

# ANGEWANDTE CHEMIE

86. Jahrgang 1974

Heft 6

Seite 209 – 242

## Neuere Befunde zur Evolution der Proteine

Von Roger Acher<sup>[\*]</sup>

Ist es möglich, durch eine einheitliche Evolutionstheorie die große Verschiedenheit der heute vorkommenden Proteine zu erklären? Die Vorgänge, die zur Diversifikation der Proteine führten, sind zweifellos jetzt noch nicht alle bekannt, doch macht der lineare Aufbau der Polypeptid-Ketten Reaktionen an den Genen wie Addition, Einfügung, Auslassung und Fusion sehr wahrscheinlich. Durch Kombination dieser Abwandlungen läßt sich die Entwicklung zahlloser, völlig verschiedener Proteine aus einem gemeinsamen Ur-Molekül erklären.

### 1. Einleitung

Wenn man davon ausgeht, daß die Proteine aufgrund ihrer Biosynthese (Übertragung der Information von den Gen-Desoxyribonucleinsäuren auf die Messenger-Ribonucleinsäuren und deren Übersetzung in die Polypeptid-Ketten<sup>[1]</sup>) indirekt die Genstruktur widerspiegeln, so läßt eine Änderung ihrer Aminosäure-Sequenz weitreichende Schlüsse über Ereignisse zu, die die Gene beeinflußt haben oder noch beeinflussen. Die neodarwinistische Evolutionstheorie nimmt an, daß sich die Gene im Laufe der Zeit durch Mutationen verändert haben, wie man sie auch heute beobachtet, d. h. durch kleine, zufällige Mutationsschritte, und daß dann durch die natürliche Auslese die Entwicklungslinien zustandekamen.

Diese Theorie berücksichtigt die Tatsache, daß die Abwandlung der Organismen, wie die Paläontologie gezeigt hat, ungerichtet verläuft und dennoch progressive Ergebnisse hervorbringt, wobei jeweils die jüngste Art am besten angepaßt ist. Diese Betrachtungsweise beruht demnach im wesentlichen auf morphologischen Beobachtungen. Es ist zu betonen, daß die natürliche Auslese zwar nicht alle Organe gleichmäßig, sondern z. B. einmal den Bewegungsapparat, ein anderes Mal die Atmungsorgane usw. betrifft und damit zu ungleichmäßigen Evolutionsgeschwindigkeiten der Organe führt, daß aber zwischen der Evolution der einzelnen Teile eine Koordination bestanden haben muß, welche die Entstehung so hoch entwickelter Tiere wie Säugetiere und Vögel ermöglichte.

Die Evolution von Molekülen, d. h. von Genen und damit auch von Proteinen, bietet gegenüber der Evolution der

Organismen aus der Sicht des Darwinismus keine neuen theoretischen Probleme. Auch hier läßt sich die Kombination von Mutation und Selektion zur Erklärung heranziehen, und es ist anzunehmen, daß die Bereitstellung einer Anzahl von Enzymen etwa für die Histidin-Biosynthese nicht schwieriger ist als z. B. der Zusammenschluß spezieller Bauelemente zu einem Auge, das akkomodieren kann. Die Evolutionstheorie hat allerdings auch ihre schwachen Punkte: Dazu zählen die willkürliche Wahl des Organs oder des Moleküls, das in Zukunft eine Rolle spielen soll, die Tatsache, daß der Vorteil einer Initialmutation eigentlich zu gering ist, um eine Selektion darauf aufzubauen, die Schwierigkeit, die Koordination der individuellen Evolutionsschritte zu erklären, usw.

Es wurde auch die Frage gestellt, ob sich die Proteine tatsächlich entwickeln, d. h. Strukturen zunehmender biologischer Wirksamkeit durchlaufen, oder ob der Selektionsvorgang auf der morphologischen Ebene stehenbleibt (wobei die Desoxyribonucleinsäuren der Gene sich zwangsläufig entwickeln). Im letzteren Falle würden die zwischen den Proteinen der einzelnen Arten beobachteten Abweichungen („Substitutionen“, d. h. Ersatz von Aminosäure-Resten durch andere) keinen Hinweis auf eine Evolution bieten. Es fällt jedoch schwer anzunehmen, daß sich die Evolution auf der Ebene der Proteine nur quantitativ, anatomisch dagegen auch qualitativ äußern soll. Allerdings ist die Zahl der Proteine bei einem Säugetier wesentlich größer als bei einem Bakterium<sup>[2]</sup>; eine zelluläre Differenzierung muß daher zu einer Vervielfachung und Spezialisierung der Proteine geführt haben.

Eine besondere Schwierigkeit bei der Untersuchung der Protein-Evolution besteht darin, daß sich ein Molekül anders verhält als ein Organ, bei dem die Funktion mehr oder weniger auch in der äußeren Form zum Ausdruck kommt und das

[\*] Prof. Dr. R. Acher  
Laboratoire de Chimie Biologique, Université de Paris VI  
96, Boulevard Raspail, Paris 6<sup>e</sup> (Frankreich)

damit eine vergleichende Bewertung von einer Art zur anderen gestattet. Ein derartiges Vorgehen ist selbst in den recht seltenen Fällen nicht möglich, in denen die dreidimensionale Struktur des Protein-Moleküls bekannt ist, denn die Konformation eines Proteins ergibt sich nicht von selbst aus seiner Funktion.

Die „biologische“ Funktion eines Proteins ist durch seine Fähigkeit, sich mit anderen Molekülen zu verbinden, d. h. an spezifischen Wechselwirkungen teilzunehmen, bestimmt. Bei diesen Wechselwirkungen sind einerseits die Morphologie des Proteins und andererseits die Art der exponierten chemischen Gruppen wesentlich, die über die Fähigkeit zum Kontakt bzw. zur Bindung entscheidet. Da die löslichen Proteine mehr oder minder globuläre Makromoleküle sind, müssen einige der Reste im Inneren verborgen sein und können nicht direkt an äußeren Wechselwirkungen teilnehmen; diese Reste beteiligen sich jedoch an Wechselwirkungen im Inneren des globulären Moleküls und sichern damit seine Stabilität. Die Einflüsse der Evolution auf das Protein werden sich demgemäß danach unterscheiden, ob sich die betroffenen Reste außen oder innen befinden; für eine Interpretation der Substitutionen muß deshalb die dreidimensionale Struktur des Proteins bekannt sein. Ebenso muß man auch die Konformation derjenigen Moleküle kennen, die mit dem Protein reagieren (Substrate, Rezeptoren, Antigene usw.), da sich die natürliche Auslese nach der Funktion, die sich aus der Wechselwirkung ergibt, richtet. Da sich biologische Eigenschaften bisher nicht auf atomarer Ebene erkennen lassen, kann man die Bedeutung der Aminosäure-Substitutionen bis jetzt nur unvollkommen erklären. Man muß sich im allgemeinen darauf beschränken, die strukturellen Veränderungen der Proteine zu katalogisieren, in der Hoffnung, dadurch weitere Details über ihre Wirkungsweise kennenzulernen.

Es versteht sich von selbst, daß nur Proteine der heute lebenden Arten untersucht werden können, d. h. von Arten, die am Ende ihrer Entwicklungslinie stehen. Man kann postulieren, daß sich die physiologischen Funktionen und damit auch die beteiligten Proteine in Rhythmen entwickelt haben, die denen der allgemeinen Morphologie annähernd parallel verlaufen. Demnach wären die Proteine der lebenden niederen Arten relativ nahe mit den Ur-Proteinen verwandt. Diese Annahme läßt sich beweisen, indem man morphologisch ähnliche Arten untersucht, bei denen man aber durch paläontologische Befunde weiß, daß sie Entwicklungslinien angehören, die sich schon vor langer Zeit getrennt haben.

Beispielsweise gehört der primitive Knochenfisch *Polypterus bichir* in die Reihe der Polypteroidei, die sich vor ungefähr 300 Millionen Jahren von der Linie der Actinopterygier (heutige Knochenfische) getrennt hat<sup>[4]</sup>. Trotzdem hat er die gleichen Neurohypophysen-Hormone wie die Actinopterygier, während die Säugetiere und die Vögel, die sich vor etwa demselben Zeitraum trennten, verschiedene Hypophysen-Hormone aufweisen. Die Annahme, daß morphologische und molekulare Evolution parallel verlaufen, muß aber noch durch weitere Beispiele erhärtet werden, bevor man sie als gesichert ansehen darf.

Große Fortschritte gelangen in den letzten Jahren bei der Reinigung und der Sequenzanalyse der Proteine; die Zahl und die Größe der bekannten Proteine haben sich seit 1966<sup>[5]</sup> wesentlich erhöht. Durch die automatische Bestimmung von Aminosäure-Sequenzen wird der „Atlas of Protein Sequence and Structure“<sup>[6]</sup>, dessen Inhalt sich schon jetzt jährlich ver-

doppelt, in seinem Umfang weiterhin beträchtlich anwachsen. Die Zahl der lebenden Arten liegt ungefähr zwischen 5 und 10 Millionen<sup>[7]</sup>; wenn man davon ausgeht, daß jede Art spezifische Proteine besitzt, und wenn man willkürlich 100000 Proteine je Art ansetzt, so gibt es 500 bis 1000 Milliarden Proteine, an denen die Biochemiker ihre Ausdauer und die ihrer Geräte erproben können. Allerdings gestattet die Aufklärung der Sequenzen solange nur sehr vorsichtige Schlüsse auf die Zusammenhänge zwischen Struktur und biologischer Wirkung, bis ähnliche Fortschritte bei der Aufklärung der Konformationen erzielt worden sind.

Da alle bisher bekannten erblichen Veränderungen Mutationen sind und da die Evolution zwangsläufig Abwandlungen im genetischen Material einschließt, sind die heute zu beobachtenden „Mutationen“ der Proteine für die Aufklärung der Vorgänge, die bei ihrer Evolution möglicherweise beteiligt waren, besonders geeignet. In diesem Fortschrittsbericht werden die Arten der Struktur-Abwandlungen, wie sie sowohl durch die molekulare Genetik als auch durch die vergleichende Biochemie aufgezeigt wurden, anhand von Beispielen erläutert. Während die Unterscheidung und Klassifizierung biologischer Arten relativ einfach ist, kann man das nicht von Molekül-Arten sagen. Die Sequenz eines Proteins wird durch ein Struktur-Gen bestimmt, das seinerseits aus einer Sequenz von Desoxyribonucleotiden besteht. Das Schicksal dieses Struktur-Gens ist bis zu einem gewissen Grade vom Schicksal der biologischen Art unabhängig. Die Bildung einer neuen biologischen Art aus einer älteren kann ohne Modifizierung des Struktur-Gens stattfinden. Umgekehrt kann sich das Struktur-Gen verdoppeln (Duplikation) und zu zwei Protein-Linien führen, ohne daß dadurch zwei neue biologische Arten entstehen. Ein Protein ist in erster Linie durch seine biologische Funktion gekennzeichnet; durch Isolierung von Proteinen aus verschiedenen Arten aufgrund dieses Kriteriums erhält man eine „Familie“ von „homologen“ Proteinen, welche einander strukturell ähnlich sind. Wenn man mehrere Proteine, die offensichtlich die gleiche Funktion haben, sich aber in ihren Aminosäure-Sequenzen unterscheiden („isologe“ Proteine), bei ein und derselben Art findet („Polymorphie“), so dürfte jedes dieser Proteine einem gesonderten Struktur-Gen entsprechen; in diesem Fall kann es schwierig werden, die Reihen der homologen Proteine aus denen der isologen Proteine herauszufinden, besonders wenn die Polymorphie nicht bei allen miteinander verglichenen Arten auftritt. Darüber hinaus können innerhalb einer biologischen Art strukturell ähnliche Proteine verschiedene Funktionen haben; dieser Befund legt die Annahme nahe, daß die Struktur-Gene Abkömmlinge eines gemeinsamen Ur-Gens sind, aus dem sie durch Duplikation und Mutation hervorgegangen sind. Die Struktur-Analogien sind in diesem Fall weniger stark als bei den isologen Proteinen ausgeprägt. Solche chemisch ähnlichen, aber funktionell verschiedenen Proteine können als „analoge“ Proteine bezeichnet werden.

## 2. Familien homologer Proteine

Man kennt mehrere Fälle, bei denen sich ein Protein aufgrund seiner Funktion und seiner speziellen Strukturelemente durch die gesamte Reihe der Wirbeltiere verfolgen läßt. Die Paläontologie lehrt, daß sich die Wirbeltiere während der letzten 500 Millionen Jahre entwickelt haben; für die Abwandlungen ihrer Proteine stand somit ausreichend Zeit zur Verfügung, um

allgemeine Trends sichtbar werden zu lassen. Bisher kennt man jedoch nur wenige Protein-Familien, die von den niederen Wirbeltieren bis hin zu den Säugetieren vorkommen: Cytochrom c, die  $\alpha$ - und  $\beta$ -Ketten des Hämoglobins, die Fibrinopeptide und die Neurohypophysen-Hormone. Nach den bisherigen Beobachtungen ändern sich die Polypeptid-Ketten in der Länge meist nur sehr wenig; die häufigste Abwandlung ist die Substitution von Aminosäure-Resten. Ein solcher Austausch einer Aminosäure kann schon durch Substitution einer einzigen Base im Nucleotidtriplett, welches den betreffenden Aminosäure-Rest im Gen codiert (Punktmutation), zustande kommen, und diese Tatsache erklärt auch, warum das Phänomen so häufig beobachtet wird.

## 2.1. Substitutionen von Aminosäuren

Der Vergleich der Aminosäure-Sequenzen innerhalb einer Protein-Familie zeigt, daß Reste in bestimmten Positionen unveränderlich sind, während andere mehr oder weniger häufig ausgetauscht werden. Bei den Substitutionen muß die allgemeine Konformation des Moleküls, welche seine Funktion bestimmt, auf jeden Fall erhalten bleiben. Es ist deshalb interessant, zunächst die Gründe dafür zu untersuchen, warum bestimmte Positionen der Polypeptid-Kette nicht verändert werden.

### 2.1.1. Die unveränderliche Funktion hydrophober Reste

Durch kristallographische Untersuchung der dreidimensionalen Strukturen war es möglich, bei Proteinen einer Familie die unveränderlichen Stellen räumlich zu lokalisieren und zu erkennen, daß sie für die konformative Stabilität des Moleküls notwendig sind.

#### 2.1.1.1. Cytochrom c

Cytochrom c ist ein elektronen-transportierendes Protein, das sich in den Mitochondrien aerober Organismen findet. Ungeachtet seiner Herkunft besteht das Molekül stets aus einer einzigen, aus 104 bis 112 Resten aufgebauten Polypeptid-Kette, welche kovalent an immer die gleiche prosthetische Gruppe gebunden ist. Das Eisen-Atom ist zwei- oder dreiwertig, je nachdem, ob das Molekül mit Cytochrom-Reduktase oder mit Cytochrom-Oxidase reagiert. Diese doppelte Wechselwirkung sollte eine recht eindeutig festgelegte Konformation voraussetzen. Da die Cytochrome c aus allen untersuchten Arten in ähnlicher Weise mit Rinder-Cytochrom-Oxidase reagieren, durfte man erwarten, daß ihre Konformationen nahe verwandt sind. Dies wurde durch Röntgen-Analyse von Pferde- und Makrelen-Fe<sup>III</sup>-Cytochrom bestätigt<sup>[8]</sup>. Bei Wirbeltieren enthält die Kette immer 104 Reste mit einer acetylierten N-terminalen Gruppe; bei Insekten ist die Kette auf 108 Reste, bei Pilzen auf 108 bis 110 und bei Pflanzen auf 112 Reste verlängert, wobei die zusätzlichen Reste stets am N-terminalen Ende angefügt sind<sup>[8]</sup> (Abb. 1). Die Zahl der unveränderlichen Positionen beträgt bei 29 untersuchten Arten – vom Menschen bis zum Weizen – 35, d. h. ungefähr ein Drittel. Statistische Berechnungen sprechen dafür, daß die Zahl der unveränderlichen Reste nicht unter 32 sinken wird<sup>[8]</sup>. Eine Sequenz von elf Resten (Positionen 70 bis 80 in der Zählung der Säugetier-Cytochrome c) ist unveränderlich. Diese Reste beteiligen sich einerseits an der Bildung einer zentralen hydrophoben Tasche und an der Bindung mit dem Häm, das in dieser Tasche

eingeschlossen ist, andererseits an der Bindung an Enzyme, z. B. Cytochrom-Oxidase.

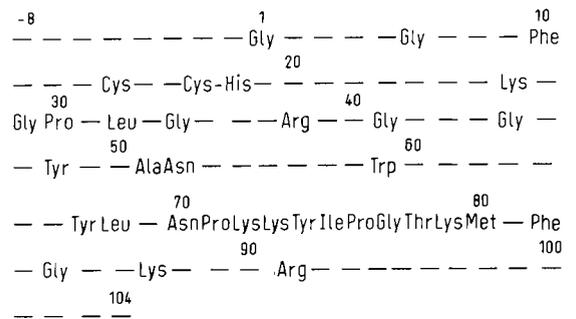


Abb. 1. Strukturschema der Cytochrome c. Die angegebenen Reste sind in den bisher bekannten Cytochromen c unveränderlich (nach [9]).

*Eingeschlossene unveränderliche Reste:* Das Häm ist an die Peptid-Kette über zwei Sulfid-Brücken gebunden und verknüpft die Cystein-Reste 14 und 17. Das Häm-Eisen selbst ist koordinativ sechsfach gebunden, und zwar an die vier Stickstoff-Atome des Porphyrins, an ein Stickstoff-Atom des Imidazol-Rings von Histidin-18 und an das Schwefel-Atom von Methionin-80. Das Häm ist außerdem fest in eine hydrophobe Tasche eingeschlossen (Abb. 2). Die lange unveränderliche Sequenz 70–80 ist derartig gefaltet, daß sie eine Wand der unpolaren Tasche bildet; die hydrophoben Reste in unmittel-

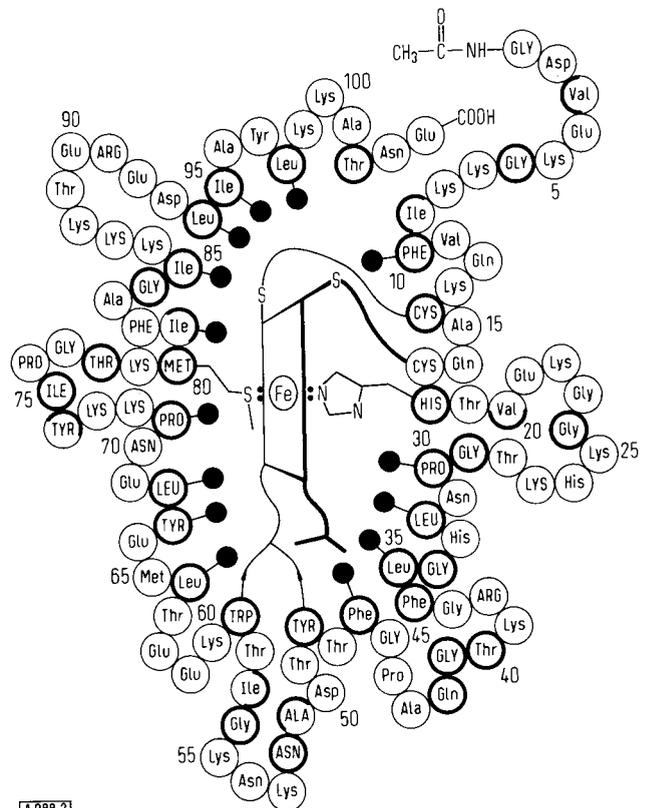


Abb. 2. Einschluß des Häms im Cytochrom-c-Molekül (schematisch). Die dicken Kreise zeigen Reste an, die im Inneren des Moleküls liegen; die mit ihnen verbundenen schwarzen Punkte markieren Seitenketten dieser Reste, die gegen das Häm stoßen. Die dünnen Kreise sind Reste, die sich an der Außenseite des Moleküls befinden, und die dicken Halbkreise entsprechen Gruppen, die halbverborgen an der Oberfläche liegen. Die Pfeile von Tryptophan-59 und Tyrosin-48 zur Propionsäure-Gruppe bedeuten Wasserstoff-Bindungen. Die durch Großbuchstaben gekennzeichneten Reste sind in den Proteinen von 29 Arten absolut konstant (nach [8]).

telbarer Nachbarschaft des Häms auf beiden Seiten sind entweder absolut unveränderlich oder unterliegen nur konservativen Substitutionen (d. h. solchen, bei denen der hydrophobe Charakter erhalten bleibt). Von den 25 hydrophoben Resten (Val, Leu, Ile, Met, Phe, Pro) in Pferde-Cytochrom c sind neun unveränderlich, zehn werden konservativ substituiert und sechs durch nichthydrophobe Reste ersetzt (Tabelle 1). Die Tasche enthält keine geladenen Reste.

Tabelle 1. Veränderlichkeit der Aminosäure-Reste bei Cytochromen c (Vergleich von Pferde-Cytochrom c mit den Cytochromen c von 28 anderen Arten).

Art der Reste	Reste insgesamt	unveränderlich	konservativ substituierbar [a]	vollständig substituierbar
sauer (Asp, Glu)	12	0	1	11 (92 %)
basisch (Lys, Arg)	21	7	1	13 (62 %)
ambivalent (Asn, Gln, Ser, Thr, His, Tyr, Trp, Cys, Ala)	34	11	13	10 (30 %)
hydrophob (Val, Leu, Ile, Met, Phe, Pro)	25	9	10	6 (24 %)
Glycin	12	8	3	1 (8 %)
Summe	104	35	28	41

[a] Konservative Substitutionen umfassen jeweils den Austausch innerhalb der Gruppen, die in der ersten Spalte links aufgeführt sind; auch der Austausch von Ala bzw. Gly mit den Resten der hydrophoben bzw. ambivalenten Gruppe wird als konservativ angesehen (nach [8]).

**Exponierte unveränderliche Reste:** Die polaren Reste an der Oberfläche der globulären Proteine spielen entweder bei Wechselwirkungen eine Rolle (aktive Zentren), oder sie sichern die korrekte Faltung der Kette, indem sie im wässrigen Medium verbleiben, oder sie sorgen – im Falle von geladenen Resten – für ein geeignetes Gleichgewicht der basischen und sauren Gruppen, so daß ein optimaler pH-Wert für das Molekül aufrecht erhalten wird. Tabelle 1 zeigt, daß recht viele basische Reste (Lysin und Arginin) im Gegensatz zu den sauren Resten unveränderlich sind. Es ist bekannt, daß an der Wechselwirkung zwischen Cytochrom c und Cytochrom-Oxidase einer oder mehrere Lysin-Reste des Cytochroms c teilnehmen, besonders Lysin-13 in Pferde-Cytochrom, da eine Blockierung durch Acetylierung oder Succinylierung die Bindung verhindert.

Röntgen-Untersuchungen zeigen, daß die sauren und basischen Reste an der Oberfläche des Moleküls getrennt liegen, und zwar umgeben zwei basische Zonen eine saure Zone; durch die Substitutionen wird diese allgemeine Anordnung nur geringfügig verändert. Diese für globuläre Proteine ungewöhnliche Verteilung ermöglicht zweifellos die Erkennung anderer Makromoleküle (Reduktase, Oxidase, Mitochondrien-Membran) und die Assoziation mit ihnen. Lysin-13, das an der Bindung an die Cytochrom-Oxidase teilnimmt, liegt direkt über der Tasche, in der das Häm untergebracht ist. Diese Position ist in allen Cytochromen c entweder von einem Lysin-Rest oder einem Arginin-Rest besetzt (konservative Substitution).

**Unveränderlichkeit von Glycin und Prolin:** Von den 12 Glycin-Resten in Pferde-Cytochrom c sind acht unveränderlich und drei unterliegen konservativen Substitutionen (Tabelle 1). Die-

se Reste, die keine sterische Hinderung verursachen, nehmen Positionen ein, in denen nur sie die allgemeine Konformation erhalten können, speziell an Ecken und an engen Kontaktstellen der Ketten. Von den vier Prolin-Resten sind drei unveränderlich und an Knickpunkten gelegen.

### 2.1.1.2. Die Globine

Bei den Globinen lassen sich drei Familien unterscheiden, die als Abwandlungen eines allgemeineren Strukturprinzips anzusehen sind: die Myoglobine sowie die  $\alpha$ - und die  $\beta$ -Hämoglobine.

**Myoglobine:** Myoglobin ist ein sauerstoff-speicherndes Protein, das sich in den Muskeln der Wirbeltiere findet. Die vollständigen Sequenzen der Myoglobine sind bis jetzt nur bei etwa einem Dutzend Säugetier-Arten bekannt; außerdem kennt man die N-terminalen Sequenzen der Myoglobine von Hühnern und von Quastenflossern. Es ist zweckmäßig, die Myoglobine direkt nach den Cytochromen c zu besprechen, da die dreidimensionale Struktur des Pottwal-Myoglobins bekannt ist<sup>[10, 12a]</sup> und da Myoglobine ebenfalls nur aus einer einzigen, mit einem Häm kombinierten Peptid-Kette bestehen. Von den 153 Aminosäure-Resten sind bis jetzt 99 als unveränderlich erkannt<sup>[11, 12]</sup>. Es ist jedoch schwierig anzugeben, wieviele von diesen tatsächlich nicht austauschbar sind, da man die Myoglobine niederer Wirbeltiere bisher praktisch nicht kennt. Das Häm, das nicht durch kovalente Bindungen mit der Kette verknüpft ist, enthält koordinativ sechswertiges Eisen, das an ein Stickstoff-Atom des Imidazol-Rings von Histidin-93 und an molekularen Sauerstoff gebunden ist. Außerdem ist das Häm in eine mit hydrophoben Resten ausgekleidete Tasche eingebettet.

Die Kette hat acht helicale Segmente und bildet annähernd eine kompakte Kugel mit den hydrophoben Gruppen im Inneren und nahezu allen polaren und geladenen Gruppen an der Oberfläche. Human- und Pottwal-Myoglobin unterscheiden sich in 26 Resten. Wenn man die dreidimensionale Struktur des Pottwal-Myoglobins zugrundelegt, so haben 20 Substitutionen an der Oberfläche des Moleküls stattgefunden, zwei an der Oberfläche von Taschen und vier im Inneren des Proteins. Drei der Unterschiede im Inneren ergeben sich durch hydrophobe Substitutionen (Nr. 28: Ile-Val; Nr. 99: Val-Ile; Nr. 142: Met-Ile). Die vierte (Nr. 110: Ala-Cys), die auch schon im Inneren von  $\alpha$ -Hämoglobin beobachtet wurde, betrifft möglicherweise einen unreaktiven Rest<sup>[11]</sup>. Das hydrophobe Innere des Proteins ist also bemerkenswert unveränderlich.

Die Myoglobin-Oberfläche hat offensichtlich keine speziellen Aufgaben zu erfüllen (Wechselwirkungen mit anderen Proteinen usw.), und man sollte daher annehmen, daß sich die Reste dort besonders leicht austauschen lassen. Dies scheint indessen nicht der Fall zu sein.

Human-Myoglobin hat dreimal so viele Reste mit den Myoglobinen anderer Säugetiere gemeinsam wie mit den  $\alpha$ - und  $\beta$ -Ketten des Human-Hämoglobins; es leuchtet ein, daß die Myoglobine als eigene Familie angesehen werden. Die allgemeine Konformation des Myoglobins findet sich jedoch nicht nur in den Ketten der Wirbeltier-Hämoglobine, sondern auch in den monomeren Hämoglobinen bestimmter wirbelloser Tiere.  **$\alpha$ - und  $\beta$ -Hämoglobine<sup>[17]</sup>:** Das normale Hämoglobin aller Wirbeltiere, mit Ausnahme der Cyclostomen (Rundmäuler), besteht aus vier Polypeptid-Untereinheiten, von denen zwei

dem  $\alpha$ - und zwei dem  $\beta$ -Typ angehören; diese bilden ein annähernd kugelförmiges Gebilde mit 50–60 Å Durchmesser. Bei manchen Arten werden zwei oder drei Hämoglobine gefunden, die sich voneinander in der Sequenz der  $\alpha$ - oder der  $\beta$ -Kette unterscheiden; es ist daher z. B. schwer zu sagen, welche von den zwei oder drei  $\alpha$ -Ketten zur normalen Reihe gehört. Grundsätzlich liegen jedoch wenigstens zwei recht verschiedene Familien vor, die des  $\alpha$ - und des  $\beta$ -Hämoglobins, die sich nach dem Auftreten der Cyclostomen und vor dem Auftreten der Fische voneinander getrennt haben müssen. Wenn sich die Aminosäure-Reste im Hämoglobin nicht austauschen lassen, so kann dies nicht nur, wie beim Myoglobin, durch die Notwendigkeit eines hydrophoben inneren Teils und durch die Polypeptid-Häm-Bindungen bedingt sein, sondern auch durch Kontakte zwischen den  $\alpha$ - und  $\beta$ -Ketten, die für die Erhaltung der tetrameren Struktur des Hämoglobins erforderlich sind.

Bei beiden Protein-Familien wurden bereits zahlreiche Moleküle aus der Reihe der Säugetiere in ihrer Sequenz aufgeklärt<sup>[6, 13]</sup>, ferner auch die  $\alpha$ -Ketten einer Vogel-<sup>[14]</sup> und einer Fisch-Art<sup>[15]</sup> sowie die  $\beta$ -Kette einer Amphibien-Art<sup>[16]</sup>. *Perutz et al.*<sup>[18]</sup> bestimmten die dreidimensionale Struktur des Pferde-Hämoglobins und ermittelten die Details über den tetrameren Aufbau des Moleküls. Hierbei wurde bestätigt, daß die  $\alpha$ - und  $\beta$ -Ketten eine ähnliche Sekundär-Struktur wie Myoglobin haben, obwohl sie sich in der Aminosäure-Zusammensetzung stark unterscheiden. Jede Kette hat sieben oder acht helicale Segmente sowie eine Tasche für das Häm.

*Perutz* lokalisierte die Reste, die an das Häm gebunden sind, sowie auch die Reste, die sich an den  $\alpha_1$ - $\beta_1$ - und  $\alpha_1$ - $\beta_2$ -Wechselwirkungen („kooperative Effekte“) beteiligen. In der  $\alpha$ -Kette (141 Reste) sind 18 Reste in Kontakt mit dem Porphyrin-Kern des Häms; Histidin-87 ist koordinativ an das Eisen gebunden. 16 Reste sind an den  $\alpha_1$ - $\beta_1$ -Kontakten und zehn an den  $\alpha_1$ - $\beta_2$ -Kontakten beteiligt (Abb. 3). In den  $\beta$ -Ketten (146 Reste) sind 20 Reste in Kontakt mit dem Porphyrin-Kern des Häms; Histidin-92 ist koordinativ an das Eisen gebunden. 18 Reste sind an den  $\beta_1$ - $\alpha_1$ -Kontakten und neun an den  $\beta_1$ - $\alpha_2$ -Kontak-

ten beteiligt. Die Bindungen zwischen den Peptid-Ketten und dem Häm oder zwischen den Peptid-Ketten verschiedenen Typs sind im allgemeinen unpolar; Bindungen zwischen den gleichartigen Ketten kommen selten vor und sind polar.

*Perutz und Lehmann*<sup>[19]</sup> fanden, daß Mutationen des Human-Hämoglobins, welche Reste in Kontakt mit dem Häm oder in  $\alpha$ - $\beta$ -Kontakten betreffen, fast immer pathologische Auswirkungen haben; dagegen führen Mutationen, die oberflächlich liegende, im allgemeinen polare Reste betreffen, gewöhnlich zu keinem klinischen Befund. Diese Beobachtung sollte mit der Tatsache korreliert werden, daß Reste, die in Kontakt mit dem Häm stehen oder an  $\alpha$ - $\beta$ -Wechselwirkungen beteiligt sind, meist generell unveränderlich sind (zumindest bei den Säugetieren). In der Familie der  $\beta$ -Ketten sind z. B. bei den Säugetieren 70 von 146 Resten unveränderlich; diese Zahl sinkt aber auf 54, wenn man auch die  $\beta$ -Kette des Frosches *Rana esculenta* einbezieht<sup>[16]</sup>. Unter den 21 Resten, die in Kontakt mit dem Häm stehen, sind 16 unveränderlich (76%), und von den neun Resten, die sich an den lebenswichtigen  $\alpha_1$ - $\beta_2$ -Kontakten beteiligen (Abb. 3), sind sieben unveränderlich (87%); dagegen sind von 17 Resten, die an den  $\alpha_1$ - $\beta_1$ -Wechselwirkungen teilnehmen, bei allen bisher bekannten  $\beta$ -Ketten nur acht nicht austauschbar (45%). Von den 100 Resten ohne Kontakt werden 25 nicht verändert. Zwei lange Sequenzen (die Reste 30–40 und 96–108) sind fast vollkommen unveränderlich; dieser sehr ungewöhnliche Befund läßt sich durch die direkte Aufeinanderfolge von Resten, die an wesentlichen Bindungen beteiligt sind, erklären.

### 2.1.1.3. Die Pankreas-Serin-Proteasen

Bei vier proteolytischen Enzymen des Pankreas sind sowohl die Sequenzen<sup>[20–23]</sup> als auch die dreidimensionalen Strukturen<sup>[24–26]</sup> bekannt: bei Chymotrypsin A und B, Trypsin und Elastase. Obwohl diese Enzyme phylogenetisch zu verschiedenen Familien gehören, ist wegen der Homologie ihrer Sequenzen und wegen der Art ihrer aktiven Zentren anzunehmen, daß sich ihre Struktur-Gene von einem gemeinsamen Ur-Gen ableiten<sup>[27]</sup>. Bemerkenswert ist, daß die übereinstimmenden Reste gewöhnlich im hydrophoben Innenraum liegen<sup>[26]</sup>

Alle chemischen und kristallographischen Befunde deuten daher darauf hin, daß in globulären Proteinen die Reste, die sich am Bau des zentralen Kerns, an der Bildung von Taschen für prosthetische Gruppen von der Art des Häms oder an Wechselwirkungen zwischen Untereinheiten beteiligen, gewöhnlich hydrophob sind und sich im Laufe der Evolution nicht verändern. Dagegen sind die an der Oberfläche gelegenen Reste hydrophil und können leicht ausgetauscht werden. Man kann daraus ableiten, daß Änderungen, die den inneren Teil betreffen, die allgemeine Konformation stark beeinflussen und entweder zum Verschwinden der Funktion und zur Eliminierung des Proteins führten oder – als Folge einer neuen Konformation und dem Auftreten einer neuen Protein-Familie – eine sehr stark abgewandelte Funktion entstehen ließen. Die Abstammung läßt sich in diesem Fall chemisch und biologisch nur schwer nachweisen.

### 2.1.2. Die veränderliche Sequenz von Proteinen

Die kristallographischen und vergleichenden chemischen Untersuchungen an Proteinen zeigen, daß das Innere des kugelförmigen Proteins widerstandsfähig gegen eine evolutionäre Abwandlung ist, während an der Oberfläche Aminosäuren in

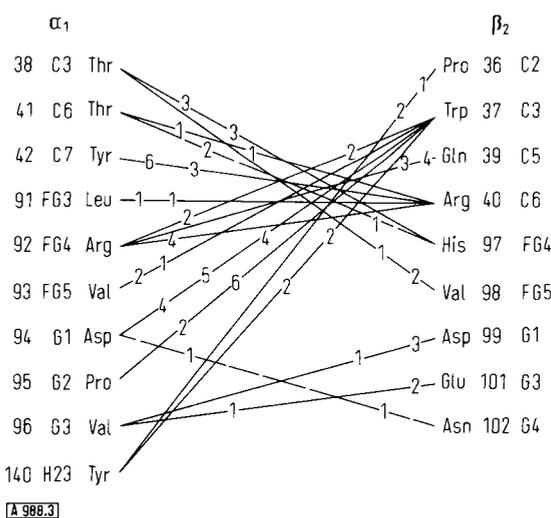


Abb. 3.

Kontakte zwischen den Resten ungleicher Ketten ( $\alpha_1$ - $\beta_2$ ) beim Hämoglobin. Die Kontakte umfassen zehn Reste von  $\alpha_1$  und neun Reste von  $\beta_2$ , die alle – mit Ausnahme der Reste Nr. 38 und 96 von  $\alpha_1$  und der Reste Nr. 39 und 101 von  $\beta_2$  – unveränderlich sind. Durchgezogene Linien entsprechen van-der-Waals-Bindungen; gestrichelte Linien zeigen Wasserstoff-Bindungen an. Die Zahlen bei den Linien geben an, wieviele Atome aus den beiden Resten von jeder Seite beteiligt sind (insgesamt 80 Atome) (nach [18]).

unterschiedlichem Maße ausgetauscht werden können. Die biologische Bedeutung der Moleküloberfläche ändert sich von Stelle zu Stelle; so sind die aktiven Zentren der Enzyme (für Katalyse, Spezifität, Coenzym-Anlagerung usw.), bei polymeren Enzymen die für die Wechselwirkung zwischen den Untereinheiten wichtigen Zonen, bei Protein-Hormonen die Zentren für die Anlagerung an die Rezeptoren usw. zweifellos weniger variabel als andere Teile, deren Funktion sich z. B. auf das Vorhandensein hydrophiler Gruppen beschränkt, die indirekt für die korrekte Faltung der Kette verantwortlich sind.

Die wenigen Enzyme, deren dreidimensionale Strukturen bekannt sind und bei denen man die Funktion der Oberfläche untersuchen kann, wurden leider noch nicht bei verschiedenen Arten verglichen. Die Kontaktzonen zwischen Cytochrom c und Cytochrom-Reduktase oder Cytochrom-Oxidase sind unbekannt; Myoglobin und Hämoglobin scheinen keine biologischen Funktionen an der Oberfläche zu haben. Interessant ist jedoch unter diesem Gesichtspunkt der Vergleich von proteolytischen Enzymen.

### 2.1.2.1. Biologische Bedeutung der Substitutionen

Wenn man annimmt, daß die molekulare Evolution, ebenso wie die morphologische Evolution, einen Fortschritt bedeutet, so kommt den Substitutionen an aktiven Zentren fundamentale Bedeutung zu. Es liegt nahe, „neutrale“ und „aktive“ Substitutionen zu unterscheiden: „Neutrale“ Substitutionen lassen die Funktion unverändert, kommen aber wahrscheinlich häufiger vor und zeigen bis zu einem gewissen Grade den Zeitraum an, vor dem sich zwei Arten getrennt haben. Durch „aktive“ Substitutionen werden dagegen die biologischen Eigenschaften eines Moleküls verändert.

Chemisch muß man zwischen konservativen Substitutionen (Austausch z. B. eines basischen oder eines hydrophoben Restes gegen einen gleichartigen Rest) und verändernden Substitutionen unterscheiden; es ist jedoch manchmal schwierig, eine Substitution der einen oder anderen Kategorie zuzuordnen. Wenn man sehr viele Sequenzen untersucht<sup>[6]</sup>, so fällt auf, daß Substitutionen chemisch äquivalenter Aminosäuren, wie Arg/Lys, Val/Ile und Ser/Thr, häufiger als andere Fälle vorkommen scheinen. Die Frage ist aber, ob diese Substitutionen wirklich als Veränderungen auf funktioneller Ebene anzusehen sind. Immerhin entspricht der ausgetauschten Aminosäure aber auf genetischer Ebene ein verändertes Triplet<sup>[2]</sup>, und eine konservative Substitution setzt ebenso wie eine verändernde Substitution eine Mutation voraus. Die Art der Substitutionen hängt also vom genetischen Code (benachbarte Triplets) und von den physikalisch-chemischen Eigenschaften der Reste (benachbarte funktionelle Gruppen und Konfigurationen) ab. Der zweite Faktor wird mit zunehmender Spezifität des Restes für die Funktion wichtiger, und im Extremfall ist der Rest nicht austauschbar.

Wenn man die „Mutabilität“ einer Aminosäure, ungeachtet ihrer Position, bei einer großen Zahl von Proteinen betrachtet, so fällt auf, daß einige Aminosäuren (wie Serin) sehr häufig, andere (wie Cystein und Tryptophan) dagegen nur selten ausgetauscht werden<sup>[6]</sup>. Die Möglichkeit der Substitution kann davon abhängen, ob unter den 20 natürlichen Aminosäuren geeignete Analoga vorhanden sind. Sehr nützlich wäre es zweifellos auch, die Gründe dafür zu ermitteln, warum bei der Evolution die Zahl der in Proteinen enthaltenen Aminosäuren

auf 20 beschränkt blieb und warum einige von diesen praktisch übereinstimmende Paare bilden.

### 2.1.2.2. Ungleichmäßige Substitutionsgeschwindigkeiten bei Protein-Familien

Die morphologische Evolution verläuft nicht mit gleichmäßiger Geschwindigkeit; eine Entwicklungslinie kann Perioden intensiver Veränderungen, Perioden geringer Veränderungen und auch Zeiten totalen Stillstandes zeigen (z. B. bei den lebenden Fossilien). Da jeder Organismus die verschiedensten Proteine enthält, könnte man fragen, ob sich alle Protein-Familien in einem mehr oder weniger ähnlichen Rhythmus ändern oder ob jede Familie ihr eigenes Schicksal hat. Tatsächlich entwickelten sich die Organe verschieden schnell, zweifellos in Abhängigkeit von ihrer Funktion, und so muß man das gleiche auch für die Proteine annehmen.

Wenn man a priori voraussetzt, daß die Struktur-Gene dieselbe Variabilität haben, müßte die Substitutionsgeschwindigkeit innerhalb einer Protein-Familie davon abhängen, wieviele Reste am Aufbau eines aktiven Zentrums beteiligt sind. Diese Hypothese kann bisher nicht überprüft werden; dagegen kann man die Zahl an Punktmutationen, die in jeder Familie vorgekommen sind, dadurch abschätzen, daß man die Substitutionen, den genetischen Code und ggf. notwendige „Zwischenstufen“ bei den Substitutionen berücksichtigt, die erhaltene Zahl auf 100 Reste umrechnet und sie auf einen Zeitraum von 100 Millionen Jahren bezieht (Tabelle 2)<sup>[6]</sup>. Obwohl es sich hierbei nur um grobe Näherungswerte handelt, die vor allem deshalb kritisiert werden können, weil eine gleichmäßige Änderungsgeschwindigkeit innerhalb einer Protein-Linie angenommen wird, ergibt sich klar, daß sich einige Protein-Familien wesentlich stärker als andere verändern.

Tabelle 2. Anzahl n der angenommenen Punktmutationen bei Proteinen, bezogen auf 100 Reste und 100 Millionen Jahre [a].

Proteine	n
Fibrinopeptide	90
Wachstumshormon	37
Pankreas-Ribonuclease	33
Immunglobuline	32
κ-Kette, C-Region	39
κ-Kette, V-Regionen	33
γ-Kette, C-Regionen	31
λ-Kette, C-Region	27
Lactalbumin	25
Hämoglobin-Ketten	14
Myoglobin	13
Pankreastrepsin-Inhibitor	11
Tierisches Lysozym	10
Gastrin	8
Melanotropin β	7
Encephalitogenes Protein der Myelin-Membran	7
Trypsinogen	5
Insulin	4
Cytochrom c	3
Glycerinaldehyd-3-phosphat-Dehydrogenase	2
Histon IV	0.06

[a] Für die meisten Proteine wurde die Zeit ausgehend von dem Zeitpunkt, zu dem sich die Säugetier-Ordnungen trennten (etwa vor 75 Millionen Jahren) extrapoliert. Für Lysozym, Trypsinogen, Insulin, Glycerinaldehyd-3-phosphat-Dehydrogenase und Histon IV wurden folgende Divergenzzeiten von der Linie der Säugetiere zugrundegelegt: Vögel (300 Millionen Jahre), Fische (400 Millionen Jahre), Insekten (900 Millionen Jahre) und Pflanzen (1500 Millionen Jahre) (nach [6]).

Es ist schwer zu verstehen, warum die Familie der Myoglobine stabiler als diejenige der α- und β-Hämoglobin-Ketten ist,

obwohl im Myoglobin offensichtlich weniger funktionale Zwänge als in den Hämoglobinen vorliegen; auch läßt sich nicht erklären, weshalb die  $\beta$ -Familie variabler als die  $\alpha$ -Familie ist, denn die beiden Ketten scheinen die gleichen Aufgaben zu haben. Cytochrom c ist besonders stabil, aber noch rätselhafter wirkt die frappierende Unveränderlichkeit der Histone: Sie unterscheiden sich bei Erbsen und bei Rindern nur um zwei Reste, obwohl sich deren Entwicklungslinien schon vor 1500 Millionen Jahren getrennt haben.

### 2.1.2.3. Stammesgeschichte und Protein-Evolution

Die Paläontologie hat gezeigt, wann ungefähr sich die Linien der Wirbeltiere voneinander getrennt haben, und es ist daher möglich, die Zeiträume seit der Trennung der Linien und die Zahl der beobachteten Substitutionen bei mehreren Protein-Familien miteinander zu vergleichen. Offensichtlich erhöht sich die Zahl der Substitutionen mit zunehmenden Zeiträumen, und unter zusätzlichen Annahmen lassen sich phylogenetische Stammbäume berechnen, die nur auf der Zahl der Substitutionen basieren<sup>16, 281</sup>. Diese Stammbäume stimmen gut mit den paläontologisch ermittelten überein.

Die Tatsache, daß bei einer gegebenen Protein-Familie, z.B. den  $\beta$ -Hämoglobinen (Tabelle 3), die Zahl der Substitutionen zwischen zwei Arten eine Funktion des Zeitraums seit der

Tabelle 3. Anzahl n der Abweichungen (Substitutionen) in der Aminosäure-Sequenz der  $\beta$ -Hämoglobine mehrerer Tierarten und paläontologischer Zeitraum t seit der Trennung von der gemeinsamen Stammlinie.

Zoologische Einteilung	t [10 <sup>6</sup> a]	n [%]
<i>Klasse Mammalia (Säugetiere)</i>		
<i>Unterklasse Eutheria</i>		
Ordnung Primaten:		
Hominiden		
(Mensch)	0	0
(Gorilla)	20	1
Cercopithecidae		
(Rhesus-Affe)	60	5
Ordnung Artiodactyla:		
Ruminantia		
(Schaf)	75	18
Suiformes		
(Schwein)	75	16
Tylopodes		
(Lama)	75	14
Ordnung Perissodactyla:		
(Pferd)	75	17
Ordnung Rodentia:		
(Maus)	75	18
<i>Unterklasse Metatheria</i>		
Ordnung Marsupialia:		
(Känguruh)	130	26
<i>Klasse Amphibia (Amphibien)</i>		
Ordnung Anura:		
(Frosch)	320	46

Trennung der beiden Arten ist und nicht von den ausgewählten Arten abhängt, legt die Vermutung nahe, daß die beobachteten Substitutionen vom Standpunkt der natürlichen Auslese aus im wesentlichen „neutral“ sind. Die Sequenzen der Proteine wären damit biologische „Uhren“. Da die zoologischen Gruppen aufgrund morphologischer Eigenschaften abgegrenzt werden, die sich mit zunehmendem Zeitraum seit Trennung der Linien stärker differenzieren, müßte der Anteil an Substitutionen bezogen auf eine Vergleichs-Art prinzipiell die Identifizierung der Klasse oder Unterklasse derjenigen Art erlauben,

von der das Protein erhalten wurde. Dabei ist es jedoch notwendig, eine rezente Art zum Vergleich heranzuziehen. Das Neunauge z. B. gehört einer Linie (Cyclostomen) an, die sich von derjenigen, die zu den anderen Wirbeltieren führte, vor etwa 500 Millionen Jahren getrennt hat; würde man die Proteine der Cyclostomen als Bezugssubstanzen wählen, so sollte sich bei allen anderen Wirbeltieren der gleiche Anteil an Substitutionen ergeben.

Es sei darauf hingewiesen, daß ungefähr gleiche Anteile an Substitutionen, die für ähnliche Zeiträume gefunden werden, nicht den Schluß zulassen, daß die Polypeptid-Ketten an den gleichen Positionen verändert wurden. Auch Ausnahmen wurden beobachtet: Beispielsweise haben die Vögel die gleichen Neurohypophysen-Hormone wie die Amphibien, von denen sie sich vor ungefähr 350 Millionen Jahren ableiteten, aber andere Neurohypophysen-Hormone als die Säugetiere, von denen sie sich erst viel später trennten. Die Abwandlung von Proteinen, ein passives Phänomen, das häufig auftritt, darf nicht darüber hinwegtäuschen, daß eine wirkliche Evolution definitionsgemäß erst dann vorliegt, wenn die Abwandlung des Proteins von einer Änderung seiner Funktion begleitet wird.

## 2.2. Länge der Polypeptid-Kette

Bei der Untersuchung einer Familie von homologen Proteinen fällt auf, daß die Zahl der Reste praktisch immer konstant bleibt: Alle Cytochrome c der Wirbeltiere enthalten 104 Aminosäure-Reste, alle  $\alpha$ -Hämoglobine 141 oder 142, alle Neurohypophysen-Hormone neun Reste usw. Um die Bedeutung dieser Konstanz zu ermessen, ist es notwendig, auf die Protein-Biosynthese einzugehen; sie umfaßt die Transkription des Struktur-Gens, d. h. einer Desoxyribonucleinsäure (DNA), in eine Messenger-Ribonucleinsäure (mRNA) und anschließend die Übersetzung der mRNA in eine Aminosäure-Sequenz. Eine mRNA codiert mehrere Proteine; die Übersetzung eines jeden umfaßt das Ablesen einer Sequenz von Triplets, ausgehend von einem Anfangstriplett (AUG oder GUG) bis zu einem Endtriplett (UAA, UAG oder UGA). Neben der RNA-Polymerase der Transfer-Ribonucleinsäuren (tRNA), welche die Anfangs- und Endsignale erkennen, nehmen auch Proteinfaktoren, die bisher nicht befriedigend identifiziert wurden, an diesem Prozeß teil<sup>129</sup>. Es ist merkwürdig, daß ein Vorgang, an dem so viele Moleküle beteiligt sind und der daher für Mutationen besonders anfällig zu sein scheint, zu einem praktisch konstanten Ergebnis führen kann. Die wenigen beobachteten Abweichungen sind daher besonders aufschlußreich.

### 2.2.1. Verkürzung und Verlängerung der Kette

Die bekannten Familien homologer Proteine umfassen nur wenige Beispiele aus niederen Arten. In einigen Fällen (Cytochrom c) scheint im Laufe der Evolution eine geringe Verkürzung stattgefunden zu haben.

#### 2.2.1.1. N-terminale Änderungen der Kette

Die Cytochrome c von Pflanzen, Pilzen und Insekten enthalten 8, 6–4 bzw. 4 Reste mehr als die Cytochrome c der Wirbeltiere, und zwar verlängern diese Reste den N-terminalen Teil der Kette<sup>91</sup>. Es scheint, als ob das Anfangssignal verschoben wurde, so daß der spätere Beginn bei den höheren Arten eine kürzere Kette ergibt.

Dieses Phänomen läßt sich jedoch nicht verallgemeinern; z. B. haben die Hämoglobine der wirbellosen Tiere und entsprechende Proteine der Pflanzen ungefähr die gleiche Anzahl an Resten wie die  $\alpha$ - oder  $\beta$ -Ketten der Wirbeltier-Hämoglobine. Eine N-terminale Auslassung von sechs Resten wurde beim  $\beta$ -Hämoglobin des Frosches beobachtet<sup>[16]</sup>, hat aber offenbar keinerlei evolutionäre Bedeutung. Immerhin scheint es, als ob die N-terminalen Enden der Ketten einer Abwandlung eher zugänglich sind als der zentrale Teil.

#### 2.2.1.2. C-terminale Änderungen der Kette

Am C-terminalen Ende scheinen Verkürzungen oder Verlängerungen seltener als am N-terminalen vorzukommen; offenbar läßt sich das Endsignal weniger leicht verschieben, oder seine Verschiebung ist unvereinbar mit der Erhaltung eines funktionsfähigen Moleküls.

Auf eine vorzeitige Beendigung der Biosynthese des Corticotropins könnte das Vorkommen eines Peptids in Schweinen und anderen Säugetieren hinweisen, das wegen seiner speziellen biologischen Eigenschaften als  $\alpha$ -Melanocyten-stimulierendes Hormon ( $\alpha$ -MSH) bezeichnet wird; dieses hat die Sequenz der ersten 13 Reste des Corticotropins, welches seinerseits aus 39 Resten besteht.

In einem unlängst isolierten, abnormalen Human-Hämoglobin (Constant Spring-(CS-)Hämoglobin) enthält die  $\alpha$ -Kette 172 anstelle der üblichen 141 Reste. Es ist bemerkenswert, daß dies einer Verlängerung um 31 Reste am C-terminalen Ende entspricht, als ob das Endsignal in Richtung auf eine Verlängerung verschoben worden wäre<sup>[30]</sup>. Eine C-terminale Verlängerung um zehn Reste wurde auch in der  $\beta$ -Kette eines Human-Hämoglobins beobachtet<sup>[31]</sup>. Obwohl das CS-Hämoglobin im untersuchten Fall nur 1–2% des gesamten Hämoglobins ausmachte und die Moleküle abnormal sind, könnte dieser Vorgang auch bei der Evolution vorgekommen sein.

#### 2.2.2. Einfügungen und Auslassungen im Inneren der Kette

Beim Vergleich von homologen Proteinen findet man manchmal, daß bei einer Art gegenüber dem allgemeinen Molekülschema ein oder zwei Reste zusätzlich vorhanden sind oder fehlen. Beispielsweise hat Karpfen- $\alpha$ -Hämoglobin 142 anstatt 141 Reste (Einfügung), und die  $\beta$ -Hämoglobine von Rindern und Schafen haben 145 anstatt 146 Reste (Auslassung). Die Homologie der Sequenzen gestattet es, die Änderung im Inneren der Kette zu lokalisieren. Einfügungen und Auslassungen dieser Art wurden auch in abnormalen Human-Hämoglobinen gefunden; diese Beobachtungen lassen sich durch reziprokes Überkreuzen (crossing-over) innerhalb der Gene, durch das die Peptid-Sequenzen beeinflußt werden, erklären. Die Einfügungen und Auslassungen im Inneren der Peptid-Kette betreffen offenbar höchstens fünf, oft aber weniger aufeinanderfolgende Reste. Ihr Einfluß auf die allgemeine Konformation und die Funktion scheint im großen und ganzen gering zu sein.

### 3. Polymorphie: Isoproteine

Durch äußerst sorgfältige Reinigung von Proteinen konnte gezeigt werden, daß gleiches Gewebe bei der gleichen Art sehr oft zwei oder mehrere Moleküle enthält, die offensichtlich die gleiche Funktion ausüben und sehr ähnliche Aminosäure-

Sequenzen haben. Diese Polymorphie ist nicht auf eine unvollständige Dominanz zurückzuführen, denn jedes Protein entspricht einem besonderen Struktur-Gen. Solche Moleküle werden als Isoproteine (Isoenzyme, Isohormone) bezeichnet, und man nimmt an, daß sie durch Duplikation eines früher vorhandenen Struktur-Gens hervorgegangen sind. Das Ausmaß der Polymorphie hängt vom Protein und von der Art ab; es ist jedoch sehr ungewöhnlich, wenn ein Protein keine Polymorphie aufweist. Cytochrom c scheint z. B. bei Wirbeltieren keine Polymorphie zu zeigen; in Hefe kommt es dagegen als Iso-1-cytochrom c und als Iso-2-cytochrom c vor<sup>[32]</sup>.

Die biologische Bedeutung dieser Gen-Vervielfachung ist noch unklar. Möglicherweise schützt das Vorhandensein von zwei oder mehreren Genen, die Proteine mit ähnlichen Funktionen produzieren, die Art gegen eine inaktivierende Mutation, welche bei einem einzigen Gen tödlich sein würde. Es ist aber auch denkbar, daß aufeinanderfolgende Mutationen schließlich dahin führen, daß eines der Proteine eine andere Funktion ausüben kann, und daß damit die Duplikation der erste Schritt bei der Bildung eines neuen Proteins ist. Es sei jedoch darauf hingewiesen, daß die Mengenverhältnisse der Isoproteine meist beträchtlich variieren, und daß das Verschwinden des häufigsten Isoproteins durch Mutation nicht von einer entsprechenden Zunahme der weniger häufigen Proteine begleitet zu sein scheint.

Die Isoproteine sind nicht sehr eindeutig definiert: da aber die Proteine gewöhnlich nach ihrer Funktion benannt werden, sind sie allgemein als Proteine anzusehen, die praktisch gleiche Wirkungen, aber etwas abweichende Aminosäure-Sequenzen haben.

#### 3.1. Isoproteine im engeren Sinne: Die Chymotrypsinogene

Die Chymotrypsinogene A und B wurden in Rinder-Pankreas gefunden. Sie sind etwa gleich groß und bestehen beide aus einer Polypeptid-Kette von 245 Resten; sehr wahrscheinlich haben sie auch ähnliche Konformationen<sup>[21]</sup>. Die Art der Aktivierung durch Trypsin unter Trennung der Arg-15—Ile-16-Bindung ist bei beiden Trypsinogenen ähnlich, jedoch wird Chymotrypsinogen B unter vergleichbaren Bedingungen 15mal schneller aktiviert. Da der einzige Unterschied in den ersten 35 Resten der Ketten ein Alanin im B-Protein anstelle eines Serins im A-Protein ist, scheint es, als ob die Aktivierungsregion im B-Protein wegen der Abwesenheit einer Wasserstoff-Bindung flexibler wäre. Zwei der drei autokatalytischen Spaltungen, die der Aktivierung beim Chymotrypsinogen A folgen, werden auch bei der Aktivierung des B-Proteins gefunden; die Spezifitäten stimmen bei beiden Enzymen praktisch überein.

Es ist bemerkenswert, daß die beiden Chymotrypsinogene, die funktionell so sehr ähnlich sind, sich durch 52 Reste (entsprechend 21% ihrer Sequenz) unterscheiden und in annähernd gleichen Mengen vom Pankreas produziert werden. Bisher läßt sich kein Grund für diese Duplikation angeben. Tatsächlich muß aber eine derartige Duplikation sogar zweimal vorgekommen sein, denn es wurde auch ein weniger gut charakterisiertes Chymotrypsinogen C isoliert. Die drei Enzyme wurden auch bei Schweinen gefunden; andere Arten sind noch nicht gut genug untersucht, um Schlüsse darüber zuzulassen, ob es sich hierbei um ein allgemein bei Wirbeltieren vorkommendes Phänomen handelt.

### 3.2. Isoproteine im weiteren Sinne: Die Globine

Die Hämoglobine der Wirbeltiere (mit Ausnahme der Cyclostomen) sind tetramere Moleküle, die je zwei Ketten vom  $\alpha$ - und  $\beta$ -Typ enthalten ( $\alpha_2\beta_2$ ). Die  $\alpha$ - und  $\beta$ -Ketten sind Isoproteine in dem Sinne, daß ihre Funktionen einander offensichtlich ähneln, während ihre Sequenzen verwandt, aber ungleich sind. Die vollständigen Sequenzen der Ketten sind für etwa 20 Wirbeltier-Arten, vorwiegend Säugetiere, bekannt, und man kann annehmen, daß die Duplikation ein allgemeines Phänomen ist. Die Polymorphie kommt sogar auch bei solchen Arten vor, die monomeres Hämoglobin haben, z. B. bei Neunaugen (Cyclostomen) und wirbellosen Tieren sowie bei den entsprechenden Proteinen der Pflanzen.

Die  $\alpha$ - und  $\beta$ -Ketten sind ungefähr gleich lang ( $\alpha$ -Kette: 141 Reste;  $\beta$ -Kette: 146 Reste) und haben sehr ähnliche Konformationen, die mit der des Myoglobins eng verwandt sind<sup>[18]</sup>. Beim Menschen finden sich 57% Substitution zwischen der  $\alpha$ - und der  $\beta$ -Kette. Auch eine kleine Menge eines Hämoglobins der Formel  $\alpha_2\delta_2$  wurde gefunden, wobei die  $\delta$ -Kette wie die  $\beta$ -Kette 146 Reste enthält und sich von dieser durch nur 7% Substitution unterscheidet<sup>[16]</sup>. Der menschliche Fetus enthält ein spezielles Hämoglobin der Formel  $\alpha_2\gamma_2$ ; die  $\gamma$ -Kette besteht wie die  $\beta$ -Kette aus 146 Resten und zeichnet sich gegenüber dieser durch 27% Substitution aus. Weitere Ketten,  $\epsilon$  und  $\zeta$ , kommen ebenfalls im Embryo vor, doch wurden sie noch nicht aufgeklärt. Auch bei anderen Säugetier-Arten ist es nicht ungewöhnlich, daß das erwachsene Tier zwei oder mehrere Hämoglobine enthält, die sich entweder in der  $\alpha$ - oder in der  $\beta$ -Kette unterscheiden; da innerhalb einer Art immer wenigstens ein Vertreter vom  $\alpha$ -Typ und vom  $\beta$ -Typ vorliegt, können stets mehrere Proteine in unterschiedlichen Mengen auftreten, von denen jedes einem anderen Struktur-Gen entspricht.

Es ist daher möglich, hier zwischen zwei Graden an Polymorphie und zwischen zwei Kategorien von Isoproteinen zu unterscheiden, nämlich zwischen den Isoproteinen im engeren Sinn, welche gleichviele Reste enthalten (z. B. die Gruppen  $\beta$ ,  $\delta$  und  $\gamma$ ), und den Isoproteinen im weiteren Sinn, die in der Anzahl ihrer Reste differieren und die auch untereinander eine stärkere Substitution aufweisen. Wenn man davon ausgeht, daß der Anteil an Substitution ein Maßstab für die Zeit ist, die seit der Duplikation verstrichen ist, d. h. für die „evolutionäre Distanz“, so kann man annehmen, daß beim

Menschen z. B. die  $\alpha$ - $\beta$ -Duplikation mit 57% Substitution zuerst stattfand; es folgten die  $\beta$ - $\gamma$ -Duplikation mit 27% Substitution und zuletzt die  $\beta$ - $\delta$ -Duplikation mit 7% Substitution. Da Myoglobin im Muskel darüber hinaus eine Funktion hat, die derjenigen des Hämoglobins weitgehend entspricht, und da es in der Kettenlänge (153 Reste) und besonders in der Konformation den  $\alpha$ - und  $\beta$ -Ketten sehr ähnlich ist<sup>[18]</sup>, kann man Myoglobin auch als Isoprotein im weiteren Sinne ansehen. Human-Myoglobin unterscheidet sich von der  $\alpha$ -Kette des Menschen durch 74% Substitution<sup>[13,31]</sup>; die Duplikation müßte daher früher als die  $\alpha$ - $\beta$ -Duplikation stattgefunden haben. Es ist interessant, daß der Längenunterschied nur dann auftritt, wenn die Substitution mehr als 50% beträgt; das bedeutet nach obiger Interpretation, daß die Duplikation sehr lange zurückliegt. Geht man davon aus, daß die erste Duplikation bei den Globinen die Myoglobin- $\alpha$ -Kette-Duplikation gewesen ist<sup>[31]</sup>, so fällt auf, daß das Protein, das zum Myoglobin wurde, sich nicht weiter dupliziert hat, und daß es auch weniger Substitutionen unterlag als das Protein, das zu den Hämoglobin-Ketten wurde. Die Faktoren, die dazu führen, daß sich ein Struktur-Gen leichter als ein anderes dupliziert, sind bisher ebenso unbekannt wie auch die Faktoren, durch die die Substitutionen in der einen Linie stärker als in der anderen begünstigt werden.

## 4. Die biologische Differenzierung der Proteine

### 4.1. Differenzierung durch Substitutionen: Analoge Proteine

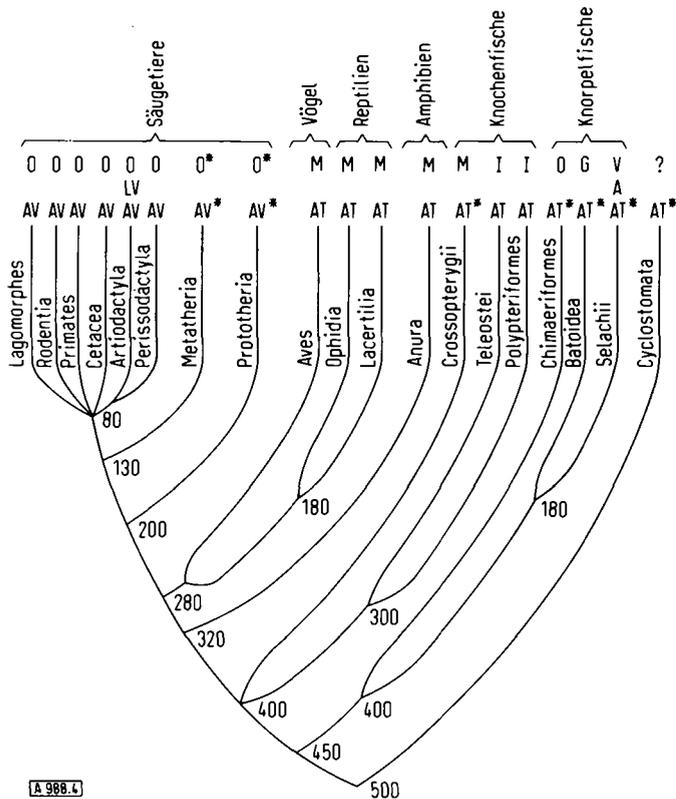
Manche Proteine mit unterschiedlichen biologischen Eigenschaften sind einander in der Anzahl der Reste und in den Aminosäure-Sequenzen so ähnlich, daß man sie wohl als Abkömmlinge eines gemeinsamen Stamm-Moleküls ansehen darf. Diese funktionell verschiedenen, aber chemisch verwandten Proteine werden im folgenden als analoge Proteine bezeichnet. Das Ausmaß der Ähnlichkeit variiert offensichtlich innerhalb jeder Gruppe; es kann im Extremfall auch strittig sein, ob zwischen den Proteinen eine Beziehung besteht.

#### 4.1.1. Die Neurohypophysen-Hormone

Obwohl die Neurohypophysen-Hormone zu den Peptiden und nicht zu den Proteinen gehören, sollen sie als Beispiele für die Umwandlung von isologen in analoge Proteine dienen.

Tabelle 4. Schema für die hypothetische Entwicklung der Neurohypophysen-Hormone aus dem Stamm-Molekül. Durch eine Gen-Duplikation und mehrere anschließende Einzelsubstitutionen in den Positionen 3, 4 oder 8 entstehen zwei getrennte Molekül-Linien.

	Stamm-Molekül											
	1	2	3	4	5	6	7	8	9			
	Cys-Tyr-?-?-Asn-Cys-Pro-?-Gly(NH <sub>2</sub> )											
	Duplikation											
Knorpelfische (Rochen) (Hundshai)	Glumitocin	Ile <sup>3</sup> -Ser <sup>4</sup> -Gln <sup>8</sup>	Arg-Vasotocin	Ile <sup>3</sup> -Gln <sup>4</sup> -Arg <sup>8</sup>	Asparagtocin	Ile <sup>3</sup> -Asn <sup>4</sup> -Leu <sup>8</sup>	Arg-Vasotocin	Ile <sup>3</sup> -Gln <sup>4</sup> -Arg <sup>8</sup>	Valitocin	Ile <sup>3</sup> -Gln <sup>4</sup> -Val <sup>8</sup>	Arg-Vasotocin	Ile <sup>3</sup> -Gln <sup>4</sup> -Arg <sup>8</sup>
Knochenfische	Isotocin	Ile <sup>3</sup> -Ser <sup>4</sup> -Ile <sup>8</sup>	Arg-Vasotocin	Ile <sup>3</sup> -Gln <sup>4</sup> -Arg <sup>8</sup>	Mesotocin	Ile <sup>3</sup> -Gln <sup>4</sup> -Ile <sup>8</sup>	Arg-Vasotocin	Ile <sup>3</sup> -Gln <sup>4</sup> -Arg <sup>8</sup>				
Amphibien, Reptilien, Vögel [34]												
Säugetiere (außer Schwein) Schwein	Oxytocin	Ile <sup>3</sup> -Gln <sup>4</sup> -Leu <sup>8</sup>	Arg-Vasopressin	Phe <sup>3</sup> -Gln <sup>4</sup> -Arg <sup>8</sup>	Oxytocin	Ile <sup>3</sup> -Gln <sup>4</sup> -Leu <sup>8</sup>	Lys-Vasopressin	Phe <sup>3</sup> -Gln <sup>4</sup> -Lys <sup>8</sup>				



A 988.4

Abb. 4. Entwicklung der Neurohypophysen-Hormone und Evolution der Arten. Die Buchstaben geben die Hormone an, die bei rezenten Vertretern der einzelnen Tier-Gruppen gefunden werden (ein Sternchen bedeutet, daß die Hormone nicht chemisch, sondern pharmakologisch identifiziert wurden): O = Oxytocin, AV = Arginin-Vasopressin, LV = Lysin-Vasopressin, M = Mesotocin, I = Isotocin, A = Aspartocin, V = Valitocin, G = Glunitocin, AT = Arginin-Vasotocin. Die Zahlen geben die paläontologischen Zeiträume seit Trennung der Linien in Millionen Jahren an (Formeln der Hormone siehe Tabelle 4).

Bisher wurden neun Neurohypophysen-Hormone, sämtlich Nonapeptide, bei etwa 40 Arten aus allen Klassen der Wirbeltiere (mit Ausnahme der Cyclostomen) identifiziert (Tabelle 4, Abb. 4).

Die biologischen Wirkungen der aus den niederen Wirbeltieren isolierten Peptide wurden vor allem an Säugetieren gemessen, so daß man ihre tatsächliche Funktion nicht genau kennt. Es ist jedoch interessant, daß in den Klassen, die direkt vor

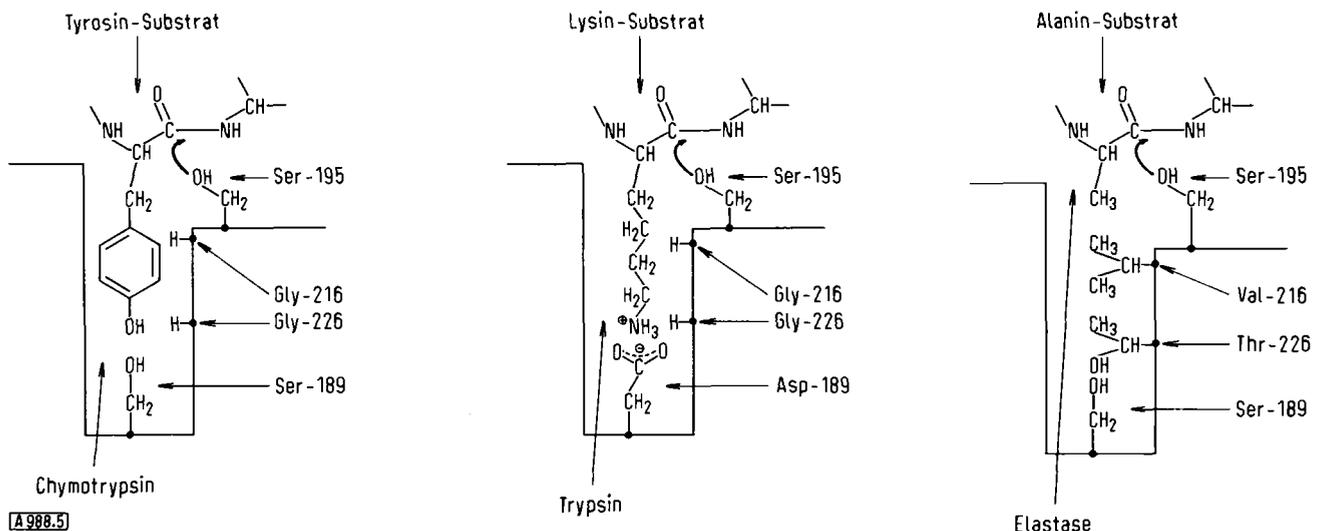
den Säugetieren auftraten, also bei den Amphibien und den Reptilien, sowohl Mesotocin (Ile<sup>8</sup>-Oxytocin) als auch Arginin-Vasotocin (Arg<sup>8</sup>-Oxytocin) eine ausgeprägte Oxytocin-Wirkung haben und daß man sie als Isopeptide ansehen kann. Vasotocin nimmt jedoch auch an der Regulierung des Wasser- und Elektrolyt-Haushalts teil und kann damit auch auf solche Rezeptoren einwirken, auf die Mesotocin keinen Einfluß hat. Bei den Säugetieren treten zwei neue Moleküle auf, Oxytocin und im allgemeinen Arginin-Vasopressin (Phe<sup>3</sup>-Arg<sup>8</sup>-Oxytocin), welche den Uterus und die Milchdrüse bzw. die Niere beeinflussen. Die Wirkung von Oxytocin auf die Niere und von Arginin-Vasopressin auf den Uterus und die Milchdrüse sind vernachlässigbar gering; die Ähnlichkeit der beiden Peptide ergab sich erst bei der Bestimmung ihrer chemischen Struktur.

Man darf annehmen, daß die Differenzierung der Funktionen bei den Proteinen ähnlich wie bei den Peptiden (wenn auch komplizierter) verlaufen ist: Nach der Duplikation wurden die Ketten durch Substitutionen und durch Abwandlung der Länge fortschreitend verändert, bis sich bei einem bestimmten Grad der Unterschiedlichkeit auch die Funktionen voneinander trennten – obwohl die Grundstrukturen der Proteine immer noch verwandt sind.

#### 4.1.2. Die Pankreas-Serin-Proteasen

Drei proteolytische Enzyme des Pankreas – Chymotrypsin, Trypsin und Elastase – haben zwar verschiedene Spezifitäten, wurden aber, zunächst durch Bestimmung der Aminosäure-Sequenzen<sup>[20-23]</sup> und dann durch Vergleich der Konformationen<sup>[24-26]</sup> als analog erkannt. Die Kettenlängen unterscheiden sich nur wenig: 241 Reste beim Rinder-Chymotrypsin, 223 Reste beim Rinder-Trypsin und 240 Reste bei der Schweine-Elastase. Die Konformationen aller drei Enzyme sind bemerkenswert ähnlich; die Unterschiede in der Spezifität lassen sich durch Abwandlungen in der Konformation der Substrat-Bindungsstelle erklären<sup>[3,4a]</sup>.

Chymotrypsin, Trypsin und Elastase haben eine Tasche, in welche die Seitenkette der Aminosäure hineinragt, die über die Spezifität entscheidet (Abb. 5). Die Position 216 am Tascheneingang wird bei Chymotrypsin und Trypsin von einem Glycin-Rest eingenommen, der die Einführung einer langen,



A 988.5

Abb. 5. Unterschiede bei den Substrat-Bindungsstellen (langer Pfeil von unten) von Chymotrypsin, Trypsin und Elastase, schematisch. Der nucleophile Angriff des Serin- $\gamma$ -Sauerstoffs im aktiven Zentrum auf das Substrat ist durch den gebogenen Pfeil angedeutet (nach [34a]).

sperrigen Seitenkette gestattet; dagegen liegt bei der Elastase ein Valin-Rest vor, der den Zugang zur Tasche auf kleine aliphatische Reste beschränkt (Numerierung bezogen auf Rinder-Chymotrypsinogen A). Die Position 189 am Grunde der Tasche ist beim Trypsin durch Asparaginsäure besetzt, die mit basischen Resten eine Ionen-Bindung zuläßt; beim Chymotrypsin wird diese Position dagegen durch Serin eingenommen, das Peptid-Bindungen an neutralen, hydrophoben Resten spaltet<sup>[26]</sup>. Die unterschiedlichen Spezifitäten scheinen hier auf unterschiedlichen Resten (Substitutionen) in der Kette zu beruhen. Durch Untersuchungen an den „Trypsin-ähnlichen“ und „Chymotrypsin-ähnlichen“ Enzymen der niederen Wirbeltiere gelingt es vielleicht, die Evolution der Substrat-Tasche zurückzuverfolgen. Die drei Pankreas-Serin-Proteasen haben alle die hydrolytische Wirkungsweise beibehalten; ihre funktionellen Unterschiede beschränken sich offensichtlich nur auf die Spezifität der Substrat-Bindung.

#### 4.2. Differenzierung durch Hybridisierung

Wenn zwei ähnliche Struktur-Gene (die z. B. zwei Isoproteine codieren) im gleichen Chromosom liegen, ist eine fehlerhafte Anordnung der Chromatiden während der Chromosomen-Paarung mit reziprokem Überkreuzen (crossing-over) der nicht-allelen Gene möglich. Nach der Trennung besitzt jedes der beiden Chromosomen ein anderes Hybrid-Gen: Das eine Hybrid enthält den Anfangsteil des einen und den Endteil des anderen Eltern-Gens, während beim anderen Hybrid die umgekehrten Teile miteinander verbunden sind. Ein solches Hybrid codiert eine Polypeptid-Kette, welche die N-terminale Sequenz des einen und die C-terminale Sequenz des anderen Proteins hat.

Ein typisches Beispiel hierfür ist die Hybridisierung zwischen den  $\beta$ - und  $\delta$ -Ketten des Hämoglobins. Menschliches Lepore-Hämoglobin ist ein durch Mutation entstandenes Hämoglobin, das aus einer normalen  $\alpha$ -Kette und einer hybridisierten  $\delta$ - $\beta$ -Kette besteht. Anti-Lepore-Hämoglobin enthält eine normale  $\alpha$ -Kette und eine hybridisierte  $\beta$ - $\delta$ -Kette. Die  $\beta$ - und die  $\delta$ -Kette haben beide je 146 Reste wie auch die Hybriden  $\beta$ - $\delta$  und  $\delta$ - $\beta$ . Beim Boston-Lepore-Hämoglobin entsprechen die Reste 1–87 der  $\delta$ -Kette, die Reste 116–146 dagegen der

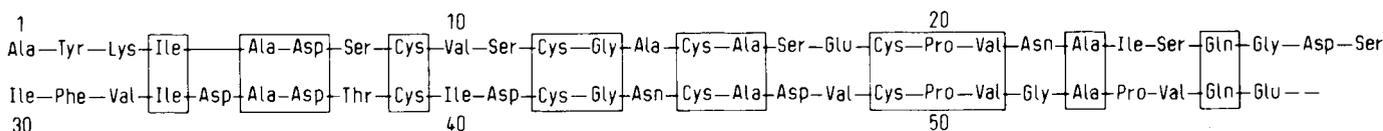


Abb. 6. Homologe Bereiche bei den beiden Teilen des Ferredoxin-Moleküls von *Clostridium pasteurianum*, die die Annahme einer intragenen Duplikation nahelegen (nach [6]).

$\beta$ -Kette. Da aber die Reste 88–115 der  $\delta$ - und  $\beta$ -Kette übereinstimmen, läßt sich die Überkreuzungsstelle nicht genau lokalisieren.

Bei Miyada-Anti-Lepore-Hämoglobin hat die  $\beta$ - $\delta$ -Kette die Reste 1–12 der  $\beta$ -Kette und die Reste 22–146 der  $\delta$ -Kette; da die Sequenz 13–21 bei beiden Ketten gleich ist, ist es auch hier nicht möglich, die Überkreuzungsstelle genau festzustellen.

Eine Hybridisierung wurde auch bei den aus Heringssperma erhaltenen Clupeinen postuliert<sup>[6]</sup>. Clupein Z stimmt mit Clu-

pein YI in der Sequenz 1–6 und mit Clupein YII in der Sequenz 6–30 überein.

Die Hybridisierung zwischen zwei Genen ist zwar vielleicht auf Gene gleicher Länge beschränkt, d. h. auf solche, die durch eine frühere Duplikation hervorgegangen sind; im Vergleich zur Differenzierung durch Punktmutation, bei der jeweils nur eine einzige Aminosäure substituiert wird, gestattet sie jedoch eine wesentlich schnellere Differenzierung.

#### 4.3. Differenzierung durch Duplikationen innerhalb des Gens

Einige Proteine haben innerhalb ihrer Polypeptid-Ketten Regionen, die sich zu wiederholen scheinen; dieser Befund legt die Vermutung nahe, daß sich zwei Tochter-Gene, die durch Duplikation des Ur-Struktur-Gens entstanden waren, vereinigten und von da an als ein einziges Gen wirkten. Die Kettenlänge wurde dadurch verdoppelt. Obwohl sich eine solche intragene Duplikation nicht sehr gut verstehen läßt, könnte diese Auslegung die diskontinuierliche Verlängerung der Polypeptid-Ketten und eine gewisse Periodizität innerhalb der Aminosäure-Sequenzen erklären.

##### 4.3.1. Ferredoxin

Ferredoxin ist ein eisenhaltiges, häm-freies Protein, das sich in photosynthetisierenden anaeroben Bakterien, Blaualgen, Grünalgen und höheren Pflanzen findet. Es nimmt an zahlreichen biochemischen Prozessen teil, z. B. an der Photosynthese, der Stickstoff-Fixierung, der Sulfat-Reduktion, an Dehydrogenierungen und an Redox-Reaktionen.

Bakterien-Ferredoxin besteht aus einer Kette mit 55 Resten und zeigt eine gewisse innere Periodizität; z. B. weisen die beiden Hälften der Ferredoxin-Kette von *Clostridium pasteurianum* (Abb. 6) 44 % homologe Anteile auf. Diese Erscheinung wird auch beim Ferredoxin anderer Bakterien beobachtet und deutet auf eine intragene Duplikation hin.

Das Ferredoxin höherer Pflanzen, z. B. Spinat, hat eine nahezu doppelt so lange Kette (97 Reste) wie das Bakterien-Ferredoxin; das Ferredoxin der Alge *Chromatium sp.* weist eine mittlere Länge (81 Reste) auf. Alle diese Ferredoxine haben homologe Regionen. Die Verlängerung gegenüber dem bakteriellen Fer-

redoxin scheint bei den Pflanzen-Ferredoxinen am N-terminalen und bei den Algen (*Chromatium sp.*) am C-terminalen Ende stattgefunden zu haben. Alle Beobachtungen sprechen für eine Verlängerung durch intragene Duplikation, wobei man eine Grundeinheit von 25–28 Resten zugrunde legen kann.

##### 4.3.2. Die Immunglobuline

Eine intragene Duplikation scheint auch bei den Immunglobulinen stattgefunden zu haben. Immunglobulin-Moleküle sind

aus zwei gleichen leichten (215–220 Reste) und zwei gleichen schweren Ketten (446 Reste) symmetrisch aufgebaut; die vier Ketten werden durch Disulfid-Brücken zusammengehalten. Die Tatsache, daß die schwere Kette nahezu genau doppelt so lang ist wie die leichte Kette, legt natürlich die Annahme einer Duplikation nahe. Darüber hinaus liegt aber offenbar auch innerhalb der leichten Kette zwischen den N- und C-terminalen Teilen eine Homologie vor. Im N-terminalen Teil scheinen jedoch Substitutionen vorzukommen; dieser wird daher als variabel (V) bezeichnet, der stabilere C-terminale Teil wird dagegen konstant (C) genannt. Die schwere Kette scheint aus einer V-Einheit und drei C-Einheiten zu bestehen. Man nimmt an, daß die Verlängerung diskontinuierlich in Blöcken von je ungefähr 110 Resten vor sich gegangen ist.

## 5. Fusion von Proteinen

Die oben genannten Beispiele zeigen, wie sich die Proteine im Laufe der Evolution ausgehend von einem Ur-Struktur-Gen durch Duplikationen, Punktmutationen und Hybridisierung zu Molekülen unterschiedlicher Größe, Sequenz und Funktion entwickelt haben können. Neue Proteine sind aber möglicherweise auch dadurch entstanden, daß sich bereits wohl differenzierte Proteine mit anderen Polypeptid-Ketten vereinigten. Eine solche Fusion kann durch Vereinigung benachbarter Struktur-Gene resultieren, kommt aber offenbar auch ohne Fusion der Gene zustande; in diesem Fall wird jeder Teil der Sequenz durch ein autonomes Struktur-Gen bestimmt.

### 5.1. Bildung bifunktionaler Enzyme durch Fusion benachbarter Gene

Die Biosynthese des Histidins wird bei *Salmonella typhimurium* durch ein Operon kontrolliert, in dem neun benachbarte Struktur-Gene diejenigen Enzyme codieren, die an der Synthese der Aminosäure beteiligt sind. Das zweite Struktur-Gen des Operons codiert die Histidinol-Dehydrogenase, das dritte die Imidazolacetolphosphat:L-Glutamat-Aminotransferase. Zwischen dem Endsignal für das eine Enzym und dem Startsignal für das andere befindet sich eine Sequenz von intercistronischen Nucleotiden, die normalerweise nicht abgelesen wird. Bei einer doppelten Mutation wurden nun aber die Unterbrechungssignale nicht erkannt, und es entstand ein einziges Gen aus den Sequenzen der beiden ursprünglichen Gene und aus der Sequenz der intercistronischen Nucleotide; dieses neue Gen ergab bei der Transkription ein Protein mit einem Molekular-Gewicht von 140000. (Die Molekulargewichte der Dehydrogenase und der Aminotransferase betragen 59000 bzw. 84000<sup>[35, 36]</sup>.) Das neue Protein hat zwei katalytische Wirkungen und ist dimer wie die ursprünglichen Enzyme; dies legt die Vermutung nahe, daß sich die Konformationen der beiden Proteine trotz ihrer kovalenten Verknüpfung erhalten haben. Man könnte sich vorstellen, daß der verbindende Polypeptid-Teil, der der Sequenz der intercistronischen Nucleotide entspricht, besonders flexibel ist.

Möglicherweise sind einige große Enzyme, z. B.  $\beta$ -Galaktosidase und DNA-Polymerase, durch Fusionen dieser Art entstanden. Bei einer proteolytischen Spaltung müßte die DNA-Polymerase zwei Fragmente ergeben, von denen jedes eine der Wirkungen des ursprünglichen Enzyms hat. Durch Fusion

benachbarter Gene läßt sich die Entstehung komplizierter Proteine im Verlauf der Evolution offenbar plausibel erklären.

## 5.2. Fusion von Proteinen ohne Gen-Fusion

Die Immunglobuline (s. Abschnitt 4.3.2), die bei den Immunreaktionen eine wesentliche Rolle spielen, sind symmetrische Moleküle, die aus je einem Paar schwerer und leichter Ketten bestehen. Beim Menschen haben die Immunglobuline G (IgG), A (IgA), M (IgM), D (IgD) und E (IgE) eine schwere Kette des  $\gamma$ -,  $\alpha$ -,  $\mu$ -,  $\delta$ - bzw.  $\epsilon$ -Typs. Diese schweren Ketten sind mit leichten Ketten der Typen  $\kappa$  und  $\lambda$  verbunden. Bei Mäusen gehören 97% der leichten Ketten dem  $\kappa$ -Typ an, beim Menschen liegen 66%  $\kappa$ -Typ und 34%  $\lambda$ -Typ vor<sup>[6, 37]</sup>. Die leichten Ketten (220 Reste) bestehen aus dem variablen (V) N-terminalen und dem konstanten (C) C-terminalen Teil. Die bisher untersuchten Ketten vom  $\kappa$ -Typ verhielten sich, als ob die konstante Region ( $C_k$ ) durch ein einziges Gen codiert wird; dagegen könnte die  $V_k$ -Region drei Genen entsprechen, da sich bei ihr drei Unterarten unterscheiden lassen. Die Aminosäure-Sequenz einer bestimmten leichten Kette würde demnach durch zwei getrennte Gene, d. h. durch ein Gen für jeden Teil, kontrolliert werden. Ähnlich ist bei den schweren Ketten das N-terminale Viertel als variable Region und der Rest als konstante Region anzusehen; diese beiden Regionen müßten in ähnlicher Weise getrennten Genen entsprechen.

## 6. Polymere Proteine und fixierte Proteine

Bei allen bisher beschriebenen Beispielen wurden Evolutionsmöglichkeiten für einzelne Polypeptid-Ketten besprochen, bei denen man annimmt, daß sie ihre Funktion unabhängig ausüben. Neben solchen Molekülen sind aber für die Zell-Physiologie auch kompliziertere Protein-Systeme notwendig.

Die Enzyme für die unerläßliche Regulierung der Stoffwechsel-Ketten<sup>[38]</sup> bestehen stets aus mehreren Untereinheiten, die durch nicht-kovalente Bindungen verknüpft sind („polymere Proteine“). Diese sogenannten „allosterischen“ Regulierungsenzyme sind im allgemeinen gegenüber dem Endprodukt einer Enzym-Kette, welches *kein* Struktur-analogon ihres Substrates ist, empfindlich. Die Zahl der biologischen Funktionen, bei denen sich die Einwirkung von derartigen Systemen nachweisen läßt, nimmt ständig zu; Beispiele sind der Sauerstoff-Transport (Hämoglobin), die Muskel-Kontraktion (Myosin, Actin) und die Wirkung spezieller Hormone (Thyreotropin).

Die Moleküle scheinen nach einem allgemeinen Schema aufgebaut zu sein, wobei die Zahl der Untereinheiten variabel ist, jedoch sehr groß sein kann. Außer bei Hämoglobin<sup>[18]</sup> ist bisher nur sehr wenig darüber bekannt, wie die Untereinheiten miteinander verbunden sind. Auf jeden Fall ist die Anpassung der Konformationen der Untereinheiten, welche die Entstehung eines polymeren Moleküls ermöglicht, ein evolutionäres Problem, wenn man davon ausgeht, daß das assoziierte Molekül aus den freien monomeren Proteinen (z. B. Hämoglobin aus den  $\alpha$ - und  $\beta$ -Ketten) gebildet wird. Vom Standpunkt der natürlichen Auslese werden dem Molekül Bedingungen bezüglich der Beschaffenheit der Kontaktstellen auferlegt; der Übergang von einem „monomeren“ zu einem „polymeren“ Protein macht ein neues Element, das der strukturellen Koordination, erforderlich.

Die Zelle als Ganzes kann als ein mikroheterogenes System angesehen werden, in welchem die Oberfläche stärkere Wirkungen ausübt als ihrem statistischen Gewicht entspricht; im Extremfall gehen die Wirkungen ausschließlich von der Oberfläche aus<sup>[39]</sup>. Proteine und besonders Enzyme, die z. B. in den Membran-Strukturen fixiert vorliegen, sind damit von der Matrix ebenso abhängig wie von ihren eigenen biologischen Funktionen. Wegen der Vielfalt der beteiligten Moleküle (Lipide, Saccharide usw.) ist anzunehmen, daß die „Mikroumgebung“ eine nachhaltige Wirkung auf die Konformation der fixierten Proteine hat. Auf dem Niveau derart komplizierter Strukturen ist es jedoch schwer, den Einfluß der natürlichen Auslese festzustellen.

## 7. Schluß

Für den Nachweis einer Verwandtschaft von Protein-Molekülen erscheint es zweckmäßig, auch die Zell-Differenzierung und die Molekül-Differenzierung miteinander zu vergleichen. Beispielsweise entwickelten sich die Zellen sowohl des Hypophysen-Vorderlappens als auch des Hypophysen-Mittellappens aus dem Ectoderm; es ist daher interessant, daß  $\alpha$ -MSH (13 Reste), das von den Zellen des Mittellappens synthetisiert wird, dem N-terminalen Teil des Corticotropins entspricht, welches im Vorderlappen entsteht. Das von den  $\alpha$ -Zellen der Langerhansschen Inseln im Pankreas erzeugte, hyperglykämisch wirkende Hormon Glucagon (29 Reste) und das von der Duodenal-Schleimhaut produzierte Hormon Sekretin (27 Reste), das die Abscheidung des Pankreas-Saftes anregt, enthalten 15 Aminosäuren gemeinsam und haben 7 konservativ substituierte Reste. Die Ähnlichkeit der beiden Proteine beruht möglicherweise darauf, daß sich der Pankreas aus dem Duodenum entwickelte. Eine gewisse Ähnlichkeit fand sich auch bei den Proteinen, die von den aus dem Endoderm hervorgegangenen Drüsen sezerniert werden<sup>[40]</sup>. Unter dem Begriff „innersekretorische Proteine“ wurden so verschiedene Proteine wie die proteolytischen Enzyme des Pankreas, die in der Leber entstehenden Blutgerinnungsfaktoren sowie die Pankreas- und Magen-Hormone zusammengefaßt.

Obwohl sich heute postulieren läßt, daß die Evolution der Proteine über spezifische chemische Vorgänge verlaufen ist, sollte man nicht vergessen, daß sie nur ein Aspekt im Gesamtkomplex der Evolution der Organismen ist.

Eingegangen am 1. März 1973 [A 988]  
Übersetzt von Dipl.-Chem. Johanna Förster, Ludwigshafen

[1] F. Crick, Nature 227, 561 (1970).

[2] J. D. Watson: Molecular Biology of the Gene. Benjamin, New York 1965.

[3] V. M. Ingram: The Hemoglobin in Genetics and Evolution. Columbia University Press, New York 1963.

[4] A. S. Romer: Vertebrates Palaeontology. University of Chicago Press, 1966.

[5] R. Acher, Angew. Chem. 78, 856 (1966); Angew. Chem. internat. Edit. 5, 798 (1966).

[6] M. O. Dayhoff: Atlas of Protein Sequence and Structure. Vol. 5. National Biomedical Research Foundation, Washington 1972.

[7] E. Mayr: Principles of Systematic Zoology. McGraw-Hill, New York 1969.

[8] R. E. Dickerson, T. Takano, D. Eisenberg, O. B. Kallai, L. Samson, A. Cooper u. E. Margoliash, J. Biol. Chem. 246, 1511 (1971).

[9] E. L. Smith: The Evolution of Proteins. The Harvey Lectures, Series 62. Academic Press, New York 1968, S. 231.

[10] J. C. Kendrew, Science 139, 1259 (1963).

[11] A. E. R. Herrera u. H. Lehmann, Nature New Biol. 232, 149 (1971).

[12] M. Dautrevaux u. Y. Boulanger, Bull. Soc. Chim. Biol. 49, 949 (1967).

[12a] R. E. Dickerson in H. Neurath: The Proteins, 2. Aufl., Vol. II. Academic Press, New York 1964, S. 603ff.

[13] G. M. Air, E. O. P. Thompson, B. J. Richardson u. G. B. Sharman, Nature 229, 391 (1971).

[14] G. Matsuda, H. Takei, K. C. Wu u. T. Shiozawa, Int. J. Protein Res. 3, 173 (1971).

[15] K. Hilde u. G. Braunitzer, Hoppe-Seylers Z. Physiol. Chem. 349, 433 (1968).

[16] J. P. Chauvet u. R. Acher, Biochemistry 11, 916 (1972).

[17] Übersicht: G. Buse, Angew. Chem. 83, 735 (1971); Angew. Chem. internat. Edit. 10, 663 (1971).

[18] M. F. Perutz, H. Muirhead, J. M. Cox u. L. G. G. Goaman, Nature 219, 131 (1968).

[19] M. F. Perutz u. H. Lehmann, Nature 219, 902 (1968).

[20] B. S. Hartley u. D. L. Kauffman, Biochem. J. 101, 229 (1966).

[21] L. B. Smillie, A. Furka, N. Nagabhusan, K. J. Stevenson u. C. O. Parkes, Nature 218, 343 (1968).

[22] K. A. Walsh u. H. Neurath, Proc. Nat. Acad. Sci. USA 52, 884 (1964).

[23] D. M. Shotton u. B. S. Hartley, Nature 225, 802 (1970).

[24] D. M. Blow, Biochem. J. 112, 261 (1969).

[25] R. M. Stroud, L. M. Kay u. R. E. Dickerson, Cold Spring Harbor Symp. Quant. Biol. 36, 29 (1971).

[26] D. M. Shotton u. H. C. Watson, Nature 225, 811 (1970).

[27] H. Neurath u. R. A. Bradshaw, Accounts Chem. Res. 3, 249 (1970).

[28] W. M. Fitch u. E. Margoliash, Science 155, 279 (1967).

[29] V. M. Ingram u. M. Grunberg-Manago: Biosynthèse des Macromolécules. Ediscience, Paris 1970.

[30] J. B. Clegg, D. J. Weatherall u. P. F. Milner, Nature 234, 337 (1971).

[31] G. Flatz, J. L. Kinderlerer, J. V. Kilmartin u. H. Lehmann, Lancet I, 732 (1971).

[32] P. P. Slonimski, R. Acher, G. Pere, A. Sels u. M. Somlo, Colloq. Int. Cent. Nat. Rech. Sci. (1965).

[33] A. E. R. Herrera u. H. Lehmann, Nature New Biol. 232, 149 (1971).

[34] R. Acher, J. Chauvet u. M. T. Chauvet, Eur. J. Biochem. 29, 12 (1972).

[34a] D. Shotton in H. Fritz u. H. Tschesche: Proteinase Inhibitors. de Gruyter, Berlin 1971, S. 47ff.

[35] J. Yourno, T. Kohno u. J. R. Roth, Nature 228, 820 (1970).

[36] M. M. Rechler u. C. B. Bruni, J. Biol. Chem. 246, 1806 (1972).

[37] H. M. Grey in E. Schoffeniels, Biochemical Evolution and the Origin of Life. North-Holland Publishing Company, Amsterdam 1971.

[38] G. Cohen: Le métabolisme cellulaire et sa régulation. Hermann, Paris 1967.

[39] E. Katchalski, I. Silman u. R. Goldman, Advan. Enzymol. 34, 445 (1971).

[40] J. W. Adelson, Nature 229, 321 (1971).