Reprinted from:
A. Plückthun, Evolution im Reagenzglas.
In: Darwins langer Arm - Evolutionstheorie heute,
Heinz-Ulrich Reyer, Paul Schmid-Hempel (Hrsg.)
vdf Hochschulverlag AG an der ETH Zürich, 2011, pp. 89-105

Andreas Plückthun

Evolution im Reagenzglas

Charles Darwins Werk hat Generationen von Evolutionsbiologen inspiriert, die Welt um uns zu erklären. Auf den ersten Blick unerwartet, und ohne dass ihm selbst das damals bewusst gewesen sein kann, hat er auch ein intellektuelles Gerüst geliefert, mit dem der Mensch Biomoleküle optimieren kann. In den letzten Jahren ist diese evolutionäre Strategie zu einer wichtigen Quelle von biopharmazeutischen Wirkstoffen geworden. Aber erst kürzlich konnte man Technologien entwickeln, die effizient genug sind, um das evolutionäre Potenzial optimal zu nutzen, und die Evolution im Reagenzglas ermöglichen. Wir werden deshalb ohne Zweifel eine weitere rasante Beschleunigung dieses Fortschritts sehen.

Die Grundlage für diese Technologien bilden Darwins sorgfältige Beobachtungen der Natur, insbesondere auf seiner fünf Jahre dauernden Reise mit der HMS *Beagle*. Sie führten ihn zu einer umfassenden Theorie über die Vielfalt der Lebensformen. Sein Hauptwerk von 1859 *On the origin of species by means of natural selection* erklärt die Aufspaltung der Organismen in verschiedene Arten und ihre Entwicklung als Folge von Variation und natürlicher Selektion (Abb. 1). Der entscheidende Punkt – nicht immer genügend betont – der evolutionären Strategie in der Natur ist die *immer wiederkehrende* Abfolge von Veränderung (Diversifizierung) und Auswahl (Selektion) der geeignetsten Varianten. Es ist ein *zyklischer* Prozess (Abb. 1d). Das heisst, die Evolution benutzt die Politik der kleinen Schritte, derer es allerdings sehr viele braucht. Jegliches Verfahren der Evolution im Labor muss deshalb in der Lage sein, die wiederkehrende Abfolge von Diversifizierung und Selektion in die Praxis umzusetzen – wenigstens ein paar Male hintereinander.

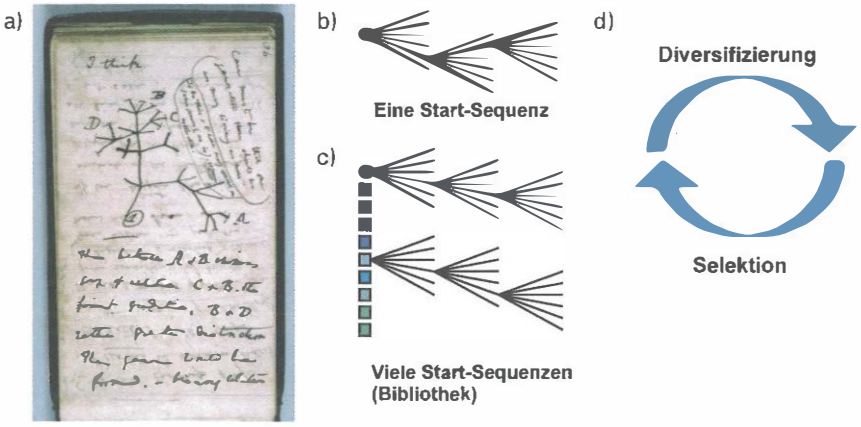

Abb. 1: Evolution als zyklische Abfolge von Diversifizierung und Selektion. (a) Skizze aus Darwins Notizbuch «B» von 1837. (b) Schematische Darstellung einer molekularen Evolution, die von einem einzigen Startmolekül beginnt. (c) Schematische Darstellung einer molekularen Evolution, die von einer grossen vorgefertigten Vielzahl von Molekülen, einer Bibliothek, beginnt. (d) Evolution als zyklische Wiederholung von Diversifizierung und Selektion

Zu Darwins Zeiten wäre eine Diskussion dieser Gedanken ausgeschlossen gewesen, denn es fehlte jegliches Wissen über die molekularen Grundlagen des Lebens. Darwin untersuchte *Merkmale (Phänotypen)* von Pflanzen und Tieren und zog aus seinen Beobachtungen den Schluss, dass Arten – jede mit ihren eigenen Merkmalen – durch Diversifizierung und Selektion entstanden sind. Ein Beispiel ist seine legendäre Analyse der Finkenschnäbel, deren Form er als Anpassung an die Nahrung deutete.

Heute wissen wir, dass solche Merkmale in der Erbsubstanz, der DNA, codiert sind. Zufällige Veränderungen (Mutationen) im Erbgut – die meisten ohne jede Konsequenz, einige durchaus schädlich und nur ganz wenige wirklich nützlich – führen zu unterschiedlichen Merkmalen, von denen sich jene durchsetzen, die ihren Trägern einen Überlebens- und Fortpflanzungsvorteil verschaffen.

Vom Merkmal zum Molekül

Es ist nun erforderlich, den Begriff ‹Merkmal› genauer anzuschauen und auf eine biochemische Grundlage zu stellen. Die Erbsubstanz DNA lässt sich in Abschnitte unterteilen, die Gene. Diese wiederum werden in eine Arbeitskopie (die mRNA) umgeschrieben, mit dem Vorteil, dass jedes der Gene unab-

hängig genau in der Menge, an dem Ort und zu der Zeit «angeschaltet» werden kann, wie sein Produkt benötigt wird. Die Bausteine der mRNA werden übersetzt in Aminosäuren, welche die Bausteine der Proteine (Eiweisse) darstellen. Die Abfolge (Sequenz) der mRNA-Bausteine wird dabei übersetzt in eine Abfolge (Sequenz) von Proteinbausteinen (Abb. 2).

Es sind die Proteine, welche die eigentliche Maschinerie der Zelle bilden (Abb. 2a). Bestimmte Proteine (die Enzyme) zerkleinern die Nahrung zu einzelnen Molekülen und bauen daraus wieder andere Moleküle auf, welche die Zelle benötigt; weitere geben der Zelle eine bestimmte Struktur (Strukturproteine), wieder andere sorgen für Bewegungen innerhalb der Zelle (Motorproteine) oder binden Botenstoffe (Rezeptoren). Der grösste Teil der Proteine jedoch ist an der Regulation anderer Proteine beteiligt, ein-

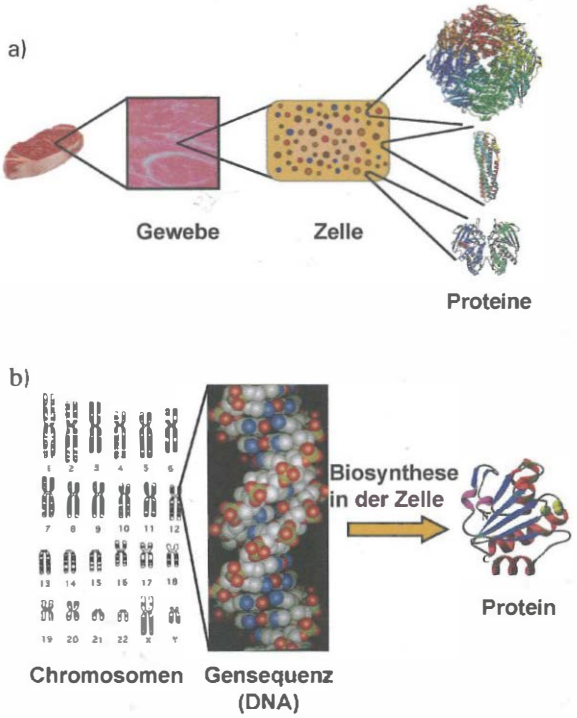

Abb. 2: Proteine – die Bausteine des Lebens. (a) Proteine üben praktisch alle Funktionen einer Zelle aus. Die Gewebe sind aus Zellen aufgebaut, welche die Proteine enthalten. Die Abfolge (Sequenz) der Aminosäuren legt sowohl Form als auch Funktion der Proteine fest sowie mit welchen anderen Faktoren sie interagieren. (b) Die Abfolge der Bausteine der Aminosäuren wird durch die Abfolge der Bausteine in den Genen (DNA) bestimmt. Die ca. 23'000 Gene des Menschen sind in 23 Chromosomen angeordnet.

schliesslich der Steuerung, wie viel von einem bestimmten Protein hergestellt werden soll. Auch eine Zelle besteht also grösstenteils aus Verwaltung mit nur wenig Produktion!

Der Mensch besitzt ca. 23'000 Gene, aus denen aber, durch Herstellung von Varianten, mehr als 100'000 verschiedene mRNA-Arbeitskopien entstehen, aus welchen wiederum vermutlich 500'000 bis 1'000'000 verschiedene Proteine gebildet werden können. Jede der verschiedenen Arten von Körperzellen (z.B. Muskelzellen, Nervenzellen oder Hautzellen) stellt ca. 100'000 verschiedene Proteine her, die jeder Zelle ihre charakteristischen Eigenschaften geben. Es ist offensichtlich, dass sich ein unvorstellbar komplexes Netz- und Regelwerk ergibt. Es wäre eine anmassende Übertreibung, zu behaupten, wir verstünden davon auch nur einen kleinen Teil.

Die Abfolge der Proteinbausteine in der Kette (also die Aminosäuresequenz) bestimmt, ob dieses Protein ein Enzym, ein Motorprotein, ein Rezeptor oder ein Regulator ist. Die Sequenz der Aminosäuren legt nämlich die dreidimensionale Form der Proteinkette fest (Abb. 2) und damit ihre Fähigkeiten. Kann sie z.B. eine chemische Reaktion steuern oder stattdessen an ein anderes Protein binden und dieses als Hemmstoff (Inhibitor) an seiner Tätigkeit hindern? Ein typisches Protein hat eine Kettenlänge von 100 bis 300 Aminosäuren, von denen es 20 verschiedene gibt. Es ist sofort ersichtlich, dass die Zahl aller möglichen Anordnungen bereits bei einem Protein mit 100 Aminosäuren ($20^{100} = 10^{130}$) die Zahl der Atome im gesamten Universum (ca. 10^{80}) um 50 Zehnerpotenzen übersteigt. Es ist deshalb keine grosse Übertreibung, die Zahl prinzipiell möglicher Proteine als «unendlich» zu bezeichnen.

Welche der 20 verschiedenen Aminosäuren an welcher Stelle im Protein sitzt, hat einen grossen Einfluss auf seine Funktion. Es ist leicht vorstellbar, dass der Austausch einer Aminosäure in vielen Fällen die Funktion zerstören wird. Ich werde aber nun die Frage diskutieren, wie ein richtiger Austausch an der richtigen Stelle die Funktion *verbessern* kann, indem er das Protein an eine neue Aufgabe anpasst.

Ein komplexes Merkmal: Alles nur eine Frage der richtigen Aminosäuren

Als Beispiel soll ein Merkmal dienen, das im Verlauf der Evolution mehrfach «erfunden» wurde und bei dem die Verknüpfung zwischen dem Vorteil für

ein Tier und den molekularen Hintergründen gut ersichtlich ist: das Farbsehen. Die Fähigkeit, Farben zu unterscheiden, liefert Bilder, die wesentlich reicher an Informationen sind, um Objekte zu erkennen und zu bewerten, als es ein monochromatisches Sehen in «Schwarz-Weiss» vermag.

Der Sehvorgang in der Retina (Netzhaut) des Auges (Abb. 3) beruht auf der Absorption (Aufnahme) von Photonen (Lichtquanten) mithilfe eines Farbstoffes, des Retinals, das in einem Rezeptor, dem Opsin, eingebettet ist. Ganz allgemein sind Rezeptoren Proteine, oft auf der Oberfläche von Zellen, welche Reize und Botenstoffe erkennen und z.B. in Ner-

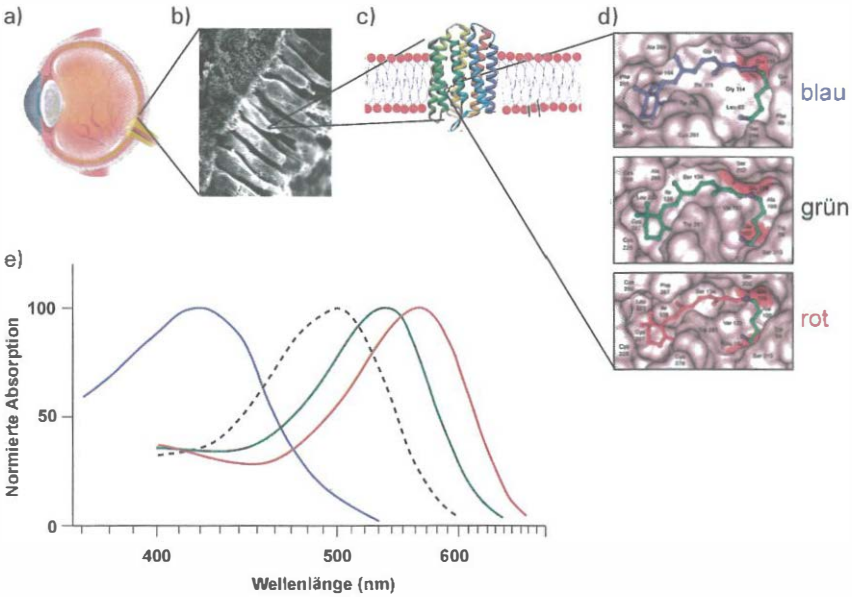

Abb. 3: Evolution des Farbsehens als Optimierung eines Proteins. (a) *In der Retina (Netzhaut) des Auges sind die Zapfen angeordnet, die das Farbsehen vermitteln.* (b) *Elektronenmikroskopische Aufnahme der Retina; ein kleinerer Zapfen (für das Farbsehen) ist hervorgehoben vor den grösseren Stäbchen (für das Sehen in der Dämmerung).* (c) *Opsin aus den Zapfen ist das Rezeptorprotein, welches das Farbsehen vermittelt, angeordnet in der Zellmembran. Es enthält Retinal, den lichtabsorbierenden Farbstoff. Wenn ein Photon auf diesen Farbstoff triff, ändert er seine Form.* (d) *Die Evolution hat drei verschiedene Opsine hervorgebracht, die den Farbstoff in leicht unterschiedlicher Umgebung binden. Auf diese Art kann Licht unterschiedlicher Wellenlänge wahrgenommen werden (blau, grün, rot), weil unterschiedlich viel Energie für die Formänderung benötigt wird. Die entscheidenden Unterschiede bestehen in den Proteinbausteinen (Aminosäuren), die den Farbstoff umgeben (einige davon sind als rote Flecken auf grauer Oberfläche hervorgehoben).* (e) *Durch Vergleich der Signale von den drei Rezeptoren kann das Gehirn die Wellenlänge jedes Photons bestimmen und somit die Farbe erkennen. (Die gestrichelte Linie gibt die Absorption der Stäbchen an, die das «Schwarz-Weiss»-Sehen bei schwachem Licht erlauben.)*

venimpulse umwandeln. Im spezifischen Fall des Sehvorganges ändert das Retinal nach Absorption des Photons seine Form, was vom Opsin-Rezeptor weitergeleitet und schliesslich als Nervenimpuls wahrgenommen wird. Die Wellenlänge (also die «Farbe»), bei der der Rezeptor maximal anspricht, hängt von der Umgebung des Retinals ab (Abb. 3d). Genauer gesagt, die Aminosäuren in der direkten Nachbarschaft bestimmen, wie viel Energie die Lichtquanten besitzen müssen, um das Retinal anzuregen. Daraus folgt, dass ein Austausch dieser Aminosäuren einen Rezeptor schaffen kann, der bei einer anderen Wellenlänge optimal arbeitet (Teller et al. 2003). Im Verlauf der Evolution wurde das Gen für Opsin verdoppelt und anschliessend wurde eine der Kopien noch einmal verdoppelt. Jede der so entstandenen drei Kopien wurde durch Austausch der Aminosäuren auf eine andere Wellenlänge optimiert. So hat der Mensch drei verschiedene Arten von Zapfen, die je einen der drei verschiedenen Rezeptoren herstellen. In Kombination decken diese drei Rezeptoren ein breites Spektrum an Lichtwellenlängen ab, und durch Vergleich ihrer relativen Anregung kann das Gehirn jedem Bildpunkt eine «Farbe» zuordnen (Abb. 3e).

In letzter Instanz musste von der Natur dazu das biochemische Problem gelöst werden, welche Aminosäuren um das Retinal angeordnet werden müssen, um eine Anregung bei einer bestimmten Wellenlänge zu erreichen. Wir müssen heute davon ausgehen, dass dies durch Versuch und Irrtum geschah – eben Darwins Prinzip der Diversifizierung und Selektion. Durch den Einbau von wenigen zufälligen «Fehlern» in das Erbmaterial (DNA) – gelegentlich eben auch in das Gen für Opsin –, ausgelöst durch Umwelteinflüsse und die spontane Fehlerrate bei der Vervielfältigung der DNA während der Zellteilung, wird genau diese Diversifizierung erreicht, die Voraussetzung für eine Evolution ist. Bei einem solch komplexen Prozess wie dem Sehvorgang wird die Selektion im Überlebensvorteil derjenigen Tiere liegen, deren Rezeptoren trotz «Fehlern» funktionsfähig bleiben und bei denen eine Variante die Unterscheidung einer zusätzlichen Farbe ermöglicht.

Wir kennen den genauen historischen Verlauf dieser Evolution natürlich nicht, aber können davon ausgehen, dass er recht langsam voranging. Sehr viele Generationen sind notwendig, um eine solche Anpassung zu erreichen.

Von der Beobachtung der Natur zur Imitation ihrer Prinzipien

Der biochemisch geschulte Betrachter eines Rezeptors wie des Opsins (Abb. 3c) wird mit Bewunderung die extrem genaue Feinabstimmung erkennen, mit der jeder Teil des Rezeptors für die Anregung durch die genaue Energie der entsprechenden Lichtquanten optimiert ist. Alles dies geschah durch die Wahl der richtigen Aminosäure an der richtigen Stelle in der Proteinkette, für jeden der drei verschiedenen Rezeptoren. Und dennoch müssen wir davon ausgehen, dass dies alles ein Produkt eines Zufallsprozesses ist, genauer, einer Suchstrategie, die durch eine endlose Aufeinanderfolge von zufälliger Diversifizierung und Selektion zu beschreiben ist (Abb. 1d). Die für den Biochemiker wichtigste Erkenntnis ist daher das ungeheure Potenzial, das in dieser Suchstrategie steckt, die wir Evolution nennen.

Die Anpassung ganzer Organismen an unsere Erwartungen und Erfordernisse, nämlich die Züchtung, ist eine der grossen Kulturleistungen des Menschen. Darwin hat auch dieses Prinzip der ‹künstlichen Selektion› bereits 1868 in seinem Werk *The variation of animals and plants under domestication* erkannt und diskutiert. In praktisch keinem Fall ist aber bisher genau bekannt, worin die molekularen Ursachen der gezüchteten Eigenschaften liegen. Und für die praktischen Anwendungen, z.B. in der Landwirtschaft, ist diese Kenntnis vielleicht auch nicht so wichtig.

Wäre es aber möglich, diese evolutionäre Suchstrategie in das Labor zu übernehmen? Könnte man in künstlicher Weise eine Aufeinanderfolge von Diversifizierung und Selektion erreichen? Und wäre es möglich, dies gezielt auf bestimmte Proteinmoleküle anzuwenden, also gleichsam eine «molekulare Züchtung» von Proteinmolekülen durchzuführen? Solche massgeschneiderten Moleküle, die durch eine gerichtete Evolution entstehen, könnten als pharmazeutische Wirkstoffe eingesetzt werden. In der «Optimierung» eines ganzen Organismus hat der Züchter keinerlei Einfluss, welches der Gene und Proteine sich verändern wird, um ein bestimmtes Merkmal zu erzeugen. In der Optimierung eines bestimmten Proteins für eine biomedizinische Anwendung hingegen muss natürlich genau dieses Protein durch Veränderung seiner Aminosäurensequenz angepasst werden.

Wir müssen also das Problem lösen, das Prinzip der Züchtung auf einzelne Moleküle zu übertragen. Ausserdem ist es erforderlich, dass wir die zyklische Aufeinanderfolge von Diversifizierung und Selektion erheblich schneller ablaufen lassen als in den historischen Dimensionen der Evolu-

tion der Arten. Schliesslich müssen wir noch exakt definieren, welche Eigenschaften wir eigentlich optimieren wollen und dazu passend eine molekulare Selektionsstrategie entwerfen.

Warum sollte man Proteine evolvieren?

Wenngleich über die letzten Jahre in einer Reihe von Labors solche Anstrengungen mit unterschiedlichsten Zielen unternommen wurden, werde ich mich hier auf die Optimierung von Proteinen konzentrieren, die eine therapeutische Anwendung haben können.

Unter den Proteinen mit therapeutischem Potenzial sind diejenigen am einfachsten zu verstehen, die ein anderes Molekül spezifisch binden. So wird beispielsweise das Zellwachstum durch körpereigene Proteine gesteuert, die sogenannten Wachstumsfaktoren. Diese lagern sich an entsprechende Rezeptoren auf der Zelloberfläche an und geben so der Zelle das Signal zur Vermehrung. Gerät dieses System ausser Kontrolle, können sich Zellen unkontrolliert vermehren und zur Entwicklung eines Tumors führen. In vielen Fällen besteht der Defekt einfach darin, dass die veränderte Zelle wesentlich *mehr* von diesem Rezeptor herstellt. Das kann durch Schädigung der DNA geschehen – z. B. durch die ständige Zufuhr von mutagenen Substanzen, wie sie z. B. im Zigarettenrauch vorhanden sind. Wenn dabei zufällig die Genregion betroffen ist, welche die Menge dieses Rezeptors kontrolliert, kann eine solche Mutation zu einer höheren Anzahl von Rezeptoren pro Zelle führen. Weiterhin sind in Tumoren Mutationen häufig anzutreffen, bei denen ein Rezeptor seine Regulation verliert, und somit *immer* aktiv ist – auch ohne Signal von aussen.

Solche Rezeptoren für Wachstumsfaktoren sind deshalb sehr interessant zur Entwicklung von neuen Tumortherapien. Es wird also ein Protein benötigt, das einen solchen Rezeptor spezifisch erkennt und an seiner Funktion hindert – inmitten der zahllosen anderen Proteine, die ein Organismus enthält. Ein solches Erkennungsprotein würde, wenn man es in die Blutbahn spritzt, an den Tumorzellen hängen bleiben, die diesen Rezeptor in grossen Mengen auf ihrer Oberfläche tragen. Natürlich ist das eine sehr vereinfachte Darstellung, aber eine extrem spezifische Erkennung des Zielmoleküls (also des Rezeptors) wird in jedem Falle die Voraussetzung zum Erfolg sein.

Wir haben somit das technische Problem zu lösen, wie man innerhalb einer vertretbaren Zeitspanne (wenige Wochen bis Monate) ein Protein

erzeugt, das spezifisch an ein anderes bindet. Trotz grosser Fortschritte bei der Aufklärung der dreidimensionalen Struktur von Proteinen (die erste wurde genau 100 Jahre nach Darwins *On the origin of species* bestimmt) ist es heute, noch einmal 50 Jahre später, immer noch nicht möglich, ein solches spezifisch bindendes Protein vollständig am Computer zu entwerfen. Die ungeheure Zahl von Möglichkeiten, wie Proteine aneinander binden können, die gegenseitige Anpassung ihrer Strukturen und die winzig kleinen Energieunterschiede all dieser Anordnungen überfordern die Genauigkeit der heutigen Rechenmethoden.

Wir sind daher auf eine experimentelle Lösung des Problems angewiesen. In der Tat ist es ein Optimierungsproblem, für das die evolutionäre Methodik die ideale Strategie darstellt. Die Zielvorgaben sind klar: Das gesuchte Protein muss den gewünschten Rezeptor erkennen, mit einer ausreichenden Energie binden und ihn an seiner Funktion hindern; es muss stabil sein und sollte weiterhin möglichst einen ähnlichen Rezeptor, der eine ganz andere Funktion hat, gerade *nicht* an seiner Funktion hindern. Nun gilt es, diese Anforderungen in einen experimentellen Plan umzusetzen.

Die Proteinbibliothek

In der natürlichen Evolution sind mit grosser Wahrscheinlichkeit die vielen Proteinvarianten der heutigen Lebensformen aus einer *geringen* Zahl von Urtypen hervorgegangen. Ich habe das oben am Beispiel des Opsins diskutiert, dessen Veränderung zum Wahrnehmen verschiedener Farben geführt hat. Unser Ziel muss es sein, diesen Prozess erheblich zu beschleunigen, d. h. den Weg der Evolution abzukürzen. Deshalb ist es von Vorteil, mit extrem vielen Ausgangsformen zu beginnen. Eine Billion (10^{12}) wäre ein guter Ansatz.

Am Anfang des Experiments steht demnach die Entwicklung einer «Bibliothek», also der Herstellung einer Sammlung von Kandidatenproteinen. Eine solche Bibliothek wäre idealerweise eine Mischung von Proteinen, welche die Natur aus einem Urprotein hervorgebracht haben *könnte* (Abb. 1c), die alle eine ähnliche dreidimensionale Struktur haben und bei denen jedes Mitglied der Mischung ein Zielmolekül erkennen *könnte* (wir wissen aber nicht, welches).

Diese Proteine sollten allesamt Eigenschaften aufweisen, die ihren späteren Einsatz in der biomedizinischen Anwendung erlauben, also effiziente

und kostengünstige Herstellung, hohe Stabilität sowie leichte Modifizierbarkeit, um an das Protein weitere therapeutische Substanzen zu koppeln. Insbesondere sollten diese Proteine aber in der Lage sein, mit hoher Wahrscheinlichkeit vom Experimentator vorbestimmte Zielmoleküle (z. B. Rezeptoren) spezifisch zu erkennen. Es ist heute noch nicht möglich, eine solche Proteinbibliothek zu entwerfen, ohne sich von der Natur inspirieren zu lassen. Man sucht deshalb in den Datenbanken nach natürlich vorkommenden Proteinklassen, die in der Lage sind, andere Proteine spezifisch zu erkennen.

Historisch gesehen waren die Antikörper die erste Proteinklasse, mit der solche Bibliotheken erzeugt wurden (Plückthun und Moroney 2005). Alle Antikörper zusammen bilden eine Proteinbibliothek, die der Körper selbst herstellt, um daraus jeweils diejenigen auszuwählen, die zur Bekämpfung einer durch ein Bakterium oder ein Virus verursachten Infektion am besten geeignet sind. Leider sind synthetische Antikörpermoleküle im Labor nicht immer problemlos herzustellen.

In jüngster Zeit wurden in meinem Labor deshalb neben synthetischen Antikörpern auch Bibliotheken ganz anderer Proteine hergestellt (Binz et al. 2003, 2004) – mit dem Ziel, stabile Proteine zu erzeugen, mit denen sich biotechnologisch gut arbeiten lässt. Als Ausgangspunkt fand eine Klasse von Proteinen Verwendung, die von der Natur vielfältig für die Aufgabe eingesetzt wird, andere Proteine spezifisch zu binden: die Ankyrine (Abb. 4).

Um eine solche Proteinbibliothek zu entwerfen, macht man sich wiederum die Evolution zunutze. Nicht nur hat jedes Lebewesen die Gene für mehrere Hundert verschiedener solcher Ankyrin-Moleküle in seinem Erbgut, sondern die verschiedenen Arten der Lebewesen haben alle auch leicht veränderte Formen, mit etwas anderer Sequenz der Aminosäuren. Kennt man nun die Aminosäuresequenzen vieler solcher Mitglieder der Proteinklasse, so kann man eine ‹Konsensus-Sequenz› aufstellen, die gleichsam einen Durchschnittswert dieser Proteinklasse darstellt. Man stellt dabei fest, dass fast alle Mitglieder dieser Klasse an einigen Stellen der Sequenz identisch sind, an anderen stark voneinander abweichen. Die Stellen, an denen fast alle Mitglieder gleich sind, müssen also für die dreidimensionale Struktur von besonders grosser Bedeutung sein. Wenn man nun noch wenigstens von einigen Mitgliedern dieser Familie ihre dreidimensionale Struktur kennt und daher weiss, wo genau sie an ihr Zielmolekül binden können (Abb. 4), so weiss man, welche Teile des Proteins besonders tauglich für die Bindung eines Zielmoleküls sind.

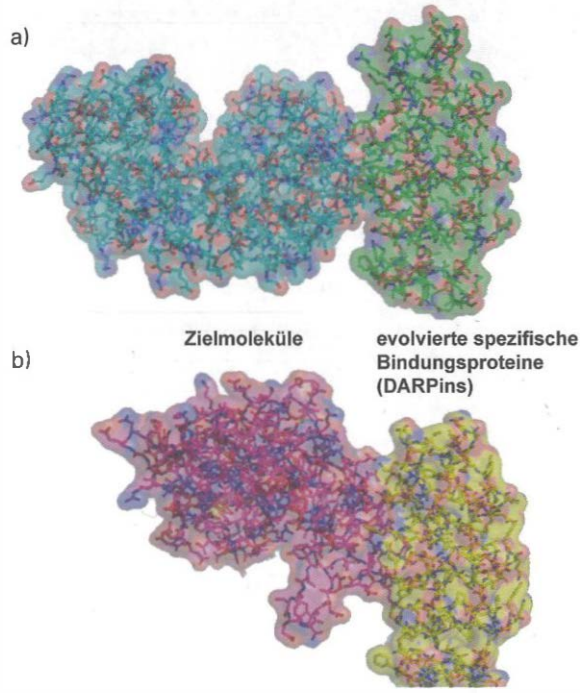

Abb. 4: Anpassung künstlicher evolvierter Proteinmoleküle an ihr Zielmolekül. Zwei evolvierte Proteinmoleküle (DARPins (für ‹Designed Ankyrin Repeat Proteins›), für Details, siehe Text) (jeweils rechts: **a**, *grün,* **b**, *gelb) erkennen spezifisch ausschliesslich das Zielmolekül, auf das sie selektioniert wurden (jeweils links:* **a**, *blaugrün;* **b**, *magenta). Sie passen wie ein Schlüssel zum Schloss. Obwohl sie ähnlich aussehen, unterscheiden sie sich in den Proteinbausteinen (Aminosäuren), die den direkten Kontakt mit dem Zielmolekül eingehen.*

Damit haben wir nun die Information für den Entwurf einer Proteinbibliothek beisammen: Wir haben eine natürliche Proteinfamilie identifiziert, die aufgrund ihrer Eigenschaften für eine biomedizinische Anwendung besonders geeignet ist, in unserem Fall die Ankyrine. Wir können durch Vergleich aller natürlichen Ankyrin-Sequenzen bestimmen, welche Aminosäuren unbedingt erhalten bleiben müssen, damit die neuen Proteine ebenfalls die gewünschte ursprüngliche Struktur einnehmen. Wir können schliesslich aus der Kenntnis der dreidimensionalen Struktur natürlicher Ankyrine mit ihren Bindungspartnern entnehmen, an welchen Stellen des Ankyrins diese typischerweise binden. An diesen Bindungsstellen muss die Proteinbibliothek «zufällige» Aminosäuren enthalten; das heisst, jedes Mitglied der Proteinbibliothek muss einen anderen Satz von Aminosäuren enthalten. Wenn die Bib-

liothek der Proteine genügend gross ist, so die Überlegung, wird es für jedes Zielmolekül immer ein paar Mitglieder geben, bei denen die Oberfläche genau passt, sodass eine Bindung erfolgt (Abb. 4): Ein Komplex ist entstanden.

Wenngleich oben betont wurde, dass es heute noch nicht möglich ist, den endgültigen Komplex am Computer masszuschneidern, so spielt bei der Planung der *Proteinbibliothek* das Computerdesign eine wichtige Rolle: Durch Vergleich der Strukturen der natürlichen Proteine sieht man, welche Teile des Proteins für den Zusammenhalt der Struktur entscheidend sind, welche Teile keinesfalls variiert werden sollten und welche Teile des Proteins variiert werden könnten.

Von der Proteinbibliothek im Computer zur DNA-Bibliothek in der Realität

Das Problem ist nun, wie man das geschilderte Grundprinzip in die Praxis umsetzt und wie man Mitglieder der Bibliothek findet, die spezifisch binden – die sprichwörtliche Nadel im Heuhaufen. Wir werden sehen, wie sich die Lösung dieses Problems verknüpfen lässt mit der Evolution im Reagenzglas, also der zyklischen Abfolge von Diversifizierung und Selektion, so wie sie Darwin als Naturprinzip erkannt hat.

Die Proteinbibliothek muss als Bibliothek von DNA-Molekülen erzeugt werden. Das heisst, die gewünschten Sequenzen der Proteinbibliothek müssen «zurück übersetzt» werden in die DNA-Sequenzen, welche die entsprechende Proteinbibliothek erzeugen würde. Doch wie geschieht das in der Praxis (Abb. 5)? Die DNA wird völlig synthetisch hergestellt, aus kleinen Stücken (Oligonukleotiden), die an einer Maschine aus einzelnen Nukleotidbausteinen zusammengebaut werden. Die kleinen Stücke werden dann zum gesamten Gen verbunden. An den Stellen, an denen verschiedene Mitglieder der Bibliothek sich später in der Aminosäure unterscheiden sollen, werden die DNA-Bausteine gemischt. So wird bei der Übersetzung in Proteine ebenfalls eine Mischung entstehen, weil jedes einzelne DNA-Molekül zu einem anderen Protein führt. Da sich die Bausteine bei der maschinellen DNA-Synthese in beliebiger Weise mischen lassen, kann so letztlich eine Proteinbibliothek erzeugt werden, bei der an vorherbestimmten Stellen in der Proteinkette eine beliebige Mischung von Aminosäuren vorliegt. Dummerweise liegen danach aber sämtliche DNA-Stücke (eine Billion synthetischer Gene) in ein und demselben Reagenzglas vor und wir wissen nicht, welches Gen wir brauchen!

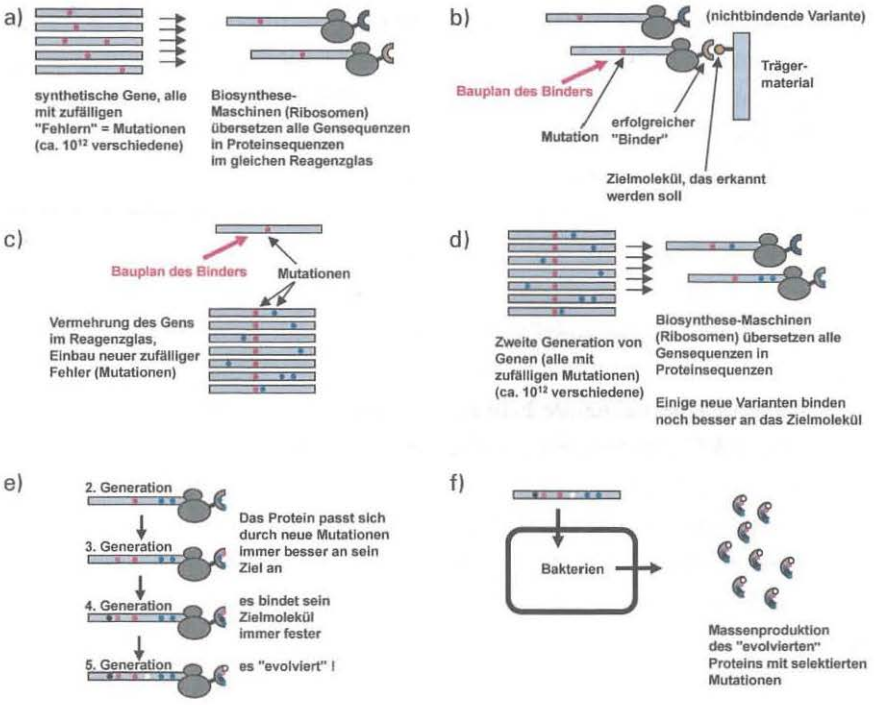

Abb. 5: Evolution im Reagenzglas (nach dem Verfahren des ‹Ribosome Display›). (a) Synthetische Gene werden erzeugt, die zufällige Fehler (Mutationen) enthalten (rote Punkte). Diese werden von einem Zellextrakt, der Ribosomen enthält (die Proteinsynthese-Maschine) in Proteine übersetzt. Das Entscheidende an dem Verfahren ist, dass sowohl die Gene als auch die Proteine an den Ribosomen verbleiben (für Details, siehe Kasten 1). (b) Das zu erkennende Zielmolekül (z. B. ein krankheitsrelevanter Rezeptor) wird an ein Trägermaterial gekoppelt und mit der Gesamtheit der Komplexe aus (a) gemischt. Es bleiben nur diejenigen hängen, die das Zielmolekül erkennen, inklusive ihres Bauplans (Gen). (c) Dieses Gen wird vervielfältigt und dabei werden wiederum neue zufällige Fehler (Mutationen) eingebaut (blaue Punkte). (d) Diese zweite Generation wird wie in (b) wieder in Protein übersetzt, sodass sie mit ihrem Bauplan verbunden bleiben. (e) Durch Wiederholung der Schritte werden sukzessive nützliche Mutationen angesammelt (farbige Punkte). (f) Das Gen der Endstufe wird in Bakterien eingeschleust, um das erzeugte Protein in grossen Mengen zu produzieren.

Die Nadel im Heuhaufen: Das Konzept

Die entscheidende Herausforderung ist nun, das Protein, welches an das gewünschte Zielmolekül bindet, und das dazugehörige Gen zu identifizieren. Dazu wurde in meinem Labor ein Verfahren entwickelt, das völlig ohne lebende Zellen auskommt (Hanes und Plückthun 1997). Ich werde hier nur

das Konzept beschreiben; die technische Durchführung ist in Kasten 1 grob skizziert und in Zahnd et al. (2007) genau beschrieben.

Der Bestand unserer Bibliothek umfasst mindestens 100 Milliarden (10^{11}) verschiedene Proteinmoleküle, meist sogar eine Billion (10^{12}). Jedes Proteinmolekül wird dabei nur ein einziges Mal vorkommen. Der entscheidende Trick ist, das Proteinmolekül mit seiner eigenen Erbinformation zu verbinden (Abb. 5). Wenn man nun die gesamte Billion von Molekülen mit einer genügenden Menge des Zielmoleküls in Kontakt bringt, so bleiben diejenigen Proteine am Zielmolekül hängen, welche mit diesem genau «zusammenpassen» (Abb. 5b). Da nun aber an jedem dieser Protein-Moleküle auch dessen eigene entsprechende Erbinformation hängt, haben wir zum Schluss am Zielmolekül eine Sammlung aller Erbinformationen von «erfolgreichen» Proteinen! Das heisst, wir haben die Baupläne all derjenigen Proteine isoliert, welche die Aufgabe erfüllen, an das Zielmolekül zu binden. Diese «Baupläne der Binder» sind die Sequenzen der Gene von erfolgreichen Proteinen. Sie lassen sich anschliessend in Bakterien einschleusen und die Bakterien stellen dann das erfolgreiche Protein her.

Was hat das nun mit Darwin zu tun? Leider ist die oben beschriebene Methode nicht perfekt. Es bleiben am Anfang nicht nur Moleküle hängen, die an das Zielmolekül binden, sondern auch solche, die an das Trägermaterial binden oder an eine sonstige Komponente. Um erfolgreich zu sein, müssen wir das Verfahren also mehrfach anwenden (Abb. 5c). Dazu müssen wir die Gene der bindenden Proteine isolieren und vervielfältigen und dann erneut die Proteine mit ihren Genen verbinden. Wenn man die Gene vervielfältigt, so kann man dort zufällige Fehler einbauen. Einige dieser Fehler werden, wie bei der Evolution in der Natur, keinen Einfluss haben, andere werden das Protein völlig unbrauchbar machen, aber einige wenige werden die Bindung des Proteins auch *verbessern*. Wenn es uns gelingt, die am besten bindenden Varianten mit der im Kasten am Ende dieses Kapitels skizzierten Technik zu isolieren, so erzeugen wir eine verbesserte Population. Wir evolvieren sie durch eine zyklische Abfolge von Diversifizierung und Selektion. Wir wenden das von Darwin beschriebene Prinzip im Reagenzglas an.

Und der praktische Nutzen?

Die skizzierte Strategie beschreibt einen Weg, wie man Proteine erzeugen kann, die prinzipiell an jedes beliebige Zielmolekül spezifisch binden

können. Damit ergeben sich vielfältige Möglichkeiten für Forschung und Anwendung (Plückthun 2009). So können in einer Zelle beispielsweise andere Proteine an ihrer Wirkung gehindert und so die Funktionen der Zelle getestet werden – wie reagiert sie, wenn diese eine Funktion ausfällt? Bis solche Anwendungen über die Forschung hinaus in die Therapie gelangen, sind aber noch einige andere Hürden zu überwinden, z. B. wie man solche Proteinmoleküle in die Zelle bringt. Auch das ist ein Gegenstand aktueller Forschung.

Dennoch ist beim Erscheinen dieses Buches das erste Proteinmolekül, ein erstes DARPin (für ‹Designed Ankyrin Repeat Protein› (vgl. Abb. 4), das wie oben beschrieben erzeugt wurde, bereits in der klinischen Erprobung im Menschen (www.molecularpartners.com). Ziel ist die Behandlung der Makuladegeneration, einer Sehstörung, die vor allem im Alter auftritt. Ursache ist ein Faktor, VEGF, der zum unkontrollierten Wachstum von Blutgefässen unter der Netzhaut führt, was die Sehschärfe erheblich beeinträchtigt. Indem nun ein Protein in das Auge gespritzt wird, welches diesen Wachstumsfaktor «abfängt», kann die Bildung dieser unerwünschten Blutgefässe gestoppt und sogar zurückgebildet werden – die Sehkraft nimmt wieder zu. Es ist offensichtlich, dass das DARPin so fest wie möglich an den Wachstumsfaktor binden muss, um zu verhindern, dass dieser an seinen Rezeptor andocken kann. Nur so kann das unerwünschte Wachstum der Blutgefässe unterbunden werden. Die perfekte Passform des Proteins konnte nur durch die Evolution im Reagenzglas erreicht werden (vgl. Abb. 4).

Schlussfolgerungen

In letzter Instanz konnte also das darwinsche Prinzip der Diversifizierung und der Selektion zur Optimierung eines therapeutisch wertvollen Proteins verwendet werden. Wenngleich vor 150 Jahren nicht einmal im Ansatz vorstellbar war, was heute Realität ist, so hat Darwin ein intellektuelles Gerüst geschaffen, mit dem der Ursprung der Arten, die Prinzipien der Züchtung und die Herstellung neuer Biomoleküle verstanden werden kann (Abb. 1a).

Wir fassen die Evolution heute als eine Suchstrategie auf, in der durch stetige Aufeinanderfolge von zufälliger Diversifizierung und Selektion eine optimale Anpassung erreicht werden kann. Wenngleich die Anfänge gemacht wurden, diese Prinzipien ins Labor zu übertragen, so sind dies immer noch bescheidene Schritte im Vergleich zu den Veränderungen, die bei der Entste-

hung des Lebens und der Arten abliefen. Der augenfälligste und wichtigste Unterschied ist die Anzahl der Zyklen von Diversifizierung und Selektion. Im Labor ist eine Handvoll Zyklen (Generationen) über ein paar Wochen bis Monate der Stand der Technik. In den ca. 3,5 Milliarden Jahren seit Entstehung des Lebens auf der Erde sind jedoch Billionen von Generationen entstanden. Es ist also keine allzu gewagte Prognose, dass jede Automatisierung, welche die Anzahl Zyklen im Labor zu erhöhen vermag, zu Molekülen führen könnte, an die wir heute noch nicht zu denken wagen. Aber auch sie werden Darwins Regeln folgen.

Die Nadel im Heuhaufen: Was wirklich passiert

Die Proteinbibliothek ist eigentlich zunächst eine Sammlung von Genen, die aus synthetischen Oligonukleotiden zusammengebaut wurden und sich, wie im Text beschrieben, an wenigen Positionen unterscheiden. Diese Gene werden zunächst mithilfe der Polymerase-Kettenreaktion (PCR) vervielfältigt. Im Reagenzglas werden sie dann von einem Enzym in die Arbeitskopie, die mRNA, umgeschrieben. Am Ende dieses Schritts haben wir 100 Milliarden bis 1 Billion *verschiedener* mRNA-Moleküle, und von jeder Sorte ein paar Tausend Exemplare. Nun geben wir einen Zellextrakt dazu, der Ribosomen enthält – jene Maschinerie, die mRNA in Proteine übersetzt. Die Ribosomen laufen an der mRNA entlang bis zum Ende und die dabei entstehende Proteinkette kommt aus dem Ribosom heraus, bleibt aber mit dem Ribosom verbunden. Erst durch ein Stop-Codon würde sie getrennt und – das ist der entscheidende Punkt – genau dieses fehlt in dieser mRNA, weil sie so entworfen wurde.

Die Proteinkette, deren grösster Teil sich nun ausserhalb des Ribosoms befindet, nimmt ihre dreidimensionale Struktur ein – sie faltet sich. So ist sie in der Lage, ein Zielmolekül zu erkennen, aber nur, wenn sie die richtige Sequenz dazu hat (Abb. 4). Nur wenige Mitglieder (ca. 1 aus einer Milliarde oder 1'000 aus einer Bibliothek von 1 Billion Mitgliedern) sind typischerweise dazu in der Lage.

Das Zielmolekül wird dazu an ein Trägermaterial gekoppelt, z.B. winzige magnetische Kugeln (Abb. 5b). So können die Proteine, die das Zielmolekül erkennen (samt mRNA und Ribosom) angereichert werden: Die riesige Mehrzahl der nichtbindenden Proteine wird weggewaschen und die bindenden Proteine werden mithilfe der magnetischen Kugeln, an

denen sie hängen, in ein neues Gefäss gebracht. Das ist die Selektion. Der Wissenschaftler kann an dieser Stelle genau bestimmen, welche Moleküle angereichert werden sollen und welche nicht. Er bestimmt den Selektionsdruck durch die genaue Versuchsanordnung.

Im neuen Reagenzglas werden die Komplexe aufgetrennt, die mRNA wird isoliert, mithilfe von Enzymen zurückübersetzt in DNA und mit der Polymerase-Kettenreaktion vervielfältigt. Danach kann eine neue Runde beginnen. Indem bei dieser Vervielfältigung mit Absicht Fehler eingebaut werden, wird Darwins Prinzip von Diversifizierung und Selektion umgesetzt (Abb. 5). Sucht man z. B. immer wieder nach den am besten bindenden Varianten, so erzeugt man durch diesen Prozess der Evolution eine Anpassung an das Zielmolekül mit immer besserer Bindungsenergie. Der Schlüssel zum Erfolg liegt also in der *zyklischen* Abfolge von Diversifizierung und Selektion (Abb. 1d).

Literatur

Binz HK, Amstutz P, Kohl A, Stumpp MT, Briand C, Forrer P, Grütter MG, Plückthun A (2004) High-affinity binders selected from designed ankyrin repeat protein libraries. Nature Biotechnology 22: 575–582

Binz HK, Stumpp MT, Forrer P, Amstutz P, Plückthun A (2003) Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. Journal of Molecular Biology 332: 489–503

Hanes J, Plückthun A (1997) *In vitro* selection and evolution of functional proteins by using ribosome display. Proceedings of the National Academy of Science of the USA 94: 4937–4942

Plückthun A (2009) Alternative scaffolds: expanding the options of antibodies. In: Little M (Hrsg.) Recombinant antibodies for immunotherapy. Cambridge University Press, Cambridge, S. 243–271

Plückthun A, Moroney SE (2005) Modern antibody technology: The impact on drug development. In: Knäblein J (Hrsg.) Modern Biopharmaceuticals Vol. 3. Wiley-VCH, Weinheim, S. 1147–1186

Teller DC, Stenkamp RE, Palczewski K (2003) Evolutionary analysis of rhodopsin and cone pigments: connecting the three-dimensional structure with spectral tuning and signal transfer. FEBS Lett 555: 151–159

Zahnd C, Amstutz P, Plückthun A (2007) Ribosome display: selecting and evolving proteins in vitro that specifically bind to a target. Nature Methods 4: 269–279