo-Charakteristika bei Verabreichung an das menschliche Subjekt. In dieser Beziehung ist es äußerst gewünscht, GM-CSF mit verringertem oder abwesendem Potenzial, eine Immunantwort in dem menschlichen Subjekt hervorzurufen, bereitzustellen.

#### ZUSAMMENFASSUNG UND BESCHREIBUNG DER ERFINDUNG

**[0018]** Die vorliegende Erfindung ermöglicht modifizierte Formen des humanen Granulozyten-Makrophagen-Kolonie-stimulierenden Faktors (GM-CSF), in dem die Immuneigenschaft mittels verringerter oder entfernter Anzahlen an möglichen T-Zell-Epitopen modifiziert ist.

**[0019]** Die Erfindung offenbart Sequenzen, die innerhalb der GM-CSF-Primärsequenz identifiziert wurden, die wegen eines MHC-Klasse-II-Bindungspotenzials mögliche T-Zell-Epitope sind. Diese Offenbarung betrifft insbesondere das humane GM-CSF-Protein mit 127 Aminosäureresten. Die Erfindung offenbart auch spezifische Positionen innerhalb der Primärsequenz des Moleküls, welche gemäß der Erfindung durch spezifische Aminosäuresubstitution, -addition oder -deletion verändert werden sollen, ohne die biologische Aktivität im Prinzip zu beeinflussen. In Fällen, in denen der Verlust der Immunisierungskraft nur durch einen gleichzeitigen Verlust an biologischer Aktivität erreicht werden kann, ist es möglich, genannte Aktivität durch weitere Veränderungen innerhalb der Aminosäuresequenz des Proteins wiederherzustellen.

**[0020]** Die Erfindung offenbart weiterhin Verfahren zur Herstellung derartiger modifizierter Moleküle, und vor

allem Verfahren zur Identifizierung genannter T-Zell-Epitope, welche Veränderung erfordern, um immunogene Stellen zu verringern oder zu entfernen.

**[0021]** Das erfindungsgemäße Protein sollte eine erhöhte Halbwertszeit im menschlichen Subjekt aufweisen und sollte bei chronischen und wiederauftretenden Krankheitsbildern, wie es bei einer Anzahl von Indikationen für GM-CSF der Fall ist, von besonderem Nutzen sein. Die vorliegende Erfindung ermöglicht modifizierte Formen von GM-CSF-Proteinen, von denen erwartet wird, dass sie in vivo verbesserte Eigenschaften aufweisen. Diese modifizierten GM-CSF-Moleküle können in pharmazeutischen Zusammensetzungen verwendet werden.

**[0022]** Zusammenfassend betrifft die Erfindung die folgenden Themen:

- ein modifiziertes Molekül mit der biologischen Aktivität von humanem GM-CSF, und das im Wesentlichen nicht immunogen oder weniger immunogen ist als jegliches nicht-modifiziertes Molekül mit der gleichen biologischen Aktivität bei Verwendung in vivo;
- ein entsprechend spezifiziertes Molekül, wobei genannter Verlust der Immunisierungskraft durch Entfernen von einem oder mehreren T-Zell-Epitopen erreicht wird, die vom ursprünglich nicht-modifizierten Molekül abstammen;
- ein entsprechend spezifiziertes Molekül, wobei genannter Verlust der Immunisierungskraft durch Verringerung der Anzahl an MHC-Allotypen erreicht wird, die von dem genannten Molekül abstammende Peptide binden können;
- ein entsprechend spezifiziertes Molekül, wobei ein T-Zell-Epitop entfernt ist;
- ein entsprechend spezifiziertes Molekül, wobei genannte, ursprünglich vorhandene T-Zell-Epitope MHC-Klasse-II-Liganden oder Peptidsequenzen sind, die die Fähigkeit zur Stimulierung oder Bindung von T-Zellen über Präsentieren auf Klasse-II zeigen;
- ein entsprechend spezifiziertes Molekül, wobei genannte Peptidsequenzen ausgewählt werden aus der Gruppe wie in Tabelle 1 dargestellt;
- ein entsprechend spezifiziertes Molekül, wobei 1–9 Aminosäurereste, vorzugsweise ein Aminosäurerest, in irgendeinem der ursprünglich anwesenden T-Zell-Epitope verändert werden;
- ein entsprechend spezifiziertes Molekül, wobei die Veränderung der Aminosäurereste Substitution, Addition oder Deletion von ursprünglich anwesendem(en) Aminosäurerest(en) durch andere(n) Aminosäurerest(e) an spezifischer(en) Position(en) ist;
- ein entsprechend spezifiziertes Molekül, wobei ein oder mehrere der Aminosäurerests substitutionen wie in Tabelle 2 angegeben ausgeführt werden;
- ein entsprechend spezifiziertes Molekül, wobei (zusätzlich) eine oder mehrere der Aminosäurerests substitutionen wie in Tabelle 3 angegeben ausgeführt werden, zur Verringerung der Anzahl an MHC-Allotypen, die von genanntem Molekül abstammende Peptide binden können;
- ein entsprechend spezifiziertes Molekül, wobei, gegebenenfalls, zusätzlich weiterhin Veränderung gewöhnlich durch Substitution, Addition oder Deletion von spezifischer(en) Aminosäure(n) ausgeführt wird, um biologische Aktivität des genannten Moleküls wiederherzustellen;
- eine DNA-Sequenz oder -molekül, das irgendeines der genannten, spezifizierten, modifizierten Moleküle wie oben und unten definiert kodiert;
- eine pharmazeutische Zusammensetzung, enthaltend ein modifiziertes Molekül mit der biologischen Aktivität von GM-CSF wie oben und/oder in den Ansprüchen definiert, gegebenenfalls zusammen mit einem pharmazeutisch annehmbaren Träger, Verdünnungsmittel oder Hilfsmittel;
- ein Verfahren zur Herstellung eines modifizierten Moleküls mit der biologischen Aktivität von GM-CSF wie in irgendeinem der Ansprüche der oben genannten Ansprüche definiert, umfassend die folgenden Schritte: (i) Bestimmung der Aminosäuresequenz des Polypeptids oder Teils davon; (ii) Identifizierung eines oder mehrerer möglicher T-Zell-Epitope innerhalb der Aminosäuresequenz des Proteins durch irgendein Verfahren, einschließlich Bestimmung der Bindung der Peptide an MHC-Moleküle unter Verwendung von in-vitro- oder in-silico-Techniken oder biologischen Assays; (iii) Entwurf neuer Sequenzvarianten mit einer oder mehreren Aminosäuren innerhalb der identifizierten möglichen T-Zell-Epitope, die auf eine solche Weise modifiziert wurden, um im Wesentlichen die Aktivität des T-Zell-Epitops, wie durch Bindung der Peptide an MHC-Moleküle unter Verwendung von in-vitro- oder in-silico-Techniken oder biologischen Assays bestimmt, zu verringern oder zu entfernen; (iv) Konstruieren derartiger Sequenzvarianten durch DNA-Rekombinationstechniken und Testen genannter Varianten, um eine oder mehrere Varianten mit erwünschten Eigenschaften zu identifizieren, und (v) gegebenenfalls Wiederholung der Schritte (ii)–(iv);
- ein entsprechend spezifiziertes Verfahren, wobei Schritt (iii) durch Substitution, Addition oder Deletion von 1–9 Aminosäureresten in irgendeinem der ursprünglich anwesenden T-Zell-Epitope ausgeführt wird;
- ein entsprechend spezifiziertes Verfahren, wobei die Veränderung in Bezug auf eine homologe Proteinsequenz und/oder in-silico-Modellierungstechniken gemacht wird;
- ein entsprechend spezifiziertes Verfahren, wobei der oben genannte Schritt (ii) durch folgende Schritte

ausgeführt wird: (a) Auswählen einer Region des Peptids mit bekannter Aminosäurerestsequenz; (b) aufeinanderfolgendes Abtasten überlappender Aminosäurerestsegmente mit zuvor bestimmter gleichmäßiger Größe, die aus mindestens drei Aminosäureresten bestehen, aus der ausgewählten Region; (c) Berechnen des MHC-Klasse-II-Molekül-Bindungswerts für jedes abgetastete Segment durch Aufsummieren zugewiesener Werte für jede in genanntem abgetasteten Aminosäurerestsegment vorhandene hydrophobe Aminosäurerest-Seitenkette; und (d) identifizieren mindestens eines der genannten Segmente, die für eine Modifizierung geeignet sind, auf Basis des berechneten MHC-Klasse-II-Molekül-Bindungswerts für dieses Segment, um den Gesamt-MHC-Klasse-II-Bindungswert des Peptids zu verändern, ohne im Wesentlichen den therapeutischen Nutzen des Peptids zu verringern; wobei Schritt (c) vorzugsweise unter Verwendung einer Böhm-Bewertungsfunktion ausgeführt wird, die so modifiziert ist, dass sie 12-6 van-der-Waals-Ligand-Proteinenergie-Abstoßungsterm und Ligandenkonformationsenergie-Term enthält, indem (1) eine erste Datenbank von MHC-Klasse-II-Molekülmodellen bereitgestellt wird, (2) eine zweite Datenbank zulässiger Peptidgrundgerüste für die MHC-Klasse-II-Moleküle bereitgestellt wird, (3) ein Modell aus der ersten Datenbank ausgewählt wird, (4) ein zulässiges Peptidgrundgerüst aus der zweiten Datenbank ausgewählt wird, (5) in jedem abgetasteten Segment vorhandene Aminosäurerest-Seitenketten identifiziert werden, (6) in jedem abgetasteten Segment der Bindungsaffinitätswert für alle vorhandenen Aminosäurerest-Seitenketten ermittelt wird und die Schritte (1) bis (5) für jedes Modell und jedes Grundgerüst wiederholt werden;

- ein 13mer-T-Zell-Epitop-Peptid mit einer möglichen MHC-Klasse-II-Bindungsaktivität und erzeugt aus immunogenetisch nicht-modifiziertem GM-CSF, ausgewählt aus der Gruppe wie in Tabelle 1 dargestellt, und seine Verwendung für die Herstellung von GM-CSF mit im Wesentlichen keiner oder geringerer Immunisierungskraft als jegliches nicht-modifiziertes Molekül mit der gleichen biologischen Aktivität bei Verwendung in vivo;
- eine Peptidsequenz bestehend aus mindestens 9 aufeinander folgenden Aminosäureresten eines 13mer-T-Zell-Epitop-Peptids wie oben spezifiziert und seine Verwendung zur Herstellung von GM-CSF mit im Wesentlichen keiner oder geringerer Immunisierungskraft als jegliches nicht-modifiziertes Molekül mit der gleichen biologischen Aktivität bei Verwendung in vivo;
- ein immunogenisch modifiziertes Molekül mit der biologischen Aktivität des humanen GM-CSF, erhältlich durch irgendeines der Verfahren wie oben und unten spezifiziert;

**[0023]** Der Begriff „T-Zell-Epitop“ bedeutet gemäß dem Verständnis dieser Erfindung eine Aminosäuresequenz, die an MHC-Klasse-II binden kann, T-Zellen stimulieren und/oder auch T-Zellen in einem Komplex mit MHC-Klasse-II binden kann (ohne notwendigerweise messbare Aktivierung). Der Begriff „Peptid“ wie hier und in den angehängte Ansprüchen verwendet, ist eine Verbindung, die zwei oder mehr Aminosäuren umfasst. Die Aminosäuren werden über eine Peptidbindung (hier unten definiert) miteinander verbunden. Es gibt 20 unterschiedliche, natürlich vorkommende Aminosäuren, die an der biologischen Herstellung von Peptiden beteiligt sind, und jegliche Anzahl von ihnen kann in jeglicher Reihenfolge verbunden werden, um eine Peptidkette oder -ring zu bilden. Die natürlich vorkommenden, in der biologischen Herstellung von Peptiden verwendeten Aminosäuren haben alle L-Konfiguration. Synthetische Peptide können unter Anwendung konventioneller Syntheseverfahren unter Verwendung von L-Aminosäuren, D-Aminosäuren, oder verschiedenen Kombinationen von Aminosäuren der beiden unterschiedlichen Konfigurationen hergestellt werden. Einige Peptide enthalten nur wenige Aminosäureeinheiten. Kurze Peptide, z. B., mit weniger als zehn Aminosäureeinheiten, werden manchmal als „Oligopeptide“ bezeichnet. Andere Peptide enthalten eine große Anzahl an Aminosäureresten, z. B. bis zu 100 oder mehr, und werden als „Polypeptide“ bezeichnet. Aufgrund der Konvention kann ein „Polypeptid“ als jegliche Peptidkette betrachtet werden, die drei oder mehr Aminosäuren enthält, während „Oligopeptid“ üblicherweise als eine besondere Art von „kurzem“ Polypeptid betrachtet wird. So, wie hier verwendet, ist es selbstverständlich, dass jeglicher Bezug auf ein „Polypeptid“ auch ein Oligopeptid umfasst. Weiterhin umfasst jeglicher Bezug auf ein „Peptid“ Polypeptide, Oligopeptide und Proteine. Jede unterschiedliche Anordnung von Aminosäuren bildet verschiedene Polypeptide oder Proteine. Die Anzahl an Polypeptiden – und so die Anzahl an verschiedenen Proteinen –, die gebildet werden kann, ist praktisch unbegrenzt.

**[0024]** „Alpha-Kohlenstoff“ (C $\alpha$ ) ist das Kohlenstoffatom des Kohlenstoff-Wasserstoff-(CH)Bestandteils, der sich in der Peptidkette befindet. Eine „Seitenkette“ ist eine an dem C $\alpha$  hängende Gruppe, die eine einfache oder komplexe Gruppe oder Einheit enthalten kann, mit physikalischen Abmessungen, die deutlich variieren können, verglichen mit den Abmessungen des Peptids.

**[0025]** Die Erfindung kann auf jegliche GM-CSF-Spezies an Molekül mit im Wesentlichen den gleichen primären Aminosäuresequenzen wie den hier offenbarten angewandt werden und umfasst daher immer GM-CSF-Moleküle, die durch Mittel der Gentechnik oder andere Prozesse abgeleitet wurden, und können mehr oder weniger als 127 Aminosäurereste enthalten.

**[0026]** GM-CSF-Proteine, wie solchen von anderen Säugerquellen identifizierte, haben viele der Peptidsequenzen der vorliegenden Offenbarung gemeinsam und haben viele Peptidsequenzen mit im Wesentlichen der gleichen Sequenz gemeinsam, wie denen des offenbarten Listing. Derartige Proteinsequenzen fallen daher genauso unter den Bereich der vorliegenden Erfindung.

**[0027]** Die Erfindung ist dazu gedacht, die praktische Realität zu überwinden, dass lösliche Proteine, die in autologe Organismen eingeführt wurden, eine Immunantwort auslösen können, die zur Entwicklung von Wirtsantikörpern führt, die das lösliche Protein binden. Ein Beispiel unter anderen ist Interferon alpha 2, für das ein Anteil der menschlichen Patienten Antikörper produziert, obwohl dieses Protein endogen hergestellt wird [Russo, D. et al (1996) *ibid*; Stein, R. et al (1988) *ibid*]. Es ist wahrscheinlich, dass die gleiche Situation die therapeutische Verwendung von GM-CSF betrifft und die vorliegende Erfindung trachtet danach, dies durch Bereitstellung von GM-CSF-Proteinen mit veränderter Neigung, eine Immunantwort bei Verabreichung an einen menschlichen Wirt hervorzurufen, anzugehen.

**[0028]** Das allgemeine Verfahren der vorliegenden Erfindung, das zu dem modifizierten GM-CSF führt, enthält die folgenden Schritte:

- (a) Bestimmen der Aminosäuresequenz des Polypeptids oder eines Teils davon,
- (b) Identifizieren eines oder mehrerer möglicher T-Zell-Epitope innerhalb der Aminosäuresequenz des Proteins durch jegliches Verfahren einschließlich Bestimmung der Bindung der Peptide an MHC-Moleküle unter Verwendung von In-vitro- oder In-silico-Techniken oder biologischen Assays;
- (c) Entwurf neuer Sequenzvarianten mit einer oder mehreren Aminosäuren innerhalb der identifizierten möglichen T-Zell-Epitope, die auf eine solche Weise modifiziert wurden, um die Aktivität des T-Zell-Epitops, wie durch die Bindung der Peptide an MHC-Moleküle unter Verwendung von In-vitro- oder In-silico-Techniken oder biologischen Assays bestimmt, im Wesentlichen zu verringern oder zu entfernen. Derartige Sequenzvarianten werden auf eine solche Weise geschaffen, um die Bildung neuer möglicher T-Zell-Epitope durch die Sequenzvariationen zu vermeiden, sofern die neuen möglichen T-Zell-Epitope nicht ihrerseits auf eine solche Weise modifiziert werden, dass die Aktivität des T-Zell-Epitops im Wesentlichen verringert oder entfernt wird; und
- (d) Aufbau derartiger Sequenzvarianten mit DNA-Rekombinationstechniken und Testen genannter Varianten, um eine oder mehrere Varianten mit wünschenswerten Eigenschaften gemäß wohlbekannter Rekombinationstechniken zu identifizieren.

**[0029]** Die Identifizierung von möglichen T-Zell-Epitopen gemäß Schritt (b) kann gemäß der Verfahren, die vorher beim Stand der Technik beschreiben wurden, ausgeführt werden. Geeignete Verfahren werden in WO 98/59244; WO 98/52976; WO 00/34317 offenbart und können vorzugsweise verwendet werden, um die Bindungsneigung von GM-CSF-abstammenden Peptiden an MHC-Klasse-II-Molekülen zu identifizieren.

**[0030]** Ein anderes sehr effizientes Verfahren zur Identifizierung von T-Zell-Epitopen durch Berechnung wird in dem BEISPIEL beschrieben, das gemäß dieser Erfindung eine bevorzugte Ausführungsform ist.

**[0031]** In der Ausführung wird eine Anzahl an unterschiedlichen GM-CSF-Proteinen hergestellt und auf die gewünschten Immun- und funktionellen Eigenschaften untersucht. Die Variantenproteine werden vorzugsweise durch DNA-Rekombinationstechniken hergestellt, obwohl andere Verfahren, einschließlich chemischer Synthese der GM-CSF-Fragmente, in Erwägung gezogen werden können.

**[0032]** Die Ergebnisse einer Analyse gemäß Schritt (b) des obigen Schemas und die humane GM-CSF-Proteinsequenz betreffend sind die 127 Aminosäurereste in Tabelle 1 aufgeführt.

Tabelle 1: Peptidsequenzen in humanem GM-CSF mit möglicher humaner MHC-Klasse-II-Bindungsaktivität.

PSPSTQPWEHVNA, QPWEHVNAIQEAR, EHVNAIQEARRLL, HVNAIQEARRLLN,  
 VNAIQEARRLLNL, NAIQEARRLLNLS, RLLNLSRDTAAE, RLLNLSRDTAAEM,  
 LNLSRDTAAEMNE, RDTAAEMNETVEV, AEMNETVEVISEM, NETVEVISEMFDL,  
 ETVEVISEMFDLQ, VEVISEMFDLQEP, EVISEMFDLQEP, ISEMFDLQEP, ISEMFDLQEP, ISEMFDLQEP,  
 SEMFDLQEP, EMFDLQEP, MFDLQEP, FDLQEP, FDLQEP, FDLQEP,  
 EPTCLQTRLELYK, TCLQTRLELYKQ, QTRLELYKQGLRG, TRLELYKQGLRGS,  
 LELYKQGLRGS, ELYKQGLRGS, QGLRGS, RGS, RGS, RGS,  
 GSLTKLKGPLTMM, SLTKLKGPLTMM, TKLKGPLTMM, KGPLTMM, KGPLTMM, KGPLTMM,  
 GPLTMM, PLTMM, LMM, TMM, TMM, TMM,  
 SHYKQHCPPTPET, CPPTPETS, PETS, CAT, CAT, CAT,  
 QTITFESFKENLK, ITFESFKENLK, ESFKENLK, SFKENLK, SFKENLK, SFKENLK,  
 ENLKDFLLV, NLKDFLLV, KDFLLV, DFLLV, DFLLV, DFLLV,  
 LLVIPFDCWEPVQ

**[0033]** Peptide sind 13mere, Aminosäuren werden unter Verwendung des Einbuchstabencodes identifiziert.

**[0034]** Die Ergebnisse eines Entwurfs und Konstrukts gemäß Schritt (c) und (d) des obigen Schemas und das modifizierte Molekül dieser Erfindung betreffend sind in Tabellen 2 und 3 aufgeführt.



Tabelle 2: Substitutionen, die zur Entfernung von möglichen T-Zell-Epitopen von humanem GM-CSF (WT = Wildtyp) führen.

Rest #	WT Rest	Substitution													
16	V	A	C	D	E	G	H	K	N	P	Q	R	S	T	
19	I	A	C	D	E	G	H	K	N	P	Q	R	S	T	
25	L	A	C	D	E	G	H	K	N	P	Q	R	S	T	
26	L	A	C	D	E	G	H	K	N	P	Q	R	S	T	
28	L	A	C	D	E	F	G	H	K	N	P	Q	R	S	T
36	M	A	C	D	E	G	H	K	N	P	Q	R	S	T	
40	V	A	C	D	E	G	H	K	N	P	Q	R	S	T	
43	I	A	C	D	E	G	H	K	N	P	Q	R	S	T	
46	M	A	C	D	E	G	H	K	N	P	Q	R	S	T	
47	F	A	C	D	E	G	H	K	N	P	Q	R	S	T	
49	L	A	C	D	E	G	H	K	N	P	Q	R	S	T	
55	L	A	C	D	E	G	H	K	N	P	Q	R	S	T	
59	L	A	C	D	E	G	H	K	N	P	Q	R	S	T	
61	L	A	C	D	E	G	H	K	N	P	Q	R	S	T	
62	Y	A	C	D	E	G	H	K	N	P	Q	R	S	T	
66	L	A	C	D	E	G	H	K	N	P	Q	R	S	T	
70	L	A	C	D	E	G	H	K	N	P	Q	R	S	T	
73	L	A	C	D	E	G	H	K	N	P	Q	R	S	T	
77	L	A	C	D	E	G	H	K	N	P	Q	R	S	T	
79	M	A	C	D	E	G	H	K	N	P	Q	R	S	T	
80	M	A	C	D	E	G	H	K	N	P	Q	R	S	T	
84	Y	C	D	E	G	H	N	P	R	S	T				
101	I	A	C	D	E	G	H	K	N	P	Q	R	S	T	
106	F	A	C	D	E	G	H	K	N	P	Q	R	S	T	
110	L	A	C	D	E	G	H	K	N	P	Q	R	S	T	
113	F	A	C	D	E	G	H	K	N	P	Q	R	S	T	
114	L	A	C	D	E	G	H	K	N	P	Q	R	S	T	
115	L	A	C	D	E	G	H	K	N	P	Q	R	S	T	
117	I	A	C	D	E	G	H	K	N	P	Q	R	S	T	

Tabelle 3: Zusätzliche Substitutionen, die zur Entfernung eines möglichen T-Zell-Epitops für 1 oder mehrere MHC-Alloypen führen

Rest #	WT Rest	Substitution																
15	H	A	C	F	G	I	L	M	P	V	W	Y						
18	A	F	H	K	L	N	P	Q	R	S	T	W	Y					
20	Q	T																
21	E	F	I	P	V	W	Y											
22	A	D	E	F	H	I	K	N	P	Q	R	S	T	V	W			
24	R	A	C	F	G	I	L	M	P	V	W	Y						
26	L	F	I	M	V	W	Y											
31	D	H																
34	A	H	K	N	P	Q	R	S	T	V	W	Y						
35	E	A	C	G	P													
36	M	W	Y															
37	N	A	C	G	P													
38	E	A	C	G	P													
42	V	A	C	D	E	G	H	K	M	N	P	Q	R	S	T	W		
45	E	A	C	F	G	I	L	M	P	V	W	Y						
47	F	W																
49	L	W	Y															
50	Q	P																
60	E	A	C	G	P													
61	L	F	I	M														
63	K	A	C	G	I	M	P	Y										
64	Q	A	C	G	P													
66	L	F	I	M	V													
67	R	A	C	G	P													
69	S	T																
70	L	M	W															
71	T	A	C	G	P													
72	K	T																
74	K	T																
75	G	H	P															
77	L	F	I	W	Y													
78	T	A	C	G	P	W	Y											
82	S	A	C	F	G	M	P	V	W	Y								
85	K	H	P															
87	H	A	C	F	G	I	M	P	W	Y								
88	C	D	E	H	K	N	P	Q	R	S	T	W						
109	N	T																
121	C	P	Y															
122	W	T																

**[0035]** Die Erfindung betrifft GM-CSF-Analoga, in denen Substitutionen von mindestens einem Aminosäurerest an Positionen gemacht wurden, die zu einer wesentlichen Verringerung der Aktivität von oder zur Entfernung eines oder mehrerer möglicher T-Zell-Epitope aus dem Protein führen. Eine oder mehrere Aminosäuresubstitutionen an bestimmten Punkten innerhalb irgendeinem der möglichen MHC-Klasse-II-Liganden, die in Tabelle 1 identifiziert wurden, können zu einem GM-CSF-Molekül mit einem verringertem immunogenen Potenzial bei Verabreichung als Therapeutikum an den menschlichen Wirt führen. Vorzugsweise werden die Aminosäuresubstitutionen an geeigneten Punkten innerhalb der Peptidsequenz gemacht, von denen das Erreichen einer wesentlichen Verringerung oder Entfernung der Aktivität des T-Zell-Epitops vorausgesagt wurde. In der Ausführung wird ein geeigneter Punkt vorzugsweise mit einer Aminosäurerestbindung innerhalb einer der Taschen übereinstimmen, die innerhalb der MHC-Klasse-II-Bindungsfurche bereitgestellt ist.

**[0036]** Am meisten ist bevorzugt, die Bindung innerhalb der ersten Tasche des Spalts an der so genannten P1- oder P1-Anker-Position des Peptids zu verändern. Die Qualität der Bindungswechselwirkung zwischen dem P1-Ankerrest des Peptids und der ersten Tasche der MHC-Klasse-II-Bindungsfurche gilt als eine Hauptdeterminante der Gesamtbindungsaffinität des gesamten Peptids. Eine geeignete Substitution an dieser Stelle des Peptids wird für einen innerhalb der Tasche weniger bereitwillig aufgenommenen Rest sein, zum Beispiel Substitution durch einen hydrophileren Rest. Aminosäurereste in dem Peptid an Positionen, die mit Bindung innerhalb anderer Taschenregionen innerhalb des MHC-Bindungsspalts übereinstimmen, werden auch betrachtet und fallen unter den Bereich der vorliegenden.

**[0037]** Selbstverständlich sind einzelne Aminosäuresubstitutionen innerhalb eines gegebenen möglichen T-Zell-Epitops der am meisten bevorzugte Weg, durch den das Epitop entfernt werden kann. Kombinationen aus Substitution innerhalb eines einzelnen Epitops können in Erwägung gezogen werden und können zum Beispiel insbesondere geeignet sein, wo individuell definierte Epitope miteinander überlappen. Darüber hinaus können Aminosäuresubstitutionen entweder einzeln innerhalb eines gegebenen Epitops oder in Kombination innerhalb eines einzelnen Epitops an Positionen gemacht werden, die nicht mit den „Taschenresten“ bezogen auf die MHC-Klasse-II-Bindungsfurche übereinstimmen, sondern an jeglichem Punkt innerhalb der Peptidsequenz. Substitutionen können bezogen auf eine homologe Struktur oder Strukturverfahren gemacht werden, die unter Verwendung von auf dem Fachgebiet bekannten In-silico-Techniken erzeugt werden, basierend auf bekannten Strukturmerkmalen des Moleküls gemäß dieser Erfindung. Alle derartigen Substitutionen fallen in den Bereich der vorliegenden Erfindung.

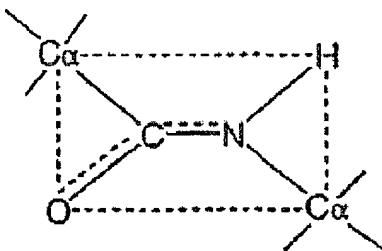
**[0038]** Andere Aminosäuresubstitutionen als innerhalb der oben identifizierten Peptide können insbesondere in Erwägung gezogen werden, wenn sie im Zusammenhang mit Substitution(en) innerhalb eines aufgelisteten Peptids gemacht werden. Zum Beispiel kann eine Veränderung in Erwägung gezogen werden, um die Struktur oder biologische Aktivität des Variantenmoleküls wiederherzustellen. Derartige ausgleichende Veränderungen und Veränderungen, um Deletion oder Addition eines bestimmten Aminosäurerestes aus dem GM-CSF-Polypeptid zu umfassen, die in einer Variante mit gewünschter Aktivität resultieren, und zusammen mit Veränderungen in jeglichem der offenbarten Peptide fallen unter den Bereich der vorliegenden.

**[0039]** In sofern als diese Erfindung modifizierten GM-CSF betrifft, sollten Zusammensetzungen, die derartige modifizierte GM-CSF-Proteine oder Fragmente von modifizierten GM-CSF-Proteinen enthalten, und verwandte Zusammensetzungen als innerhalb des Bereichs der Erfindung betrachtet werden. In einem anderen Aspekt betrifft die vorliegende Erfindung Nucleinsäuren, die modifizierte GM-CSF-Einheiten kodieren. In einem weiteren Aspekt betrifft die vorliegende Erfindung Verfahren zur therapeutischen Behandlung von Menschen unter Verwendung von modifizierten GM-CSF-Proteinen.

#### BEISPIEL

**[0040]** Es gibt eine Anzahl an Faktoren, die wichtige Rollen bei der Bestimmung der Gesamtstruktur eines Proteins oder Polypeptids spielen. Zunächst ist die Peptidbindung, d.h. die Bindung, die die Aminosäuren in der Kette zusammenfügt, eine kovalente Bindung. Diese Bindung ist in ihrer Struktur planar, im Wesentlichen ein substituiertes Amid. Ein „Amid“ ist eines aus einer Gruppe von organischen Verbindungen, die die -CONH-Gruppierung enthalten.

**[0041]** Die planare Peptidbindung, die  $C\alpha$  der benachbarten Aminosäuren verbindet, kann wie unten dargestellt, dargestellt werden:



**[0042]** Da die O=C- und die C-N-Atome in einer relativ steifen Ebene liegen, tritt um diese Achsen herum keine freie Drehung auf. Daher wird eine Ebene, wie sie schematisch durch die gestrichelte Linie dargestellt wird, manchmal als eine „Amid-“ oder „Peptidebene“ bezeichnet, eine Ebene, in der Sauerstoff-(O), Kohlenstoff-(C), Stickstoff-(N) und Wasserstoff-(H)-Atome des Peptidgrundgerüsts liegen. An gegenüberliegenden Ecken dieser Amidebene befinden sich die  $C\alpha$ -Atome. Da es im Wesentlichen keine Rotation um die O=C- und C-N-Atome

me in der Peptid- oder Amidebene gibt, enthält eine Polypeptidkette so eine Reihe von planaren Peptidverknüpfungen, die die  $C\alpha$ -Atome verbinden. Ein zweiter Faktor, der eine wichtige Rolle bei der Definition der Gesamtstruktur oder Konformation eines Polypeptids oder Proteins spielt, ist der Rotationswinkel von jeder Amidebene bezogen auf die gemeinsame  $C\alpha$ -Verknüpfung. Die Begriffe „Rotationswinkel“ und „Torsionswinkel“ werden hier als äquivalente Begriffe betrachtet. Unter der Annahme, dass O-, C-, N- und H-Atome in der Amidebene verbleiben (was üblicherweise eine gültige Annahme ist, obwohl es einige geringfügige Abweichungen von der Planarität dieser Atome für einige Konformationen geben kann) definieren diese Rotationswinkel die N- und R-Konformation des Peptidgrundgerüsts, d.h. die Struktur, so wie sie zwischen benachbarten Resten vorliegt. Diese beiden Winkel sind bekannt als  $\Phi$  und  $\psi$ . Ein Satz dieser Winkel  $\Phi_i$ ,  $\psi_i$ , wobei der tiefgestellte Index  $i$  einen bestimmten Rest einer Polypeptidkette darstellt, definiert so effizient die Sekundärstruktur des Polypeptids. Die bei der Definition der  $\Phi$ -,  $\psi$ -Winkel verwendete Konvention, d.h. die Referenzpunkte, an denen die Amidebenen einen Null-Grad-Winkel bilden, und die Definition, welcher Winkel  $\Phi$  ist und welcher Winkel  $\psi$  ist, für ein gegebenes Polypeptid, werden in der Literatur definiert. Siehe, z. B., Ramachandran et al. Adv. Prot. Chem. 23: 283–437 (1968), auf Seiten 285–94, wobei diese Seiten durch Literaturhinweis hier aufgenommen sind.

**[0043]** Das vorliegende Verfahren kann auf jegliches Protein angewandt werden und basiert zum Teil auf der Entdeckung, dass bei Menschen die primäre Tasche-1-Ankerposition der MHC-Klasse-II-Bindungsfurchen eine gut entworfene Spezifität für bestimmte Aminosäureseitenketten haben. Die Spezifität dieser Tasche wird durch die Identität der Aminosäure an Position 86 der beta-Kette des MHC-Klasse-II-Moleküls bestimmt. Diese Stelle befindet sich am Boden der Tasche 1 und bestimmt die Größe der Seitenkette, die durch diese Tasche aufgenommen werden kann. Marshall, K.W., J. Immunol., 152: 4946–4956 (1994). Wenn dieser Rest ein Glycin ist, dann können alle hydrophoben aliphatischen und aromatischen Aminosäuren (hydrophobe aliphatische sind: Valin, Leucin, Isoleucin, Methionin und aromatische sind: Phenylalanin, Tyrosin und Tryptophan) in die Tasche aufgenommen werden, wobei aromatische Seitenketten bevorzugt werden. Ist dieser Taschenrest ein Valin, dann ragt die Seitenkette dieser Aminosäure in die Tasche hinein und begrenzt die Größe der Peptidseitenketten, die aufgenommen werden können, so dass nur hydrophobe aliphatische Seitenketten aufgenommen werden können. Daher gibt es in einer Aminosäurerestsequenz, wo immer eine Aminosäure mit einem hydrophoben aliphatischen oder aromatischen Seitenkette gefunden wird, die Möglichkeit für die Anwesenheit eines MHC-Klasse-II-beschränkten T-Zell-Epitops. Ist die Seitenkette hydrophob aliphatisch, ist es jedoch etwa zweimal so wahrscheinlich, dass sie mit einem T-Zell-Epitop einhergeht, als bei einer aromatischen Seitenkette (unter der Annahme einer in etwa gleichen Verteilung der Tasche-1-Arten in der globalen Population).

**[0044]** Ein rechengestütztes Verfahren zur Ausführung der vorliegenden Erfindung beschreibt die Wahrscheinlichkeit von Peptidregionen, T-Zell-Epitope zu enthalten, wie folgt: (1) Die Primärsequenz eines Peptidsegments von vorbestimmter Länge wird gescannt und alle vorhandenen, hydrophoben aliphatischen und aromatischen Seitenketten werden identifiziert. (2) Den hydrophoben aliphatischen Seitenketten wird ein größerer Wert zugeordnet als den aromatischen Seitenketten, vorzugsweise etwa ein zweimal so großer Wert wie der den aromatischen Seitenketten zugewiesene, z. B., ein Wert von 2 für eine hydrophobe aliphatische Seitenkette und ein Wert von 1 für eine aromatische Seitenkette. (3) Der als vorliegend bestimmte Wert wird für jedes überlappende Aminosäurerestsegment (Fenster) von vorbestimmter einheitlicher Länge innerhalb des Peptids aufsummiert und der Gesamtwert für ein bestimmtes Segment (Fenster) wird einem einzelnen Aminosäurerest an einer intermediären Position des Segments (Fenster) zugeordnet, vorzugsweise einem Rest in etwa am Mittelpunkt des abgetasteten Segments (Fenster). Diese Prozedur wird für jedes abgetastete, überlappende Aminosäurerestsegment (Fenster) wiederholt. So wird jedem Aminosäurerest des Peptids ein Wert zugeordnet, der in Bezug steht zur Wahrscheinlichkeit, dass ein T-Zell-Epitop in diesem bestimmten Segment (Fenster) anwesend ist. (4) Die oben wie in Schritt 3 beschrieben, berechneten und zugewiesenen Werte können gegen die Aminosäure-Koordinaten der gesamten untersuchten Aminosäurerestsequenz aufgetragen werden. (5) Von allen Teilen der Sequenz, die einen Wert von einem vorbestimmten Wert, z. B., einen Wert von 1, haben, wird angenommen, dass sie ein T-Zell-Epitop enthalten, und können dann, wenn gewünscht, modifiziert werden.

**[0045]** Dieser besondere Aspekt der vorliegenden Erfindung stellt ein allgemeines Verfahren bereit, durch das die Regionen von Peptiden, die wahrscheinlich T-Zell-Epitope enthalten, beschrieben werden können. Modifizierungen an dem Peptid in diesen Regionen haben das Potenzial, die MHC-Klasse-II-Bindungseigenschaften zu modifizieren.

**[0046]** Gemäß einem weiteren Aspekt der vorliegenden Erfindung können T-Zell-Epitope mit größerer Genauigkeit vorausgesagt werden, indem ein technisch ausgereifteres, rechengestütztes Verfahren verwendet

wird, das die Wechselwirkungen der Peptide mit Modellen von MHC-Klasse-II-Allelen berücksichtigt. Die rechengestützte Vorhersage von T-Zell-Epitopen, die innerhalb eines Peptids vorkommen, gemäß diesem besonderen Aspekt, zieht die Konstruktion von Modellen von mindestens 42 MHC-Klasse-II-Allelen in Erwägung, basierend auf den Strukturen aller bekannten MHC-Klasse-II-Molekülen und einem Verfahren zur Verwendung dieser Modelle bei der rechengestützten Identifizierung von T-Zell-Epitopen, die Konstruktion von Peptidgrundgerüstbibliotheken für jedes Modell, um die bekannte Variabilität in den jeweiligen Peptidgrundgerüst- $\alpha$ -Kohlenstoff-(C $\alpha$ )-Positionen zu berücksichtigen, die Konstruktion von Aminosäureseitenketten-Konformationsbibliotheken für jede Grundgerüstankopplungsstelle mit jedem Modell für jede der 20 Aminosäurealternativen an Positionen, die für die Wechselwirkung zwischen Peptid und MHC-Klasse-II-Molekül kritisch ist, und die Verwendung dieser Bibliotheken an Grundgerüsten und Seitenkettenkonformationen zusammen mit einer Bewertungsfunktion, um das optimale Grundgerüst und die optimale Seitenkettenkonformation für ein bestimmtes Peptid auszuwählen, das mit einem bestimmten MHC-Klasse-II-Molekül gekoppelt ist, und die Ableitung eines Bindungswerts aus dieser Wechselwirkung.

**[0047]** Modelle von MHC-Klasse-II-Molekülen können über Homologie-Modellierung aus einer Anzahl an ähnlichen Strukturen abgeleitet werden, die in der Brookhaven Protein Data Bank ("PDB") gefunden werden. Diese können durch Verwendung einer halbautomatischen Homologie-Modellierungssoftware (Modeller, Sali A. & Blundel T.L., 1993. *J. Mol Biol* 234: 779–815) hergestellt werden, die eine simulierte Hybridisierungsfunktion zusammen mit der CHARMM Force-Field-For-Energy-Minimierung einbezieht (erhältlich von Molecular Simulations Inc., San Diego, Ca.). Es können auch alternative Modellierungsverfahren verwendet werden.

**[0048]** Das vorliegende Verfahren unterscheidet sich deutlich von anderen rechengestützten Verfahren, bei denen Bibliotheken von experimentell abgeleiteten Bindungsdaten von jeder Aminosäurealternative an jeder Position in der Bindungsfurche für einen kleinen Satz an MHC-Klasse-II-Molekülen verwendet werden (Marshall, K.W., et al., *Biomed. Pept. Proteins Nucleic Acids*, 1(3): 157–162) (1995), oder noch anderen rechengestützten Verfahren, bei denen ähnliche experimentelle Bindungsdaten verwendet werden, um die Bindungseigenschaften von bestimmten Arten von Bindungstaschen innerhalb der Furche zu definieren, wiederum unter Verwendung eines relativ kleinen Subsets an MHC-Klasse-II-Molekülen, und bei denen dann Taschenarten aus dieser Taschenbibliothek „gemischt und gematched“ werden, um künstlich weitere „virtuelle“ MHC-Klasse-II-Moleküle zu erzeugen (Sturniolo T., et al., *Nat. Biotech*, 17(6): 555–561) (1999). Beide vorgenannten Verfahren leiden an dem Hauptnachteil, dass aufgrund der Komplexität der Assays und der Notwendigkeit, eine große Anzahl an Peptidvarianten zu synthetisieren, nur eine kleine Anzahl an MHC-Klasse-II-Molekülen experimentell gescannt werden kann. Daher kann das erste Verfahren nach dem Stand der Technik nur Vorhersagen für eine kleine Anzahl an MHC-Klasse-II-Molekülen treffen. Das zweite Verfahren nach dem Stand der Technik geht auch von der Annahme aus, dass eine Tasche, die mit ähnlichen Aminosäuren ausgekleidet ist, in einem Molekül die gleichen Bindungseigenschaften haben wird, wenn sie sich zusammen mit einem anderen Klasse-II-Allel zusammen befindet, und leidet weiterhin an Nachteilen, dass nur solche MHC-Klasse-II-Moleküle „virtuell“ erzeugt werden können, die Taschen enthalten, die innerhalb der Taschenbibliothek enthalten sind. Bei Verwendung des hier beschriebenen Modellierungsansatzes kann die Struktur von jeglicher Anzahl und Art von MHC-Klasse-II-Molekülen hergeleitet werden, daher können Allele spezifisch ausgewählt werden, um für die globale Population typisch zu sein. Zusätzlich kann die Anzahl an gescannten MHC-Klasse-II-Molekülen erhöht werden, indem weitere Modelle gemacht werden, ohne dass zusätzliche Daten über komplizierte Experimente erzeugt werden müssen.

**[0049]** Die Verwendung einer Grundgerüstbibliothek berücksichtigt die Variation in den Positionen des C $\alpha$ -Atoms der verschiedenen, gescannten Peptide, wenn diese mit bestimmten MHC-Klasse-II-Molekülen gekoppelt sind. Dies steht wiederum im Gegensatz zu den alternativen rechengestützten Verfahren nach dem Stand der Technik, die oben beschrieben wurden, die auf der Verwendung von vereinfachten Peptidgrundgerüsten für das Scannen der Aminosäurebindung in bestimmten Taschen beruhen. Diese vereinfachten Grundgerüste sind wahrscheinlich nicht typisch für Grundgerüstkonformationen, die in „realen“ Peptiden gefunden werden, was zu Ungenauigkeiten bei der Vorhersage der Peptidbindung führt. Die vorliegende Grundgerüstbibliothek wurde durch Überlagerung der Grundgerüste aller an MHC-Klasse-II-Moleküle gebundenen Peptide erzeugt, die innerhalb der Protein-Datenbank gefunden wurden, und durch Aufzeichnen des quadratischen Mittels (OMW) der Abweichung zwischen den C $\alpha$ -Atomen von jeder der elf Aminosäuren, die innerhalb der Bindungsfurche lokalisiert sind. Während diese Bibliothek aus einer kleinen Anzahl an geeigneten, erhältlichen Maus- und humanen Strukturen (gegenwärtig 13) abgeleitet werden kann, wird, um die Möglichkeit einer noch größeren Variabilität zu berücksichtigen, die QMW-Zahl für jede C $\alpha$ -Position um 50 % erhöht. Die mittlere C $\alpha$ -Position von jeder Aminosäure wird dann bestimmt und eine Sphäre um diesen Punkt herumgezogen, deren Radius der QMW-Abweichung an dieser Position plus 50 % entspricht. Diese Sphäre stellt alle erlaubten C $\alpha$ -Positionen dar.

**[0050]** Ausgehend von dem C $\alpha$  mit der geringsten OMW-Abweichung (die von der oben erwähnten Aminosäure in Tasche 1, entspricht Position 2 der 11 Reste in der Bindungsfurche), wird die Sphäre in ein dreidimensionales Gitter unterteilt und jeder Scheitelpunkt innerhalb des Gitters wird dann als ein möglicher Ort für ein C $\alpha$  dieser Aminosäure verwendet. Die nachfolgende Amidebene, entsprechend der Peptidbindung an die nachfolgende Aminosäure, wird auf jede dieser Cas gepfropft und die  $\Phi$ - und  $\psi$ -Winkel werden schrittweise um festgelegte Intervalle gedreht, um das nachfolgende C $\alpha$  zu positionieren. Wenn das nachfolgende C $\alpha$  in die „Sphäre der erlaubten Positionen“ für dieses C $\alpha$  fällt, dann wird die Orientierung des Dipeptids akzeptiert, wohingegen dieses Dipeptid verworfen wird, wenn es außerhalb der Sphäre fällt.

**[0051]** Dieser Prozess wird dann für jede der nachfolgenden C $\alpha$ -Positionen wiederholt, so dass das Peptid aus dem C $\alpha$ -Keim der Tasche 1 herauswächst, bis alle neun nachfolgenden Cas aus allen möglichen Permutationen der vorhergehenden Cas positioniert wurden. Dieser Prozess wird dann ein weiteres Mal für das einzelne C $\alpha$ , das Tasche 1 vorausgeht, wiederholt, um eine Bibliothek von Grundgerüst-C $\alpha$ -Positionen, die innerhalb der Bindungsfurche lokalisiert sind, zu erzeugen. Die Anzahl an erzeugten Grundgerüsten hängt von mehreren Faktoren ab: Der Größe der „Sphäre an erlaubten Positionen“; der Feinheit der Gitterunterteilung der „primären Sphäre“ an der Tasche-1-Position; der Feinheit der schrittweisen Drehung der  $\Phi$ - und  $\psi$ -Winkel, die für das Positionieren der nachfolgenden Cas verwendet werden. Unter Verwendung dieses Prozesses kann eine große Bibliothek an Grundgerüsten geschaffen werden. Je größer die Grundgerüstbibliothek, desto wahrscheinlicher wird es, dass eine optimale Passung für ein bestimmtes Peptid innerhalb der Bindungsfurche eines MHC-Klasse-II-Moleküls gefunden wird. In sofern, dass nicht alle Grundgerüste für ein Ankoppeln an alle Modelle der MHC-Klasse-II-Moleküle geeignet sein werden, aufgrund von Zusammenstößen mit Aminosäuren der Bindungsdomänen, wird für jedes Allel ein Subset der Bibliothek geschaffen, das Grundgerüste enthält, die durch dieses Allel aufgenommen werden können. Die Verwendung der Grundgerüstbibliothek zusammen mit den Modellen der MHC-Klasse-II-Molekülen schafft eine erschöpfende Datenbank, die aus erlaubten Seitenkettenkonformationen für jede Aminosäure in jeder Position der Bindungsfurche für jedes MHC-Klasse-II-Molekül besteht, das an jedes erlaubte Grundgerüst angekoppelt ist. Dieser Datensatz wird unter Verwendung einer einfachen sterischen Überlappungsfunktion erzeugt, bei der ein MHC-Klasse-II-Molekül an ein Grundgerüst gekoppelt ist und eine Aminosäureseitenkette auf das Grundgerüst an der gewünschten Position gepfropft wird. Jede der drehbaren Bindungen der Seitenkette wird schrittweise in festgelegten Intervallen gedreht und die resultierenden Positionen der Atome abhängig von dieser untersuchten Bindung. Die Wechselwirkung des Atoms mit Atomen der Seitenketten der Bindungsfurche wird festgestellt und Positionen werden entweder akzeptiert oder verworfen, gemäß der folgenden Kriterien: Die Gesamtsumme der Überlappung aller bisher positionierten Atome darf einen vorbestimmten Wert nicht übersteigen. So ist die Genauigkeit der Konformationsuche eine Funktion des Intervalls, das bei der schrittweisen Drehung der Bindung verwendet wird, und des vorbestimmten Grenzwerts für die vollständige Überlappung. Dieser letzte Wert kann klein sein, wenn bekannt ist, dass eine bestimmte Tasche steif ist, jedoch kann die Genauigkeit gelockert werden, wenn bekannt ist, dass die Positionen der Taschenseitenketten relativ flexibel sind. So können Zugeständnisse gemacht werden, um die Variationen bei der Flexibilität innerhalb der Taschen der Bindungsfurche zu imitieren. Diese Konformationsuche wird dann für jede Aminosäure an jeder Position von jedem Grundgerüst wiederholt, wenn es an jedes der MHC-Klasse-II-Moleküle gekoppelt ist, um eine erschöpfende Datenbank der Seitenkettenkonformationen zu schaffen.

**[0052]** Es wird ein geeigneter mathematischer Ausdruck verwendet, um die Bindungsenergie zwischen Modellen von MHC-Klasse-II-Molekülen in Zusammenhang mit Peptid-Liganden-Konformationen abzuschätzen, die empirisch durch Scannen der großen, oben beschriebenen Datenbank an Grundgerüst-/Seitenkettenkonformationen abgeleitet werden müssen. So wird ein Protein auf mögliche T-Zell-Epitope gescannt, indem jedes mögliche Peptid einer Länge, die zwischen 9 und 20 Aminosäuren variiert (obwohl die Länge für jeden Scan konstant gehalten wird) den folgenden Berechnungen unterzogen wird: Ein MHC-Klasse-II-Molekül wird zusammen mit einem Peptidgrundgerüst ausgewählt, unter Berücksichtigung, dass das Molekül und die Seitenketten, die der gewünschten Peptidsequenz entsprechen, aufgepfropft sind. Atomidentität und interatomare Abstandsdaten bezüglich einer bestimmten Seitenketten an einer bestimmten Position auf dem Grundgerüst werden für jede erlaubte Konformation dieser Aminosäure (erhalten aus der oben beschriebenen Datenbank) gesammelt. Dies wird für jede Seitenkette entlang des Grundgerüsts wiederholt und Peptidpunktzahlen werden unter Verwendung einer Bewertungsfunktion abgeleitet. Die beste Punktzahl für dieses Grundgerüst wird zurückgehalten und der Prozess wird für jedes erlaubte Grundgerüst für das ausgewählte Modell wiederholt. Die Punktzahlen aus allen erlaubten Grundgerüsten werden verglichen und die höchste Punktzahl wird als Peptidpunktzahl für das gewünschte Peptid in diesem MHC-Klasse-II-Modell erachtet. Dieser Prozess wird dann für jedes Modell mit jedem möglichen Peptid, das aus dem gescannten Protein abgeleitet wurde, wiederholt und die Punktzahlen für Peptide versus Modelle werden angezeigt.

**[0053]** Im Zusammenhang der vorliegenden Erfindung ist jeder für die Bindungsaffinitätsberechnung vorgelegte Ligand ein Aminosäuresegment, das aus einem Peptid oder Protein ausgewählt wird, wie oben diskutiert. So ist der Ligand ein ausgewählter Strang an Aminosäuren mit etwa 9 bis 20 Aminosäuren an Länge, abgeleitet von einem Peptid, Polypeptid oder Protein von bekannter Sequenz. Die Begriffe „Aminosäuren“ und „Reste“ werden hier im Folgenden als äquivalente Begriffe betrachtet. Der Ligand, in der Form der aufeinander folgenden Aminosäuren des zu untersuchenden Peptids, das aus der Grundgerüstbibliothek auf ein Grundgerüst gepfropft wurde, wird in den Bindungsspalt eines MHC-Klasse-II-Moleküls aus der MHC-Klasse-II-Molekülbibliothek über die Koordinaten der  $\text{C}\alpha$ -Atome auf dem Peptidgrundgerüst positioniert und eine erlaubte Konformation für jede Seitenkette wird aus der Datenbank der erlaubten Konformationen ausgewählt. Die relevanten Atomidentitäten und interatomaren Abstände werden ebenfalls aus dieser Datenbank bezogen und verwendet, um die Peptidbindungszahl zu berechnen. Liganden mit einer höheren Bindungsaffinität für die MHC-Klasse-II-Bindungsfurche werden als Kandidaten für eine ortsgerichtete Mutagenese markiert. Aminosäuresubstitutionen werden in dem markierten Liganden vorgenommen (und so in dem interessierenden Protein), welches dann erneut unter Verwendung der Bewertungsfunktion getestet wird, um Veränderungen zu bestimmen, die die Bindungsaffinität unter einen vorbestimmten Schwellenwert verringern. Diese Veränderungen können dann in das interessierende Protein eingebaut werden, um T-Zell-Epitope zu entfernen.

**[0054]** Bindung zwischen dem Peptidliganden und der Bindungsfurche von MHC-Klasse-II-Molekülen umfasst nicht-kovalente Wechselwirkung, einschließlich, aber nicht beschränkt auf: Wasserstoffbindungen, elektrostatische Wechselwirkungen, hydrophobe (lipophile) Wechselwirkungen und Van-der-Waals-Wechselwirkungen. Diese sind in der Peptid-Bewertungsfunktion eingeschlossen, wie unten ausführlich beschrieben. Es sollte selbstverständlich sein, dass eine Wasserstoffbindung eine nicht-kovalente Bindung ist, die zwischen polaren oder geladenen Gruppen gebildet werden kann und aus einem Wasserstoffatom besteht, das von zwei anderen Atomen geteilt wird. Der Wasserstoff des Wasserstoffdonors hat eine positive Ladung, während der Wasserstoffakzeptor eine partielle negative Ladung hat. Bei Peptid/Protein-Wechselwirkungen können Wasserstoffbindungsdonoren entweder Stickstoffe mit gebundenem Wasserstoff oder Wasserstoffe gebunden an Sauerstoff oder Stickstoff sein. Wasserstoffbindungsakzeptoratome können nicht an Wasserstoff gebundene Sauerstoffe sein, Stickstoffe ohne gebundene Wasserstoffe und ein oder zwei Bindungen, oder Schwefel mit nur einer Bindung. Bestimmte Atome wie an Wasserstoffe gebundene Sauerstoffe oder Iminstickstoffe (z. B.  $\text{C}=\text{NH}$ ) können sowohl Wasserstoffakzeptoren wie auch -donoren sein. Wasserstoffbindungsenergien betragen zwischen 3 und 7 kcal/mol und sind viel stärker als Van-der-Waals-Bindungen, aber schwächer als kovalente Bindungen. Wasserstoffbindungen sind auch hoch gerichtet und sind dann am stärksten, wenn das Donoratom, Wasserstoffatom und Akzeptoratom kollinear sind. Elektrostatische Bindungen werden zwischen entgegengesetzt geladenen Innenpaaren gebildet und die Stärke der Wechselwirkung ist gemäß dem Coulomb-Gesetz umgekehrt proportional zum Quadrat des Abstands zwischen den Atomen. Der optimale Abstand zwischen Ionenpaaren beträgt etwa 2,8 Å. In Protein/Peptid-Wechselwirkungen können elektrostatische Bindungen zwischen Arginin, Histidin oder Lysin und Aspartat oder Glutamat gebildet werden. Die Stärke der Bindung wird vom pKa der ionisierenden Gruppe und der Dielektrizitätskonstante des Mediums abhängen, obwohl sie annähernd gleichwertig in Stärke zu Wasserstoffbindungen sind.

**[0055]** Lipophile Wechselwirkungen sind günstige hydrophobe-hydrophobe Kontakte, die zwischen dem Protein und dem Peptidliganden auftreten. Gewöhnlich werden diese zwischen hydrophoben Aminosäure-Seitenketten des Peptids auftreten, das innerhalb der Taschen der Bindungsfurche begraben ist, so dass sie keinem Lösungsmittel ausgesetzt sind. Es ist äußerst ungünstig, die hydrophoben Reste Lösungsmittel auszusetzen, da die umgebenden Lösungsmittelmoleküle dazu gezwungen werden, miteinander Wasserstoffbindungen einzugehen und dadurch käfigartige Clathratstrukturen zu bilden. Die resultierende Entropieabnahme ist äußerst ungünstig. Lipophile Atome können Schwefel sein, die weder polar noch Wasserstoffakzeptoren sind, und Kohlenstoffatome, die unpolar sind.

**[0056]** Van-der-Waals-Bindungen sind unspezifische Kräfte, die zwischen Atomen gefunden werden, die 3–4 Å voneinander entfernt sind. Sie sind schwächer und weniger spezifisch als Wasserstoff- und elektrostatische Bindungen. Die Verteilung der elektrischen Ladung um ein Atom verändert sich mit der Zeit und in jedem Augenblick ist die Ladungsverteilung unsymmetrisch. Diese vorübergehende Asymmetrie in der elektronischen Ladung induziert eine gleichartige Asymmetrie in benachbarten Atomen. Die resultierenden Anziehungskräfte zwischen Atomen erreichen ein Maximum am Van-der-Waals-Kontaktabstand, aber nehmen sehr schnell bei etwa 1 Å bis etwa 2 Å ab. Umgekehrt werden, wenn Atome weniger als der Kontaktabstand voneinander getrennt werden, immer stärker werdende Abstoßungskräfte dominierend, da sich die äußeren Elektronenwolken der Atome überlagern. Obwohl die Anziehungskräfte relativ schwach sind, verglichen mit elektrostatischen und Wasserstoffbindungen (etwa 0,6 kcal/mol), können insbesondere die Abstoßungskräfte sehr wichtig bei der Bestimmung werden, ob ein Peptidligand erfolgreich an ein Protein binden kann.

**[0057]** In einer Ausführungsform wird die Böhm-Bewertungsfunktion (SCORE1-Ansatz) verwendet, um die Bindungskonstante abzuschätzen. (Böhm, H.J., J. Comput Aided Mol Des., 8(3): 243–256 (1994), welches hiermit in seiner Gesamtheit aufgenommen ist). In einer anderen Ausführungsform, wird die Bewertungsfunktion (SCORE2-Ansatz) verwendet, um die Bindungsaffinitäten als einen Indikator für einen Liganden abzuschätzen, der ein T-Zell-Epitop enthält (Böhm, H.J., J. Comput Aided Mol. Des., 12(4): 309–323 (1998), welches hiermit in seiner Gesamtheit aufgenommen ist). Die Böhm-Bewertungsfunktionen, wie sie in den oben genannten Literaturzitate beschrieben werden, werden jedoch verwendet, um die Bindungsaffinität eines Liganden an ein Protein abzuschätzen, wo es bereits bekannt ist, dass der Ligand erfolgreich an das Protein bindet, und die Struktur des Protein/Ligand-Komplex gelöst wurde, wobei die gelöste Struktur in der Protein Data Bank ("PDB") vorliegt. Daher wurde die Bewertungsfunktion mit dem Nutzen von bekannten positiven Bindungsdaten entwickelt. Um eine Unterscheidung zwischen positiven und negativen Binder zu berücksichtigen, muss zu der Gleichung ein Abstoßungsterm hinzugefügt werden. Zusätzlich wird eine befriedigendere Abschätzung der Bindungsenergie eher durch Berechnung der lipophilen Wechselwirkungen auf eine paarweise Art erreicht, als durch Verwendung des gebietsbasierenden Energieterms der oben genannten Böhm-Funktionen. Daher wird in einer bevorzugten Ausführungsform die Bindungsenergie unter Verwendung einer modifizierten Böhm-Bewertungsfunktion abgeschätzt. In der modifizierten Böhm-Bewertungsfunktion wird die Bindungsenergie zwischen Protein und Ligand ( $\Delta G_{\text{Bind}}$ ) unter Berücksichtigung der folgenden Parameter abgeschätzt: Der Verringerung der Bindungsenergie aufgrund des allgemeinen Verlustes an Translations- und Rotationsentropie des Liganden ( $\Delta G_0$ ); Beiträgen von idealen Wasserstoffbindungen ( $\Delta G_{\text{WB}}$ ), wobei mindestens ein Partner neutral ist; Beiträgen von ungestörten ionischen Wechselwirkungen ( $\Delta G_{\text{ionisch}}$ ); lipophile Wechselwirkungen zwischen lipophilen Ligandenatomen und lipophilen Akzeptoratomen ( $\Delta G_{\text{lipo}}$ ); dem Verlust an Bindungsenergie aufgrund des Einfrierens von internen Freiheitsgraden in dem Ligand, d.h. der Rotationsfreiheit um jede C-C-Bindung wird verringert ( $\Delta G_{\text{rot}}$ ); der Energie der Wechselwirkung zwischen Protein und Ligand ( $E_{\text{vdW}}$ ).

**[0058]** Berücksichtigung dieser Terme ergibt Gleichung 1:

$$(\Delta G_{\text{Bind}}) = (\Delta G_0) + (\Delta G_{\text{WB}} \times N_{\text{WB}}) + (\Delta G_{\text{ionisch}} \times N_{\text{ionisch}}) + (\Delta G_{\text{lipo}} \times N_{\text{lipo}}) + (\Delta G_{\text{rot}} + N_{\text{rot}}) + (E_{\text{vdW}}).$$

wobei N die Anzahl an qualifizierten Wechselwirkungen für einen spezifischen Term ist und in einer Ausführungsform sind  $\Delta G_0$ ,  $\Delta G_{\text{WB}}$ ,  $\Delta G_{\text{ionisch}}$ ,  $\Delta G_{\text{lipo}}$  und  $\Delta G_{\text{rot}}$  konstant, welchen die Werte gegeben wurden: 5,4, -4,7, -4,7, -0,17 bzw. 1,4.

**[0059]** Der Term  $N_{\text{WB}}$  wird gemäß Gleichung 2 berechnet:

$$N_{\text{WB}} = \sum_{\text{W-Bindungen}} f(\Delta R, \Delta \alpha) \times f(N_{\text{Nachb.}}) \times f_{\text{pcs}}$$

$f(\Delta R, \Delta \alpha)$  ist eine Straffunktion, die starke Abweichungen von Wasserstoffbindungen von der Idealität berücksichtigt, und wird gemäß Gleichung 3 berechnet:

$$f(\Delta R, \Delta - \alpha) = f1(\Delta R) \times f2(\Delta \alpha)$$

wobei:  $f1(\Delta R) = 1$  wenn  $\Delta R \leq \text{TOL}$   
 oder  $= 1 - (\Delta R - \text{TOL})/0,4$  wenn  $\Delta R \leq 0,4 + \text{TOL}$   
 oder  $= 0$  wenn  $\Delta R > 0,4 + \text{TOL}$   
 und:  $f2(\Delta \alpha) = 1$  wenn  $\Delta \alpha < 30^\circ$   
 oder  $= 1 - (\Delta \alpha - 30)/50$  wenn  $\Delta \alpha \leq 80^\circ$   
 oder  $= 0$  wenn  $\Delta \alpha > 80^\circ$

**[0060]** TOL ist die tolerierte Abweichung in der Wasserstoffbindungslänge = 0,25 Å.

**[0061]**  $\Delta R$  ist die Abweichung der H-O/N-Wasserstoffbindungslänge vom Idealwert = 1,9 Å.  $\Delta \alpha$  ist die Abweichung des Wasserstoffbindungswinkels  $\angle_{\text{N/O-H...O/N}}$  von seinem idealisierten Wert von  $180^\circ$ .

**[0062]**  $f(N_{\text{Nachb.}})$  unterscheidet zwischen konkaven und konvexen Teilen einer Proteinoberfläche und schreibt daher polaren, in Taschen gefundenen Wechselwirkungen ein größeres Gewicht zu als denen, die auf der Proteinoberfläche gefunden wurden. Diese Funktion wird berechnet gemäß der unten stehenden Gleichung 4:

$$f(N_{\text{Nachb.}}) = (N_{\text{Nachb.}}/N_{\text{Nachb.,0}})^\alpha, \text{ wobei } \alpha = 0,5$$

**[0063]**  $N_{\text{Nachb}}$  ist die Anzahl an Nichtwasserstoff-Proteinatomen, die näher als 5 Å zu irgendeinem gegebenen



Proteinatom stehen.

**[0064]**  $N_{\text{Nachb},0}$  ist eine Konstante = 25

**[0065]**  $f_{\text{pKO}}$  ist eine Funktion, die das polare Kontaktoberflächengebiet pro Wasserstoffbindung berücksichtigt und daher zwischen starken und schwachen Wasserstoffbindungen unterscheidet und ihr Wert wird gemäß den folgenden Kriterien bestimmt:

$f_{\text{pKO}} = \beta$ , wenn  $A_{\text{polar}}/N_{\text{WB}} < 10 \text{ \AA}^2$   
 oder  $f_{\text{pKO}} = 1$ , wenn  $A_{\text{polar}}/N_{\text{WB}} > 10 \text{ \AA}^2$

**[0066]**  $A_{\text{polar}}$  ist die Größe der polaren Protein-Ligand-Kontaktoberfläche.

**[0067]**  $N_{\text{WB}}$  ist die Anzahl an Wasserstoffbindungen.

**[0068]**  $\beta$  ist eine Konstante, deren Wert = 1,2

**[0069]** Für die Implementierung der modifizierten Böhm-Bewertungsfunktion werden die Beiträge von ionischen Wechselwirkungen,  $\Delta G_{\text{ionisch}}$ , auf gleichartige Weise berechnet wie die von oben beschriebenen Wasserstoffbindungen, da die gleiche Geometrieabhängigkeit angenommen wird.

**[0070]** Der Term  $N_{\text{lipo}}$  wird berechnet gemäß der unten stehenden Gleichung 5:

$$N_{\text{lipo}} = \sum_{1L} f(r_{1L})$$

**[0071]**  $f(r_{1L})$  wird für alle lipophilen Ligandenatome, l, berechnet und alle lipophilen Proteinatome, L, gemäß den folgenden Kriterien:

$f(r_{1L}) = 1$  wenn  $r_{1L} \leq R1$   $f(r_{1L}) = (r_{1L} - R1)/(R2 - R1)$  wenn  $R2 < r_{1L} < R1$

$f(r_{1L}) = 0$  wenn  $r_{1L} \geq R2$

wobei:  $R1 = r_1^{\text{vdW}} + r_L^{\text{vdW}} + 0,5$

und  $R2 = R1 + 3,0$

und  $r_1^{\text{vdW}}$  der Van-der-Waals-Radius von Atom l ist

und  $r_L^{\text{vdW}}$  der Van-der-Waals-Radius von Atom L ist

**[0072]** Der Term  $N_{\text{rot}}$  ist die Anzahl an drehbaren Bindungen der Aminosäure-Seitenkette und wird als Anzahl der acyclischen  $sp^3$ - $sp^3$ - und  $sp^3$ - $sp^2$ -Bindungen genommen.

**[0073]** Rotationen von terminalem  $-\text{CH}_3$  oder  $-\text{NH}_3$  werden nicht berücksichtigt.

**[0074]** Der endgültige Term  $E_{\text{vdW}}$  wird berechnet gemäß der unten stehenden Gleichung 6:

$$E_{\text{vdW}} = \epsilon_1 \epsilon_2 \left( (r_1^{\text{vdW}} + r_2^{\text{vdW}})^{12} / r^{12} - \{ r_1^{\text{vdW}} + r_2^{\text{vdW}} \}^6 / r^6 \right),$$

wobei:

$\epsilon_1$  und  $\epsilon_2$  von der Atomidentität abhängige Konstanten sind

$r_1^{\text{vdW}} + r_2^{\text{vdW}}$  Van-der-Waals-Atomradien sind

r der Abstand zwischen einem Paar von Atomen ist.

**[0075]** In Bezug auf Gleichung 6 werden in einer Ausführungsform den Konstanten  $\epsilon_1$  und  $\epsilon_2$  die Atomwerte gegeben: jeweils C: 0,245, N: 0,283, O: 0,316, S: 0,316, (d.h. jeweils für Atome von Kohlenstoff, Stickstoff, Sauerstoff und Schwefel). Mit Bezug auf die Gleichungen 5 und 6 werden den Van-der-Waals-Radien die Atomwerte C: 1,85, N: 1,75, O: 1,60, S: 2,00 Å gegeben.

**[0076]** Selbstverständlich werden alle in den oben stehenden Gleichungen gegebenen, vorbestimmten Werte und Konstanten innerhalb der Einschränkungen des aktuellen Wissens von Protein-Ligand-Wechselwirkungen mit besonderem Bezug auf die Art der Berechnung, die hier durchgeführt wird, bestimmt. Daher ist es möglich, dass bei weiterer Verfeinerung der Bewertungsfunktion diese Werte und Konstanten sich verändern können, so dass jeglicher numerischer Wert, der die gewünschten Ergebnisse als Abschätzung der Bindungsenergie eines Proteins an einen Liganden ergibt, verwendet werden kann und so innerhalb des Bereichs der vorliegenden Erfindung fällt. Wie oben beschrieben, wird die Bewertungsfunktion auf Daten angewandt, die aus der Datenbank von Seitenkettenkonformationen, Atomidentitäten und interatomaren Abständen extrahiert wurden.

Für die Zwecke der vorliegenden Beschreibung beträgt die Anzahl an MHC-Klasse-II-Molekülen, die in dieser Datenbank eingeschlossen sind, 42 Modelle plus vier gelöste Strukturen. Es sollte aus den obigen Beschreibungen offensichtlich sein, dass die modulare Natur dieses Aufbaus des Berechnungsverfahrens der vorliegenden Erfindung bedeutet, dass neue Modelle einfach zugefügt und mit der Peptidgrundgerüst-Bibliotheks- und Seitenkettenkonformations-Suchfunktion gescannt werden können, um zusätzliche Datensätze zu erzeugen, die durch die Peptid-Bewertungsfunktion verarbeitet werden können, wie oben beschrieben. Dies ermöglicht, dass das Repertoire an gescannten MHC-Klasse-II-Molekülen einfach vergrößert werden kann, oder dass Strukturen und zugehörige Daten ersetzt werden können, wenn Daten zur Erzeugung von genaueren Modellen der existierenden Allele erhältlich sind. Das vorliegende Vorhersageverfahren kann gegen einen Datensatz kalibriert werden, der eine große Anzahl an Peptiden umfasst, deren Affinität für verschiedene MHC-Klasse-II-Molekülen vorher experimentell bestimmt wurde. Durch Vergleich von berechneten gegen experimentelle Daten kann eine Wertgrenze bestimmt werden, oberhalb derer es bekannt ist, dass alle experimentell bestimmten T-Zell-Epitope korrekt vorhergesagt wurden. Es sollte bekannt sein, dass, obwohl die oben genannte Bewertungsfunktion relativ einfach ist, verglichen mit einigen ausgereiften, erhältlichen Methodologien, die Berechnungen äußerst schnell ausgeführt werden. Es sollte auch bekannt sein, dass das Ziel keine Berechnung der wahren Bindungsenergie an sich für jedes in der Bindungsfurche eines gewählten MHC-Klasse-II-Proteins angekoppelte Peptid ist. Das zugrunde liegende Ziel ist es, vergleichende Bindungsenergie-daten als eine Hilfe zur Vorhersage der Lokalisierung von T-Zell-Epitopen basierend auf der Primärstruktur (d.h. der Aminosäuresequenz) eines gewählten Proteins zu erhalten. Eine relativ hohe Bindungsenergie oder eine Bindungsenergie oberhalb eines gewählten Schwellenwerts würde auf die Gegenwart eines T-Zell-Epitops in dem Liganden hinweisen. Der Ligand kann dann mindestens einer Runde an Aminosäuresubstitution unterzogen werden und die Bindungsenergie neu berechnet werden. Aufgrund der schnellen Natur der Berechnungen können diese Manipulationen der Peptidsequenz interaktiv innerhalb des User-Interfaces des Programms auf kosteneffizient erhältlicher Computerhardware ausgeführt werden. Daher ist keine wesentliche Investition in Computerhardware erforderlich. Einem Fachmann wird offensichtlich sein, dass andere erhältliche Software für die gleichen Zwecke verwendet werden könnte. Insbesondere kann technisch ausgereifere Software, die Liganden in Proteinbindungsstellen ankoppeln kann, im Verbindung mit der Energieminimierung verwendet werden. Beispiele für Docking-Software sind: DOCK (Kuntz et al., *J. Mol. Biol.*, 161: 269–288 (1982)), LUDI (Böhm, H.J., *J. Comput Aided Mol Des.*, 8: 623–632 (1994)) und FLEXX (Rarey M., et al., *ISMB*, 3: 300–308 (1995)). Beispiele von Molecular-Modeling- und Manipulationssoftware umfassen: AMBER (Tripos) und CHARMM (Molecular Simulations Inc.). Die Verwendung dieser Rechenverfahren würde den Datendurchsatz des erfindungsgemäßen Verfahrens aufgrund der Längen an Verarbeitungszeit, die für das Ausführen der notwendigen Berechnungen benötigt werden, stark einschränken. Es ist dennoch möglich, dass derartige Verfahren als ein „sekundärer Screen“ verwendet werden könnten, um genauere Berechnungen der Bindungsenergien von Peptiden zu erhalten, die sich über das erfindungsgemäße Verfahren als „positive Binder“ erwiesen haben. Die Beschränkung der Verarbeitungszeit für ausgereifte molekulare Mechanik- oder molekulare Dynamikberechnungen ist eine solche, die sowohl durch den Entwurf der Software, die diese Berechnungen anstellt, wie auch durch die aktuellen Technologiebeschränkungen der Computerhardware definiert wird. Es kann vorweggenommen werden, dass es in Zukunft, mit dem Schreiben eines effizienteren Codes und der stetigen Verbesserung der Geschwindigkeit von Computerprozessoren, möglich werden könnte, derartige Berechnungen innerhalb eines besser zu bewältigenden Zeitrahmens auszuführen. Weitere Information zu Energiefunktionen, die auf Makromoleküle angewandt werden, und Berücksichtigung der verschiedenen Wechselwirkungen, die innerhalb einer gefalteten Proteinstruktur stattfinden, können gefunden werden in: Brooks, B.R., et al., *J. Comput. Chem.*, 4: 187–217 (1983) und weitere Information bezüglich allgemeiner Protein-Ligand-Wechselwirkungen können gefunden werden in: Dauber-Osguthorpe et al., *Proteins* 4(1): 31–47(1988), welche durch Literaturhinweis in ihrer Gesamtheit hier aufgenommen sind. Nützliche Hintergrundinformationen können auch gefunden werden, zum Beispiel, in Fasman, G.D., ed., *Prediction of Protein Structure and the Principles of Protein Conformation*, Plenum Press, New York, ISBN: 0-306 4313-9.

## SEQUENZPROTOKOLL

<110> Merck Patent GmbH

<120> Modifizierter Granulozyten-Macrophagen-Kolonie-stimulierender Faktor (GM-CSF) mit verminderter Immunisierungskraft

<130> 101A139

<140> PCT/EP 02/02148

<141> 2002-02-28

<150> EP 01105775.9

<151> 2002-03-08

<160> 50

<170> PatentIn version 3.1

<210> 1

<211> 127

<212> PRT

<213> Homo sapiens

<400> 1

Ala Pro Ala Arg Ser Pro Ser Pro Ser Thr Gln Pro Trp Glu His Val  
1 5 10 15

Asn Ala Ile Gln Glu Ala Arg Arg Leu Leu Asn Leu Ser Arg Asp Thr  
20 25 30

Ala Ala Glu Met Asn Glu Thr Val Glu Val Ile Ser Glu Met Phe Asp  
35 40 45

Leu Gln Glu Pro Thr Cys Leu Gln Thr Arg Leu Glu Leu Tyr Lys Gln  
50 55 60

Gly Leu Arg Gly Ser Leu Thr Lys Leu Lys Gly Pro Leu Thr Met Met  
65 70 75 80

Ala Ser His Tyr Lys Gln His Cys Pro Pro Thr Pro Glu Thr Ser Cys  
85 90 95

Ala Thr Gln Thr Ile Thr Phe Glu Ser Phe Lys Glu Asn Leu Lys Asp  
100 105 110

Phe Leu Leu Val Ile Pro Phe Asp Cys Trp Glu Pro Val Gln Glu  
115 120 125

<210> 2

<211> 13

<212> PRT

<213> Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; Peptidsequenz in humanem GM-CSF mit möglicher humaner MHC-Klasse-II-Bindungsaktivität

&lt;400&gt; 2

Pro Ser Pro Ser Thr Gln Pro Trp Glu His Val Asn Ala  
 1 5 10

&lt;210&gt; 3

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; siehe SEQ No. 2

&lt;400&gt; 3

Gln Pro Trp Glu His Val Asn Ala Ile Gln Glu Ala Arg  
 1 5 10

&lt;210&gt; 4

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; siehe SEQ No. 2

&lt;400&gt; 4

Glu His Val Asn Ala Ile Gln Glu Ala Arg Arg Leu Leu  
 1 5 10

&lt;210&gt; 5

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; siehe SEQ No. 2

&lt;400&gt; 5

His Val Asn Ala Ile Gln Glu Ala Arg Arg Leu Leu Asn  
 1 5 10

&lt;210&gt; 6

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; siehe SEQ No. 2

&lt;400&gt; 6

Val Asn Ala Ile Gln Glu Ala Arg Arg Leu Leu Asn Leu  
 1 5 10

<210> 7  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 7

Asn Ala Ile Gln Glu Ala Arg Arg Leu Leu Asn Leu Ser  
 1 5 10

<210> 8  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 8

Arg Arg Leu Leu Asn Leu Ser Arg Asp Thr Ala Ala Glu  
 1 5 10

<210> 9  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 9

Arg Leu Leu Asn Leu Ser Arg Asp Thr Ala Ala Glu Met  
 1 5 10

<210> 10  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 10

Leu Asn Leu Ser Arg Asp Thr Ala Ala Glu Met Asn Glu  
 1 5 10

<210> 11  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 11

Arg Asp Thr Ala Ala Glu Met Asn Glu Thr Val Glu Val  
1 5 10

<210> 12

<211> 13

<212> PRT

<213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 12

Ala Glu Met Asn Glu Thr Val Glu Val Ile Ser Glu Met  
1 5 10

<210> 13

<211> 13

<212> PRT

<213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 13

Asn Glu Thr Val Glu Val Ile Ser Glu Met Phe Asp Leu  
1 5 10

<210> 14

<211> 13

<212> PRT

<213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 14

Glu Thr Val Glu Val Ile Ser Glu Met Phe Asp Leu Gln  
1 5 10

<210> 15

<211> 13

<212> PRT

<213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 15

Val Glu Val Ile Ser Glu Met Phe Asp Leu Gln Glu Pro  
1 5 10

<210> 16  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 16

Glu Val Ile Ser Glu Met Phe Asp Leu Gln Glu Pro Thr  
 1 5 10

<210> 17  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 17

Ile Ser Glu Met Phe Asp Leu Gln Glu Pro Thr Cys Leu  
 1 5 10

<210> 18  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 18

Ser Glu Met Phe Asp Leu Gln Glu Pro Thr Cys Leu Gln  
 1 5 10

<210> 19  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 19

Glu Met Phe Asp Leu Gln Glu Pro Thr Cys Leu Gln Thr  
 1 5 10

<210> 20  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 20

Met Phe Asp Leu Gln Glu Pro Thr Cys Leu Gln Thr Arg  
1 5 10

<210> 21

<211> 13

<212> PRT

<213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 21

Phe Asp Leu Gln Glu Pro Thr Cys Leu Gln Thr Arg Leu  
1 5 10

<210> 22

<211> 13

<212> PRT

<213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 22

Glu Pro Thr Cys Leu Gln Thr Arg Leu Glu Leu Tyr Lys  
1 5 10

<210> 23

<211> 13

<212> PRT

<213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 23

Thr Cys Leu Gln Thr Arg Leu Glu Leu Tyr Lys Gln Gly  
1 5 10

<210> 24

<211> 13

<212> PRT

<213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 24

Gln Thr Arg Leu Glu Leu Tyr Lys Gln Gly Leu Arg Gly  
1 5 10

<210> 25

<211> 13

Seite 6



<212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 25

Thr Arg Leu Glu Leu Tyr Lys Gln Gly Leu Arg Gly Ser  
 1 5 10

<210> 26  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 26

Leu Glu Leu Tyr Lys Gln Gly Leu Arg Gly Ser Leu Thr  
 1 5 10

<210> 27  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 27

Glu Leu Tyr Lys Gln Gly Leu Arg Gly Ser Leu Thr Lys  
 1 5 10

<210> 28  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 28

Gln Gly Leu Arg Gly Ser Leu Thr Lys Leu Lys Gly Pro  
 1 5 10

<210> 29  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 29

Arg Gly Ser Leu Thr Lys Leu Lys Gly Pro Leu Thr Met  
 1 5 10 -

<210> 30  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 30

Gly Ser Leu Thr Lys Leu Lys Gly Pro Leu Thr Met Met  
 1 5 10

<210> 31  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 31

Ser Leu Thr Lys Leu Lys Gly Pro Leu Thr Met Met Ala  
 1 5 10

<210> 32  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 32

Thr Lys Leu Lys Gly Pro Leu Thr Met Met Ala Ser His  
 1 5 10

<210> 33  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 33

Lys Gly Pro Leu Thr Met Met Ala Ser His Tyr Lys Gln  
 1 5 10

<210> 34  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; siehe SEQ No. 2

&lt;400&gt; 34

Gly Pro Leu Thr Met Met Ala Ser His Tyr Lys Gln His  
 1 5 10

&lt;210&gt; 35

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; siehe SEQ No. 2

&lt;400&gt; 35

Pro Leu Thr Met Met Ala Ser His Tyr Lys Gln His Cys  
 1 5 10

&lt;210&gt; 36

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; siehe SEQ No. 2

&lt;400&gt; 36

Leu Thr Met Met Ala Ser His Tyr Lys Gln His Cys Pro  
 1 5 10

&lt;210&gt; 37

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; siehe SEQ No. 2

&lt;400&gt; 37

Thr Met Met Ala Ser His Tyr Lys Gln His Cys Pro Pro  
 1 5 10

&lt;210&gt; 38

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

&lt;223&gt; siehe SEQ No. 2

&lt;400&gt; 38

Ser His Tyr Lys Gln His Cys Pro Pro Thr Pro Glu Thr  
 1 5 10

<210> 39  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 39

Cys Pro Pro Thr Pro Glu Thr Ser Cys Ala Thr Gln Thr  
 1 5 10

<210> 40  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 40

Pro Glu Thr Ser Cys Ala Thr Gln Thr Ile Thr Phe Glu  
 1 5 10

<210> 41  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 41

Cys Ala Thr Gln Thr Ile Thr Phe Glu Ser Phe Lys Glu  
 1 5 10

<210> 42  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 42

Gln Thr Ile Thr Phe Glu Ser Phe Lys Glu Asn Leu Lys  
 1 5 10

<210> 43  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 43

Ile Thr Phe Glu Ser Phe Lys Glu Asn Leu Lys Asp Phe  
1 5 10

<210> 44

<211> 13

<212> PRT

<213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 44

Glu Ser Phe Lys Glu Asn Leu Lys Asp Phe Leu Leu Val  
1 5 10

<210> 45

<211> 13

<212> PRT

<213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 45

Ser Phe Lys Glu Asn Leu Lys Asp Phe Leu Leu Val Ile  
1 5 10

<210> 46

<211> 13

<212> PRT

<213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 46

Glu Asn Leu Lys Asp Phe Leu Leu Val Ile Pro Phe Asp  
1 5 10

<210> 47

<211> 13

<212> PRT

<213> Künstliche Sequenz

<220>

<223> siehe SEQ No. 2

<400> 47

Asn Leu Lys Asp Phe Leu Leu Val Ile Pro Phe Asp Cys  
1 5 10

<210> 48  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 48

Lys Asp Phe Leu Leu Val Ile Pro Phe Asp Cys Trp Glu  
 1 5 10

<210> 49  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 49

Asp Phe Leu Leu Val Ile Pro Phe Asp Cys Trp Glu Pro  
 1 5 10

<210> 50  
 <211> 13  
 <212> PRT  
 <213> Künstliche Sequenz

<220>  
 <223> siehe SEQ No. 2

<400> 50

Leu Leu Val Ile Pro Phe Asp Cys Trp Glu Pro Val Gln  
 1 5 10

### Patentansprüche

1. Modifizierter humaner Granulozyten-Makrophagen-Kolonie-stimulierender Faktor (GM-CSF) mit der gleichen biologischen Aktivität, aber einer reduzierten Anzahl an T-Zell-Epitopen verglichen mit nicht-modifizierten GM-CSF, wenn er in vivo verwendet wird, wobei die Verringerung oder Entfernung von T-Zell-Epitopen durch Ersetzen des Aminosäurerests an Position 16 in dem ursprünglich nicht-modifizierten GM-CSF durch die Aminosäurereste, die aus der Gruppe ausgewählt sind, bestehend aus A, D, G, K, P, R, S oder T, erreicht wird.

2. Modifizierter humaner GM-CSF nach Anspruch 1, wobei die ursprünglich vorhandenen T-Zell-Epitope Peptidsequenzen mit der Fähigkeit sind, T-Zellen über die Präsentation an Klasse II zu stimulieren oder an diese zu binden.

3. Modifiziertes humanes GM-CSF-Polypeptid nach Anspruch 1 oder 2, das durch die folgenden Schritte erhältlich ist:

- (i) Bestimmen der Aminosäuresequenz des Polypeptids oder eines Teils davon,
- (ii) Identifizieren eines oder mehrerer T-Zell-Epitope in der Aminosäuresequenz des Proteins durch
  - (a) Auswählen einer Region des Peptids mit bekannter Aminosäurerestsequenz;
  - (b) aufeinanderfolgendes Abtasten überlappender Aminosäurerestsegmente mit zuvor bestimmter gleichmäßiger Größe, die aus mindestens drei Aminosäureresten bestehen, aus der ausgewählten Region;
  - (c) Berechnen des MHC-Klasse-II-Molekül-Bindungswertes für jedes abgetastete Segment durch Aufsummieren zugewiesener Werte für jede in dem abgetasteten Aminosäurerestsegment vorhandene hydrophobe Aminosäurerest-Seitenkette, wobei der Bindungswert nach einer Böhm-Bewertungsfunktion berechnet wird, die so modifiziert ist, dass sie 12-6 van-der-Waals-Ligand-Proteinenergie-Abstoßungsterm und Ligandenkonformationsenergie-Term enthält, indem (1) eine erste Datenbank von MHC-Klasse-II-Molekülmodellen bereitgestellt

wird, (2) eine zweite Datenbank zulässiger Peptidgrundgerüste für die MHC-Klasse-II-Moleküle bereitgestellt wird, (3) ein Modell aus der ersten Datenbank ausgewählt wird, (4) ein zulässiges Peptidgrundgerüst aus der zweiten Datenbank ausgewählt wird, (5) in jedem abgetasteten Segment vorhandene Aminosäurerest-Seitenketten identifiziert werden, (6) in jedem abgetasteten Segment der Bindungsaffinitätswert für alle vorhandenen Aminosäurerest-Seitenketten ermittelt wird und die Schritte (1) bis (5) für jedes Modell und jedes Grundgerüst wiederholt werden, und

(d) identifizieren mindestens eines der Segmente, die für eine Modifikation, wie in Tabelle 1 angegeben, geeignet sind, auf Basis des berechneten MHC-Klasse-II-Molekül-Bindungswertes für dieses Segment, um den Gesamt-MHC-Klasse-II-Bindungswert des Peptids zu verändern, im Wesentlichen ohne den therapeutischen Nutzen des Peptids zu verringern,

(iii) Konstruieren neuer Sequenzvarianten durch Substitution eines Aminosäurerestes, wie in Tabelle 2 angegeben, in einem der ursprünglich vorhandenen T-Zell-Epitope, wodurch die Aktivität des T-Zell-Epitops verringert oder beseitigt wird, wie durch die Bindung der Peptide an MHC-Moleküle unter Verwendung von In-vitro- oder In-silico-Techniken oder biologischen Assays oder durch Bindung von Peptid-MHC-Komplexen an T-Zellen bestimmt wird; und

(iv) Aufbau dieser Sequenzvarianten mit DNA-Rekombinationstechniken und Testen der Varianten, um eine oder mehrere Varianten mit wünschenswerten Eigenschaften zu identifizieren.

4. DNA-Sequenz, die einen modifizierten humanen GM-CSF nach Anspruch 1 oder 2 codiert.

5. Pharmazeutische Zusammensetzung, enthaltend ein modifiziertes humanes GM-CSF-Molekül nach Anspruch 1 oder 2, gegebenenfalls zusammen mit einem pharmazeutisch annehmbaren Träger, Verdünnungsmittel oder Hilfsmittel.

Es folgt kein Blatt Zeichnungen