



Doctoral Thesis

Combinatorial design and directed evolution of AroQ-class chorismate mutases

Author(s):

Walter, Kai Uwe

Publication Date:

2001

Permanent Link:

<https://doi.org/10.3929/ethz-a-004353583> →

Rights / License:

[In Copyright - Non-Commercial Use Permitted](#) →

This page was generated automatically upon download from the [ETH Zurich Research Collection](#). For more information please consult the [Terms of use](#).

Diss. ETH Nr. 14506

**Combinatorial Design and Directed Evolution
of AroQ-Class Chorismate Mutases**

A dissertation submitted to the
Swiss Federal Institute of Technology (ETH) Zurich
for the degree of Doctor of Natural Sciences

Presented by

Kai Uwe Walter
Dipl. Chem. Ludwig-Maximilians-Universität München
born on January 6, 1971 in Munich, Germany

accepted on the recommendation of

Prof. Dr. Donald Hilvert, examiner
Prof. Dr. Rudolf Glockshuber, co-examiner

Zurich, 2001

Abstract

In living organisms almost all chemical transformations, from simple hydrolyses to C-C bond formation, are catalyzed by enzymes. These biocatalysts exhibit exacting recognition and binding properties for their substrates and generally accelerate the chemical transformation they promote by at least 10^6 -fold compared to the uncatalyzed process. Despite intense study, our understanding of how these properties are encrypted in the primary amino acid sequence is incomplete. As a result, attempts to design proteins from scratch or to redesign existing enzymes to improve or alter their properties is often unsuccessful, especially when the goal is altering the catalytic activity.

Evolutionary strategies are increasingly recognized as a valuable tool for tailoring protein properties. In contrast to rational (re-)design strategies that are based on the three-dimensional structure of the target enzyme, such evolutionary strategies rely entirely on Darwinian principles of evolution, exploiting mutation, selection and amplification in order to achieve desired functional changes. Since the amino acid constituents of a protein determine its activity, the ability to simultaneously probe millions of different alternatives and to select and characterize variants with the desired function allows us to refine our understanding of the relationship between protein structure and function.

The first part of this thesis describes work utilizing evolutionary methods to improve the catalytic activity of a topologically reengineered chorismate mutase from *Escherichia coli* (EcCM). In previous studies, the homodimeric six-helix bundle EcCM was converted into a hexamer, albeit with markedly reduced catalytic activity compared to the parent enzyme. The hexamer was originally identified by complementation of a chorismate mutase deficient strain of *E. coli*. However, the cells harboring this enzyme grow poorly under selection conditions in the absence of tyrosine and phenylalanine, making optimization through additional rounds of mutagenesis and selection possible.

VIII

With the assumption that catalytically improved variants would better complement the auxotroph, the gene encoding the hexamer was mutated by DNA shuffling and subjected to selection. After two rounds of mutagenesis and selection, several clones were identified that grew at wild-type levels. Relevant variants were isolated after each round and characterized biochemically. One variant, R2-6, contained three non-silent mutations (Ser15Asp/Leu79Phe/Thr87Ile) and exhibited a 6-fold improvement in k_{cat} and a 31-fold improvement in k_{cat}/K_m . Analytical ultracentrifugation experiments revealed that this variant, in contrast to the parent hexamer, is predominantly monomeric. In addition, this variant is more stable to thermal denaturation than the parental hexamer.

Attempts to further evolve R2-6 through an additional round of mutagenesis and selection yielded numerous fast-growing clones. Thirty of these variants were assayed for their catalytic activity and several showed further improved k_{cat}/K_m values. Interestingly, many of these variants are truncated at their C-terminus. However, these mutants are only marginally stable *in vitro* and readily form higher order aggregates.

The second part of this thesis describes the application of genetic selection to explore the probability of finding novel chorismate mutases (CMs) in protein sequence space. To that end, all helical elements of the dimeric, helical bundle chorismate mutase from *Methanococcus jannaschii* (MjCM') were replaced with simple binary-patterned modules constituted from a set of four polar and four nonpolar amino acids. The binary pattern of the designed sequence reflected the alternating order of hydrophobic and hydrophilic residues found in the secondary structure units of MjCM', thus specifying the type of residue at each position but not their specific identity. Two-stage *in vivo* selection yielded catalytically active variants possessing biophysical and kinetic properties typical of the natural enzyme even though ~80% of the protein originates from the simplified modules and more than 90% of the protein consists of only eight different amino acids. This study allows a quantitative assessment of the number of sequences compatible with a given fold and function. Moreover, new insights regarding the active site were gained, and structurally important residues for the formation of the helix bundle were identified.

It is assumed that in early evolution an alphabet of fewer than 20 amino acids was used to generate proteins. Building on the results with binary patterned CMs, the feasibility to construct an active enzyme from a highly reduced set of amino acids was tested. For this study, combinatorial site-directed mutagenesis in combination with genetic selection was applied to further simplify one of the binary patterned variants. Stepwise mutagenesis/selection at seven positions in the protein yielded catalytically active variants derived from only nine different amino acids. Detailed biochemical characterization of one variant revealed biophysical and kinetic properties similar to the natural enzyme. Despite its clear catalytic activity, nuclear magnetic resonance spectroscopy experiments suggest that this simplified variant exists predominantly as a molten globule. The results of this investigation suggest the possibility that simple molten globules constituted from a narrow set of amino acids might have served as catalysts in early evolution. In addition, the feasibility of constructing a catalyst employing a nine-letter amino acid alphabet suggests that the *de novo* design of proteins with tailored enzymatic activities may be attainable, because the conformational space of such a dramatically reduced alphabet might be computationally accessible.

The studies presented in this thesis demonstrate the utility of evolutionary strategies to improve the properties of proteins and highlight the benefit of genetic selection for the investigation of fundamental features of protein structure, function and evolution.

Zusammenfassung

In lebenden Organismen werden nahezu alle chemischen Umwandlungen, von der einfachen Hydrolyse bis zur Bildung einer C-C Bindung, von Enzymen katalysiert. Diese Biokatalysatoren erkennen und binden ihre Substrate mit hoher Spezifität und beschleunigen in der Regel die von ihnen katalysierte chemische Transformation mindestens millionenfach, verglichen mit dem unkatalysierten Prozess. Trotz intensiver Forschung ist unser Verständnis dafür, wie diese Eigenschaften in der primären Aminosäuresequenz verschlüsselt sind, unvollständig. Als Folge davon sind Versuche, Proteine von Grund auf zu konstruieren oder vorhandene Enzyme so umzugestalten, dass sie in ihren Eigenschaften verbessert oder verändert sind, häufig nicht erfolgreich. Insbesondere eine Veränderung der katalytischen Aktivität gelingt oft nicht.

Evolutionäre Strategien werden zunehmend als wertvolle Methode erkannt, um Eigenschaften von Proteinen gezielt zu verändern. Im Gegensatz zu rationalen Design-Strategien, die auf der Kenntnis der dreidimensionalen Struktur des Zielenzym basieren, stützen sich evolutionäre Strategien ausschließlich auf die Darwinschen Prinzipien der Evolution, indem Mutation, Selektion und Amplifikation ausgenutzt werden, um die gewünschten Funktionsänderungen zu erzielen. Da die Aminosäurebausteine eines Proteins über seine Aktivität bestimmen, erlaubt uns die Fähigkeit, Millionen von verschiedenen Alternativen gleichzeitig zu testen sowie Varianten mit der begehrten Funktionalität zu selektionieren und zu charakterisieren, unser Verständnis für die Beziehung zwischen der Struktur und der Funktion eines Proteins zu verfeinern.

Im ersten Teil dieser Dissertation wird der Einsatz von evolutionären Methoden zur Verbesserung der katalytischen Aktivität einer Chorismat Mutase von *Escherichia coli* (EcCM) mit umgestalteter Topologie beschrieben. In vorangegangenen Studien wurde das homodimere 6-Helixbündel-Protein EcCM in ein Hexamer umgewandelt, dessen katalytische Aktivität allerdings deutlich geringer ist als die des Ausgangsenzyms. Ursprünglich wurde das Hexamer durch Komplementierung eines Chorismat Mutase defizienten *E. coli* Stammes identifiziert. Jedoch wachsen Zellen, die dieses Enzym

beinhalten, schlecht unter Selektionsbedingungen in der Abwesenheit von Tyrosin und Phenylalanin, wodurch eine Optimierung mittels weiterer Runden von Mutagenese und Selektion möglich ist.

Unter der Annahme, dass katalytisch verbesserte Varianten den auxotrophen Stamm besser komplementieren würden, wurde das, für das Hexamer kodierende Gen durch "DNA Shuffling" mutiert und einer Selektion unterzogen. Nach zwei Runden von Mutagenese und Selektion wurden mehrere Klone identifiziert, die wie das Wildtyp Enzym wachsen. Nach jeder Runde wurden relevante Varianten isoliert und biochemisch charakterisiert. Eine der Varianten, R2-6, enthielt drei nicht-stumme Mutationen (Ser15Asp/Leu79Phe/Thr87Ile) und zeigte einen 6-fach gesteigerten k_{cat} und einen 31-fach verbesserten k_{cat}/K_m Wert gegenüber dem Wildtyp. Analytische Ultrazentrifugationsexperimente ergaben, dass R2-6 im Gegensatz zum ursprünglichen Hexamer vorwiegend als Monomer existiert. Darüber hinaus ist diese Variante stabiler als das Hexamer gegenüber thermischer Denaturierung.

In Versuchen, R2-6 durch eine zusätzliche Runde von Mutagenese und Selektion weiter zu evolvieren, konnten zahlreiche schnell wachsende Klone erhalten werden. Von diesen Varianten wurden dreißig auf ihre katalytische Aktivität getestet und einige zeigten eine weitere Verbesserung ihres k_{cat}/K_m Werts. Interessanterweise sind viele dieser Varianten C-terminal verkürzt. Jedoch sind diese Mutanten *in vitro* nicht sehr stabil und formen bereitwillig Aggregate.

Im zweiten Teil dieser Dissertation wird die Anwendung genetischer Selektion zur Untersuchung der Wahrscheinlichkeit beschrieben, neuartige Chorismat Mutasen (CMs) im Protein-Sequenzraum zu finden. Hierfür wurden alle helikalen Elemente der dimeren Chorismat Mutase von *Methanococcus jannaschii* (MjCM') mit Helixbündel-Struktur durch einfache binär gemusterte Module ersetzt, die aus einem Satz von vier polaren und vier unpolaren Aminosäuren hergestellt wurden. Das entworfene binäre Muster entspricht der alternierenden Abfolge von hydrophoben und hydrophilen Resten, wie sie in den Sekundärstruktur-Einheiten von MjCM' gefunden wird; dadurch wird die Art des Restes an jeder Position spezifiziert, aber nicht dessen spezifische Identität. Eine zweistufige *in vivo* Selektion brachte katalytisch aktive Varianten mit biophysikalischen und kinetischen Eigenschaften hervor, wie sie üblicherweise für das natürliche Enzym gefunden werden, obwohl ~80% des Proteins aus den vereinfachten Modulen besteht

und mehr als 90% des Proteins lediglich aus acht verschiedenen Aminosäuren zusammengesetzt ist. Diese Studie gestattet eine quantitative Abschätzung der Anzahl der Sequenzen, die mit einer bekannten Faltung und Funktion kompatibel sind. Außerdem konnten neue Einsichten bezüglich der aktiven Tasche gewonnen und Reste erkannt werden, die entscheidend zur Bildung der Helixbündel-Struktur beitragen.

Es wird angenommen, dass in einem frühen Stadium der Evolution ein Alphabet mit weniger als 20 Aminosäuren zur Verfügung stand, um Proteine zu erzeugen. Aufbauend auf den Resultaten mit den binär gemusterten CMs wurde die Möglichkeit getestet, ein aktives Enzym mit einem stark reduzierten Satz von Aminosäuren zu konstruieren. Für diese Untersuchung wurde kombinatorische zielgerichtete Mutagenese in Verbindung mit genetischer Selektion eingesetzt, um eine der binär gemusterten Varianten weiter zu vereinfachen. Schrittweise Mutagenese/Selektion an sieben Positionen des Proteins führte zu katalytisch aktiven Varianten, die sich lediglich aus neun unterschiedlichen Aminosäuren zusammensetzen. Die detaillierte biochemische Charakterisierung einer Variante zeigte, dass diese dem natürlichen Enzym ähnliche biophysikalische und kinetische Eigenschaften besitzt. Kernresonanz-spektroskopische Experimente deuten darauf hin, dass diese vereinfachte Variante trotz ihrer eindeutigen katalytischen Aktivität überwiegend als "*molten globule*" existiert. Die Resultate dieser Untersuchung lassen vermuten, dass einfache "*molten globules*", die aus einem begrenzten Satz von Aminosäuren zusammengesetzt sind, in einem frühen Stadium der Evolution als Katalysatoren gedient haben könnten. Die Fähigkeit, einen Katalysator unter Verwendung eines neun-Buchstaben Aminosäure-Alphabets zu konstruieren, lässt außerdem vermuten, dass das *de novo* Design von Proteinen mit maßgeschneiderten enzymatischen Eigenschaften möglich sein sollte, da der Konformationsraum eines derart dramatisch reduzierten Alphabets mit einem Computer zu berechnen sein sollte.

Die Studien, die in dieser Doktorarbeit vorgestellt werden, demonstrieren die Nützlichkeit evolutionärer Strategien für die Verbesserung der Eigenschaften von Proteinen und unterstreichen den Nutzen von genetischer Selektion zur Untersuchung von fundamentalen Merkmalen der Struktur, Funktion und Evolution von Proteinen.