Dokumentation von Peptiden

Von Gerhard Langer*, München

Teil I

Nachdem die Chemie der Peptide seinerzeit von Emil Fischer begründet wurde und ihre Weiterentwicklung dann Jahrzehnte hindurch fast völlig stagnierte, ist sie zurzeit in einer schnellen Aufwärtsentwicklung begriffen. Ständig werden neue Peptide mit neuen Wirkungen in der Natur entdeckt, und es werden diese und andere neue Peptide synthetisiert. Selbst der auf diesem Gebiet praktisch tätige Fachmann vermag die Fülle des veröffentlichten Materials kaum noch zu überblicken. Auch im Deutschen Patentamt, wo Patentanmeldungen über Peptide und Peptidsynthesen geprüft werden, sah man sich vor diese Schwierigkeiten gestellt, und so ergab sich die unabweisbare Notwendigkeit, Verfahren zur Dokumentation von Peptiden zu entwickeln, um einen Überblick über den Fortschritt auf diesem Gebiet zu sichern. Zwei derartige Verfahren wurden nun hier ausgearbeitet, und zwar dienen beide ausschließlich der Dokumentation der Peptide selbst, nicht einer Dokumentation der Verfahren zu ihrer Herstellung. Über das eine dieser beiden Verfahren soll im folgenden berichtet werden, das andere Verfahren wird Gegenstand einer nachfolgenden Veröffentlichung sein.

Von den Peptiden *stricto sensu*, d.h. den eigentlichen Eiweißpeptiden, und nur auf diese bezieht sich zunächst

* Adresse des Autors: Dr. G. Langer, De-la-Paz-Straße 39, München 19.

das hier mitgeteilte Dokumentationsverfahren, ist es bekannt, daß sie aus stets den gleichen Bausteinen, etwa 20 α-Aminocarbonsäuren, in stets wechselnder Reihenfolge zusammengesetzt sind, wobei die längsten bisher in ihrem Aufbau bekannten Peptide weit über 100 α-Aminocarbonsäure-Bausteine aufweisen.

Wenn man sich nun die Aufgabe stellt, Peptide zu dokumentieren, d.h. die vielen Hunderte oder gar schon Tausende bisher bekannter Aminocarbonsäuresequenzen in einer sinnvollen Anordnung unterzubringen, so soll dieses System nach Möglichkeit drei Bedingungen entsprechen: Die erste und wichtigste Bedingung ist hierbei die, ein System aufzustellen, nach dem nicht nur jedes bekannte, sondern vor allem auch jedes zukünftige *noch zu entdeckende Peptid* einen eindeutig definierten Platz erhält, der durch eine Codezahl ausgedrückt wird, und wobei diesem Platz nur ein einziges Peptid entspricht. Die zweite Bedingung ist die, daß die Codezahl leicht aufstellbar ist und das betreffende Peptid anhand dieser Zahl schnell und leicht mit eindeutiger Sicherheit unterzubringen und ebenso auch schnell und leicht wiederaufzufinden ist. Die dritte Bedingung ist die, daß bei der Anordnung der Peptide ihre Verwandtschaft,

ala	1
arg	2
asp	3
cys	4
glu	5
gly	6
his	7
hypr	8
ile	9
leu	10
lys	11
met	12
orn	13
phe	14
pro	15
ser	16
thr	17
try	18
tyr	19
val	20

Abb. 1

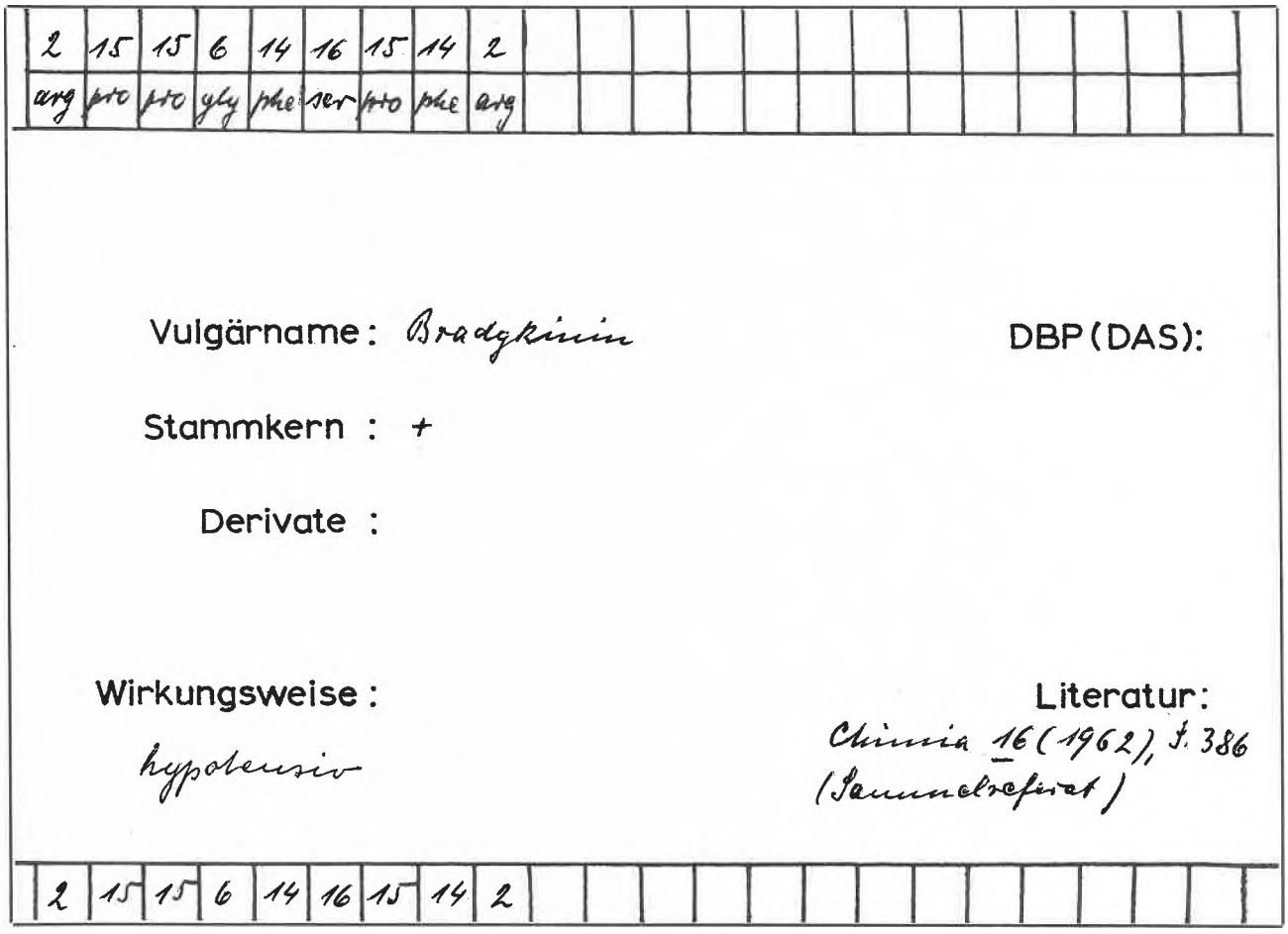


Abb. 2

die durch die Aminocarbonsäuresequenzen gegeben ist, in der Reihenfolge der Anordnung möglichst weitgehend zum Ausdruck kommt.

Nun sind die Peptide u. a. durch eine Eigenschaft ausgezeichnet, nämlich daß sie keine strukturellen Merkmale und auch keine chemischen Merkmale aufweisen, die eine irgendwie geartete Klassifizierung und somit sinnvolle Anordnung gestatten würden. Von hier aus gesehen, scheint es also nicht möglich, Peptide zu klassifizieren und zu dokumentieren, etwa in der Art, wie es in *Beilsteins Handbuch der organischen Chemie* für die Gesamtheit aller organischen Verbindungen angestrebt wird. Eine andere zur Gruppierung günstige Eigenschaft haben die Peptide indessen: es kommt ihnen eine nicht nur innerhalb Deutschlands, sondern eine international uneingeschränkt gültige einheitliche Formelschreibweise und dementsprechende Nomenklatur zu. Es läge vielleicht nahe, hierauf aufbauend die Peptide nach ihren Namen alphabetisch anzuordnen, aber abgesehen davon, daß sich hierbei eine in bezug auf die Verwandtschaft bestimmter Peptide und Peptidgruppen sinnlose Anordnung ergibt, wäre es nur mit einem außerordentlich großen Aufwand möglich, die ganz außergewöhnlich langen Namen dieser Verbindungen alphabetisch zu ordnen, und dauernde Alphabetfehler wären unvermeidbar. Selbst wenn man hierbei die Namen der Aminocarbonsäurebausteine nicht voll ausschreiben würde, sondern sich der international eingeführten Aminocarbonsäuresymbole nach BRAND und EDSALL[1] bedient, würde man, abgesehen davon, daß die Namen nicht ganz so lang sind, prinzipiell vor den gleichen Schwierigkeiten stehen.

Hier wird nun eine andere Möglichkeit der Anordnung der Peptide nach ihren Sequenzen, aufbauend auf der Reihenfolge der einzelnen Bausteine, vorgeschlagen, wobei Peptide, die gemäß der Reihenfolge ihrer Bausteine miteinander verwandt sind, einander entsprechend zugeordnet werden. Man ordnet zu diesem Zweck die als Peptidbausteine in Frage kommenden «klassischen» 20 Aminocarbonsäuren, bezeichnet durch ihre Symbole nach BRAND und EDSALL, zweckmäßig in alphabetischer Reihenfolge an und teilt ihnen in dieser Reihenfolge die Codezahlen 1 bis 20 zu; vgl. Abb. 1. Auf einer Karteikarte (Format DIN A 5; 21 × 15 cm), die am oberen Rand zwei Reihen Karos zu je 22 Feldern trägt und am unteren Rand nochmals eine Karoreihe zu 22 Feldern enthält (vgl. Abb. 2), trägt man am oberen Rand in der zweiten Reihe den Namen des Peptides (vom N-terminalen Ende beginnend) in Symbolschrift ein und in den darüber befindlichen Karos die den einzelnen Symbolen jeweils zugeordneten Codezahlen. In der am unteren Rand der Karte befindlichen Karoreihe werden die Codezahlen gleichfalls eingetragen. Nachdem man die Karten gegebenenfalls sodann mit Notizen, wie Literaturzitat, physiologische Wirksamkeit, Schmelzpunkt, Drehung usw., versehen hat, ordnet man sie hintereinander ein nach den am oberen Rand befindlichen Codezahlsequenzen, genauso, wie man sie bei einstelligen Zahlen nach dem Dezimalklassifikationssystem anordnen würde, nur mit dem einen Unterschied, daß man es beim Dezimalklassifikationssystem in jeder Stelle mit Zahlen von 0 bis 9 zu tun hat, hier in jeder Stelle mit Zahlen von 1 bis 20. Eine solche Reihenfolge von Karteikarten, bei der man also jeweils die Karten nach «steigenden» Codezahlsequenzen anordnet, sieht dann beispielsweise so aus, wie aus Abb. 3 ersichtlich, wobei die in der Abbildung unterste Karte in der Kartei vorne liegt.

Um zu vermeiden, daß man beim Einordnen einer neu hinzukommenden Karte jeweils die am oberen Kartenrand befindliche Codezahlsequenz dieser Karte aus einer gewissen Entfernung mit den Codezahlsequenzen der bereits in der Kartei befindlichen Karten vergleichen muß, was sowohl ermüdend für die Augen ist, als auch leicht zu Irrtümern Anlaß geben kann, wird, wie bereits erwähnt, die Codezahlsequenz am unteren Kartenrand bis zu einer Anzahl von 22 Codezahlen wiederholt, und

[1] *Annu. Rev. Biochem. 16* (1947) 224.

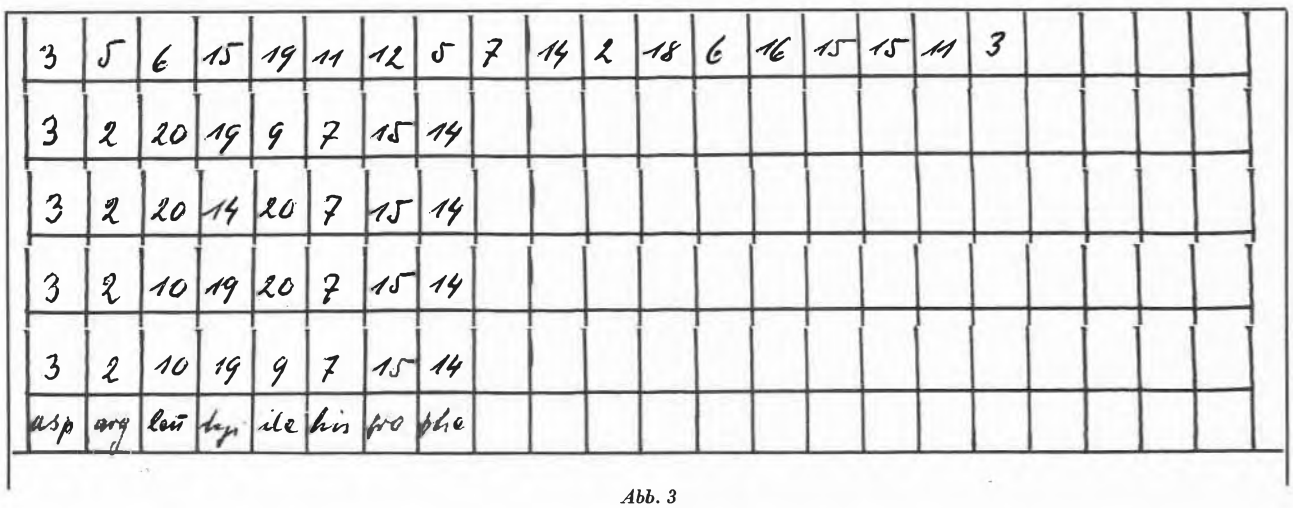


Abb. 3

so kann man den *unteren* Rand der einzuordnenden Karte auf die *oberen* Ränder der bereits eingeordneten Karten aufsetzen und so die Codezahlsequenzen unmittelbar miteinander vergleichen. Das Einordnen einer neuen Karte dauert so etwa 20 bis 30 Sekunden, und Irrtümer sind hierbei kaum möglich.

Die Vorteile dieses Dokumentationsverfahrens sind hierbei die folgenden:

1. Jedes Peptid hat einen eindeutig festgelegten Platz in der Rangfolge, der *nur* ihm, keiner anderen und vor allem auch keiner zukünftigen Aminocarbonsäuresequenz zukommt.
2. Das Dokumentationssystem ist beliebig aufnahmefähig und hat Platz für alle zukünftig noch zu beschreibenden Peptide.
3. Auch für weiterhin gegebenenfalls noch zu entdeckende Aminocarbonsäurebausteine ist das System beliebig aufnahmefähig, indem man neuentdeckte Aminocarbonsäuren in der Codezahltabelle (Abb. 1) hinten anfügt und ihnen laufend weitere Codezahlen zuerteilt.
4. Das Einordnen neuer Karten geht, wie oben bereits ausgeführt, sehr schnell, ebenso kann man jedes bereits in der Kartei befindliche Peptid leicht und sicher auffinden, indem man die Codezahlsequenz des zu suchenden Peptids anhand der Codezahltabelle aufstellt, die Codezahlsequenz am unteren Rande einer Karte aufträgt und unter Vergleich der unmittelbar über die Kartei gehaltenen Karte die Kartei durchblättert. Auch ein solches Aufsuchen dauert gewöhnlich nur 10 bis 30 Sekunden.
5. Man benötigt für dieses Dokumentationsverfahren keinerlei maschinelle Einrichtung! Es ist also überall dort angebracht, wo teure Hollerithmaschinen oder elektronische Geräte nicht zur Verfügung stehen.

Es ist nun zu berücksichtigen, daß ein jedes Peptid theoretisch in einer größeren Anzahl von optischen Isomeren existieren kann. So sind z.B., wenn man unterstellt, daß jeder einzelne Aminocarbonsäurebaustein sowohl in seiner D-, L- als auch DL-Form auftreten kann, allein für ein Hexapeptid bereits 729, für ein Decapeptid aber schon 59049 optisch isomere Variationen denkbar. Jedes dieser optisch isomeren Peptide kann darüber hinaus zu einer gewissen Anzahl von Derivaten abgewandelt werden, wie sie vor allem bei der Synthese von Peptiden auftreten, z.B. N-Carbobenzoxy-, N-*p*-Toluolsulfonyl-, Carboxylester-Derivate u.a.m. Diese unübersehbaren Mengen optischer und auch vielfacher funktioneller Derivate der Peptide können nun von einem Dokumentationssystem nicht systematisch vorgesehen werden und werden auch von dem hier vorgetragenen Dokumentationssystem nicht direkt erfaßt. Dokumentiert wird hier jeweils lediglich sozusagen das Stammpeptid, der «Stammkern», also das freie Peptid ohne jede funktionelle Veränderung und unabhängig von seinem optischen Verhalten. Da die meisten Peptide aber praktisch nur in einer oder sehr wenigen optischen Modifikationen auftreten und die Zahl der zum Zwecke ihrer Synthese hergestellten Derivate in jedem Einzelfall gleichfalls nicht groß ist, so dürfte es nicht schwierig sein, diese Abwandlungen des «Stammkerns» hinter der Karteikarte des jeweiligen Stammkerns einzuordnen und sinnvoll zu gruppieren.

Für Peptide, die mehr als 22 Aminocarbonsäurebausteine aufweisen, finden Karteikarten Verwendung, die am oberen Rande 4 Karoreihen tragen. Bei sehr langen Peptidketten dürfte es im allgemeinen indessen nicht notwendig sein, mehr als 22 bzw. 44 Stellen zu codieren und in die Karoreihen einzutragen, weil es kaum Peptide geben wird, die zwar verschieden voneinander sind, deren unterschiedliche Struktur aber erst *später* als in der 22. oder 44. Stelle auftritt.

Es waren nun noch lediglich einige Sonderregeln aufzustellen zwecks sinnvoller Einordnung zyklischer Peptide. Bei homodet zyklischen Peptiden, wie Gramicidin S beispielsweise, geht man so vor, daß man die einzelnen Aminocarbonsäurebausteine im Ring mit den Codezahlen versieht und dann das Peptid *linear* schreibt und entsprechend einordnet. Die zyklische Struktur kennzeichnet man hierbei durch eine entsprechend angebrachte Brücke, die man der besseren Auffälligkeit wegen zweckmäßig rot auszieht. Da nun zyklische Peptide einmal aus der einen, ein anderes Mal aus einer anderen linearen Sequenz durch Ringschluß synthetisiert werden können, wäre es nicht sinnvoll, als ersten Baustein an den Beginn der Sequenz denjenigen zu setzen, der von dem jeweiligen Autor entsprechend seiner Synthese dorthin gesetzt wurde, sondern man wird hier zweckmäßig nach dem «Prinzip der spätesten Systemstelle» verfah-

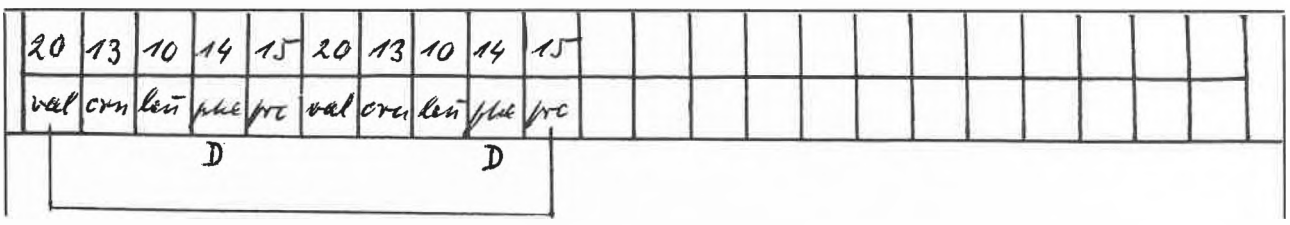


Abb. 4. Gramicidin S

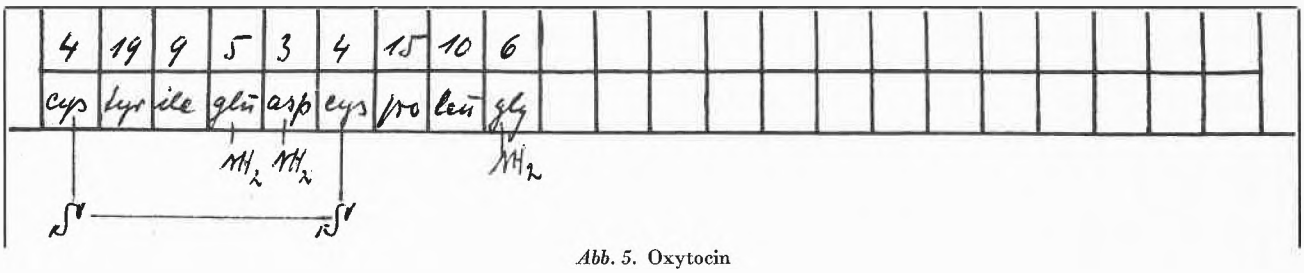


Abb. 5. Oxytocin

ren, indem man dasjenige Symbol bzw. diejenige Symbolsequenz, die in der Rangfolge der Codezahlen den höchsten (= «spätesten») Wert haben, an den Anfang setzt; vgl. hierzu die Codierung von Gramicidin S in Abb. 4.

Heterodet zyklische Peptide, wie z. B. Oxytocin, codiert man so wie das offenkettige noch nicht zum Disulfid oxydierte Octapeptid (vgl. Abb. 5) und legt es hinter letzterem als Derivat ab. Ein heterodet zyklisches Peptid wie Insulin behandelt man nach dem gleichen System derart, daß man ohne Rücksicht auf die disulfidischen zyklischen Strukturen diejenige der beiden Peptidsequenzen, die mit der höchsten (= spätesten) Codezahlsequenz beginnt, an den Anfang stellt und hiernach einordnet. Man dokumentiert also nach der B-Kette, die mit der Codezalsequenz 14–20–3– beginnt, und nicht nach der A-Kette, die mit der Codezahlsequenz 6–9–20– beginnt.

Das hier vorgeschlagene Dokumentationssystem hat sich im Deutschen Patentamt gut bewährt. Es wurde zunächst nur für «reine» Eiweißpeptide entwickelt. Inwieweit es sich auch für Depsipeptide und andere «gemischte» Peptide verwenden bzw. ausbauen läßt, wird Gegenstand weiterer Untersuchungen sein.

Teil II

Für die Dokumentation von Peptiden wurde auch ein Verfahren entwickelt, das mit Schlitzlochkarten arbeitet. Ausgehend von den 20 Aminocarbonsäurebausteinen, aus denen sich die Eiweißverbindungen aufbauen, wäre es naheliegend, diesen Bausteinen je eine Codezahl zuzuerteilen und für das jeweilige zu dokumentierende Peptid die Codezahlen der betreffenden, das Peptid zusammensetzenden Aminocarbonsäuren dann auf einer Lochkarte abzulochen. Indessen wäre hierdurch der Aufbau des einzelnen Peptides mittels einer Schlitzlochkarte nicht genügend präzisiert, insofern als ein Peptid durchaus nicht nur durch seine Aminocarbonsäurebausteine *als solche* charakterisiert ist, sondern darüber hinaus die *Reihenfolge,* in der diese Aminocarbonsäuren miteinander verknüpft sind, eine ausschlaggebende Rolle spielt. So sind beispielsweise Peptide, wie Alanyl-valyl-histidin, Histidyl-alanyl-valin und Valyl-histidyl-alanin, die aus den gleichen drei Aminocarbonsäuren bestehen, Verbindungen mit sehr verschiedenen chemischen, physikalischen und gegebenenfalls auch vor allem pharmakologischen Eigenschaften. Wenn man ferner als ausschließliche Kennzeichnen eines Peptides, das also *bestimmte* Aminocarbonsäuren in *bestimmter* Reihenfolge enthält, nur dessen einzelne Aminocarbonsäuren codieren und entsprechend in der Lochkarte ablochen würde, so würde bei späterem Selektieren nicht nur das jeweils gesuchte Peptid herausfallen, sondern man würde als «Ballastkarten» dazu auch alle diejenigen Peptide erhalten, die die gleichen Aminocarbonsäuren in anderer Reihenfolge enthalten.

Um dieser Schwierigkeit abzuhelfen, war zunächst daran gedacht, die *Reihenfolge* der Codezahlen, und damit die Reihenfolge der einzelnen Aminocarbonsäuren im Peptidmolekül beim Codieren dadurch festzulegen, daß man zwischen die Codezahlen für die einzelnen Aminocarbonsäuren weitere Codezahlen schaltet, durch die die Reihenfolge der Codezahlen und damit auch der Aminocarbonsäuren zwangsläufig gegeben wird. Das ließe sich in gewissem Umfang z. B. dadurch bewerkstelligen, daß man zwischen je zwei Codezahlen zweier Aminocarbonsäuren nach einem bestimmten Schlüssel, der hier nicht näher erläutert zu werden braucht, zwei zweistellige oder eine dreistellige weitere Codezahl als «Brückenzahl» schaltet und auch entsprechend ablocht.

Das würde, allgemein ausgedrückt, bedeuten, daß man für das jeweilige zu dokumentierende, aus n Aminocarbonsäurebausteinen bestehende Peptid, in dem einen Fall, bei Zwischenschaltung einer dreistelligen Codezahl als Brückenzahl $2n$-1 Codezahlen und die gleiche Anzahl von Lochungen benötigt, im anderen Fall, bei Zwischenschaltung zweier zweistelliger Brückenzahlen sogar $3n$-2 Codezahlen und ebensoviele Lochungen aufwenden muß.

Kleinere Peptidketten, vielleicht bis zu einer Anzahl von maximal 8 bis 10 Aminocarbonsäurebausteinen, ließen sich nach einem solchen Verfahren sicherlich sinnvoll codieren und auf einer Schlitzlochkarte lochen. Damit würden sich die Möglichkeiten eines solchen Verfahrens indessen erschöpfen, denn man muß berücksichtigen, daß bei dieser Methode ein Peptid mittlerer Kettenlänge von beispielsweise 15 Aminocarbonsäurebausteinen bereits eine Anzahl von $2 \times 15 - 1 = 29$ bzw.

von $3 \times 15 - 2 = 43$ (!) Lochungen auf *einer einzelnen* Karte und beim Selektieren auch die gleiche Anzahl von Nadelungen benötigt. Daß ein solches Verfahren beim Vorliegen von Peptiden größerer Kettenlängen praktisch nicht mehr durchführbar ist, liegt auf der Hand. Die Schlitzlochkarten üblicher Dimension (21×15 cm = DIN A5) würden zwischen den einzelnen Schlitzen nur noch wenige feste Papierflächen enthalten und zum großen Teil lediglich aus voneinander durch Schlitze getrennten Papierstreifen bestehen. Die Karten würden jede Haltbarkeit in sich verlieren, würden sich beim Selektieren verziehen, und schließlich würden auch zuviele «Ballastkarten» beim Selektieren anfallen.

Um diese Schwierigkeiten zu vermeiden, wurde nun ein anderer Weg eingeschlagen. Hierbei wird nicht für jede *einzelne* Aminocarbonsäure eine eigene Codezahl aufgestellt, die dann durch eine oder zwei weitere Codezahlen mit der nächsten Codezahl verknüpft werden muß, sondern es wird jeweils für *zwei aufeinanderfolgende* Aminocarbonsäuren in der im Einzelfall vorliegenden Reihenfolge *eine einzige* Codezahl eingeführt. Die einzelne Codezahl dokumentiert hierbei also sowohl die Aminocarbonsäure *a*, als auch die Aminocarbonsäure *b*, als auch deren vorgegebene Verknüpfungsrichtung $a \rightarrow b$. Das sieht dann in der Praxis folgendermaßen aus: das Hexapeptid $a \cdot b \cdot c \cdot d \cdot e \cdot f$ z.B. zerlegt man vom N-terminalen Ende ausgehend in die 5 Aminocarbonsäurepaare $a \cdot b$, $b \cdot c$, $c \cdot d$, $d \cdot e$ und $e \cdot f$. Wenn man diesen 5 Paaren je *eine* Codezahl zuerteilt und diese entsprechend ablocht, so ist das gesamte Peptid hiermit in der Schlitzlochkarte gekennzeichnet, und zwar sowohl *nach der Art* der Aminocarbonsäurebausteine als auch vor allem *in der Reihenfolge* der einzelnen Bausteine, die ja in der jeweiligen Codezahl zwangsläufig und eindeutig zum Ausdruck kommt. Und dies geschieht bei einem Peptid, bestehend aus *n* Aminocarbonsäuren, mittels einer Anzahl von nur *n*-1 Codezahlen.

Da nun bei der Kombination der einzelnen Aminocarbonsäuren zu Paaren nicht nur Paarungen zwischen *verschiedenenen* Aminocarbonsäuren sowohl in der einen wie auch in der umgekehrten Richtung möglich sind, sondern auch Paarungen zwischen jeweils *gleichen* Aminocarbonsäuren, so kann man aus den 20 bekannten «klassischen» Aminocarbonsäuren der Eiweißstoffe $20 \times 20 = 400$ Aminocarbonsäurepaare aufstellen, denen dann jeweils eine entsprechende Codezahl zuerteilt wird*.

Um nun beim Selektieren das gesuchte Peptid ohne Ballastkarten zu erhalten, ist es lediglich noch notwendig, außer den Codezahlen für die Aminocarbonsäurepaare noch zusätzlich eine Codezahl für die *Gesamtzahl* der Aminocarbonsäuren des betreffenden Peptides abzulochen. Für ein Peptid von *n* Bausteinen werden hiernach also insgesamt *n* Codezahlen benötigt.

Dieses Codierverfahren weist nun noch einen großen Vorteil auf: man kann hiermit auch nach bestimmten Aminocarbonsäuresequenzen suchen, unabhängig von der Kettenlänge eines Peptides. Man hat beispielsweise herausgefunden, daß das Peptid der Sequenz $a \cdot b \cdot c \cdot d \cdot e \cdot f$ oder die Sequenz als solche eine bestimmte spezifische physiologische Wirkung aufweist, und der Fachmann will feststellen, ob auch andere, größere Peptide in ihrem Molekül diese Sequenz enthalten und ob diese Peptide gegebenenfalls auch die gleiche oder eine ähnliche physiologische Wirkung zeigen. Oder aber es handelt sich darum, diese bestimmte Sequenz zu synthetisieren, und es soll festgestellt werden, ob diese Sequenz bereits bei anderen Peptiden vorkommt und nach welchem Verfahren und mit welchen Ausbeuten sie gegebenenfalls einmal synthetisiert worden ist.

Zu diesem Zweck stellt man die Codezahlen der gesuchten Sequenz fest, beispielsweise die 5 Codezahlen der Hexapeptidsequenz $a \cdot b \cdot c \cdot d \cdot e \cdot f$, und nadelt diese 5 Zahlen im Selektiergerät, in diesem Fall jedoch *ohne* die das Hexapeptid kennzeichnende und gewissermaßen begrenzende Gesamtzahl 6 der Aminocarbonsäuren zu nadeln. Jetzt werden beim Selektieren außer der Karte des Hexapeptids $a \cdot b \cdot c \cdot d \cdot e \cdot f$ auch alle diejenigen Karten herausfallen, die die Sequenz $a \cdot b \cdot c \cdot d \cdot e \cdot f$ an irgendeiner Stelle in ihrem größeren Peptidmolekül enthalten, unabhängig davon, ob sich die gesuchte Sequenz am Anfang, in der Mitte oder am Ende der Peptidkette befindet. Einschränkend ist hierbei lediglich zu sagen, daß in diesem Fall allerdings auch Karten größerer Peptide herausfallen können, wenn diese in ihrem Molekül *sämtliche* der Einzelpaarungen, also die Aminocarbonsäurepaare $a \cdot b$, $b \cdot c$, $c \cdot d$, $d \cdot e$ und $e \cdot f$, als solche enthalten, auch wenn diese gar nicht oder nur teilweise miteinander verknüpft sind. In diesem Fall muß man beim Selektieren also eine oder einige «Ballastkarten» mit in Kauf nehmen, die man dann aussortieren muß.

Es kann nun vorkommen, besonders wenn es sich um längere Peptidketten handelt, daß sich eine und dieselbe Paarung zweier Aminocarbonsäurebausteine im Molekül wiederholt, oder mit anderen Worten, daß dann die gleiche Codezahl, die auf der Lochkarte schon einmal abgelocht wurde, wiederkehrt. Wenn nun auch eine nochmalige Ablochung einer bereits schon einmal abgelochten Codezahl zum zweitenmal nicht möglich ist, so entsteht hierdurch keineswegs ein Fehler, und es kann trotzdem außer der gesuchten Peptidkarte kein anderes, vielleicht «niedrigeres» Peptid beim Selektieren herausfallen, denn das wird durch das gleichzeitige Ablochen der Codezahl für die *Gesamtzahl* der Aminocarbonsäurebausteine verhindert.

Aus diesem letzteren Grunde und dem vorher geschilderten Verfahren der Selektierung einzelner Se-

* Für das hier vorgeschlagene Codierverfahren haben sich Schlitzlochkarten der Firma Allform GmbH, Berlin, so z.B. der Vordruck GS060, Format DINA4 (= 21,0 × 29,7 cm), als besonders geeignet erwiesen. Diese Handlochkarte besitzt 1360 Lochstellen. Bei Direktverschlüsselung kann also ein Begriffsverzeichnis von 1360 Merkmalen aufgestellt werden, die theoretisch alle auf *einer* Karte abgelocht werden können, was aber in der Praxis kaum vorkommen dürfte.

quenzen wird mithin deutlich, weshalb beim Codieren von Peptiden nach dem vorliegenden Verfahren außer den n-1 Codezahlen für die Aminocarbonsäurepaare noch eine Codezahl für die Gesamtzahl der Aminocarbonsäurebausteine abgelocht werden muß.

Das vorgeschlagene Codierverfahren ist somit weitgehend abgesichert gegen einen Anfall von Ballastkarten.

Eine besondere Vereinfachung läßt sich gegebenenfalls einführen bei der Codierung von Peptiden mit ungewöhnlich langen Ketten. Wenn die Kettenlängen eine Größe erreichen von mehr als 40 bis 50 Aminocarbonsäurebausteinen, so ist es unter Umständen überhaupt nicht notwendig, die einzelnen Paare der Sequenz zu codieren und abzulochen, sondern hier dürfte es im allgemeinen genügen, einfach nur die Gesamtzahl der Bausteine zu codieren und abzulochen, weil in diesen Bereichen erst eine derart geringe Zahl von Peptiden in ihrer Struktur erforscht bzw. synthetisiert ist, daß beim Selektieren nach der Gesamtzahl der Bausteine Ballastkarten kaum zu erwarten sind. Allerdings muß man in diesem Fall auf das oben beschriebene Selektieren einzelner Sequenzen verzichten.

Außer den Codezahlen für die Aminocarbonsäurepaare bzw. den aus ihnen sich ableitenden Codezahlsequenzen der einzelnen Peptide können selbstverständlich gewünschtenfalls noch weitere Merkmale codiert und abgelocht werden.

Hierfür kommen in erster Linie weitere strukturelle Merkmale in Frage, wie vor allem zyklischer Bau mancher Peptide. Heterodet zyklische Peptide, z.B. Oxytocin, werden so behandelt, daß man die Aminocarbonsäuresequenz wie oben beschrieben codiert, und dann zusätzlich eine Codezahl für das Merkmal einer S–S-Brücke einführt. Der Beginn der Codierung erfolgt auch hier grundsätzlich am N-terminalen Ende der Aminocarbonsäuresequenz. Gewünschtenfalls kann auch die Möglichkeit vorgesehen werden, durch die Codezahl zum Ausdruck zu bringen, daß eine S—S-Brücke mehrfach auftritt, z.B. dreimal wie beim Insulin.

Etwas anders liegen die Verhältnisse bei homodet zyklischen Peptiden. Zunächst einmal haben homodet zyklische Peptide grundsätzlich eine Codezahl mehr als offenkettige oder heterodet zyklische Peptide, denn wie sich aus der Formel ohne weiteres ablesen läßt, enthalten homodet zyklische Peptide eine Aminocarbonsäurepaarung mehr als die offenkettigen oder die heterodet zyklischen Peptide. Sodann läßt sich für das Merkmal der homodet zyklischen Struktur ebenfalls eine Codezahl einsetzen. Im Gegensatz zum Verfahren bei heterodet zyklischen Peptiden ist es aber bei homodet zyklischen Peptiden ohne Bedeutung, bei welchem der Aminocarbonsäurepaare man mit der Codierung beginnt. Man erhält stets die gleichen Codezahlen bzw. Ablochungen, unabhängig von der Stelle im Ring, an der man mit der Codierung beginnt.

Selbstverständlich wird man bei heterodet wie auch bei homodet zyklischen Peptiden auch die Gesamtzahl der Aminocarbonsäurebausteine codieren und ablochen. Bei vielen homodet zyklischen Peptiden ist dies sogar unbedingt *notwendig*. Bekanntlich weisen manche homodet zyklische Peptide, beispielsweise Gramicidin S, zweimal die gleiche Aminocarbonsäuresequenz im Ring auf. Da aber jedes Aminocarbonsäurepaar, wie oben bereits ausgeführt, nur einmal gelocht werden kann, kann beim Selektieren ein kleiner Ring, der die gesamte Sequenz nur einmal enthält, von dem doppelt so großen Ring, der die gleiche Sequenz zweimal enthält, nur durch die zusätzliche Codezahl für die Gesamtzahl der Aminocarbonsäurebausteine unterschieden werden und so das Herausfallen von Ballastkarten vermieden werden.

Schließlich wird es noch wichtig sein, für die verschiedenen physiologischen Wirksamkeiten der Peptide eigene Codezahlen aufzustellen, so daß man das gesamte Lochkartenmaterial auch nach diesem Gesichtspunkt hin selektieren kann. Hierbei ist es möglich, mit einer einzigen Nadelung beispielsweise sämtliche Peptide mit oxytocischer oder bakterizider Wirksamkeit zu selektieren.

Für weitere in das vorliegende Codierverfahren eventuell noch aufzunehmende Aminocarbonsäuren, sei es, daß noch die eine oder andere Aminocarbonsäure als Baustein echter Peptide entdeckt wird, sei es, daß man beabsichtigt, auch peptidfremde Aminocarbonsäuresequenzen nach diesem System zu codieren, kann folgendes Verfahren vorgeschlagen werden: Man teilt die noch unbekannten, für eine spätere Aufnahme vorgesehenen Aminocarbonsäuren in drei Gruppen ein, nämlich saure, basische und neutrale, genannt Xs, Xb und Xn, und stellt für diese drei Gruppen ebenfalls die Codezahlen für die Paarungen mit den bisher bekannten 20 Aminocarbonsäuren auf. Man erhält dann Codezahlen für $23 \times 23 = 529$ Paarungen. Hiermit verhindert man, daß die Anzahl der Codezahlen unzweckmäßig stark anschwillt, denn man muß ja berücksichtigen, daß eine Schlitzlochkarte in ihrer Aufnahmefähigkeit für die Anzahl der zu codierenden Merkmale begrenzt ist. Die hier verwendeten Schlitzlochkarten können indessen ohne weiteres Codezahlen für $30 \times 30 = 900$ Aminocarbonsäurepaare aufnehmen.

Das vorgeschlagene Verfahren hat sich im Deutschen Patentamt auf dem Gebiet der Peptide sehr gut bewährt.