

# Screening- und Selektionsmethoden für die Analyse von Proteinfunktionen in großem Maßstab

Hening Lin und Virginia W. Cornish\*

Hochdurchsatz-Testsysteme erwecken gewaltige Erwartungen für das Protein-Engineering und die funktionelle Proteomik. Mit aussagekräftigen Tests sollte es möglich sein, beispielsweise eine stereoselektive Esterase für die chemische Synthese oder eine ortsspezifische Endonuklease für die biomedizinische Forschung zu entwickeln. Ganze cDNA-Bibliotheken, die alle

exprimierten Proteine eines bestimmten Organismus oder einer Zelllinie codieren, sollten einfach durch eine Batterie biochemischer Testsysteme geschleust werden, um die Funktion jedes einzelnen Proteins zu bestimmen. In diesem Aufsatz beschreiben wir, welche Arten von Testsystemen entwickelt wurden und wie nahe wir den Zielen sind, einerseits Proteine mit

neuen Aktivitäten zu entwickeln und andererseits den Tausenden von Proteinen, aus denen jedes Genom besteht, schnell eine Funktion zuzuordnen.

**Stichwörter:** Gerichtete Evolution • Hochdurchsatz-Screening • Protein-Engineering • Proteomik • Selektion

## 1. Einleitung

Hochdurchsatz-Testsysteme sind vielversprechende Verfahren für die Entwicklung von Proteinen mit neuen Eigenschaften. Wegen der Komplexität der Proteinfunktionen ist es sehr schwierig, Proteine mit neuen Funktionen herzustellen, geschweige denn neue Funktionen durch gerichtete Mutationen ins aktive Zentrum eines Proteins einzuführen. Daher versucht man inzwischen, Millionen von Mutanten gleichzeitig zu erzeugen und diese dann auf die gewünschte Aktivität zu testen. Die Fortschritte in der Molekularbiologie haben es viel einfacher gemacht, Proteinvarianten auf DNA-Ebene zu erzeugen. Da man es immer mit DNA zu tun hat, kann unabhängig von der Faltung oder der gewünschten Funktion des Proteins stets die gleiche Methode genutzt werden. Im Unterschied dazu müssen bei Bibliotheken niedermolekularer Substanzen für jede neue Grundstruktur neue Synthesewege entwickelt werden. Noch schwieriger allerdings ist es, Hochdurchsatz-Testsysteme zu entwickeln, mit denen Millionen von Proteinen innerhalb kurzer Zeit getestet werden können. In diesem Aufsatz konzentrieren wir uns daher auf die Fortschritte der letzten beiden Jahrzehnte bei der Entwicklung von Hochdurchsatz-Testsystemen für Proteinfunktionen.

[\*] Prof. Dr. V. W. Cornish, H. Lin  
Department of Chemistry, Columbia University  
New York, NY 10027 (USA)  
Fax: (+001) 212-932-1289  
E-mail: vc114@columbia.edu

Hochdurchsatz-Testsysteme für Proteinfunktionen werden außer beim Protein-Engineering bereits seit langem in der Enzymologie und der Wirkstoffsuche genutzt. In neuerer Zeit wurden sie auch zur entscheidenden Methode in der Proteomik. Die Fähigkeit, Millionen von Proteinvarianten zu erzeugen und ihre Aktivität zu vergleichen, ermöglicht uns nicht nur, Proteine mit neuen Eigenschaften zu entwerfen, sondern auch die Struktur und Funktion von Proteinen zu verstehen. Die pharmazeutische Industrie ist schon seit Jahrzehnten auf Hochdurchsatz-Testsysteme angewiesen, um niedermolekulare Hemmstoffe von Proteinen zu identifizieren. Seit dem Abschluss verschiedener Genomsequenzierungsprojekte in der jüngsten Vergangenheit ist es eine Herausforderung, die Funktionen der Proteine aufzuklären. Ein Genomsequenzierungsprojekt ist nur von begrenztem Nutzen, wenn wir die Proteinfunktionen nur einzeln aufklären können. Während einige Proteinfamilien anhand von Sequenzhomologien identifiziert werden können, werden Hochdurchsatz-Tests benötigt, um die Substratspezifität dieser Proteine einzugrenzen. Hochdurchsatz-Tests sind außerdem unverzichtbar, um die Aktivität des Großteils aller Proteine mit unbekannten Funktionen zu bestimmen.

Um Hochdurchsatz-Systeme für Proteinfunktionen zu entwickeln, müssen zwei Probleme überwunden werden: Erstens werden, weil DNA viel einfacher amplifiziert und sequenziert werden kann als Proteine, Methoden benötigt, um jedes Protein mit seiner spezifischen DNA-Sequenz zu verknüpfen (Abbildung 1). Ebenso wie dies zur Codierung jedes Elements einer Bibliothek niedermolekularer Substanzen genutzt

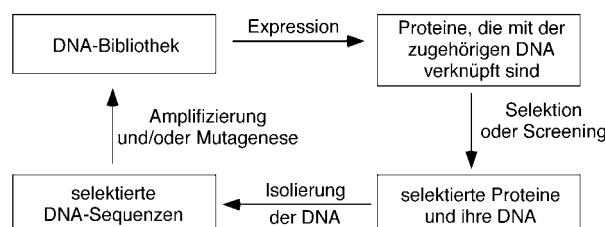


Abbildung 1. Allgemeine Strategie zur Analyse der Proteinfunktion in großem Maßstab.

werden kann, soll die DNA als Markierung für Proteinbibliotheken dienen. DNA ist ein ideales Etikett für die Proteinquenz, weil sie auch *in vivo* als Code für die Proteinquenz dient. Zweitens wird ein Hochdurchsatz-Testsystem gebraucht, das mit dieser Verbindung zwischen Protein und DNA kompatibel ist (Abbildung 1).

Die meisten Fortschritte wurden in den letzten beiden Jahrzehnten bei dem zuerst genannten Gebiet erzielt – der Entwicklung von Methoden, um jedes Protein mit seiner spezifischen DNA-Sequenz zu verknüpfen. Verschiedene raffinierte Lösungen stehen inzwischen zur Verfügung. Bei der Phagen-Display-Technik besteht buchstäblich eine physische Verbindung zwischen Protein und DNA. Bei zellulären Tests entsteht die Verbindung durch die Kompartimentierung in Einzelzellen, denn das Protein wird nur synthetisiert, wenn das Stück DNA, das dieses Protein codiert, in der Zelle

vorhanden ist. In Mikrotiterplatten-Tests haben Protein und DNA eine gemeinsame „räumliche Adresse“: Einfacher ausgedrückt: Das Protein in Spalte F, Reihe 11 wird durch die DNA in einer zweiten Platte in Spalte F, Reihe 11 codiert. Obwohl inzwischen verschiedene Methoden zur Markierung von Proteinbibliotheken mit ihrer DNA-Sequenz verfügbar sind, können auch vollkommen neue Ansätze von Vorteil sein. So kann es zum Beispiel interessant sein, über Methoden wie die Massenspektrometrie nachzudenken, durch die eine DNA-Markierung überflüssig werden könnte.

Die Verknüpfung eines Proteins mit seiner DNA-Sequenz ist allerdings nur der erste Schritt. Die eigentlichen Tests müssen noch entwickelt werden (Tabelle 1). In diesem Aufsatz werden mit „Selektion“ Methoden bezeichnet, mit denen nur Proteine mit der gewünschten Funktion untersucht werden, „Screening“ bezieht sich auf Methoden, bei denen alle Proteine untersucht werden müssen und „Testsysteme“ wird als Sammelbegriff für Selektionen und Screenings genutzt. Manche Methoden, wie die Ribosomen-Display-Technik, bieten sich besonders für Bindungstests an. Der Protein-DNA-Komplex wird mit einer Affinitätsmatrix zusammengebracht und Proteine, die an die Matrix binden, werden anschließend mit einer konzentrierten Salzlösung oder einer Ligandenlösung eluiert. Dieses Schema gilt unabhängig von der speziellen Protein-Ligand-Wechselwirkung. Enzymatische Umsetzungen stellen eine höhere Hürde dar. Bislang wurden meist Reaktionen untersucht, in denen durch

*Virginia W. Cornish wurde 1969 in Louisville, KY, geboren. Sie absolvierte ihr Grundstudium an der Columbia University. In dieser Zeit forschte sie im Labor von Professor Ronald Breslow an dem Rezeptor einer neuen Klasse von Antitumormitteln, die von der Arbeitsgruppe entwickelt wurde. Sie erwarb den B.A. der Columbia University in Biochemie im Jahr 1991 mit der Note „summa cum laude“. Dann wandte sie sich nach Westen und wechselte ins Forschungslabor von Professor Peter Schultz im Chemistry Department der University of California in Berkeley. Dort war sie an der Entwicklung einer neuen Methode beteiligt, mit der synthetische Aminosäuren mit der Synthesemaschinerie der Zelle in Proteine eingebaut werden können. In Berkeley wurde sie National Foundation Predoctoral Fellow und erhielt Stipendien der American Chemical Society Division of Organic Chemistry und des Howard Hughes Medical Institute. 1996 wurde sie NSF Postdoctoral Fellow am Biology Department am MIT unter Leitung von Professor Bob Sauer. Dort begann sie ein unabhängiges Projekt, das die Basis ihrer Laborarbeit in Columbia wurde. Virginia Cornish wurde 1999 Mitglied der Fakultät an der Columbia University. In ihrem Labor wird mit einer Kombination von synthetischer Chemie und Molekularbiologie gearbeitet, um ein zelluläres funktionsbasiertes Testsystem zum parallelen Screening großer Molekülsammlungen zu entwickeln. Sie erhielt den Beckman Young Investigator Award, einen Burroughs-Wellcome Fund New Investigator Award in Toxikologie, einen Camille and Henry Dreyfus New Faculty Award und einen National Science Foundation Career Award.*

*Hening Lin wurde 1975 in Tsingdao, China, geboren. Er absolvierte sein Grundstudium an der Tsinghua University in Beijing, China und schloss 1998 mit dem B.A. in Chemie ab. Er wechselte dann als einer ihrer ersten Studenten in die Arbeitsgruppe von Virginia W. Cornish an die Columbia University. Inzwischen ist er im fünften Jahr seines Graduiertenstudiums und arbeitet an der Entwicklung eines zellulären, funktionsbasierten Testsystems zum parallelen Screening großer Molekülsammlungen. Er ist der Arun Guthikonda Memorial Fellow des Jahres 2001.*



V. W. Cornish



H. Lin

Tabelle 1. Die Verknüpfung von Genotyp und Phänotyp in verschiedenen Testsystemen.

Methode	Verknüpfung zwischen Genotyp und Phänotyp	Testsystem (Selektion oder Screening)
1 Phagen-Display-Technik	Phagenpartikel	Bindung an eine Affinitätsmatrix
2 Ribosomen-Display-Technik	Ribosomenkomplex	Bindung an eine Affinitätsmatrix
3 mRNA-Peptid-Fusion	Peptid-mRNA-Fusion	Bindung an eine Affinitätsmatrix
4 Peptid am Plasmid	Peptid-Plasmid-Komplex	Bindung an eine Affinitätsmatrix
5 Zelloberflächen-Display-Technik	Zelle	Fluoreszenz-aktivierte Zellsortierung (FACS)
6 In-vitro-Kompartimentierung	Wasser-in-Öl-Tröpfchen	PCR
7 Genetisches Testsystem	Zelle	Komplementation, kolorimetrische Tests
8 Mikrotiterplatten und Proteinchips	Räumliche Adressierung	Radiometrisch, UV/Vis-Absorption oder Fluoreszenztest

die chemischen Eigenschaften der Reaktanten ein intrinsisches Messsignal entsteht – das Produkt fluoresziert, das Substrat ist ein Peptid oder das Produkt ist ein essentieller Metabolit. Nach den jüngsten Erfolgen wandten die Forscher ihre Aufmerksamkeit nun Testsystemen zu, die reaktionsunabhängig und mit einer weiten Palette chemischer Reaktionen vereinbar sind.

In diesem Aufsatz haben wir uns nicht auf eine bestimmte Methode konzentriert, sondern stattdessen versucht, die verschiedenen Methoden zu vergleichen, ihre Unterschiede herauszuarbeiten und zu untersuchen, welche Methode am besten für welches Problem geeignet ist. Für Hochdurchsatz-Systeme ist die Kernfrage natürlich die nach den erreichbaren Zahlen. Bei Untersuchungen der Proteinfunktion oder bei Tests aller Proteine einer vorgegebenen Zelllinie, wissen wir, wie groß die Zahlen sind. Beispielsweise codiert *Saccharomyces cerevisiae* etwa 6000 verschiedene Proteine und das menschliche Genom besteht aus etwa 30000 Proteinen. Schwieriger ist es abzuschätzen, wie viele Tests für die Entwicklung von Proteinen mit neuen Eigenschaften notwendig sind, weil die Antwort stark von der Art der Fragestellung abhängt. Wir wissen nicht, wie viele Varianten getestet werden müssen, um die Spezifität eines Enzyms zu verändern oder ein Enzym zu entwickeln, das eine neue chemische Reaktion katalysiert. An dieser Stelle sei nochmals betont, dass bereits die potenzielle Diversität eines kleinen monomeren Proteins überwältigend ist. Geht man davon aus, dass 32 DNA-Codons, also die Hälfte des genetischen Codes, ausreicht, um alle 20 proteinogenen Aminosäuren zu codieren, ergibt die Gesamtheit aller möglichen Kombinationen aller 20 Aminosäuren an jeder Position eines 200 Reste langen Proteins eine zu testende Bibliothek von  $32^{200}$  Varianten. Wenn wir auch die Notwendigkeit von Insertionen und Deletionen hinnehmen, wächst die Komplexität des Problems noch weiter. Werden allerdings alle 20 Aminosäuren in jeder Position eines 200 Reste langen Proteins unabhängig voneinander getestet, beträgt die Größe der Bibliothek nur  $32 \times 200$ . Als Anhaltspunkt dazu sei erwähnt, dass 1 mg Plasmid-DNA, eine vernünftige DNA-Menge, die sich routinemäßig in großem Maßstab isolieren lässt, aus etwa  $10^{14}$  Molekülen besteht. Wenn die DNA begrenzender Faktor ist, können also ca.  $10^{12}$  Positionen unabhängig voneinander zufallsbedingt variiert werden. Werden die Positionen aber alle gleichzeitig variiert, reicht die DNA nur für 9 Positionen. Mit einigen Methoden gelingt es inzwischen, sich der Grenze anzunähern, bei der die DNA begrenzender Faktor ist, doch in den meisten Fällen begrenzen andere Schritte im Prozess die erreichbaren Zahlen. Bei jeder Methode zur Verknüpfung eines Proteins

mit seiner DNA-Sequenz wird daher untersucht, welche Schritte den erreichbaren Durchsatz begrenzen, wie groß dieser ist und welche Möglichkeiten bestehen, die limitierenden Schritte zu verbessern. Wir beginnen gerade erst, ein Gefühl dafür zu entwickeln, wie viele Reste variiert werden müssen und ob diese Variationen unabhängig voneinander durchgeführt werden können, um die Proteinspezifität zu verändern. Wenn allerdings die Funktion des Proteins verändert werden soll, haben wir noch nicht einmal ein Gefühl für dieses Problem. Die Beispiele in diesem Aufsatz sind entsprechend auch immer mit einem Blick auf den Durchsatz dargestellt, der für die Beantwortung der jeweiligen Fragen notwendig ist.

Außerdem konzentrieren wir uns auf die allgemeine Verwendbarkeit der Methoden und den notwendigen technischen Aufwand. Bezüglich der allgemeinen Nutzung gibt es zwei Gesichtspunkte: Der erste ist, ob die Methode mit einer großen Anzahl von Proteinen mit unterschiedlichen Strukturen kompatibel ist: Kleine monomere cytosolische Proteine sind meist leicht in den Griff zu bekommen und in einer Vielzahl von Formaten testbar. Oligomere, membrangebundene und komplexere Proteine lassen sich oft nur mit Schwierigkeiten außerhalb der Zelle exprimieren. Beim Protein-Engineering ist dies ein weniger gravierendes Problem, denn man kann einfach von einem leicht zu handhabenden Protein ausgehen. Um allerdings alle exprimierten Proteine einer vorgegebenen Zelllinie zu testen, muss man eine große Variationsbreite von verschiedenen Proteintypen erfassen können. Der andere Punkt bei der allgemeinen Nutzung liegt in der Chemie und dem Testsystem. Im Idealfall wünscht man sich Methoden, die für ein breites Feld molekularer Wechselwirkungen und chemischer Reaktionen geeignet sind. Indem wir einige Anwendungen der verschiedenen Methoden diskutieren, möchten wir zeigen, welche besser für Bindungs-Tests geeignet sind, welche für Tests der Aktivität von Enzymen und für welche Methoden man Tests entwickeln kann, die reaktionsunabhängig sind. Wie bei jeder Methode bestimmt auch hier der technische Aufwand einer Methode in vielen Fällen die Verwendbarkeit. Wenn möglich sollten Methoden auch in nicht spezialisierten Labors eingesetzt werden können. Mit ausgewählten und detailliert ausgeführten Beispielen behandeln wir Gesichtspunkte wie den Bedarf nach spezieller Ausrüstung für das Display der Proteine oder die Testdurchführung, wie häufig falsch positive Ergebnisse sind, wie leicht sie mit Sekundär-Testsystemen eliminiert werden können und wie einfach die DNA am Ende des Tests für weitere Mutageneserunden oder für eine Sequenzierung amplifiziert werden kann.

Diesen Aufsatz haben wir nach den verschiedenen Methoden zur Kopplung eines Proteins mit seiner DNA-Sequenz gegliedert. Die Methoden sind danach eingeteilt, ob sie auf einer direkten Verbindung, Kompartimentierung oder räumlichen Trennung beruhen. In einem kurzen Abschnitt am Ende jedes Kapitels wird ein Ausblick auf neue Ansätze gegeben. Für jede Methode beschreiben wir das Arbeitsprinzip, entscheidende technische Verbesserungen und ausgewählte Anwendungen. An ihnen werden die Fragen nach Zahlen, allgemeiner Einsetzbarkeit und technischem Aufwand diskutiert. Es gibt beeindruckende technische Fortschritte auf allen Gebieten, die wir in diesem Aufsatz nicht im Entferntesten würdigen können und so sind jedem Kapitel Literaturhinweise mit detaillierteren Übersichtsartikeln für jedes Feld vorangestellt.

## 2. Die Methode der direkten Verknüpfung

Die wahrscheinlich direkteste Verknüpfung eines Proteins mit seiner DNA-Sequenz besteht in einer Verbindung zwischen jedem Protein und der DNA, die es codiert. Die Verbindung kann direkt sein (wie im mRNA-Display oder bei Peptiden auf einem Plasmid) oder indirekt wie im Phagen- und Zelloberflächen-Display. Es gibt viele Strategien zur Herstellung solcher direkter Verbindungen. Sie eignen sich hauptsächlich für Bindungstests, und es gibt bereits mehrere eindrucksvolle Beispiele für die Entwicklung von Proteinen mit neuen Funktionen. Vor kurzem gelang in einigen Labors auch die Messung katalytischer Aktivitäten. Die erreichbaren Bibliotheksgrößen liegen am nächsten bei der durch die DNA gegebenen Grenze von  $10^{14}$  verschiedenen Molekülen. Als technische Randbedingungen müssen die Stabilität der Verbindung und die Vereinbarkeit mit vielen strukturell verschiedenen Proteinen beachtet werden.

### 2.1. Phagen-Display-Technik

Die Phagen-Display-Technik ist eine der ersten Methoden, die zur direkten Verbindung eines Proteins mit seiner DNA-Sequenz eingeführt wurde. Es überrascht nicht, dass mit dieser Technik dann einige aufregende Fortschritte erzielt wurden und sie ausführlich in der Literatur beschrieben wurde.<sup>[1-4]</sup> Das Phagenpartikel stellt die Verknüpfung zwischen jedem Protein und seiner charakteristischen DNA-Sequenz her. Gewöhnlich wird für die Phagen-Display-Technik der Phage M13 verwendet; dieser besteht aus einem einsträngigen DNA-Molekül (dem Phagengenom), welches von Tausenden von Kopien der Phagenhüllproteine umgeben ist (Abbildung 2 A). Diese Hüllproteine werden von dem einsträngigen DNA-Genom codiert, das sie einkapseln. Bei der Phagen-Display-Technik wird die DNA, die das interessierende Protein codiert, mit einem der Hüllproteingene fusioniert, sodass die DNA-Sequenz zu einem Teil des Phagengenoms wird und das Protein auf der Phagenoberfläche als Fusionsprotein mit dem Hüllprotein exprimiert wird. Die Hüllproteinschicht besteht aus Tausenden von Kopien des Haupt-Hüllproteins pVIII, fünf Kopien von pIII

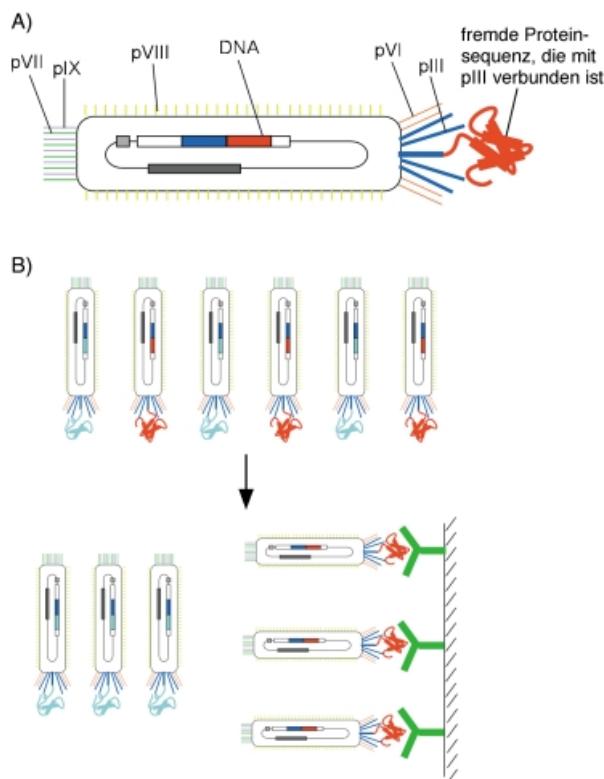


Abbildung 2. A) Struktur des filamentösen M13-Phagen, der bei der Phagen-Display-Technik eingesetzt wird. Ein Phagenpartikel besteht aus einem einsträngigen DNA-Molekül und wird von mehreren Hüllproteinen umgeben: pVIII, das Haupthüllprotein, pVII und pIX an einem Ende und pIII und pVI am anderen Ende. Eine fremde Proteineinsequenz kann mit jedem der Hüllproteine fusioniert werden. Hier ist ein Protein (in rot) dargestellt, das mit dem N-Terminus von pIII (in blau) fusioniert wurde. B) Anreicherung von Phagen durch Panning.

und pVI an einem Ende des Phagenpartikels und fünf Kopien von pVII und pIX am anderen Ende (Abbildung 2 A). 1985 kam Smith auf die Idee, mit Phagenpartikeln Protein-Bibliotheken anzureichern, indem er affinitätschromatographisch die Partikel reinigte, die die fremde Peptidsequenz an ihrer Oberfläche exprimierten, in diesem speziellen Fall Fragmente der EcoRI-Endonuclease.<sup>[5]</sup> Fragmente der Eco-RI-Endonuclease wurden ins pIII-Hüllprotein inseriert, indem das entsprechende Fragment des EcoRI-Gens im Leserahmen in das pIII-Gen des Phagengenoms kloniert wurde. Die so modifizierten Phagen behielten ihre Lebensfähigkeit und Infektiosität. Allerdings wurde die Infektiosität durch Vorinkubation mit anti-EcoRI-Antikörpern deutlich verringert. Zusammengenommen konnte durch diese Arbeiten gezeigt werden, dass fremde Proteineinsequenzen auf der Phagenoberfläche in immunologisch zugänglicher Form präsentiert werden können. Schließlich wurden anti-EcoRI-Antikörper an eine Polystyrol-Petrischale adsorbiert und zur selektiven Anreicherung der Phagen, die das EcoRI-Fragment präsentieren, aus Wildtyp-Phagen genutzt.

#### 2.1.1. Technische Betrachtungen

Die Größe von Phagen-Display-Bibliotheken wird durch die Transformationseffizienz von Bakterien begrenzt. In

einem typischen Phagen-Display-Experiment wird eine Phagemid-DNA-Bibliothek zunächst *in vitro* hergestellt und dann in geeignete *Escherichia coli*-Zellen transformiert. Nimmt man an, dass mit Elektroporation 10 µg DNA ( $10^{12}$  DNA-Moleküle) in 200 µL kompetente Zellen ( $10^{11}$  Zellen) transformiert werden, kann man  $10^{10}$  Zellen erhalten, von denen jede ein spezifisches DNA-Molekül enthält. In der Praxis werden im Allgemeinen Bibliotheken mit etwa  $10^8$  Elementen hergestellt. Vor kurzem nutzten Sblattero und Bradbury die Rekombination in Bakterien, um eine scFv-Bibliothek in Phagen mit  $3 \times 10^{11}$  Elementen herzustellen.<sup>[6]</sup> Theoretisch kann die Elektroporation mit 1 mg DNA und 20 mL geeigneten Zellen so optimiert werden, dass eine Bibliothek mit  $10^{12}$  Elementen entsteht.<sup>[7]</sup> In den Transformantnen werden die Proteine des Phagemids exprimiert und die Phagenpartikel zusammengesetzt, wenn das Fusionsprotein diesen Vorgang nicht beeinträchtigt. Die Phagenpartikel werden dann von den transfizierten Bakterien geerntet und auf Proteine mit spezifischen Bindungseigenschaften selektioniert, indem man sie durch eine Affinitätsmatrix passieren lässt (ein Verfahren, das man „Panning“ nennt). Nicht bindende Phagen werden weggewaschen und bindende Phagen können mit Salzlösungen hoher Ionenstärke, mit sauren Lösungen oder mit Lösungen des Liganden eluiert werden (Abbildung 2B). Die ausgewählten Phagen können durch eine neue Infektionsrunde in Bakterien amplifiziert und unter stringenteren Selektionsbedingungen weiter angereichert werden. Zum Schluss werden die selektierten Klone durch Sequenzierung und In-vitro-Bindungstests analysiert.

Seit der ersten Veröffentlichung von Smith gab es mehrere wichtige technische Verbesserungen der Phagen-Display-Technik. So dienten Coinfektionen mit Helperphagen dazu, die Qualität des Display-Experiments zu sichern.<sup>[8]</sup> Neben dem Hüllprotein pIII wurden auch andere Hüllproteine, darunter pVIII,<sup>[9]</sup> pVII und pIX<sup>[10]</sup> herangezogen, um Qualität und Orientierung zu kontrollieren und um gleichzeitig mehrere Komponenten auf der Phagenoberfläche zu präsentieren. Erstmalig verfolgten Duenas und Borrebaeck<sup>[11, 12]</sup> als besonders geschickte Variation die Verwendung von selektiv infektiösen Phagen (SIPs), bei der sie den modularen Aufbau des Hüllproteins pIII nutzten. pIII besteht aus drei Domänen – N1 (68 Aminosäuren), N2 (132 Aminosäuren) und CT (149 Aminosäuren) (Abbildung 3A). N1 vermittelt die Infektion von *E. coli*. N2 erkennt spezifisch den F-Pilus und verbessert die Effizienz der Infektion, ist aber nicht essenziell. Die CT-Domäne bildet einen Teil der Hülle und spielt bei der Morphogenese des Phagen eine Rolle. In ihrem ersten Artikel berichteten Duenas und Borrebaeck über die Präsentation von Antikörperfragmenten von Anti-Hühnereiweiß-Lysozym (HEL) auf der Phagenoberfläche als Fusionsprotein für pIII, dem die N1-Domäne fehlte. Dadurch waren die Phagen nicht mehr infektiös (Abbildung 3B). Wurden die Phagen allerdings in *E. coli*-Zellen zusammengesetzt, die ein HEL-N1-Fusionsprotein exprimierten, wurde die Funktion von pIII und damit die Infektiosität des Phagen durch die Bindung zwischen dem anti-HEL-Antikörperfragment und HEL rekonstituiert (Abbildung 3C). So wurden die Phagen, die den anti-HEL-Antikörper präsentierte, bei der Infektion von Bakterien um den Faktor  $10^5$  angereichert. SIP hat gegenüber

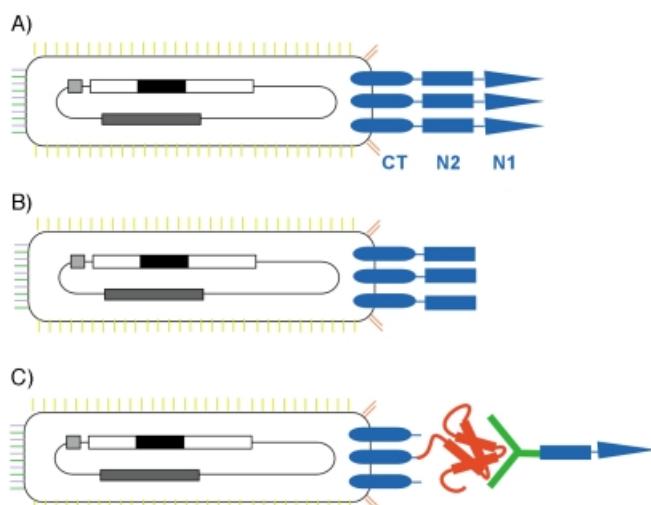


Abbildung 3. Selektiv infektiöse Phagen: A) Bei Wildtypphagen besteht das Hüllprotein pIII aus drei Domänen: den N-terminalen Domänen N1 und N2 und der C-terminalen Domäne CT. B) Phagenpartikel ohne die Domäne N1 sind nicht infektiös. Phagenpartikel, denen N1 und N2 fehlen (nicht abgebildet), sind ebenfalls nicht infektiös. C) Die Infektiosität kann restituiert werden, wenn ein Protein, das mit CT fusioniert ist, mit einem Protein, das mit N1-N2 fusioniert ist, in Kontakt tritt.

der traditionellen Phagen-Display-Technik zwei wichtige Vorteile. Zum einen geht es schneller als die herkömmliche Variante, weil der Waschschritt vermieden wird. Zum anderen überwindet es eine wichtige technische Einschränkung der Phagen-Display-Technik, dass nämlich viele Proteine nicht als Fusionen mit pIII auf der Phagenoberfläche exprimiert werden können.

### 2.1.2. Protein-Protein-Wechselwirkungen

Kurze Peptidsequenzen waren unter den ersten „Proteinen“, die auf Phagen präsentiert wurden. Peptidbibliotheken auf Phagen wurden bei der Suche nach optimalen Epitop-Sequenzen,<sup>[13]</sup> Peptidliganden<sup>[14]</sup> und Enzyminhibitoren<sup>[15]</sup> eingesetzt und auch, um Peptide zu finden, die zwischen Halbleiter-Monoschichten mit feinen Strukturunterschieden unterscheiden können.<sup>[16]</sup> Phagen-Peptid-Bibliotheken sind inzwischen kommerziell erhältlich (Ph.D., New England Biolabs). In der Hoffnung, auf die zeitraubende Hybridoma-Technik verzichten zu können, wurde die Methode sehr intensiv bei der Konstruktion von Antikörpern getestet.<sup>[17, 18]</sup> Ein Hauptproblem war die erfolgreiche Ausstellung von Antikörpern auf der Phagenoberfläche (die verschiedenen verwendeten Antikörperfragmente sind in Abbildung 4 zusammengestellt). Ein Antikörper besteht aus zwei Polypeptidketten, einer leichten (L) und einer schweren (H) Kette mit der Stöchiometrie  $L_2H_2$ . Jede Kette besitzt eine variable Region ( $V_L, V_H$ ) und eine konstante Region ( $C_L, C_{H1}, C_{H2}$  und  $C_{H3}$ ). Die leichte und die schwere Kette sind untereinander durch Disulfidbrücken verbunden, ebenso wie auch die beiden schweren Ketten untereinander. Erstmals wurde 1990 gezeigt, dass die variablen Domänen der schweren und der leichten Antikörperkette als Fv-Fragment in einer Kette auf der Oberfläche von fd-Phagen exprimiert werden können.<sup>[19]</sup> Nur wenig später wurde auch über die Phagen-

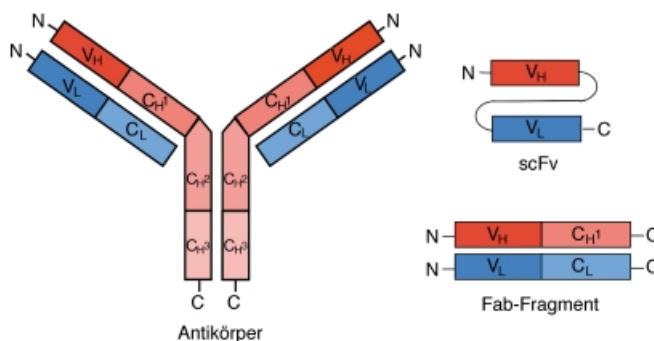


Abbildung 4. Antikörperfragmente, präsentiert auf einem Phagen. Antikörper können als scFv (Einzelketten-Fv-Fragment, *single chain Fv fragment*), bei dem die variable Domäne der schweren Kette mit der variablen Domäne der leichten Kette fusioniert wurde, oder als Fab-Fragment, bei dem die leichte Kette ( $V_L$ - $C_L$ ) und die N-terminale Hälfte der schweren Kette ( $V_H$ - $C_{H1}$ ) als Heterodimer vorliegen, auf Phagen präsentiert werden.

Display-Technik von Fab-Fragmenten berichtet.<sup>[20, 21]</sup> Nachdem dieses Hindernis erst einmal überwunden war, folgten viele weitere erfolgreiche Anwendungen. Verschiedene Phagen-Antikörper-Bibliotheken, darunter immunisierte, nicht-immunisierte und präparativ hergestellte Bibliotheken, wurden zur Isolierung neuer Antikörper und zur Verbesserung der Affinität existierender Antikörper gegen vorgegebene Antigene genutzt. Sogar katalytische Antikörper wurden mit der Phagen-Display-Technik selektiert.<sup>[22–24]</sup>

Die Phagen-Display-Technik ist offenbar gut geeignet für Anwendungen, bei denen eine gemeinsame Proteinfaltung eine große Rolle spielen, wie Antikörper und Peptide. Weniger klar ist allerdings, wie gut die Technik sich z. B. für die Zuordnung einer Funktion zu Proteinen einer cDNA-Bibliothek mit Tausenden von unterschiedlichen Proteinen einsetzen lässt. Shanmugavelu et al. berichteten vor kurzem über die Isolierung eines neuen Proteins, das an die Juvenilhormon-Esterase (JHE) bindet, aus einer cDNA-Bibliothek aus Insektengewebe mit der Phagen-Display-Technik.<sup>[25]</sup> Das Juvenilhormon (JH) kontrolliert zusammen mit Ecdysteroiden die Genexpression in Insekten. JHE spielt eine entscheidende Rolle bei der Insektenentwicklung, weil es JH abbaut. Die Autoren wollten Proteine identifizieren, die mit JHE wechselwirken und möglicherweise seine Aktivität regulieren. Die Größe der Bibliothek war unkritisch, weil die cDNA-Bibliothek nur etwa  $10^4$  Proteine umfasste. Die Herausforderung bestand dagegen in der Expression mehrerer Tausend verschiedener Proteine, die in der cDNA-Bibliothek aus dem Insektengewebe vorkommen. Wegen der großen Anzahl von Stop-Codons in der cDNA-Bibliothek konnte diese nicht als N-terminale Fusion mit pIII oder pVIII exprimiert werden. Stattdessen wurden die cDNAs als C-terminale Fusionsproteine mit der Fos-Leucin-Zipper-Domäne exprimiert, während die Jun-Leucin-Zipper-Domäne an den N-Terminus von pIII fusioniert wurde. Durch Heterodimerisierung von Fos und Jun wurden die Proteine effizient an der Phagenoberfläche präsentiert. In vier Anreicherungsrunden erhielt man Proteine, die an JHE (immobilisiert in den Vertiefungen einer 24-Loch-Polystyrol-Mikrotiterplatte) binden. In den ersten drei Runden der Selektion wurden die

gebundenen Phagen mit saurem Puffer eluiert. In der vierten Runde wurde gereinigte JHE verwendet. Mit den eluierten Phagen wurde dann *E. coli* infiziert. Phagenpartikel, die aus Einzelkolonien isoliert wurden, wurden in einer 96er Mikrotiterplatte mit ELISA auf JHE-Bindung getestet. Anzumerken ist, dass die mehrfachen Anreicherungsrunden nicht notwendig sind, um mehr Proteine testen zu können, sondern weil das Signal-Rausch-Verhältnis bei den Waschschriften ziemlich niedrig ist. Neun positive Klone wurden sequenziert, von denen einer das Protein P29 codierte, das bis dahin noch nicht charakterisiert war.

### 2.1.3. Protein-DNA-Wechselwirkungen

Nach den Antikörpern waren DNA-Bindeproteine eine der ersten Proteinklassen, bei denen die Bindungsspezifität mit der Phagen-Display-Technik verändert wurde.<sup>[26, 27]</sup> Die Pionierarbeit leisteten dabei Rebar und Pabo.<sup>[28]</sup> Sie präsentierten die drei Zinkfinger des Zif268-Proteins auf der Oberfläche des filamentösen Phagen als pIII-Fusionsprotein. Die vier Aminosäuren im ersten Zinkfinger, die aufgrund früherer Ergebnisse für die DNA-Erkennung am wichtigsten sind, wurden zufallsgesteuert variiert. So entstand eine Bibliothek aus  $10^6$  Elementen, die durch einen Pool von  $10^{10}$  Phagenpartikeln gut abgedeckt wurde. In dieser Bibliothek wurde auf drei biotinylierte doppelhelicale DNA-Sequenzen selektiert, die auf Streptavidin-beschichteten Mikrotiterplatten immobilisiert waren. Die DNA-Sequenzen (GACC, GCAC und CCTG) unterschieden sich von der Wildtyp-Sequenz (GCGC) an der Stelle, die von dem ersten Finger erkannt wird. Zinkfinger, die eine der modifizierten DNA-Sequenzen mit  $K_D$ -Werten im niedrigen picomolaren Bereich und mit einer 10 bis 100fachen Präferenz gegenüber der Wildtyp-Sequenz binden konnten, wurden für GACC und GCAC identifiziert, nicht aber für CCTG. Auch hier waren wegen des hohen Hintergrundes an nichtspezifischen Phagenpartikeln, die durch den Waschschnitt mitgeschleppt wurden (Abbildung 13B), fünf Selektionsrunden nötig. Für den Verlauf der Anreicherung mit der Zinkfinger-Bibliothek gaben die Autoren an, dass die Retentionsraten bei 0.01 % begannen und bei erfolgreicher Selektion bis auf ca. 1 % anstiegen. Diese und ähnliche Ergebnisse zeigen, dass bei einem nativen Protein, dessen Erkennungsmechanismus gut bekannt ist, gezielte Mutationen zur Erzeugung von Bibliotheken mit ca.  $10^6$  Elementen ausreichen, um die Proteinspezifität zu modifizieren.

### 2.1.4. Wechselwirkungen zwischen Proteinen und niedermolekularen Substanzen

Die nächste Frage ist, wie viele weitere Varianten getestet werden müssen, wenn man über die Veränderung der Spezifität hinaus zur De-novo-Entwicklung von Erkennung übergeht. Während Antikörper natürlich routinemäßig auf Bindung unterschiedlicher Moleküle hin entwickelt worden sind, geht man allgemein davon aus, dass die Anpassungsfähigkeit des Antikörpergerüsts einmalig ist. Es gab verschiedene Erfolge bei Protein-Protein-Wechselwirkungen,<sup>[29]</sup> doch ein entscheidender Durchbruch war erst der Nachweis, dass

ein Lipocalin-Protein, nämlich das Bilin bindende Protein (BBP) des Schmetterlings *Pieris brassicae*, zur De-novo-Bindung niedermolekularer Substanzen weiterentwickelt werden konnte.<sup>[30, 31]</sup> Die Lipocaline haben die Form eines Fasses aus acht antiparallel angeordneten  $\beta$ -Faltblatt-Strängen, wobei die vier Schleifen an einem Ende des Fasses den Eingang zur Bindungstasche bilden. Die sechzehn Reste, die mit dem Bilin direkt in Kontakt stehen und die in diesen vier Schleifen aufgereiht sind, wurden vollständig zufallsbedingt variiert. Da nur  $10^8$  individuelle Transformanten erhalten wurden, wurde nur ein sehr kleiner Bruchteil aller möglichen  $10^{24}$  Lipocalinvarianten getestet. Die mutierten BBPs wurden mit pIII fusioniert und monovalent auf Phagen ausgestellt. Selektioniert wurde durch Anreicherungsschritte der Phagen-Fusionsbibliothek gegen Fluorescein, das mit BSA oder RNase konjugiert war. Nach sechs Anreicherungsrunden zur Steigerung des Prozentsatzes an Fluorescein-spezifischen Phagen wurden BBP-Mutanten isoliert, die eine Affinität von  $10^{-7}\text{ M}$  und eine hohe Spezifität für Fluorescein aufwiesen.<sup>[30]</sup> Auf ähnliche Weise wurden auch BBP-Mutanten mit hoher Affinität und Spezifität für das Steroid Digoxigenin gefunden.<sup>[31]</sup> Weitere Experimente sind nötig, um zu testen, ob die Lipocalin-Falte ähnlich wie die IgG-Falte eine bevorzugte Gerüststruktur ist, oder ob tatsächlich Bibliotheken mit etwa  $10^8$  Elementen ausreichen, um molekulare Erkennung de novo zu entwickeln.

### 2.1.5. Enzymkatalyse

Während sich die Phagen-Display-Technik besonders für die Selektionen, die auf Bindungen basieren, anbietet, ist es eine größere Herausforderung, diese vielseitige Technik an die Katalyse anzupassen. Antikörper, die eine Vielzahl von Transformationen katalysieren, wurden durch die Weiterentwicklung von katalytischen Antikörpern gegen die Analoga von Übergangszuständen gewonnen. Theoretisch sollte es nicht schwierig sein, diese Methode an die Phagen-Display-Technik anzupassen, denn dort beruht die Selektion normalerweise auf einer Bindung. Beispielsweise katalysiert der Antikörper 17E8, der durch Immunisierung mit einem Analogon des Übergangszustandes von Norleucinphosphat gewonnen wurde, die Hydrolyse von Aminosäure-Phenylestern.<sup>[32]</sup> Baca et al.<sup>[23]</sup> verpflanzten die Antigen bindende Schleife von 17E8 auf das Gerüst eines menschlichen Antikörpers, um die Effizienz seiner rekombinanten Expression zu verbessern, und präsentierte es auf pIII als Fab-Fragment. Mit ortsspezifischer Mutagenese wurden humanisierte 17E8 (hu17E8)-Bibliotheken hergestellt. Nach fünf bis sechs Anreicherungsrunden wurden Mutanten selektioniert, die das Übergangszustandsanalogon von Norleucinphosphat binden. Daraus wurde eine Mutante mit einer doppelt so hohen katalytischen Aktivität identifiziert. Diese Mutante band allerdings das Analogon des Übergangszustandes weniger fest als der ursprüngliche hu17E8, während alle isolierten Mutanten mit höherer Bindungsaffinität eine niedrigere katalytische Effizienz aufwiesen. Dieses Beispiel verdeutlicht, dass die Bindung eines Analogons des Übergangszustandes und die katalytische Aktivität nicht dasselbe sind. Auffällig ist, dass es trotz der weiten Verbreitung des Ansatzes bei

katalytischen Antikörpern nur wenige Beispiele gibt, bei denen Analoga von Übergangszuständen genutzt wurden, um Katalysatoren mit der Phagen-Display-Technik herzustellen. Die Aufmerksamkeit scheint sich auf die direkte Selektion von katalytischer Aktivität zu richten.

Auch Suizidsubstrate wurden bei dem Affinitätsanreicherungsschritt eingesetzt, um Enzyme aufgrund ihrer katalytischen Aktivität zu selektieren. Soumillion et al.<sup>[33]</sup> synthetisierten einen  $\beta$ -Lactamase-Suizidinhibitor, der über eine Disulfidbrücke mit Biotin verbunden war. Die  $\beta$ -Lactamase war mit pIII fusioniert, wobei sich zwischen den Domänen eine Protease-Spaltstelle befand. Phagenpartikel mit aktiven und inaktiven  $\beta$ -Lactamasen wurden mit dem Suizidinhibitor inkubiert; aktive Enzyme waren danach mit Biotin markiert. Der Selektionsschritt mit Streptavidin-Kügelchen erfasste selektiv Phagenpartikel, die aktive Enzyme trugen. Auch ein ähnlicher Selektionsansatz, bei dem die Enzymaktivität mit der Infektiosität des Phagen gekoppelt wurde, wurde beschrieben.<sup>[34]</sup> Suizidsubstrate wurden zwar eingesetzt, um das Wildtyp-Enzym zu isolieren, doch es ist unklar, ob sie auch zwischen Enzymen mit unterschiedlicher katalytischer Effizienz unterscheiden können.

Wirklich benötigt wird ein direkter Test auf katalytische Aktivität, im Idealfall ein Test, der reaktionsunabhängig ist. Dazu wurden mehrere verwandte Methoden,<sup>[35–39]</sup> die auf der Bildung oder Spaltung einer Bindung an einem festen Träger oder der differentiellen Erkennung von Substrat und Produkt beruhen, veröffentlicht. In einem solchen Beispiel wurde eine unspezifische Protease, die Staphylokokken-Nuclease (SNase), aus einer Phagenbibliothek angereichert, die nicht verwandte Proteine enthielt.<sup>[35]</sup> Das Selektionsprinzip ist in Abbildung 5 illustriert: SNase und das DNA-Substrat wurden beide an die Phagenoberfläche gebunden. SNase, ein Enzym, das RNA und DNA hydrolysiert, lag fusioniert mit pIII auf einem Phagenvektor vor. Das DNA-Substrat wurde mithilfe eines Fos-Jun-Leucin-Zipper-Heterodimers eingeführt. Ein Helperphage stellte die pIII-Fos-Fusion zur Verfügung, und die Biotin-Substrat-DNA-Jun-Fusion wurde chemisch synthetisiert. Die phagengebundene SNase lag zunächst in der inaktiven Form vor, weil im Puffer kein  $\text{Ca}^{2+}$  enthalten war, und wurde dann mit der Jun-Substrat-Biotin-Fusion inkubiert und an Streptavidinkügelchen adsorbiert. Nachdem die nicht-gebundenen Phagenpartikel weggeschwommen waren, wurde durch  $\text{Ca}^{2+}$ -Zusatz zum Puffer die SNase aktiviert. Phagen, die eine aktive SNase trugen, spalteten das Substrat und lösten damit das Phagenpartikel vom festen Träger. Phagen, die SNase trugen, konnten in einer Runde aus einer 1:100-Mischung mit Phagen, die ein Fab-Fragment trugen, hundertfach angereichert werden. Vor kurzem wurde diese Methode genutzt, um DNA-Polymerase aus *Thermus aquaticus* in RNA-Polymerase zu konvertieren.<sup>[38]</sup> Die Verwendung von Leucin-Zipper-Peptiden zum Einbau des Substrats sollte es vereinfachen, die Chemie zu variieren. Der momentan zur Verfügung stehende Assay ist nur bei Metall-abhängigen Ionen einsetzbar, bei denen die Enzymaktivität während der Selektionsschritte ein- und ausgeschaltet werden kann. Es ist nicht klar, ob dieses Schalten so schnell verläuft, dass zwischen Enzymen unterschiedlicher Aktivität differenziert werden kann.

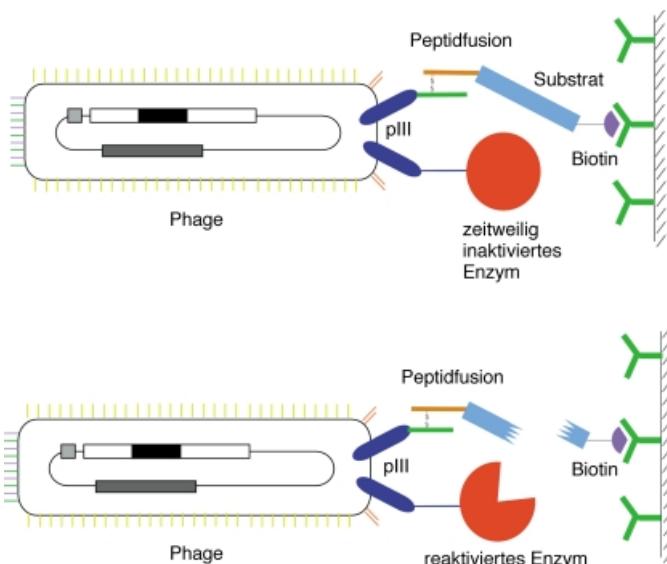


Abbildung 5. Selektion bezüglich der Enzymaktivität mit der Phagen-Display-Technik.<sup>[35]</sup> Das Substrat, ein Oligonukleotid, wurde mit einem Peptid-Heterodimer an das Phagenpartikel gebunden. Die Fusion von pIII und dem Peptid wurde durch einen Helperphagen ermöglicht. Die SNase wurde mit einem Phagemid-Vektor als pIII-Fusionprotein präsentiert. Sie war zunächst inaktiv, da kein  $\text{Ca}^{2+}$  im Medium vorhanden war (oben). Dann wurden die Phagen mit dem Komplex aus Peptid-Substrat und Biotin inkubiert und an Streptavidinkügelchen adsorbiert. Nach dem Wegwaschen nichtgebundener Phagen wurde die SNase durch Zugabe von  $\text{Ca}^{2+}$  aktiviert. Phagen, die aktive SNase-Enzyme präsentierte, spalteten das DNA-Substrat und lösten die Phagen damit vom festen Untergrund ab (unten).

## 2.2. Ribosomen-Display-Technik und mRNA-Peptid-Fusion

Bei der in diesem Abschnitt vorgestellten Methode besteht die Verknüpfung zwischen Phänotyp und Genotyp in dem nichtkovalenten oder kovalenten Komplex, der zwischen dem translatierten Peptid und der codierenden mRNA gebildet wird (Abbildung 6).<sup>[40]</sup> Erstmals wurde sie 1994 von

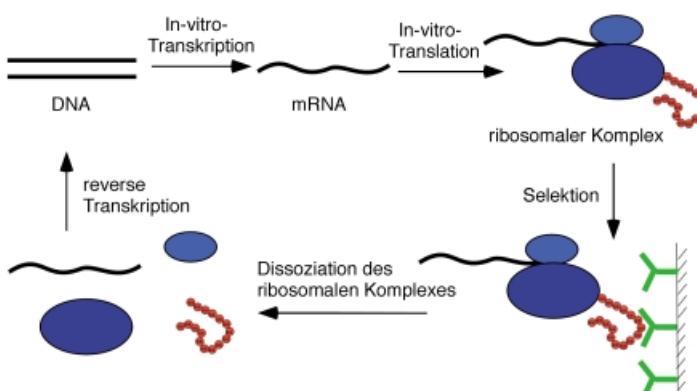


Abbildung 6. Ribosomen-Display-Technik.<sup>[41, 42]</sup> Die Polypeptide werden mit einem In-vitro-Transkriptions-/Translationssystem synthetisiert. Nach Ende der Translation sind Polypeptid und mRNA mit dem Ribosom vergesellschaftet. Die Ribosomenkomplexe werden dann durch Bindung der Polypeptide an einen festen Untergrund selektiert. Die ausgewählten Ribosomenkomplexe werden dissoziiert und die mRNA revers zu DNA transkribiert; die DNA wird sequenziert oder weiteren Runden von Mutagenese und Selektion unterworfen.

Mattheakis et al. vorgestellt,<sup>[41]</sup> und später wurde sie von Hannes und Plückthum optimiert.<sup>[42]</sup> Mit einem gekoppelten *E. coli*-S30-Transkriptions-/Translationssystem transkribierten und translatierten Mattheakis et al. in vitro eine Bibliothek von DNA-Sequenzen, die kurze Peptide codierten. Als die Translationsreaktion durch Zugabe von Chloramphenicol gestoppt wurde, blieben das translatierte Peptid und die mRNA mit dem ribosomalen Komplex assoziiert. Diese Komplexe wurden mit immobilisierten Antikörpern gegen die Peptide selektiert. Danach wurde die mRNA von dem Komplex dissoziiert, zur entsprechenden cDNA rücktranskribiert und amplifiziert, wobei Gene für Peptide angereichert wurden, die von den Antikörpern erkannt werden. Diese ausgewählten DNA-Sequenzen können weiteren Transkriptions-/Translationsrunden und weiterer Selektion unterworfen werden.

### 2.2.1. Technische Betrachtungen

Mattheakis et al. setzten in ihrem Experiment<sup>[41]</sup> 400 ng DNA ( $10^{12}$  Moleküle) ein, die in  $3 \times 10^{12}$  mRNA-Moleküle transkribiert wurden; in den 50  $\mu\text{L}$  des Transkriptions-/Translations-Reaktionsansatzes befanden sich etwa  $3 \times 10^{12}$  Ribosomen in einer Konzentration von etwa 0.1  $\mu\text{M}$ . Am Ende der Transkriptions-/Translationsreaktion waren ungefähr 27% der mRNA-Moleküle mit Ribosomen komplexiert. Die Ausbeute an gebildeten Ribosomenkomplexen hängt jedoch von der Länge des transkribierten Peptides ab. Bei ganzen Proteinen ist auch unter optimierten Bedingungen die Ausbeute nur etwa 0.2%.<sup>[42]</sup> Bei kurzen Peptiden ergibt 1 mg DNA also eine Bibliothek mit  $10^{15}$  Elementen, während sie bei großen Proteinen nur  $10^{12}$  Elemente umfasst.

Ein Problem des Ribosomen-Displays besteht in der Instabilität des ternären Ribosomen-mRNA-Peptid-Komplexes. Es ist schwierig, diesen Komplex während der Selektionsschritte intakt zu halten. Ein großer technischer Fortschritt war daher die Einführung einer kovalenten Bindung zwischen dem Peptid und der mRNA<sup>[43, 44]</sup> mithilfe von Puromycin, einem Antibiotikum, das das Aminoacyl-Ende der tRNA nachahmt (Abbildung 7A). In Abbildung 7B ist zu erkennen, dass Puromycin an das 3'-Ende einer mRNA angehängt wurde, deren Stopcodon durch einen DNA-Spacer ersetzt wurde. Wenn die mRNA in vitro translatiert wurde, hielt das Ribosom an dem DNA-Spacer, sodass das Puromycin-Molekül am Ende der mRNA die A-Position des Ribosoms besetzte und dort mit dem wachsenden Polypeptid reagieren konnte. Diese Methode wird auch als mRNA-Display bezeichnet. Optimierte Bedingungen für die Puromycin-Kopplungsreaktion, unter denen die Effizienz, mit der Puromycin an das wachsende Polypeptid bindet, zunimmt, wurden kürzlich veröffentlicht.<sup>[45, 46]</sup>

Ein typisches mRNA-Display-Experiment ist in Abbildung 7C illustriert: Eine DNA-Bibliothek wird zuerst in vitro in die mRNA-Bibliothek transkribiert. Anschließend wird die gesamte mRNA an eine chemisch synthetisierte 30 Basenpaare lange DNA-Sequenz mit Puromycin am 3'-Ende gebunden und von diesen mRNA-DNA-Puromycin-Hybridmolekülen eine In-vitro-Translation durchgeführt. Durch das Puromycin werden kovalente Komplexe zwischen der mRNA

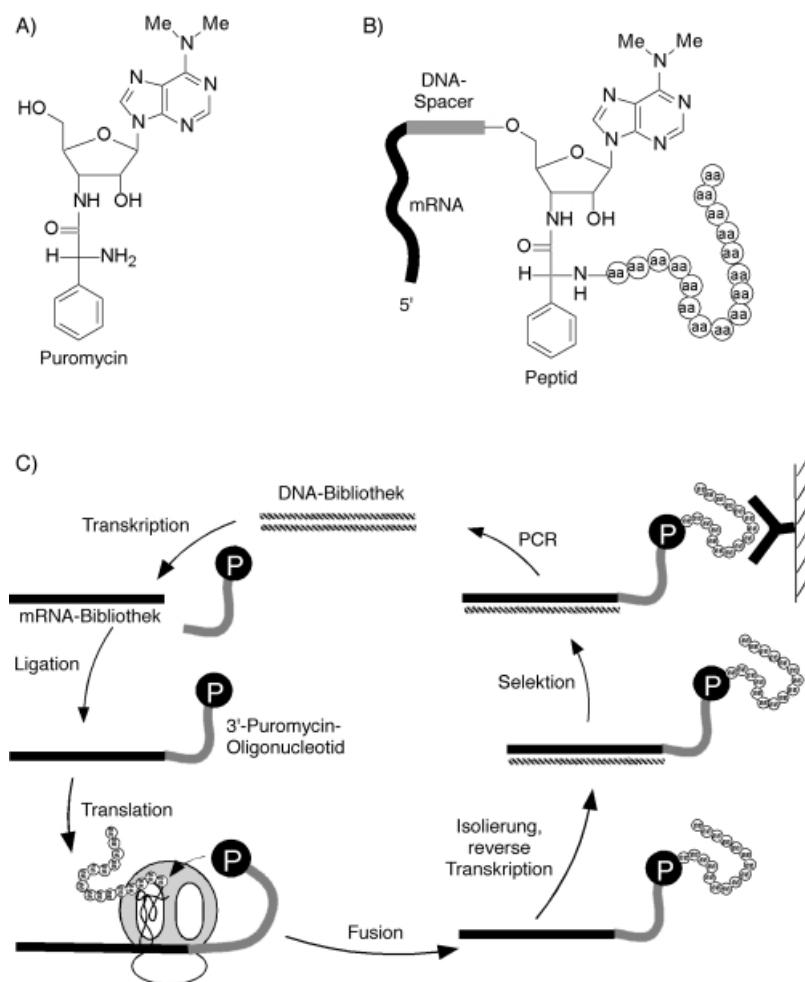


Abbildung 7. A) Struktur des Aminoacyl-tRNA-Analogons Puromycin. B) Kovalenter Komplex zwischen der mRNA und dem Peptid über Puromycin. C) Typischer Ablauf eines mRNA-Peptid-Fusionsexperiments. Wiedergabe mit Genehmigung nach Lit. [46].

und den Peptiden gebildet. Diese werden aus dem Reaktionsgemisch isoliert, revers transkribiert und danach selektiert, ob sie an ein immobilisiertes Antigen oder einen Liganden binden. Die selektierten Komplexe können dann mit PCR für die Sequenzierung amplifiziert oder weiteren Mutations- und Selektionsrunden unterworfen werden.

Auch die mRNA-Peptid-Fusionsmethode erhöht die Bibliotheksgröße. In dem ursprünglichen Artikel von Roberts und Szostak<sup>[43]</sup> wurden 5 pmol mRNA-Moleküle in einer 25 µL Translationsreaktion eingesetzt; 0.7 % der mRNA wurden in mRNA-Peptid-Fusionen eingebaut. Die Bibliothek umfasste also nur  $2.1 \times 10^{10}$  Elemente. Seit dieser ersten Veröffentlichung wurden die Kopplungsbedingungen optimiert, und es wurden Kopplungsausbeuten bis zu 40 % bei kurzen Peptiden erreicht.<sup>[46]</sup> Der mögliche Nachteil dieser Methode ist, dass alle Schritte *in vitro* durchgeführt werden, sodass es schwierig ist, Membranproteine oder Proteine, die posttranslational modifiziert werden müssen, zu exprimieren.

### 2.2.2. Protein-Protein-Wechselwirkungen

Wie die Phagen-Display-Technik wurde auch die Ribosomen-Display-Technik zunächst bei der Entwicklung von

Antikörpern angewendet. So erzeugten Hanes et al. scFvs, die eine Affinität gegen Insulin im picomolaren Bereich aufweisen, aus einer synthetischen Antikörperbibliothek.<sup>[47]</sup> Es gibt zwar nur wenige Berichte über die Ribosomen-Display-Technik bei der Entwicklung von Proteinen mit neuen Funktionen oder Aktivitäten, doch liegt dies vermutlich daran, dass die Methode noch relativ neu ist und diese Art der Anwendungen gegenwärtig noch bearbeitet werden. Theoretisch sollte die Ribosomen-Display-Technik für den gleichen Typ von Problemen geeignet sein wie die Phagen-Display-Technik.

Szostak und seine Mitarbeiter nutzten die Diversität der mRNA-Phagen-Display-Technik aus, um hochaffine Bindungsproteine aus zufällig zusammengesetzten Peptidbibliotheken zu isolieren.<sup>[48]</sup> Sie konstruierten eine Bibliothek aus 88 Aminosäureresten langen Peptiden nach dem Zufallsprinzip mit der optimierten Puromycin-Kupplungsmethode.<sup>[45, 46]</sup> Bei einer geschätzten Bibliotheksgröße von etwa  $7 \times 10^{12}$  wurde nur ein kleiner Anteil aller möglichen 88mere untersucht. Nach sieben Panning-Runden in Bezug auf die Affinität gegen immobilisiertes Streptavidin wurden Peptide mit einem  $K_D$  für Streptavidin von nur 5 nm isoliert. Auf ähnliche Weise wurden auch ATP-bindende Polypeptide mit nanomolarem  $K_D$  aus einer Peptid-Zufallsbibliothek selektiert.<sup>[49]</sup> Wie bei der Phagen-Display-Technik werden die wiederholten Panning-Runden nicht durchgeführt, um mehr Varianten zu testen, sondern um die RNA-Peptid-Fusionen mit echten Treffern anzureichern, denn ein großer Prozentsatz falsch

positiver Hits wird durch jede Selektionsrunde mitgeschleppt. Diese Ergebnisse sind überraschend und werfen die interessante Möglichkeit auf, dass die Evolution von Struktur und Funktion von Proteinen viel einfacher lösbar ist als man erwarten würde. Ausgehend von den Ergebnissen bei der Selektion ATP-bindender Peptide schätzen die Autoren, dass die Häufigkeit der funktionalen Polypeptide in der Größenordnung von  $10^{-11}$  liegt und dass diese Peptide nur mit der Ribosomen-Display-Technik isoliert werden konnten, weil damit Bibliotheken mit etwa  $10^{12}$  Elementen getestet werden können.

### 2.3. Peptide auf Plasmiden

Bei dieser Methode wird die nichtkovalente Wechselwirkung zwischen einem DNA-bindenden Protein und der Plasmid-DNA genutzt, um das Protein mit seiner DNA-Sequenz zu koppeln. Das erste dafür eingesetzte DNA-bindende Protein war der Lac-Repressor.<sup>[50]</sup> Der Lac-Repressor wird vom *lacI*-Gen codiert und bindet sehr fest an die spezifische DNA-Sequenz *lacO*. Wie in Abbildung 8 dargestellt, wurde eine Bibliothek aus  $10^8$  Dodecapeptiden an den

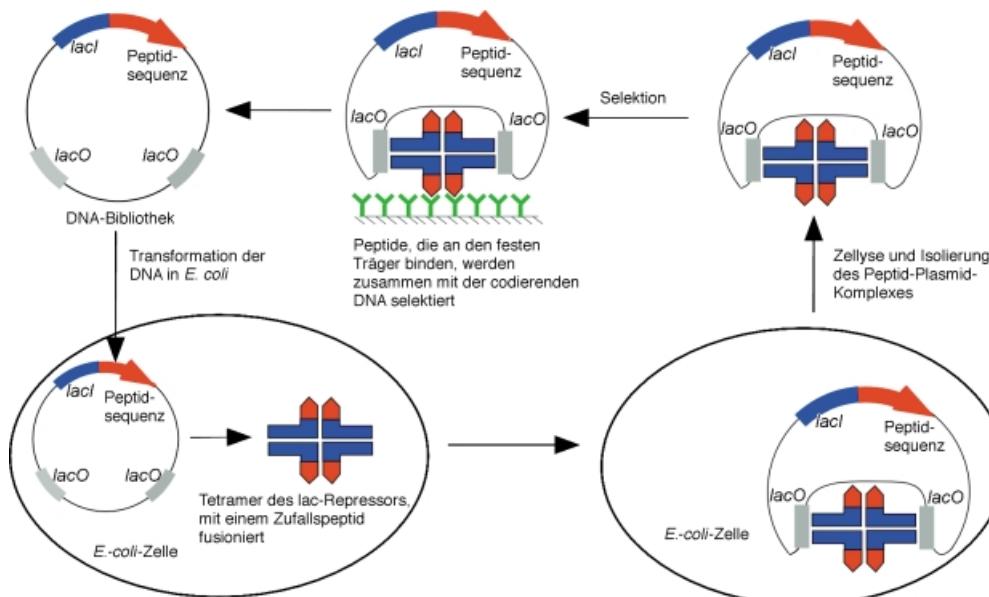


Abbildung 8. Peptid auf Plasmiden.<sup>[50]</sup> Zufällige DNA-Sequenzen werden mit dem *lacI*-Gen in einem Plasmid fusioniert, das außerdem *lacO*-Stellen trägt. Diese DNA-Bibliothek wird dann in *E. coli* transformiert. Zufallspeptide (rot) werden als Fusion mit dem Lac-Repressor LacI (blau) synthetisiert und binden an die *lacO*-Stellen des Plasmids als Tetramere.

C-Terminus des Lac-Repressors gebunden, indem synthetische Oligonukleotide stromabwärts zum *lacI*-Gen in das Plasmid pMC5 kloniert wurden, welches außerdem *lacO*-Sequenzen in Tandem-Anordnung enthielt. Im Zellinneren wurde durch das Plasmid die Synthese der Lac-Repressor-Dodecapeptid-Fusionskomplexe ausgelöst, die daraufhin Tetramere bilden und an die *lacO*-Sequenz des Plasmids binden. Nach der Zelllyse wurden die Peptid-Plasmid-Komplexe in Bezug auf die Bindungsfähigkeit der Peptide an einen immobilisierten Antikörper selektiert.

### 2.3.1. Technische Betrachtungen

Bei dieser Methode werden die Peptide in einem multivalenten Format präsentiert, da der Lac-Repressor vor der Bindung der DNA tetramerisiert und jedes Plasmid bis zu zwei Tetrameren binden kann. Daher eignet sich diese Methode zur Selektion von Bindungen mit niedriger bis mittlerer Affinität. Um sie auch an die Selektion hochaffinender Bindungspartner anzupassen, entwickelten Gates et al.<sup>[51]</sup> ein „Kopfstück-Dimer“ des Lac-Repressors als DNA-bindende Domäne. Das Kopfstück-Dimer, das weniger Kopien der Peptid-Bibliothek präsentiert, war zur Selektion höher affiner Bindepoteine geeignet, obwohl der DNA-Bindungsmechanismus unklar ist. Die Bibliotheksgröße dieser Methode ist der, die man mit der Phagen-Display-Technik erhält, ähnlich, denn sie wird ebenfalls durch die DNA-Transformationseffizienz von Bakterien begrenzt.

### 2.3.2. Protein-Protein-Wechselwirkungen

Theoretisch kann diese Methode für die gleichen Anwendungen eingesetzt werden wie die Phagen-Display-Technik; dennoch wurde sie bislang hauptsächlich zum Durch-

mustern von Peptid-Bibliotheken genutzt. 1997 berichteten Cwirla et al. über die erfolgreiche Isolierung eines Peptidagonisten für den Thrombopoietin-Rezeptor mit Wirkung im nanomolaren Bereich.<sup>[52]</sup> Zum Vergleich wandten die Autoren nicht nur die Peptid-auf-Plasmid-Methode an, sondern auch die Phagen- und Ribosomen-Display-Technik. Zunächst selektierten sie mit der pVIII-Phagen-Display-Technik und dem Lac-Repressor, die beide multivalente Display-Formate haben, Peptide mit  $IC_{50}$ -Werten zwischen 20 nm und 60 nm. Dann setzten sie ihre Selektion mit einem Kopfstück-Dimer und mit der Ribosomen-Display-Technik fort, gingen also zu einem monovalenten Display-Format über, und reicherten damit Peptidmutanten mit  $IC_{50}$ -Werten zwischen 20 und 60 nm an. Abschließend konnten sie mit der Kopfstück-Dimer-Methode das hochaffine Peptid AF12505 mit einem  $IC_{50}$  von 2 nm isolieren. Das Dimere von AF12505 hatte einen  $IC_{50}$ -Wert von 0.5 nm und war so wirksam wie rekombinantes humanes Thrombopoietin. Die Autoren schlossen aus ihren Versuchen – was allerdings von den Verfassern dieses Aufsatzes nicht nachvollzogen werden kann –, dass die Peptid-auf-Plasmid-Selektion von den drei verglichenen Methoden die effektivste sei, um hochaffine bindende Peptide zu gewinnen. Das geschilderte Anwendungsbeispiel verdeutlicht, wie wichtig es ist, die Valenz des Displays kontrollieren zu können. Hohe Valenz ist zu Beginn nötig, um niederaffine Liganden zu identifizieren, während in den Schlussrunden der Selektion eine niedrige Valenz die Unterscheidung der höchstaaffinen Liganden gewährleisten kann.

### 2.4. Zelloberflächen-Display

Methoden zur Präsentation von Proteinen auf der Oberfläche lebender Zellen sind schon lange bekannt, obwohl sie zunächst nicht zur Durchmusterung von Proteinbibliotheken entwickelt wurden.<sup>[53]</sup> Eine verbreitete Anwendung von Bakterien-Display ist z. B. die Entwicklung von Systemen zur Gewinnung von Lebendimpfstoffen. Das Prinzip ist ziemlich einfach: Ein Protein oder eine Proteinbibliothek wird an ein Membranprotein fusioniert, das als Anker dient, um die Proteine auf der Zelloberfläche zu präsentieren. Die Selektion oder Durchmusterung, normalerweise durch fluoreszenzaktiviertes Zellsortieren (FACS), wird dann mit den Zellen, die die Proteine präsentieren, durchgeführt. In Bakterien ist das am meisten genutzte Ankerprotein ein Lpp-OmpA-Hybrid, also ein Lipoprotein, fusioniert an die

Aminosäuren 46–159 des Protein A der äußeren Membran.<sup>[54]</sup> In neuerer Zeit wurde ein Hefesystem entwickelt, das das  $\alpha$ -Agglutinin, einen Zelloberflächen-Rezeptor, der an der Konjugation der Hefen beteiligt ist, als Anker nutzt.<sup>[55]</sup> Auch Systeme mit Säugerzellen wurden bereits entwickelt.<sup>[56]</sup>

#### 2.4.1. Protein-Protein-Wechselwirkungen

Obwohl auch diese Methode für die gleichen Anwendungen brauchbar sein sollte wie die Phagen-Display-Technik, wurde die Zelloberflächen-Display-Technik bislang hauptsächlich für die Präsentation von Peptid-,<sup>[57]</sup> Antikörper-<sup>[58]</sup> und T-Zell-Rezeptor-Bibliotheken<sup>[59]</sup> genutzt. Eines der eindrucksvollsten Beispiele findet sich in einer kürzlich erschienenen Publikation von Boder et al., in der die Erzeugung von scFvs mit femtomolarer Affinität für Fluorescein beschrieben wird.<sup>[60]</sup> Der Fluorescein bindende Antikörper scFv 4-4-20 wurde durch fehlertolerantes DNA-Shuffling mutagenisiert und über  $\alpha$ -Agglutinin auf der Oberfläche von Hefezellen präsentiert. Die Zellen wurden mit dem Fluorescein-Biotin-Komplex inkubiert, gewaschen und dann mit 5-Aminofluorescein als kompetitivem Liganden inkubiert. Zellen, die an Fluorescein-Biotin gebunden blieben, wurden anschließend mit Streptavidin-R-Phycoerythrin markiert und durch Fluoreszenz-aktivierte Zellsortierung (FACS) isoliert. Nach drei weiteren Mutagenese- und Screeningrunden wurde ein scFv mit einem  $K_D$  von 48 fm isoliert; diese Dissoziationskonstante liegt sogar niedriger als die für die Wechselwirkung zwischen Biotin und Streptavidin. Vermutlich liegt es am niedrigen Hintergrund der FACS-Methode, dass die Bibliothek weiter mutagenisiert und mit FACS gescreent werden kann, um so Antikörper mit steigender Affinität zu erzeugen.

#### 2.4.2. Enzymkatalyse

Vor kurzem wurde das Oberflächen-Display auf einer Bakterienzelle zum Screening nach enzymatischer Aktivität genutzt. Olsen et al. präsentierte eine mutagenisierte *E. coli*-OmpT-Protein-Bibliothek auf der Oberfläche von *E. coli*-Zellen.<sup>[61]</sup> In diesem Fall war kein Ankerprotein erforderlich, weil OmpT selbst ein membranassoziiertes Protein ist. OmpT ist eine Serinprotease, die eine Funktion bei der mikrobiellen Pathogenität zugeschrieben wird, und hat eine starke Präferenz für die basischen Aminosäurereste Arg und Lys in den Positionen P1 und P1'. Die OmpT-Bibliothek wurde nach Proteasen mit geänderter Spezifität für Substrate mit Arg und Val in den Positionen P1 und P1' durchsucht. Das Substrat (Abbildung 9) war ein Peptid, derivatisiert mit einem Fluorophor (F1) und einem löschen Fluorophor (Q), das als Partner für einen intramolekularen Fluoreszenz-Resonanz-Energietransfer (FRET) dient. Das Substrat trägt an der F1-Position drei positive Ladungen. Wenn das Peptid mit *E. coli*-Zellen inkubiert wurde, die an der Oberfläche negativ geladen sind, wurde es adsorbiert. Zellen, an deren Oberfläche Mutanten von OmpT präsentiert wurden, die gegenüber dem Substrat Arg-Val aktiv waren, spalteten das Peptid, setzten das löschen Fluorophor Q frei und änderten so das Fluoreszenzprofil der Zellen. Zellen, die aktives Enzym präsentierte, konnten dann durch FACS isoliert

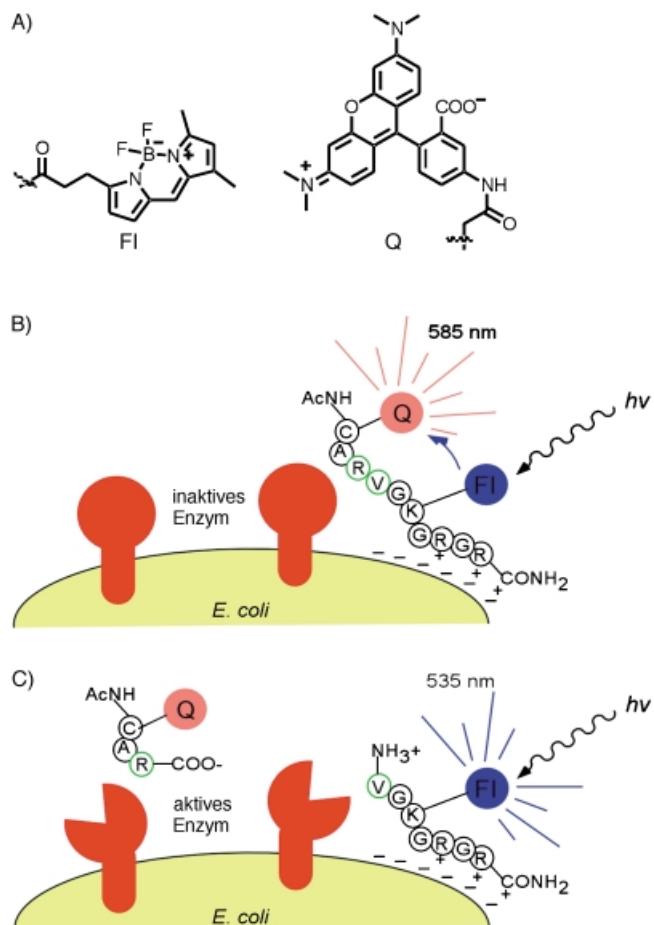


Abbildung 9. Screening-Verfahren auf Proteaseaktivität mit Bakterien-Display.<sup>[61]</sup> A) Strukturen des Fluorophors (F1) und des löschen Fluorophors (Q), der an das Peptidsubstrat gebunden ist. B) und C) Die Enzymbibliothek wird auf der Oberfläche von *E. coli* präsentiert, während das Substrat mit F1 und Q durch die ionischen Wechselwirkungen zwischen der negativ geladenen Zelloberfläche und dem positiv geladenen Substrat angezogen wird. Wenn ein inaktives Enzym präsentiert wird, löscht Q effektiv die Fluoreszenz von F1 und fluoresziert selbst bei 585 nm (B). Wird ein aktives Enzym präsentiert, spaltet dieses das Substrat und setzt Q vom Substrat frei; dadurch fluoresziert F1 bei 535 nm (C). Aufgrund der gemessenen Fluoreszenz der Zellen können aktive Enzyme isoliert werden.

werden. Mit dieser Methode durchsuchten die Autoren  $1.9 \times 10^6$  Zellen und isolierten 352 positive Klone. Diese 352 Treffer wurden in einem Test in 96er Mikrotiterplatten weiter auf Proteaseaktivität analysiert. Die aktivste OmpT-Mutante hatte eine 60fach bessere katalytische Aktivität gegenüber dem Arg-Val-Substrat im Vergleich zu dem nicht mutagenisierten OmpT. Die Methode ähnelt der Phagen-Display-Technik, mit dem Unterschied, dass der Test auf FRET aufgebaut ist anstatt auf Spaltung/Bildung einer Bindung an einen festen Träger. Dieser Ansatz verläuft nach einem allgemeingültigen Prinzip und sollte auf eine große Anzahl bindungsbildender und -spaltender Reaktionen anwendbar sein. Wahrscheinlich werden die chemischen Möglichkeiten etwas eingeschränkt durch die Abhängigkeit von FRET. FRET hängt sehr stark vom Abstand zwischen F1 und Q ab und wird außerdem durch die Orientierung von Donor- und Acceptor-molekülen zueinander beeinflusst. Will man diese

Methode auch nutzen, um einen ganzen Bereich katalytischer Aktivitäten abzudecken, muss man das Verhältnis der Enzymmoleküle zu den Substratmolekülen auf der Zelloberfläche und die Diffusionsraten genauer bestimmen.

## 2.5. Zusammenfassung

Die Phagen-Display-Technik ist ebenso wie die neueren Methoden zur direkten Verbindung eines Proteins mit seiner DNA-Sequenz eine robuste Methode zur gezielten Variation der Bindungsspezifität eines bestehenden Proteingerüsts. Während durch die DNA selbst die Zahl der Proteinvarianten, die getestet werden können, auf ca.  $10^{14}$  begrenzt wird, werden die Phagen-Display-Technik, Peptid-auf-Plasmid-Verfahren und Zelloberflächen-Display-Technik darüber hinaus durch die Transformationseffizienz eingeschränkt. Theoretisch können mit der Transformation von *E.-coli* Bibliotheken aus etwa  $10^{10}$  Elementen erreicht werden, doch wird in der Praxis von Bibliotheken mit etwa  $10^8$  Elementen berichtet. Ribosomen-Display erscheint am aussichtsreichsten, um die Grenze zu erreichen, die durch die DNA selbst gesetzt ist, und mit der Puromycin-Kopplungsmethode kann man bereits jetzt Bibliotheken mit  $10^{12}$  Elementen einfach erzeugen. Überraschend ist, dass diese Bibliotheksgrößen offenbar ausreichen, um die Proteinfunktion zu modifizieren. Beginnt man mit Proteinen, deren Wirkungsmechanismus gut verstanden ist, haben sich Bibliotheken in der Größenordnung von  $10^6$  bis  $10^8$  Elementen als ausreichend erwiesen, um Proteine mit neuen Spezifitäten für DNA oder niedermolekulare Substanzen zu erzeugen. In weiteren Untersuchungen muss jetzt abgesichert werden, dass diese Ergebnisse allgemein anwendbar sind und dass die Entwicklung von Proteinen mit neuen Spezifitäten tatsächlich ein lösbares Problem ist. Es bleibt jedoch noch immer unklar, wie viele Proteinvarianten getestet werden müssen, um eine neue Funktion zu entwickeln. Alle beschriebenen Methoden schränken das zu präsentierende Protein ein; die Begrenzungen sind am besten im Falle des Phagen-Display dokumentiert. Es ist zwar keine Fragestellung des Protein-Engineering, bei dem man wahrscheinlich mit einem gut bekannten Gerüst beginnen würde, aber viel Arbeit muss noch investiert werden, um herauszubekommen, welcher Anteil der Proteine aus einer vorgegebenen cDNA-Bibliothek heraus exprimiert und dann tatsächlich in aktiver Form mit den verschiedenen Methoden präsentiert wird. Diese Methoden sind natürlich auf Tests abgestimmt, die auf molekularer Erkennung beruhen, und das gleiche Anreicherungsverfahren kann unabhängig von der Art der Protein-Molekül-Wechselwirkung eingesetzt werden. Mit Ausnahme des Peptid-auf-Plasmid-Verfahrens ist die Kupplung zwischen Protein und DNA ziemlich robust und bedingt nur wenige Einschränkungen der Anreicherungsbedingungen. Es wird jedoch eine Herausforderung sein, diese Methoden an die Entdeckung katalytischer Aktivität anzupassen. Einige originelle Ansätze, die auf dem Bruch oder der Bildung einer Bindung mit einem festen Träger oder FRET beruhen, sind in der letzten Zeit veröffentlicht worden. Die Schwierigkeit wird darin bestehen, in diesen ja/nein-Tests Enzyme mit einem großen Bereich

katalytischer Effizienz auseinander zu halten. Während in die Herstellung der Bibliotheken noch einige Arbeit gesteckt werden muss, sind die Methoden selbst, seien es die Phagenpartikel oder das In-vitro-Transkriptions-/Translationssystem, robust und in vielen Labors etabliert. Das Signal-Rausch-Verhältnis in den Anreicherungsschritten ist schlecht und daher sind zahlreiche Runden für Bibliotheks-Amplifizierung und Anreicherung notwendig, um echte Wechselwirkungen zu identifizieren. Dies macht es wenig praktikabel, mehrere Mutageneserunden auszuführen, sodass insgesamt die Zahl der letztlich getesteten Proteinvarianten begrenzt ist. Bei den meisten Methoden kann die DNA für die Sequenzierung oder weitere Mutageneserunden durch die Infektion von Bakterien mit den Phagenpartikeln, die Retransformation von Plasmiden oder die Anzucht (Vermehrung) der Zellen einfach amplifiziert werden. Einzige Ausnahme ist das Ribosomen-Display, bei dem die RNA vor der Amplifizierung in cDNA umgeschrieben werden muss.

## 3. Kompartimentierungsmethoden

Ein ähnlicher Ansatz zur Kupplung von Protein und zugehöriger DNA-Sequenz ist, jedes Protein mit seiner DNA in einem eigenen Kompartiment einzuschließen. Dies wird normalerweise dadurch erreicht, dass Plasmid-DNA, die das Protein codiert, in eine Zelle eingebracht wird. In jüngster Zeit wurden auch Wasser-in-Öl-Emulsionen als künstliche Kompartimente eingeführt, die man in diesem Zusammenhang als „abgespeckte“ Zelle betrachten kann. Bei zellulären Testsystemen liegt der Vorteil darin, dass molekularbiologische und genetische Techniken gut etabliert sind, auch wenn die Größe der Bibliotheken durch die Transformationseffizienz der Zellen begrenzt ist. Wasser-in-Öl-Emulsionen dagegen haben wie die mRNA-Display-Technik das Potenzial, bis zu der Grenze vorzustoßen, die die DNA für die Bibliotheksgröße darstellt. Beide Ansätze scheinen gut für Enzymtests geeignet zu sein. Die meisten Fortschritte bei der Entwicklung von Enzymen mit neuen Eigenschaften wurden mit zellulären Testsystemen erzielt.

### 3.1. Zelluläre Testsysteme

Vor allem Genetiker haben schon immer Hochdurchsatz-Screening betrieben, sodass man von ihren Erfahrungen viel lernen kann. In zellulären Tests wird das Protein mit der DNA, die es codiert, verbunden, weil die DNA die Synthese des Proteins in der Zelle steuert. Dadurch stellt die Zelle ein Kompartiment dar, in dem jedes Protein mit seiner DNA verbunden ist. Die Testsysteme sind sehr effizient, weil Millionen von Plasmid-DNA-Molekülen, von denen jedes ein unterschiedliches Protein oder eine Proteinvariante codiert, in einem Ansatz in die Zellen transferiert werden können. Durch die Statistik des DNA-Transfers und den Gebrauch von Selektionsmarkern wird gesichert, dass jede Zelle am Ende nur ein einzelnes Plasmid enthält, das eine Proteinvariante codiert. Die Zellen können einfach dadurch voneinander getrennt werden, dass man sie bei entsprechen-

der Verdünnung auf festem Medium kultiviert. Am Ende des Experiments kann die DNA, die das Protein codiert, für die Sequenzierung oder weitere Mutageneserunden mit molekulärbiologischen Standardmethoden einfach aus der Zelle extrahiert werden. In einem inzwischen klassischen Experiment gelang es Beadle und Tatum in den frühen 40er Jahren, durch genetische Analysen Enzyme für die Biosynthese essenzieller Metaboliten zu identifizieren (Abbildung 10).<sup>[62]</sup> Sie induzierten in der chromosomal DNA des Ascomyce-

ten *Neurospora crassa* durch Röntgenbestrahlung zufällige Mutationen. Etwa 2000 einzelne Mutantenstämme wurden in Voll- und Minimalmedium auf Wachstum getestet, wobei das Minimalmedium nur solche Nährstoffe (z. B. Biotin) enthielt, die *Neurospora* nicht selbst synthetisieren kann. Drei Mutantenstämme wurden identifiziert, deren Wachstumsrate sich in Vollmedium nicht von der des Wildtyps unterschied, die aber in Minimalmedium nur schlecht oder gar nicht wuchsen. Durch Zugabe einzelner Nährstoffe zum Medium und anschließendem Vergleich der Wachstumsraten konnte für die drei Stämme gezeigt werden, dass sie die Fähigkeit verloren hatten, Vitamin B<sub>6</sub>, Vitamin B<sub>1</sub> bzw. p-Aminobenzoësäure zu synthetisieren. Dieser grundlegende Ansatz, die DNA, die die Proteine codiert, zu mutieren und dann auf einen nachweisbaren Phänotyp wie das Zellwachstum zu testen, ist die Basis für die genetische Analyse. Die meisten Gene, deren Funktion bekannt ist, sind mit Methoden dieses Typs identifiziert

worden, und die Funktionen Tausender Proteine sind auf ähnliche Weise bestimmt worden. Beadles und Tatums Leistung war der Nachweis, dass für die mutierten Stämme eine Enzymaktivität und nicht nur der Zelltod als Testprinzip dienen konnte. Wie wir in diesem Abschnitt diskutieren werden, besteht die Herausforderung noch immer darin, geschickte Testsysteme für die interessierenden Aktivitäten zu entwickeln.

### 3.1.1. Technische Betrachtungen

Die Fortschritte der Molekulärbiologie eröffneten eine ganze Reihe origineller genetischer Kunstgriffe, mit der die Handhabbarkeit zellulärer Tests verbessert wurde. Eine der größten Hürden bei der genetischen Analyse ist der Rücksluss auf das Gen mit der interessierenden Mutation. So waren nach dem oben beschriebenen Experiment von Beadle und Tatum noch einige weitere Jahre experimenteller Arbeit nötig, um zu bestimmen, welche Gene im *Neurospora*-Genom die Mutation trugen, die die Cofaktor-Biosynthese beeinträchtigten und die daher biosynthetische Enzyme codierten. Eine raffinierte Variante wurde von Holden et al. 1995 eingeführt:<sup>[63]</sup> Sie ersetzten jeden offenen Leserahmen in einem Genom durch eine einmalige 20 Basenpaare lange Markierung anstatt nur die chromosomale DNA zu mutieren. So kann das Gen, das für den Phänotyp einer bestimmten Mutante verantwortlich ist, schnell anhand dieser 20 Basenpaare zugeordnet werden. In den 50er Jahren führten Benzer et al. die Strategie der Komplementation ein, die für das geschilderte Problem einen allgemeinen Lösungsansatz bot.<sup>[64]</sup> Bei der genetischen Komplementation wird ein Pool von Plasmid-DNA, dessen Inserts aus chromosomaler Wildtyp-DNA stammt, in die mutierte Zelllinie eingeführt. Das DNA-Fragment, das die Mutation komplementiert, wird anhand der Wiederherstellung des Wildtyp-Phänotyps identifiziert. Im oben beschriebenen Beispiel von *Neurospora* würde Plasmid-DNA, die Fragmente des *Neurospora*-Wildtypgenoms trägt, in die *Neurospora*-Mutante eingeführt werden, die Defekte bei der Vitamin-B<sub>6</sub>-Biosynthese zeigte. Nur *Neurospora*-Mutanten mit einer Kopie des Wildtyp-Vitamin-B<sub>6</sub>-Biosynthesegens würde ohne Vitamin-B<sub>6</sub>-Supplementation wachsen, sodass Plasmid-DNA, die das Vitamin-B<sub>6</sub>-Biosynthesegen codiert, leicht erkannt werden könnte. Das Vorliegen verschiedener vollständiger Genomsequenzen erleichtert die genetische Komplementation noch weiter, weil die Plasmid-DNA-Sequenz mit der vollständigen Genomsequenz verglichen werden kann. Außerdem war es dadurch verschiedenen Arbeitsgruppen möglich, Bibliotheken mit Plasmiden herzustellen, die nicht einfach zufällige Fragmente chromosomaler DNA enthielten, sondern stattdessen ganz gezielt jeden offenen Leserahmen, der in einem gegebenen Organismus oder einer Zelllinie exprimiert wird, tragen.<sup>[65]</sup>

Eine wichtige Einschränkung der genetischen Analyse ist, dass sie nur solche natürlichen zellulären Aktivitäten erfasst, die durchgemustert werden können oder die man selektieren kann. Durch „Zwei-Hybrid“-Testsysteme können diese effizienten genetischen Tests auf viele Aktivitäten, die auf unterschiedlichsten chemischen Reaktionen beruhen, angewendet werden. Das Zwei-Hybrid-System, das ursprünglich

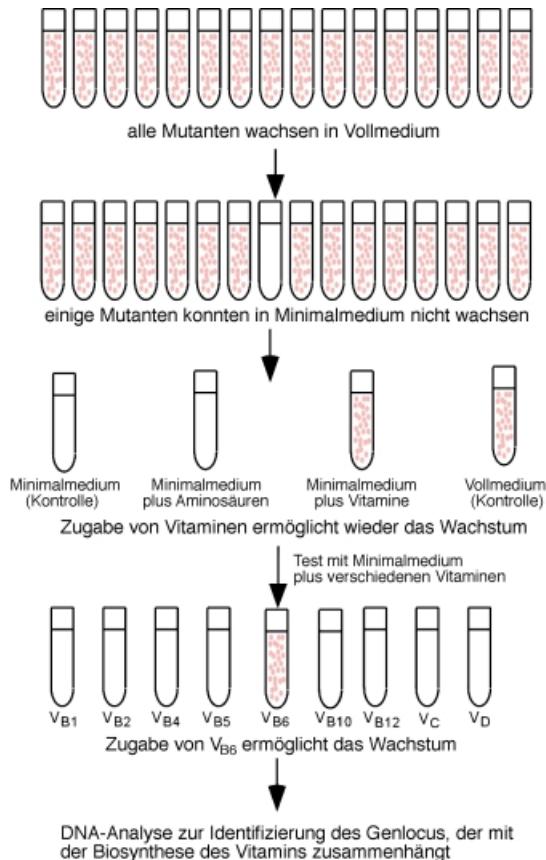


Abbildung 10. Genetische Methode zur Identifizierung von Genen, die für die Biosynthese essenzieller Metaboliten verantwortlich sind.<sup>[62]</sup>

ten *Neurospora crassa* durch Röntgenbestrahlung zufällige Mutationen. Etwa 2000 einzelne Mutantenstämme wurden in Voll- und Minimalmedium auf Wachstum getestet, wobei das Minimalmedium nur solche Nährstoffe (z. B. Biotin) enthielt, die *Neurospora* nicht selbst synthetisieren kann. Drei Mutantenstämme wurden identifiziert, deren Wachstumsrate sich in Vollmedium nicht von der des Wildtyps unterschied, die aber in Minimalmedium nur schlecht oder gar nicht wuchsen. Durch Zugabe einzelner Nährstoffe zum Medium und anschließendem Vergleich der Wachstumsraten konnte für die drei Stämme gezeigt werden, dass sie die Fähigkeit verloren hatten, Vitamin B<sub>6</sub>, Vitamin B<sub>1</sub> bzw. p-Aminobenzoësäure zu synthetisieren. Dieser grundlegende Ansatz, die DNA, die die Proteine codiert, zu mutieren und dann auf einen nachweisbaren Phänotyp wie das Zellwachstum zu testen, ist die Basis für die genetische Analyse. Die meisten Gene, deren Funktion bekannt ist, sind mit Methoden dieses Typs identifiziert

als Methode eingeführt worden war, um potenzielle Protein-Protein-Wechselwirkungen *in vivo* zu testen, wurde inzwischen angepasst, um Protein-DNA-, Protein-RNA-Wechselwirkungen sowie Wechselwirkungen zwischen Proteinen und niedermolekularen Stoffen und sogar katalytische Aktivitäten zu untersuchen.<sup>[66–72]</sup> Es beruht auf der Rekonstitution eukaryotischer Transkriptionsaktivatoren durch die Dimerisierung einer DNA-bindenden (DBD) und einer transkriptionsaktivierenden (AD) Domäne. Wenn die beiden interessierenden Proteine in Kontakt treten, wie dies in Abbildung 11 A dargestellt ist, dimerisieren sie DBD und AD auf effiziente Weise, veranlassen damit die Bildung der Transkriptionsmaschinerie und aktivieren die Transkription eines stromabwärts liegenden Reportergens. Der Test wurde erstmals an zwei Hefeproteinen demonstriert, der Serin-Threonin-Proteinkinase SNF1 und dem SNF1-Aktivatorprotein SNF4, von denen bekannt ist, dass sie *in vivo* assoziiert sind.<sup>[73]</sup> SNF1 wurde mit der GAL4-DNA-bindenden Domäne (DBD) fusioniert und SNF4 mit der GAL4-Aktivierungsdomäne (AD). Eine GAL4-Bindungssequenz wurde stromaufwärts von einem  $\beta$ -Galactosidase-Reportergen platziert. Plasmide, die die Proteinfusionen und das Reportergen

codierten, wurden in Hefe eingeführt, und das Syntheseniveau der  $\beta$ -Galactosidase wurde mit biochemischen Standardverfahren gemessen. Mit Kontrollexperimenten wurde gesichert, dass weder DBD noch AD für sich noch die einzelnen Fusionsproteine eine  $\beta$ -Galactosidase-Synthese oberhalb des Hintergrundniveaus induzierten. Die  $\beta$ -Galactosidase-Synthese wurde um das 200fache gesteigert, wenn die DBD-SNF1- und die SNF4-AD-Fusionsproteine gleichzeitig eingeführt wurden. Zum Vergleich: Das direkte DBD-AD-Fusionsprotein aktivierte die  $\beta$ -Galactosidase-Synthese um das 4000fache.

### 3.1.2. Protein-Protein-Wechselwirkungen

Die herkömmlichen genetischen Tests und das Zwei-Hybrid-System in Hefe wurden in erster Linie genutzt, um natürlich vorkommende Protein-Protein-Wechselwirkungen zu identifizieren. Eine neue und besonders eindrucksvolle Anwendung ist die automatisierte Identifizierung aller möglichen Protein-Protein-Wechselwirkungen in *S. cerevisiae*.<sup>[65]</sup> Jeder offene Leserahmen, der ein Protein codiert – in *S. cerevisiae* sind dies etwa 6000 – wurde sowohl mit der

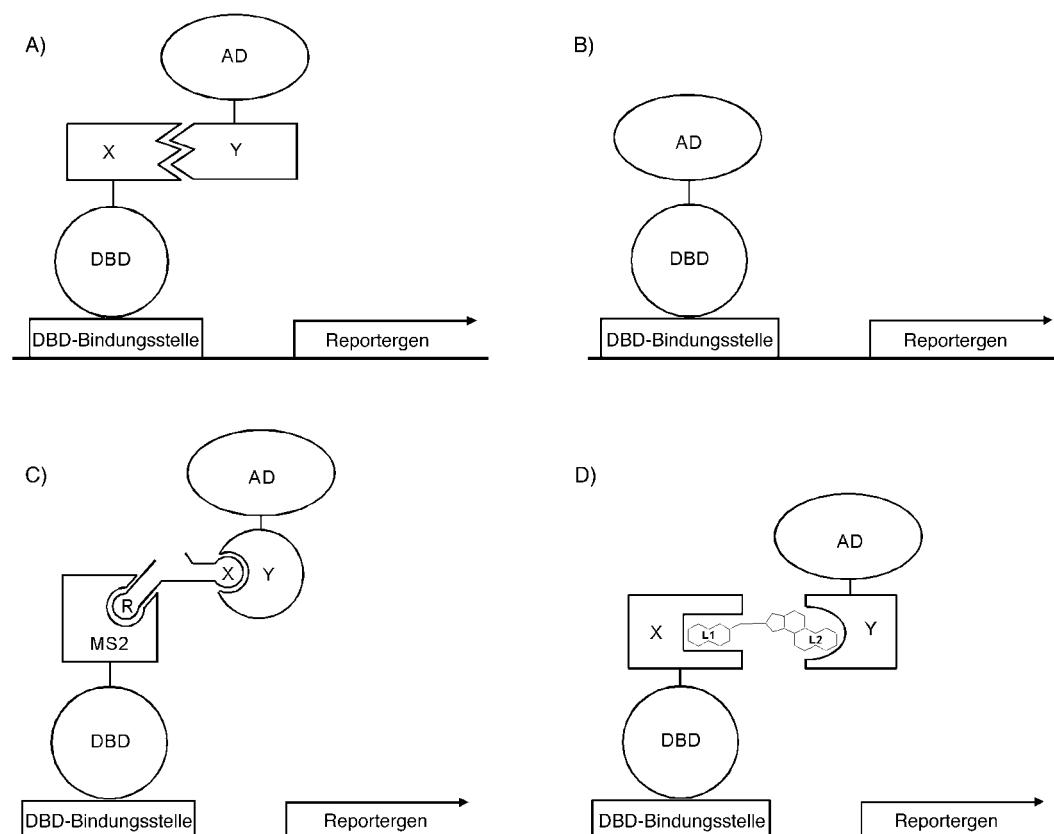


Abbildung 11. Verschiedene Hefe-*n*-Hybridsysteme, die entwickelt wurden, um Protein-Protein-, Protein-DNA-, Protein-RNA-Wechselwirkungen und Wechselwirkungen zwischen Proteinen und niedermolekularen Substanzen zu untersuchen. A) In der Originalversion des Hefe-Zwei-Hybrid-Systems wird die Transkriptionsaktivierung des Reportergens durch die Anlagerung der AD an die Promotorregion durch direkte Wechselwirkung der Proteine X und Y ausgelöst, da Protein X mit einer DBD und Protein Y mit der AD fusioniert ist. B) Im Ein-Hybrid-System ist die AD direkt mit der DBD fusioniert. Dieses System kann entweder genutzt werden, um DBDs zu testen, die an eine spezifische DNA-Sequenz binden können, oder die In-vivo-Bindungsstelle einer vorgegebenen DBD zu identifizieren. C) Im Vergleich zum Hefe-Zwei-Hybrid-System hat das Drei-Hybrid-System zum Nachweis von RNA-Protein-Wechselwirkungen eine weitere Komponente: ein hybrides RNA-Molekül. Eine Hälfte der Hybrid-RNA ist eine bekannte RNA (R), die an das MS2-Hüllprotein mit hoher Affinität bindet und als Anker dient. Die andere Hälfte ist RNA X, deren Wechselwirkung mit Protein Y getestet werden soll. D) Eine andere Version des Hefe-Drei-Hybrid-Systems dient dazu, die Wechselwirkung zwischen niedermolekularen Verbindungen und Proteinen nachzuweisen. Ligand L1, der mit Protein X wechselwirkt, ist kovalent mit Ligand L2 verbunden. Wenn dann L2 an Y bindet, wird die Transkriptionsaktivierung des Reportergens rekonstituiert.

Gal4-DBD als auch mit der Gal4-AD fusioniert, und dann wurden die beiden Fusionsbibliotheken gegeneinander getestet. Die größte Herausforderung war, alle Kombinationen der 6000 DBD- und der 6000 AD-Fusionen in Hefe zu transformieren und dann alle Zellen zu testen. Eine Bibliothek von  $10^7$  Elementen liegt an der Obergrenze der Transformationseffizienz bei Hefe und ist daher theoretisch erreichbar. Uetz und Glot et al. verglichen zwei Methoden: In einer kreuzten sie gezielt haploide MAT $\alpha$ -Zellen, die 192 DBD-Fusionsproteine enthielten, mit haploiden MAT $\alpha$ -Zellen, die 6000 AD-Fusionsproteine enthielten, in einem räumlich adressierbaren Format (Mikrotiterplatte) und testeten jede Vertiefung über eine *HIS3*-Wachstumsselektion. In einer zweiten Methode wurden MAT $\alpha$ -Zellen (mit den 6000 DBD-Fusionsproteinen) mit MAT $\alpha$ -Zellen (mit den 6000 AD-Fusionen) gekreuzt; nur die diploiden Zellen, die in einem *LEU2*-Selektionsmedium überlebten, wurden in einem Gitter ausgebracht und einzeln analysiert. Interessanterweise ergaben sich mehr „Treffer“ in dem räumlich adressierbaren Format, was bedeutet, dass die Bedeutung unterschätzt wird, die der Optimierung der Randbedingungen bei neuen Methoden für ein Hochdurchsatz-Screening und dem Problem, falsch Positive und falsch Negative in genetischen Tests zu erkennen, zukommt.

Wenn das Zwei-Hybrid-Verfahren auch in erster Linie zur Entdeckung natürlicher Protein-Protein-Wechselwirkungen diente, sollte es auch gut zur Entwicklung von Proteinen geeignet sein. Brent et al.<sup>[74]</sup> wiesen nach, dass das Zwei-Hybrid-System zur Identifizierung von Peptidaptameren aus einer Bibliothek zufälliger Peptidsequenzen dienen kann, die die Cyclin-abhängige Kinase 2 hemmen (Abbildung 12). Die Bibliothek aus Peptiden mit 20 Aminosäureresten wurde nicht einfach als Verlängerung von AD präsentiert, sondern im aktiven Zentrum des *E. coli*-Thioredoxin (TrxA). Diese Bibliothek in der TrxA-Schleife wurde mit der B42-AD fusioniert und das Cdk2-Gen mit der LexA-DBD. In einer einzigen Testrunde wurden  $6 \times 10^6$  B42-TrxA-Transformanten, also nur ein sehr kleiner Prozentsatz der  $10^{27}$  möglichen 20mere, auf Bindung an LexA-Cdk2 getestet. Aus diesem Ansatz wurden 66 Kolonien isoliert, die die Transkription von *LEU2* und dem Reportergen *lacZ* aktivierten. Diese Kolonien enthielten interessanterweise nur 14 verschiedene Peptidsequenzen, die Cdk2 mit hoher Affinität banden. Mit Oberflächen-Plasmonresonanz konnte bestimmt werden, dass die Peptidaptamere mit  $K_D$ -Werten von 30 bis 120 nm an Cdk2 binden (Tabelle 2). In Kinase-Hemmtests hatten die Peptidaptamere IC<sub>50</sub>-Werte für den Cdk2/Cyclin-E-Kinase-Komplex von 1 bis 100 nm. Überraschend war bei diesem Experiment, dass die Liganden mit nanomolaren Affinitäten in einer

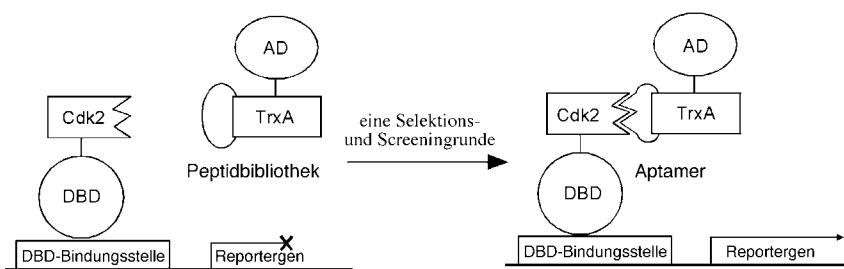


Abbildung 12. Mit dem Hefe-Zwei-Hybrid-System lassen sich Peptide entwickeln, die Zielproteine mit hoher Affinität binden.

Tabelle 2. Sequenzen und Bindungsaffinitäten der 14 Aptamere, die mit einem Hefe-Zwei-Hybrid-System isoliert wurden.<sup>[74]</sup>

Aptamer	$K_D$ [nm]	Aminosäure-Sequenz
pep1	n.b. <sup>[a]</sup>	ELRHRLGRAL SEDMVRLAW GPTSHCATVP GRSDLWRVIR FL
pep2	$64 \pm 16$	LVCKSYRLDW EAGALFRSLF
pep3	$112 \pm 1$	YRWQQGVVPS NMASCSFRCQ
pep4	n.b.	SSFLWLLMV KSIKRAAWEL GPSSAWNTSG WASLSDFY
pep5	$52 \pm 3$	SVRMRYGIDA FFDLGGLLHG
pep6	n.b.	RVKLGYSFWA QSSLRCISVG
pep7	n.b.	OLYAGCYLGIV VIASSLSIRV
pep8	$38 \pm 5$	YSFVHGHFFN FRVSWREMLA
pep9	n.b.	QQRVFSPSW FTCAGTSDFW GPEPLFDWTR D
pep10	$105 \pm 10$	QVWSWLALGW RWLRRYGWNM
pep11	$87 \pm 7$	WRRMELDAEI RWVKPISPLE
pep12	n.b.	RPLTGRWVVW GRRHEECGLT
pep13	n.b.	PVCCMMYGH TAPHSVFNVD
pep14	n.b.	WSPELLRAMV AFRWLLERRP

[a] n.b.: nicht bestimmt.

einzigsten Selektionsrunde isoliert wurden. Wie bei der Phagen-Display-Technik weiter oben beschrieben, lässt auch dieses Beispiel vermuten, dass Bibliotheksgrößen von  $10^6$  bis  $10^8$  Elementen ausreichen sollten, um die Proteinspezifität zu modifizieren. Im Unterschied zur Phagen-Display-Technik ist jedoch nur eine Selektionsrunde erforderlich, um echte hochaffine Liganden zu erkennen. Ähnliche Ergebnisse wurden mit Peptidaptameren in traditionellen genetischen Selektionsansätzen erhalten.<sup>[75]</sup>

### 3.1.3. Protein-DNA-Wechselwirkungen

Bereits früh wurde klar, dass so, wie das Zwei-Hybrid-System in Hefe zur Identifizierung von Protein-Protein-Wechselwirkungen dienen kann, Transkriptionsaktivatoren direkt in einem „Ein-Hybrid-System“ zur Entdeckung von DNA-Protein-Wechselwirkungen tauglich sein sollten (Abbildung 11B). In Wirklichkeit wurde dieser Typ von Versuchen bereits durchgeführt, bevor das Konzept des Ein-Hybrid-Systems als solches entwickelt war. So wurde zum Beispiel schon 1983 mit einem transkriptionsbasierten Selektionsverfahren eine His6 → Pro Mnt-Variante erzeugt, die bevorzugt einen mutierten Mnt-Operator bindet.<sup>[76]</sup> Etwas später passten Pabo et al.<sup>[77]</sup> ein bakterielles Zwei-Hybrid-System an, um Zink-Finger-Varianten mit definierten DNA-Bindungsspezifitäten zu entwickeln (Abbildung 13).<sup>[78]</sup> In diesem Testsystem werden die Protein-Protein-Wechselwirkungen auf der Basis der Dimerisierung eines DNA-bindenden Proteins und der α-Domäne der RNA-Polymerase

A)

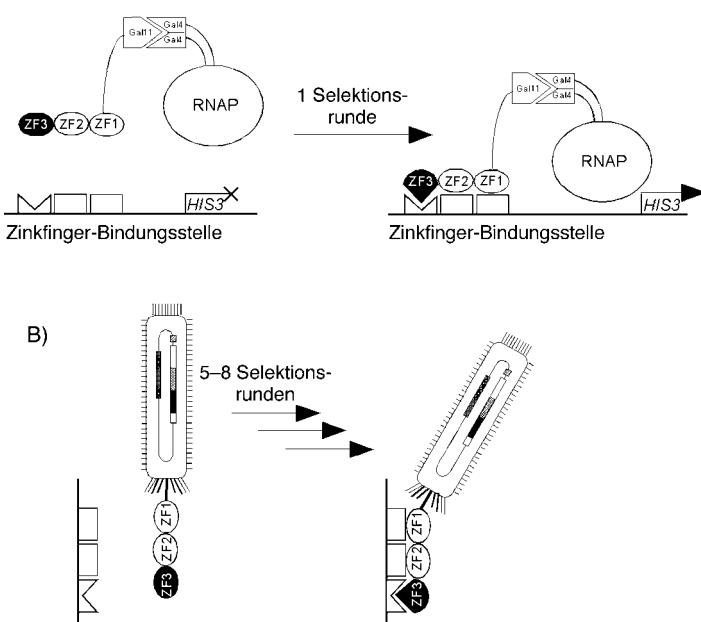


Abbildung 13. Die Entwicklung von Zinkfingern, die für die DNA-Sequenz AAA spezifisch sind, mithilfe eines Zwei-Hybrid-Testsystems<sup>[78]</sup> und einer Phagen-Display-Selektion.<sup>[26, 79]</sup> A) Das bakterielle RNA-Polymerasesystem (RNAP) wurde als Grundlage des Zwei-Hybrid-Systems genutzt. Zinkfinger 1, 2 und 3 aus dem Zif268-Protein wurden mit dem Gal11-Protein fusioniert. Das Gal4-Protein, das Gal11 mit hoher Affinität bindet, wurde mit der N-terminalen Domäne der  $\alpha$ -Untereinheit von RNAP fusioniert. An der Bindungsstelle für ZF3 wurde die Teilsequenz AAA erzeugt und sechs Reste in ZF3 wurden zufällig variiert. Falls ZF3 an die AAA-Position mit hoher Affinität bindet, wird der RNAP-Komplex herangeholt und aktiviert die Transkription des *HIS3*-Reportergens. In nur einer Selektionsrunde wurden mehrere Proteine identifiziert, die spezifisch an die Zielsequenz in der DNA binden. Diese Sequenzen können zwei charakteristisch unterschiedlichen Kategorien zugeordnet werden (Tabelle 3). B) Ein ähnliches Experiment wurde mit der Phagen-Display-Technik durchgeführt; dabei wurden ZF1, 2 und 3 des Zif268-Proteins auf der Phagenoberfläche als pIII-Fusionsprotein präsentiert.

(RNAP) und der dadurch erfolgenden Aktivierung der RNAP-abhängigen Transkription nachgewiesen. Um ein Messsignal für Protein-DNA-Wechselwirkungen zu erzeugen, wurden drei Tandem-Zinkfinger mit Gal11 fusioniert und Gal4, das hochaffin an Gal11 bindet, wurde mit der  $\alpha$ -Domäne der RNAP fusioniert. Außerdem wurde ein Reportersystem konstruiert, in dem, falls die Zinkfinger mit hoher Affinität an die gewünschte DNA-Sequenz binden, die Transkription eines *HIS3*-Reportergens aktiviert wird. Wie zuvor für die Peptidaptamere beschrieben, konnten hier Zinkfinger mit neuer DNA-Bindungsspezifität aus einer Bibliothek von etwa  $2 \times 10^8$  Zinkfinger-Varianten in einer einzigen Selektionsrunde isoliert werden. Die Autoren, die in der Vergangenheit routinemäßig die Phagen-Display-Technik verwendet haben,<sup>[28, 79]</sup> weisen darauf hin, dass die Selektion viel einfacher abläuft als bei der Phagen-Display-Technik, bei der mehrere Selektions- und Amplifizierungsrunden in ihrem Labor für das gleiche Evolutionsexperiment erforderlich waren (Abbildung 13, Tabelle 3).

### 3.1.4. Wechselwirkungen zwischen Proteinen und niedermolekularen Substanzen

Das Zwei-Hybrid-System wurde auf Wechselwirkungen zwischen Proteinen und niedermolekularen Systemen und zwischen Proteinen und RNA ausgedehnt, indem ein überbrückendes kleines Moleköl oder eine RNA mit eingebracht wurde und so ein „Hefe-Drei-Hybrid-System“ entstand (Abbildung 11C, D).<sup>[80–82]</sup> Die eine Hälfte des Brückenmoleküls wird als Anker genutzt, und die andere Hälfte des Moleküls ist der interessierende Molekölteil. Licitra und Liu bauten mit dem Steroid Dexamethason als Anker ein Hefe-Drei-Hybrid-System mit niedermolekularer Substanz auf und wiesen nach, dass dieses zur Suche nach der Zielstruktur eines Wirkstoffes in vivo dienen kann. Zum Nachweis der Machbarkeit zeigten sie, dass FKBP12 mit Dex-FK506 als überbrückendes kleines Moleköl aus einer Jurkat-cDNA-Bibliothek extrahiert werden konnte. Dex-FK506 dimerisierte spezifisch ein LexA-Glucocorticoid-Rezeptor(GR)-Fusionsprotein und eine cDNA-B42-AD-Bibliothek. Zuerst wurde mit einer *LEU2*-Wachstumsselektion nach Proteinen gesucht, die an FK506 binden. Dann wurden die Treffer bestätigt, indem die Transkriptionsniveaus eines *lacZ*-Reportergens mit und ohne das niedermolekulare Dex-FK506 verglichen wurden. In den Drei-Hybrid-Systemen ist das überbrückende Moleköl geeignet, um falsch-positive Treffer auszusondern. Bemerkenswert ist, dass nur FKBP12 identifiziert wurde, obwohl es in der Jurkat-cDNA-Bibliothek verschiedene FKBP12 ist. FKBP12 ist das FK506-bindende Protein mit der höchsten Affinität und bindet FK506 mit einem  $K_D$  von 0.4 nm. Der Schlüssel zum Erfolg eines Drei-Hybrid-Testsystems ist vermutlich der Anker. Anker mit hoher Affinität zur DBD-Rezeptor-Fusion werden benötigt, wenn Wechselwirkungen zwischen niedermolekularen Substanzen bzw. RNA und Proteinen über einen größeren Affinitätsbereich detektiert werden sollen. So haben wir in unserem Labor beispielsweise Methotrexat (Mtx), das Dihydrofolatreductase (DHFR) mit picomolarer Affinität bindet, für eine Anwendung in diesem System entwickelt.<sup>[83]</sup>

### 3.1.5. Enzymkatalyse

Genetische Methoden liefern offenbar besonders wirksame Such- und Selektionsverfahren für Enzymaktivitäten. Der Trick dabei ist, eine Aktivität zu finden, nach der man suchen oder selektieren kann, z. B. eine Reaktion mit einem fluoreszierenden Produkt. Arnold et al. stellten mit einem genetischen Screeningverfahren beispielsweise Cytochrom-P450-Varianten her, die Wasserstoffperoxid anstelle von molekularem Sauerstoff und NADH als Cosubstrat umsetzen (Abbildung 14).<sup>[84]</sup> Dazu wurden zufällige Mutationen in das Cytochrom-P450-Gen mit fehlertoleranter PCR eingeführt. Naphthalin wurde als Substrat in einem *E. coli*-Stamm genutzt, der so verändert war, dass er jedes hydroxylierte Naphthalinderivat oxidativ zu einem fluoreszierenden Dimer kupplte. Zellen, die P450-Varianten mit der gewünschten Aktivität enthielten, konnten durch digitale Fluoreszenzbildgebung identifiziert werden, wenn sie auf einer Petrischale mit naphthalinhaltigem Substrat wuchsen. Da ein

Tabelle 3. Vergleich der an die Target-DNA-Sequenz AAA bindenden Zinkfinger-Sequenzen, die mit einem bakteriellen Zwei-Hybrid-System und einer Phagen-Display-Methode isoliert wurden.<sup>[78, 79]</sup>

Bakterielles Zwei-Hybrid-System (1 Selektionsrunde)							Phagen-Display-Technik (5 – 8 Selektionsrunden)					
-1	1	2	3	5	6	-1	1	2	3	5	6	
Q	R	G	N	L	V	Q	R	T	N	I	T	
Q	K	T	N	M	V	Q	Q	H	N	K	L	
Q	K	Y	N	I	L	Q	R	N	N	L	L	
Q	R	Y	N	V	V	Q	A	N	N	R	T	
Q	K	G	N	M	V	Q	K	T	N	L	N	
Q	K	G	N	H	V	Q	H	G	N	V	A	
Q	K	G	N	M	V	Q	K	T	N	L	T	
Q	R	G	N	K	V	Q	K	T	N	D	T	
Q	R	G	N	K	T	Q	K	H	N	Q	V	
Q	L	G	N	M	V	Q	P	G	N	Q	T	
Q	K	G	N	K	V	Q	K	T	N	E	H	
Consensussequenz: <sup>[a]</sup>	Q	+	g	N	–	v	Q	–	–	N	–	–
N	S	G	A	Y	N	N	S	G	N	H	T	
N	S	G	A	W	N	N	S	G	A	S	N	
N	S	G	A	F	N	N	S	G	A	A	N	
N	S	G	T	H	N	N	S	G	A	T	N	
N	T	G	A	Y	N	N	S	G	A	T	N	
Consensussequenz: <sup>[a]</sup>	N	s	G	a	–	N	N	S	G	a	–	n
G	S	G	A	Y	N	A	T	G	A	H	N	

[a] Großbuchstaben bedeuten gemeinsame Aminosäuren in allen verglichenen Sequenzen, Kleinbuchstaben zeigen eine starke Präferenz für die jeweilige Aminosäure an; „+“ steht für eine bevorzugt positiv geladene Aminosäure und „–“ steht dort, wo keine bevorzugte Aminosäure bestimmt werden konnte.

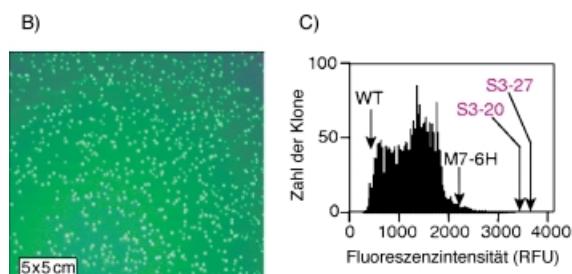
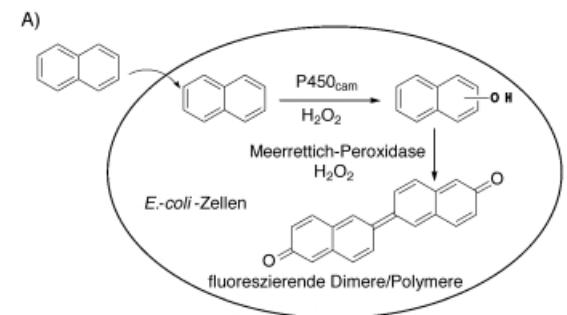


Abbildung 14. Genetisches Screening-Verfahren zur Entwicklung von Cytochrom-P450-Varianten, die Wasserstoffperoxid als Cosubstrat umsetzen.<sup>[84]</sup> A) Die Screening-Strategie: Das Screening beruht auf dem Nachweis fluoreszierender Produkte, die bei der ersten Hydroxylierung von Naphthalin durch P450<sub>cam</sub>-Varianten entstehen und dann mit Meerrettich-Peroxidase oxidativ gekuppelt werden. B) Digitale Fluoreszenzdarstellung von *E. coli*-Zellen, die P450<sub>cam</sub>-Varianten tragen. Zellen mit P450<sub>cam</sub>-Mutanten wurden auf Platten mit Naphthalin und Wasserstoffperoxid angezogen. Anschließend wurde die digitale Bildauswertung der Fluoreszenzintensitäten durchgeführt. Die Zellen, die eine aktive P450<sub>cam</sub>-Mutante enthielten und daher größere Fluoreszenzintensitäten aufwiesen, wurden dann analysiert. C) Histogramm, in dem die Fluoreszenzintensitäten von ca. 70000 Klonen der zweiten Generation von Mutanten aufgetragen sind. Die Fluoreszenzintensität des Wildtyps P450<sub>cam</sub>, die Mutation M7-6H aus der ersten Generation und die Mutanten S3-20 und S3-27 aus der zweiten Generation sind hervorgehoben. Wiedergabe mit Genehmigung nach Lit. [84].

Screening- statt eines Selektionsverfahrens angewendet wurde, war das Experiment auf eine kleinere Zahl von Proteinvarianten begrenzt (nämlich etwa 20000 pro Platte) als im Falle einer Selektion in Bezug auf das Wachstum. In der ersten Runde, in der 200000 Kolonien durchmustert wurden, zeigte eine große Zahl von Klonen eine erhöhte Fluoreszenz. Drei Mutanten aus der Gruppe von Kolonien mit der höchsten Fluoreszenz wurden charakterisiert und wiesen eine etwa zehnfach höhere Aktivität auf als das Wildtyp-P450-Cytochrom. In einer zweiten Runde wurden fünf aktive Mutanten über den Staggered-Extension-Process (StEP) rekombiniert, und die Nachkommen wurden durchmustert. Nun wurden wieder verschiedene Varianten isoliert, die etwa 20-mal aktiver waren als das Wildtypenzym. Man mag spekulieren, ob die Aktivität mit einer größeren Zahl von Mutantentests noch weiter verbessert werden können. Problematisch hätte dabei aber die Unterscheidung zwischen Proteinen mit einer kleinen Aktivitätserhöhung und solchen mit signifikanter Verbesserung werden können.

Nicht nur bei der genetischen Analyse, sondern auch bei der Entwicklung von Proteinen mit neuen Funktionen erscheint die Komplementation als naheliegende Wahl. Die Idee ist, dass man nicht Fragmente des Chromosoms auf Komplementation eines Mutantenphänotyps testet, sondern ein Protein mutiert und es weiterentwickelt, damit es einen vorgegebenen Phänotyp komplementiert.<sup>[85, 86]</sup> Yano et al. wendeten diese Strategie an, um eine Aspartat-Aminotransferase in eine Aminotransferase für verzweigte Aminosäuren umzuwandeln. Dieses Beispiel illustriert die Durchschlagskraft dieser Methode.<sup>[87]</sup> Zuerst konstruierten sie einen *E. coli*-Stamm mit einer Knock-out-Mutation im Wildtyp-Aminotransferase-Gen für verzweigte Aminosäuren und zeigten, dass dieser Stamm ohne Anreicherung des Mediums mit Val,

Ile und Leu nicht lebensfähig war. Das Wildtyp-Aspartat-Aminotransferase-Gen wurde mutiert, durch DNA-Shuffling rekombiniert und dann in den *E.-coli*-Selektionsstamm eingeführt. DNA-Shuffling, eine auf der PCR beruhende Methode, kann zur Einführung von Punktmutationen und zur Erzeugung rekombinierter Genprodukte dienen. Fünf Mutagenese- und Selektionsrunden wurden durchgeführt. Bei jeder Runde wurden  $10^6$ – $10^7$  Mutanten untersucht. Die Stringenz der Selektion konnte einfach gesteigert werden, indem entweder die Konzentration von 2-Oxoalatin im Wachstumsmedium, das Expressionsniveau des Enzyms oder die Inkubationszeit geändert wurde. Nach der letzten Runde wurde ein mutiertes Enzym mit 13

Punktmutationen isoliert, das einen  $10^5$ -fach höheren  $k_{cat}/K_M$ -Wert für  $\beta$ -verzweigte Aminosäuren und eine 30-fache Abnahme des  $k_{cat}/K_M$ -Wertes für Aspartat hatte. Die Verbesserung beim  $k_{cat}/K_M$ -Wert ist eindrucksvoll und spiegelt wahrscheinlich die Bedeutung wider, die die Modulation der Selektionsstringenz hat – zunächst um schlechte Katalysatoren zu entdecken und dann, um auf hocheffiziente Katalysatoren zu selektieren. Die Autoren gehen auch die Frage an, wie schwer es ist, die Spezifität zu verändern. Indem alle 13 Mutationen einzeln getestet wurden, zeigten sie, dass tatsächlich nur 6 Punktmutationen notwendig waren, um die Aktivität für das neue Substrat  $10^5$ -fach zu steigern. Interessanterweise zeigten die Ergebnisse auch, dass alle 6 Mutationen sich einzeln günstig auswirkten – so besteht die Möglichkeit, dass jede Position in einem Protein unabhängig variiert werden kann.

### 3.2. Liposomenbasierte Testsysteme

Für die natürliche Selektion, die treibende Kraft der Darwinschen Evolution, ist die Kompartimentierung der Zellen entscheidend. Tawfik und Griffiths<sup>[91]</sup> ahmten die zelluläre Kompartimentierung in einem einfacheren In-vitro-System, einer Wasser-in-Öl-Emulsion, nach. Durch die Zugabe einer In-vitro-Transkriptions-/Translations-Reaktionsmischung zu einer gerührten Suspension von Mineralöl mit Detergentien entstand eine Emulsion mit einem durchschnittlichen Tröpfchen-Durchmesser in der Größenordnung bakterieller Zellen. Innerhalb des wässrigen Kompartiments der Emulsion konnte dann DNA transkribiert und translatiert werden. Die Bedingungen wurden so kontrolliert, dass im Durchschnitt jedes Wasser-in-Öl-Tröpfchen nur ein einzelnes Gen enthielt. Eine Mischung von zwei Genen, dem *M.-HaeIII*-Gen, das die DNA-Methyltransferase HaeIII codiert, und dem *folA*-Gen, das die DHFR codiert, wurde in den wässrigen

Kompartimenten transkribiert und translatiert. Dann wurde die Emulsion entmischt, und die DNA in der wässrigen Phase wurde mit der HaeIII-Endonuklease geschnitten und dann mit PCR amplifiziert. Weil die DNA-Methyltransferase die HaeIII-Schnittstellen in den Tröpfchen, die das *M.-HaeIII*-Gen enthielten, methylierte, blieben nur diese *M.-HaeIII*-Gene unversehrt und wurden durch PCR amplifiziert. Mit einer Selektionsrunde wurde eine Mischung der Gene *M.-HaeIII* zu *folA* von 1:1000 angereichert bis zu einem Verhältnis von 1:1 (Abbildung 15).<sup>[91]</sup> Eine ähnliche Methode zur Selektion Nickel-bindender Proteine ist auch bekannt.<sup>[92]</sup>

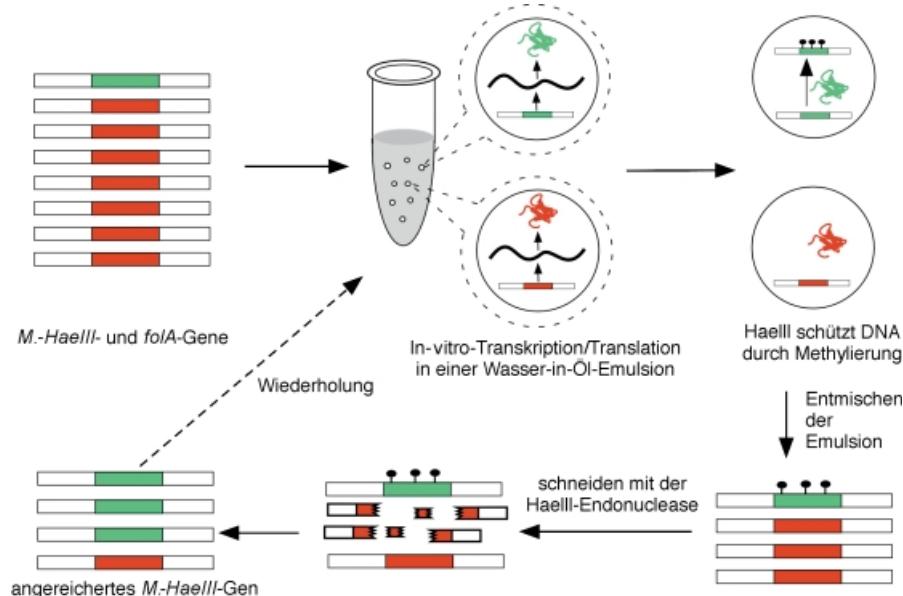


Abbildung 15. In-vitro-Kompartimentierung.<sup>[91]</sup> Eine Mischung von zwei Genen, *M.-HaeIII* und *folA*, wurde im wässrigen Kompartiment einer Wasser-in-Öl-Emulsion transkribiert und translatiert. Das *M.-HaeIII*-Genprodukt, die DNA-Methyltransferase HaeIII, methylierte die DNA und verhinderte die Spaltung durch die HaeIII-Endonuklease. Nach Trennung der Emulsion und Spaltung der DNA mit der HaeIII-Endonuklease blieb nur das *M.-HaeIII*-Gen unversehrt, das dann mit PCR amplifiziert und so selektiv angereichert wurde.

Da alle Schritte der In-vitro-Kompartimentierung tatsächlich in vitro ablaufen, hat diese Methode das Potenzial, bis an die Grenze vorzustoßen, an der die DNA begrenzend für die Bibliotheksgröße wird. Im ersten Artikel von Tawfik und Griffiths wird beschrieben, dass  $50 \mu\text{L}$  eines In-vitro-Transkriptions-/Translations-Reaktionsgemisches zu  $950 \mu\text{L}$  Mineralöl mit Detergentien gegeben wurden. Der durchschnittliche Durchmesser der gebildeten Tröpfchen lag bei  $2.6 \mu\text{m}$ . Die Emulsion enthielt also etwa  $10^{10}$  Tröpfchen pro mL, was bedeutet, dass etwa  $10^{10}$  einzelne Proteine bei diesem Volumen getestet werden können. Die In-vitro-Kompartimentierung liefert im Unterschied zu Methoden wie der RNA-Display-Technik ein geeignetes Format, um katalytische Reaktionen mit größerer Wechselzahl nachzuweisen. Bislang sind die Tests auf Enzyme zur Modifizierung oder Amplifizierung von DNA beschränkt.<sup>[93]</sup> Die Herausforderung wird sein, die Methode auf chemische Reaktionen auszudehnen, die nichts mit DNA zu tun haben.

### 3.3. Zusammenfassung

Genetische Komplementation und In-vitro-Kompartimentierung sind gegenwärtig die wirkungsvollsten Testsysteme für Enzymkatalyse. Die Zahl der Proteinevarianten, die mit genetischen Tests wie der Phagen-Display-Technik untersucht werden kann, ist durch die Effizienz der DNA-Transformation begrenzt. Bei der In-vitro-Kompartimentierung wie der RNA-Display-Technik dagegen ist lediglich die DNA limitierend. Bibliotheksgrößen von etwa  $10^6$  bis  $10^8$  Elementen erscheinen auch hier ausreichend, um Enzyme mit neuen Funktionen zu entwickeln. Es gibt verschiedene eindrucksvolle Anwendungsbeispiele für die Modifizierung von Enzymspezifitäten mit In-vivo-Komplementation. Der Vorteil von In-vivo-Tests liegt besonders bei cDNA-Bibliotheken darin, dass das Protein in seiner natürlichen Umgebung exprimiert wird. Während sie ursprünglich zur Komplementation der natürlichen Enzymfunktion eingesetzt wurden, konnte mithilfe der Ein-, Zwei- und Drei-Hybrid-Systeme eine Ausweitung der Tests auf Protein-DNA-, Protein-Protein- und Protein-RNA-Wechselwirkungen sowie auf Wechselwirkungen zwischen Proteinen und niedermolekularen Substanzen erreicht werden. Die nächste Aufgabe ist, diese Tests auch an Reaktionen anzupassen, die nicht in der Zelle vorkommen. Eine neue Entwicklung ist die In-vitro-Kompartimentierung. Diese Methode wird so adaptiert werden müssen, dass sie andere Messgrößen ablesen kann als die der DNA-Chemie. Genetische Tests sind technisch geradlinig, weil sie auf den Fortschritten der Molekularbiologie aufbauen. Die DNA aus positiv getesteten Zellen kann einfach extrahiert werden und weiteren Mutageneserunden oder der Sequenzierung unterworfen werden. Das Hauptproblem bei genetischen Tests ist, falsch positive und falsch negative Treffer auszusondern. Entscheidend dafür ist ein aussagekräftiges sekundäres Testsystem. Die In-vitro-Kompartimentierung ähnelt technisch gesehen dem RNA-Display. Offen ist aber noch, als wie robust sich die Emulsionstechnik erweisen wird.

## 4. Methoden auf der Basis räumlicher Zuordnungen

Methoden, die auf räumlichen Zuordnungen beruhen, verbinden die Identität eines Peptids oder eines Proteins mit einer eindeutigen räumlichen Position wie der Vertiefung einer Mikrotiterplatte oder einer Stelle auf einem festen Träger, sodass jeder Proteintreffer zur zugehörigen DNA-Sequenz zurückverfolgt werden kann.<sup>[94–97]</sup> Mikrotiter- oder „96-well“-Platten werden schon lange für diesen Zweck genutzt. In jüngerer Zeit wurden auch Fortschritte bei der Derivatisierung fester Träger mit Proteinen, so genannten Proteinchips, gemacht.

### 4.1. Mikrotiterplatten-Tests

96er Mikrotiterplatten werden in der biochemischen Forschung seit den 50er Jahren genutzt.<sup>[98]</sup> Routinemäßig fanden sie bei ELISA-Tests (*enzyme-linked immunosorbent assay*)

Anwendung, seit diese in den 70er Jahren erstmals eingeführt wurden.<sup>[99]</sup> ELISA ist eine empfindliche Methode, bei der die Antikörperbindung eines Antigens über die katalytische Aktivität eines Enzyms nachgewiesen wird, das mit dem Antikörper fusioniert ist. Als Antigen wird die zu untersuchende Substanz bezeichnet, mit der zuvor Mäuse immunisiert wurden. Das Antigen wird an die Vertiefungen einer 96er Mikrotiterplatte adsorbiert und dann mit dem enzymgebundenen Antikörper inkubiert. Ungebundene Antikörper werden wegwaschen und gebundene Antikörper nachgewiesen, indem die enzymkatalysierte Farbreaktion verfolgt wird. Das Durchmustern von Peptid- oder Proteinbibliotheken anhand einer biochemischen Funktion in einer Mikrotiterplatte ist der Durchführung eines ELISA-Tests sehr ähnlich. Peptid- oder Proteinlösungen können auf die Vertiefungen einer Mikrotiterplatte verteilt und geeignete biochemische Tests dann in jeder Vertiefung durchgeführt werden.

#### 4.1.1. Technische Betrachtungen

In Mikrotiterplatten-Tests müssen Tausende von Proteinproben, seien es Zellkulturen, Rohextrakte oder gereinigte Proteine, zunächst vorbereitet und dann in Tausende von Vertiefungen der Mikrotiterplatten überführt werden. Anschließend können biochemische Tests in jeder Vertiefung durchgeführt werden, um die Proteine mit der gewünschten Funktion zu identifizieren. Daher ist eines der Hauptprobleme, Tausende von Proteinen schnell und kostengünstig zu reinigen. Wenn die gewünschte Aktivität im Rohextrakt der Zellen getestet werden kann, reduziert sich dies auf eine technische Frage. Die Zahl der Aktivitäten, die getestet werden könnten, wäre jedoch erheblich größer, wenn Methoden zur Massenreinigung von Proteinen zur Verfügung ständen. Martzen et al. lösten dieses Problem, indem sie die Proteine als GST-Fusionsproteine reinigten. In diesem speziellen Fall reinigten sie alle 6000 offenen Leserahmen (ORFs), die in der Hefe *S. cerevisiae* exprimiert werden, als GST-Fusions-Proteine und durchmusterten diese ORF-Bibliotheken nach verschiedenen enzymatischen Aktivitäten.<sup>[100]</sup> Der Trick ist die Art, wie die Bibliothek aus GST-Fusions-Proteinen konstruiert wurde. Die ORFs wurden in einem zweistufigen PCR-Verfahren amplifiziert, sodass im zweiten PCR-Schritt alle ORFs mit einem universellen Primer amplifiziert werden konnten. Dadurch kann die ORF-Bibliothek anschließend einfach in ein neues Expressionsformat kloniert werden. In diesem Beispiel wurden die ORFs unter die Kontrolle eines kupferinduzierbaren Promoters gebracht. Die entstehenden 6144 Hefestämme, von denen jedes ein Hefe-ORF als GST-Fusions-Protein enthielt, wurden in 64 96er Mikrotiterplatten übertragen. Zunächst wurden die Hefestämme in 64 Pools, die alle je 96 Vertiefungen einer Mikrotiterplatte enthielten, angezüchtet, die GST-ORF-Fusionsproteine wurden gereinigt und in den 64 Pools getestet. Dann wurde der positive Pool aufgelöst, indem der gleiche Test mit jeder Reihe und jeder Spalte der 96er Platte als einem Pool durchgeführt wurde. So konnte das aktive ORF-GST-Fusionprotein identifiziert werden. Getestet wurden die Proteine auf Aktivitäten für tRNA-Ligase, 2'-Phosphotrans-

ferase und cyclische Phosphodiesterase mit dünnenschichtchromatographischen Tests mit radioaktiven Substraten. So wurden fünf Enzyme kloniert, von denen drei vorher noch unbekannt waren.

Es sei an dieser Stelle noch einmal daran erinnert, dass durch die Mikrotiterplatten zwar ein Protein mit seiner DNA-Sequenz verknüpft werden kann, dass die Platten allein jedoch nicht für Screening- und Selektionsverfahren geeignet sind. Der Vorteil besteht allerdings darin, dass viele traditionelle biochemische Tests einfach an dieses Format angepasst und automatisiert werden können.<sup>[101–103]</sup> Verbreitet sind SPA-Verfahren (scintillation proximity assays),<sup>[104]</sup> UV/Vis-Absorptionsmessungen und Chemolumineszenz- und Fluoreszenztests.<sup>[105]</sup> Fluoreszenztests sind besonders brauchbar und Variationen wie FRET und Fluoreszenzpolarisation werden allgemein genutzt. Der Trend bei Mikrotiterplatten-Tests geht zur Miniaturisierung, und hochdichte Mikrotiterplatten mit 384 oder 1536 Vertiefungen sind inzwischen Standard. Wegen ihrer hohen Empfindlichkeit und der Geschwindigkeit gewinnen Fluoreszenzmessungen immer mehr Bedeutung bei miniaturisierten Hochdurchsatz-Testsystemen. Eigen et al.<sup>[106]</sup> wandten die Zwei-Farben-Fluoreszenz-Korrelationsspektroskopie an, um katalytischen Umsatz durch Restriktionsendonukleasen nachzuweisen. Sie nannten die Methode RAPID-FCS (*rapid assay processing by integration of dual-color FCS*). Ein 66 bp langes Oligonucleotid mit Schnittstellen für Endonukleasen wurde an einem Ende mit dem Fluoreszenzfarbstoff Cy5, am anderen Ende mit Rhodamingrün markiert. Mit Kreuzkorrelationsanalysen wurden gespaltene und ungespaltene Oligonucleotidproben unterschieden. Eine Analysezeit von ca. 1 Sekunde reichte aus, um unterscheidbare Kreuzkorrelationswerte zu erhalten. Für die Messung genügte ein Probenvolumen von wenigen Mikrolitern. Mit dieser kurzen Analysezeit und dem geringen Probenvolumen ist die RAPID-FCS gut geeignet für ein Hochdurchsatz-Screening.

#### 4.1.2. Enzymkatalyse

Testsysteme in Mikrotiterplatten sind besonders gut für Messungen der enzymatischen Aktivität geeignet, weil jede Reaktion räumlich isoliert abläuft. Hilvert et al. entwickelten mit einem Sekundär-Screening in einer 96er Platte einen Antikörper, der eine Umlagerungsreaktion katalysiert. Sie durchmustereten Tausende monoklonaler Antikörper gegen ein Analogon des Übergangszustandes und konnten einen Antikörper isolieren, der den Zerfall von substituierten Benzisoxazolen zu 2-Cyanphenolen um den Faktor  $10^8$  beschleunigte.<sup>[107]</sup> Zunächst wurden dazu monoklonale Antikörper mit der Standard-Hybridoma-Technik gegen das Analogon des Übergangszustandes erzeugt. Dann wurden etwa 1000 Antikörper in 96er Mikrotiterplatten überführt und anhand der Absorptionszunahme bei 340 nm auf Grund der Bildung des Produktes 2-Cyanphenol nach der gewünschten Aktivität durchmustert (Abbildung 16). Zwei dieser Antikörper, die Geschwindigkeitssteigerungen um  $3.4 \times 10^8$  bzw.  $3.5 \times 10^7$  im Vergleich zur unkatalysierten Reaktion hervorriefen, wurden weiter analysiert (Tabelle 4). Normalerweise werden nur 10–100 monoklonale Antikörper nach

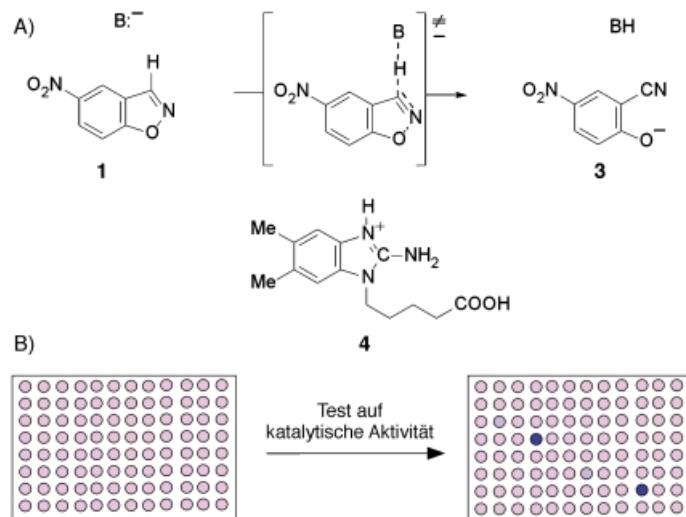


Abbildung 16. Erzeugung hoch effizienter katalytischer Antikörper mit einem Übergangszustandsanalogon und einem Screening auf Mikrotiterplatten.<sup>[107]</sup> A) Struktur des Übergangszustandsanalogons 4, mit dem die katalytischen Antikörper gewonnen wurden, die die Transformation von 1 zu 3 katalysieren. B) Die hergestellten Antikörper wurden in eine Mikrotiterplatte überführt und dort auf katalytische Aktivität getestet, indem die Absorptionsänderung bei 340 nm bei der Umwandlung von 1 in 3 verfolgt wurde. Zwei hochaktive Antikörper wurden isoliert. Wiedergabe mit Genehmigung nach Lit. [107]

Tabelle 4. Die kinetischen Parameter der katalytischen Antikörper, die von Hilvert et al.<sup>[107]</sup> hergestellt wurden.

Katalysator	$k_{\text{cat}}$ [s <sup>-1</sup> ]	$K_m$ [μM]	$k_{\text{cat}}/K_m$ [s <sup>-1</sup> M <sup>-1</sup> ]	$\frac{k_{\text{cat}}/K_m}{k_{\text{AcO}^-}}$	pKa	$k_{\text{cat}}/k_{\text{AcO}^-}$ [M]
34E4	0.66	120	$5.5 \times 10^3$	$3.4 \times 10^8$	6.0	41 000
35F10	0.35	630	$5.6 \times 10^2$	$3.5 \times 10^7$	5.5	22 000

der gesuchten Aktivität durchmustert und dabei jeder Antikörper in einem separaten kinetischen Test analysiert. Die  $10^8$ fache Geschwindigkeitszunahme, die in diesem Fall erzielt wurde, ist viel größer als das, was im Allgemeinen mit katalytischen Antikörpern erreicht wird. Das ist ein Beleg für den Nutzen eines solch umfangreichen sekundären Screenings und der direkten Suche nach katalytischer Aktivität.

Mikrotiterplatten-Tests wurden nicht nur zum Design von Antikörpern genutzt, sondern auch zum Protein-Design allgemein. Giver et al. durchmusterten beispielsweise eine Bibliothek mit mutagenisierten p-Nitrobenzylesterasen aus *Bacillus subtilis* (pNBE), um hitzestabile pNBES zu entwickeln und um die Beziehung zwischen Thermostabilität und katalytischer Aktivität aufzuklären.<sup>[108]</sup> Die Proteinbibliothek (Abbildung 17) wurde durch mutagenisierende PCR und durch DNA-Shuffling hergestellt und dann in *E. coli* exprimiert. Von jeder Generation wurden 500–2000 Kolonien ausgesucht und in 96er Mikrotiterplatten überführt. pNBE-Aktivität und Thermostabilität der Proteine wurden mit einem Standard-p-Nitrobenzylacetat-Test gemessen, in dem der Enzymdurchsatz anhand der Bildung eines farbigen Produktes nach Zusatz von Zellrohextrakt bestimmt wird. Sechs Mutagenese- und Screeningrunden stabilisierten die pNBE deutlich mit einem Anstieg von  $T_m$  um  $14^\circ\text{C}$ . Interessanterweise änderte sich dabei die Enzymaktivität

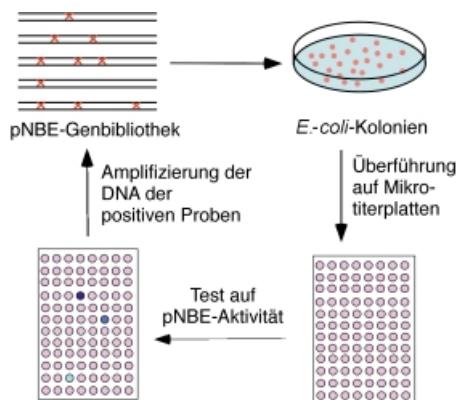


Abbildung 17. Evolution einer thermostabilen pNBE.<sup>[108]</sup> Die Proteinbibliothek, die aus dem Wildtyp-pNBE-Gen durch mutagenisierende PCR und DNA-Shuffling gewonnen wurde, wurde in *E. coli* exprimiert. Einzelne Kolonien wurden ausgewählt und in Mikrotiterplatten überführt, wo die Aktivität und die Thermostabilität der Proteine mit einem Standard-*p*-Nitrobenzylacetat-Test bestimmt wurde. Die DNA aus Klonen mit erhöhter Thermostabilität wurde isoliert und weiteren Mutagenese- und Screeningrunden unterworfen. Nach sechs Runden war der  $T_m$ -Wert des Enzyms um 14 °C gestiegen, während die Aktivität bei Raumtemperatur unverändert war.

der Mutante bei Raumtemperatur nicht. Diese Anwendung war wahrscheinlich erfolgreich, weil eine überschaubare Problemstellung – Steigerung der Thermostabilität eines bekannten Enzyms – gewählt wurde, die mit einer Bibliothek mit etwa  $10^4$  Elementen bearbeitet werden konnte und weil ein einfacher Zelltest zur Messung der Aktivität der Esterase verfügbar war.

Bibliotheken mit räumlich zuordbaren Elementen scheinen besonders gut geeignet zu sein für das Screening von cDNA-Bibliotheken, bei denen weniger Proteine getestet werden müssen als beim Entwickeln eines Proteins mit einer neuen Aktivität. Haushalter et al. konnten mit In-vitro-Expressionsklonierung neue DNA-Glycosylasen aus einer *Xenopus*-Embryo-cDNA-Bibliothek identifizieren.<sup>[109]</sup> Die cDNA-Bibliothek wurde in 120 Pools aufgeteilt, von denen jeder ungefähr 100 cDNAs enthielt. Diese cDNA-Pools wurden in vitro in den Vertiefungen von Mikrotiterplatten transkribiert und translatiert. Die translatierten Proteine wurden mit  $^{32}\text{P}$ -markierten, Mechanismus-basierten Glycosylase-Inhibitoren inkubiert und dann durch native Polyacrylamid-Gelelektrophorese aufgetrennt. Der Pool mit deutlich sichtbarer Inhibitor-Bindung wurde schrittweise weiter aufgetrennt und getestet, bis ein einzelner aktiver Klon identifiziert war. Dieser Klon wurde charakterisiert und dabei nachgewiesen, dass er eine neue Einzelstrang-selektive monofunktionale Uracil-DNA-Glycosylase-Aktivität aufweist.

## 4.2. Proteinchips

Ermutigt durch die Erfolge mit DNA-Mikrochips richten sich die Anstrengungen seit einiger Zeit darauf, Protein-Mikrochips zu entwickeln. Auf diesen sind Proteine auf einem festen Träger in hoher Dichte immobilisiert, wodurch Proteinfunktionen in Größenordnungen des Genoms untersucht werden können. Proteine wurden beispielsweise auf Poly-

vinylidendifluoridmembran,<sup>[110]</sup> auf Nitrocellulosemembran,<sup>[111]</sup> auf Glasträgern<sup>[112, 113]</sup> und in Polyacrylamidgel<sup>[114]</sup> immobilisiert. Macbeath und Schreiber<sup>[112]</sup> fixierten mit einem Präzisionsroboter verschiedene Proteine durch Aufdrucken kovalent auf einem Glasträger in einem hochdichten Format. Die Glasträger wurden zuerst mit einem Aldehyd-haltigen Silan-Reagens derivatisiert. Dann wurden Proteinproben in wenigen Nanolitern Volumen auf die Träger gedruckt. Die Aldehydgruppen auf dem Glas reagierten mit primären Aminen der Proteine unter Bildung von Schiff-Basen und fixierten so die Proteine kovalent auf der festen Oberfläche. Damit kann eine Dichte von 1600 Proteinspots pro Quadratzentimeter erreicht werden. Die Brauchbarkeit dieser Proteinchipps wurde auch bei der Suche nach bekannten Wechselwirkungen zwischen Proteinen und zwischen Proteinen und niedermolekularen Substanzen und der Identifizierung von Proteinsubstraten von Proteinkinasen nachgewiesen.

### 4.2.1. Technische Betrachtungen

Im Unterschied zu Tests auf Mikrotiterplatten, in denen rohe Zellsäfte eingesetzt werden können, müssen die Proteine für Tests mit Protein-Mikrochips gereinigt und dann auf die Chips aufgedruckt werden. Daher ist die Reinigung Tausender von Proteinen der entscheidende limitierende Schritt dieser Methode. Verglichen mit den Mikrotiterplatten-Systemen erreicht man mit Proteinchipps einen höheren Durchsatz. Außerdem werden nur Nanoliter der Lösung gereinigten Proteins für jeden Proteinspot auf dem Chip benötigt. Ein anderer Vorteil von Proteinchipps gegenüber Mikrotiterplatten ist, dass Waschschritte die Lokalisierung des Proteins nicht zerstören. Andererseits eignen sich immobilisierte Proteine, wie in diesen Untersuchungen zum Nachweis des Prinzips gezeigt wurde,<sup>[111, 114, 115]</sup> besser für Bindungs- als für Enzymtests. Eine geschickte Lösung, die die Vorteile beider Methoden verbindet, ist die Immobilisierung von Proteinen in Nanovertiefungen.<sup>[116]</sup> Ein besonders interessanter Ansatz aus neuerer Zeit ist die von Zauddin und Sabatini durchgeführte Immobilisierung von cDNAs auf festem Träger und die anschließende Transfektion von Zellen in großem Umfang.<sup>[117]</sup>

### 4.2.2. Anwendungen

Bislang wurden meist Methoden entwickelt, um einige Proteine auf einen festen Träger aufzubringen. Ein eindrucks voller Fortschritt gelang vor kurzen mit dem Druck und der Anwendung eines Chips, der das gesamte Hefeproteom präsentierte. Zhu et al.<sup>[118]</sup> klonierten 5800 offene Leserahmen von Hefe in einen Hefeexpressionsvektor mit hoher Kopienzahl in Form von GST-His6-Fusionsproteinen unter Kontrolle des *GAL1*-Promotors. Die Proteine wurden mit Glutathion-Agarose-Kügelchen in 96er Mikrotiterplatten gereinigt. Anschließend wurden sie auf Glasträger gedruckt, die entweder mit einem Aldehyd oder mit Nickel beschichtet waren. Die Aldehydgruppen reagieren mit den primären Aminen am Protein, während Nickel am Proteinende Komplexe mit dem His6-Tag bildet, deshalb werden die Proteine

an der Glasoberfläche fixiert. Der Nachweis des GST-Tags auf dem Proteinchip zeigte, dass mehr als 90 % der Proteine exprimiert, gereinigt und aufgedruckt worden waren. Calmodulin-bindende Proteine wurden nach Inkubation des Proteinchips mit biotinyliertem Calmodulin durch den Nachweis von gebundenem biotinyliertem Calmodulin mit Cy3-konjugiertem Streptavidin nachgewiesen, nachdem nicht-gebundenes biotinyliertes Calmodulin weg gewaschen worden war (Abbildung 18). Mit diesem Experiment wurden sechs be-

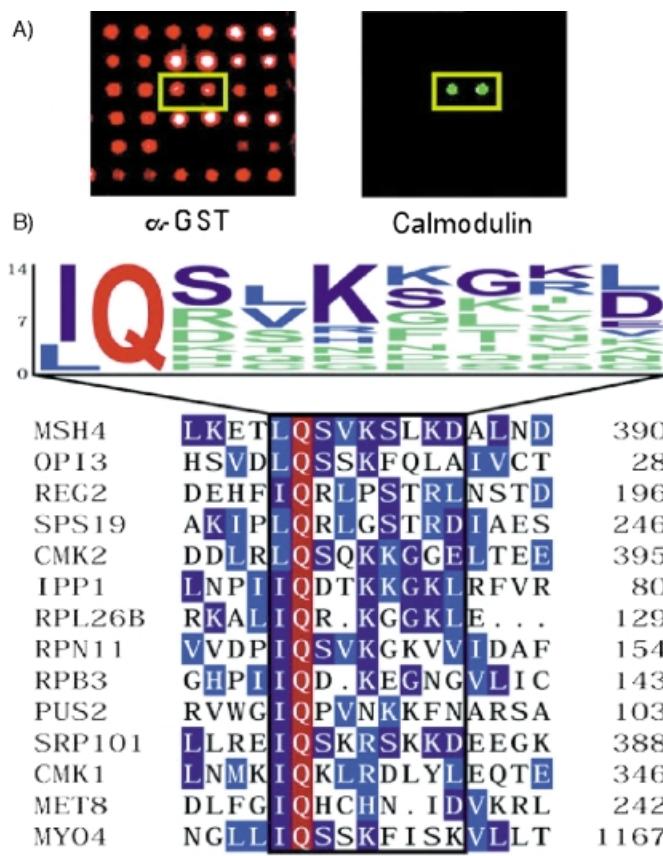


Abbildung 18. Test auf Calmodulin bindende Proteine auf einem Proteinchip.<sup>[118]</sup> A) Links ist ein Kontroll-Proteomchip abgebildet, der mit anti-GST (rot) behandelt wurde. Dieses Kontrollexperiment zeigt, dass mehr als 90 % der Proteine erfolgreich exprimiert, gereinigt und auf die Glaspfälzchen übertragen worden waren. Auf dem Chipausschnitt rechts sind positive Signale (in grün) zu sehen, die mit einer biotinylierten Calmodulinprobe entdeckt wurden (Doppelbestimmung). B) 39 positive Proteine wurden identifiziert, von denen 14 ein Motiv gemeinsam haben, dessen Konsensussequenz lautet (I/O)QXK(K/X)GB. X steht für einen beliebigen Aminosäurerest, B für einen basischen Rest. Die Buchstabengröße ist ein Maß für die relative Häufigkeit der markierten Aminosäure. Wiedergabe mit Genehmigung nach Lit. [118].

kannte und 33 unbekannte potenzielle Calmodulin-bindende Proteine identifiziert. Sechs bekannte Calmodulin-bindende Proteine wurden nicht gefunden. Auf ähnliche Weise wurden auch Phospholipid bindende Proteine nachgewiesen. Wichtig ist, dass die Autoren durch die Bestimmung der GST-Markierung und der Calmodulin-bindenden Protein bestätigten, dass der Chip erfolgreich einen hohen Prozentsatz der cDNA-Bibliothek präsentierte.

### 4.3. Zusammenfassung

Der Vorteil von Verfahren mit räumlich adressierbaren Bibliotheken scheint zu sein, dass sie vergleichsweise einfach an traditionelle biochemische Tests für die Proteinfunktion anzupassen sind. Vor allem bei Anwendungen, bei denen die Proteine gereinigt werden müssen, sind räumlich adressierbare Bibliotheken realistischerweise auf etwa 10<sup>4</sup> Proteinvarianten beschränkt. Es gibt zwar Anwendungen beim Protein-Engineering, doch scheinen Methoden mit räumlich adressierbaren Bibliotheken am besten geeignet, um cDNA-Bibliotheken für die Bestimmung der Proteinfunktion zu testen. Wie alle In-vitro-Methoden sind diese Tests auf Proteine beschränkt, die außerhalb der Zelle stabil sind. Außerdem ist noch unklar, wie der feste Träger möglicherweise die Proteinfunktion beeinflusst. Eine große Zahl von chemischen Reaktionen kann mit diesen Methoden nachgewiesen werden, weil man viele traditionelle Testsysteme verwenden kann. Proteine, die in die Vertiefungen von Mikrotiterplatten fixiert wurden, sind für Tests bezüglich der Bindung und der Enzymaktivität verwendbar. Der Nachteil räumlich adressierbarer Methoden ist, dass sie technisch aufwändig sind. Reinigen und Aufbringen der Proteine ist sehr arbeitsintensiv. Außerdem muss man davon ausgehen, dass Proteinchips bei verschiedenen Tests nicht so robust sind wie DNA-Chips. Auch muss am Ende des Tests jedes Protein zu seiner spezifischen DNA zurückverfolgt werden.

### 5. Direkte Proteinsequenzierung

Eine andere Möglichkeit ist, dass die Sequenzierungsverfahren von Proteinen bis zu einem Punkt vorangetrieben werden, an dem eine Kupplung mit der DNA unnötig ist. In diesem Abschnitt werden wir die jüngsten Fortschritte der Massenspektrometrie betrachten, die in die Nähe dieses Ziels kommen.

#### 5.1. Technische Betrachtungen

Die Massenspektrometrie (MS) entwickelt sich zu einem wichtigen Werkzeug für die Proteomik-Forschung. Treibende Kraft dabei sind die Fortschritte von MS und Genomsequenzierung. Zu den Fortschritten der MS gehören die Entwicklung neuer Ionisationstechniken wie die Electrospray-Ionisation (ESI) und die matrixgestützte Laserdesorption und -ionisation (MALDI).<sup>[119-121]</sup> MS kann zur Bestimmung der Proteinsequenz eingesetzt werden, weil das Protein in kleine Peptide fragmentiert werden kann, deren Massen oder Sequenzen durch MS zugänglich sind, anschließend kann die Proteinsequenz von den Massen oder Sequenzen der Peptidfragmente mithilfe der Genomsequenz abgeleitet werden. Daher ist die Kenntnis der Genomsequenz Voraussetzung für diese Art von MS-Methoden. Henzel et al. führte diese Methode in den späten 80ern ein.<sup>[122]</sup> Sie reinigten *E. coli*-Proteinextrakte mit zweidimensionaler Gelelektrophorese. Die Proteine wurde mit Elektroblotting auf eine Poly(vinylidendifluorid)-Membran (PVDF) übertragen und mit

Coomassie-Blau gefärbt. Einzelne Proteinspots wurden ausgeschnitten, in situ mit Dithiothreitol reduziert, mit Iodessigsäure alkyliert und mit Trypsin verdaut. Die so gewonnenen Peptidfragmente wurden dann mit MALDI analysiert. Die Massen der Peptidfragmente wurden mit dem Computerprogramm FRAGFIT verarbeitet, das nach Peptidsequenzen sucht, die mit den erhaltenen Massen aus einer Proteinsequenz-Datenbank übereinstimmen. Auf diese Weise kann die Identität der Proteine, die durch 2D-Gelektrophorese getrennt wurden, abgeleitet werden. Von den zehn Proteinspots, die analysiert wurden, konnte die Identität von neun Spots durch N-terminale Sequenzanalyse bestätigt werden. Inzwischen gibt es auch Veröffentlichungen, in denen Methoden zur Verbesserung der Empfindlichkeit und der Treffsicherheit bei der Identifizierung beschrieben werden.<sup>[123]</sup>

## 5.2. Anwendungen

MS-Methoden können zur Identifizierung von Proteinen eingesetzt werden, die einen Proteinkomplex bilden. Ein Vorteil davon ist, dass in einem Schritt der gesamte Komplex und nicht eine einzelne Protein-Protein-Wechselwirkung untersucht wird. Link et al. identifizierten mit MS 75 von 78 vorhergesagten ribosomalen Hefeproteinen aus dem 80S-Ribosomenkomplex.<sup>[124]</sup> Der Proteinkomplex wurde gereinigt, denaturiert, reduziert und zu Peptidfragmenten verdaut. Diese wurden dann mit zweidimensionaler Säulenchromatographie aufgetrennt und in ein Massenspektrometer eluiert. Die Peptidsequenzen wurden mithilfe der Tandem-Massenpektrometrie (MS/MS) bestimmt und in translatierten Genomdatenbanken gesucht. Mit 120 µg gereinigtem Proteinkomplex wurden 95 Polypeptide identifiziert, von denen 90 aus verschiedenen ribosomalen Genen stammten und die 75 der 78 Proteine repräsentieren (einige Proteine werden von zwei separaten Genen mit > 98 % identischer Aminosäuresequenz codiert). Außerdem wurde YMR116C (BEL1), ein Protein, das bislang noch nicht bekannt war, ebenfalls als Teil des Heferibosoms identifiziert.

Mit quantitativen MS-Methoden kann die Höhe der Proteinexpression in verschiedenen Zellzuständen bestimmt werden. Diese Methode ähnelt den DNA-Mikroarrays, nur dass hierbei die tatsächlichen Proteinmengen und nicht die mRNA-Mengen erfasst werden. Ein umfassendes Profil der Genexpression in verschiedenen Zellstadien kann helfen, Proteine und Gene zu identifizieren, die an einem bestimmten biologischen Stoffwechselweg beteiligt sind. Gygi et al. entwickelten eine Methode mit isotopencodierten Affinitätsmarkierungen (*isotope-coded affinity tags* – ICAT), um relative Proteinexpressionshöhen in unterschiedlichen Zellstadien zu bestimmen.<sup>[125]</sup> Die ICAT besteht aus einem Biotinmolekül, einem Linker, in den unterschiedliche Isotope eingebaut werden können, und einer thiolspezifischen reaktiven Gruppe (Abbildung 19). Lösliche Proteine aus Hefezellen wurden in Medium mit entweder 2 % Galactose oder 2 % Ethanol als Kohlenstoffquelle gezogen, dann isoliert und reduziert. Mit zwei ICAT-Reagentien, eines mit einem schweren Isotop und eines mit einem leichten Isotop, wurden die beiden Proteinproben markiert und dann gemischt und verdaut. Cystein

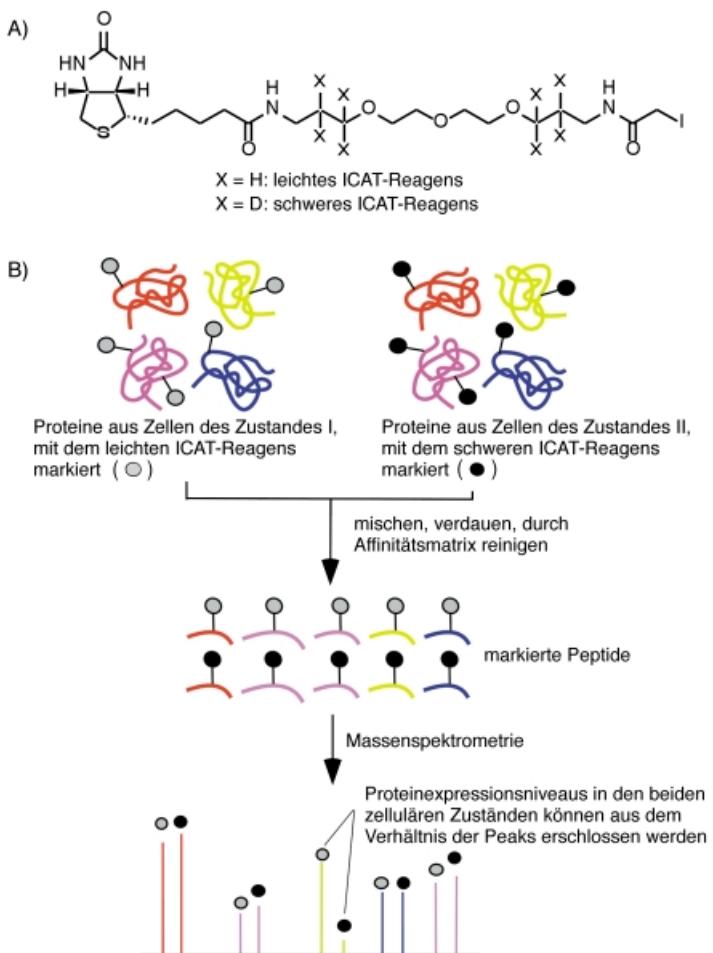


Abbildung 19. MS-Verfahren zur Bestimmung der relativen Proteinexpressionshöhe bei verschiedenen Zellzuständen.<sup>[125]</sup> A) Struktur der ICAT. B) Isolierung und Reduzierung löslicher Proteine aus Zellen in zwei unterschiedlichen Zuständen. Mit den beiden ICAT-Reagentien, eines mit schwerem Isotop, eines mit leichtem Isotop, wird je eine der beiden Proteinproben markiert. Dann werden die beiden Proteinproben gemischt und verdaut. Cystein enthaltende Peptide, die mit dem ICAT-Reagens markiert sind, werden durch Avidin-Affinitätschromatographie isoliert. Nach Trennung durch Mikrokapillar-Flüssigchromatographie werden die markierten Peptide massenspektrometrisch analysiert. Die relativen Proteinexpressionshöhen bei den beiden Zellzuständen können aus dem Verhältnis korrespondierender Peaks aus den beiden Proben gemessen werden. Die Identität des Proteins erhält man, indem man zunächst die Peptidsequenz mit Tandem-MS bestimmt und dann in einer Datenbank nach der Sequenz sucht.

enthaltende Peptide wurden mit Avidin-Affinitätschromatographie isoliert und mit Mikrokapillar-Flüssigkeitschromatographie gekoppelt mit MS/MS sequenziert. Anhand dieser Peptidsequenzen wurden die fünf vollständigen Proteinsequenzen in Sequenzdatenbanken gesucht. Die relativen Expressionshöhen bei den zwei unterschiedlichen Zellzuständen wurden durch Vergleich der Intensitäten jeweils von Paaren der Peptid-Molekülen mit identischen Sequenzen, die mit ICAT-Reagentien mit verschiedenen Isotopen markiert worden waren, erhalten. Die Expressionshöhen von 34 Genen aus Hefe, die in 2 % Galactose oder 2 % Ethanol als Kohlenstoffquelle gewachsen waren, wurden miteinander verglichen; die Ergebnisse korrelierten gut mit dem, was vom Hefestoffwechsel bekannt ist.

Wie bei der Phagen- und Ribosomen-Display-Technik ist es auch bei MS-Methoden viel schwieriger, nicht Bindungen, sondern Enzymaktivitäten zu testen. Cravatt et al. wiesen kürzlich mit einem spezifischen Inhibitor Serinhydrolasen in Gewebextrakten mit einer Methode nach, die sie „Aktivitätsbasiertes Proteinprofilierung“ nannten.<sup>[89]</sup> In ihrem speziellen Fall nutzten sie ein biotinyliertes Fluorophosphonat (FP-biotin). Fluorophosphonat-Derivate sind selektive und irreversible Inhibitoren von Serinhydrolasen, und sie reagieren nur mit Enzymen, die sich in einem katalytisch aktiven Zustand befinden. Rohextrakte von Rattengehirn wurden mit FP-Biotin inkubiert, und anschließend wurden die Proteine einer Standard-SDS-Polyacrylamidgelektrophorese mit nachfolgendem Western-Blot unterworfen. Für die Trennung wurde ein Konjugat aus Avidin und Meerrettich-Peroxidase genutzt. Mehr als zehn Proteine wurden entdeckt, von denen zwei, ein 75-kDa-Protein und ein 85-kDa-Protein, mit MS analysiert wurden. Chemische Sonden für Cystein-Proteasen wurden ebenfalls beschrieben.<sup>[90]</sup>

### 5.3. Zusammenfassung

MS-Methoden sind noch sehr jung, sodass sich in den nächsten Jahren an der Technik noch viel ändern kann. Theoretisch sind MS-Testsysteme nur durch die Masse der Proteine limitiert; dies macht Bibliotheken möglich, die sogar größer sind als die bei der RNA-Display-Technik. Dennoch wurden bisher nur relativ kleine Bibliotheken getestet. MS-Methoden haben keine inhärenten Einschränkungen bezüglich des Typs der zu testenden Proteine, und sie sind kompatibel mit In-vitro- und In-vivo-Testsystemen. Es ist noch nicht klar, welche Typen von Testsystemen am besten mit MS-Methoden verknüpft werden, dafür gibt viele Möglichkeiten. Andererseits sind auch viele technische Fragestellungen noch zu lösen, beispielsweise, wie man Proteine mit sehr niedrigen zellulären Konzentrationen nachweisen kann, oder wie man die DNA, die einen Protein-Hit codiert, für weitere Mutageneserunden regeneriert. Neben MS gibt es vielleicht auch noch andere Methoden für eine direkte Sequenzierung von Proteinen, die bislang übersehen wurden.

## 6. Zusammenfassung und Ausblick

Die Entwicklung robuster Methoden zum Test von Proteinfunktionen während der letzten beiden Jahrzehnte lässt auf einen Durchbruch beim Protein-Engineering und in der Proteomik-Forschung in naher Zukunft hoffen. Verschiedene Methoden, um ein Protein mit seiner DNA-Sequenz zu kuppeln, sind inzwischen gut etabliert. Sie reichen von Methoden wie der Phagen-Display-Technik, bei dem eine physische Verbindung zwischen dem Protein und seiner DNA-Sequenz hergestellt wird, über Tests in Mikrotiterplatten, bei denen jedes Protein einfach unabhängig vom anderen untersucht wird, bis hin zu zellulären Testsystemen, bei denen jede Zelle ein einzelnes Plasmid trägt, das eine individuelle Proteinvariante codiert. Diese Methoden sind bereits eindrucksvoll, doch deuten die letzten Fortschritte bei Pro-

teinchips und Massenspektrometrie darauf hin, dass neue Methoden noch weitere Vorteile erbringen können. Schon jetzt erzielen die Forscher ihre ersten Erfolge bei der Modifizierung der Funktionen bestehender Proteine. Es soll dabei nicht verschwiegen werden, dass diese Erfolge zum Teil damit zusammenhängen, dass passende Reaktionen ausgewählt wurden, zu denen es eine bequeme Selektion oder ein gutes Screeningsystem gibt, z. B. eine Reaktion, deren Produkt fluoresziert oder essenziell für die Zelle ist. Die nächste Herausforderung ist es, allgemeine Testsysteme für enzymatische Aktivität zu entwickeln, die die Etablierung einer neuen Chemie ermöglicht, und diese Methoden nicht nur beim Protein-Engineering sondern auch in der Proteomik-Forschung anzuwenden.

*Wir danken S. Licht, C. Noren, T. Yano und den Mitarbeitern unseres Labors für ihre hilfreichen Kommentare während der Vorbereitung dieses Manuskriptes.*

Eingegangen am 6. Juni 2001 [A478]  
Übersetzt von Dr. Burkard Neuß, Jülich

- [1] G. P. Smith, V. A. Petrenko, *Chem. Rev.* **1997**, *97*, 391.
- [2] S. S. Sidhu, *Curr. Opin. Biotechnol.* **2000**, *11*, 610.
- [3] D. J. Rodi, L. Makowski, *Curr. Opin. Biotechnol.* **1999**, *10*, 87.
- [4] P. Forrer, S. Jung, A. Plückthun, *Curr. Opin. Struct. Biol.* **1999**, *9*, 514.
- [5] G. P. Smith, *Science* **1985**, *228*, 1315.
- [6] D. Sblattero, A. Bradbury, *Nat. Biotechnol.* **2000**, *18*, 75.
- [7] S. S. Sidhu, H. B. Lowman, B. C. Cunningham, J. A. Wells, *Methods Enzymol.* **2000**, *328*, 333.
- [8] S. Bass, R. Green, J. A. Wells, *Proteins: Struct. Funct. Genet.* **1990**, *8*, 309.
- [9] J. Greenwood, A. E. Willis, R. N. Perham, *J. Mol. Biol.* **1991**, *220*, 821.
- [10] C. Gao, S. Mao, C. L. Lo, P. Wirsching, R. A. Lerner, K. D. Janda, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 6025.
- [11] M. Duenas, C. A. Borrebaeck, *Bio/Technology* **1994**, *12*, 999.
- [12] K. Gramatikoff, O. Georgiev, W. Schaffner, *Nucleic Acids Res.* **1994**, *22*, 5761.
- [13] J. K. Scott, G. P. Smith, *Science* **1990**, *249*, 386.
- [14] N. C. Wrighton, F. X. Farrell, R. Chang, A. K. Kashyap, F. P. Barbone, L. S. Mulcahy, D. L. Johnson, R. W. Barrett, L. K. Jolliffe, W. J. Dower, *Science* **1996**, *273*, 458.
- [15] R. Hyde-DeRuysscher, L. A. Paige, D. J. Christensen, N. Hyde-DeRuysscher, A. Lim, Z. L. Fredericks, J. Kranz, P. Gallant, J. Zhang, S. M. Rocklage, D. M. Fowlkes, P. A. Wendler, P. T. Hamilton, *Chem. Biol.* **2000**, *7*, 17.
- [16] S. R. Whaley, D. S. English, E. Hu, P. F. Barbara, A. M. Belcher, *Nature* **2000**, *405*, 665.
- [17] W. Dall'Acqua, P. Carter, *Curr. Opin. Struct. Biol.* **1998**, *8*, 443.
- [18] H. R. Hoogenboom, P. Chames, *Immunol. Today* **2000**, *21*, 371.
- [19] J. McCafferty, A. D. Griffiths, G. Winter, D. J. Chiswell, *Nature* **1990**, *348*, 552.
- [20] A. S. Kang, C. F. Barbas III, K. D. Janda, S. J. Benkovic, R. A. Lerner, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 4363.
- [21] C. F. Barbas, III, A. S. Kang, R. A. Lerner, S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 7978.
- [22] J. Yu, S. Y. Choi, K.-D. Moon, H.-H. Chung, H. J. Youn, S. Jeong, H. Park, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 2880.
- [23] M. Baca, T. S. Scanlan, R. C. Stephenson, J. A. Wells, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 10063.
- [24] K. D. Janda, L.-C. Lo, C.-H. L. Lo, M.-M. Sim, R. Wang, C.-H. Wong, R. A. Lerner, *Science* **1997**, *275*, 945.
- [25] M. Shanmugavelu, A. R. Baytan, J. D. Chesnut, B. C. Bonning, *J. Biol. Chem.* **2000**, *275*, 1802.
- [26] Y. Choo, A. Klug, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11163.
- [27] A. C. Jamieson, S.-H. Kim, J. A. Wells, *Biochemistry* **1994**, *33*, 5689.

- [28] E. J. Rebar, C. O. Pabo, *Science* **1994**, 263, 671.
- [29] G. A. Weiss, H. B. Lowman, *Chem. Biol.* **2000**, 7, 547.
- [30] G. Beste, F. S. Schmidt, T. Stibora, A. Skerra, *Proc. Natl. Acad. Sci. USA* **1999**, 96, 1898.
- [31] S. Schlehuber, G. Beste, A. Skerra, *J. Mol. Biol.* **2000**, 297, 1105.
- [32] J. C. Guo, W. Huang, T. S. Scanlan, *J. Am. Chem. Soc.* **1994**, 116, 6062.
- [33] P. Soumillion, L. Jespers, M. Bouchet, J. Marchand-Brynaert, G. Winter, J. Fastrez, *J. Mol. Biol.* **1994**, 237, 415.
- [34] C. Gao, C.-H. Lin, C.-H. L. Lo, S. Mao, P. Wirsching, R. A. Lerner, K. D. Janda, *Proc. Natl. Acad. Sci. USA* **1997**, 94, 11777.
- [35] H. Pedersen, S. Holder, D. P. Sutherlin, U. Schwitter, D. S. King, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **1998**, 95, 10523.
- [36] S. Demartis, A. Huber, F. Viti, L. Lozzi, L. Giovannoni, P. Neri, G. Winter, D. Neri, *J. Mol. Biol.* **1999**, 286, 617.
- [37] J.-L. Jestin, P. Kristensen, G. Winter, *Angew. Chem.* **1999**, 111, 1196; *Angew. Chem. Int. Ed.* **1999**, 38, 1124.
- [38] S. Atwell, J. A. Wells, *Proc. Natl. Acad. Sci. USA* **1999**, 96, 9497.
- [39] I. Ponsard, M. Galleni, P. Soumillion, J. Fastrez, *ChemBioChem* **2001**, 2, 253.
- [40] R. W. Roberts, *Curr. Opin. Chem. Biol.* **1999**, 3, 268.
- [41] L. C. Mattheakis, R. R. Bhatt, W. J. Dower, *Proc. Natl. Acad. Sci. USA* **1994**, 91, 9022.
- [42] J. Hanes, A. Plückthun, *Proc. Natl. Acad. Sci. USA* **1997**, 94, 4937.
- [43] R. W. Roberts, J. W. Szostak, *Proc. Natl. Acad. Sci. USA* **1997**, 94, 12297.
- [44] N. Nemoto, E. Miyamoto-Sato, Y. Husimi, H. Yanagawa, *FEBS Lett.* **1997**, 414, 405.
- [45] G. Cho, A. D. Keefe, R. Liu, D. S. Wilson, J. W. Szostak, *J. Mol. Biol.* **2000**, 297, 309.
- [46] R. Liu, J. E. Barrick, J. W. Szostak, R. W. Roberts, *Methods Enzymol.* **2000**, 318, 268.
- [47] J. Hanes, C. Schaffitzel, A. Knappik, A. Plückthun, *Nat. Biotechnol.* **2000**, 18, 1287.
- [48] D. S. Wilson, A. D. Keefe, J. W. Szostak, *Proc. Natl. Acad. Sci. USA* **2001**, 98, 3750.
- [49] A. D. Keefe, J. W. Szostak, *Nature* **2001**, 410, 715.
- [50] M. G. Cull, J. F. Miller, P. J. Schatz, *Proc. Natl. Acad. Sci. USA* **1992**, 89, 1865.
- [51] C. M. Gates, W. P. C. Stemmer, R. Kaptein, P. J. Schatz, *J. Mol. Biol.* **1996**, 255, 373.
- [52] S. E. Cwirla, P. Balasubramanian, D. J. Duffin, C. R. Wagstrom, C. M. Gates, S. C. Singer, A. M. Davis, R. L. Tansik, L. C. Mattheakis, C. M. Boytos, P. J. Schatz, D. P. Baccanari, N. C. Wrighton, R. W. Barrett, W. J. Dower, *Science* **1997**, 276, 1696.
- [53] S. Ståhl, M. Uhlén, *Trends Biotechnol.* **1997**, 15, 185.
- [54] J. A. Francisco, C. F. Earhart, G. Georgiou, *Proc. Natl. Acad. Sci. USA* **1992**, 89, 2713.
- [55] E. T. Boder, K. D. Wittrup, *Nat. Biotechnol.* **1997**, 15, 553.
- [56] W. Ernst, R. Grabherr, D. Wegner, N. Borth, A. Grassauer, H. Katinger, *Nucleic Acids Res.* **1998**, 26, 1718.
- [57] A. Christmann, K. Walter, A. Wentzel, R. Kratzner, H. Kolmar, *Protein Eng.* **1999**, 12, 797.
- [58] P. S. Daugherty, G. Chen, M. J. Olsen, B. L. Iverson, G. Georgiou, *Protein Eng.* **1998**, 11, 825.
- [59] M. C. Kieke, E. V. Shusta, E. T. Boder, L. Teyton, K. D. Wittrup, D. M. Kranz, *Proc. Natl. Acad. Sci. USA* **1999**, 96, 5651.
- [60] E. T. Boder, K. S. Midelfort, K. D. Wittrup, *Proc. Natl. Acad. Sci. USA* **2000**, 97, 10701.
- [61] M. J. Olsen, D. Stephens, D. Griffiths, P. Daugherty, G. Georgiou, B. L. Iverson, *Nat. Biotechnol.* **2000**, 18, 1071.
- [62] G. W. Beadle, E. L. Tatum, *Proc. Natl. Acad. Sci. USA* **1941**, 27, 499.
- [63] M. Hensel, J. E. Shea, C. Gleeson, M. D. Jones, E. Dalton, D. W. Holden, *Science* **1995**, 269, 400.
- [64] S. Benzer, *Sci. Am.* **1962**, 72.
- [65] P. Uetz, L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, J. M. Rothberg, *Nature* **2000**, 403, 623.
- [66] B. L. Drees, *Curr. Opin. Chem. Biol.* **1999**, 3, 64.
- [67] D. S. McNabb, L. Guarente, *Curr. Opin. Biotechnol.* **1996**, 7, 554.
- [68] E. Warbrick, *Structure* **1997**, 5, 13.
- [69] H. Lin, V. W. Cornish, *Angew. Chem.* **2001**, 113, 895; *Angew. Chem. Int. Ed.* **2001**, 40, 871.
- [70] R. K. Brachmann, J. D. Boeke, *Curr. Opin. Biotechnol.* **1997**, 8, 561.
- [71] M. Vidal, P. Legrain, *Nucleic Acids Res.* **1999**, 27, 919.
- [72] P. Colas, R. Brent, *Trends Biotechnol.* **1998**, 16, 355.
- [73] S. Fields, O. Song, *Nature* **1989**, 340, 245.
- [74] P. Colas, B. Cohen, T. Jessen, I. Grishina, J. McCoy, R. Brent, *Nature* **1996**, 380, 548.
- [75] T. C. Norman, D. L. Smith, P. K. Sorger, B. L. Drees, S. M. O'Rourke, T. R. Hughes, C. J. Roberts, S. H. Friend, S. Fields, A. W. Murray, *Science* **1999**, 285, 591.
- [76] P. Youderian, A. Vershon, S. Bouvier, R. T. Sauer, M. M. Susskind, *Cell* **1983**, 35, 777.
- [77] S. L. Dove, J. K. Joung, A. Hochschild, *Nature* **1997**, 386, 627.
- [78] J. K. Joung, E. I. Ramm, C. O. Pabo, *Proc. Natl. Acad. Sci. USA* **2000**, 97, 7382.
- [79] H. A. Greisman, C. O. Pabo, *Science* **1997**, 275, 657.
- [80] D. M. Spencer, T. J. Wandless, S. L. Schreiber, G. R. Crabtree, *Science* **1993**, 262, 1019.
- [81] E. J. Licitra, J. O. Liu, *Proc. Natl. Acad. Sci. USA* **1996**, 93, 12817.
- [82] E. C. Griffith, E. J. Licitra, J. O. Liu, *Methods Enzymol.* **2000**, 328, 89.
- [83] H. Lin, W. M. Abida, R. T. Sauer, V. W. Cornish, *J. Am. Chem. Soc.* **2000**, 122, 4247.
- [84] H. Joo, Z. Lin, F. H. Arnold, *Nature* **1999**, 399, 670.
- [85] B. Hall in *Evolutionary Biology*, Bd. 15 (Hrsg.: M. Hecht, B. Wallace, G. Pierce), Plenum, New York, **1982**, S. 85.
- [86] J. D. Hermes, S. C. Blacklow, J. R. Knowles, *Proc. Natl. Acad. Sci. USA* **1990**, 87, 696.
- [87] T. Yano, S. Oue, H. Kagamiyama, *Proc. Natl. Acad. Sci. USA* **1998**, 95, 5511.
- [88] G. Xia, L. Chen, T. Sera, M. Fa, P. G. Schultz, F. E. Romesberg, *Proc. Natl. Acad. Sci. USA* **2002**, 99, 6597.
- [89] Y. Liu, M. P. Patricelli, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **1999**, 96, 14694.
- [90] D. Greenbaum, K. F. Medzihradzky, A. Burlingame, M. Bogoy, *Chem. Biol.* **2000**, 7, 569.
- [91] D. S. Tawfik, A. D. Griffiths, *Nat. Biotechnol.* **1998**, 16, 652.
- [92] N. Doi, H. Yanagawa, *FEBS Lett.* **1999**, 457, 227.
- [93] F. J. Ghadessy, J. L. Ong, P. Holliger, *Proc. Natl. Acad. Sci. USA* **2001**, 98, 4552.
- [94] A. Q. Emili, G. Cagney, *Nat. Biotechnol.* **2000**, 18, 393.
- [95] G. Walter, K. Büsow, D. Cahill, A. Lueking, H. Lehrach, *Curr. Opin. Microbiol.* **2000**, 3, 298.
- [96] E. J. Grayhack, E. M. Phizicky, *Curr. Opin. Chem. Biol.* **2001**, 5, 34.
- [97] H. Zhu, M. Snyder, *Curr. Opin. Chem. Biol.* **2001**, 5, 40.
- [98] R. Manns, [http://www.messebasel.ch/miptec/mission\\_history.htm](http://www.messebasel.ch/miptec/mission_history.htm), 1999.
- [99] A. Voller, D. Bidwell, G. Huldt, E. Engvall, *Bull. WHO* **1974**, 51, 209.
- [100] M. R. Martzen, S. M. McCraith, S. L. Spinelli, F. M. Torres, S. Fields, E. J. Grayhack, E. M. Phizicky, *Science* **1999**, 286, 1153.
- [101] S. A. Sundberg, *Curr. Opin. Biotechnol.* **2000**, 11, 47.
- [102] R. P. Hertzberg, A. J. Pope, *Curr. Opin. Chem. Biol.* **2000**, 4, 445.
- [103] J. E. González, P. A. Negulescu, *Curr. Opin. Biotechnol.* **1998**, 9, 624.
- [104] N. Bosworth, P. Towers, *Nature* **1989**, 341, 167.
- [105] A. J. Pope, U. Haupps, K. J. Moore, *Drug Discovery Today* **1999**, 4, 350.
- [106] A. Koltermann, U. Kettling, J. Bieschke, T. Winkler, M. Eigen, *Proc. Natl. Acad. Sci. USA* **1998**, 95, 1421.
- [107] S. N. Thorn, R. G. Daniels, M. M. Auditor, D. Hilvert, *Nature* **1995**, 373, 228.
- [108] L. Giver, A. Gershenson, P.-O. Freskgard, F. H. Arnold, *Proc. Natl. Acad. Sci. USA* **1998**, 95, 12809.
- [109] K. A. Haushalter, P. T. Stukenberg, M. W. Kirschner, G. L. Verdine, *Curr. Biol.* **1999**, 9, 174.
- [110] A. Lueking, M. Horn, H. Eickhoff, K. Büsow, H. Lehrach, G. Walter, *Anal. Biochem.* **1999**, 270, 103.
- [111] H. Ge, *Nucleic Acids Res.* **2000**, 28, e3.
- [112] G. MacBeath, S. L. Schreiber, *Science* **2000**, 289, 1760.
- [113] C. A. Rowe, L. M. Tender, M. J. Feldstein, J. P. Golden, S. B. Scruggs, B. D. MacCraith, J. J. Cras, F. S. Ligler, *Anal. Chem.* **1999**, 71, 3846.
- [114] P. Arenkov, A. Kukhtin, A. Gemmell, S. Voloshchuk, V. Chupeeva, A. Mirzabekov, *Anal. Biochem.* **2000**, 278, 123

- [115] B. Schweitzer, S. Wiltshire, J. Lambert, S. O'Malley, K. Kukanskis, Z. Zhu, S. F. Kingsmore, P. M. Lizardi, D. C. Ward, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 10113.
- [116] H. Zhu, J. F. Klemic, S. Chang, P. Bertone, A. Casamayor, K. G. Klemic, D. Smith, M. Gerstein, M. Reed, M. Snyder, *Nat. Genet.* **2000**, *26*, 283.
- [117] J. Ziauddin, D. M. Sabatini, *Nature* **2001**, *411*, 107.
- [118] H. Zhu, M. Bilgin, R. Bangham, D. Hall, A. Casamayor, P. Bertone, N. Lan, R. Jansen, S. Bidlingmaier, T. Houfek, T. Mitchell, P. Miller, R. A. Dean, M. Gerstein, M. Snyder, *Science* **2001**, *293*, 2101.
- [119] J. S. Andersen, M. Mann, *FEBS Lett.* **2000**, *480*, 25.
- [120] M. J. Chalmers, S. J. Gaskell, *Curr. Opin. Biotechnol.* **2000**, *11*, 384.
- [121] J. Godovac-Zimmermann, L. R. Brown, *Mass Spectrom. Rev.* **2001**, *20*, 1.
- [122] W. J. Henzel, T. M. Billeci, J. T. Stults, S. C. Wong, C. Grimley, C. Watanabe, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 5011.
- [123] A. Shevchenko, O. N. Jensen, A. V. Podtelejnikov, F. Sagliocco, M. Wilm, O. Vorm, P. Mortensen, A. Shevchenko, H. Boucherie, M. Mann, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 14440.
- [124] A. J. Link, J. Eng, D. M. Schieltz, E. Carmack, G. J. Mize, D. R. Morris, B. M. Garvik, I. J. R. Yates, *Nat. Biotechnol.* **1999**, *17*, 676.
- [125] S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, R. Aebersold, *Nat. Biotechnol.* **1999**, *17*, 994.