

Genetische Selektion – eine Strategie zur Untersuchung und Herstellung von Enzymen

Sean V. Taylor, Peter Kast und Donald Hilvert*

Natürliche Enzyme sind über Jahrtausende durch den graduellen Prozess der darwinistischen Evolution entstanden. Die Grundelemente der Evolution – Mutation, Selektion und Vervielfältigung – können auch im Labor eingesetzt werden, um Proteinkatalysatoren in entsprechend kürzeren Zeiträumen herzustellen und zu charakterisieren. Genetische In-vivo-Selektionsstrategien ermöglichen eine er-

schöpfende Analyse von Proteinbibliotheken mit 10^{10} unterschiedlichen Mitgliedern; und mit In-vitro-Methoden lassen sich sogar noch größere Ensembles studieren. Folglich können evolutionäre Ansätze statistisch wertvolle Einsichten in komplexe und oftmals subtile Wechselwirkungen geben, die die Proteinfaltung, die Struktur und den katalytischen Mechanismus beeinflussen. Solche Me-

thoden werden zunehmend auch als Hilfsmittel beim Proteindesign eingesetzt, da sie den Zugang zu neuen Proteinen mit maßgeschneiderten katalytischen Aktivitäten und Selektivitäten eröffnen.

Stichwörter: Enzyme • Enzymkatalyse • Genetische Selektion • Gesteuerte Evolution • Proteindesign • Proteinstrukturen

1. Einleitung

Evolution bezeichnet den langsamem, kontinuierlichen Prozess, dem alle Lebewesen unterliegen und in dessen Verlauf sie sich diversifizieren und an Komplexität gewinnen. In sich wiederholenden Cyclen von Mutation, Selektion und Vervielfältigung reichern sich so neue Merkmale in einer Population von Organismen an.^[1] Diejenigen Merkmale, die einen Vorteil unter den vorherrschenden Umweltbedingungen verschaffen, werden von einer Generation an die nächste weitergegeben. Bereits seit Urzeiten hat sich die Menschheit direkt der Evolution bedient, um Pflanzen und Tiere mit nützlichen Eigenschaften zu züchten. Die Kreuzung von Individuen mit vorteilhaften Charakteristika nutzt mit Erfolg die sexuelle Rekombination, eine der mächtigsten evolutiven Strategien, um neue Varianten hervorzu bringen. Aus diesen Kreuzungen werden die Nachkommen mit den attraktivsten Eigenschaften für die weitere Züchtung ausgewählt, wodurch die Entwicklung in eine bestimmte Richtung gesteuert wird.

Biologen und Chemiker haben in neuerer Zeit damit begonnen, evolutionäre Strategien einzusetzen, um die Eigenschaften einzelner Moleküle anstelle ganzer Organismen

gezielt zu verändern. Eine große Zahl an Methoden wurde entwickelt, um Diversität in Molekülpopulationen zu erzielen. Abhängig vom jeweiligen Experiment erfordert die Mutagenese die degenerierte Oligonucleotid-gesteuerte oder fehlerbehaftete (error-prone) DNA-Synthese,^[2, 3] das Mischen mutierter DNA-Fragmente^[4, 5] oder die kombinatorische Synthese von Bibliotheken chemischer Verbindungen.^[6] Aus den resultierenden molekularen Ensembles müssen die gewünschten Mitglieder durch Selektions- oder Screeningverfahren identifiziert werden. Im Amplifikationsschritt können die selbstreplizierenden Eigenschaften der ausgewählten Moleküle genutzt werden (z. B. im Falle von Nucleinsäuren). Alternativ dazu kann die gewünschte Verbindung durch chemische Synthese in großem Maßstab hergestellt werden.

In der Praxis ist die Auswahlprozedur der kritischste Schritt jedes im Labor durchgeföhrten Evolutionsexperimentes. Wie isoliert man effizient seltene, aber nützliche Varianten aus einer komplexen Mischung mit wenigen wünschenswerten Molekülen? Ein Chemiker, der nach dem optimalen oktameren Peptidinhibitor für ein bestimmtes Enzym strebt, würde vielleicht damit beginnen, ein bekanntes Peptid an einer Stelle systematisch zu substituieren. Beschränkte man die erlaubten Substitutionen auf die 20 Standardaminoäuren, wäre die individuelle Prüfung der 20 unterschiedlichen oktameren Peptide, um den Inhibitor mit der höchsten Affinität zu finden, relativ trivial. Hingegen würde die gleichzeitige, willkürliche Variation von zwei Positionen des Peptids bereits 400 (20^2) Varianten liefern. Das Durchmustern und Testen einer solchen Bibliothek wäre um einiges auf-

[*] Prof. D. Hilvert, Dr. S. V. Taylor, Dr. P. Kast
Laboratorium für Organische Chemie
ETH Zürich
8093 Zürich (Schweiz)
Fax: (+41) 1-632-1486
E-mail: hilvert@org.chem.ethz.ch

wändiger, aber immer noch handhabbar. Jedoch wäre es sicherlich eine nahezu unlösbare Aufgabe, den besten Inhibitor in einer kombinatorischen Bibliothek vollständig randomisierter Oktamerpeptide mit 20^8 ($= 2.6 \times 10^{10}$) verschiedenen Mitgliedern zu erkennen, es sei denn, man verfügte über eine extrem effiziente und empfindliche Sortiermethode.

Die genetische Selektion ist zurzeit vielleicht die effizienteste zur Verfügung stehende Technik, um große Populationen biologischer Makromoleküle zu analysieren. In diesem Übersichtsartikel konzentrieren wir uns auf die Verwendung von Selektionsprinzipien als Werkzeug für die mechanistische Enzymologie und das Proteindesign. Eine erschöpfende Abhandlung der unzähligen bislang entwickelten genetischen Selektionssysteme würde allerdings den Rahmen des Artikels sprengen, weshalb wir uns hier auf eine kleine Auswahl beschränken. Den interessierten Leser möchten wir deshalb auch auf andere, in den letzten Jahren erschienene Übersichtsartikel zu diesem Thema verweisen.^[7-17]

2. Selektion versus Screening

Die Identifizierung interessanter Varianten in großen kombinatorischen Bibliotheken (oder Genbanken) kann entweder durch die individuelle Prüfung aller Mitglieder erreicht werden oder aber indem man Bedingungen verwendet, unter denen nur die interessanten Varianten auftreten.

Leider werden die Begriffe „Screening“ für die erstgenannte Strategie und „Selektion“ für die zweite oft falsch angewandt. Der prinzipielle Unterschied zwischen den beiden Ansätzen ist in Abbildung 1 schematisch dargestellt.

Generell erfordern Screeningstrategien eine aktive Untersuchung aller in der Bibliothek vorhandenen Varianten, wie es in Abbildung 1 A und B aus der Sicht einer Honigbiene dargestellt ist. Begehrte, nektarproduzierende Blüten müssen aus einem großen Meer von Blüten, die keinen Nektar produzieren (unerwünschte Bibliotheksmitglieder), aufgespürt werden. Falls erwünschte und unerwünschte Varianten

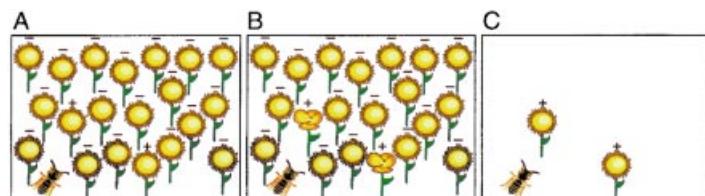


Abbildung 1. Unterschiedliche Arten der Suche nach nektarproduzierenden Blüten (mit + gekennzeichnet): A) Zufallsscreening: Die Biene muss jede einzelne Blume kontrollieren, um die produzierenden von der großen Masse der nichtproduzierenden Blüten (mit – gekennzeichnet) zu unterscheiden. B) Erleichtertes Screening: Die Biene kann nektarproduzierende von nichtproduzierenden Blüten anhand ihres charakteristischen Phänotyps unterscheiden. C) Selektion: Die Biene trifft lediglich auf die begehrten Blüten, weil die Nektarproduktion in diesem Fall eine Voraussetzung für das Pflanzenwachstum war.

Donald Hilvert promovierte 1983 in Chemie an der Columbia University, New York. Nach einem Postdoc-Aufenthalt an der Rockefeller University, New York, folgte er einem Ruf an das Scripps Research Institute in La Jolla, Kalifornien, wo er seine Interessen im Bereich der biologischen Chemie weiterverfolgte. 1997 trat er seine gegenwärtige Professur für Chemie an der Eidgenössischen Technischen Hochschule (ETH) Zürich an. Seine Forschungsarbeiten konzentrieren sich auf die Funktion von Enzymen und die Nachahmung der Eigenschaften dieser bemerkenswerten Katalysatoren im Labor. Seine Bemühungen auf diesem Gebiet wurden mit einer Reihe von Ehrungen ausgezeichnet, unter anderem mit dem Pfizer Award in Enzyme Chemistry.



D. Hilvert

P. Kast

S. V. Taylor

Peter Kast erhielt 1986 sein Diplom in Naturwissenschaften (Dipl.-Natw. ETH) von der ETH Zürich und promovierte 1991 am Institut für Mikrobiologie der ETH. Anschließend arbeitete er ein Jahr als Postdoktorand an der ETH und von 1992 bis 1995 in den Departments of Chemistry and Molecular Biology am Scripps Research Institute in La Jolla, Kalifornien, wo er 1995 zum Assistant Professor ernannt wurde. Seit Ende 1997 ist er am Laboratorium für Organische Chemie der ETH Zürich tätig und setzt dort seine Arbeiten über die Bildung von Proteinen mittels gesteuerter molekularer Evolution unter Einsatz genetisch veränderter bakterieller Systeme fort.

Sean V. Taylor erhielt 1992 seinen B.S.-Abschluss in Chemie an der Penn State University und promovierte 1998 mit einer unter Anleitung von Tadhg Begley im Department of Chemistry an der Cornell University angefertigten Arbeit über die Biosynthese von Thiamin. Im Sommer 1998 schloss er sich der Arbeitsgruppe von D. Hilvert an der ETH Zürich als Postdoktorand an. Im Mittelpunkt seiner gegenwärtigen Arbeit steht die Evolution von Proteinen und das Proteindesign.

schwer voneinander zu unterscheiden sind, dann muss das Screening sequentiell, in zufälliger Auswahl (blind) durchgeführt werden (Abbildung 1 A). Ein „erleichtertes Screening“ ist manchmal möglich, sofern die gewünschten Varianten ein unterschiedliches Erscheinungsbild (einen speziellen „Phänotyp“) aufweisen (Abbildung 1B). Dennoch liegen zu jeder Zeit auch unnütze Bibliotheksmitglieder vor, die das Signal/Rausch-Verhältnis senken und um materielle und arbeitstechnische Ressourcen konkurrieren.

Demgegenüber werden bei Selektionsstrategien Bedingungen angewandt, die ausschließlich das Überleben der gewünschten Varianten begünstigen und so den Auswahlprozess imitieren, der mit echter darwinistischer Evolution assoziiert wird. Abbildung 1C zeigt, wie leicht die Biene unter Selektionsbedingungen – alle nicht-nektarproduzierenden Blüten wurden eliminiert – auf nektarproduzierende Blüten trifft.

2.1. Klassische Lösungen des Sortierproblems

Selektions- und/oder Screeningmethoden sind weit verbreitet in der Molekularbiologie, vor allem bei der Klonierung bzw. dem Einbau von DNA-Fragmenten oder Genen in zirkuläre DNA-Moleküle, so genannte Vektoren, die innerhalb der Bakterien durch die zelluläre DNA-Synthesemaschinerie repliziert werden können.^[18, 19] Der Klonierungsvorgang beinhaltet das enzymatische Verbinden (Ligation) der Enden des einzufügenden Fragments mit den entsprechenden komplementären Enden eines passend zurechtgeschnittenen Vektors.^[20] Das resultierende zirkuläre Molekül, ein Plasmid, wird dann durch einen „Transformation“ genannten Prozess in Bakterien (gewöhnlich der Spezies *Escherichia coli*) eingeschleust. Allerdings nimmt nur ein kleiner Teil der Zellen während der Transformation erfolgreich DNA auf. Um auf diese transformierten Bakterien zu selektieren, wird die Zellsuspension auf in Petrischalen befindliches Agarmedium ausgebracht, das ein bestimmtes Antibiotikum enthält. Weil der Vektor zusätzlich ein spezifisches Gen aufweist, das die Resistenz gegen dieses Antibiotikum vermittelt, können lediglich die Zellen überleben, sich vermehren und zu einer Kolonie auf dem Agarmedium heranwachsen, die den ursprünglichen Vektor oder das neue Plasmid aufgenommen haben. Da jede der ungefähr 10^7 Zellen einer Kolonie die gleiche DNA enthält, gehören sie alle zum selben Klon.

Nach diesem sehr wirkungsvollen anfänglichen Selektionsschritt bleibt dem Molekularbiologen die Aufgabe, diejenigen Klone zu identifizieren, die das gewünschte Plasmid und nicht nur das ursprüngliche Vektormolekül (ohne das eingefügte Gen) enthalten.^[19] Abhängig von der Komplexität der Klonierungsstrategie und der Effizienz des Ligationsschrittes verfügen möglicherweise nur relativ wenige der antibiotikaresistenten Klone über das gewünschte DNA-Fragment. Die in Abbildung 2 gezeigten Beispiele illustrieren die Überlegenheit einer In-vivo-Selektionsmethode unter Einsatz eines Klonierungsvektors für positive Selektion gegenüber zufälligem oder erleichtertem Screening, um Klone mit der begehrten genetischen Zusammensetzung („Genotyp“) zu erkennen.

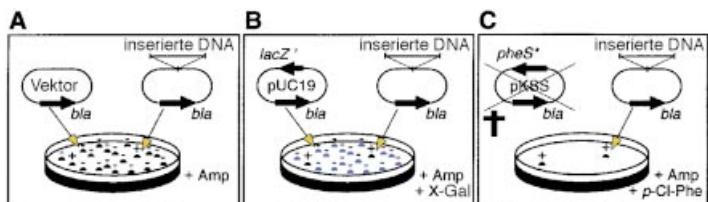


Abbildung 2. Suchstrategien zur Unterscheidung zwischen Kolonien, die Plasmide mit einem inserierten Gen aufweisen (mit + gekennzeichnet) von solchen mit der reinen Vektor-DNA (mit – gekennzeichnet). Die Kolonien werden aus einzelnen Bakterien nach der Plattierung auf nährstofffreie Agarplatten im Anschluss an die Transformation erhalten.^[20] Das bla-Gen kodiert β-Lactamase, das die Transformanten resistent gegen das Antibiotikum Ampicillin (Amp) macht. A) Zufallscreening: Der Phänotyp aller Kolonien ist identisch. Die DNA von vielen zufällig ausgewählten Klonen muss isoliert und analysiert werden, um das gewünschte Plasmid zu finden. B) Erleichtertes Screening: Der Klonierungsvektor pUC19^[223] enthält einen Abschnitt (*lacZ'*) des β-Galactosidase kodierenden *lacZ*-Gens. Das kodierte Polypeptid (das α-Fragment des Enzyms) ist in der Lage, sich mit dem Produkt eines zellulären *lacZ*-Fragments zu einer funktionellen β-Galactosidase zusammenzulagern.^[224] Diese spaltet 5-Brom-4-chlor-3-indolyl-β-D-galactopyranosid (X-Gal), wodurch ein blaues Indigoderivat ausfällt.^[225] Somit entwickeln Kolonien, die den Vektor enthalten, eine Blaufärbung auf X-Gal-haltigen Agarplatten. Im Unterschied dazu bilden Klone, die das gewünschte Plasmid enthalten, farblose Kolonien, weil das klonierte Gen in die das α-Fragment kodierende Region inseriert wurde, wodurch die β-Galactosidase-Aktivität eliminiert wird. Die Klone, welche als Phänotyp grauweiß sind (hier als schwarze Punkte gezeigt), müssen aus dem Hintergrund der blaugefärbten Kolonien ausgewählt werden, um den korrekten Genotyp zu erhalten. C) Selektion: Der positive-Selektionsvektor pKSS trägt ein mutiertes *pheS*-Gen (*pheS**), das die Substratspezifität der kodierten Phenylalanyl-tRNA-Synthetase-α-Untereinheit erweitert.^[226] In Gegenwart des Phenylalaninanalogons *p*-Chlorphenylalanin können Bakterien, die das mutierte *pheS*-Gen von pKSS exprimieren, nicht überleben, weil der Einbau des Substratanalogons in zelluläre Proteine eine allgemeine Fehlfunktion dieser Proteine zur Folge hat. Jedoch können transformierte Zellen auf Agarplatten, die Ampicillin und *p*-Chlorphenylalanin enthalten, zu Kolonien heranwachsen, wenn das unter diesen Umständen tödlich wirkende *pheS**-Gen von pKSS durch die Insertion eines klonierten DNA-Fragments zerstört wird. Somit weisen alle wachsenden Kolonien der ausplattierten Transformationsmischung den richtigen Genotyp auf.

2.2. Vorteile und Grenzen der Selektion

Der wichtigste Vorteil der Selektion gegenüber dem Screening ist die Möglichkeit der simultanen Analyse von wesentlich mehr Bibliotheksmitgliedern. Dies basiert auf der Tatsache, dass uninteressante Varianten nie in Erscheinung treten. Die Konsequenz daraus ist, dass die Bibliotheken sehr viel schneller und mit einem höheren Durchsatz untersucht werden können. Mit den besten zurzeit verfügbaren Screeningverfahren, bei denen fluoreszierende oder chromogene Substanzen eingesetzt werden,^[21] beträgt die maximal prüfbare Zahl an Bibliotheksmitgliedern ungefähr 10^5 .^[22] Dagegen können durch genetische Selektion bis zu 10^{10} Klone in einem einzigen In-vivo-Experiment in *E.-coli*-Zellen beurteilt werden. Diese Zahl ist eine Obergrenze, welche auf den mit konventionellen Methoden erreichbaren Transformationseffizienzen basiert.^[23] Allerdings hängt die absolute obere Grenze, zumindest prinzipiell, lediglich vom Maßstab der Transformation ab.

Trotzdem ist es eine Herausforderung, geeignete Selektionsstrategien für beliebige katalytische Umsetzungen zu entwickeln. Insbesondere kann die Realisierung von In-vivo-

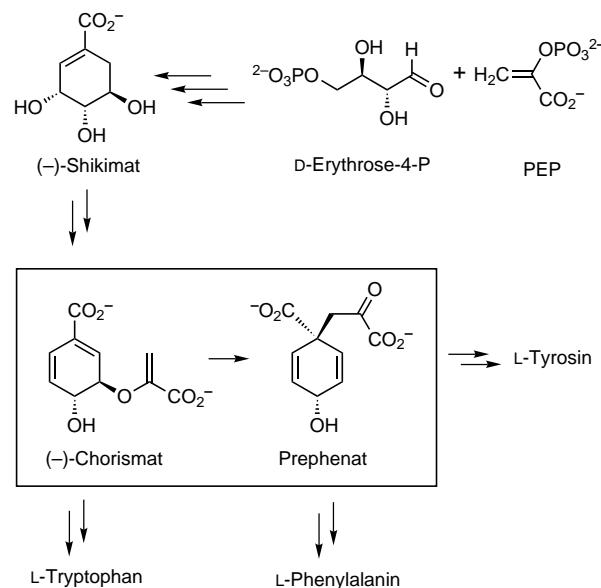
Selektionsschemata für Enzyme, die in nichtnatürlichen Umgebungen wie organischen Lösungsmitteln aktiv sein sollen,^[24, 25] schwierig oder sogar unmöglich sein. Die Koppelung der angestrebten Reaktion an das Überleben der Zellen im Selektionsschritt bedarf unter Umständen der Entwicklung von komplexen, nichttrivialen und intelligenten Versuchsanordnungen. Jedes Selektionssystem, besonders ein sehr kompliziertes, birgt stets die Möglichkeit, dass überlebensfähige, aber unvorhergesehene oder gar unerwünschte Lösungen der gestellten Überlebensaufgabe auftauchen.^[26, 27] Im Falle einer zu großen Häufung von falsch-positiven Resultaten könnte ein effizientes Screening vornötigen sein, oder das System müsste umgestaltet werden, um den unerwünschten Hintergrund durch die Einführung eines weiteren Selektionsschrittes zu eliminieren.

Speziell im Zusammenhang mit sehr großen kombinatorischen Bibliotheken gewinnt die Selektion gegenüber dem Screening enorm an Bedeutung. Sobald man von der Randomisierung eines acht Reste aufweisenden Peptids zu einem Protein aus 100 Aminosäuren (der typischen Größe einer kleinen funktionellen Domäne) übergeht, steigt die Zahl der möglichen Sequenzpermutationen auf astronomische 20^{100} . Das Vorhaben, auch nur einen winzigen Bruchteil dieses Sequenzraumes mittels gesteuerten molekularen Evolutionsexperimenten zu untersuchen, erfordert eine Selektion, selbst wenn die Entwicklung eines geeigneten Systems zunächst beträchtlich mehr Aufwand bedeutet als die Einrichtung eines Screeningverfahrens.

3. Genetische Selektion neuartiger Chorismat-Mutasen

Die Anwendung evolutionärer Strategien im Labor kann mit Proteinen, die einfache Stoffwechselreaktionen katalysieren, illustriert werden. Eine der einfachsten Reaktionen dieser Art ist die Umwandlung von Chorismat in Prephenat (Schema 1), eine [3,3]-sigmatrope Umlagerung. Diese Umsetzung ist eine Schlüsselreaktion im Shikimat-Biosyntheseweg, der in Pflanzen und niederen Organismen zu den aromatischen Aminosäuren führt.^[28, 29] Chorismat-Mutassen beschleunigen diese Reaktion mehr als millionenfach.^[30]

Chorismat-Mutassen werden seit langem intensiv untersucht, nicht zuletzt, weil weitere Beispiele für enzymkatalysierte pericyclische Reaktionen selten sind. Obwohl Mutasen unterschiedlicher Herkunft oft wenig oder keine Sequenzähnlichkeit aufweisen, sind ihre kinetischen Eigenschaften ähnlich.^[30] Die Strukturunterschiede bestehen bis hin zu den Tertiär- und Quartärstrukturen, wie aus dem Vergleich der Chorismat-Mutassen von *Bacillus subtilis* (BsCM),^[31, 32] *E. coli* (EcCM)^[33] und der Hefe *Saccharomyces cerevisiae* (ScCM)^[34, 35] hervorgeht (Abbildung 3). BsCM, ein Mitglied der AroH-Klasse von Chorismat-Mutassen, ist ein symmetrisches, als Pseudo- α/β -Fass gepacktes Homotrimer. Es verfügt über drei identische aktive Taschen, die an den Grenzflächen der Untereinheiten lokalisiert sind (Abbildung 3A). EcCM und ScCM (Abbildung 3B bzw. C) sind hingegen entfernt verwandte Homodimere, die der vollständig α -helicalen AroQ-Klasse von Chorismat-Mutassen ange-



Schema 1. Shikimat-Biosyntheseweg zu aromatischen Aminosäuren in Pflanzen und niederen Organismen. Die [3,3]-sigmatrope Umlagerung von Chorismat zu Prephenat ist im Kasten dargestellt. PEP = Phosphoenolpyruvat.

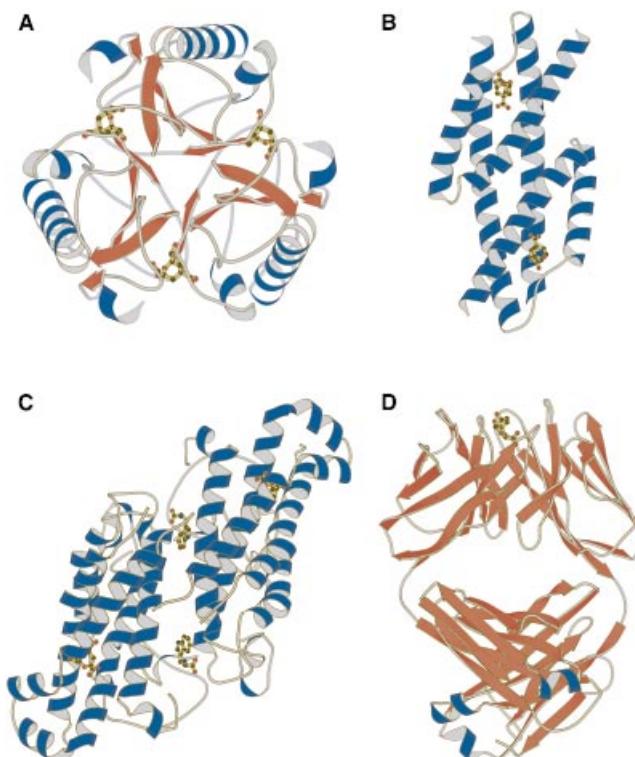


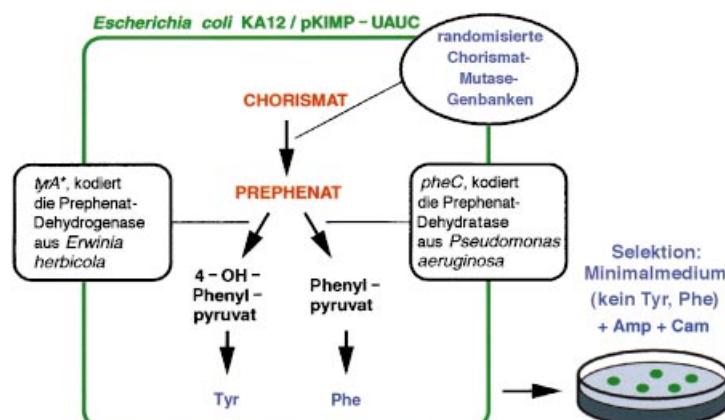
Abbildung 3. Bändermodelle der drei natürlich vorkommenden Chorismat-Mutasen und eines katalytischen Antikörpers. Helices sind blau, β -Faltblätter rot dargestellt. In jeder aktiven Tasche ist ein Molekül des Übergangszustands-analogen Inhibitors **1** (Schema 3) als Kugel-Stab-Modell mit Kohlenstoffatomen in Schwarz und Sauerstoffatomen in Rot gezeigt. A) Monofunktionelle Chorismat-Mutase von *B. subtilis* (BsCM).^[31, 32] B) Chorismat-Mutase-Domäne der difunktionalen Chorismat-Mutase/Prephenat-Dehydratase von *E. coli* (EcCM).^[33] C) Allosterisch regulierte Chorismat-Mutase der Hefe *S. cerevisiae* (ScCM).^[34] Die allosterischen Stellen des Enzyms befinden sich an der Grenzfläche zwischen den Monomerbausteinen des Dimers. In dieser Struktur sind diese Stellen mit dem Aktivator Tryptophan besetzt. D) Katalytischer Antikörper 1F7^[42] der unter Verwendung des Haptens **1** erzeugt wurde.^[39]

hören.^[36, 37] Die Suche nach Homologen in Sequenzdatenbanken hat weitere Proteine zu Tage gefördert, die einer der genannten Strukturklassen angehören.^[36–38] Aber auch katalytische Antikörper mit bescheidener Chorismat-Mutase-Aktivität konnten erhalten werden.^[39–41] Wie in Abbildung 3D für den Antikörper 1F7 gezeigt, weisen sie die typische Immunoglobulin-Faltung auf.^[42]

Die architektonische Vielfalt, die die Natur verwendet, um diese relativ einfache chemische Transformation auszuführen, ist beeindruckend und wirft viele faszinierende Fragen auf. Wie verrichten die einzelnen Enzyme ihre Arbeit? Welche Gemeinsamkeiten und Unterschiede gibt es? Warum sind natürlich vorkommende Chorismat-Mutasen 10⁴-mal effizienter als der katalytische Antikörper? Kann man auf der Grundlage von mechanistischen Erkenntnissen den Antikörper verbessern? Ganz allgemein, welche entscheidenden chemischen Faktoren liegen den einzelnen Strukturen zu Grunde? Die einzelnen katalytischen Domänen der BsCM-, EcCM- und Antikörper-Proteine sind alle etwa 100 Reste lang. Warum nimmt das Polypeptid in einem Fall eine gemischte α/β -Konformation, in einem anderen eine vollständig α -helicale Struktur und im dritten eine vollständige β -Topologie an? Der Einsatz von Zufallsmutationen und genetischer Selektion in vivo ermöglicht eine schnelle und effiziente Untersuchung des Zusammenhangs zwischen Sequenz, Struktur und Funktion.

3.1. Das Selektionssystem

In den letzten zehn Jahren haben wir einige Hefe-^[43, 44] und *E. coli*-Stämme^[45, 46] hergestellt, denen die Chorismat-Mutasefunktion fehlt. Ein typisches bakterielles Selektionssystem ist in Schema 2 dargestellt. Es basiert auf dem *E. coli*-Stamm



Schema 2. Ein *E. coli*-Selektionssystem für Chorismat-Mutase-Aktivität.^[45] Für dieses Selektionssystem musste der Stoffwechselweg zwischen Chorismat und Phenylpyruvat sowie 4-OH-Phenylpyruvat umgebaut werden. Die difunktionalen Enzyme Chorismat-Mutase/Prephenat-Dehydratase und Chorismat-Mutase/Prephenat-Dehydrogenase kodierende Gene wurden entfernt. Monofunktionelle Versionen der Dehydrogenase und der Dehydratase werden mit dem Plasmid pKIMP-UAUC zur Verfügung gestellt. Potentielle Chorismat-Mutasen aus randomisierten Genbanken werden aufgrund ihrer Fähigkeit bewertet, den genetischen Defekt zu komplementieren, sodass die Zellen in Abwesenheit von Tyrosin (Tyr) und Phenylalanin (Phe) wachsen können. Das Genbankplasmid und pKIMP-UAUC weisen Gene auf, die Resistenz gegen Ampicillin (Amp) bzw. Chloramphenicol (Cam) vermitteln.

KA12,^[46] dem die chromosomal Gene für beide difunktionalen Chorismat-Mutasen (Chorismat-Mutase/Prephenat-Dehydrogenase und Chorismat-Mutase/Prephenat-Dehydratase) entfernt wurden. Monofunktionelle Versionen der Prephenat-Dehydratase^[47] und der Prephenat-Dehydrogenase,^[48] die von anderen Organismen stammen, werden durch das Plasmid pKIMP-UAUC zur Verfügung gestellt, sodass den Zellen schließlich nur die Chorismat-Mutase-Aktivität fehlt.^[45]

Aufgrund des fehlenden Enzyms können die KA12/pKIMP-UAUC-Bakterien nicht wachsen, sofern nicht zusätzlich Tyrosin und Phenylalanin im Wachstumsmedium zur Verfügung gestellt werden. Den Fehler im Metabolismus dieser Bakterien kann man sich als eine Unterbrechung in einem Stromkreis vorstellen. Dieser genetische Defekt lässt sich beheben, indem man den Zellen ein zusätzliches Plasmid, das eine natürliche Chorismat-Mutase kodiert, zur Verfügung stellt. Alternativ kann aber auch eine Mutante des natürlichen Enzyms oder sogar ein unverwandtes Polypeptid in die Zellen eingeführt werden. Sollte das Protein in der Lage sein, die Umwandlung von Chorismat in Prephenat zu katalysieren, dann könnten die Zellen selbstständig eigenes Tyrosin und Phenylalanin herstellen und demzufolge unter Selektionsbedingungen wachsen (d.h. ohne Zugabe von Tyrosin und Phenylalanin zum Medium). Falls das Protein die Umlagerung nicht beschleunigt, wäre eine Zellvermehrung absolut unmöglich. Daher kann man seine Aufmerksamkeit direkt auf funktionelle Klone richten, ohne Zeit und Mittel auf unproduktive Bibliotheksmitglieder verschwenden zu müssen. Darüber hinaus ist in diesem System die Verbindung zwischen dem gewünschten Phänotyp (einem Katalysator, der das Wachstum unter Selektionsbedingungen ermöglicht) und dem dafür verantwortlichen Genotyp (dem Gen, das den Katalysator kodiert) automatisch gegeben, wodurch eine einfache Isolierung, Diversifizierung und weitere Evolution von interessanten Varianten der Genbank möglich ist.

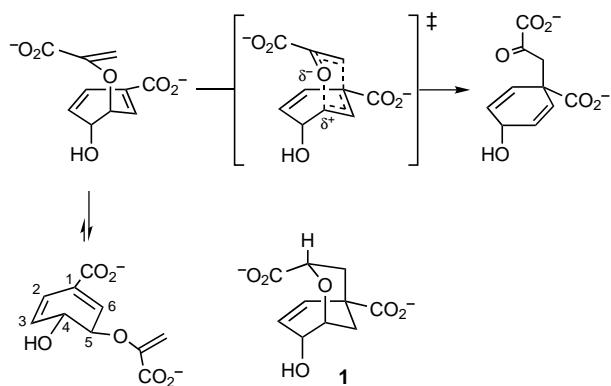
Da bereits die Verschiebung von katalytischen Resten um wenige Zehntel Angström den Unterschied zwischen vollständiger und absolut keiner Aktivität bedeuten kann, ist die enzymatische Aktivität ein extrem strenges Kriterium für erfolgreiches Design. Zieht man dies in Betracht, dann können zu vielen der oben aufgeworfenen Fragen wertvolle Antworten erhalten werden, indem man intelligent entworfene Proteinbibliotheken nach aktiven Chorismat-Mutasen durchsucht. Das Potential dieser Strategie in Forschungsgebieten, die von der mechanistischen Enzymologie bis hin zum Proteindesign reichen, lässt sich am besten durch einige Beispiele demonstrieren.

3.2. Mechanistische Untersuchungen

3.2.1. Katalytisch aktive Aminosäurereste

Unser Verständnis für die Umlagerung von Chorismat beruht zu großen Teilen auf mechanistischen Untersuchungen der unkatalysierten Reaktion^[49–52] und auf Computersimulationen.^[53–59] Für die Umlage-

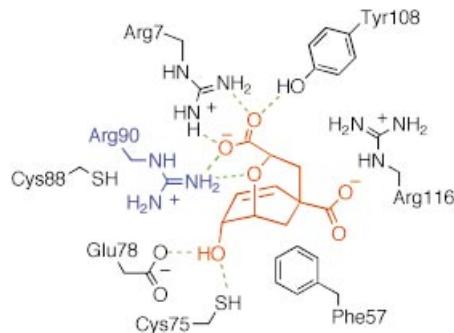
rung muss Chorismat, das in Lösung normalerweise in einer pseudodiäquatorialen Konformation vorliegt,^[51] zunächst eine pseudodioxiale Konformation einnehmen (Schema 3). Anschließend findet die exergonische Reaktion^[60] über einen Übergangszustand mit sesselähnlicher Geometrie statt.^[50] Der Vorgang scheint konzertiert, wenn auch asynchron abzulaufen, wobei die Bildung der C-C-Bindung zeitlich verzögert zum Bruch der C-O-Bindung erfolgt.^[49]



Schema 3. Chorismat bevorzugt in Lösung die pseudodiäquatoriale Konformation. Um zum pericyclischen Übergangszustand zu gelangen, muss es die ungünstigere pseudodioxiale Konformation einnehmen. Die konformativ eingeschränkte oxabicyclische Dicarbonsäure **1**, die den Übergangszustand nachahmt, ist ein wirkungsvoller Inhibitor natürlicher Chorismat-Mutassen.^[62] Antikörper, die mit Hilfe dieser Verbindung erzeugt wurden, katalysieren die Reaktion ebenfalls, wenn auch 100- bis 10000-mal schlechter als ihre natürlichen Pendants.^[39, 41]

Auch der Übergangszustand der enzymatischen Reaktion weist eine sesselähnliche Geometrie auf.^[61] Konformativ eingeschränkte Verbindungen, die diese Struktur nachahmen, wie die oxabicyclische Dicarbonsäure **1** (Schema 3), sind gute Inhibitoren für Chorismat-Mutassen.^[62–64] Die Art, wie ein Protein einen solchen Hochenergiezustand stabilisiert, ist seit langem Gegenstand zahlreicher Debatten. Vor wenigen Jahren wurden Schweratom-Isotopeneffekte zur Charakterisierung der Struktur des an BsCM gebundenen Übergangszustandes herangezogen.^[65] Ein sehr großer ¹⁸O-Isotopeneffekt an O(5) (ca. 5 %), der Stelle der Bindungsspaltung, und ein kleiner, gewöhnlicher ¹³C-Isotopeneffekt an C(1) (ca. 0.6 %), der Stelle der Bindungsbildung, belegen, dass der Umlagerungsschritt geschwindigkeitsbestimmend für dieses Enzym ist. Außerdem bestätigen die Resultate den Verlauf der enzymatischen Reaktion über einen konzertierten, aber asymmetrischen Übergangszustand. Ein Vergleich mit theoretisch ermittelten Isotopeneffekten, die aus Becke3LYP/6-31G*-Rechnungen erhalten wurden,^[65] legt jedoch nahe, dass der Übergangszustand im Enzym unter Umständen stärker polarisiert ist, d.h. die C-O-Bindungsspaltung weiter fortgeschritten und die C-C-Bindungsbildung weniger weit gediehen ist, als dies in Lösung der Fall ist.

Eine plausible Erklärung für die größere Polarisierung des gebundenen Übergangszustandes lässt sich aus der Untersuchung der Reste herleiten, die die aktive Tasche von BsCM bilden (Schema 4).^[31, 32] So befindet sich die positiv geladene Guanidiniumgruppe von Arg90 in einer Entfernung zum



Schema 4. Die aktive Tasche von BsCM: Das Übergangszustandsanalogon **1** (rot) wird durch ein ausgedehntes Netz aus Wasserstoffbrücken und elektrostatischen Wechselwirkungen gebunden. Die Seitenkette von Arg 90 (blau) ist so orientiert, dass dessen Guanidiniumgruppe im H-Brücken-Abstand zum Ether-Sauerstoffatom des Liganden angeordnet ist.

Ether-Sauerstoffatom des gebundenen Substrats, die für eine Wasserstoffbrücke geeignet ist. Die negative Partialladung, die sich im Übergangszustand am Ether-Sauerstoffatom aufbaut, kann dadurch elektrostatisch stabilisiert werden. In der Regel würde man den Beitrag eines solchen Restes zur Katalyse mittels ortsspezifischer Mutagenese ermitteln. Beispielsweise könnte man Arg 90 durch Lysin oder Methionin ersetzen, um herauszufinden, welchen Effekt man erzielt, wenn das Guanidiniumion durch ein anderes Kation ersetzt wird bzw. die positive Ladung völlig entfernt wird. Mit dem vorhandenen Selektionssystem ist jedoch eine gleichzeitige Bewertung aller 20 natürlichen Aminosäuren an Position 90 möglich. Diejenigen Aminosäuren, die ein aktives Enzym liefern, können schnell durch Sequenzierung der Gene, die die Chorismat-Mutase-Defizienz kompensieren können, identifiziert werden. Bei der Durchführung dieses Experiments wurden bemerkenswerte Resultate erhalten:^[45] Ausschließlich Arginin an Position 90 führte zu aktiven Enzymen. Sogar der konservative Austausch von Arginin gegen Lysin führte zu einem Enzym, welches *in vivo* inaktiv war. Die kinetische Charakterisierung repräsentativer Mutanten ergab, dass die Entfernung der Argininseitenkette die katalytische Effizienz um fünf bis sechs Größenordnungen herabsetzt.^[38]

Falls die Rolle von Arg 90 tatsächlich in der elektrostatischen Stabilisierung der sich im Übergangszustand am Ether-Sauerstoffatom entwickelnden negativen Ladung besteht, könnte man sich fragen, warum man keine Komplementation beobachtet, wenn Arginin durch einen anderen positiv geladenen Rest wie Lysin ersetzt wird. Eine schlechtere Produktion der Arg90Lys-Variante in *E. coli* ist eine mögliche Erklärung hierfür. Der Wachstumsphänotyp hängt direkt von der insgesamt in der Zelle vorhandenen Chorismat-Mutase-Aktivität ab, die sowohl von der spezifischen Aktivität als auch von der Konzentration des Katalysators bestimmt wird. Sollte das Enzym zwar katalytisch aktiv, aber nur in äußerst geringer Konzentration vorhanden sein, würde kein Wachstum zu beobachten sein. Auch wenn es den Anschein hat, dass die BsCM-Mutanten im Selektionssystem in ähnlichen Mengen produziert werden,^[45] sind Ausnahmen möglich.

Eine vielleicht interessantere Erklärung ist, dass die aktive Tasche des Enzyms intrinsische strukturelle Zwänge auf die

möglichen Substitutionen ausüben könnte. Die Lysinseitenkette ist kürzer als die von Arginin, und deshalb ist es möglich, dass sie nicht weit genug in die aktive Tasche hineinreichen kann, um die Ammoniumgruppe so zu positionieren, dass diese eine H-Brücke mit dem kritischen Ether-Sauerstoffatom bildet. Wenn dies zutrifft, dann sollten möglicherweise zusätzliche Mutationen innerhalb der aktiven Tasche den Ersatz von Arg 90 ohne Funktionsausfall gestatten.

Diese Möglichkeit wurde von uns getestet, indem wir gleichzeitig zwei Reste – Arg 90 und Cys 88 – der Bindungsstelle von BsCM mutierten.^[45] Die Seitenkette von Cys 88 ist ungefähr 7 Å vom gebundenen Liganden entfernt, lagert sich aber entlang der Seitenkette von Arg 90 und beeinflusst dadurch möglicherweise deren Konformation (Schema 4). Die kombinatorische Mutagenese dieser beiden Reste liefert 400 Varianten, die sich schnell mit Hilfe des Chorismat-Mutase-defizienten Selektionsstammes bewerten lassen. Die in diesem Experiment erhaltenen Resultate erwiesen sich als recht informativ (Abbildung 4).^[45] Sie deuten darauf hin, dass

benötigt, vorausgesetzt, Cys 88 wird gegen einen Lysinrest ausgetauscht.

Kinetische Untersuchungen der Cys88Lys/Arg90Ser-Doppelmutante ergaben, dass ein Lysinrest in Position 88 drei der fünf bis sechs Zehnerpotenzen an katalytischer Effizienz wiedergewinnt, welche durch die Entfernung des Argininrestes in Position 90 verloren werden.^[38] Kristallographische Studien dieser Doppelmutante ergaben, dass die Seitenkette von Lys 88 in die aktive Tasche hineinragt und sich ihre Ammoniumgruppe analog der Arg90-Guanidiniumgruppe des Wildtypenzyms in H-Brücken-Entfernung zum Ether-Sauerstoffatom des gebundenen Liganden hin ausrichtet (Abbildung 5).^[38] Somit konnte durch kombinatorische Mutagenese und Selektion eine veränderte aktive Tasche erhalten werden, die eine neue Lösung für das chemische Problem der effizienten Umlagerung von Chorismat in Prephenat beiträgt.

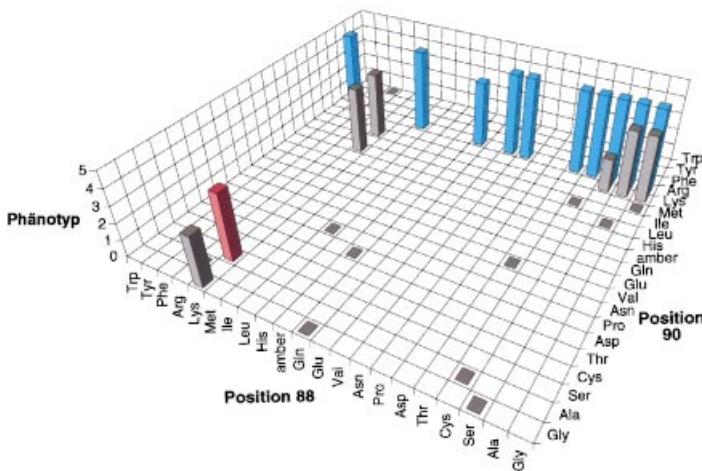


Abbildung 4. Ergebnis eines kombinatorischen Mutagenese- und Selektionsexperiments, das die gleichzeitige Randomisierung der Positionen 88 und 90 in BsCM zum Gegenstand hatte.^[45] Die Aminosäuren sind entsprechend dem Volumen ihrer Seitenkette geordnet. Aktive Varianten, deren Sequenzen aufgeklärt wurden, und einige inaktive Kontrollvarianten sind mit ihrer In-vivo-Aktivität (ihrem Phänotyp) dargestellt. Die phänotypische Bewertung spiegelt die Größe der Kolonien wieder und reicht von 0 (kein Wachstum) bis 5 (wildtypartiges Wachstum). Blaue Säulen stehen für Varianten mit dem Wildtyp-Arg 90. Die rote Säule zeigt das Ergebnis mit der Cys88Lys/Arg90Ser-Doppelmutante, die ausführlicher untersucht wurde (siehe Abbildung 5).

Cys88 nicht wesentlich für die Katalyse ist. Es kann gegen große, mittelgroße und kleine Aminosäuren ausgetauscht werden; sofern sich ein Argininrest in Position 90 befindet, werden aktive Enzyme erhalten. Ersetzt man Cys 88 durch einen kleinen Rest (beispielsweise Glycin oder Alanin), fanden sich auch aktive Doppelmutanten mit Arg 90Lys als zusätzlicher Mutation, die beweisen, dass die Guanidiniumgruppe nicht absolut notwendig für die Katalyse ist. Das vielleicht interessanteste Ergebnis dieser Experimente ist, dass man überhaupt kein kationisches Zentrum an Position 90

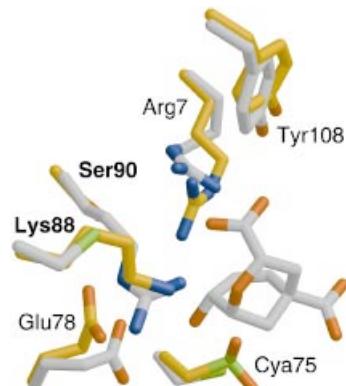
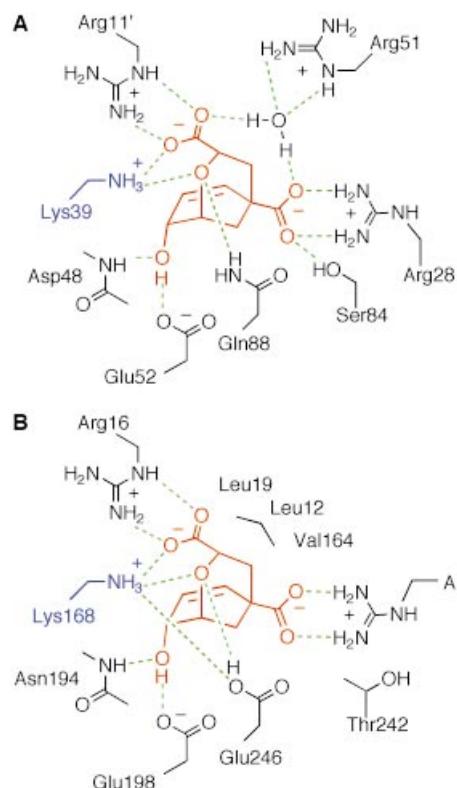


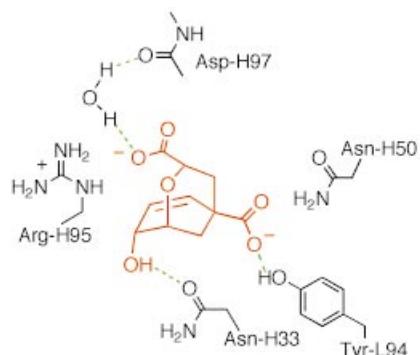
Abbildung 5. Überlagerung der Reste der Cys88Lys/Arg90Ser-Doppelmutante ohne gebundenen Liganden (Kohlenstoffatome gelb, mit Angabe der Reste) und der Reste des mit dem Inhibitor 1 komplexierten BsCM-Wildtyps (Kohlenstoffatome grau).^[38] Cys 75 ist in der kristallisierten Mutante partiell zum Sulfinat (Cya 75) oxidiert.

Die Resultate dieser Selektionsexperimente sind für den Mechanismus der enzymatischen Reaktion insofern bedeutend, als sie die kritische Rolle eines kationischen Zentrums für die Chorismat-Mutase-Aktivität untermauern. Es wurden keine aktiven Katalysatoren erhalten, wenn ein kationisches Zentrum in der Nachbarschaft des Ether-Sauerstoffatoms des Substrats fehlte. Der Schluss, dass ein kationisches Zentrum für eine hohe Chorismat-Mutase-Aktivität essentiell ist, kann daher wesentlich überzeugender gezogen werden, als dies im Anschluss an ein gewöhnliches Mutageneseexperiment, bei dem nur einzelne Substitutionen Berücksichtigung gefunden hätten, möglich gewesen wäre.^[66] Selbstverständlich können wir nicht ausschließen, dass in Zukunft durch weitergehende Mutagenese andere Anordnungen von Aminosäureresten erhalten werden, welche auch ohne ein kationisches Zentrum katalytisch aktiv sind. Es muss in diesem Zusammenhang allerdings gesagt werden, dass sowohl die Chorismat-Mutase von *E. coli*^[33] als auch die der Hefe^[34, 35] einen kationischen Lysinrest in einer ähnlichen Position in ihrer jeweiligen aktiven Tasche aufweisen (Schema 5), auch wenn ihre Tertiärstrukturen ansonsten nicht mit der von BsCM verwandt sind. Durch Mutagenese wurde gezeigt, dass dieser Lysinrest in EcCM für die Katalyse notwendig ist.^[67, 68]



Schema 5. Die aktiven Taschen von EcCM (A) und ScCM (B), komplexiert mit dem Übergangszustandsanalogon **1**. Die dem Arg-90-Rest von BsCM entsprechenden Lys-Reste sind blau dargestellt.

Ebenfalls sollte erwähnt werden, dass der 10⁴-mal weniger aktive katalytische Antikörper 1F7^[39, 40] über kein strategisch ähnlich platziertes kationisches Zentrum verfügt (Schema 6).^[42] Dies erstaunt nicht allzu sehr, wenn man die



Schema 6. Die aktive Tasche des katalytischen Antikörpers 1F7 mit gebundenem **1**. Arg-H95 befindet sich nicht im H-Brücken-Abstand zum Ether-Sauerstoffatom des Liganden.^[42]

Struktur des Übergangszustandsanalogons betrachtet, das zur Erzeugung dieses Katalysators verwendet wurde (siehe Schema 3).^[39] Trotzdem lassen sich aus dieser Beobachtung Strategien zur Produktion von Antikörpern mit wesentlich höherer Aktivität ableiten. Beispielsweise ließen sich zusätzliche negative Ladungen in ein Hapten einbringen, um neue Antikörper zu generieren, welche das katalytisch entscheidende kationische Zentrum enthalten. Alternativ dazu könn-

te man ein kationisches Zentrum auch direkt in die aktive Tasche von 1F7 mittels gerichteter Mutagenese einbauen. Ein dritter Ansatz, der auf Selektion basiert, könnte am Ende dennoch die erfolgreichste Strategie zur Optimierung der Aktivität des Antikörpers sein.

Die Aktivität des Fab-Fragmentes von 1F7 im Cytoplasma eines Chorismat-Mutase-defizienten Hefestamms wurde bereits nachgewiesen.^[43, 44] Bei einem genügend hohen Expressionsniveau ist der katalytische Antikörper in der Lage, das fehlende Enzym zu ersetzen und den Stoffwechseldefekt in Ansätzen zu beheben.^[44] Es ist daher durchaus möglich, dass man unter Selektionsdruck Varianten mit höherer katalytischer Effizienz ausfindig macht. Erste Resultate aus solchen Experimenten sehen viel versprechend aus.^[69]

3.2.2. Zufällig verkürzte Proteinvarianten

Das Fehlen von Informationen zur Struktur eines Proteins erschwert die Identifizierung geeigneter Reste für eine Mutationsanalyse. In diesen Fällen kann die Zufallsmutagenese in Verbindung mit Selektion eine hervorragend geeignete Analysemöglichkeit sein. Wir haben uns dieser Methode bedient, um herauszufinden, in welchem Maße die 17 C-terminalen Aminosäuren zur enzymatischen Aktivität von BsCM beitragen.^[70]

BsCM ist 127 Aminosäuren lang. Die ersten 115 Reste nehmen eine wohldefinierte Struktur ein, die in einer 3₁₀-Helix nahe der aktiven Tasche endet (Abbildung 3 A).^[31, 32] Die weiteren Reste sind in der Kristallstruktur des Enzyms ohne gebundenes Substrat nicht aufgelöst.^[31] Dies legt den Schluss nahe, dass sie entweder hochgradig ungeordnet sind oder mehrere Konformationen annehmen können. Kristallographisch kann nach der Bindung eines Liganden an BsCM eine höhere Ordnung im C-Terminus festgestellt werden.^[31] Aus Fourier-Transformations(FT)-IR-spektroskopischen Untersuchungen wurden ebenfalls ligandeninduzierte Konformationsänderungen abgeleitet.^[71] Diese Befunde lassen darauf schließen, dass der flexible C-Terminus eine gewisse Rolle für die Enzymfunktion spielt. Möglicherweise dient er als eine Art Deckel für die aktive Tasche, der das Substrat von der umgebenden Lösung abschottet und/oder zur Stabilisierung des Übergangszustandes beiträgt.

Um diesen Fragestellungen nachzugehen, haben wir eine Strategie entwickelt, die den C-Terminus des Enzyms nach einem Zufallsprinzip verkürzt und anschließend direkt auf die funktionellen Klone im KA12/pKIMP-UAUC-System selektiert.^[70] Wir setzten eine PCR-Methode (PCR = Polymerase-Kettenreaktion) mit zwei partiell randomisierten Oligonukleotiden ein, um zwei Genbanken zu konstruieren. In diesen waren die BsCM-Codons für die Reste 116–127 oder 111–127 (die kristallographisch beobachtete 3₁₀-Helix eingeschlossen) randomisiert. Die Genbanken waren so entworfen, dass die Häufigkeit von Stopps an jedem mutagenisierten Codon optimiert war. Auf diese Weise wurde eine ganze Serie zufällig verkürzter Proteine erhalten. Weil nur funktionelle Gene die Selektion überleben, können entbehrliche Teile des kodierten Proteins rasch und direkt erkannt werden. Die Methode eignet sich außerdem zur Erforschung nichtnatürlicher, alternativer Sequenzen, sodass man auch Reste, die

wesentliche strukturelle oder mechanistische Aufgaben erfüllen könnten, entdecken kann.

Nach der Fertigstellung und der Transformation in den bakteriellen Selektionsstamm enthielten beide Genbanken 10^4 – 10^5 Mitglieder. Fast alle Mitglieder der Genbank mit den mutagenisierten 12 C-terminalen Resten waren auf dem Selektionsmedium lebensfähig. Diese Resultate belegen, dass keiner der 12 C-terminalen Reste für die enzymatische Aktivität absolut notwendig ist. Ihre Mutation und sogar die Entfernung der letzten 11 Reste hat keinen wesentlichen Einfluss auf die In-vivo-Aktivität. Dagegen waren nur ca. 25 % der Genbankmitglieder in der Lage, auf Selektionsplatten zu wachsen, wenn die letzten 17 Reste gleichzeitig mutagenisiert wurden. Die zusätzlichen fünf Aminosäuren bilden eine wohldefinierte 3_{10} -Helix und steuern Wechselwirkungen zum restlichen Protein und seinen Liganden bei. Sequenzanalysen ergaben, dass Mutationen an diesen Positionen toleriert werden, aber dass die Wildtypreste Lys111, Ala112, Leu115 und Arg116 deutlich bevorzugt sind. Darüber hinaus wurden keine aktiven Enzyme mit weniger als 116 Aminosäuren gefunden. Offensichtlich können die Reste der 3_{10} -Helix nicht entfernt werden, ohne das gesamte Enzym zu zerstören.

Wie die kinetische Charakterisierung einiger ausgewählter BsCM-Varianten ergab, hat eine Verkürzung oder die Mutation des C-Terminus nur geringen Einfluss auf die Wechselzahl k_{kat} des Enzyms (Tabelle 1). Wenn Chorismat in der aktiven Tasche der Varianten gebunden ist, wird es annähernd so effizient zu Prephenat umgewandelt wie durch Wildtyp-BsCM. Dennoch sinkt der k_{kat}/K_M -Wert ganz erheblich (Tabelle 1; K_M = Michaelis-Konstante). Dieses Ergebnis weist darauf hin, dass der C-Terminus zur enzymatischen Effizienz durch gleichartige Bindung des Substrats und des Übergangszustands beiträgt, auch wenn er nicht direkt an der chemischen Umwandlung des gebundenen Liganden beteiligt ist.

Tabelle 1. Katalytische Parameter der verkürzten BsCM-Varianten.^[70]

BsCM-Variante	C-terminale Aminosäuresequenz	k_{kat} [s ⁻¹]	k_{kat}/K_M [M ⁻¹ s ⁻¹]
5–8	KAVVLR	26	2.8×10^3
1–3	NSNVLRP	30	2.0×10^3
V–7	KAVVLLT	26	1.6×10^3
5–11	KAVVLRPN	23	5.6×10^3
Wildtyp	KAVVLRPDLSLTKNTEL	46	6.9×10^5

3.3. Strukturuntersuchungen

Selektionsmethoden eignen sich bestens zur Untersuchung von Faktoren, die die Struktur oder Stabilität eines Proteins beeinflussen. Der Informationsgehalt einer Proteinsequenz ist typischerweise sehr uneinheitlich. Viele Reste sind äußerst tolerant gegenüber einer breiten Palette von Substitutionen, andere wiederum können nicht ohne erhebliche Auswirkungen auf die Faltung oder Funktion verändert werden.^[72, 73] Deshalb ist die Untersuchung der Beziehung zwischen Sequenz und Struktur eines Proteins ein kombinatorisches Problem, das leistungsfähige Methoden zur gleichzeitigen Bewertung vieler verschiedener möglicher Sequenzen erfor-

dert. Experimentell erreichbare Populationen von 10^{10} Molekülen sind im Vergleich zu den 20^{100} ($= 10^{130}$) Möglichkeiten für ein Polypeptid von 100 Aminosäuren Länge eigentlich verschwindend klein. Dennoch können mit Genbanken dieser Größe statistisch wertvolle Einsichten hinsichtlich der in Proteinen intrinsisch bevorzugten Sekundärstrukturen,^[74, 75] hinsichtlich der Einschränkungen für die Segmente, die die Sekundärstrukturelemente verbinden,^[76–78] und hinsichtlich der für die optimale Packung der Reste im Proteininnern notwendigen Anordnungen erhalten werden.^[79–82]

3.3.1. Einschränkungen für interhelicale Schleifen

Der Einfluss, den interhelicale Schleifen auf die Struktur eines Proteins haben, wurde an verhältnismäßig einfachen Vierhelixbündel-Proteinen erforscht, indem kombinatorische Genbanken erzeugt und auf funktionelle Varianten durchgemustert wurden. So mutagenisierten Hecht und Mitarbeiter eine drei Reste umfassende interhelicale Schleife in Cytochrome b_{562} .^[76] Alle 31 isolierten Varianten der Zufallsgenbank nahmen eine stabile, der natürlichen Faltung ähnliche Struktur ein, wie anhand der für Hämkomplexe von Cytochrome b_{562} charakteristischen Rotfärbung geschlossen wurde. Ähnliche Resultate ergaben Versuche mit mutagenisierten Schleifen in ROP,^[77, 83] einem dimeren RNA-bindenden Vierhelixbündel-Protein, obwohl die genaue Analyse einzelner Varianten auf einige Veränderungen bezüglich der Thermostabilität hindeutete.^[84] Insgesamt legen die genannten Studien nahe, dass für interhelicale Segmente nur wenige Einschränkungen bezüglich ihrer Sequenz oder Länge bestehen.

Wie allgemeingültig sind diese Schlussfolgerungen? Sind sie auch auf komplexere Proteine übertragbar? Gibt es schwerer fassbare Eigenschaften, die die Schleifenbildung beeinflussen? Die Selektion auf katalytische Aktivität, ein besonders empfindlicher Test für strukturelle Unversehrtheit, eröffnet die Möglichkeit, subtilere Sequenzvorlieben ans Licht zu bringen, als dies einfache Screeningverfahren vermögen. Wir testeten diese Annahme am Beispiel des EcCM-Enzyms (Abbildung 3B).^[85] Die beiden identischen Polypeptide, die dieses Helixbündelprotein bilden, setzen sich aus drei Helices zusammen, die durch zwei Schleifen miteinander verbunden sind (Abbildung 6A).^[33] Für die Analyse wurde die Schleife L2 ausgewählt, die die Helices H2 und H3 verbindet und am weitesten von der aktiven Tasche entfernt ist.

Zunächst wurden drei dem Lösungsmittel exponierte Reste der Schleife (Ala65-His66-His67, Abbildung 6B) randomisiert und katalytisch aktive Varianten durch In-vivo-Selektion identifiziert. Das anfängliche Ensemble enthielt 8000 (2^3) unterschiedliche Mitglieder, von denen mehr als 63 % die Chorismat-Mutase-Defizienz des bakteriellen Wirtstamms korrigierten, obgleich die Wachstumsgeschwindigkeiten stark voneinander abwichen. Der große Anteil von komplementierenden Klonen bestätigt eindeutig, dass kaum Einschränkungen für diesen Schleifenabschnitt bestehen. Trotzdem lässt sich aus den unterschiedlichen Wachstumsgeschwindigkeiten der komplementierenden Klone folgern, dass Substitutionen in der Schleife entweder die Proteinproduktion, die Stabilität oder die katalytische Aktivität beeinflussen können. In der

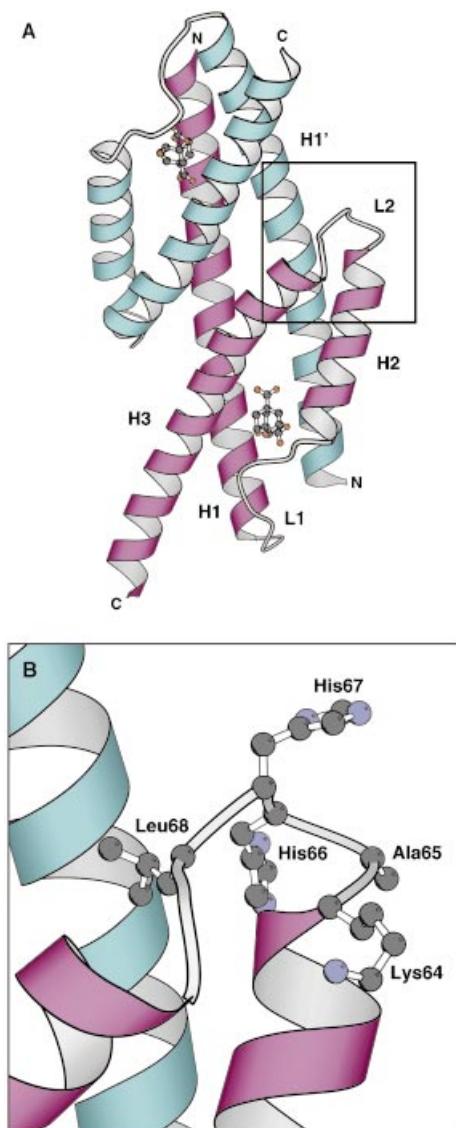


Abbildung 6. Bändermodell (A) und die Schleife L2 (B) von EcCM. Die Zufallsmutagenese der Schleife L2 und die anschließende In-vivo-Selektion auf Chorismat-Mutase-Aktivität ergaben, dass für die dem Lösungsmittel exponierten Schleifenreste kaum Einschränkungen für die Sequenz bestehen, abgesehen von einer leichten Bevorzugung hydrophiler Aminosäuren.^[85] Dagegen ist aufgrund langreichweiterer tertiärer Wechselwirkungen in Position 68 eine hydrophobe aliphatische Aminosäure unbedingt erforderlich.

Tat ergibt eine genaue Sequenzanalyse der am schnellsten wachsenden Klone eine statistisch relevante Bevorzugung hydrophiler Reste an diesen Lösungsmittel-exponierten Stellen. Obwohl auch Prolin-enthaltende Schleifen mit einem konformativ veränderten Rückgrat toleriert werden, traten diese Sequenzen lediglich in Klonen mit verminderter Aktivität auf.

Verändern benachbarte Sekundärstrukturelemente das zulässige Austauschmuster von Aminosäureresten in der Schleife? Diese Frage konnte durch die Randomisierung des gleichen Ala65-His66-His67-Dreierpeptids in Verbindung mit Lys64, das Teil der angrenzenden Helix H2 ist (Abbildung 6), untersucht werden. Eine Genbank, die alle 160000 (20^4) möglichen Mitglieder umfasst, wurde aufgebaut und der

Selektion unterworfen. Wieder war ein großer Teil der Sequenzen (>50%) in der Lage, die natürliche Sequenz funktionell zu ersetzen; somit limitiert die Sekundärstruktur an Position 64 weder das Spektrum an zulässigen Seitenketten an dieser Stelle noch beeinflusst dieser Rest die Vorliebe für bestimmte Reste in der benachbarten Schleifenregion.

Wesentlich andere Resultate erhielt man, wenn der im Protein vergrabene Schleifenrest Leu68 zusammen mit Ala65-His66-His67 randomisiert wurde. Im Unterschied zu den anderen Schleifenresten befindet sich die Seitenkette von Leu68 an der Grenzfläche zwischen L2 und H1' (Helix H1 der anderen Untereinheit) und ist deshalb vom Lösungsmittel abgeschirmt (Abbildung 6). In diesem Fall komplementierten lediglich 6–7% der 160000 möglichen Sequenzen die Chorismat-Mutase-Defizienz. Die Sequenzanalyse ergab, dass an Position 68 eine hydrophobe aliphatische Aminosäure besonders unverzichtbar ist. Daraus lässt sich folgern, dass weitreichende Tertiärstrukturwechselwirkungen zwischen Resten in einer interhelicalen Schleife und Resten an anderen Stellen im Protein, im Unterschied zum schwachen Einfluss eines an einer Sekundärstruktur beteiligten Restes, die zulässigen Sequenzsubstitutionen stark einschränken können.

Die Studie mit EcCM unterstreicht die wichtigsten Vorteile der genetischen Selektion gegenüber Screeningmethoden. Erstens können weitaus größere Genbanken mit wenig zusätzlichem Aufwand untersucht werden. Das Durchmuster von 10^5 unterschiedlichen Klonen ohne eine geeignete spektroskopische Methode wäre ein extrem mühseliges Unterfangen, wohingegen die Selektion die gesamte Mischung auf einfache und bequeme Art und Weise in aktive, weniger aktive und inaktive Enzyme trennt. Zweitens erlaubt es das strengere Kriterium, dass selektierte Varianten katalytisch aktiv sein müssen, auch nicht offensichtliche Besonderheiten zu erkennen, wie die statistisch relevante Bevorzugung von hydrophilen Resten an Lösungsmittel-exponierten Stellen, oder auch Schleifen mit veränderter Proteinrückgratkonformation zu identifizieren. Diese Feinheiten in der Proteinstruktur würden bei einer Methode mit weniger strengen Anforderungen für funktionelle Varianten möglicherweise übersehen.

3.4. Änderung der Proteintopologie

Alle Proteine setzen sich aus einem beschränkten Satz von Sekundärstrukturelementen zusammen. Aber noch beginnen wir gerade erst zu verstehen, wie wir die einzelnen Elemente kombinieren können, um wohldefinierte Tertiär- und Quartärstrukturen zu erhalten. Evolutionäre Strategien, die es gestatten, eine riesige Zahl von Alternativen auf der Basis eines selektionierbaren Phänotyps wie der katalytischen Aktivität gleichzeitig zu bewerten, sind hierbei allerdings eine große Hilfe.

3.4.1. Neue Quartärstrukturen

Der Umbau von vorhandenen Proteingerüsten ist ein realisierbarer erster Schritt hin zum größeren Ziel – dem De-novo-Design funktioneller Enzyme. Die Änderung des

Oligomerisierungszustand verschiedener Proteine, wie die Umwandlung von Monomeren in Multimere^[86] oder von Multimeren in Monomere,^[87–92] hat spezielle Aufmerksamkeit erfahren. Solche Systeme könnten helfen, die evolutionären Ursprünge von multimeren Proteinen aufzuklären.^[93, 94]

Ein oligomeres Protein in ein katalytisch funktionelles Monomer umzuwandeln ist vermutlich besonders schwierig, wenn die einzelnen Polypeptidketten entwirrt werden müssen, wie im Falle der homodimeren EcCM mit ihren kompliziert ineinander verschlungenen Untereinheiten.^[33] Dennoch schien es durch das Einfügen einer flexiblen Scharnier-Schleife („hinge-loop“) in die lange Helix H1, die sich über die gesamte Länge des EcCM-Dimers erstreckt, möglich, die Quartärstruktur des Enzyms ändern zu können. Die wichtige Rolle, die Scharnier-Schleifen für die Festlegung des Oligomerisierungszustandes spielen, wurde bereits beschrieben.^[93, 94] Obwohl Reste beider Polypeptidketten die aktive Tasche von EcCM bilden, könnte ein solches Scharnier es der N-terminalen Hälfte von H1 ermöglichen, sich auf sich selbst zurückzufalten und dabei das zweite Polypeptid so zu ersetzen, dass dadurch ein monomeres, katalytisch aktives Vierhelixbündel entsteht (Abbildung 7). Es erscheint ebenfalls möglich, dass, abhängig von der Schleifenlänge und -zusammensetzung, auch andere Oligomerisierungszustände eingenommen werden könnten.

Zur Prüfung dieser Möglichkeiten wurden Zufallsabschnitte mit einer Länge von vier bis sieben Resten in die Mitte der Helix H1 von EcCM eingefügt.^[95] Die einzelnen Genbanken (im Folgenden mit L4–L7 bezeichnet) haben eine maximale theoretische Diversität von 160 000 (2^0), ca. 3.2×10^6 (2^4), ca. 6.4×10^7 (2^6) bzw. ca. 1.28×10^9 (2^7) verschiedenen Mitgliedern. In allen Fällen lieferte die Transformation der Chorismat-Mutase-defizienten Bakterien ungefähr 10^7 Klone. Für die Genbanken L4 und L5 ergibt sich damit eine mehrfache Abdeckung aller Möglichkeiten; der Sequenzraum der Genbank L6 ist zu ca. 10 % und derjenige der Genbank L7 zu ca. 1 % abgedeckt.

Unter Selektionsbedingungen wurden keine funktionellen Klone in der Genbank L6 gefunden, aber 0.05 %, 0.002 % und 0.5 % der L4-, L5- bzw. L7-Genbankmitglieder hoben den genetischen Defekt auf. Einige selektionierte aktive Enzyme wurden biochemisch charakterisiert (Schema 7). Die Einführung von Segmenten mit vier oder sieben Aminosäuren in die Mitte der Helix H1 führte zu monomeren Enzymen mit nahezu Wildtypaktivität; allerdings waren diese Proteine nicht stabil und neigten zur Aggregation. Bei der Einführung eines Abschnitts aus fünf Aminosäuren (Cys-Phe-Pro-Trp-Asp) erhielt man hingegen eine gut handhabbare, hexamere Chorismat-Mutase (wie durch analytische Ultrazentrifugation gezeigt wurde), die ungefähr 200-mal weniger aktiv ist als das Wildtypprotein.

Vergleicht man die Eigenschaften der hexameren Spezies mit denjenigen der Monomere, dann liegt die Vermutung nahe, dass die Proteinstabilität eine wichtige treibende Kraft in der Evolution von oligomeren Proteinen ist.^[95] Jedoch sind

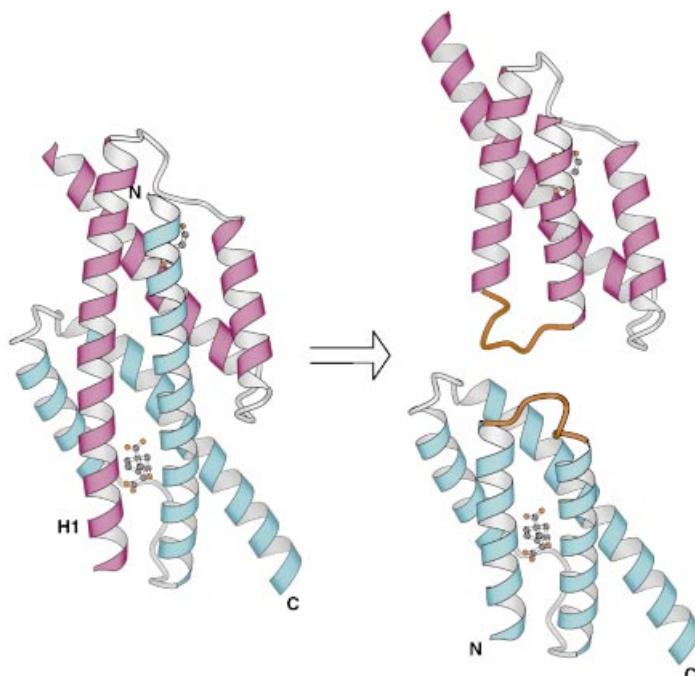
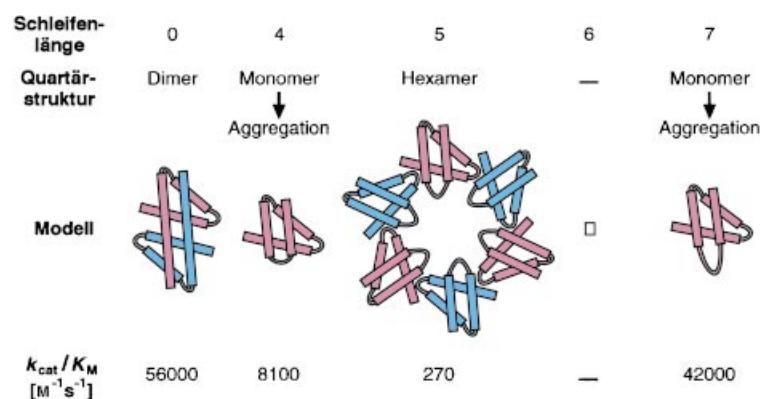


Abbildung 7. Topologische Umwandlung der dimeren EcCM in ein Monomer: Die Einführung einer flexiblen Schleife in die dimerdurchspannende Helix H1 würde es der N-terminalen Hälfte der Helix ermöglichen, sich auf sich selbst zurückzufalten, um so das andere Polypeptid zu ersetzen und dadurch einen funktionellen Katalysator mit Vierhelixbündel-Struktur zu bilden.



Schema 7. Modelle der strukturellen Organisation von EcCM-Varianten mit Scharnier-Schleifen sowie Angabe der Zahl der Reste in der eingefügten Scharnier-Schleife, des Oligomerisierungszustands der aktiven Varianten und der experimentell bestimmten k_{cat}/K_M -Werte.^[95]

noch weitere Arbeiten vonnöten, um die Organisation der Untereinheiten im Hexamer sowie den Einfluss der Quartärstruktur auf die Funktion aufzuklären.

Allgemein lässt sich festhalten, dass die Schleifenlänge vermutlich von weitaus größerer Bedeutung für die Struktur ist, als aufgrund anderer Experimente angenommen werden konnte.^[83] Die geringe Prozentzahl aktiver Varianten, die aus diesen Selektionsexperimenten gewonnen wurden, deutet darauf hin, dass relativ wenige Schleifensequenzen eine Änderung der Quartärstruktur zulassen, ohne die Struktur der aktiven Tasche zu beeinträchtigen. Darüber hinaus bedeutet das äußerst seltene Auftreten funktioneller Klone (z.B. 1 aus 50000 L5-Genbankmitgliedern), dass es extrem

schwierig gewesen wäre, aktive Enzyme allein mit Hilfe einer In-vitro-Screeningmethode für Chorismat-Mutase-Aktivität zu finden.

3.4.2. Stabile monomere Mutasen

Die topologische Umwandlung des EcCM-Homodimers in ein Monomer erfolgt zwar unter Erhaltung der aktiven Tasche des Enzyms, aber das resultierende Protein ist weniger stabil. Das Dimer scheint durch eine Ansammlung hydrophober Reste in einer Region zwischen den beiden aktiven Taschen stabilisiert zu werden;^[37] diese Wechselwirkungen fallen in einem Monomer weg. Vermutlich wird die Aggregation zusätzlich dadurch begünstigt, dass diese sonst abgeschirmten apolaren Gruppen dem Lösungsmittel ausgesetzt werden. Weil sich die stark geladene Bindungstasche aber im Innern des Helixbündels befindet,^[33] weist das Monomer selbst keinen klassischen hydrophoben Kern auf, und damit fehlt die übliche treibende Kraft für die Proteinfaltung.^[96]

Deshalb wurde ein thermostabiles Dimer als alternativer Ausgangspunkt für das Design einer stabilen monomeren Mutase ins Auge gefasst. Eine große Zahl von EcCM-Sequenzhomologen, manche von thermophilen Organismen, sind bekannt. Das hyperthermophile Archaeabakterium *Methanococcus jannaschii* beispielsweise produziert eine Chorismat-Mutase (MjCM), die 25 K stabiler als EcCM ist.^[37] Obwohl die Sequenzen lediglich zu 21 % identisch sind, sind sechs prominente Reste, die die aktive Tasche auskleiden, absolut konserviert, und die Enzyme haben ähnliche Aktivitäten. Da der hydrophobe Kern von MjCM dem von EcCM sehr ähnlich ist, müssen andere, von der Dimerisierungsregion weiter entfernte Wechselwirkungen für die zusätzliche Stabilität verantwortlich sein. Genau diese Wechselwirkungen sollten auch das gewünschte Monomer stabilisieren.

Varianten von MjCM mit veränderter Topologie wurden durch Einfügen einer flexiblen Schleife in die Mitte der Helix H1 hergestellt,^[97] wie im vorhergehenden Abschnitt für EcCM beschrieben (siehe Abbildung 7). Allerdings wurde die Schleife in diesem Fall nach dem bekannten Helix-Schleife-Helix-Motiv der *E.-coli*-Seryl-tRNA-Synthetase modelliert.^[98] Zwei Reste der Helix H1 (Leu22 und Lys23; die Nummerierung folgt der der homologen EcCM-Sequenz) wurden dupliziert, und eine Sequenz aus sechs randomisierten Resten wurde zwischen die verdoppelten Aminosäuren eingefügt. Außerdem wurden zwei Punktmutationen (Leu22a Glu und Ile79 Arg) als „negative“ Designelemente eingeführt, um die Tendenz zur Dimerisierung abzuschwächen und eine unerwünschte Aggregation von monomeren Varianten durch Verkleinerung der exponierten hydrophoben Oberfläche zu minimieren. Eine Bibliothek mit diesem Design umfasst ca. 6.4×10^7 (20^6) verschiedene Mitglieder. Die Bibliothek wurde auf genetischem Niveau zusammengesetzt, in den Chorismat-Mutase-defizienten *E.-coli*-Stamm transformiert ($>10^8$ Transformanten) und durch Selektion evaluiert.

Im Unterschied zum oben beschriebenen, auf festem Medium durchgeföhrten Selektionsexperiment ließ man die Genbank-Transformanten in Abwesenheit von Tyrosin und Phenylalanin in einer Flüssigmutter wachsen. Dadurch findet

ein direkter Wettbewerb um die Nährstoffe statt, und der kleine Teil der ursprünglichen Genbank, der funktionelle Chorismat-Mutasen kodiert (0.7 %), kann sich vermehren. Nach dreitägiger Selektion komplementierten mehr als 80 % der Klone den genetischen Mangel; dies entspricht einer über 100fachen Anreicherung.

Durch die Eliminierung von inaktiven sowie von schwach aktiven Klonen wird die Suche nach einem geeigneten Monomer wesentlich vereinfacht. Aber nicht die Topologie, sondern die Funktion war die Grundlage für die Selektion. Aus diesem Grunde war ein anschließendes, auf Größenausschluss-Chromatographie basierendes Screening ausgewählter, komplementierender Klone erforderlich. Nur einer der 26 auf diese Weise analysierter Klone wies die für ein Monomer erwartete Retentionszeit auf (Abbildung 8). In dieses Protein waren die sechs Reste Ala-Arg-Trp-Pro-Trp-Ala eingefügt. Seine monomere Form wurde mittels analytischer Ultrazentrifugation nachgewiesen. Es weist außerdem einen hohen Helixanteil auf, denaturiert chemisch und thermisch kooperativ ($\Delta G_U(H_2O) = 2.7 \text{ kcal mol}^{-1}$) und zeigt eine beträchtliche katalytische Aktivität. Sein k_{cat} -Wert für die Umlagerung von Chorismat ist identisch mit dem von MjCM, während der K_M -Wert lediglich dreimal höher ist.^[97]

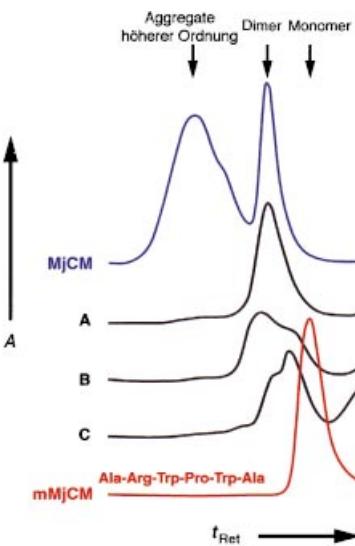


Abbildung 8. Unterscheidung der dimeren und monomeren Chorismat-Mutasen durch analytische Größenausschluss-Chromatographie.^[97] Wildtyp-MjCM (oberstes Chromatogramm) wurde als eine Mischung aus Dimeren und Aggregaten höherer Ordnung erhalten. Die meisten selektionsisierten Varianten (z.B. Chromatogramme A–C) waren Dimere oder Mischungen aus Dimeren und Monomeren. Nur eine von 26 getesteten Varianten (mMjCM, unterstes Chromatogramm) lag ausschließlich monomer vor. Es hatte die angegebene Insertionssequenz aus sechs Aminosäuren. Der monomere Charakter des Proteins in Lösung wurde durch analytische Ultrazentrifugation bestätigt.

Die Kombination aus Selektion (>100fache Anreicherung) und Screening (1 aus 26) zeigt, dass weniger als 0.05 % der möglichen Schleifensequenzen fähig sind, ein gut handhabbares, monomeres Protein zu bilden. Dieses Ergebnis widerspricht erneut der einfachen Annahme, dass die meisten interhelicalen Schleifensequenzen funktionell gleichwertig seien.^[76] Es unterstreicht außerdem eindrucksvoll die Vorteile

der Selektionsmethode. Die individuelle Musterung von 1000–10000 Proteinen, um eine Variante mit der gewünschten Aktivität und Topologie zu finden, würde einen immensen experimentellen Aufwand erfordern.

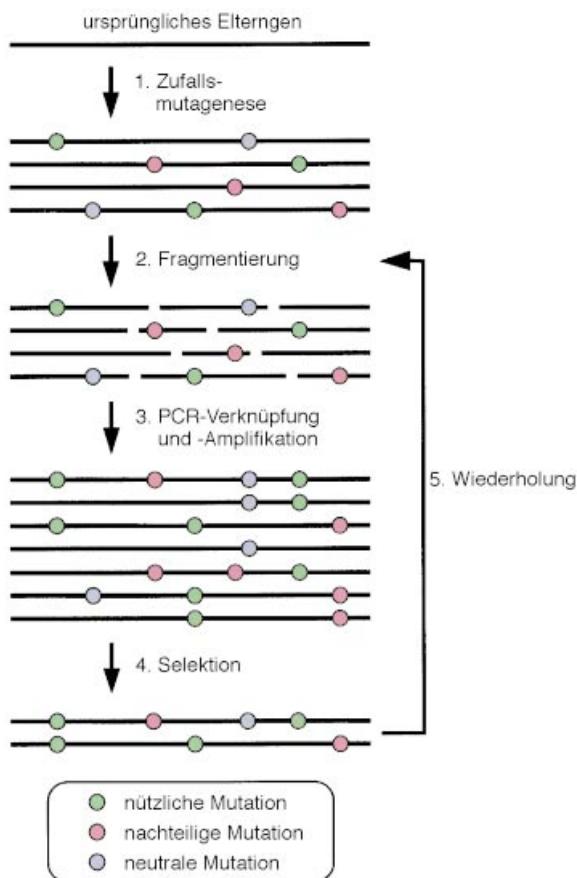
Ein weiterer entscheidender Vorteil der Selektion lässt sich z.B. daraus ableiten, dass Zellen, die das Monomer in sich trugen, wesentlich besser in Abwesenheit von Tyrosin und Phenylalanin wuchsen als Zellen mit dem MjCM-Wildtypdimer.^[97] Obwohl die genauen Ursachen für diesen Unterschied nicht bekannt sind, könnte dieser darauf zurückgeführt werden, dass das MjCM-Dimer bei der Überproduktion in *E. coli* katalytisch inaktive Aggregate bildet, während das Monomer dazu nicht neigt (Abbildung 8). Der Prozess der genetischen Selektion resultiert nicht nur in einer Maximierung der katalytischen Aktivität, sondern minimiert gleichzeitig die mit Fehlfaltung, Aggregation oder Toxizität von Proteinen einhergehenden Probleme. Diese Eigenschaft könnte für die Produktion von Biokatalysatoren zur kommerziellen Nutzung, bei der es auf stabile und einfach herzustellende Enzyme ankommt, wertvoll sein.

3.5. Verbesserung schwacher Enzymaktivität

Die eigentliche darwinistische Evolution besteht aus vielen Cycles von Mutation und Selektion. Dieser Prozess kann im Labor zur Optimierung der Eigenschaften eines ineffizienten Enzyms nachgeahmt werden, was häufig mit dem Schlagwort „gesteuerte Evolution“ („directed evolution“) bezeichnet wird.^[10]

Die in Abschnitt 3.4.1 beschriebene hexamere, aber wenig aktive Chorismat-Mutase^[95] konnte auf diese Weise verbessert werden.^[99] Unter Anwendung von DNA-Shuffling (Schema 8),^[5, 100] das die sexuelle genetische Rekombination *in vitro* nachahmt, wurden Mutationen in die Untereinheiten des Hexamers kodierende Gen eingefügt. Verbesserte Varianten wurden wie zuvor aufgrund ihrer Fähigkeit, den Fehler im Chorismat-Mutase-Stoffwechsel der Bakterien zu beheben, selektiert. Aus den am schnellsten wachsenden Zellen wurde die Plasmid-DNA isoliert, und die gesamte Prozedur wurde wiederholt.

Nach zwei Runden der Mutation und Selektion wurden einige Klone gefunden, die ähnlich wuchsen wie der Wildtyp. Die entsprechenden Enzyme wurden isoliert und biochemisch charakterisiert.^[99] Eine Variante, die drei Mutationen enthielt (Ser 15 Asp, Leu 79 Phe und Thr 87 Ile), hatte einen neunmal höheren k_{kat} - und einen 35-mal höheren k_{kat}/K_M -Wert. Der Größenausschluss-Chromatographie zufolge ist sie immer noch wesentlich größer als der dimere Wildtyp. Genauere biophysikalische Untersuchungen sind noch voraussetzt, um abschließende Aussagen über den exakten Oligomerisierungszustand und die Organisation der Untereinheiten machen zu können. Dieser Katalysator wurde einer weiteren Runde von Mutagenese und Selektion unterzogen, wobei weitere Klone mit etwas verbesserten k_{kat}/K_M -Werten identifiziert wurden. Interessanterweise fehlt vielen dieser Klone der C-terminale Histidin-Schwanz, der angehängt worden war, um die Reinigung zu vereinfachen, sowie weitere fünf bis acht Aminosäuren. Es ist durchaus möglich, dass die Entfer-



Schema 8. DNA-Shuffling.^[5, 100] Techniken wie die fehlerbehaftete (error-prone) PCR können eingesetzt werden, um ein Wildtypen zu mutagenisieren. Die resultierende Ansammlung homologer Gene enthält nützliche, nachteilige oder neutrale Punktmutationen. Die Behandlung dieses Genpools mit DNaseI liefert einen Satz von Zufallsfragmenten. Das erneute Zusammenfügen dieser Zufallssequenzen zu Genen voller Länge liefert neue Rekombinationen. Die Selektion auf Varianten mit verbesselter Funktion bildet den Ausgangspunkt für eine weitere Runde des DNA-Shuffling. Punktmutationen ergeben sich ebenfalls während des Rekombinationsprozesses, wodurch jede Runde einem echten darwinistischen Evolutionszyklus entspricht.

nung dieser C-terminalen Erweiterung der Helix H3 die für die Packung der höheren Oligomere nötigen Wechselwirkungen zwischen den Untereinheiten verbessert.

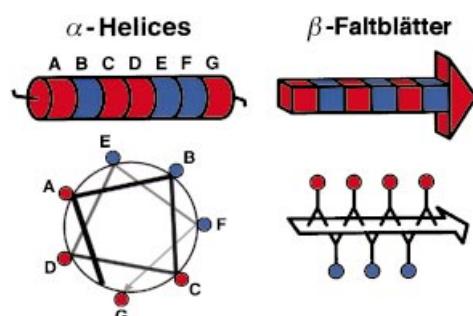
Prinzipiell könnte man die Methode der gesteuerten Evolution beliebig oft wiederholen, bis man die gewünschte Aktivität erhält. Erreicht man allerdings den Punkt, an dem der Katalysator so aktiv ist, dass die Wirtzelle so effizient wie der Wildtypstamm wächst, wird die Selektion von weiter verbesserten Varianten schwierig. Dies trifft auf das modifizierte Hexamer zu, auch wenn es immer noch um eine Größenordnung weniger aktiv ist als das homodimere MjCM.^[37] Weil die Gesamtaktivität sowohl von der spezifischen Aktivität als auch von der Konzentration des Katalysators abhängt, kann man durch Herabsetzung der Konzentration an verfügbarem Katalysator den Selektionsdruck erhöhen. In der Praxis kann die intrazelluläre Proteinkonzentration auf vielfältige Art und Weise, z.B. durch Verwendung von Plasmiden mit geringer Kopienzahl,^[101] schwachen Promotoren^[102] und ineffizienten Ribosomenbindungsstellen,^[103] verringert werden.

3.6. Proteindesign

Das De-novo-Design *funktioneller* Proteine ist ein weitaus ambitionierteres Ziel als die topologische Nachbildung eines vorhandenen Enzyms. Einige kleinere Proteingerüste wurden erfolgreich auf der Grundlage elementarer Prinzipien aufgebaut,^[104–110] oftmals mit Hilfe von Computer-Algorithmen.^[111, 112] Die Ausstattung dieser Moleküle mit Funktionen – der Fähigkeit, andere Moleküle mit hoher Selektivität zu erkennen oder chemische Reaktionen zu katalysieren – bleibt jedoch eine enorme Herausforderung.

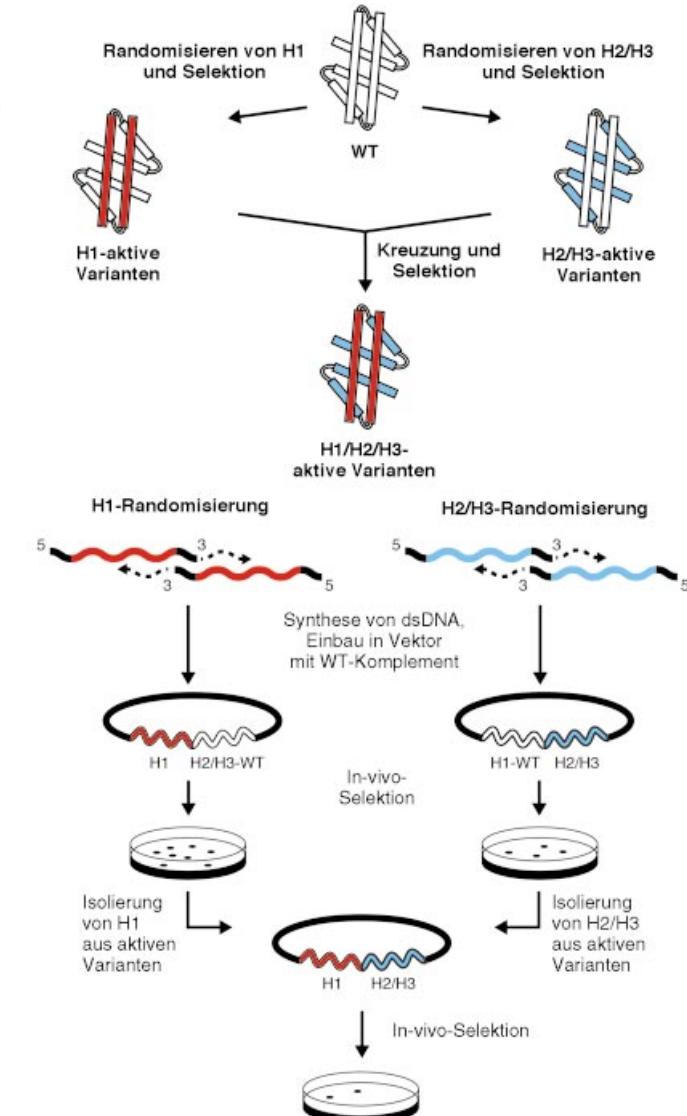
Im Prinzip sollten die gleichen Selektionsmethoden, die zur Charakterisierung der entscheidenden mechanistischen und strukturellen Faktoren von Enzymen sowie bei der Veränderung ihrer Topologie erfolgreich waren, hilfreich für das Design neuer Proteinkatalysatoren sein. Man könnte sich z.B. vorstellen, aktive Katalysatoren direkt aus großen, randomisierten Proteinbanken zu selektionieren.^[113] Allerdings ist der Sequenzraum fast unendlich groß. Eine Proteinbank mit der Masse der Erde (5.98×10^{27} g) würde bestenfalls 3.3×10^{47} unterschiedliche Sequenzen umfassen. Diese Zahl, die sogar für ein kleines Protein nur einen verschwindend kleinen Bruchteil aller Möglichkeiten repräsentiert, übersteigt die experimentell zugängliche Zahl bei weitem. Selbst die empfindlichste und effizienteste Selektionsmethode wäre zwecklos, es sei denn, die Katalysatoren wären in einem Ensemble randomisierter Varianten weitaus häufiger als angenommen vertreten.

Ein realistischer Kompromiss zwischen den beiden Ansätzen könnte darin bestehen, die grundlegenden Strukturinformationen, wie die Sequenzvorlieben von Helices und Faltblättern oder die Tendenz hydrophober Reste, sich im Proteininnern abzuschirmen, für das Design zielgerichteter randomisierter Genbanken zu nutzen, aus denen funktionelle Katalysatoren selektiert werden können. Eine mögliche allgemeine Strategie für das Proteindesign, die auf der nicht zufälligen Verteilung von Aminosäuren in gefalteten Strukturen beruht, ist z.B. die „binäre Musterung“ („binary patterning“) von polaren und unpolaren Aminosäuren (Schema 9).^[75, 114, 115] Diese Strategie wurde bereits erfolgreich eingesetzt, um kompakte Helixbündel, die viele Eigenschaften nativer Proteine aufwiesen, wie Proteaseresistenz und kooperative Faltung, zu erhalten.^[116–118]



Schema 9. Amphiphile Sekundärstrukturen: Durch die Vorgabe der Positionen von polaren (rot) und unpolaren Resten (blau) mittels einfacher binärer Muster können amphiphile helicale oder β -Faltblatt-Strukturen konstruiert werden.^[227]

Wir haben die Möglichkeit eines solchen Ansatzes für das Design aktiver Katalysatoren untersucht, indem wir systematisch alle Sekundärstrukturelemente einer homodimeren Chorismat-Mutase des AroQ-Typs (Schema 10) durch binär gemusterte Module mit Zufallssequenzen ersetzen. Die katalytischen Eigenschaften der Proteine aus den resultierenden Genbanken wurden anschließend mit Hilfe der genetischen Selektion bewertet. Auf diese Weise ließen sich auch quantitative Informationen zur Robustheit des hier zu Grunde liegenden Proteingerüsts und Einblicke in die subti-



Schema 10. Aufbau von großen binär gemusterten AroQ-Chorismat-Mutase-Genbanken. A) In einer zweistufigen Strategie wurden zunächst die H1- und H2/H3-Helices separat randomisiert und funktionelle Varianten selektiert, anschließend wurden ausschließlich die funktionellen binär gemusterten Segmente der ursprünglichen Genbanken kombiniert, und die Selektion wurde wiederholt. In den resultierenden Konstrukten war das Protein zu ca. 80 % randomisiert. B) Die binär gemusterten Gene wurden mit Hilfe synthetischer randomisierter Oligonucleotide, die entweder dem die Helix H1 (rot) oder dem die Helices H2/H3 (blau) kodierenden Abschnitt entsprachen, konstruiert und wie gezeigt mit Vektorfragmenten (schwarz) verknüpft. Im ersten Schritt wurden die randomisierten Abschnitte mit ihren Wildtypgegenstücken H2/H3-WT und H1-WT (weiß) kombiniert. Die Selektion wurde nach dem in Schema 2 dargestellten Verfahren durchgeführt. WT = MjCM'-Wildtyp; ds = doppelsträngig.

len Wechselwirkungen, die für die Bildung einer aktiven Tasche notwendig sind, gewinnen.^[119]

In der ersten Phase des Experiments wurden die den Helices H1 und H2/H3 entsprechenden Module jeweils einzeln mit den passenden helicalen Gegenstücken des thermostabilen MjCM'-Enzyms, einer Protease-resistanten Version der dimeren *M.-jannaschii*-Chorismat-Mutase,^[37] kombiniert. Dies entspricht einer Randomisierung von 37 bzw. 42 % des gesamten Proteins im Falle der H1- bzw. der H2/H3-Genbank. An jeder randomisierten Stelle wurden nur vier polare (Asn, Asp, Gln, Glu) oder vier unpolare Reste (Ile, Leu, Met und Phe) zugelassen. Die Schleifen und die katalytischen Reste wurden dabei zunächst beibehalten.

Der Chorismat-Mutase-defiziente Bakterienstamm wurde mit den binär gemusterten Genbanken transformiert, und die Fähigkeit der teilrandomisierten Proteine zur Faltung in geordnete, katalytisch aktive Strukturen wurde anhand der Komplementation der Chorismat-Mutase-Defizienz beurteilt. Ungefähr 0.02 % der Klone der H1-Genbank waren aktiv. Mit anderen Worten, ca. eines von 4500 binär gemusterten Enzymen ist ein funktionierender Katalysator. Ein etwas kleinerer Teil der H2/H3-Genbank, etwa 0.006 % (eine von 17500 Varianten), konnte den genetischen Defekt beheben. Angesichts der Häufigkeit von Katalysatoren wäre es sogar möglich gewesen, aktive Enzyme zu finden, wenn man in jeder der beiden Genbanken die drei unveränderten Reste der aktiven Tasche (Arg11, Arg28 und Lys39 in H1 sowie Arg51, Glu52 und Gln88 in H2/H3) ebenfalls mutagenisiert hätte. Selbst wenn man annimmt, dass lediglich eine einzige Aminosäure an jeder Position der aktiven Tasche toleriert wird, hätte sich die Wahrscheinlichkeit, einen Katalysator aufzuspüren, nur um den Faktor 20³ verringert. Damit hätten aktive Chorismat-Mutasen auch noch aus experimentell zugänglichen Genbanken mit mindestens 10⁸ Mitgliedern (20³ × 10⁴) isoliert werden können. Das Prinzip der binären Musterung eignet sich also bestens zum Design von Katalysatoren.

Katalysatoren aus beiden Genbanken (H1- oder H2/H3-Randomisierung) wurden isoliert und charakterisiert.^[119] Trotz des limitierten Satzes an Bausteinen, die in den randomisierten Abschnitten zum Einsatz kamen, sind die Enzyme erstaunlich aktiv. Mit mäßig erhöhten K_m -Werten und k_{kat} -Werten im Bereich von 0.2 bis 2.3 s⁻¹ ähneln sie dem MjCM-Wildtyp ($k_{\text{kat}} = 5.7 \text{ s}^{-1}$) recht gut.^[37] Während die meisten Positionen des Proteins relativ tolerant gegenüber Substitutionen sind, offenbarte die Sequenzanalyse einige Stellen, welche nur wenige Aminosäuren zulassen. So weisen alle aktiven Klone eine absolut konservierte Asn-Lys-Abfolge an den Positionen 84 und 85 auf, eine Bevorzugung, die bei unselektierten Klonen nicht ersichtlich ist. Im Nachhinein lässt sich dieser Befund anhand der EcCM-Struktur erklären.^[33] Der zu Asn84 homologe Rest bildet eine Wasserstoffbrücke zu einer der beiden Carboxylatgruppen des gebundenen Liganden, und eine Methylengruppe des benachbarten Lysinrestes geht vermutlich van-der-Waals-Wechselwirkungen mit dem Liganden ein. Offenbar wären alle anderen an diesen Positionen zugelassenen Aminosäuren nicht in der Lage gewesen, solche Wechselwirkungen beizusteuern.

Die vorselektionierten H1- und H2/H3-Module wurden auf ähnliche Weise miteinander kombiniert, wie schwere und

leichte Ketten von Immunoglobulinen im Immunsystem zu diversen Antikörperstrukturen zusammengefügt werden.^[120] In den daraus resultierenden binär gemusterten H1/H2/H3-Bibliotheken waren ca. 80 % der MjCM'-Reste randomisiert, und >90 % des Proteins bestand aus nur acht verschiedenen Aminosäuren (Schema 10A). Erstaunlicherweise war nur ungefähr eine von 10 000 möglichen Kombinationen ein aktiver Katalysator, obwohl jede einzelne der vorselektierten H1- und H2/H3-Komponenten zusammen mit einer Wildtyphälfte funktionell ist. Offenbar sind die natürlichen H1- und H2/H3-Helices von MjCM' wesentlich besser als die selektierten Abschnitte in der Lage, als Templat für korrekt gefaltete, aktive Enzyme zu fungieren. Möglicherweise tolerieren sie destabilisierende Veränderungen an den randomisierten Positionen des Proteins besser.

Wie bereits bei den individuellen H1- und H2/H3-Genbanken aufgefallen war, sind bestimmte Positionen in aktiven H1/H2/H3-Varianten hoch konserviert. Zusätzlich zu der bereits festgestellten Asn84-Lys85-Abfolge enthalten die meisten der komplementierenden Klone die Reste Ile14, Asp15 und Asp18. Diese in der H1-Bibliothek nicht besonders konservierten Aminosäuren tragen wahrscheinlich zur Stabilisierung der aktiven Tasche bei. In EcCM wechselwirkt der Rest in Position 14 direkt mit dem gebundenen Liganden über van-der-Waals-Kontakte, während die Reste 15 und 18 aus der zweiten Sphäre um die aktive Stelle mit den katalytisch wichtigen Resten Arg51 und Arg28 wechselwirken.

Die biochemische Charakterisierung einer der selektierten H1/H2/H3-Varianten bestätigt, dass es sich um ein helicale Homodimer handelt, das thermisch kooperativ denaturiert ($T_m \approx 44^\circ\text{C}$). Die Umlagerung von Chorismat zu Prephenat katalysiert sie im Vergleich zum *M.-jannaschii*-Enzym mit einem nur 15fach niedrigeren k_{kat} -Wert und einem 40fach höheren K_m -Wert.^[119] Durch die Kombination von binärer Musterung und zwei Evolutionscyclen haben wir aus einem vereinfachten Satz von Bausteinen ein neues Helixbündelprotein erhalten, dessen biophysikalischen und funktionellen Eigenschaften denen der natürlichen Chorismat-Mutasen ähneln. Die Optimierung dieses und verwandter Katalysatoren der Bibliothek durch zusätzliche Runden von Mutagenese und Selektion wird für die Identifizierung kritischer Wechselwirkungen, die die Faltung, Funktion und Stabilität beeinflussen, nützlich sein.

Experimente dieser Art können wertvolle Einblicke in die chemischen Grundlagen von Proteinstrukturen liefern. Die Ergebnisse stützen Spekulationen, denen zufolge die ersten Proteine aus einer begrenzten Zahl von Aminosäuren aufgebaut gewesen sein könnten.^[121–123] Nichtsdestotrotz unterstreicht das relativ geringe Vorkommen aktiver Enzyme in den kombinatorischen Genbanken die Schwierigkeit, helicale Gerüste mit katalytischen Eigenschaften aufzubauen – sogar wenn sowohl die Position als auch die Identität aller kritischen Reste der aktiven Tasche im Vorhinein festgelegt wurden.

Wir schätzen,^[119] dass eine Genbank mit mindestens 10²⁴ Mitgliedern nötig wäre, um einen aktiven Katalysator unter vollständig randomisierten, binär gemusterten AroQ-Templaten zu finden. Aufgrund bisheriger Erfahrungen nehmen wir an, dass andere Faltungen noch bedeutend schwerer zu konstruieren sein könnten.^[124] Weil Genbanken dieser Grö-

ßenordnung experimentell nicht zugänglich sind, werden für das De-novo-Design maßgeschneiderter Enzyme zweifellos schrittweise Ansätze vonnöten sein. Am Anfang könnte die Entwicklung eines stabilen Gerüsts stehen,^[228] das dann mit funktionellen Gruppen für Bindung und Katalyse versehen wird, und schließlich optimiert man das gesamte Ensemble hinsichtlich der gewünschten Aktivität. Es ist offensichtlich, dass ein effizientes Selektionssystem von unschätzbarem Wert, wenn nicht gar absolut notwendig für das Gelingen eines solchen Unterfangens sein wird.

4. Andere Anwendungen der In-vivo-Selektion

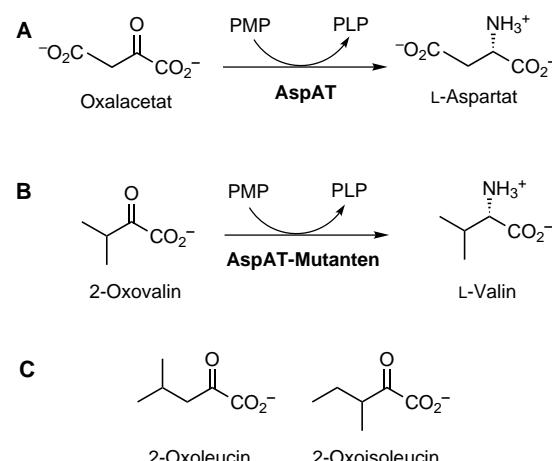
Die Prinzipien der genetischen Selektion, die in Abschnitt 3 anhand des Chorismat-Mutase-Systems illustriert wurden, sind generell auf das Studium von Proteinstrukturen und -funktionen anwendbar. Einige in den vorhergehenden Abschnitten nicht berücksichtigte, zusätzliche Beispiele sollen im Folgenden hervorgehoben werden.

4.1. Veränderung der Selektivität

Wegen ihrer außergewöhnlichen Effizienz, ihrer hohen Selektivität und der milden Reaktionsbedingungen sind Enzyme wertvolle Werkzeuge in der organischen Synthese, insbesondere für Umwandlungen, die mit konventionelleren Methoden nicht durchführbar sind oder die strukturell komplexe und biologisch wichtige Verbindungen wie Kohlenhydrate, Nucleinsäuren und Proteine betreffen.^[125] Für Chemiker, die nach allgemein anwendbaren Synthesemethoden suchen, kann die eingeschränkte Substratspezifität vieler natürlicher Enzyme jedoch eine erhebliche Limitierung darstellen. Deshalb wurden große Anstrengungen zur rationalen Umgestaltung aktiver Taschen von Enzymen unternommen. Allerdings war der Erfolg, abgesehen von wenigen bemerkenswerten Ausnahmen,^[126–128] dürfzig.

Hier bietet die Randomisierung in Verbindung mit Screening- oder Selektionsverfahren eine attraktive Alternative zur rationalen Umgestaltung einer aktiven Stelle, um deren Substratspezifität zu verändern. Die Umwandlung einer homodimeren Aspartat-Aminotransferase (AspAT) in eine effiziente Valin-Aminotransferase ist ein bemerkenswertes Beispiel dafür, was mit In-vivo-Selektion erreicht werden kann.^[129]

Die Wildtyp-AspAT katalysiert die Umwandlung von Oxalacetat und Aspartat ineinander (Schema 11), ist aber gegenüber β -verzweigten hydrophoben Aminosäuren praktisch nicht aktiv. Um die Substratspezifität zu ändern, wurde das ursprüngliche Gen durch DNA-Shuffling mutagenisiert (siehe Schema 8). Anschließend wurden Varianten selektiert, die fähig waren, einen auxotrophen *E. coli*-Stamm zu komplementieren, dem das Gen für die für Substrate mit β -verzweigter Kette spezifische Aminotransferase fehlte. Dieses Enzym ist zur Synthese von Valin, Leucin und Isoleucin notwendig. Deshalb kann der Mutantenstamm in Abwesenheit von Valin nur wachsen, wenn die Aktivität der AspAT für die Umwandlung von 2-Oxovalin zu Valin ansteigt.



Schema 11. A) Aspartat-Aminotransferase (AspAT) katalysiert die Transaminierung von Oxalacetat zu L-Aspartat unter Umwandlung des Cofaktors Pyridoxaminphosphat (PMP) zu Pyridoxalphosphat (PLP). Die Reaktion ist vollkommen reversibel. B) Mutagenese und Selektion in einem auxotrophen *E. coli*-Stamm lieferten AspAT-Varianten, die außerdem die reversible Umwandlung des β -verzweigten 2-Oxovalin zu L-Valin katalysieren.^[129] C) 2-Oxoleucin und 2-Oxoisoleucin sind ebenfalls Substrate für die mutierten AspATs.

Insgesamt wurden fünf Runden von Mutagenese und Selektion durchgeführt. Jede Runde lieferte zwischen 10^6 und 10^7 Transformanten, wobei jeweils die Plasmid-DNA der 90–100 am schnellsten wachsenden Klone für die folgende Shuffling-Runde verwendet wurde. Die Selektionskriterien für AspAT-Mutanten, die 2-Oxovalin als Substrat nutzen konnten, wurden in jeder Runde verschärft. Anfangs wurde die Inkubationszeit im Selektionsmedium verkürzt, dann wurde auf 2-Oxovalin im Wachstumsmedium verzichtet (wodurch dessen intrazelluläre Konzentration herabgesetzt wird), und schließlich wurde das Expressionsniveau der Gene der AspAT-Mutanten abgesenkt.

Eine der aus diesem Experiment hervorgegangenen Mutanten wies eine 10^5 fache Zunahme des k_{kat}/K_M -Wertes für 2-Oxovalin und eine 30fache Abnahme des k_{kat}/K_M -Wertes für saure 2-Oxosäuren im Vergleich mit dem Wildtypenzym auf (Schema 11B). Sie enthielt 13 Aminosäure-Substitutionen.^[129] Ausgehend von einer AspAT-Variante mit drei Mutationen wurde in einem weiteren Evolutionsexperiment eine neue, nahe verwandte Variante mit 17 Aminosäure-Substitutionen gefunden,^[130] deren Effizienz bezüglich β -verzweigter 2-Oxosäuren 2×10^6 fach erhöht war. Die Kristallstrukturuntersuchung dieser Variante mit gebundener Isovaleriansäure ergab, dass lediglich einer der 17 mutierten Reste mit dem Substrat wechselwirkt und einige dieser Reste sogar mehr als 10 Å von der aktiven Tasche entfernt sind. Die resultierende Umgestaltung der aktiven Tasche sowie der Grenzfläche zwischen den Untereinheiten, die letztlich für die veränderte Spezifität verantwortlich ist, ist auf die subtiles und kumulativen Einflüsse der zahlreichen Substitutionen in einiger Entfernung von der Bindungstasche zurückzuführen. Ähnliche Beobachtungen wurden in natürlichen Systemen gemacht, beispielsweise bei der Reifung der Immunantwort auf kleine Haptene.^[131–133] Die Schwierigkeit, solche Modifikationen in der zweiten Sphäre a priori vor-

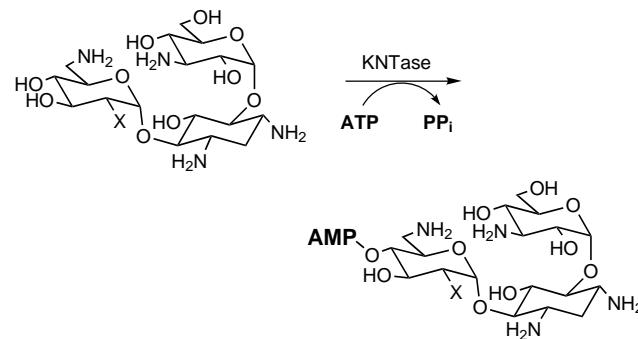
herzusagen, verdeutlicht den Nutzen der gesteuerten Evolution.

Die erfolgreiche Veränderung der Substratspezifität von β -Lactamase,^[100, 134–139] Thymidin-Kinase,^[140] Prenyl-Transferasen,^[141] OXidosqualen cyclisierenden Enzymen,^[142] N' -[(5'-Phosphoribosyl)formimino]-5-aminoimidazol-4-carboxamid-Ribonucleotid-Isomerase (ProFARI)^[143] und Aminoacyl-tRNA-Synthetasen,^[144, 145] um nur einige wenige aktuelle Beispiele zu nennen, spricht Bände über die Vielseitigkeit des Selektionsansatzes.

4.2. Erhöhung der Thermostabilität

Die mäßige Stabilität von Enzymen, die aus mesophilen Organismen isoliert werden, kann eine weitere Beschränkung für viele praktische Anwendungen sein. Aus diesem Grunde besteht ein beträchtliches Interesse an der Isolierung von Katalysatoren aus Organismen, die unter extremen Bedingungen leben.^[146] Auch Evolutionsexperimente sind geeignet, Proteine hervorzubringen, die in einem größeren Bereich unterschiedlicher experimenteller Bedingungen aktiv sind.

Eine Strategie, thermostabile Enzyme hervorzubringen, erfordert die Einführung des kodierenden Gens eines interessierenden Enzyms aus einem mesophilen Organismus in einen thermophilen. Anschließend wird auf Varianten selektiert, die auch bei erhöhten Wachstumstemperaturen aktiv bleiben.^[147] Diese Methode wurde zur Identifizierung von Kanamycin-Nucleotidyl-Transferase-Mutanten eingesetzt (Schema 12).^[148, 149]



Schema 12. Die Kanamycin-Nucleotidyltransferase (KNTase) überträgt eine Nucleosidmonophosphat-Gruppe von ATP auf die 4'-Hydroxygruppe der Antibiotika Kanamycin A ($X = \text{OH}$) und B ($X = \text{NH}_2$). Diese Reaktion macht Bakterien resistent gegen Aminoglycosid-Medikamente.

Zellen des mäßig thermophilen *Bacillus stearothermophilus*, die das mesophile Wildtypenzym enthalten, sind bei 47 °C resistent gegenüber dem Antibiotikum Kanamycin, jedoch nicht bei Temperaturen, die 55 °C übersteigen. Varianten, die auch bei Temperaturen oberhalb von 60 °C Kanamycin-resistent waren, erhielt man durch Mutagenese und nachfolgende Selektion im thermophilen Bakterium. Eine einzige Mutation, Asp80Tyr, ist für die erhöhte Resistenz des Enzyms gegenüber irreversibler thermischer Inaktivierung verantwortlich.^[148, 149] Weitere Selektionsexperimente bei 70 °C ergaben, dass Thr130Lys eine zweite stabilisierende

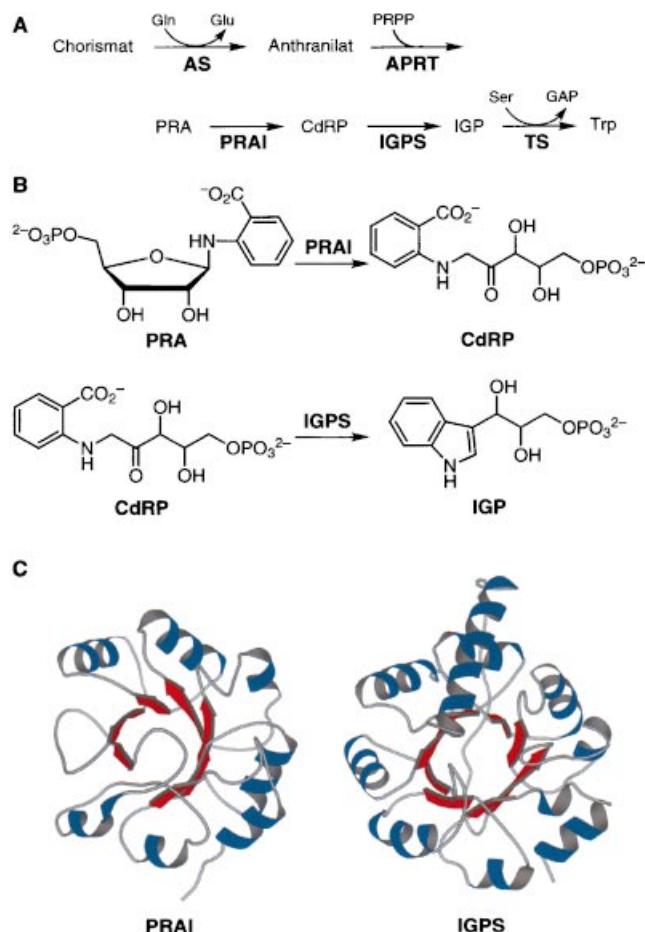
Mutation ist.^[148] Die Asp80Tyr/Thr130Lys-Doppelmutante hat eine Halbwertszeit von mehr als einer Stunde bei 60 °C, während das Wildtypenzym unter diesen Bedingungen beinahe augenblicklich denaturiert. Außerdem ist die Doppelmutante weitaus resistenter gegen proteolytischen Abbau. Ein sogar noch stabileres Enzym ($T_m = 84^\circ\text{C}$), das 19 Aminosäure-Substitutionen enthält, wurde durch wiederholte Runden von DNA-Shuffling und Selektion im extrem thermophilen Bakterium *Thermus thermophilus* erhalten.^[150] Erstaunlicherweise beinhalteten die stabilisierenden Mutationen nicht die in den vorhergehenden Studien identifizierten Aminosäuren. Darüber hinaus sind die meisten der betroffenen Reste an der Oberfläche des Proteins lokalisiert, und 5 der 19 Reste wurden gegen Prolin ausgetauscht.

Jede Enzymaktivität, auf die sich in thermophilen Bakterien selektionieren lässt, sollte diesem Ansatz zugänglich sein. Allerdings könnten sich In-vitro-Strategien, die in verschiedenen Display-Formaten eingesetzt werden können (siehe Abschnitt 5), letztlich als flexibler in Bezug auf die Stabilisierung gegenüber einer breiten Palette extremer Bedingungen erweisen. Zusätzlich zum praktischen Nutzen, der aus diesen Experimenten resultiert, können die erzielten Resultate zu einem tieferen Verständnis für die Wechselwirkungen, welche zur Proteinstabilität beitragen, führen.^[151]

4.3. Veränderung der Funktion

Eine aussichtsreiche Strategie, neue Funktionalität in bereits existierende Proteine einzuführen, besteht darin, den Prozess der divergenten Evolution zu imitieren. Die Natur beschränkt sich auf eine begrenzte Zahl von Proteinarchitekturen,^[152] die sie für ihre vielen unterschiedlichen Bedürfnisse nutzt. Die Immunoglobulin-Grundstruktur^[153] ist ein gutes Beispiel dafür, wie mit geringen Abwandlungen die Erkennung unendlich vieler fremder Antigene möglich wird. Auf die gleiche Weise, mit der Chemiker Antikörper (und damit die mikroevolutionären Prozesse des Säugetier-Immunsystems) zur Herstellung einer breiten Palette von maßgeschneiderten Katalysatoren nutzen,^[154–156] könnten vorhandene Proteingerüste als Ausgangspunkte zur Entwicklung neuer aktiver Taschen mit unterschiedlichen Funktionen dienen. Die α/β -Fassstruktur,^[157, 158] welche ungefähr 10 % aller strukturell charakterisierten Proteine zu Grunde liegt, könnte unter diesem Gesichtspunkt besonders vielseitig einsetzbar sein.^[161]

Die Umwandlung einer Indol-3-glycerinphosphat-Synthase (IGPS) in eine sehr aktive Phosphoribosylanthranilat-Isomerase (PRAI) veranschaulicht die Anwendbarkeit dieser Strategie.^[159] PRAI und IGPS katalysieren aufeinander folgende Schritte in der Biosynthese von Tryptophan (Schema 13 A): das erste die Umwandlung von *N*-5'-Phosphoribosylanthranilat (PRA) zu 1'-(2'-Carboxyphenylamino)-1'-desoxyribulose-5'-phosphat (CdRP), das zweite die Umlagerung von CdRP zu Indol-3-glycerinphosphat (IGP) unter Freisetzung von CO_2 und H_2O (Schema 13 B). Beide Enzyme sind α/β -Fässer (Schema 13 C), die in *E. coli* kovalent miteinander verbunden sind. Außerdem binden beide Enzyme CdRP.



Schema 13. A) Biosynthese von Tryptophan ausgehend von Chorismat. AS = Anthranilatsynthase; PRPP = Phosphoribosylpyrophosphat; APRT = Anthranilat-Phosphoribosyltransferase; PRA = N-5'-Phosphoribosylanthranilat; PRAI = Phosphoribosylanthranilat-Isomerase; CdRP = 1'-(2'-Carboxyphenylamino)-1'-desoxyribulose-5'-phosphat; IGPS = Indol-3-glycerinphosphat-Synthase; IGP = Indol-3-glycerinphosphat; TS = Tryptophan-Synthase; GAP = Glyceraldehyd-3-phosphat. B) Durch gesteuerte Evolution wurde eine neue PRAI erhalten, die die Umwandlung von PRA zu CdRP katalysiert.^[159] Die Ausgangsstruktur für dieses Experiment war die der IGPS, die die Cyclisierung von CdRP zu IGP katalysiert. C) Sowohl die PRAI als auch die IGPS sind Proteine mit α/β -Fassstruktur, unterscheiden sich aber leicht in ihren Schleifenstrukturen.

Die Strategie, eine IGPS in eine PRAI umzuwandeln, umfasste rationales Design, Zufallsmutationen und genetische Selektion. Ausgehend von den Kristallstrukturen der beiden Enzyme wurde zunächst das IGPS-Gerüst umgestaltet, sodass es demjenigen der PRAI stärker ähnelte. Zuerst wurde ein langer, in PRAI nicht vorhandener N-terminaler Abschnitt entfernt. Als Nächstes wurde eine 15 Reste umfassende Schleife durch Schleifen aus 4–7 Resten ersetzt (PRAI weist eine Schleife aus vier Resten auf). Die Variation der Schleifenlänge und -sequenzen ergab einen Satz unterschiedlicher Varianten, aus denen man in späteren Selektionsschritten aktive Katalysatoren auswählen konnte. Als Letztes wurde eine zweite Schleifenregion mit der PRAI-Konsensussequenz GXGGXGQ ersetzt und ein Asparaginsäurerest in Position 184 eingeführt, um als allgemeine Base in der umgebauten aktiven Tasche zu dienen.

Aktive Varianten wurden aus der resultierenden IGPS-Mutanten-Genbank durch genetische Selektion in einem PRAI-defizienten *E. coli*-Stamm, welcher Tryptophan nicht selbst herstellen kann, isoliert. Ungefähr 500 von 30 000 Mitgliedern der Genbank konnten bei geringen Tryptophan-Konzentrationen wachsen. Die Gene der selektierten Klone wurden anschließend zusätzlichen Runden von Mutationen und Selektion unterworfen, bis einige Varianten gefunden wurden, die in Abwesenheit von Tryptophan wachsen konnten. Eine von ihnen, ivePRAI, wurde detailliert charakterisiert. Das kodierte Protein war löslich und korrekt gefaltet. Darüber hinaus besaß es in vitro eine beachtliche PRAI-, aber keine IGPS-Aktivität, obwohl es nur 28% Sequenzidentität zur PRAI und 90% zur IGPS hatte. In der Tat war die Spezifitätskonstante k_{cat}/K_M von ivePRAI sechsmal höher als die der Wildtyp-PRAI, was sich auf eine erhöhte Affinität für das PRA-Substrat zurückführen ließ.^[159]

Die erfolgreiche Umwandlung einer IGPS in eine PRAI im Labor stützt die These, dass die natürlichen Enzyme des Tryptophan-Biosyntheseweges durch divergente Evolution entstanden sind.^[160] Außerdem zeigt dieses Experiment besonders deutlich, wie neue katalytische Aktivitäten relativ rasch aus bereits existierenden Bindungstaschen hervorgehen können. Auch wenn sich die Erweiterung dieser Strategie auf Reaktionen, deren Substrat keine Affinität zum Templatprotein aufweist, als schwieriger erweisen sollte, lassen die mit IGPS erzielten viel versprechenden Resultate vermuten, dass das häufig vorkommende α/β -Fassmotiv ein allgemein nützliches Gerüst zum Design neuer Enzyme sein könnte.^[161]

4.4. Optimierung mehrstufiger Prozesse

Analog der natürlichen darwinistischen Evolution können mehrere Komponenten eines Biosyntheseweges – sogar ganze Organismen – gleichzeitig in einem im Labor durchgeführten Evolutionsexperiment optimiert werden. Gewöhnlich ist die Voraussetzung für ein genetisches Selektionssystem die Konstruktion eines geeigneten Selektionsstamms durch Mutationen oder Eliminierung chromosomaler Gene. Eine andere Möglichkeit besteht darin, einen vorhandenen Organismus zu verwenden und diesen Selektionsbedingungen zu unterwerfen, die fremde Gene notwendig machen. Die Optimierung der Effizienz eines Arsenat-Entgiftungsweges,^[162] der möglicherweise in Hinblick auf biologische Altlastensanierung von Nutzen sein könnte, ist ein interessantes Beispiel für diese Strategie.

Das Arsenat-Resistenzoperon des Bakteriums *Staphylococcus aureus* umfasst die drei Gene *arsR*, *arsB* und *arsC*, die es gegenüber Arsenat- und Arsenitionen resistent machen. Das gesamte Operon wurde mittels DNA-Shuffling (siehe Schema 8) mutagenisiert und in einen gewöhnlichen *E. coli*-Stamm eingeführt. Die Selektion von Arsenat-resistenten Bakterien erfolgte durch Plattierung der Population auf Medium, das steigende Konzentrationen an Arsenat enthielt (16–400 mM). Nach drei Runden Mutagenese und Selektion wurden schnell wachsende Klone isoliert. Einer von ihnen konnte sogar in Flüssigmedium wachsen, das 500 mM Arsenat enthielt.

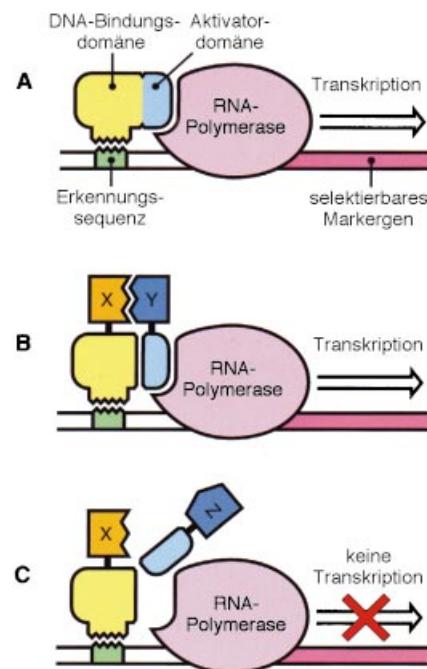
Bei der Analyse der beteiligten Operonregion wurden insgesamt 13 Mutationen entdeckt. Zehn (sieben davon waren „silent“, d.h. führten nicht zu Aminosäure-Substitutionen) traten in *arsB* auf, das eine membranassoziierte Pumpe kodiert, die Arsenit aktiv aus der Zelle transportiert; gemeinsam trugen diese Mutationen zu einer vier- bis sechsfach verbesserten Arsenitresistenz bei, vermutlich durch Verstärkung der Expression unter gleichzeitiger Verminderung der Toxizität des ArsB-Proteins. Zwei Silent-Mutationen fanden sich in *arsR*, das das Operonrepressor-Protein kodiert. Eine weitere Mutation befand sich stromabwärts, unmittelbar nach dem Operon. Interessanterweise enthielt das Arsenat-reduktase-Gen *arsC* überhaupt keine Mutationen. Dennoch war die Gesamtgeschwindigkeit der Arsenatreduktion 12fach erhöht, was mit einem erhöhten Genexpressionsniveau infolge der Mutationen im stromaufwärts liegenden *arsB*-Gen erklärt wurde.^[162]

Diese Experimente machen einmal mehr deutlich, dass auch viele Änderungen, die nicht zu einer höheren kinetischen Effizienz eines bestimmten Enzyms führen, dem Organismus unter Selektionsbedingungen zum Überleben verhelfen können. Ob die besseren Überlebenschancen auf die verbesserten katalytischen Fähigkeiten, die effizientere Genexpression oder die erhöhte Proteinstabilität zurückzuführen sind, ist für den Selektionsprozess irrelevant. In der Tat scheint die Feinabstimmung nichtkatalytischer Eigenschaften im Allgemeinen wesentlich einfacher als die Optimierung der katalytischen Aktivität zu sein. Die dafür erforderlichen Mutationen sind vermutlich in nahezu allen Systemen häufiger und werden besser toleriert.

4.5. Auffinden von neuen Liganden und Inhibitoren

Obwohl die Struktur und die Funktion von Enzymen im Mittelpunkt dieses Übersichtsartikels stehen, soll darauf hingewiesen werden, dass die genetische Selektion zunehmend auch als wertvolles Werkzeug zur Identifizierung neuer Peptidliganden für biologische Rezeptoren und zur In-vivo-Charakterisierung von Ligand-Rezeptor-Wechselwirkungen eingesetzt wird. Vor allem das Zwei-Hybrid-System in Hefe (yeast two-hybrid system)^[163–167] hat sich für diese Anwendungen als besonders nützlich erwiesen.

Das Zwei-Hybrid-System in Hefe ist in Schema 14 dargestellt. Die Transkription der Gene in Hefe basiert auf der Wechselwirkung zwischen der RNA-Polymerase und der Aktivatordomäne eines Transkriptionsfaktors. Die Aktivatordomäne wird über die DNA-Bindungsdomäne des Transkriptionsfaktors an eine DNA-Promotorregion gebunden. Im Zwei-Hybrid-System besteht die kovalente Verbindung der beiden Domänen des Transkriptionsaktivators nicht mehr. Individuelle Hybridbibliotheken lassen sich durch die Verknüpfung von einer oder beiden getrennten Domänen an ein Ensemble anderer Moleküle (DNA, RNA, Proteine oder niedermolekulare Verbindungen) erhalten. Sobald zwei der verbundenen Moleküle miteinander wechselwirken, kommen die Aktivator- und die DNA-Bindungsdomäne zusammen, sodass die Transkription vom Promoter aus stattfinden kann, welcher stromabwärts von der DNA-Bindungsstelle liegt.



Schema 14. Zwei-Hybrid-Systeme aus Hefe:^[163–167] A) Die Genexpression in Hefe benötigt die Wechselwirkung der RNA-Polymerase mit einem an die DNA gebundenen Aktivator für die Transkription. B) Zwei unterschiedliche Proteine oder Proteinbibliotheken werden mit der DNA-Bindungs- bzw. der Aktivatordomäne eines Hefe-Transkriptionsaktivators verknüpft. Eine Wechselwirkung zwischen zwei zusammenpassenden Proteinen (X und Y) rekonstituiert den Aktivator, und das Reportergen wird transkribiert. C) Bibliotheksmitglieder, die nicht miteinander wechselwirken können, können keinen funktionellen Aktivator bilden.

Die Transkription führt zur Expression eines Reportergens und ist dadurch mit einem selektionierbaren oder durchmusterbaren Phänotyp verbunden.

Am häufigsten wird das Zwei-Hybrid-System in Kombination mit einem Screening verwendet,^[164–166] wobei das Produkt der Expression des Reportergens der Zelle eine zusätzliche, leicht zu detektierende, aber für das Überleben unwesentliche Funktion verleiht. Es finden aber auch Zwei-Hybrid-Systeme Anwendung, die auf einer für die Zelle absolut notwendigen Funktion des Reportergens beruhen. Als Beispiel für eine solche, auf echter genetischer Selektion basierenden transkriptionellen Aktivierung lässt sich das Auffinden von Peptiden, die *in vivo* an die Cyclin-abhängige-Kinase 2 (CdK2) binden, anführen.^[168] Die Peptide wurden in Form einer 20 Reste umfassenden Insertionssequenz in der die aktive Tasche beinhaltenden Schleife des Thioredoxins von *E. coli* präsentiert; die besten Varianten wiesen eine Affinität zur Zielsequenz im nanomolaren Bereich auf. Ein Drei-Hybrid-System in Hefe zum Aufspüren von Ligand-Rezeptor-Wechselwirkungen ist ebenfalls beschrieben worden.^[169] Die verhältnismäßig geringe Transformationseffizienz von Hefe limitiert allerdings die Größe von Genbibliotheken, die erschöpfend untersucht werden können, aber die Entwicklung von auf *E. coli* beruhenden Zwei-Hybrid-Systemen^[170, 171] könnte die Möglichkeiten solcher Systeme deutlich erweitern.

Die genetische Selektion mit Hefe ließ sich auch ohne den Umweg über Zwei-Hybrid-Formate einsetzen, um Peptide aufzuspüren, die biologische Signalübertragungswege inhibi-

bieren.^[172] Die Selektion von Inhibitoren des Spindel-Kontrollpunktes (spindle checkpoint) und des Geschlechtspheromon-Signalübertragungsweges gelang z.B. mit Hilfe von rational entworfenen Hefestämmen. Die jeweiligen Peptidbibliotheken wurden dabei in einer an der Proteinoberfläche lokalisierten Schleife einer inaktiven Variante der *Staphylococcus*-Nuclease präsentiert. Die vermuteten intrazellulären Ziele wurden anschließend durch eine Kombination aus Zwei-Hybrid-Analyse und genetischer Aufteilung der betroffenen Biosynthesewege identifiziert.^[172] Eine ähnliche genetische Selektionsstrategie, allerdings in *E. coli*, führte kürzlich zur Identifizierung von Peptiden aus neun Resten, welche die intrazelluläre Dimerisierung der HIV-1-Protease inhibieren.^[173]

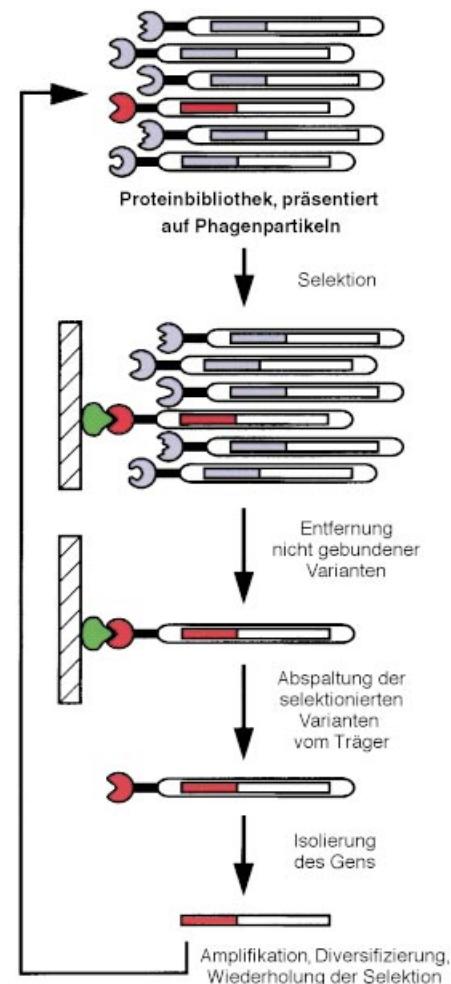
Solche Untersuchungen sind für die Entdeckung von Medikamenten von Vorteil, da sie neue Leitstrukturen und eine große Menge von Informationen über die Struktur-Funktions-Beziehungen in biologischen Rezeptoren liefern. Die erhaltenen Informationen sind für das Design von effizienten nichtpeptidischen Inhibitoren von großer Bedeutung. Die Charakterisierung der Angriffsstelle des selektierten Peptids kann außerdem neue Komponenten des untersuchten Signalübertragungsweges offenbaren, die ihrerseits hervorragende Ziele für die spezifische Inhibierung durch niedermolekulare organische Verbindungen sein könnten.

5. In-vitro-Selektionssysteme

Die in diesem Übersichtsartikel angeführten Beispiele zeigen, dass eine große Zahl von Proteinfunktionen mit einem genetisch selektionierbaren Phänotyp verknüpft werden können. Trotzdem ist es für viele interessante Anwendungen nicht trivial, eine solche Verknüpfung zu erreichen. Es kann sogar unmöglich sein (z.B. wenn nichtphysiologische Bedingungen zur Selektion benötigt würden). In diesen Fällen bietet die In-vitro-Selektion oft eine praktikable Alternative. Wie bei der In-vivo-Selektion muss auch bei der In-vitro-Selektion die genetische Information mit der Proteinfunktion in irgendeiner Art und Weise gekoppelt werden. Einige Möglichkeiten, wie dies realisiert werden kann, werden im Folgenden besprochen.

5.1. Phagen- und Zelloberflächen-Display

Die Präsentation von Proteinen (und Peptiden) an der Zelloberfläche^[174, 175] oder an Bakteriophagen^[176, 177] ist eine sehr effektive Methode, den Phänotyp mit dem Genotyp zu koppeln. Die zu untersuchenden Proteine aus einem Ensemble werden mit einem Membranprotein fusioniert oder mit einem der äußeren Capsidproteine eines filamentösen Phagen verbunden (Schema 15). Die daraus resultierende Phagenbibliothek wird einem immobilisierten (oder immobilisierbaren) Liganden angeboten, und spezifische Rezeptoren können in vitro durch Affinitätsselektion isoliert werden. Da die Zellen oder Bakteriophagen, die die selektierten Rezeptoren tragen, auch die entsprechenden kodierenden

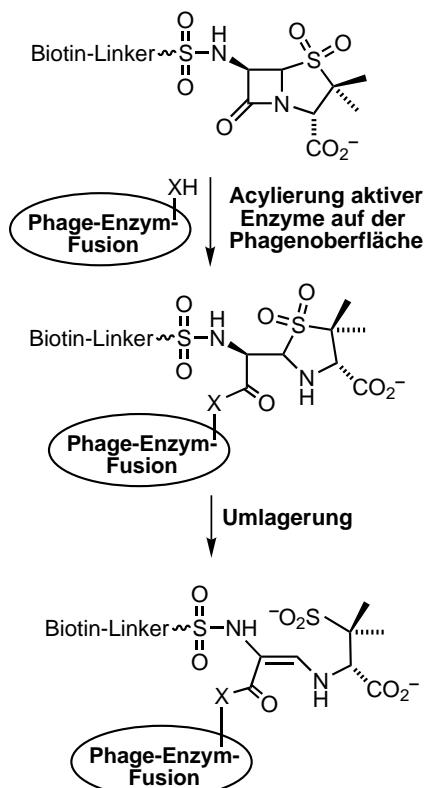


Schema 15. Selektion mit Phagendisplay:^[176, 177] Die Fusion des einen Antikörper oder ein anderes Protein kodierenden Gens mit dem Gen eines der Hüllproteine des Bakteriophagen macht es möglich, dass Proteine und ihre Varianten auf der Oberfläche des Viruspartikels präsentiert werden können. Die Selektion auf Bindung mit einem immobilisierten Liganden (grün) ermöglicht es, Proteine mit den gewünschten Bindungseigenschaften gleichzeitig mit dem zugehörigen kodierenden Gen (rot), das topologisch mit dem Protein verbunden ist, aus großen Populationen zu isolieren.

Gene beherbergen, wird im gleichen Schritt auch auf das gewünschte Gen selektiert. Es kann von der Festphase abgespalten und in vivo vervielfältigt werden. Außerdem lässt es sich weiter verändern und kann zusätzlichen Selektionsrunden unterworfen werden, um die Eigenschaften des kodierten Enzyms weiter den gewünschten Zielen anzupassen.

Display-Methoden finden breite Anwendung zur Identifizierung von Peptidliganden für diverse Rezeptoren und um auf Proteine mit neuen Bindungseigenschaften zu selektieren. Erwähnenswert in diesem Zusammenhang sind die Optimierung von DNA-bindenden Proteinen für die spezifische Erkennung jeder gewünschten DNA-Sequenz,^[178–185] die Minimierung der Immunoglobulin-Z-Domäne des Proteins A^[186] sowie die Evolution von Antikörperfragmenten mit femtomolaren Antigen-Bindungsaffinitäten.^[187] Die direkte Selektion von Proteinen aufgrund ihrer katalytischen Eigenschaften erwies sich als sehr viel schwieriger, aber einige viel versprechende Strategien sind entwickelt worden.

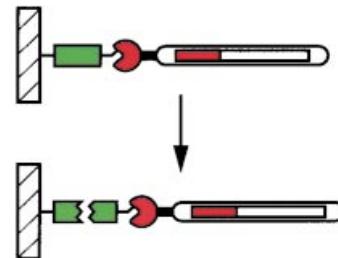
Ein Ansatz bedient sich reaktiver Liganden^[188] und auf mechanistischen Erkenntnissen beruhender Inhibitoren.^[189, 190] Damit werden Enzyme und katalytische Antikörper aus Phagenbanken angereichert. Phagenpartikel mit passender chemischer Funktionalität reagieren, entweder direkt oder nach chemischer Aktivierung, mit dem Inhibitor (bzw. dem Liganden) unter Bildung einer kovalenten Bindung. Dadurch werden aktive Phagen abgefangen und an einer Festphase immobilisiert (Schema 16). Die gleichen



Schema 16. Auffangen von Phagen-gebundenen Enzymen mit aus dem Enzymmechanismus abgeleiteten Inhibitoren: Penicillin-Sulfonderivate, die über einen Linker mit Biotin verbunden sind, eignen sich dazu, Enzyme, die auf filamentösen Phagen präsentiert werden, abzufangen, sofern die Enzyme das β -Lactam nucleophil hydrolysieren können.^[189] Enzyme, denen das Nucleophil (X) in ihrer aktiven Tasche fehlt, werden nicht kovalent modifiziert. Durch die Bindung von Biotin an immobilisiertes Streptavidin können die abgefangenen Phagen isoliert werden.

Proteingruppen, die für die Aktivierung des Inhibitors verantwortlich sind, können anschließend zur katalytischen Umwandlung eines Substrats benutzt werden. Mit dieser Methode können neue Katalysatoren entdeckt werden,^[189, 190] allerdings nur indirekt. Diese Selektionsstrategie erfordert lediglich die irreversible Modifizierung des Rezeptors – und zwar in einem einzigen Schritt – anstelle einer effizienten Umsetzung des Substrates, die es erforderlich machen könnte, mehrere Übergangszustände zu stabilisieren und eine Produktinhibierung zu vermeiden. Die Notwendigkeit, geeignete reaktive Derivate zu entwerfen und zu synthetisieren, sowie das ineffiziente Auffangen des durch den Katalysator gebildeten reaktiven Zwischenprodukts können zusätzliche praktische Limitierungen dieses Ansatzes sein.

Eine andere Strategie, um auf Katalysatoren zu selektieren, basiert auf der Colokalisation des Katalysators und des Substrates oder Produktes. So können immobilisierte Substrate mit Phagen verknüpft werden, die Enzyme präsentieren (Schema 17).^[191] Ist das Enzym aktiv, kann die intramolekulare Umwandlung eines Substrats in das entsprechende Produkt eine Möglichkeit zur Selektion von Phagen, die ein aktives Enzym tragen, sofern die Reaktion wie dargestellt zur Abspaltung des Phagen von der Festphase führt.^[191]



Schema 17. Die intramolekulare Umwandlung eines Substrats in das entsprechende Produkt ist eine Möglichkeit zur Selektion von Phagen, die ein aktives Enzym tragen, sofern die Reaktion wie dargestellt zur Abspaltung des Phagen von der Festphase führt.^[191]

molekulare Spaltung des Substrats das Phagenpartikel selektiv von der Festphase lösen, sodass das kodierende Gen isoliert und vermehrt werden kann. Analoge Systeme wurden für chemische Reaktionen entworfen, die eine Markierung am Phagen anbringen oder ein markiertes Produkt liefern, das anschließend mit dem Phagenpartikel reagiert, bevor es abdiffundiert.^[192] Die Fähigkeit einer modifizierten Protease, Peptide zu ligieren, wurde z.B. folgendermaßen optimiert: Enzymvarianten, die erfolgreich ein biotinyliertes Peptid an ihren eigenen N-Terminus anhefteten, wurden durch Streptavidin selektiv gebunden.^[193] In einer weiteren Variation dieser generellen Methode wurden gegen das Produkt gerichtete Antikörper eingesetzt, um Reaktionsprodukte zu binden, die an Phagenpartikeln verankert waren. Gleichzeitig wurden so auch die katalytisch aktiven Enzyme isoliert, die ebenfalls auf diesen Phagenpartikeln vorhanden waren.^[194]

Eine ganz andere Methode, um funktionelle Proteasen aus Proteinbibliotheken, die auf der Oberfläche bakterieller Zellen präsentiert werden, zu identifizieren, wurde kürzlich beschrieben.^[195] Sie basiert auf den elektrostatischen Wechselwirkungen, die positiv geladene Peptidsubstrate und -produkte auf der negativ geladenen Zelloberfläche festhalten. Die Spaltung des Substrats, das kovalent an fluoreszierende FRET-Gruppen (FRET = Fluoreszenz-Resonanzenergie-Transfer) gebunden ist, setzt den Fluoreszenzlöscher frei, wodurch ein Fluoreszenzsignal hervorgerufen wird. Aktive Katalysatorvarianten können dann durch Fluoreszenz-aktivierte Zellsortierung (FACS) isoliert werden. Auf diese Weise wurden Mutanten der Serin-Protease OmpT, die die Arg-Val-Bindung mit bis zu 60-mal höherer Effizienz spalten, aus Bibliotheken aus 6×10^5 randomisierten Varianten isoliert. Diese Methode, die eigentlich eher ein Hochdurchsatz-Screeningverfahren als eine echte darwinistische Selektion ist, sollte für eine Vielzahl von Proteinen, die auf der Oberfläche von Bakterien oder Hefen präsentiert werden (können), anwendbar sein.

5.2. Kompartimentierung

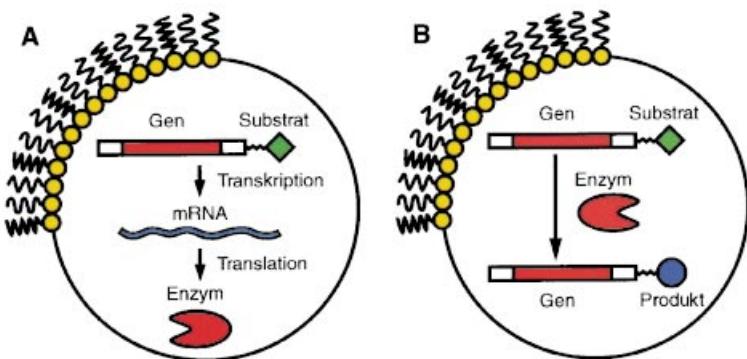
Zellfreie Transkriptions/Translations-Systeme können eine vollständige In-vitro-Alternative zum Phagendisplay darstellen. Solche Systeme enthalten alle RNAs, Proteine und niedermolekularen Verbindungen, die für die Transkription eines Gens und seine Übersetzung in ein Protein notwendig sind.^[196, 197] Durch Einkapselung dieser Komponenten in einem künstlichen, zellähnlichen Reaktionsraum kann der Genotyp topologisch mit dem Phänotyp gekoppelt werden. Beispielsweise könnte die Fähigkeit eines translatierten Enzyms, sein im gleichen Kompartiment lokalisiertes, kodierendes Gen zu verändern, als eine Selektionsmöglichkeit genutzt werden (Schema 18). In Kompartimenten, in denen ein inaktives Protein hergestellt wird, bleibt das entsprechende Gen unverändert und wird deshalb nicht vervielfältigt.

Die Anwendbarkeit dieser Methode wurde demonstriert, indem man Mischungen von Genen verwendete, die die *HaeIII*-Methyl-Transferase und ein anderes, damit nicht verwandtes Enzym kodieren. Die Methyl-Transferase methyliert die DNA und macht sie so resistent gegenüber der Spaltung durch die *HaeIII*-Endonuklease. Die Mischung wurde in eine Wasser/Öl-Emulsion eingebracht, die isolierte, wässrige Kompartimente (winzige, eingeschlossene Wassertropfen) mit einem Volumen ähnlich dem von Bakterien aufwies.^[198] Nach der Transkription/Translation in den isolierten Reaktionsräumen wurden Methyl-Transferase-kodierende Gene durch Methylierung geschützt und konnten per PCR vervielfältigt werden, wohingegen diejenigen, die das andere Enzym kodierten, durch die *HaeIII*-Endonuklease abgebaut wurden. So konnten Methyl-Transferase-Gene aus einer Mischung mit einem 10⁷fachen Überschuss der Gene, die das andere Enzym kodierten, isoliert werden.^[198]

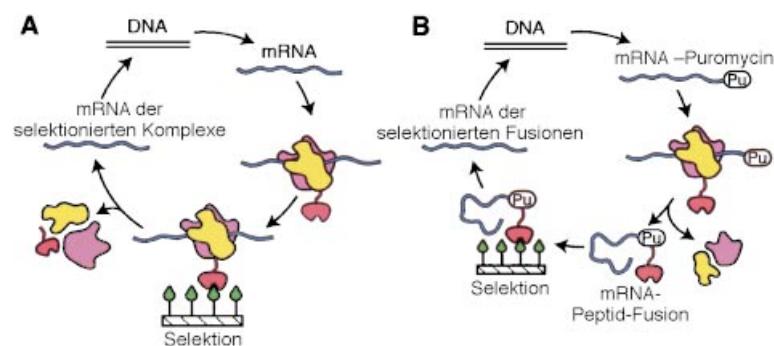
Wenn eine allgemein anwendbare Strategie für die Selektion entwickelt werden kann, dann könnten sich solche isolierten Reaktionsräume in molekularen Evolutionsexperimenten als breit anwendbar erweisen, weil damit das Einbringen in Zellen überflüssig wäre. Somit könnten weitaus größere Genbanken als in In-vivo-Experimenten zum Einsatz kommen und Schwierigkeiten, die mit der ineffizienten mikrobiologischen Herstellung einiger Proteine verbunden sind, umgangen werden.

5.3. Verknüpfung von mRNA und Protein während der Translation

Zwei weitere Methoden, das Ribosomen- (Polysomen-) Display^[199–203] und die Bildung von mRNA-Protein-Fusionen,^[204–206] können für die In-vitro-Evolution von Proteinen verwendet werden. Beide basieren auf der zellfreien Translation von mRNA-Templaten für die Herstellung von Proteinen an Ribosomen (Schema 19). Bei der natürlichen



Schema 18. Die genetische Selektion durch Kompartimentierung beruht auf der Transkription/Translation eines einzelnen Gens in einem sich in einer Wasser/Öl-Emulsion befindlichen Wassertropfen (A).^[198] Die anschließende Modifizierung des Gens durch das kodierte und colokalisierte Enzym (B) bildet die Grundlage für die Selektion. Gene, die ein inaktives Protein kodieren, werden nicht modifiziert und können im Selektionsprozess eliminiert werden.



Schema 19. Selektion durch Ribosomen-Display oder mRNA-Peptid/Protein-Fusion:^[199–203] A) Die In-vitro-Translation von mRNA an Ribosomen kann durch die Inkubation bei niedrigen Temperaturen und Zugabe von Mg²⁺ gestoppt werden. Komplexe aus Ribosomen, mRNA und frisch synthetisiertem Polypeptid, welches die gewünschte Rezeptorspezifität aufweist (rot), lassen sich durch Bindung an einen immobilisierten Liganden (grün) selektieren. Die Gewinnung und reverse Transkription der so zurückgehaltenen mRNA liefert die DNA, die nach PCR-Amplifikation charakterisiert oder für eine weitere Selektionsrunde verwendet werden kann. B) Eine mRNA-Protein-Fusion entsteht bei der Ribosomen-katalysierten Translation eines mRNA-Puromycin-Konjugats. Nach der Proteinsynthese wird das Protein (rot) im Peptidyl-Transferase-Zentrum des Ribosoms von der tRNA auf Puromycin (Pu) übertragen. Das gefaltete Protein kann danach zusammen mit seiner kodierenden mRNA In-vitro-Selektionsprozessen unterworfen werden.

Proteinsynthese werden Botschaft (Protein) und Bote (Messenger, mRNA) üblicherweise voneinander getrennt, sofern sie nicht in einem geschlossenen Reaktionsraum erfolgt. Diese Trennung lässt sich allerdings verhindern, indem das System so manipuliert wird, dass das translatierte Protein nicht vom Ribosom dissoziieren kann.^[199–203] Alternativ kann das Protein mit seiner eigenen mRNA über eine stabile, kovalente Bindung verknüpft werden.^[204–206]

Die Dissoziation der Komplexe aus Ribosomen, mRNA und translatiertem Protein lässt sich durch niedrige Temperaturen und Zugabe von Magnesiumacetat zur Pufferlösung unterdrücken.^[201] Noch bedeutender ist die Entfernung des Stopcodons, das normalerweise als Signal für Polypeptid-freisetzende Terminationsfaktoren dient und somit für die Beendigung der Proteinsynthese. Dadurch wird die Hydrolyse der an das Ribosom gebundenen Peptidyl-tRNA stark verlangsamt.^[201] Faltet sich das präsentierte Protein korrekt,

dann können auf Affinität basierende Selektionsmethoden eingesetzt werden, um Varianten zu isolieren, die niedermolekulare Verbindungen, Oligonucleotide oder andere Proteine erkennen (Schema 19A). Diese Technik wurde beispielsweise erfolgreich zur Erhöhung der Stabilität und Ligandenaffinität von Antikörpermolekülen eingesetzt.^[201, 207]

Die Zugabe von Ribonuclease-Inhibitoren zur Reaktionsmischung und der Einbau von „Stem-Loop“- (Haarnadel-) Strukturen in die mRNAs, um den Abbau der Ribonucleinsäuremoleküle zu verhindern, trugen dazu bei, die Effizienz von Experimenten dieser Art zu optimieren.^[201] Auch die Zugabe von molekularen Faltungshelfern (sog. Chaperonen), die *in vitro* die richtige Faltung der frisch synthetisierten Proteine unterstützen, kann dienlich sein.^[201]

Die Bildung kovalenter Fusionen zwischen einer mRNA und dem von ihr kodierten Protein könnte sich letztlich als eine robustere Methode zur Verknüpfung des Genotyps mit dem Phänotyp herausstellen.^[204–206] Solche Konstrukte werden z.B. durch In-vitro-Translation von mRNA-Templaten, die das Antibiotikum Puromycin an ihrem 3'-Ende tragen, erhalten (Schema 19B). Puromycin kann nach der Translation der RNA in die Bindungsstelle der Peptidyl-Transferase eindringen und sich am C-Terminus des entstehenden Proteins anheften, um ein kovalentes mRNA-Protein-Konjugat zu bilden. Anschließend kann auf das translatierte Protein, wie bereits für die anderen Formate beschrieben, selektiert werden. Da es kovalent mit seiner eigenen mRNA verknüpft ist, kann die Selektion – im Unterschied zu der in Schema 19A dargestellten Strategie – nach der Entfernung der Ribosomen und anderer Komponenten des In-vitro-Translationalssystems erfolgen.

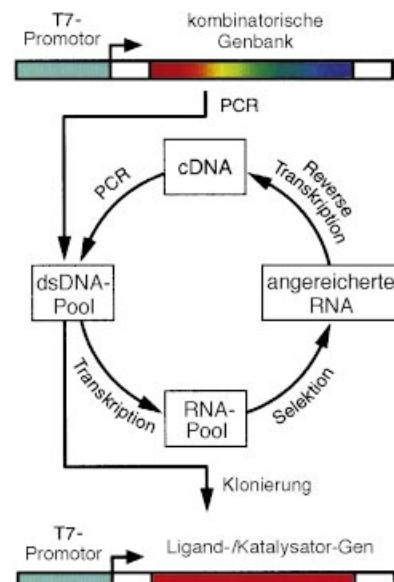
Aus einer Reihe von Gründen sind In-vitro-Transkriptions/Translations-Systeme attraktiv für die gesteuerte Evolution: Erstens kann eine breite Palette von Bedingungen für den eigentlichen Selektionsschritt verwendet werden. Zweitens können nichtnatürliche Aminosäuren über Suppressorcodons direkt in das Protein eingebaut werden.^[208] Und schließlich lassen sich *in vitro* wesentlich größere und komplexere Genbanken herstellen als dies *in vivo* möglich wäre, weil die limitierende Transformationseffizienz von *E. coli* in diesen Systemen keine Rolle spielt. Mit den heutigen Verfahren können Genbanken mit 10^{11} bis 10^{13} unterschiedlichen Mitgliedern aufgebaut werden. Eine bescheidene Maßstabsvergrößerung und eine verbesserte Genbanksynthese lassen Ensembles mit 10^{15} Mitgliedern in den Bereich des Möglichen rücken.^[205, 209]

Ähnlich den Phagen- und Zelloberflächen-Display-Techniken eignen sich diese Systeme vorzüglich für Selektionsverfahren, die auf Wechselwirkungen zwischen Rezeptor und Ligand beruhen. So wurden vor kurzem vier ATP-Binder durch die mRNA-Peptid-Fusionsstrategie ausgehend von einer Bibliothek aus 6×10^{12} Proteinen selektiert, in denen 80 benachbarte Aminosäuren randomisiert wurden.^[228] Für das rasche Auffinden von miteinander wechselwirkenden Proteinen in Organismen, die der genetischen Analyse schwer zugänglich sind, wurde vorgeschlagen, Genbanken zu verwenden, die aus genomischer oder cDNA aus natürlichen Quellen hergestellt wurden.^[205] Der Einsatz einiger der hier diskutierten In-vitro-Transkriptions/Translations-Systeme in

Kombination mit der Selektion von Proteinen auf der Grundlage ihrer katalytischen Aktivität (siehe Abschnitt 5.1) könnte darüber hinaus die gesteuerte Evolution neuer Enzymaktivitäten gestatten.

5.4. RNA-Evolutionsexperimente

Der Wert großer Genbanken zeigt sich besonders bei der direkten Selektion auf RNA- und DNA-Moleküle, die als Rezeptoren und Katalysatoren dienen können.^[210] Experimente dieser Art erfordern gewöhnlich die Herstellung und das Screening von Molekülbibliotheken aus bis zu 10^{15} unabhängigen Mitgliedern. Da Nucleinsäuren Struktur, Funktion und genetische Information in einem Molekül vereinen, ist die In-vitro-Gewinnung und -Vermehrung seltener Gene für einen gewünschten Phänotyp unkompliziert und sehr effizient (Schema 20). Deshalb können Nucleinsäure-Populationen leicht vielen Cyclen aus Selektion und Anreicherung unterworfen werden.



Schema 20. In-vitro-Evolution von RNA-Katalysatoren: Eine durch In-vitro-Transkription erzeugte kombinatorische Genbank von RNAs wird der Selektion auf Ligandenbindung oder Katalyse unterworfen. Moleküle, die die gewünschten Merkmale aufweisen, werden von inaktiven Varianten getrennt, um einen neuen, an funktionellen Sequenzen angereicherten RNA-Pool zu erhalten. Dieser wird revers in cDNA transkribiert und mit der PCR-Methode vervielfältigt, um Material für die nächste Selektionsrunde zu erhalten. Der Zyklus wird so lange wiederholt bis sich genügend aktive Sequenzen im Reservoir angereichert haben, die für eine eingehende Charakterisierung kloniert werden können. Wenn eine fehlerbehaftete PCR-Methode eingesetzt wird,^[3] repräsentiert jede Runde einen echten darwinistischen Evolutionszyklus.

RNA- und DNA-Katalysatoren wurden für eine ganze Palette von Reaktionen hergestellt, wie Bindungsspaltungen, Ligationen und Modifizierungen von Oligonucleotiden, Alkylierungen, Acyltransferreaktionen und sogar in der Natur nicht vorkommenden Diels-Alder-Cycloadditionen.^[210] Obwohl die Selektion funktioneller Nucleinsäuren einfacher zu sein scheint als die funktioneller Proteine (siehe z.B. Ab-

schnitt 3.6), ist die Effizienz der resultierenden Katalysatoren gegenüber der hochentwickelter Enzyme häufig nur recht mäßig. Der Grund dafür ist vielleicht darin zu suchen, dass sie aus nur vier Nucleotiden aufgebaut sind. Dennoch wurden in einigen Fällen außerordentlich hohe Aktivitäten erzielt.

Besonders die Übertragung von Phosphatgruppen eignet sich für die Ribozymkatalyse. Die direkte Selektion von sehr leistungsfähigen selbstligierenden Ribozymen aus großen Gemischen randomisierter Sequenzen ist in diesem Zusammenhang eine der eindrucksvollsten Errungenschaften.^[211] Der Ausgangspunkt war eine Genbank aus ca. 10^{15} verschiedenen RNAs mit jeweils 220 randomisierten Nucleotiden. Ungefähr eine von 10^{13} randomisierten Sequenzen verfügte über die gewünschte Ligasefunktion. Es ist klar, dass die Chance, einen Katalysator zu entdecken, verschwindend gering gewesen wäre, hätte die anfängliche Genbank „nur“ 10^{10} Sequenzen umfasst. Die Population ligierender Ribozyme wurde durch zusätzliche Diversifikation (Mutagenese) und Selektion weiterentwickelt. Eines der resultierenden Ribozyme wurde anschließend zu einem echten Katalysator umgebaut, der eine intermolekulare Ligation mit hoher Wechselzahl vermittelt.^[212] Seine Geschwindigkeitskonstante ist größer als 1 s^{-1} ; dies entspricht einer Beschleunigung um einen Faktor von fast 10^9 im Vergleich zur nichtkatalysierten Reaktion.

Nucleinsäureliganden, -rezeptoren und -katalysatoren sind neuartige Analoga ihrer Peptid/Protein-Pendants mit vielen potentiellen Anwendungen als Medikamente oder in der Diagnostik. Die verwendeten Strategien zu ihrer Herstellung, das Verständnis, wie sie aus randomisierten Sequenzen hervorgehen, sowie ein Einblick in ihre Vor- und Nachteile im Vergleich mit den Polypeptiden werden uns wahrscheinlich auch in Zukunft helfen, unsere Fähigkeiten zur Herstellung und Optimierung neuer Protein-Katalysatoren zu verbessern.

6. Zusammenfassung und Ausblick

Die Natur hat das Problem der Gestaltung von Proteinen durch den Prozess der darwinistischen Evolution gelöst. Jedes einzelne Protein in unseren Zellen, angefangen bei den Enzymen und Rezeptoren bis hin zu den Regulations- und Strukturproteinen, ist aus diesem Verfahren hervorgegangen. Wie die in diesem Übersichtsartikel präsentierten Beispiele ausführlich belegen, können evolutionäre Strategien auch im Labor erfolgreich zur Untersuchung der Struktur und Funktion bereits existierender Proteine sowie zur Entwicklung neuer eingesetzt werden.

Die zielgerichtete Mutagenese, bei der eine Aminosäure eines Proteins durch eine andere ausgetauscht wird, ist ein etabliertes und unverzichtbares Werkzeug in der Proteinchemie. Die Größe und Komplexität von Proteinen macht jedoch häufig eine eindeutige Interpretation der Resultate solcher Experimente unmöglich. Dies ist eine Folge der ungleichmäßigen Verteilung der Strukturinformationen innerhalb der Aminosäuresequenz des Proteins – der Tatsache, dass einige Reste Substitutionen tolerieren, wohingegen andere nicht ohne nachteilige Auswirkungen ersetzt werden können – und der zahlreichen energetisch ähnlichen Konformationen, die

jedes Polypeptid annehmen kann. Diese Faktoren und die ausgeprägte Empfindlichkeit der katalytischen Aktivität gegenüber scheinbar geringen Störungen in der Struktur machen die Charakterisierung und den (Um)bau von Enzymen im Labor so schwierig.

Im Unterschied zu Methoden, die auf Substitutionen einzelner Aminosäuren basieren, sind kombinatorische Methoden ausgezeichnet zur Bewertung des komplexen und letztlich Form und Funktion bestimmenden Zusammenspiels der Wechselwirkungen in Proteinen geeignet.^[73] Wie anhand der Chorismat-Mutasen gezeigt (Abschnitte 3.2 und 3.3) können Ensembles degenerierter Proteine subtile Tendenzen, die beim individuellen Austausch von Aminosäuren übersehen werden könnten, aufzeigen. Die erhaltenen Informationen können unser Verständnis über Proteinstruktur und -funktion verifizieren und weiter vertiefen. Evolutionäre Strategien können zusammen mit Informationen über die Struktur und den Mechanismus ebenfalls beim Proteindesign zum Zuge kommen,^[8] wie dies bei der Herstellung neuer monomerer und binär gemusterter Mutasen (Abschnitte 3.4–3.6) sowie bei der Veränderung der Substratspezifität von Aspartat-Aminotransferase (Abschnitt 4.1) und der Funktion von IGPS (Abschnitt 4.3) der Fall war. Das schrittweise Vorgehen bei evolutionären Methoden ist in diesem Zusammenhang von besonderem Vorteil, weil dadurch die schnelle Optimierung der untersuchten Molekülpopulation auf ein gewünschtes Ziel hin möglich ist. Sogar ohne detaillierte Kenntnisse der Struktur oder ohne genau zu verstehen, wie ein System funktioniert, kann die gesteuerte Evolution angewandt werden, um die Eigenschaften von Proteinen für gezielte praktische Anwendungen maßzuschneidern (Abschnitt 4.4).

Selbstverständlich benötigt man für die erfolgreiche Anwendung evolutionärer Strategien eine leistungsfähige Methode, um die vielen Mitglieder einer Bibliothek zu untersuchen und diejenigen mit den begehrten Eigenschaften, sei es die Erkennung eines bestimmten Liganden oder die Katalyse einer interessanten chemischen Umwandlung, zu identifizieren. Bei kleinen Bibliotheken lassen sich die nützlichen Varianten bereits durch ein Screeningverfahren auffinden. Auch für große Bibliotheken ist – abhängig von der Schwierigkeit der gestellten Aufgabe und der Entfernung, die im Sequenzraum zwischen der Ausgangsstruktur und dem gewünschten Endpunkt liegt^[213] – ein Screening oft mehr als ausreichend. Beispielsweise konnten enantioselektive Esterasen und Lipasen,^[214, 215] thermostabile Hydrolasen^[216, 217] sowie metallabhängige Enzyme mit neuartigen Spezifitäten^[218, 219] nach einer Zufallsmutagenese durch Screening identifiziert werden. Darüber hinaus ist zu erwarten, dass Fortschritte in der Automatisierung und Miniaturisierung die Anwendbarkeit und Verbreitung von Screeningmethoden für Hochdurchsatz-Analysen relativ großer Molekülembles fördern.^[10, 11, 195, 198]

Für Aufgabenstellungen, bei denen statistisch aussagefähige Zusammenhänge gefragt oder für die Lösungen außerordentlich selten sind, ist die genetische Selektion wahrscheinlich die Methode der Wahl. In-vivo-Selektionsstrategien gestatten die gründliche Suche in Genbanken mit 10^{10} Varianten, und noch erheblich größere Genbanken sind In-

vitro-Methoden zugänglich (Abschnitte 5.2–5.4). Die grundätzliche Herausforderung bei der Entwicklung eines Selektionsverfahrens ist die Verknüpfung der interessierenden makromolekularen Eigenschaften mit einer klar erfassbaren biologischen Reaktion, wodurch die einfache Gewinnung des verantwortlichen kodierenden Gens ermöglicht wird. Prinzipiell eignet sich hierfür jeder Stoffwechselprozess, ebenso auch Transformationen, die zur Bildung oder Zerstörung lebensnotwendiger Nährstoffe oder Toxine führen (Abschnitte 3 und 4). Raffinierte In-vitro-Systeme erweitern die Palette der Möglichkeiten zur Selektion von Molekülen mit interessanten katalytischen oder Bindungseigenschaften (Abschnitt 5).

Zwar sind Evolutionsexperimente im Labor inzwischen fest etabliert, doch könnten Verbesserungen auf einigen Gebieten den Nutzen solcher Methoden für verschiedene praktische Anwendungen noch erheblich erhöhen. Verbesserte Verfahren für die Mutagenese^[134, 220, 221] und die Transformation von Mikroorganismen^[23, 222] z.B. könnten die Herstellung von noch geschickter aufgebauten und noch größeren Proteinbanken erlauben. Noch vielseitigere und weit allgemeiner anwendbare Selektionsstrategien, die die Selektion auf beliebige Funktionen gestatten, wären ebenfalls von großem Wert. Des Weiteren ist es ein Wunschtraum, die Prinzipien der Evolution über die Nucleinsäuren und Proteine hinausgehend auf eine breitere Palette von Verbindungen auszuweiten. Die Realisierung eines solch ehrgeizigen Ziels würde die Entwicklung von neuartigen kodierten kombinatorischen Bibliotheken erfordern sowie Strategien, um die gewünschten Verbindungen aus solchen Ensembles zu isolieren und gleichzeitig Informationen über ihren Aufbau zu erhalten. In diesem Zusammenhang werden neue Methoden für die „autokatalytische“ Detektion und die selektive Vervielfältigung der „besten“ Moleküle nach jedem Evolutionszyklus benötigt.

Evolutionäre Methoden werden „rationale“ Ansätze zur Untersuchung von Makromolekülen nie verdrängen können. Vielmehr ergänzen sie die klassischen Strategien der Analyse und des Designs. In der Tat scheint die gesteuerte Evolution am leistungsfähigsten zur Erkundung und Entwicklung der Eigenschaften von Proteinen, Nucleinsäuren und anderen kodierten Verbindungen zu sein, wenn sie in Kombination mit detaillierten strukturellen und mechanistischen Informationen angewendet wird. In der „Post-genomic“-Ära werden diese Methoden sehr wahrscheinlich zunehmend an Wert für die Analyse und die Bewertung der Datenflut gewinnen, die heute bei der Untersuchung der vielen biologischen Systeme entsteht. Für die Zukunft können wir erwarten, dass die intelligente Anwendung von evolutionären Methoden helfen wird, die Zusammenhänge zwischen Sequenz und Struktur, Struktur und Funktion sowie zwischen miteinander wechselwirkenden Makromolekülen in der Zelle aufzuklären. Neue Anwendungen in der Medizin und der Industrie werden wohl in kurzem Abstand folgen.

Die Autoren danken Kinya Hotta und Gavin MacBeath für ihre Hilfe beim Erstellen der Graphiken, Kai Walter für seine Mithilfe bei der Übersetzung ins Deutsche sowie der ETH

Zürich, dem Schweizerischen Nationalfonds zur Förderung der wissenschaftlichen Forschung und Novartis Pharma für die finanzielle Unterstützung.

Eingegangen am 25. Januar 2001 [A 446]

- [1] R. Dawkins, *The Blind Watchmaker*, Longham, London **1986**.
- [2] *In vitro Mutagenesis Protocols* (Hrsg.: M. K. Trower), Humana, Totowa, NJ, **1996**.
- [3] R. C. Cadwell, G. F. Joyce, *PCR Methods Appl.* **1992**, 2, 28–33.
- [4] W. P. C. Stemmer, *Bio/Technology* **1995**, 13, 549–553.
- [5] W. P. C. Stemmer, *Proc. Natl. Acad. Sci. USA* **1994**, 91, 10747–10751.
- [6] *Methods Enzymol.* **1996**, 267 (Hrsg.: J. N. Abelson).
- [7] P. Kast, D. Hilvert, *Pure Appl. Chem.* **1996**, 68, 2017–2024.
- [8] P. Kast, D. Hilvert, *Curr. Opin. Struct. Biol.* **1997**, 7, 470–479.
- [9] R. T. Sauer, *Fold. Des.* **1996**, 1, R27–R30.
- [10] F. H. Arnold, *Acc. Chem. Res.* **1998**, 31, 125–131.
- [11] F. H. Arnold, A. A. Volkov, *Curr. Opin. Chem. Biol.* **1999**, 3, 54–59.
- [12] D. C. Demirjian, P. C. Shah, F. Moris-Varas, *Top. Curr. Chem.* **1999**, 200, 1–29.
- [13] U. Kettling, A. Koltermann, M. Eigen, *Top. Curr. Microbiol. Immunol.* **1999**, 243, 173–186.
- [14] A. Koltermann, U. Kettling, *Biophys. Chem.* **1997**, 66, 159–177.
- [15] L. P. Encell, D. M. Landis, L. A. Loeb, *Nat. Biotechnol.* **1999**, 17, 143–147.
- [16] A. Skandalis, L. P. Encell, L. A. Loeb, *Chem. Biol.* **1997**, 4, 889–898.
- [17] J. Fastrez, *Mol. Biotechnol.* **1997**, 7, 37–55.
- [18] B. Lewin, *Genes VII*, Oxford University Press, London, **2000**.
- [19] J. D. Watson, M. Gilman, J. Witkowski, M. Zoller, *Recombinant DNA: A Short Course*, W. H. Freeman, San Francisco, **1992**.
- [20] J. Sambrook, D. W. Russell, *Molecular Cloning. A Laboratory Manual*, 3. Aufl., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, **2001**.
- [21] P. B. Fernandes, *Curr. Opin. Chem. Biol.* **1998**, 2, 597–603.
- [22] A. Qureshi-Emili, G. Cagney, *Nat. Biotechnol.* **2000**, 18, 393–397.
- [23] D. Hanahan, J. Jessee, F. R. Bloom, *Methods Enzymol.* **1991**, 204, 63–113.
- [24] L. You, F. H. Arnold, *Protein Eng.* **1996**, 9, 77–83.
- [25] F. H. Arnold, J. C. Moore, *Adv. Biochem. Eng. Biotechnol.* **1997**, 58, 1–14.
- [26] C. M. Gates, W. P. C. Stemmer, R. Kaptein, P. J. Schatz, *J. Mol. Biol.* **1996**, 255, 373–386.
- [27] R. R. Breaker, A. Banerji, G. F. Joyce, *Biochemistry* **1994**, 33, 11980–11986.
- [28] U. Weiss, J. M. Edwards, *The Biosynthesis of Aromatic Amino Compounds*, Wiley, New York, **1980**.
- [29] E. Haslam, *Shikimic Acid: Metabolism and Metabolites*, Wiley, New York, **1993**.
- [30] B. Ganem in *Comprehensive Natural Products Chemistry*, Vol. 5 (Hrsg.: C. D. Poulter), Elsevier, Oxford, **1999**, S. 343–369.
- [31] Y. M. Chook, J. V. Gray, H. Ke, W. N. Lipscomb, *J. Mol. Biol.* **1994**, 240, 476–500.
- [32] Y. M. Chook, H. Ke, W. N. Lipscomb, *Proc. Natl. Acad. Sci. USA* **1993**, 90, 8600–8603.
- [33] A. Y. Lee, P. A. Karplus, B. Ganem, J. Clardy, *J. Am. Chem. Soc.* **1995**, 117, 3627–3628.
- [34] N. Sträter, G. Schnappauf, G. Braus, W. N. Lipscomb, *Structure* **1997**, 5, 1437–1452.
- [35] Y. Xue, W. N. Lipscomb, R. Graf, G. Schnappauf, G. Braus, *Proc. Natl. Acad. Sci. USA* **1994**, 91, 10814–10818.
- [36] W. Gu, D. S. Williams, H. C. Aldrich, G. Xie, D. W. Gabriel, R. A. Jensen, *Microb. Comp. Genomics* **1997**, 2, 141–158.
- [37] G. MacBeath, P. Kast, D. Hilvert, *Biochemistry* **1998**, 37, 10062–10073.
- [38] P. Kast, C. Grisostomi, I. A. Chen, S. Li, U. Krengel, Y. Xue, D. Hilvert, *J. Biol. Chem.* **2000**, 275, 36832–36838.
- [39] D. Hilvert, S. H. Carpenter, K. D. Nared, M.-T. M. Auditor, *Proc. Natl. Acad. Sci. USA* **1988**, 85, 4953–4955.

- [40] D. Hilvert, K. D. Nared, *J. Am. Chem. Soc.* **1988**, *110*, 5593–5594.
- [41] D. Y. Jackson, J. W. Jacobs, R. Sugasawara, S. H. Reich, P. A. Bartlett, P. G. Schultz, *J. Am. Chem. Soc.* **1988**, *110*, 4841–4842.
- [42] M. R. Haynes, E. A. Stura, D. Hilvert, I. A. Wilson, *Science* **1994**, *263*, 646–652.
- [43] K. Bowdish, Y. Tang, J. B. Hicks, D. Hilvert, *J. Biol. Chem.* **1991**, *266*, 11901–11908.
- [44] Y. Tang, J. B. Hicks, D. Hilvert, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 8784–8786.
- [45] P. Kast, M. Asif-Ullah, N. Jiang, D. Hilvert, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 5043–5048.
- [46] P. Kast, M. Asif-Ullah, D. Hilvert, *Tetrahedron Lett.* **1996**, *37*, 2691–2694.
- [47] G. Zhao, T. Xia, R. S. Fischer, R. A. Jensen, *J. Biol. Chem.* **1992**, *267*, 2487–2493.
- [48] T. Xia, G. Zhao, R. S. Fischer, R. A. Jensen, *J. Gen. Microbiol.* **1992**, *138*, 1309–1316.
- [49] L. Addadi, E. K. Jaffe, J. R. Knowles, *Biochemistry* **1983**, *22*, 4494–4501.
- [50] S. D. Copley, J. R. Knowles, *J. Am. Chem. Soc.* **1985**, *107*, 5306–5308.
- [51] S. D. Copley, J. R. Knowles, *J. Am. Chem. Soc.* **1987**, *109*, 5008–5013.
- [52] J. J. Gajewski, J. Jurayj, D. R. Kimbrough, M. E. Ganem, B. Ganem, B. K. Carpenter, *J. Am. Chem. Soc.* **1987**, *109*, 1170–1186.
- [53] O. Wiest, K. N. Houk, *J. Org. Chem.* **1994**, *59*, 7582–7584.
- [54] O. Wiest, K. N. Houk, *J. Am. Chem. Soc.* **1995**, *117*, 11628–11639.
- [55] P. D. Lyne, A. J. Mulholland, W. G. Richards, *J. Am. Chem. Soc.* **1995**, *117*, 11345–11350.
- [56] M. M. Davidson, I. R. Gould, I. H. Hillier, *J. Chem. Soc. Perkin Trans.* **1996**, *2*, 525–532.
- [57] M. M. Davidson, I. H. Hillier, *J. Chem. Soc. Perkin Trans.* **1994**, *2*, 1415–1417.
- [58] D. L. Severence, W. L. Jorgensen, *J. Am. Chem. Soc.* **1992**, *114*, 10966–10968.
- [59] H. A. Carlson, W. L. Jorgensen, *J. Am. Chem. Soc.* **1996**, *118*, 8475–8484.
- [60] P. Kast, Y. B. Tewari, O. Wiest, D. Hilvert, K. N. Houk, R. N. Goldberg, *J. Phys. Chem. B* **1997**, *101*, 10976–10982.
- [61] S. G. Sogo, T. S. Widlanski, J. H. Hoare, C. E. Grimshaw, G. A. Berchtold, J. R. Knowles, *J. Am. Chem. Soc.* **1984**, *106*, 2701–2703.
- [62] P. A. Bartlett, C. R. Johnson, *J. Am. Chem. Soc.* **1985**, *107*, 7792–7793.
- [63] P. A. Bartlett, Y. Nakagawa, C. R. Johnson, S. H. Reich, A. Luis, *J. Org. Chem.* **1988**, *53*, 3195–3210.
- [64] J. V. Gray, D. Eren, J. R. Knowles, *Biochemistry* **1990**, *29*, 8872–8878.
- [65] D. J. Gustin, P. Mattei, P. Kast, O. Wiest, L. Lee, W. W. Cleland, D. Hilvert, *J. Am. Chem. Soc.* **1999**, *121*, 1756–1757.
- [66] S. T. Cload, D. R. Liu, R. M. Pastor, P. G. Schultz, *J. Am. Chem. Soc.* **1996**, *118*, 1787–1788.
- [67] D. R. Liu, S. T. Cload, R. M. Pastor, P. G. Schultz, *J. Am. Chem. Soc.* **1996**, *118*, 1789–1790.
- [68] S. Zhang, P. Kongsaeree, J. Clardy, D. B. Wilson, B. Ganem, *Bioorg. Med. Chem.* **1996**, *4*, 1015–1020.
- [69] „Evolutionary Studies with a Catalytic Antibody“: Y. Tang, Dissertation, The Scripps Research Institute, **1996**.
- [70] M. Gamper, D. Hilvert, P. Kast, *Biochemistry* **2000**, *39*, 14087–14094.
- [71] J. V. Gray, J. R. Knowles, *Biochemistry* **1994**, *33*, 9953–9959.
- [72] M. H. J. Cordes, A. R. Davidson, R. T. Sauer, *Curr. Opin. Struct. Biol.* **1996**, *6*, 3–10.
- [73] J. F. Reidhaar-Olson, R. T. Sauer, *Science* **1988**, *241*, 53–57.
- [74] M. Blaber, X. Zhang, B. W. Matthews, *Science* **1993**, *260*, 1637–1640.
- [75] H. Xiong, B. L. Buckwalter, H. M. Shieh, M. H. Hecht, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 6349–6353.
- [76] A. P. Brunet, E. S. Huang, M. E. Huffine, J. E. Loeb, R. J. Weltman, M. H. Hecht, *Nature* **1993**, *364*, 355–358.
- [77] L. Castagnoli, C. Vetriani, G. Cesareni, *J. Mol. Biol.* **1994**, *237*, 378–387.
- [78] J. A. Ybe, M. H. Hecht, *Protein Sci.* **1996**, *5*, 814–824.
- [79] W. A. Lim, R. T. Sauer, *Nature* **1989**, *339*, 31–36.
- [80] M. E. Milla, R. T. Sauer, *Biochemistry* **1995**, *34*, 3344–3351.
- [81] E. P. Baldwin, O. Hajiseyedjavadi, W. A. Baase, B. W. Matthews, *Science* **1993**, *262*, 1715–1718.
- [82] Z. L. Fredericks, G. J. Pielak, *Biochemistry* **1993**, *32*, 929–936.
- [83] P. F. Predki, L. Regan, *Biochemistry* **1995**, *34*, 9834–9839.
- [84] P. F. Predki, V. Agrawal, A. T. Brünger, L. Regan, *Nat. Struct. Biol.* **1996**, *3*, 54–58.
- [85] G. MacBeath, P. Kast, D. Hilvert, *Protein Sci.* **1998**, *7*, 325–335.
- [86] S. M. Green, A. G. Gittis, A. K. Meeker, E. E. Lattman, *Nat. Struct. Biol.* **1995**, *2*, 746–751.
- [87] R. R. Dickason, D. P. Huston, *Nature* **1996**, *379*, 652–655.
- [88] R. A. Albright, M. C. Mossing, B. W. Matthews, *Biochemistry* **1996**, *35*, 735–742.
- [89] T. V. Borchert, R. Abagyan, R. Jaenicke, R. K. Wierenga, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 1515–1518.
- [90] M. C. Mossing, R. T. Sauer, *Science* **1990**, *250*, 1712–1715.
- [91] W. Schliebs, N. Thanki, R. Jaenicke, R. K. Wierenga, *Biochemistry* **1997**, *36*, 9655–9662.
- [92] X. Shao, P. Hensley, C. R. Matthews, *Biochemistry* **1997**, *36*, 9941–9949.
- [93] M. J. Bennett, S. Choe, D. Eisenberg, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3127–3131.
- [94] M. J. Bennett, M. P. Schlunegger, D. Eisenberg, *Protein Sci.* **1995**, *4*, 2455–2468.
- [95] G. MacBeath, P. Kast, D. Hilvert, *Protein Sci.* **1998**, *7*, 1757–1767.
- [96] L. Lins, R. Brasseur, *FASEB J.* **1995**, *9*, 535–540.
- [97] G. MacBeath, P. Kast, D. Hilvert, *Science* **1998**, *279*, 1958–1961.
- [98] M. G. Oakley, P. S. Kim, *Biochemistry* **1997**, *36*, 2544–2549.
- [99] K. U. Walter, S. V. Taylor, G. Mäder, P. Kast, D. Hilvert, unveröffentlichte Ergebnisse.
- [100] W. P. C. Stemmer, *Nature* **1994**, *370*, 389–391.
- [101] M. Couturier, F. Bex, W. K. Maas, *Microbiol. Rev.* **1988**, *52*, 375–395.
- [102] T. A. Y. Ayoubi, W. J. M. VanDeVen, *FASEB J.* **1996**, *10*, 453–460.
- [103] H. A. DeBoer, A. S. Hui, *Methods Enzymol.* **1990**, *185*, 103–114.
- [104] G. R. Dieckmann, D. K. McRorie, D. L. Tierney, L. M. Utschig, C. P. Singer, T. V. O'Halloran, J. E. Penner-Hahn, W. F. DeGrado, V. L. Pecoraro, *J. Am. Chem. Soc.* **1997**, *119*, 6195–6196.
- [105] S. F. Betz, W. F. DeGrado, *Biochemistry* **1996**, *35*, 6955–6962.
- [106] S. Dalal, S. Balasubramanian, L. Regan, *Nat. Struct. Biol.* **1997**, *4*, 548–552.
- [107] J. E. Rozelle, A. Tropsha, B. W. Erickson, *Protein Sci.* **1994**, *3*, 345–355.
- [108] K. M. Gernert, M. C. Surles, T. H. Labean, J. S. Richardson, D. C. Richardson, *Protein Sci.* **1995**, *4*, 2252–2260.
- [109] P. B. Harbury, J. J. Plecs, B. Tidor, T. Alber, P. S. Kim, *Science* **1998**, *282*, 1462–1467.
- [110] X. Jiang, E. J. Bishop, R. S. Farid, *J. Am. Chem. Soc.* **1997**, *119*, 838–839.
- [111] B. I. Dahiyat, S. L. Mayo, *Science* **1997**, *278*, 82–87.
- [112] S. M. Malakauskas, S. L. Mayo, *Nat. Struct. Biol.* **1998**, *5*, 470–475.
- [113] A. R. Davidson, R. T. Sauer, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 2146–2150.
- [114] M. W. West, M. H. Hecht, *Protein Sci.* **1995**, *4*, 2032–2039.
- [115] J. U. Bowie, R. T. Sauer, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 2152–2156.
- [116] S. Kamtekar, J. M. Schiffer, H. Xiong, J. M. Babik, M. H. Hecht, *Science* **1993**, *262*, 1680–1685.
- [117] S. Roy, G. Ratnaswamy, J. A. Boice, R. Fairman, G. McLendon, M. H. Hecht, *J. Am. Chem. Soc.* **1997**, *119*, 5302–5306.
- [118] S. Roy, M. H. Hecht, *Biochemistry* **2000**, *39*, 4603–4607.
- [119] S. V. Taylor, K. U. Walter, P. Kast, D. Hilvert, *Proc. Natl. Acad. Sci. USA* **2001**, im Druck.
- [120] E. A. Kabat, *Structural Concepts in Immunology and Immunochemistry*, Holt, Rinehart, and Winston, New York, **1976**.
- [121] J. T.-F. Wong, *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 1909–1912.
- [122] C. Woese, *The Genetic Code*, Harper & Row, New York, **1967**.
- [123] F. H. C. Crick, *J. Mol. Biol.* **1968**, *38*, 367–379.
- [124] W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri, A. Lombardi, *Annu. Rev. Biochem.* **1999**, *68*, 779–819.
- [125] C.-H. Wong, G. M. Whitesides, *Enzymes in Synthetic Organic Chemistry*, Elsevier Science, Tarrytown, **1994**.
- [126] H. M. Wilks, K. W. Hart, R. Feeney, C. R. Dunn, H. Muirhead, W. N. Chia, D. A. Barstow, T. Atkinson, A. R. Clarke, J. J. Holbrook, *Science* **1988**, *242*, 1541–1544.
- [127] L. Hedstrom, L. Szilagyi, W. J. Rutter, *Science* **1992**, *255*, 1249–1253.

- [128] J. J. Onuffer, J. F. Kirsch, *Protein Sci.* **1995**, 4, 1750–1757.
- [129] T. Yano, S. Oue, H. Kagamiyama, *Proc. Natl. Acad. Sci. USA* **1998**, 95, 5511–5515.
- [130] S. Oue, A. Okamoto, T. Yano, H. Kagamiyama, *J. Biol. Chem.* **1999**, 274, 2344–2349.
- [131] P. A. Patten, N. S. Gray, P. L. Yang, C. B. Marks, G. J. Wedemayer, J. J. Boniface, R. C. Stevens, P. G. Schultz, *Science* **1996**, 271, 1086–1091.
- [132] G. J. Wedemayer, P. A. Patten, L. H. Wang, P. G. Schultz, R. C. Stevens, *Science* **1997**, 276, 1665–1669.
- [133] I. M. Tomlinson, G. Walter, P. T. Jones, P. H. Dear, E. L. L. Sonnhammer, G. Winter, *J. Mol. Biol.* **1996**, 256, 813–817.
- [134] A. Cramer, S. A. Raillard, E. Bermudez, W. P. C. Stemmer, *Nature* **1998**, 391, 288–291.
- [135] H. Viadiu, J. Osuna, A. L. Fink, X. Soberon, *J. Biol. Chem.* **1995**, 270, 781–787.
- [136] M. Zaccole, E. Gherardi, *J. Mol. Biol.* **1999**, 285, 775–783.
- [137] J. F. Petrosino, T. Palzkil, *J. Bacteriol.* **1996**, 178, 1821–1828.
- [138] W. Huang, J. Petrosino, M. Hirsch, P. S. Shenkin, T. Palzkil, *J. Mol. Biol.* **1996**, 258, 688–703.
- [139] C. Cantu, W. Huang, T. Palzkil, *J. Biol. Chem.* **1996**, 271, 22538–22545.
- [140] K. M. Munir, D. C. French, D. K. Dube, L. A. Loeb, *Protein Eng.* **1994**, 7, 83–89.
- [141] S.-I. Ohnuma, K. Hirooka, H. Hemmi, C. Ishida, C. Ohto, T. Nishino, *J. Biol. Chem.* **1996**, 271, 18831–18837.
- [142] E. A. Hart, L. Hua, L. B. Darr, W. K. Wilson, J. H. Pang, S. P. T. Matsuda, *J. Am. Chem. Soc.* **1999**, 121, 9887–9888.
- [143] C. Jürgens, A. Strom, D. Wegener, S. Hettwer, M. Wilmanns, R. Sterner, *Proc. Natl. Acad. Sci. USA* **2000**, 97, 9925–9930.
- [144] D. R. Liu, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **1999**, 96, 4780–4785.
- [145] D. R. Liu, T. J. Magliery, M. Pastrnak, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **1997**, 94, 10092–10097.
- [146] M. W. W. Adams, R. M. Kelly, *Trends Biotechnol.* **1998**, 16, 329–332.
- [147] T. Oshima, *Curr. Opin. Struct. Biol.* **1994**, 4, 623–628.
- [148] H. Liao, T. McKenzie, R. Hageman, *Proc. Natl. Acad. Sci. USA* **1986**, 83, 576–580.
- [149] M. Matsumura, S. Aiba, *J. Biol. Chem.* **1985**, 260, 15298–15303.
- [150] J. Hoseki, T. Yano, Y. Koyama, S. Kuramitsu, H. Kagamiyama, *J. Biochem.* **1999**, 126, 951–956.
- [151] M. Matsumura, S. Yasumura, S. Aiba, *Nature* **1986**, 323, 356–358.
- [152] Y. I. Wolf, N. V. Grishin, E. V. Koonin, *J. Mol. Biol.* **2000**, 299, 897–905.
- [153] D. R. Davies, E. A. Padlan, S. Sheriff, *Annu. Rev. Biochem.* **1990**, 59, 439–473.
- [154] P. G. Schultz, R. A. Lerner, *Science* **1995**, 269, 1835–1842.
- [155] D. Hilvert, *Annu. Rev. Biochem.* **2000**, 69, 751–793.
- [156] R. A. Lerner, S. J. Benkovic, P. G. Schultz, *Science* **1991**, 252, 659–667.
- [157] C. Bränden, J. Tooze, *Introduction to Protein Structure*, Garland, New York, **1999**.
- [158] C. I. Bränden, *Curr. Opin. Struct. Biol.* **1991**, 1, 978–983.
- [159] M. M. Altamirano, J. M. Blackburn, C. Aguayo, A. R. Fersht, *Nature* **2000**, 403, 617–622.
- [160] R. Sterner, G. R. Kleemann, H. Szadkowski, A. Lustig, M. Hennig, K. Kirschner, *Protein Sci.* **1996**, 5, 2000–2008.
- [161] G. Winter, C. Milstein, *Nature* **1991**, 349, 293–299.
- [162] A. Cramer, G. Dawes, E. Rodriguez, Jr., S. Silver, W. P. C. Stemmer, *Nat. Biotechnol.* **1997**, 15, 436–438.
- [163] S. Fields, O.-K. Song, *Nature* **1989**, 340, 245–246.
- [164] P. Colas, R. Brent, *Trends Biotechnol.* **1998**, 16, 355–363.
- [165] R. M. Frederickson, *Curr. Opin. Biotechnol.* **1998**, 9, 90–96.
- [166] R. K. Brachmann, J. D. Boeke, *Curr. Opin. Biotechnol.* **1997**, 8, 561–568.
- [167] 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, 623–627.
- [168] P. Colas, B. Cohen, T. Jessen, I. Grishina, J. McCoy, R. Brent, *Nature* **1996**, 380, 548–550.
- [169] E. J. Licitra, J. O. Liu, *Proc. Natl. Acad. Sci. USA* **1996**, 93, 12817–12821.
- [170] J. C. Hu, M. G. Kornacker, A. Hochschild, *Methods* **2000**, 20, 80–94.
- [171] G. Karimova, J. Pidoux, A. Ullmann, D. Ladant, *Proc. Natl. Acad. Sci. USA* **1998**, 95, 5752–5756.
- [172] 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, 591–595.
- [173] S.-H. Park, R. T. Raines, *Nat. Biotechnol.* **2000**, 18, 847–851.
- [174] J. A. Francisco, R. Campbell, B. L. Iverson, G. Georgiou, *Proc. Natl. Acad. Sci. USA* **1993**, 90, 10444–10448.
- [175] E. T. Boder, K. D. Wittrup, *Nat. Biotechnol.* **1997**, 15, 553–557.
- [176] *Phage Display of Peptides and Proteins* (Hrsg.: B. K. Kay, J. Winter, J. McCafferty), Academic Press, New York, **1996**.
- [177] G. P. Smith, V. A. Petrenko, *Chem. Rev.* **1997**, 97, 391–410.
- [178] S. Chusacultanachai, K. A. Glenn, A. O. Rodriguez, E. K. Read, J. F. Gardner, B. S. Katzenellenbogen, D. J. Shapiro, *J. Biol. Chem.* **1999**, 274, 23591–23598.
- [179] H. A. Greisman, C. O. Pabo, *Science* **1997**, 275, 657–661.
- [180] S. A. Wolfe, E. I. Ramm, C. O. Pabo, *Struct. Fold. Des.* **2000**, 8, 739–750.
- [181] T. E. Wilson, T. J. Fahrner, M. Johnston, J. Milbrandt, *Science* **1991**, 252, 1296–1300.
- [182] R. R. Beerli, D. J. Segal, B. Dreier, C. F. Barbas, *Proc. Natl. Acad. Sci. USA* **1998**, 95, 14628–14633.
- [183] A. C. Jamieson, S.-H. Kim, J. A. Wells, *Biochemistry* **1994**, 33, 5689–5695.
- [184] E. J. Rebar, C. O. Pabo, *Science* **1994**, 263, 671–673.
- [185] Q. Liu, D. J. Segal, J. B. Ghira, C. F. Barbas, *Proc. Natl. Acad. Sci. USA* **1997**, 94, 5525–5530.
- [186] A. C. Braisted, J. A. Wells, *Proc. Natl. Acad. Sci. USA* **1996**, 93, 5688–5692.
- [187] E. T. Boder, K. S. Midelfort, K. D. Wittrup, *Proc. Natl. Acad. Sci. USA* **2000**, 97, 10701–10705.
- [188] 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, 11777–11782.
- [189] P. Soumillion, L. Jespers, M. Bouchet, J. Marchand-Brynaert, G. Winter, J. Fastrez, *J. Mol. Biol.* **1994**, 237, 415–422.
- [190] K. D. Janda, L.-C. Lo, C.-H. L. Lo, M.-M. Sim, R. Wang, C.-H. Wong, R. A. Lerner, *Science* **1997**, 275, 945–948.
- [191] H. Pedersen, S. Hölder, D. P. Sutherlin, U. Schwitter, D. S. King, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **1998**, 95, 10523–10528.
- [192] J.-L. Jestin, P. Kristensen, G. Winter, *Angew. Chem.* **1999**, 111, 1196–1200; *Angew. Chem. Int. Ed.* **1999**, 38, 1124–1127.
- [193] S. Atwell, J. A. Wells, *Proc. Natl. Acad. Sci. USA* **1999**, 96, 9497–9502.
- [194] S. Demartis, A. Huber, F. Viti, L. Lozzi, L. Giovannoni, P. Neri, G. Winter, D. Neri, *J. Mol. Biol.* **1999**, 286, 617–633.
- [195] M. J. Olsen, D. Stephens, D. Griffiths, P. Daugherty, G. Georgiou, B. L. Iverson, *Nat. Biotechnol.* **2000**, 18, 1071–1074.
- [196] S. A. Lesley, *Methods Mol. Biol.* **1995**, 37, 265–278.
- [197] R. J. Jackson, T. Hunt, *Methods Enzymol.* **1983**, 96, 50–74.
- [198] D. S. Tawfik, A. D. Griffiths, *Nat. Biotechnol.* **1998**, 16, 652–656.
- [199] L. C. Mattheakis, R. R. Bhatt, W. J. Dower, *Proc. Natl. Acad. Sci. USA* **1994**, 91, 9022–9026.
- [200] L. C. Mattheakis, J. M. Dias, W. J. Dower, *Methods Enzymol.* **1996**, 267, 195–207.
- [201] J. Hanes, A. Plückthun, *Proc. Natl. Acad. Sci. USA* **1997**, 94, 4937–4942.
- [202] C. Schaffitzel, J. Hanes, L. Jermutus, A. Plückthun, *J. Immunol. Methods* **1999**, 231, 119–135.
- [203] M. He, M. J. Taussig, *Nucleic Acids Res.* **1997**, 25, 5132–5134.
- [204] R. W. Roberts, J. W. Szostak, *Proc. Natl. Acad. Sci. USA* **1997**, 94, 12297–12302.
- [205] R. Liu, J. E. Barrick, J. W. Szostak, R. W. Roberts, *Methods Enzymol.* **2000**, 318, 268–293.
- [206] N. Nemoto, E. Miyamoto-Sato, Y. Husimi, H. Yanagawa, *FEBS Lett.* **1997**, 414, 405–408.
- [207] J. Hanes, L. Jermutus, S. Weber-Bornhauser, H. R. Bosshard, A. Plückthun, *Proc. Natl. Acad. Sci. USA* **1998**, 95, 14130–14135.
- [208] V. W. Cornish, D. Mendel, P. G. Schultz, *Angew. Chem.* **1995**, 107, 677–689; *Angew. Chem. Int. Ed. Engl.* **1995**, 34, 621–633.

- [209] G. Cho, A. D. Keefe, R. Liu, D. S. Wilson, J. W. Szostak, *J. Mol. Biol.* **2000**, *297*, 309–319.
- [210] D. S. Wilson, J. W. Szostak, *Annu. Rev. Biochem.* **1999**, *68*, 611–647.
- [211] D. P. Bartel, J. W. Szostak, *Science* **1993**, *261*, 1411–1418.
- [212] E. H. Ekland, J. W. Szostak, D. P. Bartel, *Science* **1995**, *269*, 364–370.
- [213] J. C. Moore, H.-M. Jin, O. Kuchner, F. H. Arnold, *J. Mol. Biol.* **1997**, *272*, 336–347.
- [214] U. T. Bornscheuer, J. Altenbuchner, H. H. Meyer, *Biotechnol. Bieng.* **1998**, *58*, 554–559.
- [215] M. T. Reetz, A. Zonta, K. Schimossek, K. Liebeton, K.-E. Jaeger, *Angew. Chem.* **1997**, *109*, 2961–2963; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2830–2832.
- [216] L. Giver, A. Gershenson, P.-O. Freskgard, F. H. Arnold, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12809–12813.
- [217] H. Zhao, F. H. Arnold, *Protein Eng.* **1999**, *12*, 47–53.
- [218] H. Joo, Z. Lin, F. H. Arnold, *Nature* **1999**, *399*, 670–673.
- [219] T. Kumamaru, H. Suenaga, M. Mitsuoka, T. Watanabe, K. Furukawa, *Nat. Biotechnol.* **1998**, *16*, 663–666.
- [220] Z. Shao, H. Zhao, L. Giver, F. H. Arnold, *Nucleic Acids Res.* **1998**, *26*, 681–683.
- [221] H. Zhao, L. Giver, Z. Shao, J. A. Affholter, F. H. Arnold, *Nat. Biotechnol.* **1998**, *16*, 258–261.
- [222] J. R. Thompson, E. Register, J. Curotto, M. Kurtz, R. Kelly, *Yeast* **1998**, *14*, 565–571.
- [223] C. Yanisch-Perron, J. Vieira, J. Messing, *Gene* **1985**, *33*, 103–119.
- [224] A. Ullmann, F. Jacob, J. Monod, *J. Mol. Biol.* **1967**, *24*, 339–343.
- [225] J. P. Horwitz, J. Chua, R. J. Curby, A. J. Tomson, M. A. Da Rooge, B. E. Fisher, J. Mauricio, I. Klundt, *J. Med. Chem.* **1964**, *7*, 574–575.
- [226] P. Kast, *Gene* **1994**, *138*, 109–114.
- [227] E. T. Kaiser, F. J. Kézdy, *Science* **1984**, *223*, 249–255.
- [228] A. D. Keefe, J. W. Szostak, *Nature* **2000**, *401*, 715–718.