

Diss. ETH No. 17897

**Artificial Enzymes: From Catalytic Antibodies
Toward *de novo* Enzyme Design**

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Natural Sciences

presented by

ROGER MÜLLER

Dipl. Natw. ETH

born on April 11, 1979

citizen of Rüslikon (ZH) and Zürich (ZH)

Accepted on the recommendation of

Prof. Dr. Donald Hilvert, examiner
Prof. Dr. Peter Seeberger co-examiner

Zürich, 2008

Abstract

A major goal of enzyme design is the customized generation of novel proteins that catalyze reactions that are not found in nature. Catalysis with high rates and selectivities would be of great practical interest for applications in biotechnology an industry. Such tailor-made catalysts may be generated using different approaches, including catalytic antibody technology, directed evolution, rational protein redesign, and most recently computational enzyme design. In this doctoral thesis, strategies for exploiting functional groups in catalytic antibodies and computational design are compared and contrasted.

Catalytic antibodies are useful models of natural enzymes. Using synthetic transition state analogs as antigens, catalytic antibodies have been created for a variety of chemical reactions. To elicit specific catalytic groups (e.g. general acids/bases or nucleophiles) in antibody active sites, charge complementarity between the antibody and its hapten has proven to be a valuable tool (Chapter 1). For example, this approach has successfully produced efficient catalysts for proton abstraction reactions, such as the base-promoted Kemp elimination reaction of benzisoxazoles to give salicylonitriles.

One of the most efficient catalysts for the Kemp elimination is antibody 34E4, which promotes efficient proton transfer from 5-nitrobenzisoxazole using a precisely positioned carboxylate residue (Glu^{H50}). While 34E4 displays high catalytic efficiency relative to other catalytic antibodies, it is still relatively inefficient compared to natural enzymes. Crystallographic and pre-steady state kinetic analyses of antibody 34E4 (Chapter 2) show that the resting catalyst adopts two interconverting active-site conformations, only one of which is capable of binding substrate. In the predominant isomer, the indole side chain of Trp^{L91} occupies the binding site and blocks ligand access. Slow conformational isomerization of this residue, on the same time scale as catalytic turnover, creates a deep and narrow binding site that can accommodate substrate and promote proton transfer using Glu^{H50} as base. Although 34E4 is among the best catalysts for the deprotonation of benzisoxazoles, its efficiency appears to be significantly limited by its conformational plasticity to its active site.

Antibody 13G5 also catalyzes Kemp eliminations, but its preferred substrate is 6-glutaramidebenzisoxazole, which is highly deactivated compared to 5-nitrobenzisoxazole (Chapter 3). In contrast to antibody 34E4, 13G5 exploits concerted acid/base chemistry. Mutagenesis and X-ray studies identified Asp^{H35} as the base that abstracts the proton from the substrate. In addition, a well-defined water molecule that is activated by Glu^{L34} stabilizes the developing negative charge in the transition state.

Whereas the substitution of Glu^{L34} by glutamine causes a negligible decrease in k_{cat} , it broadens the pH-rate profile by increasing the apparent ionization constant of the basic limb by more than two $\text{p}K_{\text{a}}$ units. The amide side chain of glutamine presumably serves as alternative hydrogen bond donor to the water molecule over a broader pH range than a carboxylic acid. Introduction of alanine at position L34 similarly broadens the pH-rate profile, but unexpectedly boosts the activity of the catalyst 16-fold, presumably due to a considerable remodeling of the solvent structure around the proton-donating water. Compared to the second-order rate constant for the acetate-promoted decomposition of 6-glutamidedibenzisoxazole, the $k_{\text{cat}}/K_{\text{m}}$ of the Glu^{L34}Ala variant provides an astonishing 10^9 -fold rate advantage and the aspartate base exhibits an effective molarity of $>10^6$ M. These large effects illustrate the utility of bifunctional catalysis in an antibody active site for promoting reactions with unactivated substrates.

Although catalytic antibodies have been generated for a wide range of chemical transformations, enzyme-like efficiencies have so far proven elusive. Computational enzyme design represents a potentially promising alternative for creating new enzymes. In contrast to immunization with an imperfect transition state analog, computational enzyme design offers the possibility of computing the real transition state(s) of a desired reaction and positioning functional groups around it for optimal catalysis. These minimal active site templates (theozymes) can be used to explore appropriate protein scaffolds capable of harboring such minimal active sites. If successful, the packing of the active site is optimized through rational incorporation of space-filling residues, which firmly anchor the positioning of the catalytic residues. Promising designs are then synthesized and tested experimentally for catalysis.

The incredible diversity of natural enzyme active sites can inspire the choice of catalytic groups to be used in theozyme calculations. Using high-resolution crystal structures of well-characterized natural enzymes, we extracted minimal functional group constellations, which we call reaction archetypes. In Chapter 4, we describe the basic methodology for extracting minimal active site descriptions from a large body of structural information, taking the well-known catalytic triad of hydrolytic enzymes as an example. The triads from different enzymes and fold families are positioned nearly identically with regard to the atoms important for catalysis. However, the vectorial origin of the functional group side chains are fold-dependent and vary considerably among the different folds, offering the designer freedom in functional group orientation. For the carboxylesterases, the extracted geometrical parameters correlating the catalytic triad residues to each other and to the substrate were compared to quantum

mechanically optimized structures along the entire reaction coordinate of butyrylcholine esterase. The theoretical ranges largely overlap with the empirical ranges, although tetrahedral structures most closely match experiment. Enzyme active sites are apparently preorganized to a geometry that can be objectively and quantitatively defined as minimizing conformational reorganization while maintaining optimal transition state stabilization for every step during catalysis.

In Chapter 5, we expand these initial reaction archetypes (catalytic triads) to functional group equivalents for general acids, general bases and nucleophiles. Nature uses a broad variety of different functional groups for specific chemical tasks. For example, histidine, lysine, carboxylate, and tyrosine side chains are used as bases to promote proton transfer. In many cases, the catalytic residues are activated by neighboring functional groups, which may be integrated as minimal catalytic units. This functional diversification is likely to be useful in computational enzyme design as well, since the choice of functional group can be tailored to the geometric and electrostatic demands of individual starting scaffolds.

In comparison to catalytic antibodies, computational enzyme design offers the possibility of designing enzyme-like arrays of functional groups directly into artificial active sites. Moreover, it is not restricted to a single protein scaffold and the possible intrinsic limitations thereof. Recent exciting progress in generating artificial proteins with considerable catalytic activity has highlighted the promise of this field. In combination with *in vivo* or *in vitro* evolution strategies, computational enzyme design may prove to be the best strategy for creating artificial enzymes in the future.

Zusammenfassung

Ein bedeutendes Ziel des Enzymdesigns ist die Herstellung neuartiger Proteine, welche Reaktionen katalysieren die in der Natur nicht vorkommen. Für Anwendungen in der Biotechnologie und Industrie sind effiziente, selektive Katalysatoren von grossem praktischem Interesse. Solche massgeschneiderte Katalysatoren können auf verschiedene Arten hergestellt werden: durch katalytische Antikörper, zielgerichtete Evolution, rationale Umgestaltung existierender aktiver Taschen, oder aber seit kurzem auch durch computergestütztes Design von Enzymen. In dieser Dissertation werden Strategien zur Einbindung funktioneller Gruppen in katalytischen Antikörpern und computergestütztem Design verglichen und gegenübergestellt.

Katalytische Antikörper sind nützliche Modelle natürlich vorkommender Enzyme. Mit Hilfe synthetischer Übergangszustandsanaloge wurden katalytische Antikörper für eine Vielfalt chemischer Reaktionen generiert. Um gezielt katalytische Gruppen (z.B. generelle Säuren/Basen oder Nukleophile) in aktive Taschen eines Antikörpers einzubauen, hat sich die Methode der Ladungskomplementarität zwischen Antikörper und dessen Hapten als sehr nützlich bewiesen (Kapitel 1). Diese Methode hat unter anderem Antikörper hervorgebracht, die Deprotonierungs-Reaktionen wie z. B. die Kemp-Eliminierung katalysieren und dabei Benzisoxazole in Salicylonitrile umwandeln können.

Einer der effizientesten Katalysatoren für die Kemp-Eliminierung ist der Antikörper 34E4, welcher die Umwandlung von 5-Nitrobenzisoxazol mit Hilfe eines präzise positionierten Carboxylat-Restes (Glu^{H50}) katalysiert. Obwohl 34E4 eine sehr hohe Effizienz gegenüber anderen katalytischen Antikörpern aufweist, ist er dennoch relativ langsam im Vergleich zu natürlichen Enzymen. Kristallographische und “pre-steady state“ kinetische Untersuchungen an 34E4 (Kapitel 2) haben aufgezeigt, dass der Antikörper in zwei interkonvertierenden Konformationen vorkommt, wobei nur die eine das Substrat binden kann. Im überwiegend vorkommenden Isomer besetzt die Indolseitenkette des Trp^{L91} die Bindungstasche und blockiert so den Zugang für das Substrat. Erst die langsame Isomerisierung dieses Tryptophan-Restes, ungefähr in derselben Zeitskala wie die Substratumsetzung, kreierte eine schmale Bindungstasche, die Platz für das Substrat bietet. Obwohl 34E4 einer der besten Katalysatoren für die Deprotonierung von Benzisoxazolen ist, scheint seine Effizienz durch diese konformationelle Plastizität beträchtlich limitiert zu sein.

Der Antikörper 13G5 katalysiert auch Kemp-Eliminierungen, jedoch mit 6-Glutaramidbenzisoxazol als Substrat, welches im Vergleich zu 5-Nitrobenzisoxazol viel

weniger reaktiv ist (Kapitel 3). Im Gegensatz zu Antikörper 34E4 bedient sich 13G5 konzertierter Säure/Base Katalyse um das Substrat umzusetzen. Mutagenese- und Kristallstruktur-Studien haben Asp^{H35} als Base für die Protonabstraktion identifiziert. Zusätzlich wird die negative Ladung im Übergangszustand durch ein Wassermolekül stabilisiert, welches seinerseits von Glu^{L34} aktiviert wird. Während der Austausch dieses Glutamates den k_{cat} Wert nur unwesentlich verringert, wird die scheinbare basische Ionisierungskonstante um mehr als zwei pH-Einheiten erhöht. Die Amidseitenkette des Glutamates fungiert vermutlich als alternativer Wasserstoffbrücken-Donor zum Wassermolekül über einen weiteren pH-Bereich als die Karbonsäure. Auch der Austausch durch Alanin führt zur selben Verbreiterung des pH-Profiles, die Aktivität des Katalysators wird jedoch unerwarteter Weise ca. 16-fach gesteigert, wahrscheinlich wegen der beträchtlichen Reorganisation des Wasserstoffbrücken-Netzwerkes um das katalytische Wassermolekül. Der Vergleich der Geschwindigkeitskonstante zweiter-Ordnung der Acetat-katalysierten Zersetzung von 6-Glutaramidbenzisoxazol mit dem k_{cat}/K_m Wert der Glu^{L34}Ala Variante führt zu einer erstaunlichen 10^9 -fachen Beschleunigung. Die Aspartat-Base weist zudem eine effektive Molarität von $>10^6$ M auf. Diese enormen Effekte veranschaulichen den Nutzen der bifunktionellen Katalyse in aktiven Taschen von Antikörpern bei der Umsetzung wenig reaktiver Substrate.

Obwohl katalytische Antikörper für viele verschiedene chemische Umwandlungen generiert wurden scheinen enzym-ähnliche katalytische Aktivitäten nur schwer zugänglich. Computergestütztes Design von Enzymen stellt eine viel versprechende Alternative dar um neue Enzyme zu kreieren. Im Gegensatz zur Immunisierung mit einem mangelhaften Übergangszustandsanalog hat man bei computergestützten Methoden die Möglichkeit den realen Übergangszustand einer Reaktion zu berechnen und alle notwendigen funktionellen Gruppen für die Katalyse optimal um diesen zu platzieren. Diese auf die wesentlichen Gruppen reduzierten aktiven Taschen, so genannte Theozyme, können nachfolgend benützt werden um die katalytischen Reste in geeigneten Proteingerüsten zu platzieren. Anschliessend wird die Packung der aktiven Tasche rund um die katalytischen Reste optimiert. Vielversprechende Designs werden dann hergestellt und experimentell auf katalytische Aktivität getestet.

Die unglaubliche Vielfalt aktiver Taschen natürlicher Enzyme kann die Wahl von katalytischen Gruppen in Theozymberechnungen anregen. In Kapitel 4 haben wir minimale Konstellationen funktioneller Gruppen ("Reaction Archetypes") aus hochauflösenden Kristallstrukturen von gut charakterisierten natürlichen Enzymen extrahiert. Am Beispiel der katalytischen Triaden hydrolytischer Enzyme beschreiben

wir eine generelle Methode um minimale Konstellationen der funktionellen Gruppen auf der Basis von strukturellen Informationen aktiver Taschen von verschiedenen Enzymen zu definieren. Die katalytischen Triaden verschiedener Enzyme und Struktur-Familien ballen eng zusammen bezüglich jener Atome die wichtig für die Katalyse sind. Der vektorielle Ursprung der Seitenketten der Triaden-Reste jedoch ist struktur-abhängig und variiert beträchtlich zwischen den verschiedenen Strukturen, was wiederum viel Freiheit für den Enzymdesigner bezüglich der Orientierung der funktionellen Gruppen bedeutet. Für die Carboxylesterasen wurden die extrahierten geometrischen Parameter zusätzlich mit quantenmechanisch optimierten Strukturen entlang der ganzen Reaktionskoordinate der Butyrylcholinesterase verglichen. Die theoretischen Werte liegen vorwiegend im Rahmen der empirisch gemessenen Werte und die tetraedrischen Intermediate liegen am nächsten bei den experimentellen Werten. Die aktive Tasche des Enzyms ist scheinbar so organisiert, um einerseits die konformationelle Reorganisation zu minimieren und andererseits die optimale Stabilisierung des Übergangszustandes für jeden Schritt der Katalyse zu gewährleisten.

In Kapitel 5 weiten wir die eingeführten "Reaction Archetypes" (katalytische Triaden) zu Äquivalenten von funktionellen Gruppen für generelle Säuren, generelle Basen und Nukleophile aus. Die Natur nutzt eine grosse Vielfalt an verschiedenen funktionellen Gruppen für eine bestimmte chemische Umwandlung. Für die generelle Basen-Katalyse wählt die Natur beispielsweise zwischen Histidin-, Lysin-, Carboxylat-, und Tyrosin-Seitenketten aus, um einen effizienten Protonentransfer zu erreichen. Meistens werden diese katalytischen Reste durch benachbarte Reste aktiviert und wir haben sie als minimale katalytische Einheiten zusammengefasst. Solche strukturelle Vielfalt ist mit grosser Wahrscheinlichkeit auch nützlich für computergestütztes Design von Enzymen, da somit die Wahl der funktionellen Gruppen auf die geometrischen und elektrostatischen Gegebenheiten des Proteingerüsts abgestimmt werden kann.

Im Vergleich zu katalytischen Antikörpern eröffnet das computergestützte Design von Enzymen nicht nur die Möglichkeit enzym-ähnliche Anordnungen funktioneller Gruppen in künstliche Proteine einzubauen, weiterhin ist es auch nicht auf ein einziges Proteingerüst und dessen mögliche Limitationen begrenzt. Die kürzlich erzielten Fortschritte bei der Herstellung künstlicher Proteine unterstützen die Versprechen in dieses Gebiet. In Kombination mit *in vivo* oder *in vitro* Evolutions-Strategien scheint das computergestützte Design von Enzymen künftig die beste Strategie zur Herstellung künstlicher Enzymen zu sein.